309 V

# ENGINEERING ELECTRONICS

## *with Industrial Applications and Control*

**JOHN D. RYDER, Ph.D.**

*Dean of Engineering and*
*Professor of Electrical Engineering*
*Michigan State University*

B 1063

ANDHRA PRADESH
LIBRARY
College of Basic Courses,
Rajendranagar.
AGRICULTURAL UNIVERSITY

PROCESSED
ANGRAU Central Library
Hyderabad
12469

INTERNATIONAL STUDENT EDITION

ANGRAU CENTRAL LIBRARY
Acc: No; 12469
Date: 13.8.85

McGRAW-HILL BOOK COMPANY, INC.
New York Toronto London

KŌGAKUSHA COMPANY, LTD.
Tokyo

CHECKED 1997

CHECKED

ENGINEERING ELECTRONICS

*INTERNATIONAL STUDENT EDITION*

Exclusive rights by Kōgakusha Co., Ltd. for manufacture and export from Japan. This book cannot be re-exported from the country to which it is consigned by Kōgakusha Co., Ltd. or by McGraw-Hill Book Company, Inc. or any of its subsidiaries.

X

Copyright © 1957 by the McGraw-Hill Book Company, Inc. All rights reserved. This book, or parts thereof, may not be reproduced in any form without permission of the publishers.

*Library of Congress Catalog Card Number* 56–12272

621·381

ANGRAU
621.381
N57RYD
Acc No. 12469

TOSHO PRINTING CO., LTD., TOKYO, JAPAN

McGraw-Hill Electrical and Electronic Engineering Series

FREDERICK EMMONS TERMAN, *Consulting Editor*

W. W. HARMAN and J. G. TRUXAL, *Associate Consulting Editors*

---

1063

ANDHRA PRADESH
LIBRARY
College of Basic Courses,
Rajendranagar.
AGRICULTURAL UNIVERSITY

# ENGINEERING ELECTRONICS

## McGraw-Hill Electrical and Electronic Engineering Series

FREDERICK EMMONS TERMAN, *Consulting Editor*

W. W. HARMAN and J. G. TRUXAL, *Associate Consulting Editors*

---

*Bailey and Gault* · ALTERNATING-CURRENT MACHINERY
*Beranek* · ACOUSTICS
*Bruns and Saunders* · ANALYSIS OF FEEDBACK CONTROL SYSTEMS
*Cage* · THEORY AND APPLICATION OF INDUSTRIAL ELECTRONICS
*Cauer* · SYNTHESIS OF LINEAR COMMUNICATION NETWORKS, VOLS. I AND II
*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
*Geppert* · BASIC ELECTRON TUBES
*Glasford* · FUNDAMENTALS OF TELEVISION ENGINEERING
*Happell and Hesselberth* · ENGINEERING ELECTRONICS
*Harman* · FUNDAMENTALS OF ELECTRONIC MOTION
*Harrington* · INTRODUCTION TO ELECTROMAGNETIC ENGINEERING
*Hayt* · ENGINEERING ELECTROMAGNETICS
*Hessler and Carey* · FUNDAMENTALS OF ELECTRICAL ENGINEERING
*Hill* · ELECTRONICS IN ENGINEERING
*Johnson* · TRANSMISSION LINES AND NETWORKS
*Kraus* · ANTENNAS
*Kraus* · ELECTROMAGNETICS
*LePage* · ANALYSIS OF ALTERNATING-CURRENT CIRCUITS
*LePage and Seely* · GENERAL NETWORK ANALYSIS
*Millman and Seely* · ELECTRONICS
*Millman and Taub* · PULSE AND DIGITAL CIRCUITS
*Rodgers* · INTRODUCTION TO ELECTRIC FIELDS
*Rüdenberg* · TRANSIENT PERFORMANCE OF ELECTRIC POWER SYSTEMS
*Ryder* · ENGINEERING ELECTRONICS
*Seely* · ELECTRON-TUBE CIRCUITS
*Seely* · ELECTRONIC ENGINEERING
*Seely* · INTRODUCTION TO ELECTROMAGNETIC FIELDS
*Seely* · RADIO ELECTRONICS
*Siskind* · DIRECT-CURRENT MACHINERY
*Skilling* · ELECTRIC TRANSMISSION LINES
*Skilling* · TRANSIENT ELECTRIC CURRENTS
*Spangenberg* · FUNDAMENTALS OF ELECTRON DEVICES
*Spangenberg* · VACUUM TUBES
*Stevenson* · ELEMENTS OF POWER SYSTEM ANALYSIS
*Storer* · PASSIVE NETWORK SYNTHESIS
*Terman* · ELECTRONIC AND RADIO ENGINEERING
*Terman and Pettit* · ELECTRONIC MEASUREMENTS
*Thaler* · ELEMENTS OF SERVOMECHANISM THEORY
*Thaler and Brown* · SERVOMECHANISM ANALYSIS
*Thompson* · ALTERNATING-CURRENT AND TRANSIENT CIRCUIT ANALYSIS
*Truxal* · AUTOMATIC FEEDBACK CONTROL SYSTEM SYNTHESIS

## PREFACE

Electronics, employing basic knowledge of the dynamics of electric charges and their physical behavior in conductors, circuit and network theory, electromagnetic fields, and feedback and control theory, is profoundly influencing all areas of electrical engineering. Formerly somewhat restricted to the higher frequencies and to the transmission of information, electronics has now invaded the industrial and business areas, and a broader viewpoint is needed.

This text attempts to supply the basic knowledge required by those working in the broad nonradio portion of the electrical-engineering field. Many arbitrary decisions have been made to include or exclude material since, while the radio portion of the field of electronics is rather specifically bounded, no such bounds have been found for the industrial or engineering electronics area. Because of this undefined breadth it has not been possible to study applications in great detail, and this has aided in preventing the work from becoming a handbook of operation of specific circuits and apparatus.

The first three chapters provide a survey of electronic devices and concepts for those without previous training. Fundamentals of semiconductors and the transistor are covered in Chap. 12. Throughout, frequent recourse is had to mathematics and circuit theory to provide a quantitative understanding of the subject.

JOHN D. RYDER

# CONTENTS

*Preface* . . . . . v

1. INTRODUCTION . . . . . 1

1-1. The atom and the electron. 1-2. Development of the vacuum tube. 1-3. Gas tubes. 1-4. Solid-state developments.

2. PHYSICAL PHENOMENA IN ELECTRON TUBES . . . . . 5

2-1. Action of electrons in electric fields. 2-2. Current flow in space. 2-3. Action of electrons in magnetic fields. 2-4. Electrons in metals. 2-5. The surface-energy barrier and electron emission. 2-6. Thermionic-cathode materials. 2-7. Thermionic-cathode types. 2-8. Photoemission. 2-9. Secondary and high-field emissions. 2-10. Space-charge effects in the thermionic diode. 2-11. Current control by a grid; the vacuum triode. 2-12. Energy losses and cooling in vacuum tubes. 2-13. Ionization in gases.

3. THE VACUUM TUBE AS A CIRCUIT ELEMENT . . . . . 32

3-1. Nonlinear volt-ampere characteristics. 3-2. Vacuum-triode current and voltage notation. 3-3. Triode graphical characteristics and triode parameters. 3-4. The dynamic transfer curve. 3-5. Graphical analysis of circuit performance. 3-6. The voltage-source equivalent circuit for a triode. 3-7. The current-source equivalent circuit. 3-8. Gain of an amplifier. 3-9. Measurement of tube parameters. 3-10. The tetrode. 3-11. The pentode. 3-12. The beam tube. 3-13. The variable-$\mu$ tube. 3-14. Self-bias of amplifiers; voltage sources. 3-15. Distortion of waveform. 3-16. Classification of tube operating conditions. 3-17. The decibel in gain measurement. 3-18. The cathode-ray tube and oscillograph.

4. SMALL-SIGNAL AMPLIFIERS AS CIRCUIT ELEMENTS . . . . . 64

4-1. Basic amplifier types. 4-2. Coupling of cascaded amplifiers. 4-3. Input admittance of a triode as a grounded-cathode amplifier. 4-4. Grounded-cathode amplifier analysis. 4-5. The cathode-follower amplifier. 4-6. Design of cathode-follower circuits. 4-7. Analysis of the modified cathode follower. 4-8. The grounded-grid amplifier. 4-9. A-C operation of voltage amplifiers

5. RESPONSE OF SMALL-SIGNAL AMPLIFIERS . . . . . 87

5-1. Frequency and phase response of the grounded-cathode amplifier. 5-2. Further study of the frequency and phase response of the *RC* amplifier. 5-3. Steady-state decibel-gain plots. 5-4. *RC* circuits for alteration of bandwidth. 5-5. Factors affecting amplifier bandwidth. 5-6. Low-fre-

quency compensation of *RC* amplifiers. 5-7. High-frequency compensation of the *RC* amplifier. 5-8. Pulse response of the *RC* coupling circuit. 5-9. Relations between sinusoidal and pulse response. 5-10. Square-wave response and testing. 5-11. Fourier series. 5-12. Analysis of a recurrent rectangular pulse. 5-13. Noise in amplifiers.

6. Low-frequency Power Amplifiers . . . . . 136

6-1. Output circuits for low-frequency power amplifiers. 6-2. Load-line construction. 6-3. Load line for reactive loads. 6-4. The ideal Class A amplifier. 6-5. Power relations. 6-6. Theoretical limit on Class A efficiency. 6-7. Determination of amplitude distortion. 6-8. Intermodulation distortion. 6-9. Choice of load for optimum power output. 6-10. The distortion rule. 6-11. The push-pull amplifier. 6-12. Bias considerations in push-pull amplifiers. 6-13. Equivalent circuit for the push-pull amplifier. 6-14. Graphical analysis using the composite tube characteristics. 6-15. The Class AB push-pull amplifier. 6-16. The Class B push-pull amplifier. 6-17. Design of Class B amplifiers for large power output. 6-18. Grid circuit power requirements. 6-19. Signal sources for push-pull amplifiers.

7. Feedback in Amplifiers . . . . . 178

7-1. Feedback principles. 7-2. Negative feedback and gain stability. 7-3. Negative feedback and distortion. 7-4. Negative feedback and noise. 7-5. Feedback measured in decibels; an example. 7-6. Basic negative-feedback circuits. 7-7. Other feedback circuits. 7-8. Positive feedback; decoupling. 7-9. Stability of feedback amplifiers. 7-10. Further discussion of stability conditions.

8. Direct-coupled Amplifiers; Computing Amplifiers . . . . . 206

8-1. Early direct-coupled amplifiers. 8-2. The resistance-coupled amplifier. 8-3. Cathode coupling in d-c amplifiers. 8-4. The balanced push-pull amplifier. 8-5. Cathode-drift compensation. 8-6. Series cathode compensation; the bridge-balanced amplifier. 8-7. Vacuum-tube electrometers. 8-8. Residual drift in direct-coupled amplifiers. 8-9. Chopper-modulated amplifiers. 8-10. Chopper-stabilized direct-coupled amplifiers. 8-11. The operational amplifier. 8-12. Use of positive feedback. 8-13. Analog computation. 8-14. Lagrange's equation for dynamical systems. 8-15. The direct-analog computer. 8-16. Scale factors. 8-17. Functional or operational analog computers. 8-18. Function generators. 8-19. Nonlinear-circuit elements. 8-20. Limiting circuits. 8-21. Multiplying devices and circuits.

9. Switching Circuits; Digital Computation . . . . . 255

9-1. Circuit elements. 9-2. Tubes in switching circuits. 9-3. Response of *RC* and *RL* series circuits. 9-4. The time constant. 9-5. The differentiating circuit. 9-6. Integrating circuits. 9-7. Clipping circuits. 9-8. Clamping circuits. 9-9. Clamping in the television system. 9-10. The ringing circuit. 9-11. The blocking oscillator. 9-12. Time base or sweep generation for cathode-ray-tube deflection. 9-13. Simple *RC* sweep generation. 9-14. Thyratron sweep generators. 9-15. Triggered vacuum-tube sweep generators. 9-16. Linearization of the *RC* sweep. 9-17. Other forms of time bases. 9-18. Sweep circuits for magnetically deflected

cathode-ray tubes. 9-19. The free-running multivibrator. 9-20. Effect of bias. 9-21. Synchronization of the multivibrator. 9-22. The one-shot multivibrator. 9-23. The trigger, or flip-flop, circuit. 9-24. Scaling circuits. 9-25. Decimal scaling circuits. 9-26. The decade-counter tube. 9-27. The binary-number system. 9-28. Digital computers. 9-29. Computer input and output. 9-30. Memory, or storage, facilities. 9-31. Vacuum-tube gating circuits. 9-32. Diode switching matrices. 9-33. And, or, not circuits. 9-34. Addition logic.

10. Power Supplies and Filters; Voltage Regulators . . . . . 328

10-1. The half-wave diode rectifier. 10-2. The full-wave rectifier. 10-3. Ripple factor. 10-4. The shunt-capacitor filter. 10-5. The series-inductor filter. 10-6. The L-section, or inductor-input, filter. 10-7. The bleeder resistor and critical input inductance. 10-8. Capacitor-input filters. 10-9. Voltage-multiplying rectifiers. 10-10. Rectifier-circuit design. 10-11. Voltage-regulator (VR) tubes. 10-12. Vacuum-tube regulators. 10-13. A-c regulators. 10-14. Voltage references.

11. Oscillators and Class C Amplifiers; High-frequency Heating . . 366

11-1. Waveforms in Class C Amplifiers. 11-2. Analysis of Class C amplifier performance. 11-3. Power relations. 11-4. The resonant-load circuit. 11-5. Frequency multipliers. 11-6. Feedback in oscillators. 11-7. Circuit criteria for feedback oscillators. 11-8. Basic feedback-oscillator circuits. 11-9. Electron coupling; crystal control. 11-10. Phase-shift oscillators. 11-11. Ultrasonic oscillators. 11-12. Electromagnetic fields in metals. 11-13. Skin effect in metals. 11-14. Induction heating of cylindrical bars. 11-15. Heating of dielectric cylinders or slabs. 11-16. Power required for rapid heating. 11-17. Heating-power calculations. 11-18. Oscillators as high-voltage d-c sources.

12. Semiconductors; Transistors . . . . . 412

12-1. Semiconductors. 12-2. Energy levels and conductivity in semiconductors. 12-3. The *P-N* junction. 12-4. Junction and point-contact diode rectifiers. 12-5. The point-contact transistor. 12-6. The junction transistor. 12-7. Transistors and vacuum tubes. 12-8. Four-terminal networks. 12-9. Equivalent circuit for the common-base amplifier. 12-10. The common-emitter circuit. 12-11. The common-collector amplifier. 12-12. Summary of transistor amplifier characteristics. 12-13. Graphical analysis of large-signal transistor performance. 12-14. Bias and stabilization. 12-15. Cascaded amplifiers. 12-16. Frequency response of the $RC$ amplifier. 12-17. Push-pull and complementary symmetry circuits. 12-18. Instability in point-contact transistor circuits. 12-19. Transistor feedback oscillators. 12-20. Negative-resistance oscillators and triggers. 12-21. The tetrode transistor. 12-22. Large-area rectifiers. 12-23. Thermistors.

13. Photoelectric Devices and Applications . . . . . 467

13-1. Light. 13-2. Cathode materials. 13-3. The vacuum photoemissive cell (phototube). 13-4. The gas-filled phototube. 13-5. Phototube relays. 13-6. Phototube circuits for measurement purposes. 13-7. Use of modulated light. 13-8. Electron multipliers. 13-9. The photovoltaic cell. 13-10. The phototransistor. 13-11. Light amplifiers.

14. Power Rectification . . . . . 488

14-1. Arc discharges. 14-2. Thermionic cathodes in gas discharges. 14-3. Mercury-pool cathodes. 14-4. Mercury-vapor-tube characteristics and ratings. 14-5. Arc-back. 14-6. Basic half-wave gas-diode rectifier. 14-7. Full-wave and bridge rectifiers. 14-8. Parallel operation of gas diodes. 14-9. Three-phase half-wave rectifier relations. 14-10. *m*-phase-rectifier relations. 14-11. Transformer utility factor. 14-12. *m*-phase rectifiers with inductance filters. 14-13. Effect of transformer reactance. 14-14. The three-phase bridge circuit. 14-15. Double-Y rectifiers. 14-16. Polyphase branched connections. 14-17. Performance.

15. Power Control and Inversion . . . . . 528

15-1. The thyratron. 15-2. Ionization and deionization times. 15-3. Grid current in thyratrons. 15-4. The gas tetrode or shield-grid thyratron. 15-5. Bias control of average current. 15-6. Phase-shift control. 15-7. Combined phase and bias control. 15-8. The phase-shifting bridge. 15-9. Saturable-reactor control. 15-10. Pulse control of thyratrons. 15-11. The ignitron. 15-12. Ignitron firing circuits. 15-13. Controlled thyratron or ignitron rectifiers. 15-14. Control of alternating current. 15-15. Ignitron ratings. 15-16. Switching gas triodes with d-c anode potential. 15-17. Gas-triode counting circuits. 15-18. Inverters. 15-19. The transistor inverter.

16. Relays, Timers, and Resistance-welding Controls . . . . . 570

16-1. The contact relay. 16-2. D-c-operated time-delay relays. 16-3. A-c-operated time-delay relays. 16-4. Time-interval measurements. 16-5. The resistance-welding process. 16-6. Automatic sequence timer. 16-7. Synchronous weld control. 16-8. Follow-up, or trailing-tube, circuit. 16-9. Timing the weld. 16-10. Heat control. 16-11. Energy-storage welding systems. 16-12. Polyphase welding systems.

17. Electronic Motor Control . . . . . 588

17-1. D-C motor relations. 17-2. Thyratron supply for the d-c motor. 17-3. Speed-torque relations. 17-4. Speed regulation by armature-voltage control. 17-5. IR-drop compensation. 17-6. Field weakening. 17-7. Armature-current limiting. 17-8. Auxiliary controls. 17-9. Speed control by tachometer. 17-10. Control of low-torque a-c motors. 17-11. Phase-selective circuits.

18. Servomechanisms . . . . . 613

18-1. Servomechanisms and control. 18-2. First-order servomechanism. 18-3. Second-order systems. 18-4. Transient response of a second-order system to a step input. 18-5. Derivative control. 18-6. Error-rate control. 18-7. Integral control. 18-8. Temperature regulator with integral control. 18-9. Sinusoidal response of a second-order system. 18-10. Transfer-function analysis. 18-11. Transfer function for multiple-loop feedback. 18-12. Transfer functions of several circuits. 18-13. Transfer functions of several physical elements. 18-14. A-C motor servo system. 18-15. Stability. 18-16. Stability criteria. 18-17. Gain and phase margin. 18-18. Example: Further Nyquist plots. 18-19. Other systems.

*Index* . . . . . 655

CHAPTER 1

# INTRODUCTION

The science of electronics is fundamental to a vast area of electrical engineering. Certainly the development of this science must be given credit for the useful application of frequencies above a few hundred cycles and for the wide application of nonsinusoidal waveforms. Electronic knowledge is now employed in devices operating from zero frequency to tens of thousands of megacycles, handling powers from micromicrowatts to thousands of kilowatts, and employing a great variety of the infinitude of waveforms.

In this text it is proposed to emphasize the scientific principles of importance to the area known as industrial, or engineering, electronics, usually associated with the lower frequencies and dealing quite largely with the signals and methods leading to the control of energy in some form, as contrasted with the concept of transmission of information more often associated with the field of radio.

**1-1. The Atom and the Electron.** Present atomic theories, based on the atom model proposed by Bohr, place the mass and positive charge of an atom in a central nucleus, surrounded by moving electrons. The *nucleus* is composed of *neutrons* and *protons*. Neutrons have a mass approximately one-sixteenth that of the mass of the oxygen atom and have zero electrical charge. They occur in fixed numbers in all nuclei except that of hydrogen $H^1$. The proton has a mass slightly lighter than that of the neutron and approximately equal to the mass of the hydrogen atom $H^1$. It has a charge equal to that of the electron but positive in sign. The chemical elements are differentiated by varying numbers of protons and neutrons in the nucleus, the number of protons being equal to the atomic number, hydrogen having one, helium two, and so on. Since the atom is in electrical balance in its normal state, the number of associated electrons moving around the nucleus is equal to the number of protons.

The *electron* is the smallest known indivisible unit of electric charge. Early experiments succeeded in measuring the ratio of the charge of the electron to its mass, and in 1910 the charge was first measured, thus determining these properties of the electron. Repeated experiments have refined the values for these constants of the electron, and they are now

believed to be

$$\text{Charge} = e = 1.602 \times 10^{-19} \text{ coulomb}$$
$$\text{Mass} = m = 9.106 \times 10^{-31} \text{ kg}$$
$$\text{Charge/unit mass} = \frac{e}{m} = 1.759 \times 10^{11} \text{ coulomb/kg}$$

The electron appears to have a dual personality, at times being best thought of as a particle, and in other experiments seeming to require wave properties for a satisfactory explanation of the phenomena. This duality is not yet fully understandable. When wave properties are assigned, the electron wavelength is related to its velocity as

$$\lambda = \frac{h}{mv} \quad \text{meters}$$

where $h$ = Planck's constant = $6.624 \times 10^{-34}$ joule-sec
$m$ = electron mass, kg
$v$ = electron velocity, meters/sec

Electric current is due to movement of electric charges, frequently electrons. However, long before the existence of the electron was known, it was customary to speak of electric current as due to movement of positive charges from positive to negative in a metallic circuit external to the source. This usage is too well established to be readily overcome, even though it is now known to be reversed, with the negative electrons moving from negative to positive in the external metallic circuit. Therefore, when electric currents are discussed, the customary direction will be meant; when reference is made to electronic current, then the flow of electrons is being considered.

The *photon* is also important as a fundamental particle. It is a bundle of radiant energy, or light, the amount of energy being related to the frequency as

$$W = hf \quad \text{joules}$$

The amount of energy carried by any given photon thus depends on its frequency. This particle also shows duality in character and must be considered as having wave properties in discussing many optical experiments.

The photon, while not having a material mass, may be considered as having a radiation mass, and this equivalent mass is related to the photon energy by the Einstein equation

$$W = mc^2 \quad \text{joules}$$

where $c$ is the velocity of light = $3 \times 10^8$ meters per sec.

**1-2. Development of the Vacuum Tube.** The first use of a vacuum tube probably occurred with the discovery of the Edison effect by

Thomas A. Edison in 1883. During an experiment with an incandescent lamp Edison discovered that a current would flow through the evacuated space from a heated filament to a small metal plate when the plate was positive and ceased to flow when the plate was negative, indicating rectifying properties. This current discovered by Edison is now known to be due to electron flow from the heated filament to the positive electrode.

In 1897 J. A. Fleming, in England, after hearing of the work of Edison, applied the Edison effect to the detection or rectification of radio signals in a diode which became known as the *Fleming valve.*

In 1907 Lee de Forest added a grid, or control, electrode to the Fleming valve and made the highly sensitive relay device now known as the triode, or three-element tube. A very large amount of research in vacuum techniques, electron-emitter materials, and circuit design was required before the triode could be fully exploited. These developments included pumps capable of really high vacuums, thoriated-tungsten and oxide-coated emitters, and circuit and tube designs suited to a wide range of applications.

About 1926 the difficulties due to internal element capacities, which prevented useful amplification at frequencies higher than about 1 Mc, caused Dr. A. W. Hull to add a fourth electrode, the screen grid, to the triode, making the tetrode, or four-element tube. A few years later he added a third grid, or suppressor, to the tetrode, making the pentode, to overcome certain deficiencies in tetrode operation. This latter type has properties which make it highly desirable over a very broad frequency range, and it is now one of the most generally used tube types.

Many other specialized forms are available and will be discussed in connection with their applications.

**1-3. Gas Tubes.** Owing to poor vacuum techniques much early electronic work was done with tubes having considerable gas present. The properties of a gaseous discharge, wherein the presence of positive gas ions could neutralize the negative space charge of the electrons, were early recognized and rectifiers using mercury vapor developed.

About 1928 the *thyratron* and *grid-glow* tubes, three-electrode gas control tubes, made their appearance. These tubes made possible control of considerable amounts of electric power with high efficiency. This development was followed in 1932 by the discovery of controlled ignition on a mercury-pool cathode, resulting in the *ignitron.* This tube was capable of handling and controlling currents even higher than those possible with the thyratron. These two types have been largely responsible for the application of electronics in the power field.

**1-4. Solid-state Developments.** There exists a large class of materials which have electrical properties falling between metallic conductors and insulators. These materials are called *semiconductors* and show inter-

esting electrical properties which are largely due to a shortage of current carriers. These materials exhibit negative temperature coefficients of resistance, and it is found that they may conduct by use of negative carriers or electrons or by positive carriers (absence of electrons) called *holes*. Silicon, germanium, selenium, and titanium compounds are a few of the materials exhibiting these unusual properties.

Crystal rectifiers employing some of these materials have been in use in radio since the beginning of the century. Recently it has been discovered that control or relay properties are possible, leading to triode and tetrode forms of a semiconductor control device known as the *transistor*.

Because of the lack of need for a heated electron source and possible long life as a result, these devices are of considerable interest for various industrial applications.

## REFERENCES

1. Millikan, R. A.: "Electrons (+ and −), Protons, Photons, Neutrons, Mesotrons, and Cosmic Rays," University of Chicago Press, Chicago, 1947.
2. Stranathan, J. D.: "The Particles of Modern Physics," The Blakiston Division, McGraw-Hill Book Company, Inc., New York, 1942.
3. Crowther, J. A.: "Ions, Electrons, and Ionizing Radiations," 7th ed, Longmans, Green & Co., Inc., New York, 1939.
4. Tonks, L.: Electrical Discharge in Gases, *Trans. AIEE*, **53**: 239 (1934).
5. Bardeen, J., and W. H. Brattain: The Transistor, a Semiconductor Triode, *Phys. Rev.*, **74**: 230 (1948).
6. Becker, J. A., and J. N. Shive: The Transistor—A New Semiconductor Amplifier, *Elec. Eng.*, **68**: 215 (1949).

CHAPTER 2

# PHYSICAL PHENOMENA IN ELECTRON TUBES

Electron tubes are nonlinear-circuit devices and operate as they do because of a number of physical phenomena which occur inside the tube, such as emission of electrons, the action of electrons in fields, limitation of current by space charge, the effects of the initial velocities of the emitted electrons, the energy delivered to the electrodes, and ionization of any gas present.

These physical processes impose limits on the practical operating ranges of electron tubes. Since an understanding of tube properties, and the proper application of tubes, are both so dependent on these physical processes, it seems desirable to survey the phenomena inherent in electron tubes in this chapter.

**2-1. Action of Electrons in Electric Fields.** The movement of electrons and other charges through electron tubes is due to forces applied to these particles by the electric or magnetic fields present. In considering these forces and the motions of charge in vacuum tubes, it is usual to assume that the number of electrons per unit volume, or the *charge density*, is sufficiently small so that the repulsion of one electron on another may be neglected. It is also customary to assume a vacuum sufficiently good so that there will be no collisions with gas atoms or ions, and that gravitational forces are small with respect to forces exerted by the fields. All these assumptions can be well supported in practice.

Electric-field forces are used to accelerate electrons in most of the usual vacuum tubes. The intensity of an electric field is defined as the force (newtons) per unit positive charge (coulombs), or

$$\mathcal{E} = \frac{\mathbf{f}}{q} \qquad \text{volts/m} \tag{2-1}$$

If an electron having the electronic charge $-e$, of magnitude $1.60 \times 10^{-19}$ coulomb, is placed in this field, the force acting on the electron is

$$\mathbf{f}_e = -e\mathcal{E} \qquad \text{newtons} \tag{2-2}$$

the negative sign indicating that the force is opposite in direction to that of the field, or is directed toward the positive electrode.

From Newton's law, $\mathbf{f} = m\mathbf{a}$, the acceleration experienced by the electron in the field is

$$\mathbf{a} = \frac{d^2x}{dt^2} = \frac{-e\mathcal{E}}{m} \tag{2-3}$$

for velocities at which relativistic variation of mass may be neglected. The ordinary laws of particle dynamics then apply. If the field is varying with time, it should be noted that the acceleration will be a function of time.

The work done on the electron by the force in moving the electron between two points, $A$ and $B$, in the field, is

$$W = \int_A^B f_e \cos \alpha \, ds = -\int_A^B e\mathcal{E} \cos \alpha \, ds$$

where $\alpha$ is the angle between the direction of the force and the positive direction of movement $ds$.

This value of work is equal to the kinetic energy acquired by the electron; so

$$W = \frac{m(v^2 - v_o^2)}{2} = -e \int_A^B \mathcal{E} \, ds \tag{2-4}$$

It should be noted that the negative of the integral of the field intensity between points $A$ and $B$ is equal to the potential $V$ between these points, so that

$$W = \frac{m(v^2 - v_o^2)}{2} = Ve \qquad \text{joules} \tag{2-5}$$

Thus the energy acquired by an electron is equal to the charge on the electron times the potential difference through which the electron has moved.

If the initial velocity $v_o$ is zero, then Eq. (2-5) yields a useful result,

$$v = \sqrt{\frac{2Ve}{m}} = 5.93 \times 10^5 \sqrt{V} \tag{2-6}$$

This is the velocity reached by an electron starting at rest and traveling to a point $V$ volts above the potential of the starting point.

Equation (2-5) gives rise to a unit of energy frequently employed in electronic work. This is the electron volt and is the amount of energy received by an electron in rising through a potential of 1 volt, or

$$1 \text{ ev} = 1 \times 1.60 \times 10^{-19} = 1.60 \times 10^{-19} \text{ joule}$$

The electron volt is a much used unit in stating the energy of particles.

When speaking of particles of very high energy, the abbreviations Mev and Bev are used, meaning "million electron volts" and "billion electron volts," respectively.

**Example.** An electron with initial velocity due to an energy of 10 ev enters the electric field between two parallel plates at $A$ in Fig. 2-1 and at an angle of 30° to the electrodes. The electron is to strike point $B$ on the upper plate, 0.5 cm to the right of a vertical line through $A$. Find the potential required on the upper plate, the time of travel, the final velocity, and the energy given to the upper electrode.

Initial velocity:

$$v_o = \sqrt{\frac{2Ve}{m}} = \sqrt{2 \times 10 \times 1.76 \times 10^{11}}$$
$$= 1.87 \times 10^6 \text{ m/sec}$$

Velocity components:

$$v_{oy} = 1.87 \times 10^6 \times \sin 30° = 0.935 \times 10^6 \text{ m/sec}$$
$$v_{ox} = 1.87 \times 10^6 \times \cos 30° = 1.62 \times 10^6 \text{ m/sec}$$

The horizontal component $v_{ox}$ is not changed by the field, thus the time to travel 0.5 cm (0.005 m) horizontally, and also to reach the upper plate, is

$$t = \frac{0.005}{1.62 \times 10^6} = 3.09 \times 10^{-9} \text{ sec}$$

If upward acceleration is considered positive, then acceleration due to the electric field is

$$a = \frac{e\mathcal{E}}{m} = 1.76 \times 10^{11}\mathcal{E}$$

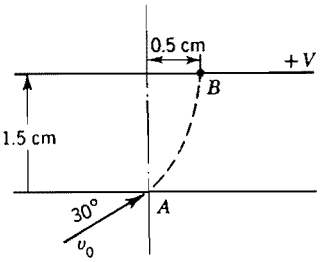



FIG. 2-1. Trajectory of electron.

Distance traveled upward $= y = \frac{at^2}{2} + v_{oy}t = 0.015$ m.

$$0.015 = \frac{1.76 \times 10^{11}\,\mathcal{E}\,(3.09 \times 10^{-9})^2}{2} + 0.935 \times 10^6 \times 3.09 \times 10^{-9}$$

$$\mathcal{E} = \frac{12.1 \times 10^{-3}}{8.40 \times 10^{-7}} = 1.44 \times 10^4 \text{ volts/m}$$

Potential $V = 1.44 \times 10^4 \times 0.015 = 216$ volts positive on upper plate.

$$\text{Upward velocity} = v_y = at + v_{oy}$$
$$= 1.76 \times 10^{11} \times 1.44 \times 10^4 \times 3.09 \times 10^{-9} + 0.935 \times 10^6$$
$$= 8.78 \times 10^6 \text{ m/sec}$$
$$\text{Total final velocity} = \sqrt{v_y^2 + v_{ox}^2}$$
$$= \sqrt{(8.78 \times 10^6)^2 + (1.62 \times 10^6)^2} = 8.93 \times 10^6 \text{ m/sec}$$

Energy delivered to the upper plate on impact is

$$W = \frac{mv^2}{2} = \frac{9.11 \times 10^{-31}(8.93 \times 10^6)^2}{2} = 3.62 \times 10^{-17} \text{ joule}$$

*Check.* Total potential through which the electron has moved is 10 + 216 = 226 volts. Total energy of the electron is therefore 226 ev, and

$$226 \text{ ev} = 226 \times 1.60 \times 10^{-19} = 3.62 \times 10^{-17} \text{ joule}$$

**2-2. Current Flow in Space.** When an electron is accelerated by the electric field, as in the configuration of Fig. 2-2, the law of conservation

of energy requires that the increase in kinetic energy of the electron must be equal to the energy delivered to the electron. In this case the source of the energy is the battery in the circuit. The electron, in moving toward the positive electrode, induces a positive charge on that electrode, or tends to force an electron into the external circuit. At the same time it induces a slight negative charge, or tends to pull an electron into the negative plate. As a result there is a small movement, or drift, of negative charge through the battery from the positive electrode to the negative electrode.

The movement of this negative charge constitutes a conventional electric current in the opposite direction, or out of the positive-battery terminal, and this represents energy supplied by the source. When the electron reaches the positive plate, it has succeeded in forcing a charge redistribution which has left the positive plate deficient by just one electron and this is supplied by the arriving electron. The system is then once more in electrical balance, and current flow stops. The electrical energy which was supplied to the electron and converted to kinetic energy is given up as heat and perhaps X rays upon striking the electrode.

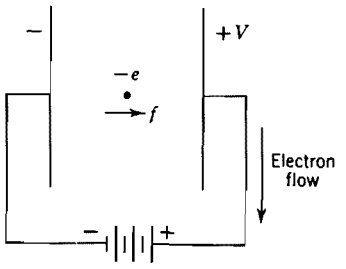



FIG. 2-2. Electron in transit in electric field.

The source of potential has contributed energy as given by

$$W = Vq = \int_{t_o}^{t} Vi\,dt = \frac{m(v^2 - v_o^2)}{2} \tag{2-7}$$

Taking the time derivative,

$$Vi = mv\frac{dv}{dt} = mva = -ve\mathcal{E}$$

$$i = -\frac{ev\mathcal{E}}{V} \quad \text{amp} \tag{2-8}$$

This is conventional current, which accounts for the minus sign. It should be noted that velocity $v$ is the result of the acceleration of the field $\mathcal{E}$, or is the velocity component parallel to the field direction.

An electron may contribute energy to an electric circuit by being shot into a retarding electric field. Since the electron will then be moving toward the negative electrode, an induced positive charge will be set up there, causing an electron flow in the external circuit from negative to positive. This may be seen as a conventional current through the source from positive to negative and represents a contribution of energy to the

source emf. The electron loses kinetic energy and will be slowed down as a result.

**Example.** Find the current flowing in the external circuit at the instant just before impact, for the electron in the example of Sec. 2-1.

The external circuit conventional current is due only to the electron-velocity component parallel to the electric field or toward the electrodes, so that

$$i = \frac{ev\mathcal{E}}{V} = \frac{1.60 \times 10^{-19} \times 8.78 \times 10^{6} \times 1.44 \times 10^{4}}{216}$$
$$= 9.45 \times 10^{-11} \text{ amp}$$

This conventional current appears to flow out of the positive-battery electrode; therefore the battery is supplying energy to the electron.

**2-3. Action of Electrons in Magnetic Fields.** The force on a current in a conductor in a magnetic field is

$$\mathbf{f}_m = \mathbf{Bil} \sin \theta \qquad \text{newtons} \tag{2-9}$$

where **B** = magnetic-flux density, webers/sq m, and $\theta$ = angle between the direction of the positive current and direction of flux. Thus the force is proportional to the component of i normal to **B**. This force acts perpendicular to the plane containing **B** and i.

Since the current $i$ is due to movement of electrons, it may be expressed in terms of charge passing a point per second as

$$i = -\frac{ne}{t} \tag{2-10}$$

where $n$ is the number of electronic charges $e$ passing in time $t$. Then

$$\mathbf{f}_m = -\frac{\mathbf{B}nel \sin \theta}{t}$$

where $l/t$ is the velocity $v$ of the electrons. Then, for a single electron in the field

$$\mathbf{f}_m = -\mathbf{Bev} \sin \theta \qquad \text{newtons} \tag{2-11}$$

In the notation of vector analysis this force equation appears as

$$\mathbf{f}_m = -e(\mathbf{v} \times \mathbf{B})$$

where the cross product is indicated. This force is always at right angles to the instantaneous velocity of the electron. The negative sign appears since the force will be opposite to that experienced by a conventional current of positive charges flowing in the direction of $v$.

The electron has components of velocity parallel and normal to the direction of **B**. These are $v_p = v \cos \theta$, and $v_n = v \sin \theta$. Thus Eq. (2-11) becomes

$$\mathbf{f}_m = -\mathbf{Bev}_n \tag{2-12}$$

In finding the direction of the force it is convenient to think of the moving electron as creating a magnetic field. It is customary to use the fingers of the right hand to determine the direction of field surrounding a conventional current whose direction is indicated by the thumb. This rule may be modified for moving electrons if the left hand is used, the thumb showing the direction of movement of the electron, and the fingers then indicating the surrounding magnetic field. In Fig. 2-3 the electron sets up a magnetic field around its path, directed out of the paper on the left of the electron. By the usual rules for forces in magnetic fields the force will be away from the side on which the field intensities add and thus will be to the right as indicated.

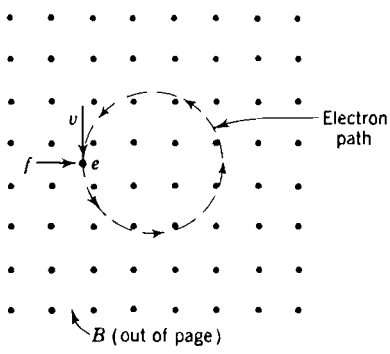



FIG. 2-3. Path of electron in a magnetic field $B$, directed out of page.

Being always perpendicular to the instantaneous velocity, the force $\mathbf{f}_m$ can perform no work on the electron, and the electron velocity $\mathbf{v}_n$ remains constant. Since $\mathbf{v}_n$ and $\mathbf{B}$ are constant, then $\mathbf{f}_m$ is constant and produces an acceleration at right angles to the velocity $\mathbf{v}_n$. This causes the electron to move in a circular path at constant speed equal to $\mathbf{v}_n$.

The centrifugal force must equal the magnetic force, or

$$\frac{mv_n^2}{r} = Bev_n$$

from which the radius of the circular path is

$$r = \frac{mv \sin \theta}{Be} \quad \text{meters} \tag{2-13}$$

Using well-known laws of dynamics, it is possible to show that the angular velocity is

$$\omega = \frac{v_n}{r} = \frac{Be}{m} \tag{2-14}$$

and that the period, or time for one revolution around the circular path, is

$$T = \frac{2\pi m}{Be} \quad \text{sec} \tag{2-15}$$

This result is independent of velocity, a high-velocity electron traveling round a large-radius circle in the same time that a low-velocity electron travels round a small-radius circle.

**Example.** An electron, having an energy of 20 ev, enters a magnetic field of density 0.02 weber/sq m, at an angle of 30° to the field. Find the position of the electron after it has made one revolution in the field.

The initial velocity is

$$v_o = \sqrt{\frac{2Ve}{m}} = \sqrt{2 \times 20 \times 1.76 \times 10^{11}} = 2.65 \times 10^6 \text{ m/sec}$$

Then
$$v_p = v_0 \cos 30° = 2.65 \times 10^6 \times 0.866 = 2.29 \times 10^6 \text{ m/sec}$$
$$v_n = v_0 \sin 30° = 2.65 \times 10^6 \times 0.500 = 1.32 \times 10^6 \text{ m/sec}$$

The time of travel for one revolution is

$$T = \frac{2\pi m}{Be} = \frac{2\pi}{0.02 \times 1.76 \times 10^{11}} = 1.78 \times 10^{-9} \text{ sec}$$

During this time the electron translates parallel to the field due to the velocity $v_p$, so that it takes a position $L$ m along the flux lines from the entering point, where $L$ is

$$L = v_p T = 2.29 \times 10^6 \times 1.78 \times 10^{-9}$$
$$= 4.07 \times 10^{-3} \, m = 0.407 \text{ cm}$$

**2-4. Electrons in Metals.** In a metal a great many atoms are very closely packed together in a regular lattice, or pattern, called a crystal. Electrons associated with each nucleus have varying amounts of total energy, divided between kinetic energy of motion and potential energy of position in the electric field of the nucleus. As in all atoms the electrons can exist only in discrete energy states, and only one electron can exist in a given energy state, or level, at a given time.

In a free atom these energy levels differ by appreciable amounts of energy, but in a solid the effect of the numerous closely packed neighboring atoms is to greatly increase the number of allowable levels and thus to decrease the difference in energy between them, and these energy differentials are very small.

Electrons having the lowest amounts of energy are rather closely held by the field of the positive nucleus, but other electrons with higher energies are able to move to energy levels farther away. In fact, many of these electrons have sufficient energy to move so far from their own nucleus that the nuclear fields of adjacent atoms overlap and cancel, and a given electron may be quite uncertain as to which atom it belongs to. These higher-energy electrons then drift in a random manner through the metal and are called *free electrons*.

If a voltage is applied across a portion of a conductor, an electric field will be created within the conductor. The free electrons are acted upon by forces due to this field, and a uniform drift in one direction is superimposed on their random velocities. This drift constitutes a charge movement, or a current, through the conductor.

The distribution of random energy among the electrons in a metal has been studied, and a function known as the *Fermi-Dirac distribution* is

believed to predict the number of electrons $dN$, having a value of energy in a small range lying between $E$ and $E + dE$, at an absolute temperature $T$. This function is

$$dN = \frac{c\sqrt{E}\,dE}{\epsilon^{(E-E_M)e/kT} + 1} = F\,dE \tag{2-16}$$

where $c$ = constant of proportionality = $6.82 \times 10^{27}$ electrons/m/ev$^{3/2}$
$E$ = energy, ev
$k$ = Boltzmann's constant = $1.38 \times 10^{-23}$ joule/°K
$T$ = temperature, °K (°C abs)
$E_M$ = see below

At absolute zero temperature this distribution predicts the parabolic curve of Fig. 2-4. This curve shows that the energy of the highest energy electron at 0°K is given the value $E_M$. Theory also predicts that

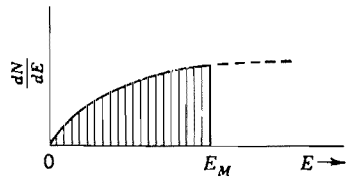



FIG. 2-4. Fermi-Dirac energy distribution for a metal at 0°K.

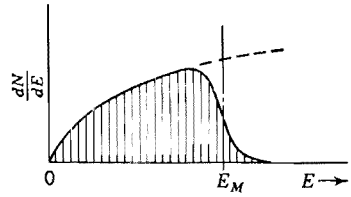



FIG. 2-5. Electron energy distribution for tungsten at 3000°K.

all possible energy levels below $E_M$ are *filled*, or occupied, by electrons at absolute-zero temperature, this being indicated by the cross-hatching in the figure. Thus absolute zero may be defined as that temperature at which all the electrons are in their lowest possible energy states.

As the temperature of the metal is raised, some of the thermal energy put into the metal is given to the higher-energy, or free, electrons and these electrons are enabled to move to higher energy levels, above $E_M$. However, since most of the heat energy contributed to the metal is taken up in thermal vibration of the atoms and only a small amount is available to increase the energy of some of the electrons, then the distribution of energy among the electrons is only slightly affected, becoming as in Fig. 2-5 for a temperature of 3000°K, for tungsten.

The value $E_M$ is a characteristic of the material and may be computed from

$$E_M = \left(\frac{3N}{2c}\right)^{2/3} = 3.64 \times 10^{-19} N^{2/3} \tag{2-17}$$

where $N$ is the number of valence electrons per unit volume of the metal. Since $N$ is a function of the particular material, then $E_M$ has values dependent on the material.

The $E_M$ energy level is sometimes called the *Fermi characteristic energy*

of a material and serves as a measure of the relative level of electron energies in a material. For metallic conductors $E_M$ always occurs at the top of a group, or *band*, of occupied, or filled, electron energy levels at 0°K. In semiconductors and insulators the situation is changed, and this will be discussed later.

**2-5. The Surface-energy Barrier and Electron Emission.** In order to free an electron from a material, or to cause it to be emitted from the surface, it is found that the electron must have an amount of energy, or an outward-directed velocity, sufficient to overcome binding forces which exist at the surface of the material. Thus, while the previous section discussed the distribution of energy among the electrons of a metal, for emission it is necessary to consider only those electrons having components of velocity normal to the surface.

When only the electrons having outward-directed velocities are considered, the distribution of energy among them becomes as in Fig. 2-6. Since the curve for the elevated temperature $T°$ approaches the axis asymptotically, there will be some electrons with high outward velocities. If the amount of energy required for an electron to escape from the surface is $E_B$, then this value may be indicated on the abscissa, or energy scale, and it can be seen that for the higher temperatures there can be electrons in the metal having sufficient outward velocity to escape. These electrons constitute the emitted current.

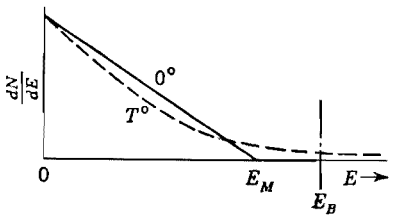



FIG. 2-6. Distribution of energy among electrons having velocity components normal to the surface.

From the figure it can be seen that energy must be contributed to the highest-energy electrons at 0°K if they are to have outward velocities sufficient for emission. In the paragraph above it was assumed that this extra energy was supplied to the electrons in the metal by raising its temperature, or by contributing thermal energy to the electrons. It is found that the required extra energy may be supplied by several different means, and the particular form of electron emission is named accordingly. Thus, if thermal energy is supplied, the process is called *thermal emission.* If the needed energy is supplied by light or radiant means or by photons striking the surface, the resultant electron flow is called *photoemission;* if the energy is mechanically supplied by bombardment of the surface with a primary beam of electrons or ions, then the process is *secondary emission;* and if supplied by the forces exerted by an intense attractive electric field, the emission is called *high-field,* or *autoelectronic,* emission.

The surface-energy barrier $E_B$ is dependent on the material, and on its surface condition, whether clean, coated with another metal, or with a coating of adsorbed gas atoms. From Fig. 2-6 it can be seen that the extra energy which must be contributed to permit emission is equal to the energy difference between $E_B$, the barrier energy, and $E_M$, the energy of the highest-energy electrons at 0°K. This amount of energy is called the *work function* of the material. It is usually expressed in electron volts and given the symbol $E_W$, so that

$$E_W = E_B - E_M \tag{2-18}$$

Because of the dependence of the work function on surface condition and cleanliness, it is difficult to obtain consistent experimental measurements of the quantity. Those given in Table 1 are believed representative of the best available data.

To determine the current which can be emitted from a given unit surface at a given temperature, it is necessary to return to the energy distribution for $T°$ in Fig. 2-6. It is then necessary to determine the equation of the distribution for a temperature $T$, integrate this equation from $E_B$ to infinity to obtain the number of electrons having energy sufficient to pass through the surface, and multiply by the average electron velocity and the charge per electron. The equation for emitted current resulting is usually given as

$$I = A_o S T^2 \epsilon^{-b_o/T} \quad \text{amp} \tag{2-19}$$

where $A_o = 4\pi mek^2/h^3$ = universal proportionality constant, amp/m/deg$^2$
$S$ = emitting surface area, sq m
$T$ = temperature of surface, °Kelvin
$b_o = cE_W/k = 11{,}600E_W$
$k$ = Boltzmann's constant = $1.380 \times 10^{-23}$ joule/°K

This expression is known as *Richardson's*, or *Dushman's, equation* after two men who independently derived the relation. Dushman showed that $A_o$ should be a universal constant having the value 120.4 for all materials, but experiment finds considerable variation, the reason being unknown but thought to be connected with surface condition.

The Dushman equation predicts the amount of current emitted from any metal of known properties at a temperature $T$. It serves as a design relation for thermal-emission devices or thermionic tubes and also points out a current limit which is imposed on their operation. This maximum current which can be obtained from a thermal electron source is said to be the *temperature saturation value.*

The Dushman equation shows a very steep relation between temperature and emitted current. Thus electron tubes must be operated in a fairly limited temperature range, the lower limit fixed by sufficient emis-

TABLE 1. VALUES OF THE EMISSION CONSTANTS FOR CERTAIN METALS

| Material | $A_o$, amp/sq m/deg$^2$ | $b_o$ | $E_W$, ev | Melting point, °K |
|---|---|---|---|---|
| Carbon | $60.2 \times 10^4$ | 54,500 | 4.7 | |
| Cesium | $16.2 \times 10^4$ | 21,000 | 1.81 | 301 |
| Molybdenum | $60.2 \times 10^4$ | 49,900 | 4.3 | 2895 |
| Nickel | $26.8 \times 10^4$ | 32,100 | 5.0 | 1725 |
| Platinum | $32 \times 10^4$ | 61,700 | 5.32 | 2047 |
| Tantalum | $60.2 \times 10^4$ | 47,600 | 4.1 | 3123 |
| Thorium | $60.2 \times 10^4$ | 39,400 | 3.4 | 2118 |
| Tungsten | $60.2 \times 10^4$ | 52,400 | 4.52 | 3655 |

sion and the upper limit determined by evaporation or melting of the emitting material. A curve of emitted current as a function of temperature is shown in Fig. 2-7 for tungsten in its normal operation range.

The emission constants $A_o$ and $b_o$ may be experimentally determined if Dushman's equation is rewritten as

$$\log \frac{i}{T^2} = \log A_o S - \frac{0.4343 b_o}{T} \quad (2\text{-}20)$$

If the term $\log i/T^2$ is plotted against $1/T$, the result should be a straight line with slope of $-0.4343b_o$ and an intercept of $\log A_o S$, from which $b_o$ and $A_o$ may be computed.

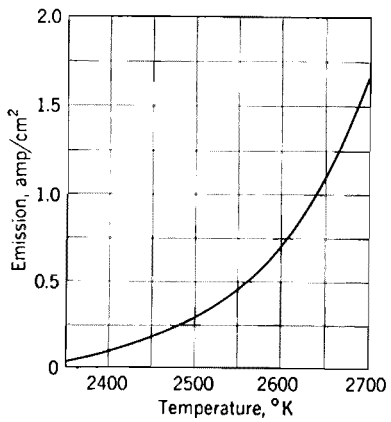



FIG. 2-7. Emitted current density vs. temperature for tungsten.

**2-6. Thermionic-cathode Materials.** Since the electron emitter in an electron tube is ordinarily operated as the negative electrode, it is called a *cathode*. The requirements for materials suited to thermionic-cathode, or emitter, application are such as to limit the field of choice. While a low work function and mechanical stability at high temperatures are the desirable features of a material, they are properties which are not combined in available materials.

*Tungsten* was the material chosen for the early forms of thermionic tubes. While its work function is not low, its high melting point permits its operation at temperatures around 2500°K, and reasonable electron emission and mechanical life are obtained. It is subject to poisoning by oxygen, water vapor, and some other gases and must be used in tubes which are well evacuated. To maintain the operating temperature of 2500°K, rather large amounts of heating power are required, and the

*relative emission efficiency* of tungsten is considered as in the range of 4 to 20 ma of emission per watt of heating power.

Because of its low emission efficiency when compared with other materials, tungsten is now used as an emitter in only a few types of rectifier diodes, some high-voltage power tubes, and X-ray tubes.

*Thoriated tungsten* was developed by Langmuir and Rogers in 1913 as an improved electron-emitter material. One or two per cent of thorium oxide is added to the tungsten powder before the wire-drawing process. During evacuation of the tube the oxide is broken down to pure thorium, and some of this diffuses through the wire and forms a monatomic layer of thorium on the surface.

An intense electric field is developed between the adsorbed thorium atoms and the base tungsten, this field reducing the work function and aiding electrons to escape. The relative emission efficiency is between 50 and 100 ma per watt of heating power at operating temperatures approximating 1900°K.

Excess operating temperature will cause evaporation of thorium atoms from the surface faster than the thorium can diffuse to the surface, and temperatures below normal reduce the breakdown rate of the thorium oxide, and insufficient thorium is made available to cover the surface. Close control of the operating temperature ensures a tube life limited only by the reserve supply of thorium present in the material.

Thoriated tungsten surfaces are susceptible to damage from bombardment by gas ions and are used under high-vacuum conditions. Most medium- and large-size transmitting power tubes employ thoriated tungsten wire as a cathode.

*Oxide-coated* emitters were discovered by Wehnelt in 1904 but not generally applied until about 1922. Such a cathode consists of a base of nickel, konel, or other alloys, coated with a layer of oxides, usually of barium and strontium. Under evacuation some of the oxide is broken down to pure metal, and this forms an adsorbed surface layer which is electropositive to the base, aiding in electron emission and reducing the work function.

Oxide coatings are susceptible to oxygen poisoning and to damage by gas-ion bombardment. Consequently they are usually employed only in tubes with operating voltages below 1000 volts, to reduce possible ionic energies. Emission efficiency is usually considered as in the range of 100 to 10,000 ma per watt of heating power at operating temperatures of about 1000°K.

While Dushman's equation constants are given in Table 2, yet oxide materials do not usually show a definite saturation value of current, and pulsed values of 100 amp per sq cm can be obtained.

Oxide materials are ordinarily coated on the base materials by spraying

processes. Because of this ease in manufacturing and because of the low heating power required, oxide cathodes are used in almost all small electron tubes and probably represent more than 95 per cent of all tubes produced.

TABLE 2. CHARACTERISTICS OF USUAL CATHODE MATERIALS

| Material | $A_o$ | $b_o$ | $E_W$ ev | Emission eff, ma/watt | Operating temperature, °K |
|---|---|---|---|---|---|
| Tungsten | $60.2 \times 10^4$ | 52,400 | 4.52 | 4–20 | 2500–2600 |
| Thorium on tungsten | $3.0 \times 10^4$ | 30,500 | 2.6 | 50–100 | 1900–2000 |
| Oxide-coated | $0.01 \times 10^4$ | 11,600 | 1.0 | 100–10,000 | 1000–1200 |

**Example.** A flat disk-shaped oxide cathode with one side coated has a diameter of 3 mm. If the observed emission or saturation current is 5.70 ma, find the operating temperature of the cathode.

$$\text{Surface} = \frac{\pi d^2}{4} = \frac{\pi \times 0.003^2}{4} = 0.00000707 \text{ sq m}$$

$$i = A_o S T^2 \epsilon^{b_o/T}$$

$$0.0057 = 0.01 \times 10^4 \times 7.07 \times 10^{-8} T^2 \epsilon^{-11,600/T}$$

$$T^2 \epsilon^{-11,600/T} = 8.05$$

Taking logarithms to the base 10,

$$2 \log T - \frac{11,600}{T} \times 0.4343 = 0.906$$

$$2 \log T - \frac{5030}{T} = 0.906$$

It is necessary to assume values of $T$ to solve this equation. Since oxide-coated cathodes operate near 1000°K, assume this as a first $T$ value.

$$2 \times 3.0 - 5.03 = 0.906$$

$$0.97 \neq 0.906$$

Thus the initial assumption is high. Try 990°K.

$$2 \times 2.996 - 5.080 = 0.906$$

$$5.992 - 5.080 \neq 0.906$$

Thus 990°K is still very slightly high. Trying 988°K,

$$2 \times 2.995 - 5.091 = 0.904 \cong 0.906$$

The operating temperature is very slightly above 988°K.

**2-7. Thermionic-cathode Types.** There are three general types of thermionic cathodes: *filamentary*, or directly heated, *equipotential*, or indirectly heated, and the *heat-shielded* cathode.

In the directly heated construction a wire, or filament, of tungsten, thoriated tungsten, or oxide-coated nickel is heated by passage of an

electric current, and emission takes place directly from the wire. Owing to the voltage drop produced by the heating current all points on the filament are not at the same potential. Several such types are illustrated in Fig. 2-8.

The indirectly heated form of cathode is made by coating a small sleeve of nickel or other base metal with oxides. An insulated tungsten heater is then inserted in the sleeve, emission occurring from the outer surface of the sleeve while heat is indirectly supplied by current in the heater. The functions of heating and emission are separated and the emitter is entirely equipotential. Various forms of heater wires are used, usually coated with a baked-on alumina or magnesia insulation before insertion in the cathode sleeve.

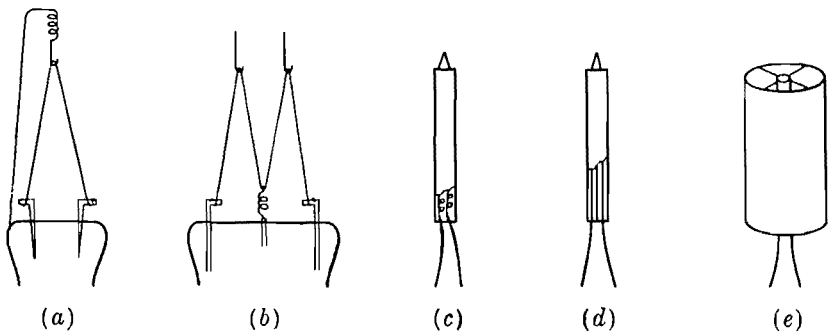



FIG. 2-8. Cathode types. (*a*) Inverted V; (*b*) inverted W; (*c*) spiral heater cathode; (*d*) folded heater; (*e*) heat-shielded cathode.

This construction has considerable thermal inertia, and heat can be supplied by alternating current without cyclic temperature variations which would produce fluctuations in emission. Indirectly heated cathodes are generally employed in receiving tubes for this reason.

A strip of oxide-coated filament may be folded or corrugated to make a heat-shielded cathode, with decreased heat radiation for a given amount of emitting surface. It may then be surrounded by a polished-nickel heat reflector further to reduce the radiation loss. An equipotential heat-shielded form is usually designed with oxide-coated fins radiating from a central sleeve containing the heater wire, the whole surrounded by a nickel heat shield.

The purpose of heat shielding is to reduce the heat radiated from a cathode, or to reduce the power input for a given electron emission, thereby improving the relative emission efficiency.

Since the heat-shielded construction leads to internal pockets and cavities from which it is difficult to extract the electrons and in which the electrons would build up negative space charges and repel further

emission, these cathodes are employed only in gaseous tubes wherein positive gas ions are available to neutralize the pockets of negative charge.

**2-8. Photoemission.** Light, or radiant energy, may be thought of as consisting of chunks, or bundles, of energy, according to the more usual wave theories. It is found that the energy per bundle is related to the frequency of the light by

$$W = hf = \text{quantum energy, joules} \tag{2-21}$$

where $h$ = Planck's constant = $6.624 \times 10^{-34}$ joule-sec. Each bundle of energy is called a *photon*, and the amount of energy is called a quantum.

If the quantum energy carried by a photon is equal to or greater than the work function of a surface, then it is possible for a photon to strike a surface atom and cause emission of an electron. The photon is absorbed, and if its energy was greater than the work function, the excess energy appears as kinetic energy of the emitted electron. Emission occurs in less than $10^{-9}$ sec after light strikes the surface, and since intensity of light is merely a measure of the number of photons arriving, an increase in intensity causes an increase in the emitted current.

The above theory was formally expressed by Einstein as

$$hf = eE_W + \frac{mv^2}{2} \tag{2-22}$$

This equation states that the energy of the impinging photon is transferred to the electron as energy, $eE_W$, to overcome the work function and as kinetic energy of motion of the electron after emission. Each photon does not necessarily produce emission of an electron, many photons being reflected from the surface.

Since $eE_W$ is a constant for a given surface, it can be seen that there will be a minimum value of photon energy, or a threshold frequency, $f_o$, below which emission will not take place. Thus

$$f_o = \frac{eE_W}{h} \tag{2-23}$$

and at this frequency of exciting radiation the current becomes zero.

Equation (2-23) provides a method of measuring the work functions of various materials, by measurement of the threshold photoelectric frequency. Good agreement with thermionic values can be obtained for highly clean surfaces.

Low-work-function materials are the best materials for photoelectric cathodes. Cesium, potassium, rubidium, sodium, and lithium, combined with either oxygen or hydrogen layers, give a choice of operating fre-

quencies or wavelengths throughout the visible spectrum. A typical sensitive layer is formed by cesium deposited as a monomolecular layer on a layer of cesium oxide on a silver surface. Such deposition processes are an art and are not well understood. A photocell is shown in Fig. 2-9.

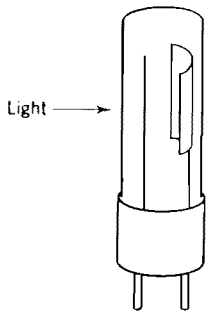



FIG. 2-9. Typical photocell construction.

**Example.** An electron is photoelectrically emitted from a surface with an energy of 0.67 ev, after the surface was exposed to light of 5200 A wavelength. What is the work function of the surface? (1 A = $10^{-10}$ m.)

$$\text{Frequency of light} = f = \frac{c}{\lambda} = \frac{3 \times 10^8}{5.2 \times 10^{-7}} = 5.76 \times 10^{14}$$

$$\text{Then} \qquad E_W = \frac{hf}{e} - \frac{mv^2}{2e}$$

but $mv^2/2e$ is the excess energy in electron volts = 0.67, so that

$$E_W = \frac{6.62 \times 10^{-34} \times 5.76 \times 10^{14}}{1.60 \times 10^{-19}} - 0.67$$

$$= 2.38 - 0.67 = 1.71 \text{ ev} = \text{work function}$$

**2-9. Secondary and High-field Emissions.** Electrons may be given sufficient energy for emission by the mechanical impact of an electron or ion striking the surface. The phenomenon is called *secondary emission*, and the emitted electrons are *secondary* electrons. The particles striking the surface and causing emission are *primary* electrons, or *primary* ions.

Secondary emission can be produced by bombardment of a material, conductor or nonconductor, by a high-energy primary beam of electrons or other particles. The secondary electrons emitted will be attracted to any electrode of higher potential in the region and constitute a flow of charge from the surface, whereas the primary beam carries charge to the surface.

The average number of secondary electrons emitted per primary electron or particle striking the surface is called the *secondary-emission coefficient* and given the symbol $\delta$. Values of $\delta$ for some surfaces may exceed 10.

Secondary emission depends on the nature and work function of the surface. The energies of the emitted electrons are usually less than 20 ev, but the number of secondary electrons, or $\delta$, is a function of the primary energy, usually having a broad maximum at 400 to 600 ev primary energy.

Surface condition, as well as low work function, is important in materials for secondary emission, since the emission is primarily a surface function. High-emission materials are composite surfaces, similar in most respects to those used for photoemission. Cesium–cesium oxide is one of the best surfaces for high $\delta$ value, whereas clean graphite is

one of low $\delta$ value. Values of $\delta$ as functions of primary beam energy are given for several materials in Fig. 2-10.

In some tubes secondary emission is depended on for a source of electrons. In such tubes a coated cathode, similar to surfaces described, is bombarded by gas ions from the space and emits electrons necessary to maintain the discharge. Such cathodes require no external heat source and are frequently called *cold cathodes*, although their operating temperatures may be high owing to bombardment.

If a high positive electric field is applied at the surface of a metal, it is found possible to pull electrons directly out of the material at ordinary temperatures. The electric field lowers the value of $E_B$, the surface barrier energy, and at field intensities of about $10^6$ volts per cm the emission increases to high values. This is *high-field*, or *autoelectronic, emission*. Partial lowering of the energy barrier by the applied field is noted at lower voltages, where it is called the Schottky effect.

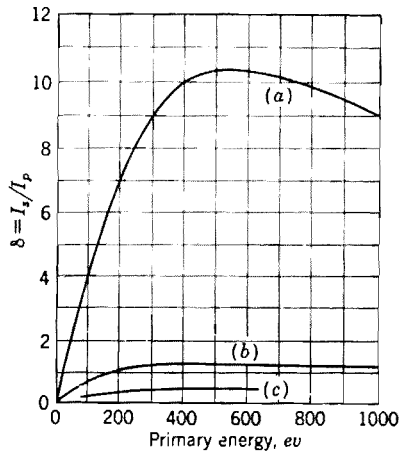



FIG. 2-10. Secondary-emission ratios for (*a*) Cs-CsO-Ag surface, (*b*) nickel, (*c*) carbonized nickel.

The electric field between two closely spaced electrodes, or a point and a plane, or between a fine central wire and a surrounding cylinder may reach high values without high potentials. Special design precautions are necessary in high-voltage tubes for X-ray and rectifier service to prevent such field build-up and possible damage by the high currents associated with high-field emission.

**2-10. Space-charge Effects in the Thermionic Diode.** The electrons emitted from a heated thermionic cathode may have energies in excess of $E_B$, and thus some may be emitted with small outward velocities. The most probable value of this excess energy is given by $2kT/e$, and for an emitter at 1000°K the most probable outward energy is only 0.17 ev. While these emission energies exist, they are usually small compared with the energies which are given to the electrons by the applied potentials in tubes.

If a thermionic cathode is sealed into an evacuated envelope and heated to emitting temperature, electrons will be emitted and ultimately fill the space surrounding the cathode with negative charge. As emission continues, the electric field due to the charges already in the space will repel some electrons back to the cathode. An equilibrium condition

will be reached when the *space charge* becomes sufficient to return electrons to the cathode at the same rate at which they are emitted, giving a net emission of zero.

If a second electrode, or *anode*, is added in the envelope and made positive with respect to the cathode, electrons will be attracted from the space to the anode, reducing the charge in the space and the repelling field, and permitting some of the electrons with highest energy of emission to travel to the anode. Equilibrium will again occur when the net number of electrons leaving the cathode is equal to the number collected by the anode.

If the applied potential be varied and a plot of anode-cathode potential be made against current, a curve as in Fig. 2-11 is obtained. This is a

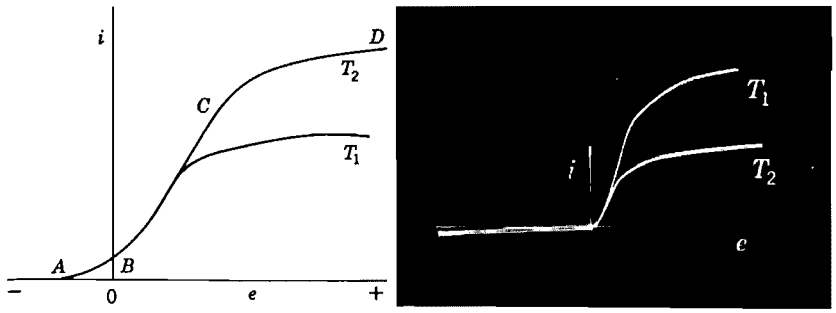



FIG. 2-11. (*a*) Volt-ampere diagram for a diode. (*b*) Oscillogram from FP-400 diode, for two temperatures of a tungsten filament.

volt-ampere diagram for a thermionic *diode*, or two-element tube, and shows that a thermionic diode is a *nonlinear*-electric-circuit element.

The curve divides into three regions. Region *A-B* shows that the emitted electrons have excess energies of emission, since a few are able to overcome the repelling forces set up by a slightly negative anode.

Region *C-D* can be explained as caused by flow of saturation current, as in this region of applied voltage the anode is able to collect essentially all the electrons emitted by the cathode. This fact is further demonstrated by the curves for temperature $T_1$ and for a lower temperature $T_2$, taken from a tungsten filament in Fig. 2-11*b*. Oxide or thoriated tungsten emitters fail to show such pronounced saturation, their emission continuing to increase considerably as the voltage of the anode is raised.

In the region *B-C* the current is limited by the presence of negative space charge, owing to insufficient applied voltage to sweep out all the electrons emitted by the cathode. Equilibrium current for any applied voltage is reached when the net number of electrons permitted to leave the cathode region equals the number collected by the anode, and the space charge is stable.

If the number of electrons leaving the cathode region exceeds the number collected, the space charge will grow until the repelling electric field is increased to the point where it reduces the number leaving the cathode. If the net number emitted is less than the number collected, the space charge will be reduced until the repelling field permits an increase in the net number leaving the cathode.

If a higher anode voltage is applied, the current collected increases, the space charge is reduced, and a new equilibrium condition is created with net emission equal to the increased rate of collection of electrons.

Thus the presence of negative charge in the space between cathode and anode limits the current obtainable at the anode. The effect of the space charge on the distribution of potential in the space is demonstrated in Fig. 2-12. With parallel-plane anode and cathode and without space charge and the cathode cold the distribution of potential is uniform as in Fig. 2-12*a*. With the temperature of the cathode high enough to provide more emission than required by the anode current the negative space charge reduces the potential at all points, and the potential distribution becomes that of Fig. 2-12*b*. A repelling field is present at the cathode surface, as indicated by the negative slope of the potential curve at that point.

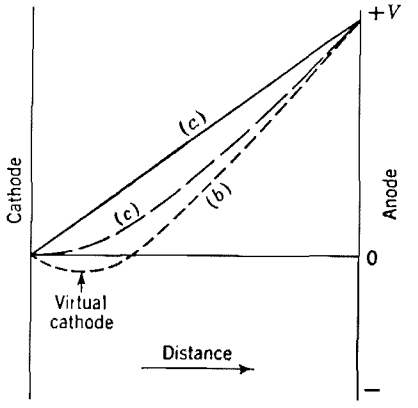



FIG. 2-12. Potential variation in a diode (*a*) without space charge, (*b*) with initial velocities and space charge, (*c*) without initial velocities.

It can be reasoned that the potential minimum or *virtual cathode* and the repelling field are present because of the small initial emission velocities of the electrons. If these velocities were zero, or if the current were just at saturation value, the potential curve would be as in Fig. 2-12*c*, with zero slope and zero field at the cathode surface.

It is possible to obtain an analytical expression for the region *B-C* of the volt-ampere curve. The work is simplified under an assumption of zero initial electron velocity, and experiment shows that this gives a satisfactory theoretical explanation of the results.

Assume infinite parallel-plane cathode and anode surfaces separated a distance $d$, the cathode being heated to provide any required emission. If $n$ is the number of electrons per square meter passing a point per second in the space and $v$ the velocity of electron movement toward the anode,

then the current density, or current per unit area of cathode or anode, is

$$J = -nev = -\rho v \tag{2-24}$$

The velocity of the electrons, neglecting initial energies, is dependent on the electron energy, or

$$\frac{mv^2}{2} = Ve \tag{2-25}$$

where $V$ is the potential at any point $x$ in the space region.

Poisson's equation predicts the rate of change of field intensity in a space filled with electric charge as

$$-\frac{d\mathcal{E}}{dx} = \frac{d^2V}{dx^2} = -\frac{\rho}{\epsilon_v} \tag{2-26}$$

where $\rho$ = the charge density, coulombs/cu m, and $\epsilon_v$ = permittivity of space = $10^{-9}/36\pi$.

Substituting Eqs. (2-24) and (2-25) into Poisson's equation gives

$$\frac{d^2V}{dx^2} = \frac{J}{\epsilon_v \sqrt{2Ve/m}}$$

Solution of this equation gives

$$J = \frac{9\epsilon_v}{4}\sqrt{\frac{2e}{m}}\frac{V^{3/2}}{x^2} \quad \text{amp/sq m}$$

To obtain the current density at the anode at distance $d$ and in terms of the cathode-anode potential $E_b$, this becomes

$$J = 2.34 \times 10^{-6}\frac{E_b^{3/2}}{d^2} \quad \text{amp/sq m} \tag{2-27}$$

This is the *Langmuir-Child*, or three-halves-power law. It shows that, in the region *B-C* of the volt-ampere curve, the current is a nonlinear function of cathode-anode potential and is independent of cathode temperature (sufficient emission assumed). Thus a cathode may be able to supply $i$ amp of emission current; yet the applied potential may be insufficient actually to attract and collect all the emitted current.

Because of the neglect of the initial electron velocities the field at the cathode surface has in effect been assumed zero. This assumption is not strictly true, as shown by Fig. 2-12*b*, but the field at the virtual cathode is zero, so that the assumption regarding electron velocities has in effect moved the cathode surface out a small amount toward the anode. Because the position of the virtual cathode is dependent on cathode temperature, the actual current collected under space-charge-limited conditions is a slight function of cathode temperature.

The Langmuir-Child law is of extreme importance in the electronic field, since almost all vacuum-tube operation occurs under space-charge-limited conditions.

For a tube with cylindrical anode and concentric cathode the space-charge-limited current is given by

$$i_b = 14.6 \times 10^{-6} \frac{l}{r_a} \frac{E_b^{3/2}}{\beta^2} \quad \text{amp} \tag{2-28}$$

where $l$ is anode length and $\beta$ is a function of $r_a/r_k$, the ratio of anode radius to cathode radius. For values of this ratio above 7 the value of $\beta^2$ may be taken as unity with small error.

In a space-charge-limited vacuum tube, the current depends on the three-halves power of the applied potential, and this is found to be a general relation for any geometry. Thus analytic proof of the nonlinear nature of a vacuum diode has been obtained.

**2-11. Current Control by a Grid; the Vacuum Triode.** In 1907 de Forest introduced a wire mesh, or grid, electrode between the cathode and anode of a diode, making the *triode*. The grid is able to exercise electrostatic control of the current to the plate, and thus the triode tube is essentially a smoothly controllable relay.

The grid, being usually maintained a few volts negative to the cathode, and being located in the space-charge cloud of electrons, adds its effect to that of the space charge in controlling the number of electrons passing to the anode, or in controlling the tube current.

Thus the field intensity at the cathode surface is some function of $e_c$, the cathode-to-grid potential. In the previous section it was shown that the field intensity at the cathode surface was also a function of $e_b$, the cathode-anode potential. However, the grid is much closer to the cathode than is the anode, and consequently 1 volt on the grid will produce a field at the cathode equal to that produced by many volts on the anode. The ratio of the anode potential required to produce a given change in cathode-surface field intensity to the grid potential required to produce the same change in field is known as the *amplification factor* of the tube, given the symbol $\mu$.

This definition states that the variation in cathode-surface field produced by a 1-volt change of anode potential is $1/\mu$ times as effective on the current as a 1-volt change in grid potential. The anode current $i_b$ may then be expressed as

$$i_b = F\left(e_c + \frac{e_b}{\mu}\right) \tag{2-29}$$

and this relation finds wide use in the study of the triode.

Experiment shows that the functional relationship expressed by Eq.

(2-29) is approximated by the form

$$i_b = K\left(e_c + \frac{e_b}{\mu}\right)^a \tag{2-30}$$

where $a$ approximates the value $\frac{3}{2}$. The constant of proportionality $K$ is known as the *perveance*, which is expressed as amperes per volt$^{3/2}$.

If the logarithm of Eq. (2-30) is taken, then

$$\ln i_b = \ln K + a \ln\left(e_c + \frac{e_b}{\mu}\right)$$

and a straight-line relation between $i_b$ and $e_c + e_b/\mu$ would be expected on log-log graph paper. Values of $a$ and $K$ can then be found from slope and intercept of the approximate straight line obtained.

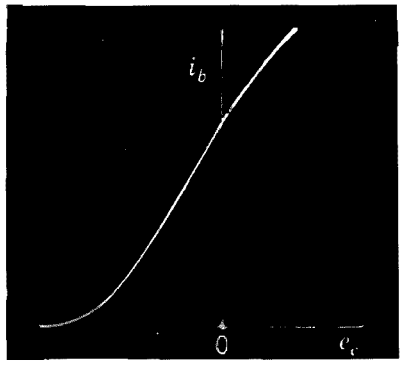



FIG. 2-13. Plot of $i_b$ vs. $e_c$ showing nonlinearity of relation for a triode.

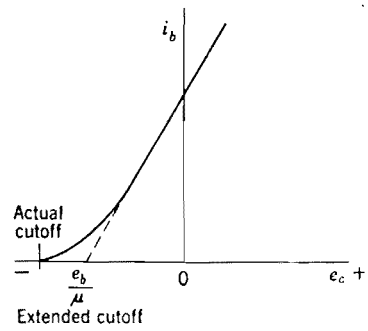



FIG. 2-14. Actual and extended cutoff.

Experimental evidence for a nonlinear relation of the form of Eq. (2-30) between $i_b$ and $e_c$ is presented by the oscillogram of Fig. 2-13, wherein $e_c$ is varied over a considerable range, $e_b$ being constant.

The oscillogram indicates that the plate current goes to zero for some value of negative grid voltage, and the voltage for which this occurs is called the *cutoff grid voltage*. Actually a tube cuts off asymptotically, and cutoff values are not precise. It is occasionally helpful to use the *extended cutoff voltage*, obtained by treating the characteristic as a straight line extended to zero current, as in Fig. 2-14. The extended cutoff value is obtained by setting $i_b$ equal to zero and letting $a = 1$ in Eq. (2-30); then the cutoff value is

$$e_c = -\frac{e_b}{\mu}$$

Since the grid is ordinarily maintained negative, except in a few applications, it will not attract electrons and grid current will be negligible.

When the control grid is made positive, however, grid current will flow and this must be allowed for in tube and circuit design.

**2-12. Energy Losses and Cooling in Vacuum Tubes.** An electron, in being accelerated from a cathode toward a positive anode, is given kinetic energy. Upon impact with the positive electrode the velocity goes to zero, and the energy is converted to heat, and to X rays in some very high-voltage tubes.

The instantaneous power delivered to the anode is $e_b i_b$. In addition the anode structure may intercept a portion of the heat radiated from the cathode. This total power must be removed from the anode by conduction or radiation, since convection in a vacuum is not possible.

Most small tubes are designed to radiate this heat and to keep the anode temperature low, since excessive temperature of the metal parts may cause emission of occluded gases, damaging the vacuum, and softening of the glass may occur at temperatures of 400 to 600°C.

The anodes of small vacuum tubes are usually made of blackened nickel or iron to improve radiation. Most medium-power tubes are made with tantalum or molybdenum anodes and are radiation-cooled by operation at high temperatures, the above metals usually reaching a cherry-red to yellow heat, which is permissible because of thorough degassing in manufacture. Graphite is also employed because of its high radiation emissivity but must not be operated at a temperature at which it will show color because of the difficulty of removing all gases in manufacture.

Very large power tubes cannot radiate their heat at practical temperatures, and conduction cooling is required. Such tubes are designed with external copper anodes, water-cooled by being placed in jackets surrounded by circulating water, or air-cooled by having air blown through copper fins attached to the external anode wall.

Since removal of the losses is so critical a matter in vacuum tubes, it is customary to rate tubes in terms of their permissible plate dissipation. The output possible from a given tube is then a function of the circuit efficiency. This is an important departure from the ratings applied to much other electrical apparatus.

**2-13. Ionization in Gases.** Atoms in gases at low pressures are so widely spaced that they can be considered independent of one another, except for occasional collisions due to random thermal motions. The outermost electrons of any atom, while having high energy, are rather loosely bound to their nucleus because of their distance from the nucleus and also because of the shielding effect of the innermost electrons. Thus bombardment of the atoms by passing free electrons may knock out, or detach, one of these outer electrons. The initial electron and the new electron remain free, the atom becomes a positive ion of one unit positive

charge, and the gas is said to be ionized if this event occurs on a large scale.

Detachment of an electron or ionization of the atom requires a fixed amount of energy sufficient to remove the electron completely from the field of the positive nucleus. This amount of energy is called the ionization energy, or when expressed in electron volts it is called the *ionization potential* of the gas. The bombarding electron must supply the ionization energy to knock out, or detach, an electron. If the bombarding electron has less energy, some of this may be used to raise an electron in the atom from a normal lower energy state to a higher energy state; if the bombarding electron has excess energy, then this excess is divided between the initial and detached electrons. A few ionization potentials are given in

TABLE 3. IONIZATION POTENTIALS FOR A FEW GASES

| Gas or vapor | Ionization potential, ev | Discharge color |
|---|---|---|
| Argon | 15.7 | Blue |
| Helium | 24.5 | Yellow |
| Mercury | 10.4 | Blue-green |
| Sodium | 5.1 | Yellow |
| Neon | 21.5 | Orange |

Table 3. It can be seen that the ionization potential is also that potential by which an electron, initially at rest, must be accelerated to have sufficient energy to ionize an atom of a particular gas.

While many positive ions are formed, many others are recombining with electrons to become neutral once more. In addition many atoms have received energy less than that needed for ionization, and electrons have been raised from normal states to higher excited energy levels without removal from the atom. When an electron returns to form a neutral atom, or when an excited electron returns to its normal level, energy is given up and is radiated as a flash of light in which frequency is related to energy by the Planck relation

$$W = hf \qquad \text{joules}$$

The presence of ionization in a gas is thus indicated by a visible glow of a typical color. The cumulative effect of the individual light flashes produces this glow and accounts for the line spectrum produced by a gaseous electrical discharge.

The positive ion is heavy and slow-moving under field forces, but, wherever it is, it is capable of neutralizing the negative space charge



ఓం శ్రీ సాయినాదాయ నమః
ఓం సాయి శ్రీ సాయి జయ జయ సాయి

# సర్వరోగ నివారణీ యజ్ఞమునకు ఆహ్వానము

గురుబంధువులారా **15-2-1995**

ప్రస్తుత కాల పరిస్థితులలోమనము అనేక రకములైన శారీరక, మానసిక బాధలకు గురి అవుతున్నాము. ఆధునికశాస్త్రవేత్తల అభిప్రాయము, పొగ, నీరు యితర వాతావరణ కాలుష్యము యొక్క ప్రభావము వలన ఈ రోగాలు సంభవించుచున్నాయని అభిప్రాయములను వెలిబుచ్చుచున్నారు కొన్ని వ్యాధులు డాక్టర్లకు కూడ అంతుచిక్కని సంఘటనలు కూడా మనము చూస్తున్నాము. అయితే వాటికి పరిష్కారము ఏమిట? అని ప్రశ్నించుకుంటే, ఈ కలియుగంలో సద్గురువులను ఆశ్రయించుటయే మార్గమని ఋషులు, పెద్దలు చెబుతున్నారు.

శ్రీ సాయినాథుని వంటి సద్గురువు ఈ ప్రపంచములోనే ఎవ్వరులేరని వారి చరిత్రద్వార మనకు తెలియుచున్నది. వారు షిరిడిలో అవతరించిన తరుణమే మొదలు రోగులను చేరదీసి, వైద్యము చేయుటయే మొదటి కార్యక్రమముగా చేబట్టి రోగనివారణ చేసినారు. శ్రీ సాయిబాబాగారీ మార్గంలోనే వైద్యము చేబట్టిన శ్రీ వర్మగారికి శ్రీ సాయినాథుడు "రోగనివారణ" అనే హోమాన్ని చేపట్టమని ఆదేశించినారు. శ్రీ సాయి ఆశీస్సులతో ఈ యజ్ఞాన్ని 15-2-1995 బుధవారము ఆరోగ్య సాయి మందిరమున. నాగార్జుననగర్‌లో జరుపుటకు నిశ్చయించడమైనది. కావున భక్తులు యధా విధిగా ఈ హోమములో పాల్గొని ఆరోగ్యసాయి ఆశీస్సులుపొంది సకుటుంబముగా ఆయురారోగ్యాలు పొందవలసినదిగా కోరుచున్నాము.

## కార్యక్రమము

తేది బుధవారము 15-2-1995 రోజున ఉ॥గం॥ 8-00 లకు హోమము ప్రారంభమగును మరియు అన్నదానము జరుగును.

తేది గురువారము 16-2-1995 రోజున సా॥గం॥ 6-00 లకు ఓం సాయి శ్రీ సాయి జయజయసాయి అనునామ సంకీర్తన జరుగును.
రా॥గం॥ 9-00 లకు హారతి, ప్రసాద వినియోగము

గమనిక : పాల్గొన దలచిన భక్తులు 10-2-1995లోగా ఈ క్రింది అడ్రసుకు విచ్చేసి వారి పేర్లను నమోదు చేసుకొనవలయును.

అడ్రసు : ఇ.నెం. 8-3-944/1/B/2,
నాగార్జుననగర్, ఎల్లారెడ్డిగూడ,
హైదరాబాదు-500 873.

ఇట్లు
ఆరోగ్య సాయి మందిరము

లోకా సమస్తా సుఖినోభవంతు

Engineering Mathematics } Grewal

~~510~~ 515 Physics — Gaurav gupta

~~530~~ 515 Elec. Engg — Thereza

due to an adjacent electron. If sufficient positive gas ions are created in a diode, the negative space charge limiting the current is neutralized and a very large current may flow.

Consider a gas enclosed in a container at a pressure of a few hundredths of a millimeter of mercury and with appropriate electrodes sealed in. There will always be a few free electrons present in the gas, presumably created by cosmic rays or other radiation. If a small potential is applied between the electrodes, the electrons will drift toward the anode. As the potential is raised, these electrons may acquire sufficient energy to ionize an atom upon collision. Thus there will exist two electrons and a positive ion, and these will move in the applied field. The two electrons acquire ionizing energy under the action of the field, collide with atoms, and there will then be four electrons and two positive ions. This process is cumulative and will continue until *breakdown* of the gas occurs.

After breakdown occurs, the space becomes filled with a *plasma* of electrons and positive ions and the potential between the electrodes falls to a value approximating the ionization potential of the gas, just sufficient to give the electrons enough energy to ionize and create positive ions.

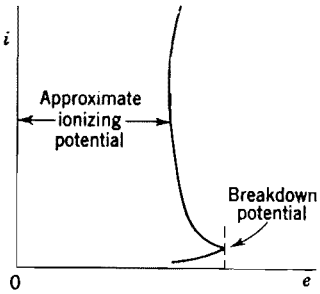



FIG. 2-15. Volt-ampere diagram of a gas discharge.

Additional electrons must be supplied if the discharge is to be *self-maintaining*. In a *glow* discharge the source of the electrons is a cathode, usually coated with high secondary-emission material, and the electrons are emitted as secondary electrons under bombardment of the positive ions on the surface. A colored light glow appears on the cathode surface, and a feature of this cathode glow is that the current density on the cathode surface remains essentially constant, the glow changing in area with changing current demands.

In an *arc* discharge the source of electrons may be a heated cathode, as in a thermionic arc, a pool of mercury, or a solid metal. In the case of the mercury-pool cathode the emission mechanism is not well understood, but it is believed to be a form of high-field emission. In an arc discharge the positive ions are important in neutralizing the electron space charge. Because of this neutralization, the usual gas-tube potentials are low and of the order of the ionization potential of the gas.

A volt-ampere diagram for such a gas discharge is shown in Fig. 2-15. It can be seen that, after the voltage reaches the breakdown value, the potential falls to the maintaining value which approximates the ionization potential. This diagram will be further discussed in Chap. 14.

## PROBLEMS

**2-1.** Two large, parallel metal electrodes are separated a distance of 0.5 cm, with the upper electrode 20 volts positive. An electron is released on the negative electrode with an upward velocity due to an energy of 12 ev. (*a*) Find the time of travel to the upper electrode. (*b*) What are the velocity and energy of the electron on impact?

**2-2.** An electron is released at the center of the negative electrode of Prob. 2-1, with a velocity of $5 \times 10^6$ m/sec, directed at 30° to the electrode. Plot the path followed by the electron, and find (*a*) the time of transit, (*b*) the distance traveled parallel to the electrode, (*c*) the energy delivered to the electrode on impact. The upper electrode is at +20 volts as before.

**2-3.** Plot the external circuit current against time for the electron of Prob. 2-2.

**2-4.** The electron of Fig. 2-16 has an initial velocity due to an energy of 10 ev, directed as shown. Find the potential $E$ to be placed on electrodes 1, 2 which will

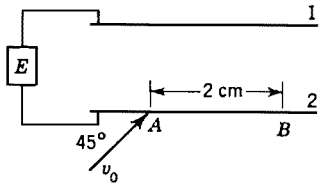



FIG. 2-16

cause the electron to reach point $B$. Plot the value of current flowing in the circuit due to the flight of the electron.

**2-5.** Source $E$ is removed in Prob. 2-4 and a magnetic field substituted. Find the flux density $B$ required to cause the electron, of initial velocity due to 10 ev, to reach point $B$. State the direction of the flux.

**2-6.** An electron with an energy in the horizontal direction of 100 ev enters the field between two horizontal plates 2 cm long and separated 1.5 cm. The lower plate is at +200 volts. Find the direction and flux density, in webers per square meter, of a magnetic field which will just prevent deflection of the electron from its path. Neglect fringing of the fields.

**2-7.** An electron of 50 ev energy enters a perpendicular magnetic field of 0.04 weber per sq m. Find the rpm of this electron and the radius of its path in the field.

**2-8.** A tungsten filament of 1 in. length and diameter of 0.006 in. is operated at 2550°K. (*a*) Calculate the emission current. (*b*) At what temperature must the same filament operate if it is made of thoriated tungsten?

**2-9.** To what value must the work function of a thoriated-tungsten surface be changed to raise the current density to 15,000 amp per sq cm at a temperature of 2550°K?

**2-10.** An oxide-coated emitter is 2.1 in. long, 0.05 in. in diameter, and operated at 1025°K. What is the temperature-limited current? If the temperature falls to 1000°K, by what percentage does the emitted current change?

**2-11.** The filament of an FP-400 tungsten-filament diode is 1.25 in. long, 0.005 in. in diameter. If the temperature-limited current is 24 ma, what is the filament temperature?

**2-12.** (*a*) Calculate the energy carried by photons of yellow light at 5893 A; of blue light of 4210 A; of ultraviolet light of 2537 A.

(*b*) If each of these photons strikes a cesium surface of 1.81 volts work function, find the electron velocities of emission, if emission occurs.

**2-13.** A sodium surface photocell has a threshold wavelength of 6550 A. This surface is struck by green light of 5300 A wavelength. If electrons are emitted, what will be the maximum velocity of emission?

**2-14.** A certain vacuum diode requires 30 volts for a current of 120 ma under space-charge-limited conditions. What voltage will cause a current of 95 ma? What is the maximum current which can flow without the anode loss exceeding 3.2 watts?

**2-15.** A cylindrical-anode diode, with a thoriated-tungsten filament 2.5 cm long and 0.05 cm in diameter, operates at a filament temperature of 1900°K. The anode diameter is 1.0 cm. If the current through the diode is to be 35 per cent of saturation value, what cathode-anode voltage will be required?

**2-16.** A cylindrical-anode diode is to use a thoriated-tungsten filament and to give a saturation current of 300 ma. The anode must be large enough to radiate the plate loss at a rate of 1.5 watts per sq cm, with saturation current flowing at 100 anode volts. The filament temperature is 1925°K. (*a*) What should the anode diameter and length be? (*b*) Find the filament diameter. Assume $\beta^2 = 1.0$.

## REFERENCES


1. Crowther, J. A.: "Ions, Electrons, and Ionizing Radiations," 7th ed., Longmans, Green & Co., Inc., New York, 1939.
2. Page, L., and N. I. Adams: "Principles of Electricity," D. Van Nostrand Company, Inc., Princeton, N.J., 1931.
3. Ryder, J. D.: "Electronic Fundamentals and Applications," Prentice-Hall, Inc., Englewood Cliffs, N.J., 1950.
4. Dushman, S.: Thermionic Emission, *Revs. Mod. Phys.*, **2**: 381 (1930).
5. Jones, H. A., and I. Langmuir: Characteristics of Tungsten Filaments as Functions of Temperature, *Gen. Elec. Rev.*, **30**: 310, 354, 408 (1927).
6. Langmuir, I.: The Electron Emission from Thoriated-tungsten Filaments, *Phys. Rev.*, **22**: 357 (1923).
7. Blewett, J. P.: Properties of Oxide-coated Cathodes, *J. Appl. Phys.*, **10**: 831 (1939).
8. Coomes, E. A.: The Pulsed Properties of Oxide Cathodes, *J. Appl. Phys.*, **17**: 647 (1946).
9. Hughes, A. L., and L. A. DuBridge: "Photoelectric Phenomena," McGraw-Hill Book Company, Inc., New York, 1932.
10. Glover, A. M.: A Review of the Development of Sensitive Phototubes, *Proc. IRE*, **29**: 413 (1941).
11. Woolridge, D. E.: Theory of Secondary Emission, *Phys. Rev.*, **56**: 562 (1939).
12. Langmuir, I.: The Effects of Space Charge and Residual Gases on Thermionic Currents in High Vacuum, *Phys. Rev.*, **2**: 450 (1913).
13. Compton, K. T., and I. Langmuir: Electrical Discharges of Gases, *Revs. Mod. Phys.*, **2**: 123 (1930).
14. Loeb, L. B.: "Fundamental Processes of Electrical Discharge in Gases," John Wiley & Sons, Inc., New York, 1939.
15. Richtmyer, F. K., and E. H. Kennard: "Introduction to Modern Physics," 4th ed., McGraw-Hill Book Company, Inc., New York, 1947.

CHAPTER 3

# THE VACUUM TUBE AS A CIRCUIT ELEMENT

The vacuum tube is in general a nonlinear-electric-circuit device. This nonlinearity imposes restrictions on the analysis of circuits containing vacuum tubes, but two different methods for solution of such circuits are available.

The first method is limited to regions of linear tube operation and replaces the tube with an equivalent linear active circuit, which then can be employed with standard methods of circuit analysis. This method is usually restricted to small input signals wherein operation is over only a small portion of the tube graphical characteristic, which may then justifiably be assumed linear. Because of the ease of use, and because a great amount of tube operation is desired to be under linear conditions, this method is extensively applied.

In the second and more general method the tube is recognized as a nonlinear element and graphical methods of analysis applied, working directly from the curves of tube performance. These methods are most applicable with large input signals and will be so employed here.

**3-1. Nonlinear Volt-ampere Characteristics.** The volt-ampere curve has been mentioned as a source of information on the properties of circuit elements. In particular a curve such as that of Fig. 3-1*a* gives graphical information concerning the resistance of the device whose characteristics are plotted.

Two types of resistance are present in electrical circuits. The first type, *static resistance*, is defined in terms of voltage and current as

$$R = \frac{v}{i} \qquad \text{ohms} \tag{3-1}$$

and, referring to point $A$ of Fig. 3-1*a*, the static resistance is given by

$$R_A = \frac{v_A}{i_A}$$

which is recognizable as the reciprocal of the slope of the dashed straight line from the origin to $A$.

The second type is *variational resistance*, which appears in circuits with

changing voltages and currents and is defined in terms of small (theoretically infinitesimal) changes of voltage and current as

$$r = \frac{dv}{di} \quad \text{ohms} \tag{3-2}$$

The latter concept may be further explained through use of Fig. 3-1*b*. If a resistor has a voltage $v_A$ applied, then the current has a value $i_A$ and operation is at point $A$. If the applied voltage is changed to $v_C$, then the current changes to $i_C$ and operation moves to point $C$ on the curve. For any intermediate value of voltage a particular operating point exists *on the curve*, so that for variation of voltage between points $A$ and $C$, or *for a small a-c component superimposed on a d-c component*, the resistor appears to have an average reciprocal slope, and therefore a resistance, given by

$$r = \frac{v_A - v_C}{i_A - i_C} = \frac{\Delta v}{\Delta i}$$

If the change in voltage is reduced and made infinitesimal, then Eq. (3-2) is obtained. This resistance for small a-c variations of current or voltage is graphically represented by the reciprocal of the slope of the curve at the operating point, the latter point being determined by any d-c component present.

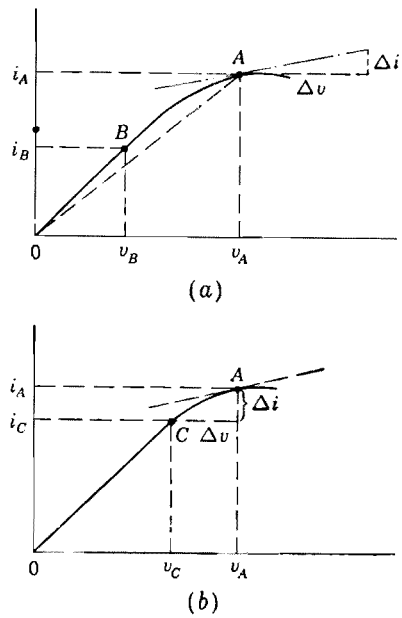



FIG. 3-1. (*a*) Volt-ampere characteristic of a resistance. (*b*) Illustrating variational resistance.

Returning to Fig. 3-1*a*, the static resistance at $A$ was $R_A = v_A / i_A$. The variational resistance is given as

$$r_A = \frac{dv}{di}$$

which is the reciprocal of the slope of the straight dashed line through $A$ tangent to the curve.

The resistance values as measured at point $B$ are identical, whether considered as static or variational, since the slope of the line from the origin to $B$ and the slope of the curve at $B$ are the same.

A resistor whose volt-ampere curve is a straight line is said to be *linear;* one with variable slope is *nonlinear*. Various other properties of graphi-

cal volt-ampere relations as applied to vacuum tubes will be discussed in succeeding sections.

**3-2. Vacuum-triode Current and Voltage Notation.** Vacuum-tube circuits involve a variety of currents and voltages, and designating symbols are required if mutual understanding is to be achieved. The Institute of Radio Engineers (IRE) has adopted certain standard circuit symbols and notation, and these will be employed in this book in so far as possible.

All tube voltages are usually measured with respect to cathode $K$ as reference. This is indicated in Fig. 3-2, which represents a simple form of triode amplifier stage. Small, or lower-case, letters are used to indicate varying or instantaneous quantities, and capital letters denote d-c or rms values. Because of the three triode elements, grid, cathode, and plate or anode, frequent use is made of the letters $g$, $k$, and $p$ as subscripts. Also, in the days before rectified a-c power supplies became available, the batteries supplying the filament, plate, and grid circuits were referred to as $A$, $B$, and $C$ respectively, and these letters have been preserved as subscripts in notation. Thus $c$ and $g$ ordinarily refer to grid-circuit quantities, $b$ and $p$ to plate-circuit quantities, and $k$ and $f$ to the cathode or filament circuit.

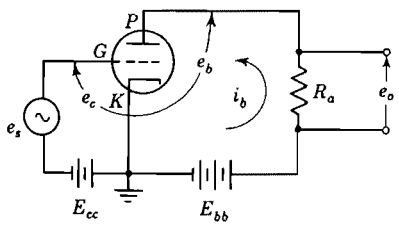


FIG. 3-2. Simple triode amplifier circuit.

Some of the more common symbols to be employed are:

$e_c$ = instantaneous total voltage rise from cathode to grid
$e_g$ = instantaneous value of a-c component of voltage rise from cathode to grid
$E_c$ = average or quiescent value of voltage rise from cathode to grid
$E_{cc}$ = grid-circuit d-c supply-voltage rise from cathode
$e_s$ = instantaneous value of a-c input voltage
$E_s$ = effective or rms value of a-c input voltage
$E_g$ = effective or rms value of the a-c component of cathode-to-grid voltage rise
$E_f$ = filament or heater terminal voltage
$I_f$ = filament or heater current
$i_b$ = instantaneous total anode current
$i_p$ = instantaneous value of a-c component of anode current
$I_b$ = average or quiescent value of anode current
$I_{bs}$ = average value of anode current with signal present. This symbol will be used only when it is wished to emphasize that $I_b$ and $I_{bs}$ are different, as in nonlinear operation of a tube

$I_p$ = effective or rms value of a-c component of anode current
$e_b$ = instantaneous total voltage rise from cathode to anode
$e_p$ = instantaneous value of a-c component of cathode-to-anode voltage rise = $e_o$
$E_b$ = average or quiescent value of anode voltage rise from cathode
$E_p$ = effective or rms value of a-c component of cathode-anode voltage
$E_{bb}$ = anode-circuit d-c supply-voltage rise from cathode
$e_o$ = instantaneous value of load or output voltage, as a rise from cathode to anode = $e_p$

For tubes with multiple grids a system of numerical subscripts is used, as $e_{c1}$, $e_{c2}$, . . . , the number indicating the particular grid meant, with the grid nearest the cathode counted as 1. If no subscript is used, then only one grid is implied or the grid used for control of the anode current by the input signal is meant.

In analysis of amplifiers, as in Fig. 3-2, it is usual to assume that the input signal $e_s$ is sinusoidal and has an rms value $E_s$. For the circuit shown this a-c input is also $e_g$, the a-c component of cathode-grid voltage. This is not always the case in more complex circuits. It is then possible to write $e_g$, the rise of potential from cathode to grid, as the sum of the d-c grid bias $E_{cc}$ and the instantaneous alternating value of input $e_g$. That is,

$$e_c = E_{cc} + e_g \qquad (3\text{-}3)$$

In most circuits the grid-bias voltage $E_{cc}$ is made negative to the grid. When a numerical value is introduced into equations involving $E_{cc}$ the negative sign should also be supplied if appropriate, since $E_{cc}$ was defined as a voltage rise to the grid. With $e_s$ sinusoidal, $e_c$ as a time function is plotted in Fig. 3-3*a*, with $E_{cc}$ negative to the grid.

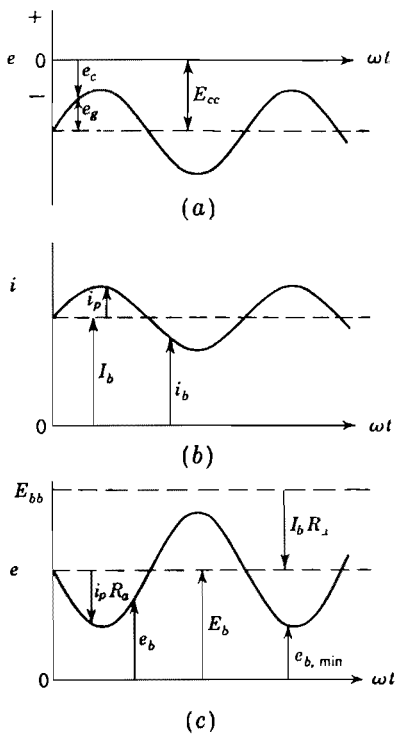



FIG. 3-3. Current and voltage relations.

Because of the control of anode current by the grid voltage, the voltage $e_s$ will cause a variation of anode current. Anode-current conditions

are shown in Fig. 3-3*b*, and

$$i_b = I_b + i_p \tag{3-4}$$

becomes the defining relation.

The quiescent, or no-signal, condition leads to the following voltage relation in the anode circuit,

$$E_b = E_{bb} - I_b R_a \tag{3-5}$$

which is modified by the appearance of $i_p$ when the input $e_s$ is present. The instantaneous cathode-anode voltage then is

$$e_b = E_{bb} - (I_b + i_p)R_a$$

from which the varying component of the output-voltage rise is

$$e_o = -i_p R_a = e_b - E_b \tag{3-6}$$

or

$$e_b = E_b + e_o$$

and these relations are plotted in Fig. 3-3*c*.

It should be noted that the plate current is in phase with the grid voltage. As the anode current increases, the load voltage also increases in magnitude, and thus the tube voltage decreases, since

$$(I_b + i_p)R_a + e_b = \text{constant} = E_{bb}$$

Thus grid voltage and tube voltage are out of phase.

**3-3. Triode Graphical Characteristics and Triode Parameters.** A convenient presentation of performance characteristics of any nonlinear-circuit element is the graphical form. It has already been shown that the vacuum tube is, in general, a nonlinear device and that its characteristics were related as

$$i_b = f(\mu e_c + e_b)$$

in Eq. (2-29). The characteristics of a triode are a function of three variables $i_b$, $e_c$, and $e_b$, and if plotted these characteristics would appear as a surface. For convenience in representation on a plane it is customary to employ the three projections of this surface on the coordinate planes, giving three families of characteristic curves.

If the projection on the $i_b$, $e_b$ plane is chosen, the *plate characteristics* will be obtained, showing the variation of plate current with plate voltage for various grid-voltage values. A value of $e_c$ is chosen and kept constant for each curve of the family. Typical triode plate characteristics are shown in Fig. 3-4*a* and *b*.

Since these curves are volt-ampere plate-circuit plots, the slope of the curves represents the reciprocal of a resistance. As was reasoned in Sec. 3-2, the variational resistance is of importance for a-c or small vari-

ations of current, and this is the usual manner of tube operation in amplifiers. Thus the reciprocal of the slope of the curve at the operating point $Q$ in Fig. 3-4*a* is of interest and is

$$\frac{1}{\text{Slope}} = r_p = \frac{\partial e_b}{\partial i_b}\bigg]_{e_c = k} \tag{3-7}$$

and this is the defining relation for $r_p$, the *plate resistance to variations in current.* Its reciprocal, *plate conductance,* is given the symbol $g_p$.

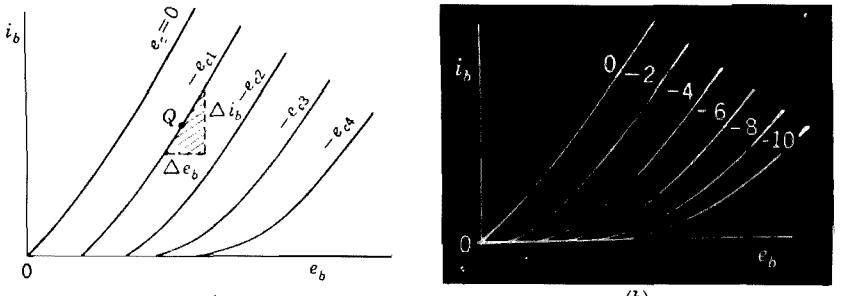



FIG. 3-4. (*a*) Plate characteristics of a triode. (*b*) Oscillogram of plate characteristics of a 6J5 triode. Length of $e_b$ axis = 400 volts.

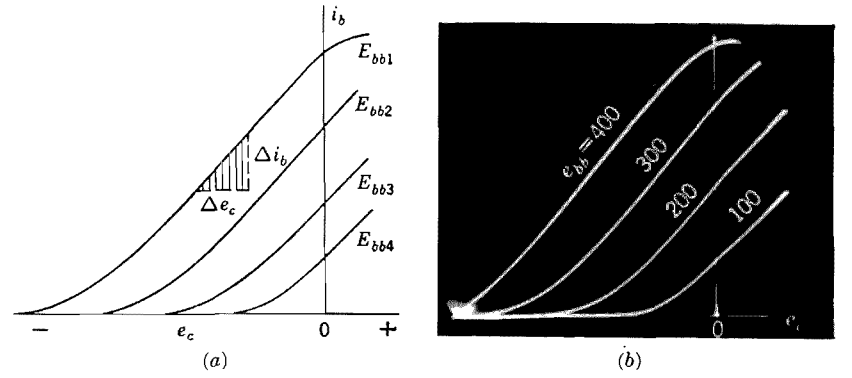



FIG. 3-5. (*a*) Transfer characteristics for a triode. (*b*) Transfer-characteristic oscillogram for a 6J5 triode.

The slope of the curves is not constant, but a region of approximate constant slope exists, and it will be possible to treat the vacuum tube as a linear device within this region. Plate resistance is a function of tube design, and triodes with $r_p$ values from 300 to above 100,000 ohms are available.

A second family of curves called the mutual, or *transfer, characteristics* relates $e_c$ and $i_b$, for constant values of $e_b$. A transfer-characteristic family is shown in Fig. 3-5*a* and *b*. Here again the matter of importance is the

*change* of $i_b$ produced by a *change* in $e_c$, and for a constant $e_b$ value such a change takes place along the curve. The slope of the curve, or the rate of change of $i_b$ with respect to $e_c$, is then

$$\text{Slope} = g_m = \frac{\partial i_b}{\partial e_c}\bigg]_{e_b = k} \tag{3-8}$$

The partial derivative is used again to indicate that the third variable $e_b$ is held constant.

The dimensions of the derivative are those of conductance, and since the curve relates current in one circuit and voltage in another, a transfer conductance is implied. Thus the parameter $g_m$ is called the *grid-plate transconductance*, or the mutual conductance. Transconductance is usually stated in micromhos, or $10^{-6}$ mho, and tubes are available with $g_m$ values from a few hundred to above 40,000 $\mu$mhos.

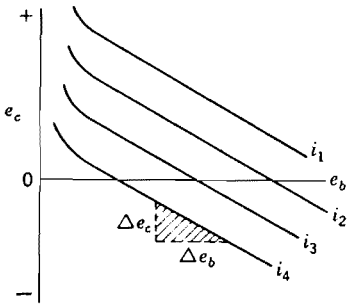



FIG. 3-6. Constant-current characteristics for a triode.

The third family of curves is shown in Fig. 3-6. These are the *constant-current* characteristics, showing the effect of $e_c$ and $e_b$ in holding $i_b$ constant. In Sec. 2-11 a parameter $\mu$ was defined as the ratio of the effect of grid voltage to the effect of plate voltage on the cathode surface electric field, or on the current. It may be noted that $g_m$ measures the effect of the grid voltage on plate current and $g_p$ measures the effect of the plate voltage on plate current. Then $\mu$, $g_m$, and $g_p$ must be related such that

$$\mu = \frac{g_m}{g_p} = \frac{\partial i_b/\partial e_c}{\partial i_b/\partial e_b} = -\frac{\partial e_b}{\partial e_c}\bigg]_{i_b = k} \tag{3-9}$$

It can be seen that this definition of $\mu$ shows the parameter to be equal to the slope of the constant-current curves of Fig. 3-6. The negative sign indicates that the direction of grid-voltage change must be opposite to the direction of plate-voltage change for $i_b$ to remain constant. The constant-current curve family has applications in performance calculations for certain forms of power amplifiers. The parameter $\mu$ has been called the *amplification factor*.

Equation (3-9) is more frequently written as

$$\mu = g_m r_p \tag{3-10}$$

which relates the three tube parameters at the point of operation and

imposes limitations on the tube designer, because of the relations between the tube parameters and mechanical tube design.

Thus it is interesting to note that:

1. The factor $\mu$ increases with grid-plate spacing but is independent of cathode-grid spacing.
2. The amplification factor $\mu$ increases as the grid-wire spacing is decreased.
3. The plate resistance increases with increase in cathode-plate spacing.
4. The plate resistance increases with decrease in grid-wire spacing.
5. The transconductance $g_m$ increases as the spacing between cathode and grid is decreased.
6. The transconductance increases with anode current and thus with cathode size.

Because of items 1 to 4 it may be reasoned that a high-$\mu$ tube is also a high-$r_p$ tube. Item 5 accounts for the cathode-grid spacings as small as 0.001 in. found in some designs of high-$g_m$ tubes.

**3-4. The Dynamic Transfer Curve.** The transfer characteristics of Fig. 3-5 were taken at constant anode voltage, without a load in the plate circuit, and are called the *static curves*. When a load is present and the tube gives useful plate-circuit output, it is $E_{bb}$, the supply voltage, which is constant and the tube voltage varies. To predict actual circuit performance, it is necessary to draw a *dynamic curve* for a given load and constant $E_{bb}$. This is obtained graphically by subtracting the quantity $i_b R_a$ from each ordinate value of the particular static curve chosen. A dynamic curve is shown in Fig. 3-7 for a 6J5 triode, superimposed on the static characteristics of the tube.

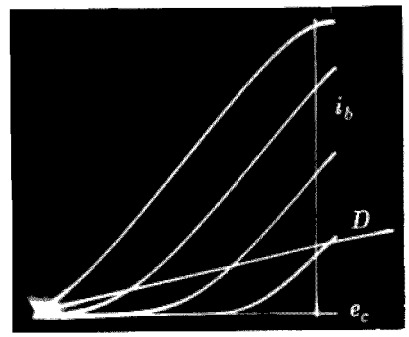



FIG. 3-7. Transfer characteristics for a 6J5 triode, showing a dynamic curve $D$ for $E_{bb}$ = 400 volts, $R_L$ = 20,000 ohms, $\mu$ = 20, $r_p$ = 7700 ohms rated.

The slope of the dynamic curve will always be less than the static curve, and the dynamic curve will be more nearly linear. This last point is of considerable importance, since the dynamic curve can be used to determine graphically output-current waveforms for a given load, as in Fig. 3-8. Linearity of the dynamic curve is important if the output-current waveform is not to be distorted, or the output-current waveform is to be identical with the input waveform of grid voltage.

The point marked $Q$ is called the *operating point*, or *quiescent point*, at which the tube operates with zero input signal. It is fixed by the

choice of $E_{cc}$ and $E_{bb}$ voltages. Since the dynamic curve can be plotted only point by point and is generally a curve, it is not often used for graphical analysis. It is, however, frequently employed as a starting point in theoretical analyses of vacuum-tube operation.

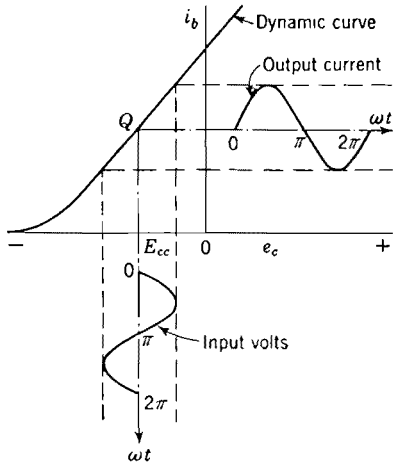



FIG. 3-8. Graphical determination of output-current waveform.

**3-5. Graphical Analysis of Circuit Performance.** The plate circuit in Fig. 3-9, composed of tube, $E_{bb}$, and load $R_a$, is a series circuit

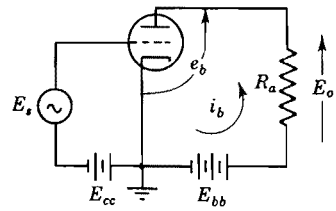



FIG. 3-9. Triode amplifier.

with a current $i_b$ flowing through all elements. Looking at the portion of the circuit external to the tube, the anode voltage is

$$e_b = E_{bb} - i_b R_a$$

from which

$$i_b = \frac{E_{bb}}{R_a} - \frac{e_b}{R_a} \tag{3-11}$$

Equation (2-29) implied the functional relation between $i_b$ and $e_b$ inside the tube as

$$i_b = f(\mu e_c + e_b) \tag{3-12}$$

Equations (3-11) and (3-12) both involve the same current $i_b$. Simultaneous analytic solution of the equations is not possible because of the general form of Eq. (3-12). However, this equation is represented by the plate family of curves, and a simultaneous graphical solution of the two equations is possible. This is known as the *load-line* method of solution.

Equation (3-11) is the expression for a straight line through the points

$$i_b = 0 \qquad e_b = E_{bb}$$
$$i_b = \frac{E_{bb}}{R_a} \qquad e_b = 0$$

and having a slope of $-1/R_a$. Use of this line, superimposed on the plate family of a given tube, will predict the current flowing for any value of $e_c$, with the given load and $E_{bb}$.

Such a *d-c load line* is drawn in Fig. 3-10 for $E_{bb} = 300$ volts, $R_a = 7500$ ohms, so that the current intercept is $E_{bb}/R_a = 4$ ma and the voltage intercept is $E_{bb} = 300$ volts. If the grid bias $E_{cc}$ now be set at $-10$ volts, the $Q$ point, or point of operation under quiescent conditions or no input signal, will fall at the intersection fixed by the d-c load line and the

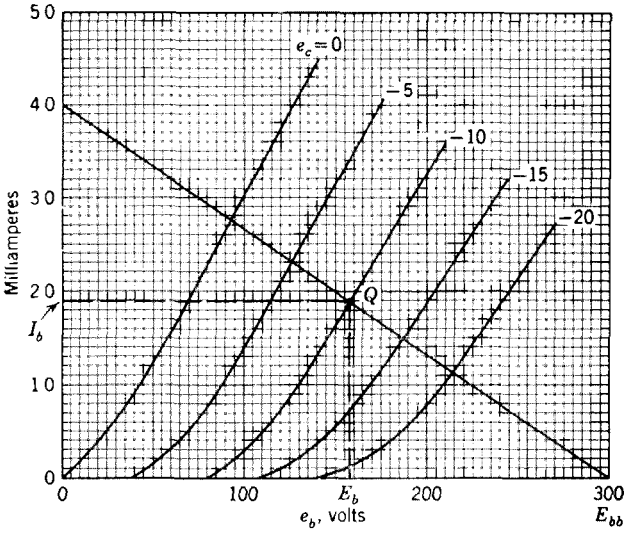



FIG. 3-10. Drawing the d-c load line.

characteristic for $e_c = -10$ volts. The operating conditions at the $Q$ point then are found to be $I_b = 1.9$ ma, tube voltage $= E_b = 157$ volts, and thus the load voltage is $300 - 157 = 143$ volts.

Ordinarily a varying or a-c input-signal voltage is used. If amplifiers are to be cascaded, that is, if a second tube is to use the output of the first as an input voltage, it is necessary that the a-c voltage be separated from the d-c voltage in the output circuit. This is accomplished by some form of coupling circuit, one form of which is shown in Fig. 3-11, composed of $C$ to block the d-c voltage, $R_a$ as a plate-load resistor, and $R_g$ as a d-c path in the output circuit. Thus the d-c and a-c current paths in the composite load circuit are not identical. The tube operating point continues to be fixed by $R_a$ and the d-c load line, but an additional load line is needed if an a-c input signal is present.

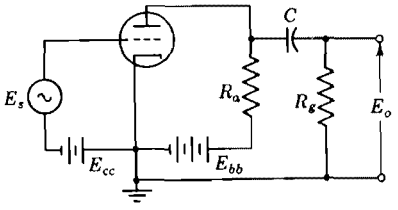



FIG. 3-11. Triode amplifier with *RC* coupling added.

If the reactance of $C$ is negligible, and $R_g = 15{,}000$ ohms, $R_a = 7500$

ohms as before, then the a-c load resistance is 5000 ohms. For $E_s = 0$ the tube is under quiescent conditions, or at the $Q$ point fixed by the d-c load line, and $E_{cc} = -10$ volts, and this must also be a point on the *a-c load line.* A line passed through $Q$ having a slope equal to $-1/R_{ac}$ (or $-1/5000$ for the circuit under discussion) will be the a-c load line. Such a line may be readily drawn, as in Fig. 3-12, by constructing a line having intercepts of 100 volts and $100/R_{ac} = {}^{100}/_{5000} = 0.002$ amp, and this is shown dotted in the figure. This line has the proper slope of $-1/5000$, and the a-c load line may then be drawn through the $Q$ point

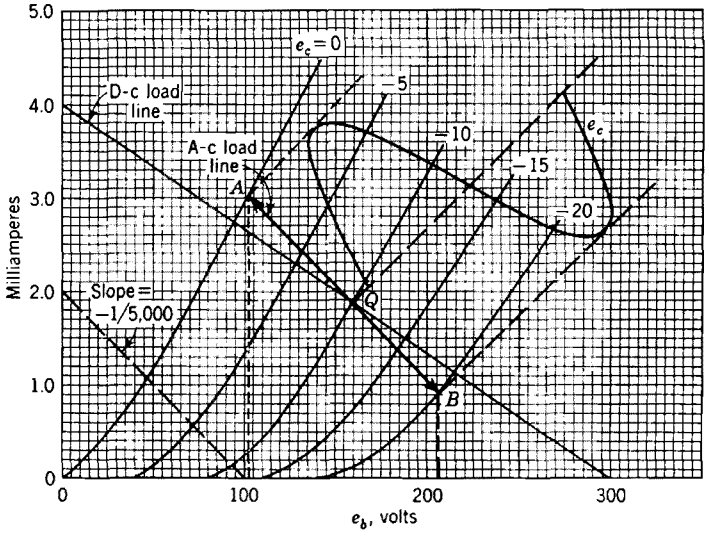



FIG. 3-12. Method of drawing the d-c and a-c load lines.

parallel to the dashed construction line. All points on the a-c load line represent instantaneous conditions for varying or a-c grid voltages.

Thus, if an a-c input sinusoidal signal having a peak of 10 volts be applied to the grid as shown, the instantaneous operating point will move back and forth along the a-c load line, up to $e_c = 0$ at the positive peak $A$ with maximum current of 3.0 ma and back to $e_c = -20$ at the negative peak $B$ with a minimum current of 0.9 ma. A waveform of output current can be plotted point by point if desired.

The a-c output voltage can be obtained as

$$E_o = \frac{E_{max} - E_{min}}{2\sqrt{2}} = \frac{206 - 102}{2\sqrt{2}} = 36.7 \text{ rms volts}$$

since the full voltage swing $A$ to $B$ is equal to the double peak value. Since $\sqrt{2}$ is the relation between peak and rms values, the result is

correct only for sine-wave outputs. The fundamental rms component of a-c plate current also is

$$I_p = \frac{I_{max} - I_{min}}{2\sqrt{2}} = \frac{3.0 - 0.9}{2\sqrt{2}} = 0.388 \text{ ma}$$

The *magnitude of the voltage gain* of the amplifier may also be found as

$$|\text{Gain}| = |A| = \frac{E_o}{E_g} = \frac{36.7}{7.07} = 5.2$$

Thus the load line allows determination of considerable information by graphical means. The method is usually employed and accurate only for large signals, since measurements from a graph cannot be well made for small voltage swings.

General rules for drawing the load line are:

1. Draw a d-c load line on the plate characteristics with $y$ intercept $= E_{bb}/R_a$ and $x$ intercept $= E_{bb}$.
2. Locate the $Q$ point on this line.
3. Lay out a construction a-c slope of proper amount, using $E_{bb} = 100$ and $i = 100/R_{ac}$ as intercepts.
4. Pass the a-c load line through $Q$, with slope parallel to the construction line of (3).

**3-6. The Voltage-source Equivalent Circuit for a Triode.** Much amplifier performance is with small input voltages, and linear or distortionless performance is desired. The tube graphical characteristics have considerable regions of linearity, in which the slopes or tube parameters are constant. If operation is confined to such regions, an analytic method is available for circuit solution of such linear-tube operation.

Equation (2-29), giving anode current as a function of $e_c$ and $e_b$, may be written in more general form as

$$i_b = f(e_c, e_b) \tag{3-13}$$

The total differential may be taken as

$$di_b = \frac{\partial i_b}{\partial e_c} de_c + \frac{\partial i_b}{\partial e_b} de_b \tag{3-14}$$

It is change in $i_b$ that is of interest, but the change must be finite, not infinitesimal, to be practical. It is permissible to substitute small finite changes for the above infinitesimals, *provided* that the partial-derivative terms are constants over the region of the $\Delta$ changes. Since the partial-derivative terms can be recognized as

$$\frac{\partial i_b}{\partial e_c} = g_m \qquad \frac{\partial i_b}{\partial e_b} = \frac{1}{r_p}$$

then the limitation of constancy implies that changes in $i_b$, $e_c$, or $e_b$ are to be confined to regions of the graphical tube characteristics having constant slope or where the curves are linear.

Substituting in Eq. (3-14) gives

$$\Delta i_b = g_m \Delta e_c + \frac{1}{r_p} \Delta e_b$$

Noting that the variation components of $e_b$, $e_c$, and $i_b$ are $e_o$, $e_g$, and $i_p$, then

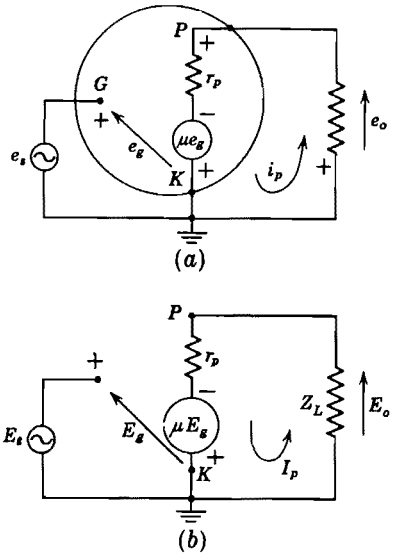



FIG. 3-13. (*a*) Equivalent circuit for the triode. (*b*) Same, with effective currents and voltages.

$$i_p = g_m e_g + \frac{1}{r_p} e_o$$

or
$$-\mu e_g + i_p r_p - e_o = 0 \quad (3\text{-}15)$$

This illustrates the relation between *varying* components of voltage in the circuit, *all d-c values being eliminated.* Actually the d-c sources supply the energy and fix the $Q$ point, and thus determine the particular values of $\mu$ and $r_p$ in use.

Equation (3-15) might also be written for the $i_p$ mesh of Fig. 3-13*a*, relating the voltages and currents of the circuit, and therefore this circuit is equivalent to the triode tube in so far as voltages and currents are concerned. That is, the triode may be replaced by an equivalent generator $\mu e_g$ with positive terminal to cathode, plus an internal resistance $r_p$. It should be noted that both $e_g$ and $e_o$ are defined as rises in voltage from reference or cathode, and the current $i_p$ flows in a direction consistent with the indicated generator polarity.

If the voltage and current variation is sinusoidal, the equivalent-circuit equation may be written in effective values as

$$-\mu \mathbf{E}_g + \mathbf{I}_p r_p = \mathbf{E}_o \quad (3\text{-}16)$$

This applies to the circuit of Fig. 3-13*b*, which is equivalent to a simple vacuum-tube amplifier with sinusoidal input and output.

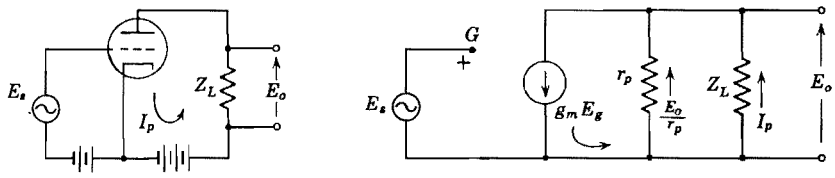
Thus a vacuum triode, operating in a linear region of its characteristics, or with constant values of $\mu$, $r_p$, and $g_m$, may be replaced with an equivalent voltage-source circuit, for a-c or varying current and voltage quantities. Conventional circuit-analysis methods may then be employed.

**3-7. The Current-source Equivalent Circuit.** Norton's theorem permits the transformation of the voltage-source equivalent circuit into a *current-source* equivalent circuit for a triode. By rearranging Eq. (3-16)

$$g_m\mathbf{E}_g = \mathbf{I}_p - \frac{\mathbf{E}_o}{r_p} \tag{3-17}$$

This equation is pertinent to the circuit of Fig. 3-14*b*, wherein a source current $g_m\mathbf{E}_g$ supplies two branches with current values of $\mathbf{I}_p$ and $-\mathbf{E}_o/r_p$. Here again it is desirable to remember that $E_o$ is a rise above reference, thus accounting for the negative sign.

The tube is then equivalent to a current source of value $g_m\mathbf{E}_g$, in parallel with a resistor $r_p$ or a conductance $g_p$. Direction of current flow is consistent with direction of the original voltage generator. Thus a second, or current-source, form of equivalent circuit is available for the




(*a*) (*b*)

FIG. 3-14. (*a*) Triode amplifier. (*b*) Current-source equivalent for (*a*).

vacuum tube. In many circuit applications, in which parallel paths exist in the plate circuit, this form will be found most convenient. It will also be the most useful form for the pentode tube.

**3-8. Gain of an Amplifier.** The increase in output voltage over the input voltage value is an important reason for employing amplifiers. This increase, or *voltage gain*, is defined as the ratio of output-voltage rise to input-voltage rise and is in general a complex number, since the two voltages may not be in phase. Thus

$$\text{Voltage gain} = \mathbf{G} = \frac{\mathbf{E}_o}{\mathbf{E}_s} \tag{3-18}$$

where $\mathbf{E}_s$ is the signal input to the circuit and may not be the actual grid-cathode voltage.

By noting that in Fig. 3-14*b*

$$\mathbf{I}_p = -\frac{\mathbf{E}_o}{\mathbf{Z}_L}$$

the gain for the circuit may be expressed by use of Eq. 3-17 as

$$\mathbf{G} = \frac{\mathbf{E}_o}{\mathbf{E}_s} = -\frac{\mu\mathbf{Z}_L}{r_p + \mathbf{Z}_L} \tag{3-19}$$

which will have a phase angle in addition to the 180° shift implied by the minus sign.

**3-9. Measurement of Tube Parameters.** The tube parameters $\mu$, $g_m$, and $r_p$ are defined as the slopes of the characteristic curves at the desired operating point, but accurate measurement of the values of the parameters directly from the curves is difficult. Circuits are available, however, which allow accurate measurement by bridge or null methods. The analysis of these circuits follows as an application of the equivalent circuits of the preceding section.

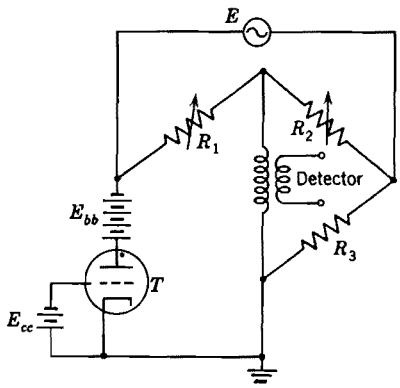



FIG. 3-15. Wheatstone bridge for measurement of $r_p$.

The bridge of Fig. 3-15 can be used for measurement of $r_p$ at $Q$ points fixed by selection of $E_{bb}$ and $E_{cc}$. When balanced for no output at the detector

$$\frac{r_p}{R_1} = \frac{R_3}{R_2} \tag{3-20}$$

and then

$$r_p = \frac{R_1 R_3}{R_2} \tag{3-21}$$

Frequencies of less than 1000 cycles are ordinarily used for the driving source $E$. The resistance of $E_{bb}$ must be negligible.

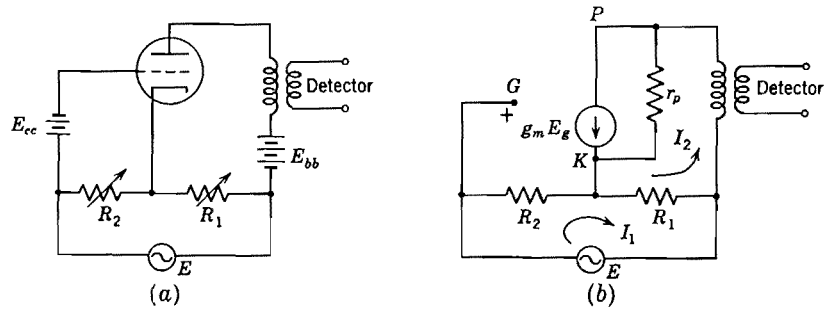



FIG. 3-16. Miller bridge and its current-source equivalent for measurement of $\mu$.

The Miller bridge of Fig. 3-16, when balanced for zero detector output, or $I_2 = 0$, gives a measurement of $\mu$. For this condition

$$g_m \mathbf{E}_g r_p = \mathbf{I}_1 R_1 \tag{3-22}$$

Since $\mathbf{E}_g = \mathbf{I}_1 R_2$, then

$$\mu = g_m r_p = \frac{R_1}{R_2} \tag{3-23}$$

A third circuit is available for measurement of $g_m$ as in Fig. 3-17. For balance, or zero detector output,

$$g_m \mathbf{E}_g \frac{r_p R_3}{r_p + R_3} = \mathbf{I}_1 R_1$$

Since $\mathbf{E}_g = \mathbf{I}_1 R_2$, then

$$g_m = \frac{R_1(r_p + R_3)}{r_p R_2 R_3} \tag{3-24}$$

If the measurement is to be independent of $r_p$, then, by choosing $r_p \gg R_3$, Eq. (3-24) reduces to

$$g_m = \frac{R_1}{R_2 R_3} \tag{3-25}$$

ANDHRA 1063 ... Courses, ... AGRICULTURAL UNIVERSITY

Headphones may be used as a detector, although a cathode-ray oscillograph with amplifier may show more sensitivity to the balance condition.

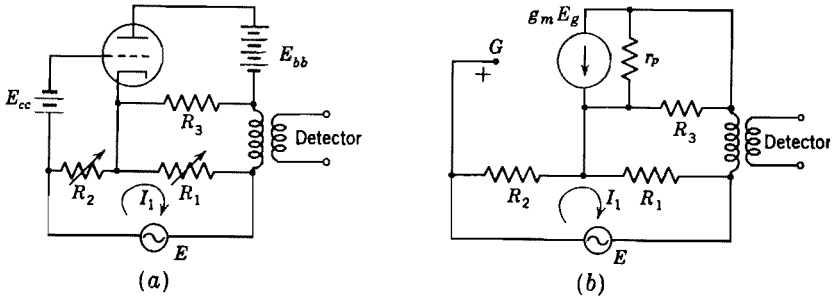



FIG. 3-17. (*a*) Bridge for measurement of $g_m$. (*b*) Current-source equivalent.

**3-10. The Tetrode.** The internal capacities inherently present in the triode become appreciably small reactances at frequencies of a megacycle or so and introduce circuit difficulties which will be further discussed in Chap. 4. These difficulties and instability seriously limit the usefulness of a triode as an amplifier at these frequencies.

In an attempt to overcome the deficiencies of the triode Hull introduced the *screen-grid tube* in 1927. Being a four-electrode tube, it is a *tetrode*. The second, or screen, grid is placed between the first, or control, grid and the anode, to screen or electrically shield the control grid from the anode. By so doing the grid-plate capacity $C_{gp}$ is reduced from a value of 3 to 10 micromicrofarads ($\mu\mu$f) to 0.01 $\mu\mu$f or less for a tetrode. This materially extends the frequency range of operation.

By operation of the screen at a positive d-c potential and at cathode a-c potential the screen takes over the electron-accelerating function of the anode and still retains its own shielding property. The effect of the anode potential on the electric field at the cathode is much reduced

by the shielding effect of the screen, and the anode current becomes almost independent of anode potential. Thus the plate resistance of a tetrode is very high, of the order of hundreds of thousands of ohms. Actually the tetrode operates much as a triode, made up of cathode, control grid, and screen grid as anode, with the tetrode anode serving merely as an electron collector.

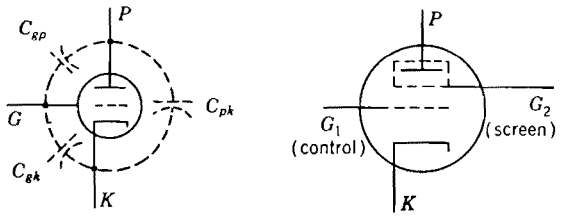



FIG. 3-18. (*a*) Interelectrode capacities of a triode. (*b*) Diagrammatic representation of a tetrode.

A set of plate characteristics for a tetrode are shown in Fig. 3-19. The badly warped region at reduced anode voltage occurs when anode potential falls below screen potential. Secondary electrons emitted from the anode by primary-electron bombardment are then attracted to the more positive screen, representing a current opposite to the direction of anode current. As can be seen from the curves, the secondary-emission current can exceed the current received from the cathode and cause a reversal of anode current. This effect can be reduced by carbonizing the anode surface, thus reducing the secondary-emission coefficient.

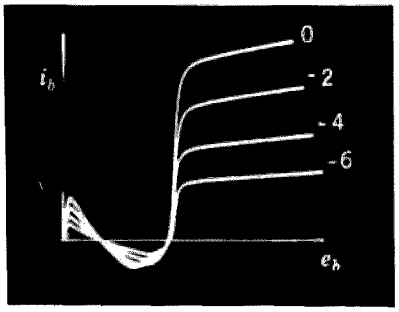


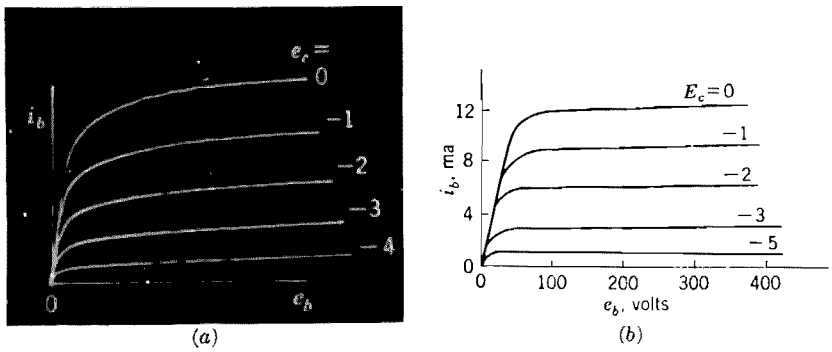

FIG. 3-19. Oscillogram of plate characteristics for a tetrode (type 24A), showing reversed anode current. $E_{c2} = 150$ volts.

Because of the limited linear region of operation available in the characteristics, the tetrode was soon superseded by the pentode and is not now in general use in its original form.

**3-11. The Pentode.** In 1930 the pentode, or five-element tube, was introduced, to overcome the deficiencies of the tetrode. Having three grids, control, screen, and suppressor, it eliminated the kinks in the tetrode curves at low anode voltage, still further improved the internal shielding, and raised the plate resistance above a megohm for usual tubes. The amplification factor also became very high.

A set of plate characteristics for a pentode appear in Fig. 3-20, with

apparent superiority in linearity over those of the tetrode. These curves show the effect of the suppressor grid in eliminating secondary emission from the anode. This grid is placed between screen and anode and operated at cathode a-c and d-c potential. The primary-electron stream has sufficient energy to override the retarding field of the suppressor




(*a*) (*b*)

FIG. 3-20. (*a*) Oscillogram of plate characteristics of a 6J7 pentode; $E_{c2}$ = 90 volts. (*b*) Typical pentode plate characteristics; $E_{c2}$ = 90 volts.

and to reach the anode. The secondary electrons emitted from the anode face a negative, or repelling, field due to the suppressor and return to the anode, with no effect on the anode current. Figure 3-21 shows by direct comparison the improvement in characteristics given by the pentode over the tetrode.

In operation in the linear region equivalent circuits similar to those found for the triode may be employed. It is most convenient to utilize the current-source circuit, since, as can be seen from the characteristics, a pentode operates much like a constant-current source, independent of applied plate voltage.

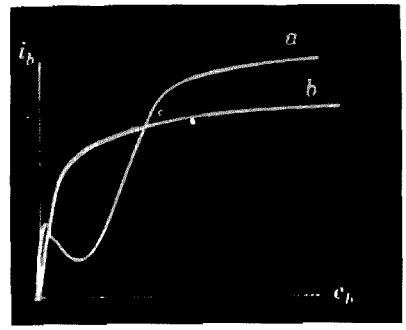



FIG. 3-21. Performance of a 6SJ7, connected (*a*) as a tetrode, (*b*) as a pentode.

Since $r_p$ exceeds a megohm in value for most pentodes, use of resistance loads large with respect to $r_p$ will require excessive values of $E_{bb}$, owing to the $I_bR_a$ drop. If resonant-load circuits are employed, it will be found impossible to obtain resonant impedances of the order of $r_p$. Because of these difficulties, plate loads of 100,000 to 200,000 ohms are usual. Since the load value is small with respect to $r_p$, the current will be determined largely by the generator resistance and not the load.

Referring to the current-source equivalent circuit, and since $r_p \gg Z_L$,

the $r_p$ branch may be dropped from the circuit since the current may be assumed as flowing entirely through the $Z_L$ branch. This leads to the pentode equivalent circuit of Fig. 3-22. This circuit is again based on an assumption of linear operation and on $r_p \gg Z_L$. Since tube capacities have not yet been considered, the circuit applies only for low frequencies.

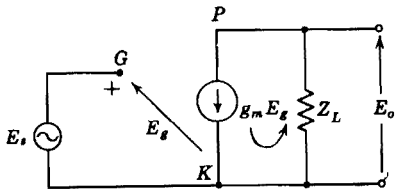



FIG. 3-22. Current-source equivalent circuit for a high-$r_p$ pentode.

The gain for a pentode amplifier is then calculated as

$$\mathbf{E}_o = -g_m \mathbf{E}_g Z_L$$

and

$$\mathbf{A} = \frac{\mathbf{E}_o}{\mathbf{E}_g} = -g_m Z_L \qquad (3\text{-}26)$$

which is simpler than the corresponding triode expression. The graphic load-line method is also applicable to the pentode.

With $g_m$ that of the 6AK5 pentode of 5100 $\mu$mhos and a load of 25,000 ohms, a gain of 127 per stage can be realized. Because of the high gain and low $C_{gp}$ the pentode has wide application in all frequency ranges.

**3-12. The Beam Tube.** Design changes in the pentode have developed a modified tetrode, the *beam tube*, having improved pentode characteristics. For large signal and power outputs the curved low-voltage regions of the pentode (Fig. 3-20*a*) limit the operation. They are due to interference of the screen with the primary-electron stream. In addition the high screen currents of the pentode cause screen power losses and reduce the over-all efficiency.

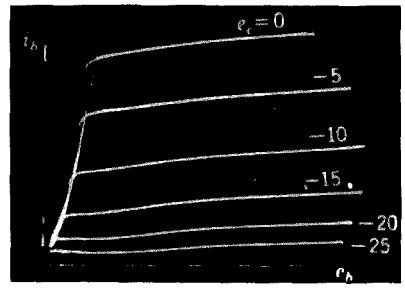



FIG. 3-23. Plate characteristics for the 6L6 beam tetrode; $E_{c2}$ = 230 volts.

In the beam tube the screen-grid wires are aligned behind, or in the electronic shadow of, the control-grid wires, reducing interference of the screen with the primary electrons. The suppressor grid is replaced with beam-forming plates at each side of the cathode, forcing the electron flow to occur in limited areas. The high electron density in front of the anode produces a negative space charge or a reduced-potential region sufficient to provide suppressor action on the secondary electrons. Actually there is no suppressor grid, no interference with the primary electrons, and the plate characteristics are straightened out, as shown in Fig. 3-23. Only at low anode voltage and current is there a slight dip, reminiscent of the characteristics of the early tetrodes.

The beam tube has advantages over the pentode when large signals

and power outputs are desired, since a lower grid input is required. This implies an increased value of $g_m$, which is realized. Since the beam tube is ordinarily used for large signals, the graphical load-line method of performance is most satisfactory.

**3-13. The Variable-$\mu$ Tube.** The gain of a pentode amplifier has been shown to be stated by $-g_m Z_L$. The parameter $g_m$ is a function of grid bias $E_{cc}$ since, as the operating point is moved down the transfer curve by increase of bias, the slope of the curve lessens and the value of $g_m$ decreases. This provides a means of smoothly varying the gain of an amplifier by shift of the d-c grid bias. However, for a triode or ordinary pentode the abrupt curvature at the bottom of the curve may introduce considerable distortion.

The *variable-$\mu$*, or *remote-cutoff*, pentode was designed to handle the gain-control function with large input voltages without distortion. As can be seen in Fig. 3-24, the dynamic grid transfer characteristic is smooth and regular over a wide variation of slope or of $g_m$.

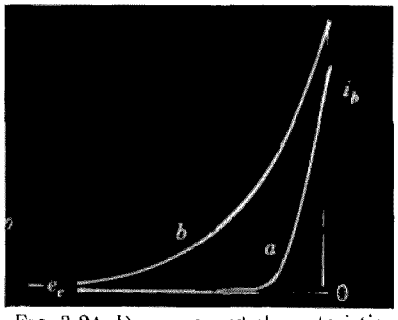



FIG. 3-24. Dynamic grid characteristics for (*a*) 6SJ7 sharp-cutoff pentode, (*b*) 6SK7 remote-cutoff or variable-$\mu$ pentode.

The result is achieved by winding the grid wires with variable pitch, these being close-wound at the ends and coarse-wound at the middle. When low negative grid bias is applied, the tube has some high value of $g_m$ or $\mu$ obtained as an average of the effects of the coarse- and fine-spaced portions of the grid. As the bias is increased negatively, the closely spaced portions of the grid reach cutoff and the $g_m$ or $\mu$ takes a lower value as an average of the effects of the coarser-spaced grid wires.

The transconductance for a remote-cutoff pentode may be 2000 $\mu$mhos at $E_{cc} = -3$ volts, giving a gain of 100 with a load of 50,000 ohms. At $E_{cc} = -35$ volts the $g_m$ value is reduced to 10 $\mu$mhos and the gain to 0.5.

**3-14. Self-bias of Amplifiers; Voltage Sources.** The definition for total anode current has been given as

$$i_b = I_b + i_p$$

If this current flows from cathode to ground through a parallel $RC$ circuit as in Fig. 3-25, the d-c component will produce a potential, positive to cathode, equal to

$$E_k = R_k I_b \qquad (3\text{-}27)$$

Since this is also negative to grid, it may be er

ANGRAU Central Library
Hyderabad
12469

negative grid bias $E_{cc}$ and the value of $R_k$ chosen so that

$$R_k = \frac{E_{cc}}{I_b} \tag{3-28}$$

This is usual practice in amplifier circuits.

Capacitor $C_k$ is chosen to have a reactance small with respect to $R_k$ so that the two current components may be assumed to divide as shown, and no a-c voltage will appear across the parallel combination. If $X_{ck} \ll R_k$ at the lowest frequency of interest, the inequality will be satisfied for all higher frequencies. A value of $X_{ck} = R_k/10$ may often be sufficient, although cost may frequently dictate smaller values of $C_k$.

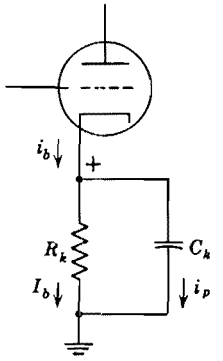



FIG. 3-25. Analysis of current flow in cathode-bias circuit.

If $C_k$ is of proper size, the whole $R_k$, $C_k$ bias circuit may be assumed to disappear from equivalent circuits used to analyze performance. Operation with $C$ absent or small will be discussed in Chap. 5.

The use of cathode bias in triode and pentode circuits is shown in Fig. 3-26. In addition a method of obtaining screen-grid voltage from the anode supply is shown for the pentode in Fig. 3-26*b*. The value of $R_s$ is chosen to drop the anode voltage to the desired screen value, owing to flow of screen current. Capacitor $C_s$ bypasses the screen direct to cathode and should represent a reactance of only a few hundred ohms at the lowest operating frequency.

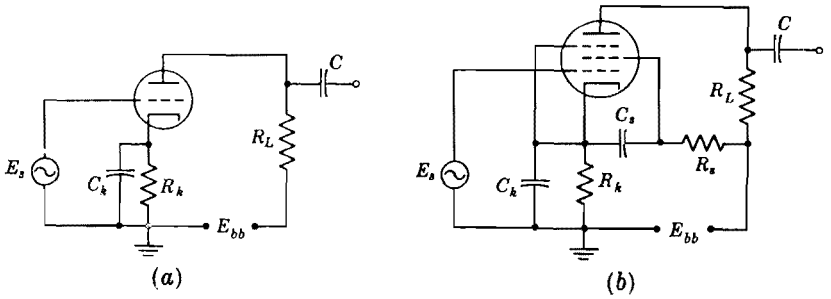



FIG. 3-26. (*a*) Triode amplifier with self-bias. (*b*) Pentode amplifier with self-bias and screen supply.

Anode voltages are usually supplied from rectified and filtered a-c power supplies, even though batteries may be shown in many circuits for simplicity.

**3-15. Distortion of Waveform.** Any difference between input and output waveforms of an amplifier indicates that *distortion* is present.

In many applications it is desired that the vacuum tube amplify the input signal and produce an undistorted output voltage, or one whose waveform is identical with the input. In other applications distortion is deliberately introduced since it is desired to produce an altered, or distorted, output.

Three types of distortion are found in vacuum-tube circuits as:

1. Frequency distortion
2. Nonlinear, or amplitude, distortion
3. Phase, or delay, distortion

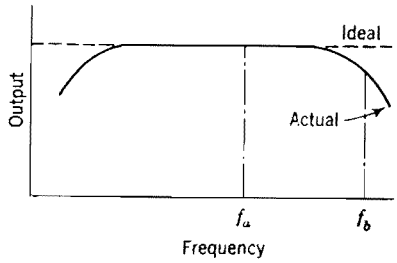



FIG. 3-27. Amplifier gain-frequency curve.

*Frequency distortion* is produced by unequal amplification of the different component frequencies of a given signal. Referring to Fig. 3-27, it can be seen that an amplifier having a horizontal gain-frequency plot as shown by the dashed line will be ideal, or will have no frequency distortion. An actual amplifier curve may droop at one or both ends of the frequency axis, and the gain for frequencies $f_a$ and $f_b$ will differ. If a composite waveform having both frequencies present as

$$e_s = K_a \sin \omega_a t + K_b \sin \omega_b t$$

is applied, and if the gain for $f_a$ is $n$ and for $f_b$ is $0.8n$, then the output will be

$$e_o = nK_a \sin \omega_a t + 0.8nK_b \sin \omega_b t$$

and obviously the output waveform will differ from the input, and frequency distortion is present.

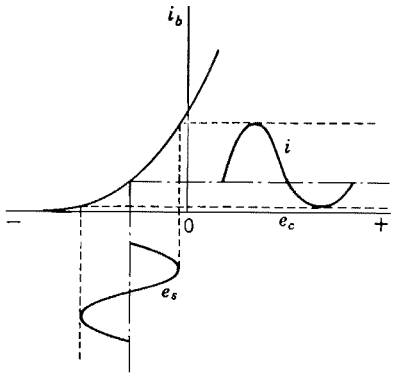



FIG. 3-28. Nonlinear distortion as caused by a curved dynamic transfer characteristic.

Frequency distortion is due to the presence of reactive elements associated with the tubes or circuits.

*Nonlinear*, or *amplitude, distortion* is produced if the input-output relation of the tube and circuit is not linear, or if different amplitudes of input are not amplified by equal factors. If the dynamic transfer curve of a vacuum tube is not linear, then nonlinear distortion will occur as shown in Fig. 3-28. Although the grid input voltage $e_s$ is sinusoidal, yet the output current is distorted, indicating the presence of harmonic frequencies in the output.

When a number of input frequencies are simultaneously present, a non-

linear characteristic also produces *intermodulation distortion*. In addition to harmonics the curved characteristic introduces frequencies which are sums and differences of all the frequencies present. These sum and difference frequencies bear no harmonic relationship to the original input and are treated by the ear as noise.

Nonlinear distortion is usually caused by improper operating conditions for the tube.

*Phase*, or *delay*, *distortion* occurs if the phase relation between various frequency components of the input signal is not maintained. If the phase angle $\theta$ by which the various component frequencies of an input signal are shifted is proportional to frequency, then the phase of all components is shifted, or delayed, by the same time interval and the whole wave shifts, or is delayed, as a unit, with no phase distortion.

The important criterion in connection with phase distortion is that $d\theta/d\omega$ be constant to give zero phase distortion. Phase distortion is caused by the presence of reactive circuit elements.

**3-16. Classification of Tube Operating Conditions.** Vacuum-tube operating conditions are universally prescribed by the designations *Class A*, *AB*, *B*, or *C*. These designations are graphically illustrated in Fig. 3-29, showing the relation of bias, input signal, and output-current waveform, utilizing the dynamic transfer curve.

In Fig. 3-29, under Class A operation, the grid-bias voltage $E_{cc}$ is so set that operation occurs over a linear region of the curve, or $E_{cc}$ is chosen somewhat less than one-half the cutoff-voltage value. The input grid signal is sufficiently small so that the grid remains negative at all times. As a result plate current flows continuously, and the output waveform is similar to the input waveform, giving low distortion because of the linear characteristic.

Voltage amplification is high, but power output small because of limited voltage and current swings. The power efficiency is low, being theoretically limited to 50 per cent, but in practice being in the range of 2 to 25 per cent. Class A amplifiers are used for voltage gain and where small amounts of power with low distortion are required.

In Fig. 3-29*b*, indicating Class B operation, the grid-bias voltage is set at approximately the cutoff value, and the input signal may drive the grid positive at times. Plate current flows for one-half of the time, or for 180° of the input cycle, and the pulses of output current are similar in waveform to the positive half cycles of input voltage. Since the negative half cycles are absent, the distortion is high.

Because of the large voltage swings possible, the power output is high and the power efficiency theoretically reaches 78.5 per cent, practical circuits giving 50 to 60 per cent.

Class AB operation is with bias intermediate to Class A and Class $B_1$

and plate current flows for more than 180° and less than 360° of the input cycle. Distortion, power output, and efficiency are between Class A and Class B limits.

Class C operation, as in Fig. 3-29, requires a grid-bias voltage greater than cutoff, usually two or more times cutoff. The input signal is large

APAU CENTRAL LIBRARY
Acc: No: 1246
Date: 13.8.85

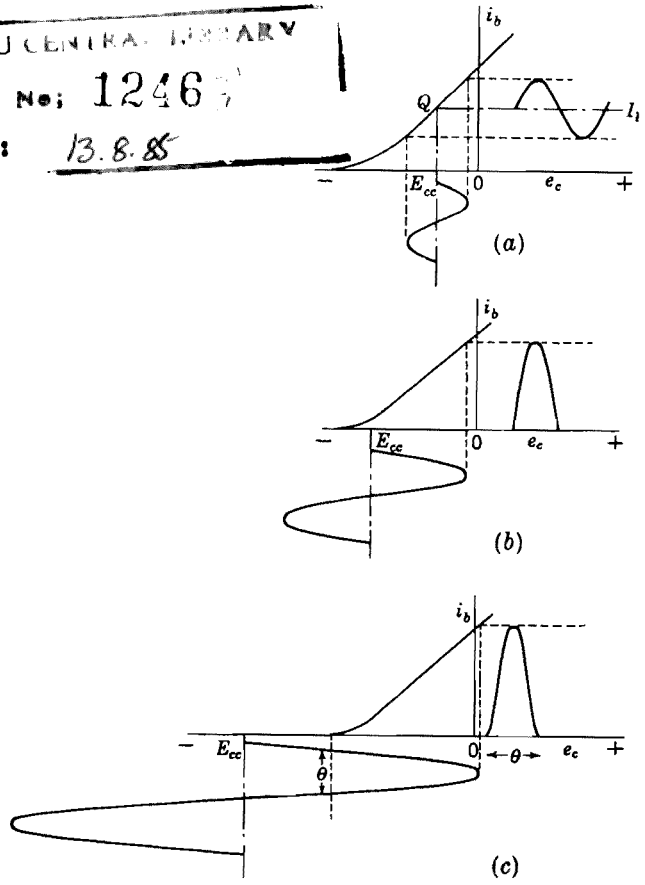



FIG. 3-29. (*a*) Class A operation. (*b*) Class B. (*c*) Class C.

and may drive the grid considerably positive. The plate current flows for an angle $\theta$ less than 180° of the input cycle. The distortion is very high, but power output may also be high for a given tube. Efficiency theoretically reaches 100 per cent, practice yielding values of 65 to 85 per cent. Because of the high distortion, Class C operation is ordinarily applied only at high frequencies, where the distortion products can be removed by simple forms of filter circuits.

**3-17. The Decibel in Gain Measurement.** It will be found of considerable value in the analysis of circuit performance to employ logarithmic units of amplifier gain and power output. Also, since the ear is the final receptor for the output of much electronic equipment, and since the ear hears sound intensities on a logarithmic scale, a logarithmic unit is desirable from this standpoint.

A unit for logarithmic measurement of *power ratios* has been established and named the *bel*, after Alexander Graham Bell. A unit one-tenth as large has been found more practical, and named the *decibel*, where

$$\text{Number of decibels} = \text{db} = 10 \log \frac{P_2}{P_1} \tag{3-29}$$

It should be emphasized that the decibel is a unit of power ratio and not directly a unit of absolute power. However, if a certain reference power, or *zero level*, for $P_1$ is adopted and stated, then computation of absolute power follows, the results being stated as so many decibels above or below zero level. A variety of reference levels has been proposed; one which has become common is 0.001 watt.

If input and output of an amplifier are measured across the same resistance values, then by the definition of power

$$\text{db} = 10 \log \frac{E_2^2}{E_1^2} = 20 \log \frac{E_2}{E_1} \tag{3-30}$$

If the grid-cathode circuit of a tube is considered an open circuit and the output voltage $E_o$ is supplied to a circuit also drawing no power, then the *decibel gain* of a voltage amplifier is also defined by Eq. (3-30), where

$$\text{Gain in db} = 20 \log \frac{E_o}{E_s} \tag{3-31}$$

A negative sign implies a power loss. The above equation has become a definition for decibel gain of voltage amplifiers.

In this text "log" will be taken to mean logarithm to the base 10, whereas "ln" will be used for Naperian-base logarithms.

**Example 1.** The output of an amplifier is 3.5 watts under one condition and 7.0 watts under another. Find the decibel output level, 0.001 watt reference.

Under the first condition

$$10 \log \frac{3.5}{0.001} = 10 \times 3.544 = 35.44 \text{ db above reference}$$

Under the second condition

$$10 \log \frac{7.0}{0.001} = 10 \times 3.845 = 38.45 \text{ db above reference}$$

Thus *doubling the power* has increased the decibel level by 3.01 db (log of 2).

**Example 2.** The input of an amplifier is 0.01 volt, giving an output of 2.17 volts. Find the amplifier gain in decibels. Then

$$20 \log \frac{2.17}{0.01} = 20 \times 2.336 = 46.72 \text{ db gain}$$

**Example 3.** A certain transformer has an efficiency of 72 per cent. Find the decibel transformer loss.

$$\text{Efficiency} = \frac{P_2}{P_1} = 0.72$$

Therefore

$$\text{db} = 10 \log 0.72$$

and to avoid negative logarithms, note that

$$\log \frac{1}{A} = -\log A$$

Then

$$\text{db} = -10 \log 1.39 = -10 \times 0.143 = -1.43 \text{ db}$$

The negative sign then indicates a loss.

**3-18. The Cathode-ray Tube and Oscillograph.** Waveforms of voltages and currents may be directly studied in a tube in which the point of impact of a beam of electrons is made visible. Such a tube is called

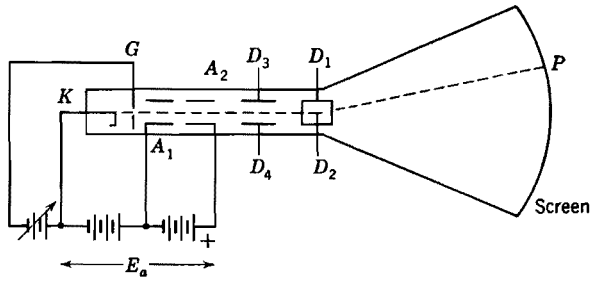



FIG. 3-30. Outline of cathode-ray tube.

the *cathode-ray tube,* since the electron beam was first thought to consist of rays from the cathode.

Owing to the electrical charge of the electrons in the beam, the beam may be bent, or deflected, by varying electric and magnetic fields, to trace a pattern on a fluorescent screen, which glows upon impact with the electrons. The tube may be utilized to plot the instantaneous values of one electrical potential with respect to a second, to form an $x$, $y$ plot, or indirectly to plot one electrical variable as a function of time.

Figure 3-30 diagrams such a tube, with cathode $K$ as the source of electrons. Grid $G$ has a small hole in its center and transmits a pencil of electrons, the beam current being varied by changing the bias potential with respect to cathode. This permits adjustment of spot intensity on the fluorescent screen. The pencil of electrons then passes through

holes in anodes $A_1$ and $A_2$, the pair forming an electric-field lens by which the electron beam is accurately brought to a focus as a small spot on the screen. The potential of $A_1$ is varied to provide a focusing control. The complete beam-forming assembly is known as an *electron gun.*

After leaving the gun, the beam passes between a pair of parallel plates $D_3$ and $D_4$, and then between a second pair oriented at right angles to the first pair. An electric field between $D_3$ and $D_4$ deflects the electron beam in a direction normal to the plane of the plates (vertical in the figure). Another voltage on the $D_1$, $D_2$ pair produces a deflection perpendicular to that of the $D_3$, $D_4$ pair (horizontal in the figure). Thus

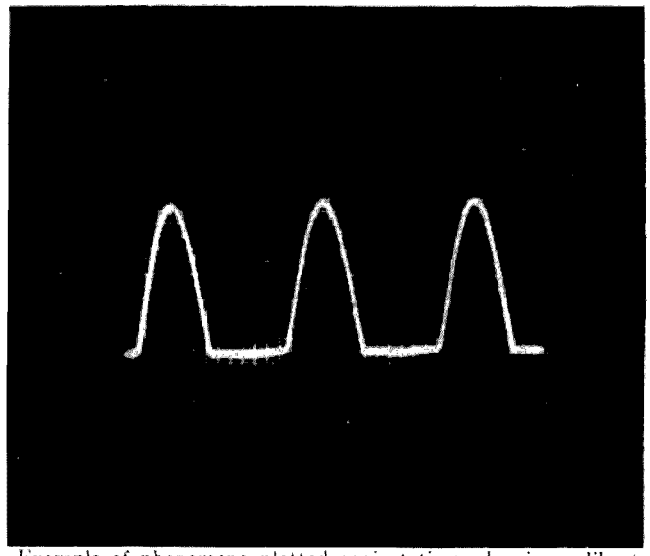


FIG. 3-31. Example of phenomena plotted against time, showing calibrated screen.

it is possible to control the position of the spot on the screen in two dimensions by means of two potentials.

The deflection achieved for the beam can be calculated from the geometry of the tube as

$$D = \frac{l_d L E_d}{2 d E_a} \quad \text{meters} \tag{3-32}$$

where $l_d$ = length of deflection plate, m
$L$ = distance, center of plate to screen, m
$E_d$ = deflecting potential
$d$ = distance between deflecting plates, m
$E_a$ = accelerating potential acting on beam

This shows the deflection proportional to the applied deflecting potential $E_d$, or to the field intensity $E_d/d$. The accelerating potential may

range from 1000 to 80,000 volts, dependent on the speed with which the electron spot is to be deflected, or the intensity of the desired trace on the screen.

By application of a potential varying linearly with time to the $x$-axis deflection plates it is possible to show potentials applied to the $y$-axis plates as functions of time. Figure 3-31 illustrates such a plot of a rectified a-c wave, showing also the ability to perform quantitative measurements through the use of a calibrated plastic screen over the fluorescent end of the tube. The linear time-base voltage will be of the saw-tooth form of Fig. 3-32, the quick drop to zero taking the retrace back across the screen so rapidly as to fail to register. In some cases a blanking negative voltage is derived from this rapidly changing portion of the sweep wave and applied to the grid $G$ of the tube to cut off, or blank, the beam and screen trace during the beam return to the time zero point.

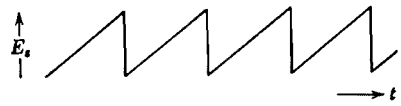



FIG. 3-32. Sweep-voltage waveform.

It is also possible to deflect the electron beam by use of a magnetic field transverse to the beam direction, the field being obtained from a coil and a varying current. The deflection obtained may be calculated from

$$D = \sqrt{\frac{e}{m}} \frac{B}{\sqrt{2E_a}} l_m L \qquad \text{meters} \tag{3-33}$$

where $e/m$ = charge-to-mass ratio of the electron = $1.759 \times 10^{11}$ coulombs/kg
$E_a$ = accelerating potential
$B$ = flux density acting, webers/sq m
$l_m$ = length of beam in magnetic field, m
$L$ = distance, center of field to screen, m

For air-cored coils $B$ is a direct function of coil current. Since the coil current is

$$i = \frac{1}{L} \int_0^t e \, dt$$

and $e$ is the potential applied to the coil, then direct observation of potentials by magnetic deflection does not give true waveforms. However, it is relatively easy to generate a coil current as a linear time function, and magnetic deflection is frequently used for providing the linear time base or sweep action. This is the case in most television receivers.

The fluorescent-screen material is carefully compounded of natural and synthetic minerals to give the desired color response and phosphorescent

properties. Screens are available in which the image persists for varying times after the excitation ceases, so that lines produced by very rapid or very infrequent traces may be retained for easy visual observation. Persistence values from a few microseconds to several minutes are available. For the longer times the effect is produced by a two-layer screen, in which a short-persistence material with a high output in the blue and

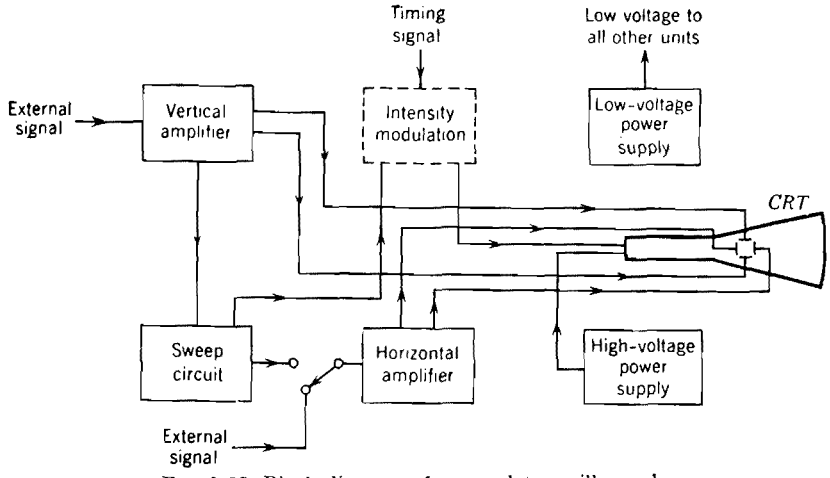



FIG. 3-33. Block diagram of a complete oscillograph.

ultraviolet is excited by the beam, the ultraviolet light then exciting the second layer to a long-time phosphorescent action in the yellow portion of the spectrum. The usual laboratory oscillograph is normally fitted with the P1 green screen of medium persistence of 20 to 30 msec, giving good visual acuity and reasonable persistence for visual observation.

A complete oscillograph, diagrammed in Fig. 3-33, will include amplifiers

TABLE 4. FLUORESCENT-SCREEN MATERIALS

| Type phosphor | Color | Persistence | Application |
|---|---|---|---|
| P1 | Green | Medium, 20 msec | General-purpose oscillograph |
| P2 | Blue-green followed by yellow | Blue, short<br>Yellow, several minutes | Transient visualization |
| P4 | White | Medium, 33 msec | Television |
| P5 | Blue | Very short, 20 $\mu$sec | Fast photography |
| P7 | Blue, then yellow | Blue, short<br>Yellow, 1 min | Radar |
| P11 | Blue | Short, 1 to 2 msec | Similar to P5 |
| P15 | Blue-green | Very short | High-speed phenomena |

for both $x$ and $y$ deflecting voltages, a generator for the sweep-voltage waveform with a considerable range of frequencies available, and power supplies for accelerating and focus potentials and for the various amplifiers. The amplifiers for the deflection potentials must have sufficiently wide frequency response as not to distort even at the highest expected input frequency.

Recording of the screen traces by photography is relatively easy and has been employed for many illustrations of transients in this book. Use of the P1 or P5 screens is desirable, both being suited to panchromatic film. Use of such film with speed rating of ASA 100, and exposures of $\frac{1}{10}$ sec at $f/5.6$, will usually result in satisfactory negatives. Development of the film in a high-contrast developer and printing on No. 5 contrast paper will aid in building up image contrast.

## PROBLEMS

**3-1.** A special tungsten-filament lamp has a volt-ampere curve given by the equation

$$I = 0.012V^{1.9} \quad \text{amp}$$

Find the static resistance, variational resistance, and voltage across the lamp for currents of 0.62 and 1.15 amp.

**3-2.** The following data are taken on a certain triode:

| $i_b$, ma | $e_c$, volts | $e_b$, volts |
|---|---|---|
| 10.0 | −3 | 155 |
| 10.0 | −5 | 190 |
| 9.0 | −3 | 145 |
| 8.3 | −2 | 155 |

Find the values of $\mu$, $g_m$, and $r_p$ in the region, assuming the data define linear curves.

**3-3.** It is possible to expand the function of Eq. (3-12) in a Taylor's series, and for a particular tube the first two terms are

$$i_b = 57 \times 10^{-6}(12e_c + e_b) + 0.24 \times 10^{-6}(12e_c + e_b)^2$$

and further terms are negligible. If the $Q$ point is chosen at $e_c = -6$ volts and $e_b = 180$ volts, find (*a*) value of $i_b$, (*b*) $r_p$, (*c*) $g_m$.

**3-4.** A triode has characteristics shown in Fig. 3-34. For a load of 125,000 ohms, $I_b = 0.8$ ma, $E_{cc} = -1.0$ volts find (*a*) $E_{bb}$, (*b*) output a-c voltage for $E_s = 0.7$ volt rms, (*c*) value of $E_b$.

**3-5.** What load resistance will place the $Q$ point at $I_b = 0.8$ ma, $E_b = 235$ volts if $E_{bb} = 350$ volts, for the tube of Fig. 3-34? Graphically calculate the gain of an amplifier operated with this load and $Q$ point; find the value of $E_{cc}$ and $I_{max}$ and $I_{min}$ for a 2-volt peak-to-peak signal.

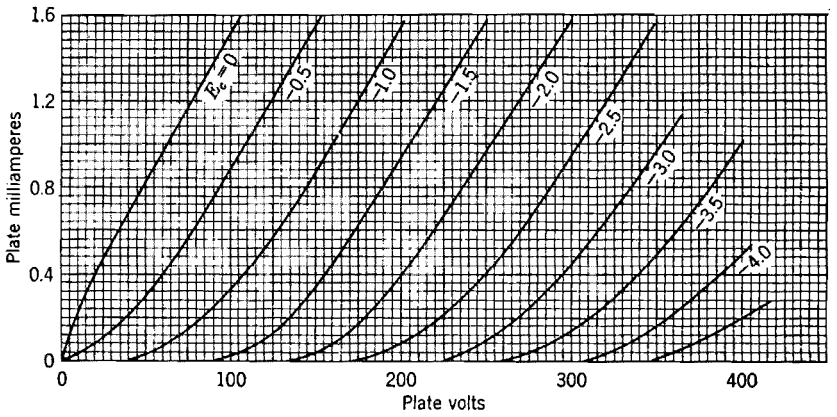



FIG. 3-34

**3-6.** Draw the current-source equivalent circuits for Fig. 3-35. Repeat, using voltage-source circuits. All capacity reactances are negligibly small.

**3-7.** In Fig. 3-35*c*, $R_1$ = 10,000, $R_2$ = 15,000 ohms, $g_m$ = 2000 μmhos, $r_p$ = 800,000 ohms. For $E_s$ = 3.0 volts find the reading of an a-c milliammeter at $I$.

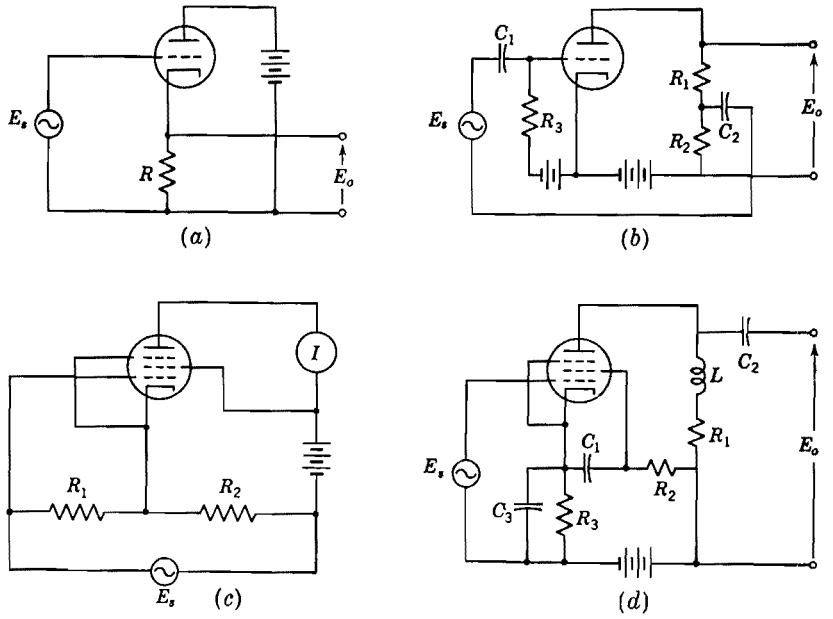



FIG. 3-35

**3-8.** Two tubes having the following parameters are connected in parallel: $\mu_1$ = 4.5, $g_{m1}$ = 1750 μmhos; $\mu_2$ = 30.5, $g_{m2}$ = 3600 μmhos. Find $\mu$, $g_m$, and $r_p$ of a single tube which will be equivalent to the parallel combination.

**3-9.** Two nonidentical tubes are operated in parallel. Determine the expressions for equivalent $\mu$, $g_m$, and $r_p$ in terms of $\mu_1$, $g_{m1}$, $r_{p1}$ and $\mu_2$, $g_{m2}$, and $r_{p2}$ of the respective tubes.

**3-10.** The tube of Fig. 3-34 is to be operated from $E_b = 250$ volts. If $I_b$ is to be 1.0 ma, specify the value of bias, the value of a cathode resistor to supply the bias, and the total power input to the tube.

**3-11.** The dynamic transfer curve of a triode can be represented by

$$i_b = a_0 + a_1 e_g + a_2 e_g^2 + a_3 e_g^3$$

If $e_g = E_m \sin \omega t$, find the amplitudes of all frequencies present in the output current.

**3-12.** The tube of Prob. 3-11 has applied as an input voltage

$$e_g = E_1 \sin \omega t + E_3 \sin 3.5\omega t$$

Find all frequencies and their amplitudes present in the output current.

**3-13.** A certain triode with an output power of 0.275 watt supplies power to a transformer of 65 per cent efficiency. What is the decibel transformer loss, and the power level in the load in decibels, 0.001 watt reference?

**3-14.** A microphone with output of $-56$ db and 200 ohms impedance supplies an amplifier of $+31$ db output level into a load of 16 ohms (0.001 watt reference). (*a*) Find the required amplifier decibel gain. (*b*) Find the amplifier output power in watts, and the load voltage and current.

## REFERENCES

1. "Standards on Electronics," Institute of Radio Engineers, New York, 1938.
2. "American Standards for Graphical Symbols for Electronic Devices," American Standards Z32.10—1944, American Standards Association, New York.
3. "Standards on Abbreviations, Graphical Symbols, Letter Symbols, and Mathematical Signs," Institute of Radio Engineers, New York, 1948.
4. Preisman, A. "Graphical Constructions for Vacuum Tube Circuits," McGraw-Hill Book Company, Inc., New York, 1943.

CHAPTER 4

# SMALL-SIGNAL AMPLIFIERS AS CIRCUIT ELEMENTS

The performance of a vacuum tube, be it triode, tetrode, or pentode, is dependent on the particular circuit configuration with which the tube is employed. As three-element devices, tubes may be connected into four-terminal input and output circuits in six different ways, only three of which have so far had much practical value. These three basic circuit forms differ as to gain, input and output impedances, and external-circuit loading, and each circuit must be studied in some detail.

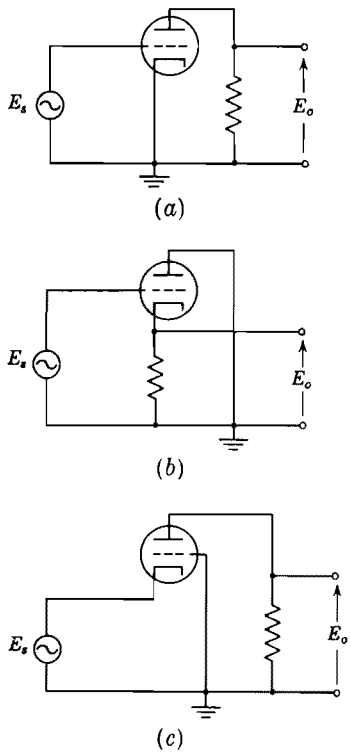



FIG. 4-1. (*a*) Grounded-cathode amplifier. (*b*) Cathode follower. (*c*) Grounded-grid amplifier.

For this study the tube will be considered as a linear-circuit device under the equivalent-circuit assumption, and therefore only small-signal inputs will be considered. Basic circuit performance will be studied in this chapter, response of these circuits as functions of input frequency and waveform being deferred to Chap. 5.

**4-1. Basic Amplifier Types.** To emphasize the fundamental differences between the three important basic circuits, they have been simplified in Fig. 4-1, showing only the a-c signal circuits, no power supplies or coupling circuits being incorporated.

In each circuit one tube element is common to both input and output circuits, and this common element is usually placed at ground or reference a-c potential. In Fig. 4-1*a* this common electrode is the cathode, and the circuit is sometimes referred to as the *grounded-cathode* amplifier. It is the most generally used type for voltage amplification, at both low and high frequencies.

In Fig. 4-1*b* the anode is grounded and output taken from the load in the cathode circuit. Since the cathode voltage is in phase with the grid input signal and tends to go up or down as the grid voltage does, the circuit is referred to as a *cathode follower*. It has a voltage gain less than unity and is of value as a current amplifier and as a step-down impedance transformer for matching of a high impedance to a lower impedance. It also supplies an output in phase with the input signal.

The third circuit, in Fig. 4-1*c*, utilizes the grid as the common or grounded element and is known as the *grounded-grid* amplifier. Input is supplied to the cathode circuit, and besides providing voltage gain the circuit serves as a step-up impedance transformer. Because of the electric-field shielding provided by the grounded grid between the cathode and anode, or input and output circuits, the circuit has applications at high frequencies, where isolation of input and output circuits is important.

While drawn for triodes, all three basic circuits may be employed with pentodes, although most pentode applications occur with the circuit of Fig. 4-1*a*. This choice is dictated by the fact that in Fig. 4-1*b* and *c* the cathode is not at zero a-c potential, and capacitive bypassing of the screen grid to place it at cathode potential becomes more difficult and expensive. In addition, one usual reason for employing pentodes is the good internal shielding between input and output circuits, and in Fig. 4-1*b* this feature is not especially required, and in Fig. 4-1*c* the grounded grid provides just such shielding with a triode.

**4-2. Coupling of Cascaded Amplifiers.** In electronic amplifiers a small input signal is successively amplified by a chain of *cascaded* amplifiers until the desired output voltage or power magnitude is reached. In cascaded circuits the output of one tube becomes the input signal for a second tube. In circuits designed to amplify a-c signals only, means must be provided to introduce the a-c voltage drop across the load into the input circuit of the second tube and at the same time to block any d-c voltages present.

Alternating-current coupling circuits as shown in Fig. 4-2 are utilized to perform these functions, and employ a resistor $R_a$ to provide a load-voltage drop and a capacitor to block out any d-c voltages, as in Fig. 4-2*a* for resistance-capacity- (*RC*-) coupled stages, an inductor and capacitor as in Fig. 4-2*b* for choke-coupled amplifiers, or a transformer as in Fig. 4-2*c*. A grid leak $R_g$ is required to provide a d-c leakage path as a return for any electrons which happen to reach the grid of the second tube. Obviously the winding of the transformer performs this function in the transformer-coupled amplifier.

The grid leak also furnishes a path for application of the bias voltage. Its resistance should be high enough so that its presence does not greatly reduce the load impedance value of $R_a$ or $\omega L$ for the first tube and low

enough so that a charge will not be built up on the grid by the few electrons reaching it. The grid leak usually has a value between 200,000 and 1,000,000 ohms.

In a cascaded amplifier the over-all voltage gain, given as a voltage ratio as in Sec. 3-8, is the product of the individual-stage gains. When gain is expressed in decibels, as in Sec. 3-17, the over-all gain is the sum

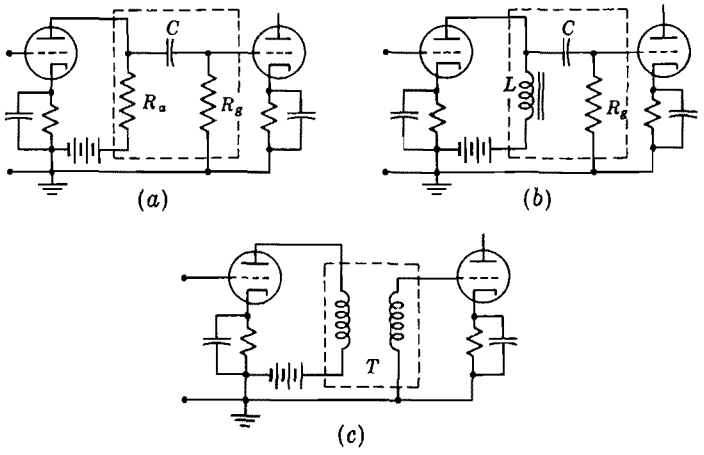



FIG. 4-2. Alternating-current-amplifier coupling circuits. (*a*) *RC* coupling. (*b*) Choke coupling. (*c*) Transformer coupling.

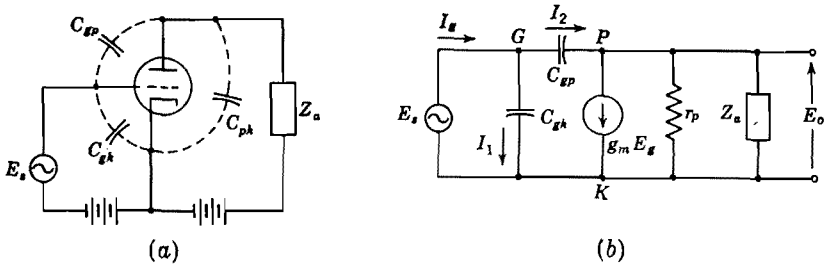



FIG. 4-3. (*a*) Triode, showing internal tube capacities. (*b*) Equivalent circuit for (*a*) with $C_{pk}$ included in $Z_a$.

of the individual-stage decibel gains. This ease of use of decibel gain is an important reason for employing the decibel unit.

Because of the presence of reactive elements in the coupling circuits and in the internal tube capacities, frequency, phase, and transient distortion are produced in amplifiers. These effects will be discussed in Chap. 5.

**4-3. Input Admittance of a Triode as a Grounded-cathode Amplifier.** While current flow due to electrons attracted to a negative control grid is negligibly small, yet the physical capacities between tube elements will

contribute to a current in the grid circuit of a triode. The presence of this current, and the capacities, may be of great importance in some applications.

The tube capacities are shown and labeled in Fig. 4-3*a*. While the internal tube capacities may approximate only a few micromicrofarads each, yet their effect on circuit operation can be considerable by reason of their connection into an active circuit.

Considering the output capacity $C_{pk}$ as part of the load $\mathbf{Z}_a$, since it is directly in parallel, leads to the equivalent circuit of Fig. 4-3*b*. The alternating current to the grid can be written as

$$\mathbf{I}_g = \mathbf{I}_1 + \mathbf{I}_2 \tag{4-1}$$

wherein

$$\mathbf{I}_1 = j\omega C_{gk}\mathbf{E}_s$$
$$\mathbf{I}_2 = -j\omega C_{gp}(\mathbf{E}_o - \mathbf{E}_s)$$

For most circuits and tubes

$$\left| \frac{r_p \mathbf{Z}_a}{r_p + \mathbf{Z}_a} \right| \ll \frac{1}{\omega C_{gp}}$$

which is equivalent to stating that the current $g_m\mathbf{E}_g = g_m\mathbf{E}_s$ is large with respect to $\mathbf{I}_2$. It is then possible to write

$$\mathbf{E}_o = -g_m\mathbf{E}_s \frac{r_p\mathbf{Z}_a}{r_p + \mathbf{Z}_a} \tag{4-2}$$

and

$$\mathbf{I}_g = j\omega C_{gk}\mathbf{E}_s + j\omega C_{gp}\mathbf{E}_s\left(1 + g_m \frac{r_p\mathbf{Z}_a}{r_p + \mathbf{Z}_a}\right) \tag{4-3}$$

Since the input admittance $\mathbf{Y}_g = \mathbf{I}_g/\mathbf{E}_s$, then

$$\mathbf{Y}_g = j\omega\left[C_{gk} + C_{gp}\left(1 + g_m \frac{r_p\mathbf{Z}_a}{r_p + \mathbf{Z}_a}\right)\right] \tag{4-4}$$

Equation (4-4) is the expression for the input admittance of a triode in the grounded-cathode circuit, under the assumptions. It can be seen that the admittance is due to considerably more than the presence of $C_{gk}$. The effect of the active nature of the circuit and the tube equivalent generator is seen by the presence of $g_m$ and $r_p$.

Since the parallel impedance in the plate circuit will be complex in general, let

$$\mathbf{Z}' = \frac{r_p\mathbf{Z}_a}{r_p + \mathbf{Z}_a} = R' \pm jX'$$

the plus or minus sign depending on the type of reactance present in the load. The input admittance will then be

$$\mathbf{Y}_g = -g_m\omega C_{gp}(\pm X') + j\omega[C_{gk} + C_{gp}(1 + g_m R')] \tag{4-5}$$

In Fig. 4-4 the input admittance of the circuit shown can be written as

$$\mathbf{Y}_g = \frac{1}{R_{in}} + j\omega C_{in}$$

and by comparison with Eq. (4-5) it can be seen that Fig. 4-4 represents the triode input circuit, where

$$R_{in} = -\frac{1}{g_m \omega C_{gp}(\pm X')} \tag{4-6}$$

$$C_{in} = C_{gk} + C_{gp}(1 + g_m R') \tag{4-7}$$

The input capacity is considerably greater than $C_{gk}$ alone, and the signs in the expression for $R_{in}$ lead to three possible cases:

| Load | Input resistance | Input capacity |
|---|---|---|
| Capacitive | Positive | $C_{in}$ |
| Resistive | Infinite, or open circuit | $C_{in}$ |
| Inductive | Negative | $C_{in}$ |

Thus for capacitive loads the grid input circuit represents a resistive load on the signal source, its value being inversely proportional to frequency. For inductive loads the negative grid input resistance indicates a feedback of power from the plate circuit through $C_{gp}$ to the grid circuit, and this may lead to instability and oscillation. That is, the negative input resistance cancels the positive grid circuit resistance and may produce zero or a net negative circuit resistance. Such a circuit is inherently an oscillator.

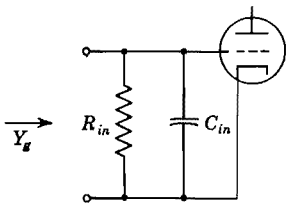



FIG. 4-4. Input circuit of a triode.

Because of the incipient oscillation caused by the above feedback of power in triode circuits, the capabilities of a triode are limited to amplification at only the lower radio frequencies. However, a pentode, in which $C_{gp}$ is made very small by design, does not produce this feedback and is capable of stable amplification at frequencies of some hundreds of megacycles, in the grounded-cathode circuit. When triodes are employed at such frequencies, the grounded-grid circuit is used because of the excellent internal shielding provided by this circuit.

**Example.** One triode of a 12AU7 double triode having $r_p = 7700$ ohms, $g_m = 2200$ $\mu$mhos, $C_{gp} = 1.5$ $\mu\mu$f, $C_{gk} = 1.6$ $\mu\mu$f, and $C_{pk} = 0.5$ $\mu\mu$f is operated with a resistive load of 40,000 ohms. Find the equivalent input-circuit quantities for this tube at frequencies of 10,000 cycles and 10 Mc, in the grounded-cathode circuit of Fig. 4-1*a*.

It was assumed that $Z_a$ was composed of $R_a$ and $C_{pk}$ in parallel. At 10,000 cycles the admittance of this combination is

$$\mathbf{Y}_a = \frac{1}{R_a} + j\omega C_{pk} = 0.25 \times 10^{-4} + j2\pi \times 10^4 \times 0.5 \times 10^{-12}$$
$$= 0.25 \times 10^{-4} + j3.14 \times 10^{-8} = 0.25 \times 10^{-4} \text{ mho}$$

Therefore $\mathbf{Z}_a = 40{,}000$ ohms

The quantity $r_p\mathbf{Z}_a/(r_p + \mathbf{Z}_a)$ is needed and can be computed as

$$\frac{r_p\mathbf{Z}_a}{r_p + \mathbf{Z}_a} = R' + jX' = \frac{7700 \times 40{,}000}{7700 + 40{,}000} = 6430 \text{ ohms}$$

Since $X'$ must be zero, then

$$R_{in} = -\frac{1}{g_m\omega C_{gp}(+X')} = \infty$$

Computing the input capacity,

$$C_{in} = C_{gk} + C_{gp}(1 + g_mR')$$
$$= 1.6 + 1.5(1 + 0.0022 \times 6430)$$
$$= 1.6 + 1.5 \times 15.2$$
$$= 24.2 \ \mu\mu\text{f} \qquad \text{at 10,000 cycles}$$

This capacity is very much larger than $C_{gk}$ alone.

At a frequency of 10 Mc

$$\mathbf{Y}_a = 0.25 \times 10^{-4} + j2\pi \times 10^7 \times 0.5 \times 10^{-12}$$
$$= 0.402 \times 10^{-4}/51°$$
$$\mathbf{Z}_a = 2.48 \times 10^4/-51°$$

Then

$$\frac{r_p\mathbf{Z}_a}{r_p + \mathbf{Z}_a} = \frac{7700 \times 24{,}800/-51°}{7700 + (15.3 - j19.3) \times 10^3}$$
$$R' + jX' = 5.03 \times 10^3/-11° = (4.9 - j0.91) \times 10^3$$

Therefore

$$R_{in} = -\frac{1}{g_m\omega C_{gp}(-X')} = \frac{1}{0.0022 \times 2\pi \times 10^7 \times 1.5 \times 10^{-12} \times 0.91 \times 10^3}$$
$$= 5300 \text{ ohms}$$
$$C_{in} = C_{gk} + C_{gp}(1 + g_mR')$$
$$= 1.6 + 1.5(1 + 0.0022 \times 4.9 \times 10^3)$$
$$= 1.6 + 1.5 \times 11.8$$
$$= 19.3 \ \mu\mu\text{f}$$

At 10 Mc the input of a 12AU7 triode with 40,000 ohms resistive plate load appears as a 19.3-$\mu\mu$f capacitor in parallel with 5300 ohms. This is usually considered an undesirably low input resistance and makes such a triode unsuited for use at such frequencies.

**4-4. Grounded-cathode Amplifier Analysis.** The *RC*-coupled amplifier of Fig. 4-5 is probably the most commonly used form of circuit, being cascaded to realize high over-all voltage gain in many types of equipment. The basic form of the triode *RC* amplifier reduces to the equivalent circuit of Fig. 4-6*a*, from which it can be seen that

$$\mathbf{I}_p = \frac{\mu\mathbf{E}_g}{r_p + \mathbf{Z}_a} \tag{4-8}$$

and by definition $\mathbf{E}_o = -\mathbf{I}_p\mathbf{Z}_a$, so that

$$\mathbf{E}_o = \frac{-\mu\mathbf{Z}_a\mathbf{E}_g}{r_p + \mathbf{Z}_a}$$

In amplifier practice the *voltage gain* as defined in Sec. 3-8 may be found for the grounded-cathode amplifier as

$$\text{Gain} = \mathbf{A} = \frac{\mathbf{E}_o}{\mathbf{E}_s} = \frac{\mathbf{E}_o}{\mathbf{E}_g} = \frac{-\mu\mathbf{Z}_a}{r_p + \mathbf{Z}_a} = \frac{-g_m\mathbf{Z}_a}{1 + \mathbf{Z}_a/r_p} \tag{4-9}$$

This gain figure is dimensionless and is generally a complex number. The phase reversal indicated by the negative sign is inherent in this type

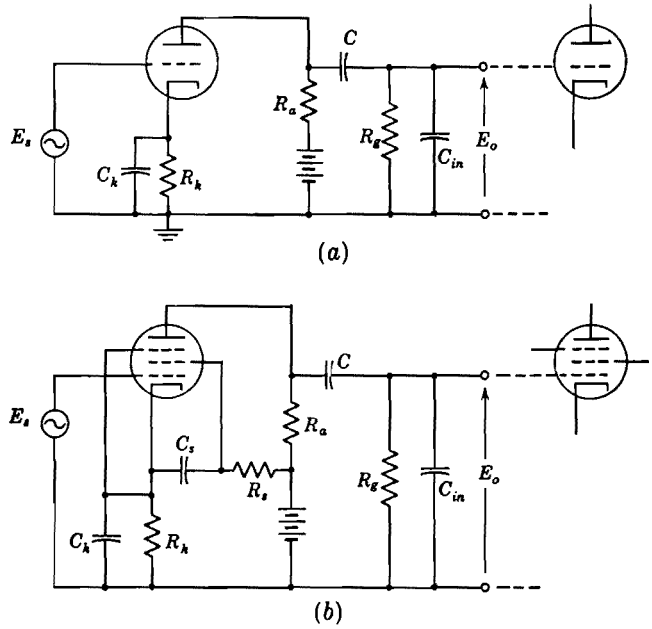



FIG. 4-5. (*a*) Triode *RC*-coupled grounded-cathode amplifier. (*b*) Same, for pentode.

of circuit; the angle found in addition to the 180° inherent shift is called the phase shift of the circuit.

Maximum voltage gain can be obtained by making $\mathbf{Z}_a$ large with respect to $r_p$, as can be seen by rewriting Eq. (4-9) as

$$\mathbf{A} = \frac{-\mu}{1 + r_p/\mathbf{Z}_a}$$

However, if resistance is used as $Z_a$, then the large d-c voltage supply needed to overcome the d-c drop in the load resistor becomes uneconomic or even unsafe, and extremely large inductive impedances are expensive

and vary in value with frequency because of the inherent capacity present in the windings. Loads of 100,000 to 250,000 ohms become approximate maximums, these usually achieving desirably high gains per stage, without need for d-c supplies exceeding 300 to 350 volts. The asymptotic approach toward $\mu$ as the potential total gain, shown in Fig. 4-7, also demonstrates that very large values of $Z_a$ are not required.

If a pentode is employed in the grounded-cathode circuit as in Fig. 4-5*b*, the equivalent circuit can be reduced to that in Fig. 4-6*b*, wherein all signal-carrying impedances in the plate circuit are lumped as $\mathbf{Z}_a$. For the reasons discussed in the paragraph above, pentode load impedances are usually very small with respect to the plate resistances because of the high values of $r_p$. However, with the high available $\mu$ values high gains are still possible. Essentially the pentode tube is a constant-current generator of value $g_m\mathbf{E}_g$, and if the plate resistance is neglected as a very large shunt value, an equivalent circuit as in Fig. 4-6*c* is obtained. The output-voltage rise is immediately obtainable as

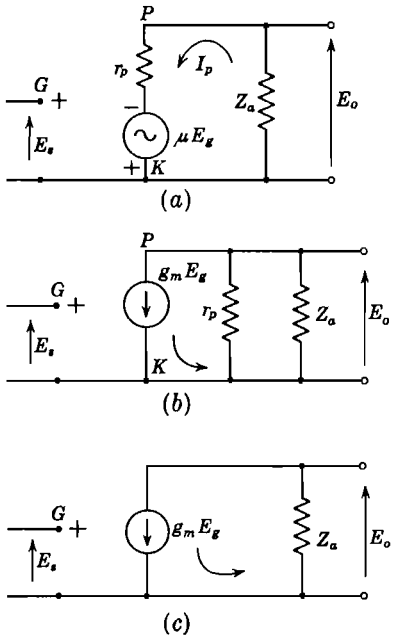



FIG. 4-6. Equivalent circuit for triode (*a* or *b*), for pentode (*b* or *c*).

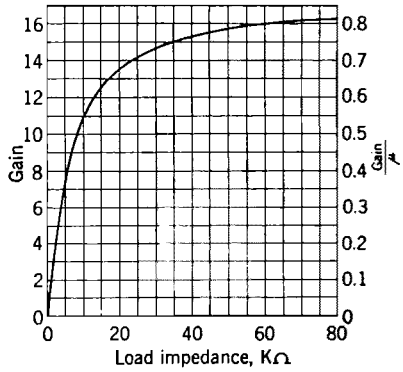



FIG. 4-7. Measured variation of gain with load impedance for a 6J5 tube with $E_{bb}$ = 250 volts, $E_{cc}$ = −6 volts, $\mu$ = 20, $r_p$ = 7700 ohms rated.

$$\mathbf{E}_o = -g_m\mathbf{E}_g\mathbf{Z}_a$$

from which the gain for a high-plate-resistance pentode in the grounded-cathode circuit is

$$\text{Gain} = \mathbf{A} = \frac{\mathbf{E}_o}{\mathbf{E}_s} = \frac{\mathbf{E}_o}{\mathbf{E}_g} = -g_m\mathbf{Z}_a \tag{4-10}$$

which is a simple expression.

It should be noted that the second form of the gain expression for the triode in Eq. (4-9), and the above equation, both show the gain proportional to the transconductance $g_m$ of the tube. This is an important criterion of use in selection of a suitable tube type for a voltage amplifier.

**Example.** An input signal of 1 volt rms at 1000 cycles is applied to the grid of the triode in Fig. 4-8. The load impedance is $\mathbf{Z}_a = 10{,}000 + j16{,}000$ ohms, and the frequency is considered low enough so that the internal tube capacities may be neglected. Find the a-c anode current $\mathbf{I}_p$ and the gain of the circuit if $\mu = 20$ and $r_p = 10{,}000$ ohms.

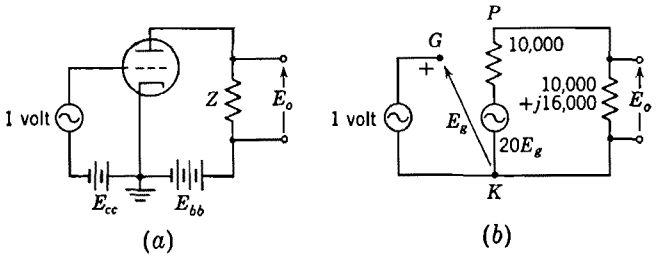



FIG. 4-8. (*a*) Triode amplifier, and (*b*) the equivalent circuit.

Referring to the circuit of Fig. 4-8*b*, the cathode-grid voltage rise is $\mathbf{E}_g = 1$, and $\mu\mathbf{E}_g = 20$. Thus by Eq. (4-8)

$$\mathbf{I}_p = \frac{\mu\mathbf{E}_g}{r_p + \mathbf{Z}_a} = \frac{20}{10{,}000 + 10{,}000 + j16{,}000}$$
$$= \frac{20}{25{,}600/\underline{38.6°}} = 0.000782/\underline{-38.6°} \text{ amp}$$

The output-voltage rise then is

$$\mathbf{E}_o = -\mathbf{I}_p\mathbf{Z}_a = -7.82 \times 10^{-4}/\underline{-38.6°} \times 18{,}900/\underline{58°}$$
$$= -14.8/\underline{19.4°} = 14.8/\underline{199.4°} \text{ volts}$$

The gain therefore is

$$\mathbf{A} = \frac{\mathbf{E}_o}{\mathbf{E}_g} = \frac{14.8/\underline{199.4°}}{1/\underline{0°}} = 14.8/\underline{199.4°}$$

The amplifier has a gain of 14.8 with a phase shift of 19.4°, in addition to the 180° phase shift which is inherent in this type of circuit.

The input impedance for both triodes and pentodes in this circuit was discussed in Sec. 4-3. By consideration of Thévenin's theorem it is apparent that the output impedance of a tube as the generator in this circuit is equal to $r_p$ shunted by $C_{pk}$. However, when a tube is called upon to supply appreciable power output, considerations of distortion of waveform enter, as discussed in Chap. 6, and the theoretical conditions for impedance matching and maximum power transfer cannot usually be applied.

**4-5. The Cathode-follower Amplifier.** To form the *cathode-follower* amplifier, the load $R_k$ is placed in the cathode circuit, and the anode becomes the common or grounded element (for signal voltages), as in Fig. 4-9.

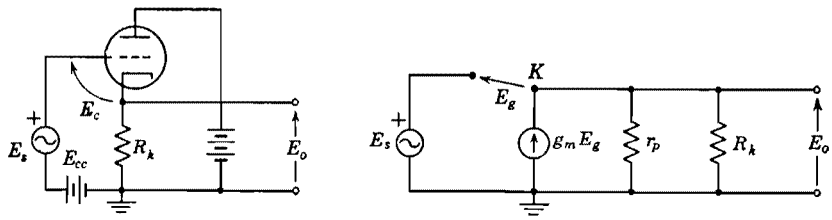



FIG. 4-9. (*a*) Cathode-follower amplifier. (*b*) Equivalent circuit for (*a*).

Neglecting the internal tube capacities for the moment, and considering the equivalent circuit of Fig. 4-9*b*, the following current summation may be written,

$$g_m \mathbf{E}_g - \mathbf{E}_o \left( \frac{1}{r_p} + \frac{1}{R_k} \right) = 0 \tag{4-11}$$

$$\mathbf{E}_g = \mathbf{E}_s - \mathbf{E}_o \tag{4-12}$$

where $\mathbf{E}_o$ is used as indicated as a rise from ground or common terminal to the cathode. Substitution gives

$$\mathbf{E}_o \left( 1 + \frac{1}{\mu} + \frac{1}{g_m R_k} \right) = \mathbf{E}_s$$

from which

$$\mathbf{A} = \frac{\mathbf{E}_o}{\mathbf{E}_s} = \frac{1}{1 + 1/\mu + 1/g_m R_k} = \frac{\mu}{1 + \mu + r_p/R_k} \tag{4-13}$$

$$= \frac{\mu R_k}{r_p + (\mu + 1) R_k} \tag{4-14}$$

Equation (4-13) shows that the gain is always less than unity, approaching $\mu/(\mu + 1)$ for $r_p \ll R_k$.

The internal plate-circuit impedance or output impedance of the amplifier is important in designing the circuit to match its load for maximum power output, and in studying the frequency response of the circuit. The output impedance may be found by inserting a current in Fig. 4-9*b* of value $\mathbf{I}_o$ and flowing into the circuit at the upper $\mathbf{E}_o$ terminal. The ratio of $\mathbf{E}_o$ to this current $\mathbf{I}_o$, taken with the internal source $\mathbf{E}_s$ short-circuited, will give the output impedance.

Writing a current summation,

$$g_m \mathbf{E}_g + \mathbf{I}_o - \mathbf{E}_o \left( \frac{1}{r_p} + \frac{1}{R_k} \right) = 0 \tag{4-15}$$

and shorting $\mathbf{E}_g$ shows that

$$\mathbf{E}_g = -\mathbf{E}_o \tag{4-16}$$

After a substitution

$$\mathbf{E}_o\left(g_m + \frac{1}{r_p} + \frac{1}{R_k}\right) = \mathbf{I}_o$$

which yields

$$\mathbf{Z}_{out} = \frac{\mathbf{E}_o}{\mathbf{I}_o} = \frac{1}{g_m + 1/r_p + 1/R_k} = \frac{r_p R_k}{r_p + (\mu + 1)R_k} \tag{4-17}$$

$$= \frac{R_k}{1 + g_m R_k + R_k/r_p} \tag{4-18}$$

This is the impedance a cathode follower presents to any load connected to its output circuit. For a pentode with $R_k \ll r_p$ it approximates

$$\mathbf{Z}_{out} = \frac{1}{g_m + 1/R_k}$$

which reduces to $1/g_m$ for usual circuits.

Equation (4-17) may be rearranged to

$$\mathbf{Z}_{out} = \frac{r_p/(\mu + 1)R_k}{r_p/(\mu + 1) + R_k} \tag{4-19}$$

which has a simple physical interpretation, since the expression is that of two resistors in parallel, as in Fig. 4-10. The output impedance of the cathode follower can be seen to be represented by the resistor $R_k$ in parallel with a resistor of value $r_p/(\mu + 1)$, which represents the effect of the tube.

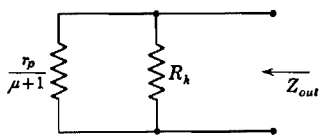



FIG. 4-10. Effective output circuit of the cathode follower.

For a 12AU7 triode having $r_p = 7700$ ohms, $\mu = 20$ the resistor effectively paralleling $R_k$ is 385 ohms. This gives an output-impedance value which makes the cathode follower valuable as a line-matching circuit or as a step-down impedance transformer over wide frequency ranges.

In parallel with this output impedance will be the plate-to-cathode capacity $C_{pk}$, and the cathode-to-heater capacity $C_{hk}$, if the heater circuit is grounded at signal frequency, as is usually the case. However, these capacities do not impose serious limitation on the high frequency limit or response of the circuit because of the small value of $Z_{out}$. For the 12AU7 mentioned above, $C_{pk} = 0.50\ \mu\mu f$, $C_{hk} = 2.5\ \mu\mu f$, and the reactance of these capacities in parallel at 13.8 Mc is 3850 ohms, or ten times $r_p/(\mu + 1)$. Thus operation at least to this frequency is possible.

The cathode follower also has an input admittance which is considerably smaller than the grounded-cathode circuit. In the equivalent cir-

cuit of Fig. 4-11, which includes the various tube capacities except $C_{pk}$, which has already been found of negligible effect on $R_k$, the currents may be written as

$$\begin{aligned} \mathbf{I}_g &= \mathbf{I}_1 + \mathbf{I}_2 \\ \mathbf{I}_1 &= j\omega C_{gk}(\mathbf{E}_s - \mathbf{E}_o) \\ \mathbf{I}_2 &= j\omega C_{gp}\mathbf{E}_s \end{aligned} \tag{4-20}$$

Substitution in Eq. (4-20), and division by $E_s$, gives

$$\mathbf{Y}_g = \frac{\mathbf{I}_g}{\mathbf{E}_s} = j\omega \left[ C_{gp} + C_{gk}\left(1 - \frac{\mathbf{E}_o}{\mathbf{E}_s}\right)\right] \tag{1-21}$$

and since $\mathbf{E}_o/\mathbf{E}_s$ is the gain $\mathbf{A}$,

$$\mathbf{Y}_g = j\omega[C_{gp} + C_{gk}(1 - \mathbf{A})] \tag{4-22}$$

With a resistive load $R_k$ and neglecting the effect of $C_{pk}$, the input admittance is due to a capacity

$$C_{in} = C_{gp} + C_{gk}(1 - |A|) \tag{4-23}$$

Since the gain $|A|$ is usually less than but near unity in magnitude, the input capacity is reduced nearly to that of $C_{gp}$ alone, a major advantage

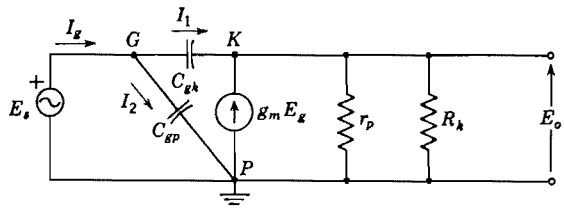



FIG. 4-11. Equivalent circuit of the cathode follower, including $C_{gp}$ and $C_{gk}$.

in reduction of input admittance at high frequencies. The circuit is thus of use where a very high impedance is to be coupled to a low impedance load, over a wide frequency range.

**4-6. Design of Cathode-follower Circuits.** Many of the design parameters of the usual grounded-cathode amplifier are arbitrary, but this is not usually true for the cathode follower. After choice of a value of $R_k$, based either on obtaining a desired gain or desired output impedance or on matching a load impedance, the rest of the circuit may be readily specified.

If the output impedance is specified, a desirable tube may be quickly chosen since

$$\mathbf{Z}_{out} \cong \frac{r_p}{\mu + 1} \cong \frac{1}{g_m} \qquad \text{for } \mu > 10$$

The ratio $R_k/r_p$ and a value for $R_k$ may then be selected from Eq. (4-13)

by use of Fig. 4-12, using the value of $\mu$ of the chosen tube, to give a desired value of gain. It can be seen from the curves that any value of $\mu > 10$ and $R_k/r_p > 1$ gives a gain greater than 0.8, and choice of $R_k$

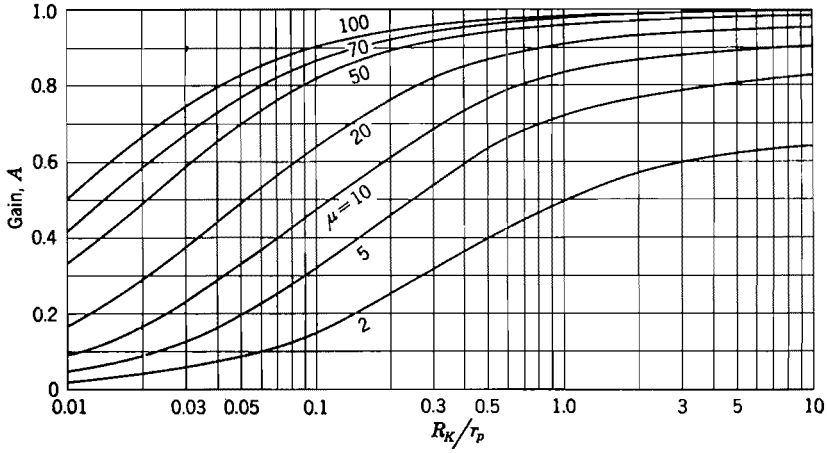



FIG. 4-12. Cathode-follower gain as a function of $R_k/r_p$ and $\mu$.

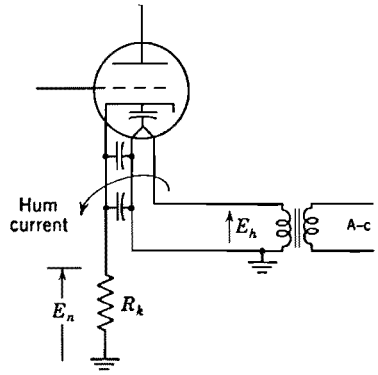



FIG. 4-13. Path of a-c hum current into $R_k$.

is not critical. However, if the cathode is heated with alternating current, large values of $R_k$ may introduce power-supply hum into the circuit through the heater-cathode capacitance, as indicated in Fig. 4-13. If this hum is undesirable, then $R_k$ values above about 50,000 ohms should usually be avoided.

**Example.** A cathode follower is designed with a cathode resistor of 50,000 ohms. The heater is supplied with 6.3 volts, 60 cycles, one side of the heater circuit is grounded, and $C_{hk} = 2.5$ $\mu\mu$f. Find the hum voltage introduced on the cathode.

The cathode resistor and $C_{hk}$ may be considered in series with 6.3 volts a-c and ground. The reactance of the capacity is

$$X_C = \frac{1}{2\pi \times 60 \times 2.5 \times 10^{-12}} = 1.06 \times 10^9 \text{ ohms}$$

The hum voltage $E_n$ then will be

$$E_n \cong \frac{R_k}{X_C} E_k = \frac{5 \times 10^4 \times 6.3}{1.06 \times 10^9}$$
$$\cong 298 \times 10^{-6} \text{ volt} = 298 \ \mu\text{volts}$$

In a cathode follower used as an impedance transformer for a low-voltage-level high-impedance transducer, such as a photoelectric cell, a voltage of this magnitude might be quite appreciable compared with the desired signal.

The tube and load have now been determined. Grid-bias voltage and the magnitude of possible input and output signals remain to be specified. A load line for $R_k$ as a load should be drawn on a set of plate characteristics for the available $E_{bb}$ value. Intersections of this line and the characteristics then determine actual $e_c$ or total grid-cathode voltages and instantaneous plate currents. Referring to Fig. 4-9, and writing a grid-circuit equation in instantaneous values, gives

$$e_c = e_s - e_o$$

Noting that $e_o = i_b R_k$, then

$$e_s = e_c + i_b R_k \tag{4-24}$$

Using as values of $e_c$ those voltages given by the grid-voltage lines on the plate family, and the currents at which the load line intersects the grid-voltage lines, a table of values of $e_c$, $i_b$, $i_b R_k = e_o$, and $e_s$ may be calculated. This may be illustrated using the load-line and plate family of Fig. 4-14.

Assuming a 6J5 triode, with $\mu = 20$, $r_p = 7700$ ohms, $R_k = 30{,}000$ ohms, and $E_{bb} = 300$ volts, a load line is drawn on the tube characteristics as in Fig. 4-14. Using the intersections of the load line and a few of the $E_c$ curves to determine $e_c$ and corresponding $i_b$ values, the accompanying

| $e_c$, volts | $i_b$, ma | $i_b R_k$, volts | $e_s$, volts |
|---|---|---|---|
| 0 | 7.4 | 222 | 222 |
| − 4 | 5.3 | 159 | 155 |
| − 8 | 3.3 | 99 | 91 |
| −12 | 1.6 | 48 | 36 |
| −16 | 0.45 | 13.5 | − 2.5 |
| −18 | 0.0 | 0.0 | −18 |

table is constructed by use of Eq. (4-24) for the $e_s$ value. For instance, the intersection of the $e_c = -4$ volt line and the load line shows a plate current of 5.3 ma. The value of $i_b R_k$ is then $30{,}000 \times 0.0053 = 159$ volts.

From Eq. (4-24)

$$e_s = -4 + 159 = 155 \text{ volts}$$

and this is recorded in the last column, as the required input signal.

Comparison of the $i_b R_k = e_o$ column and the $e_s$ column shows how the output voltage tends to follow or approach the input signal and thus explains the name of cathode follower for the circuit.

A plot of $e_s$ against $i_b R_k = e_o$ gives the line of Fig. 4-15. This is the input-output relation of the cathode follower and is linear except for a

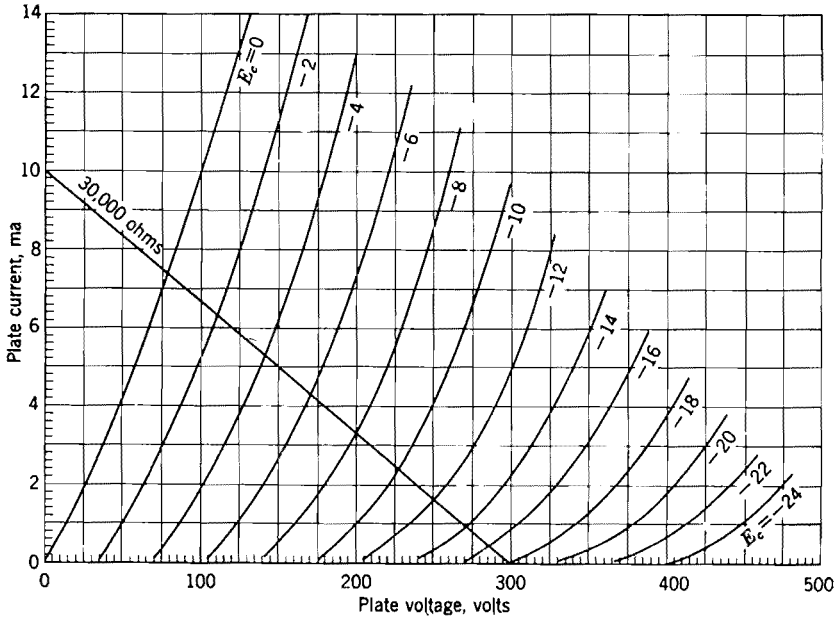



FIG. 4-14 Load line for cathode follower, $E_{bb}$ = 300 volts, $R_k$ = 30,000 ohms.

very slight curvature very near the cutoff condition. From the table it can be seen that the grid is being swung from $e_c = 0$, or zero grid-cathode volts, to cutoff at −18 volts. The table also shows that an input signal of 222 volts positive can be handled without driving the grid above zero volts or positive, but a negative signal of only −18 volts will drive the grid to cutoff. In some applications with only positive-going signals this might be satisfactory, but for a symmetrical signal a better $Q$ point may be found.

Figure 4-15 indicates a possible total input signal of 240 volts, which would represent a symmetrical signal of 120 volts peak. To handle such an input signal satisfactorily requires that the $Q$ point be located at the mid-point of the linear input-output curve. Since the line covers

an input total voltage of 240, the mid-point will be 120 volts below the positive peak value, or at 222 − 120 = +102 volts, and such a point is shown in the figure as $Q'$. The value of $e_s = 102$ volts will then be *grid-to-ground* bias, and the curve shows that at the same time $e_o = 109$ volts, which is the *cathode-to-ground* potential. Thus the actual $Q$-point *grid-to-cathode* bias is

$$e_s - e_o = 102 - 109 = -7 \text{ volts} = e_c$$

If these potential levels be supplied to grid and cathode, the $Q$ point will be at $Q'$ in Fig. 4-15, and the circuit will handle a symmetrical peak-to-peak input signal of 240 volts, giving a peak-to-peak output of 222 volts.

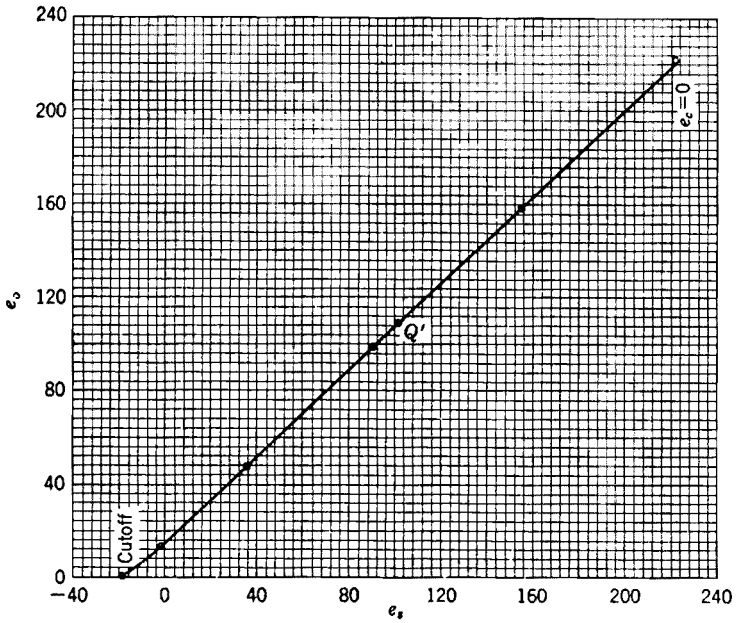



FIG. 4-15. Output vs. input volts for a cathode follower.

It is not necessary to draw the line of Fig. 4-15 or to calculate the table of values if the only design information needed is the proper bias potentials and if a knowledge of the departure from linearity is not desired. The bias values may be found directly by noting that at one extreme of operation $e_c = 0$, or the grid is about to go positive, and at the other limit $i_b = e_o = 0$, or the tube is cut off. Directly from the tube curves of Fig. 4-14, at $e_c = 0$ it is seen that the load line indicates $i_b = 7.4$ ma. Then $e_o = 0.0074 \times 30{,}000 = 222$ volts, and since $e_c = 0$, Eq. (4-24) indicates that $e_s = e_o = 222$ volts as well.

At the other extreme end of the load line the tube cuts off at $e_c = -18$

volts, and since $i_b R_k = e_o = 0$, then by Eq. (4-24), $e_s = e_c = -18$ volts. These operations supply the data in the accompanying table. Thus the

| | At zero grid volts | At cutoff |
|---|---|---|
| $e_s$ | 222 | −18 |
| $e_o$ | 222 | 0 |
| $e_c$ | 0 | −18 |

input signal will swing

$$222 - (-18) = 240 \text{ volts}$$

and the output voltage $e_o$ will swing

$$222 - 0 = 222 \text{ volts}$$

and the actual grid-cathode voltage goes from 0 to −18 volts. The $Q'$ point then should be placed at 222 − 120 = 102 volts and the other potentials found as before.

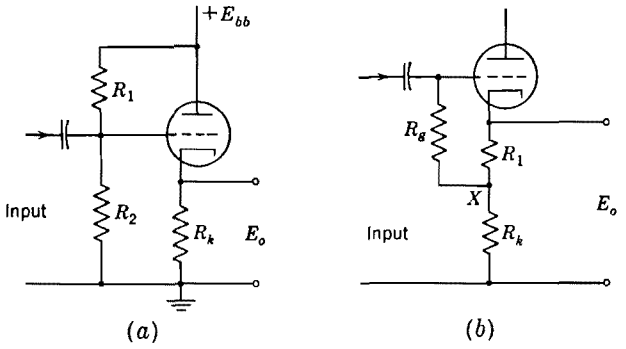



FIG. 4-16. Two methods of biasing a cathode follower.

The two circuits of Fig. 4-16 are practical methods of obtaining the necessary electrode potentials, Fig. 4-16*a* illustrating a simple method of placing the grid at +102 volts to ground, Fig. 4-16*b* showing another method of finding a grid return point at +102 volts to ground, at the same time maintaining the cathode value at the desired +109 volts above ground. The latter circuit has other interesting properties to be discussed in the next section.

In Fig. 4-16*a*, assuming $E_{bb} = 300$ volts, and making

$$R_1 + R_2 = 500{,}000 \text{ ohms}$$

as an arbitrary choice, then

$$R_2 = {}^{102}\!/_{300} \times 500{,}000 = 170{,}000 \text{ ohms}$$

by making use of the voltage-divider action of $R_1$ and $R_2$. Then

$$R_1 = 500{,}000 - 170{,}000 = 330{,}000 \text{ ohms}$$

and the grid will be maintained at 102 volts positive to ground. The cathode-to-ground potential of 109 volts $= e_c$ at the $Q'$ point and is supplied by the drop in the cathode resistor.

In Fig. 4-16*b*, at the $Q'$ point, the voltage from cathode to ground is +109 volts. The potential at point $X$ on the cathode resistor is to be +102 volts, or −7 volts to cathode. Then

$$\frac{R_1}{R_1 + R_k} = \frac{7}{109}$$

and since $R_1 + R_k = 30{,}000$ ohms, then $R_1 = 1930$ ohms.

This completes the design of the cathode-follower circuit. The gain as actually measured from the load line, and as predicted by Eq. (4-13), is

$$\mathbf{A} = {}^{222}\!/_{240} = 0.925$$

By use of the equation

$$\mathbf{A} = \frac{\mu(R_1 + R_k)}{r_p + (\mu + 1)(R_1 + R_k)} = \frac{20 \times 30{,}000}{7700 + 21 \times 30{,}000} = 0.94$$

and this gives a reasonable check.

**4-7. Analysis of the Modified Cathode Follower.** The circuit of Fig. 4-17 can be shown to have an exceedingly high input resistance, as an

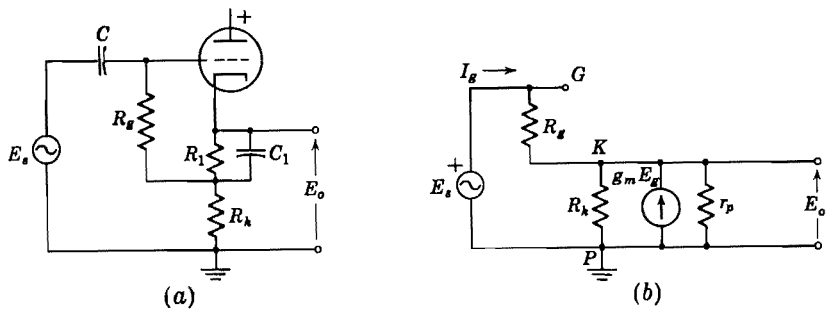



FIG. 4-17. (*a*) Modified cathode follower to provide bias, and (*b*) the equivalent circuit.

added valuable circuit property. This is a result of coupling part of the output circuit back into the input, or of adding an active element (the $g_m E_g$ generator) into the input passive network.

Resistor $R_1$, for bias, will ordinarily be small with respect to $R_k$ and can be dropped from the equivalent circuit of Fig. 4-17*b*. It is occasionally bypassed as shown, this also justifying its elimination from the

circuit of Fig. 4-17*b*. The grid leak $R_g$ will usually be made several megohms for this circuit and will be very large with respect to $R_k$. It is then possible to write a current summation at the cathode $k$ as

$$\mathbf{I}_g + g_m\mathbf{E}_g - \mathbf{E}_o\left(\frac{1}{r_p} + \frac{1}{R_k}\right) = 0 \tag{4-25}$$

and the cathode-grid voltage as

$$\mathbf{E}_g = \mathbf{E}_s - \mathbf{E}_o = \mathbf{I}_g R_g \tag{4-26}$$

To obtain the gain, Eq. (4-26) should be used in (4-25) and $\mathbf{E}_g$ and $\mathbf{I}_g$ eliminated, giving

$$\mathbf{E}_o\left(g_m + \frac{1}{R_g} + \frac{1}{R_k} + \frac{1}{r_p}\right) = \mathbf{E}_s\left(g_m + \frac{1}{R_g}\right)$$

The gain can then be obtained as

$$\mathbf{A} = \frac{\mathbf{E}_o}{\mathbf{E}_s} = \frac{1}{1 + \dfrac{R_g(1 + R_k/r_p)}{R_k(1 + g_m R_g)}} \tag{4-27}$$

As would be expected, the gain is always less than unity. If $R_k/r_p \gg 1$ and $g_m R_g \gg 1$, both usually reasonable approximations, the gain reduces to

$$\mathbf{A} \cong \frac{\mu}{\mu + 1} \tag{4-28}$$

A similar process, eliminating $\mathbf{E}_o$ and $\mathbf{E}_g$ by use of Eq. (4-26), gives

$$\mathbf{I}_g\left(1 + g_m R_g + \frac{R_g}{R_k} + \frac{R_g}{r_p}\right) = \mathbf{E}_s\left(\frac{1}{R_k} + \frac{1}{r_p}\right)$$

The input resistance is given by $\mathbf{E}_s/\mathbf{I}_g$; so

$$R_{in} = \frac{\mathbf{E}_s}{\mathbf{I}_g} = R_g + \frac{R_k(1 + g_m R_g)}{1 + R_k/r_p} \tag{4-29}$$

If the same approximations are made as for the gain, then

$$R_{in} = R_g(\mu + 1) \tag{4-30}$$

The input resistor value can thus be multiplied by a large factor. In certain types of circuits where loading of the source is important, or in other circuits where the time constant represented by $CR_{in}$ must be large, the circuit is of value.

**4-8. The Grounded-grid Amplifier.** A grounded-grid, or cathode-input, amplifier is illustrated in Fig. 4-18. This circuit has a reasonable voltage and power gain, and the grounding of the grid reduces the input

capacity to $C_{gk}$, the output capacity to $C_{gp}$, and since both can be small, then wide-band frequency response is possible. A main feature of the circuit is a step-up impedance ratio, from a low-input to a high-output impedance.

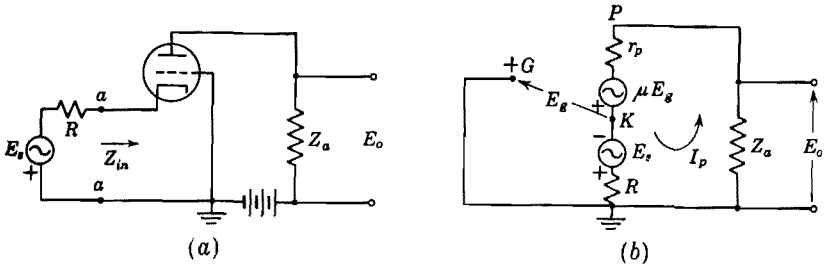



FIG. 4-18. The grounded-grid amplifier and its equivalent circuit.

Since plate current flows through the input source $\mathbf{E}_s$, its internal resistance $R$ is important and is included in the equivalent circuit in Fig. 4-18*b*. The mesh equations are

$$\mu\mathbf{E}_g + \mathbf{E}_s - \mathbf{I}_p(r_p + R + \mathbf{Z}_a) = 0 \tag{4-31}$$

$$\mathbf{E}_g - \mathbf{E}_s + \mathbf{I}_p R = 0 \tag{4-32}$$

and the plate current is

$$\mathbf{I}_p = \frac{(\mu + 1)\mathbf{E}_s}{r_p + (\mu + 1)R + \mathbf{Z}_a}$$

Since the load-voltage rise is $\mathbf{E}_o = -\mathbf{I}_p\mathbf{Z}_a$, and since $\mathbf{E}_s$ also represents a negative rise of voltage above the common terminal, the gain is

$$\mathbf{A} = \frac{\mathbf{E}_o}{\mathbf{E}_s} = \frac{(\mu + 1)\mathbf{Z}_a}{r_p + (\mu + 1)R + \mathbf{Z}_a} \tag{4-33}$$

There is no 180° phase reversal in this circuit. The tube appears to have an amplification factor of $\mu + 1$, since the input voltage not only is amplified but adds directly to the output.

Further analysis of Eq. (4-33) shows that the tube appears to have an internal resistance or effective plate resistance of

$$r_p' = r_p + (\mu + 1)R$$

This is the output impedance of the amplifier.

The input impedance into which the source generator works is of interest. This can be found, neglecting tube capacitances, as

$$\frac{\mathbf{E}_s}{\mathbf{I}_p} = \frac{r_p + (\mu + 1)R + \mathbf{Z}_a}{\mu + 1}$$

$$= R + \frac{r_p + \mathbf{Z}_a}{\mu + 1}$$

Hence the input impedance of the tube alone, as seen at the terminals $a$, $a$ of Fig. 4-18$a$, is

$$\mathbf{Z}_{in} = \frac{r_p + \mathbf{Z}_a}{\mu + 1} \tag{4-34}$$

and thus the circuit steps down the total impedance of the plate circuit by the factor $\mu + 1$. For a tube with $\mu = 20$, $r_p = 7700$ ohms, and with a plate load of 40,000 ohms, the input resistance is 2270 ohms.

Since the input and output circuits are in series, the circuit is analogous to a pump which lifts water from a low-pressure system to a high-pressure line.

**4-9. A-C Operation of Voltage Amplifiers.** Economy of manufacture occasionally dictates that the rectifier power supply be eliminated as a source of d-c plate potential and a tube be operated with a-c anode potential in a control application where the pulses of plate current cause no difficulty. Plate-current conduction then will occur in half-cycle pulses on the positive half cycle of plate voltage. The grid voltage during this half cycle will be controlling, and operation may be understood if studied during the time of positive anode voltage.

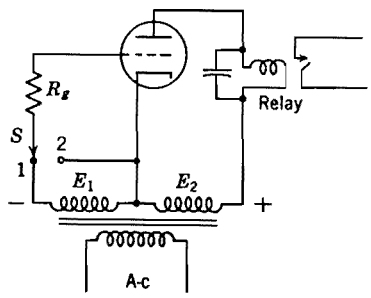



FIG. 4-19. Alternating-current-operated contact amplifier.

The plate-current waveform will be approximate half sine waves, and the average values of these waves will determine the average or d-c current present. The average may be calculated as $1/\pi$ times the peak value, approximately.

A contact amplifier, to operate a relay without carrying appreciable current at the contacts, is shown in Fig. 4-19. When contact $S$ is connected to 1, the grid is biased negatively by voltage $E_1$. If $E_1$ has a value exceeding the value of $E_2/\mu$, the tube will be cut off during this interval. When the contact is shifted to 2, the bias $E_1$ is removed and the grid voltage becomes zero, the plate-current pulses then having a sufficient average value to operate the plate-circuit relay.

If $R_g$ is 1 megohm and $E_1 = 50$ volts peak, the peak current handled by the contacts is always less than 50 $\mu$amp and will have an average value of 16 $\mu$amp during the half cycle in which the grid is positive. This small current will ensure an indefinitely long contact life. To prevent relay chatter, a capacitor of a few microfarads may be connected in parallel with the relay coil to smooth the coil current.

Other applications using a-c anode potential occur in photocell circuits and motor control circuits.

## PROBLEMS

(Throughout, it is assumed that the student has available a tube handbook giving characteristics of the various commercial types called for in the problems.)

**4-1.** A tube with $\mu = 70$, $r_p = 66{,}000$ ohms has a 2.0 volts rms signal applied in a grounded-cathode circuit. The load is 80,000 ohms resistive. Find the a-c plate current, the output a-c voltage, and the decibel gain of the amplifier.

**4-2.** The tube of Prob. 4-1 is operated as a grounded-cathode amplifier at 50,000 cycles with a load of $50{,}000 + j50{,}000$ ohms. If $C_{gp} = 2.5$ $\mu\mu$f, $C_{gk} = 4.0$ $\mu\mu$f, and $C_{pk} = 3.2$ $\mu\mu$f, calculate the input resistance and input capacitance of the tube.

**4-3.** The tubes of Fig. 4-20*a* are identical triodes with $g_m = 2100$ $\mu$mhos, $r_p = 7700$ ohms. If $I_{b1} = 4.0$ ma, $I_{b2} = 2.5$ ma, $E_{cc} = -6.0$ volts for both tubes, calculate the value of $R_k$, and find the over-all voltage gain and decibel gain of the circuit.

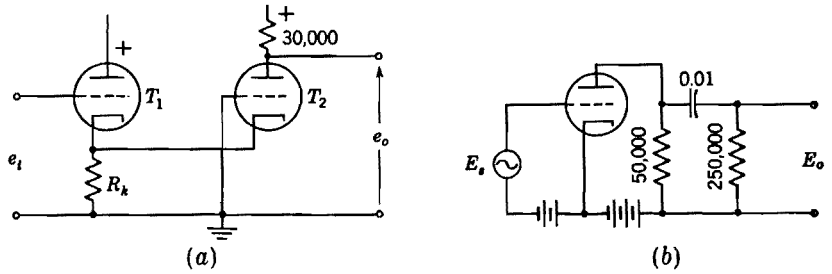



FIG. 4-20

**4-4.** In Fig. 4-20*b*, $g_m = 2000$ $\mu$mhos, $r_p = 30{,}000$ ohms, and $\mathbf{E}_s = 2.0$ volts rms. Calculate the gain and output voltage for a signal at 200 cycles. Repeat for 2000 cycles, and state the kind of distortion present.

**4-5.** A pentode has $g_m = 4500$ $\mu$mhos, $r_p = 1.5$ megohms, and the output voltage is to be 40 volts rms with an a-c plate current of 0.2 ma. (*a*) What signal voltage is required? (*b*) What percentage error in this voltage is introduced by use of the approximate equivalent circuit ($r_p$ = infinite)?

**4-6.** A 6AU6 pentode tube is used in the circuit of Fig. 4-5*b*. For a $Q$ point of $I_b = 3$ ma, $I_{c2} = 0.8$ ma, $E_{bb} = 250$ volts, $E_{c1} = -1.5$ volts, $E_{c2} = 100$ volts, and with $g_m = 4800$ $\mu$mhos, $r_p = 1.5$ megohms, calculate the values of $R_s$ and $R_k$ needed. Find the value of $R_a$ to give a voltage gain of 125.

**4-7.** A cathode-follower amplifier is to use a 12AX7 triode having $\mu = 35$, $r_p = 62{,}000$ ohms, $E_{cc} = -2$ volts. If $I_b = 0.3$ ma, specify the bias resistor, and the load resistor for a gain of 0.85, as in Fig. 4-16*b*, to handle a symmetrical voltage signal.

**4-8.** A 12AX7 (curves in any tube handbook) is to be used as a cathode follower with $E_{bb} = 250$ volts. If it is to have an output impedance of 600 ohms, specify the cathode resistor, determine the maximum allowable symmetrical input-voltage swing, find the gain, and determine the operating bias ($e_c$).

**4-9.** In Fig. 4-16*b* the tube is a 6J5 ($\mu = 20$, $r_p = 7700$ ohms), and $R_k = 25{,}000$ ohms, $R_1 = 2000$ ohms. Find the value of $R_g$ which may be used to give an input resistance

of 150 megohms. With this value of $R_g$, compute the output voltage and gain for $E_s = 2$ volts rms, neglecting the reactance of $C$.

**4-10.** A grounded-grid amplifier uses a tube having $\mu = 10$, $r_p = 9500$ ohms. If $E_s = 1$ volt rms, find (*a*) the output voltage across a 50,000 ohms resistive load if $R_1 = 200$ ohms, (*b*) the power gain of the amplifier in decibels.

**4-11.** A source has a voltage of 0.01 volt and an internal resistance of 300 ohms. Design a grounded-grid amplifier using a 6J6 ($\mu = 38$, $r_p = 7100$ ohms) to have a matching input resistance. What output voltage and decibel power gain will be obtained?

**4-12.** Referring to the circuit of Fig. 4-21, the tubes are identical triodes having $\mu = 10$, $r_p = 8000$ ohms. If the plate currents are 2.0 ma each and $E_{cc} = -8$ volts for both tubes, calculate values of $R_1$ and $R_2$, and find the over-all gain.

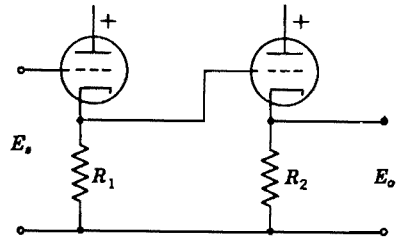



FIG. 4-21

## REFERENCES

1. Richter, W.: Cathode-follower Circuits, *Electronics*, **16**: 112 (November, 1943).
2. Jones, M. C.: Grounded-grid Radio-frequency Voltage Amplifiers, *Proc. IRE*, **32**: 423 (1944).
3. Ryder, J. D.: "Electronic Fundamentals and Applications," Prentice-Hall, Inc., Englewood Cliffs, N. J., 1950.
4. Seely, S.: "Electron-tube Circuits," McGraw-Hill Book Company, Inc., New York, 1950.

CHAPTER 5

# RESPONSE OF SMALL-SIGNAL AMPLIFIERS

In the previous chapter the amplifier in its several forms was studied in general as a circuit element, without regard to the manner in which the tube or load reactances might affect its output response with changes in input frequency or waveform. The analysis and design of amplifiers for steady-state operation over wide frequency ranges and also with pulsed inputs will now be undertaken.

The internal tube capacities, and coupling circuits employed to separate the a-c, or varying, and d-c output components, introduce reactive circuit elements and thus possibilities for frequency and phase distortion. A great variety of waveforms must be handled by amplifiers, and the frequency and phase characteristics must be carefully designed to avoid distortion in operating over wide bands of frequencies. Frequent use of pulses or rectangular waves illustrates the need for wide frequency ranges, since, from the steady-state viewpoint, harmonics above the hundredth are required to preserve rectangular and pulse forms. Intermittent pulsed waveforms also may not permit the attainment of a steady state, and the response of circuits to each isolated pulse must also be determined if waveform is to be preserved.

Thus for linear small-signal amplifiers the circuit designer must consider frequency and phase response on a sinusoidal basis and transient response to pulse and other nonsinusoidal waveforms. The design of amplifiers with suitable responses is discussed in this chapter, largely in terms of the *RC*-coupled grounded-cathode circuit, because of its wide adoption. Extension to other amplifier forms follows by application of the material of Chap. 4.

**5-1. Frequency and Phase Response of the Grounded-cathode Amplifier.** Frequency and phase distortion are due to the presence of reactive elements in amplifier circuits. For instance, in the *RC*-coupled circuit of Fig. 5-1, the tube capacitances $C_{pk}$ and $C_{in}$ and the coupling capacitor $C$ are functions of frequency and may cause distortion. Inductance may at times be introduced, and this is also a function of frequency.

Minimization of frequency and phase distortion is essential to preservation of signal waveform in amplifiers. Analysis of small-signal amplifier

operation to determine the extent of the distortion is usually undertaken by plotting gain and phase angle as functions of frequency. Ordinarily a logarithmic frequency scale is employed, in order to encompass the very wide frequency bands used. A typical plot for an *RC*-coupled amplifier is shown in Fig. 5-2, distortion being indicated at both low- and high-frequency extremes.

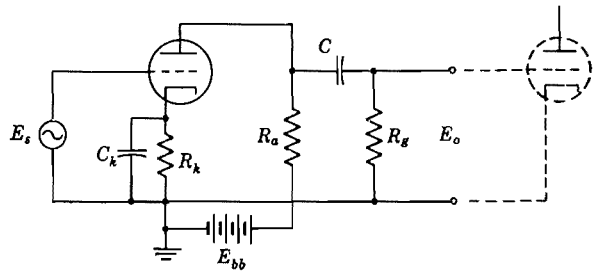



FIG. 5-1. *RC*-coupled amplifier.

The curve form indicates that the study of the *RC* amplifier may be based on a performance analysis covering three frequency ranges. Referring to the circuit of Fig. 5-1 and to the curve of Fig. 5-2, the *low-frequency* range may be defined as that in which the series-connected capacitors represent appreciable reactances and must be considered in the analysis, and the *mid-frequency* range is that in which all capacitive reactances are negligible in effect compared with their associated resistors,

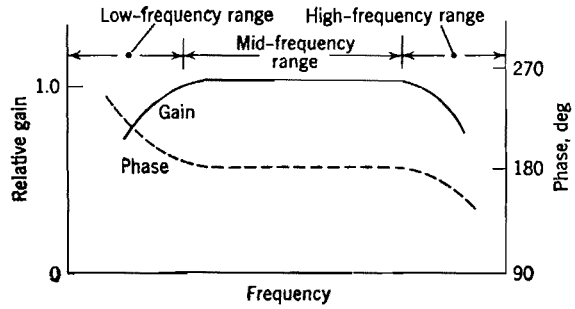



FIG. 5-2. Typical *RC* amplifier gain and phase curves.

the series capacitors representing negligibly small reactances, the parallel elements being very large reactances. The *high-frequency* range is that in which the parallel-connected capacitances become reactances small enough to have an appreciable shunting effect on circuit performance.

Figure 5-3*a* represents the complete equivalent circuit for the triode amplifier of Fig. 5-1. Three capacitive reactances are present, the series blocking capacitor $C$, which is a determining factor in low-frequency performance, and the parallel elements $C_{pk}$ and $C_{in}$, effective at high fre-

quencies since they are ordinarily small-magnitude capacitances. Capacitance $C_{pk}$ is the anode-cathode, or output, capacitance of the tube, and $C_{in}$ was calculated for this circuit in Sec. 4-3. While not shown on the diagram, the stray capacitance of wiring and components in shunt to

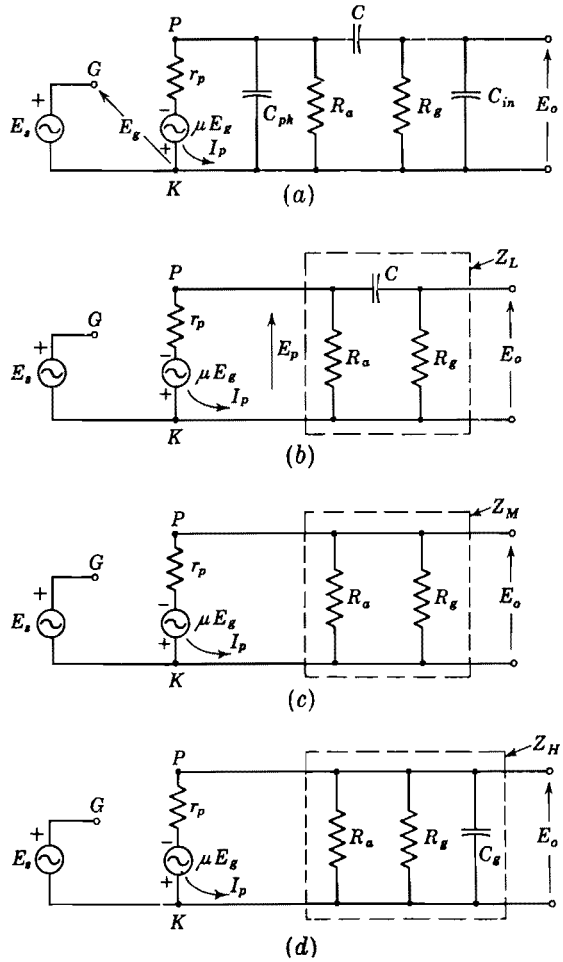



FIG. 5-3. Equivalent circuits for the $RC$ amplifier. (*a*) Exact; (*b*) low-frequency; (*c*) mid-frequency; (*d*) high-frequency.

ground may approximate 5 to 20 $\mu\mu$f and should be included. Thus in Fig. 5-3*d* the capacitance $C_g$ is

$$C_g = C_{pk} + C_{in} + C_{\text{wiring}}$$

For a pentode in the second, or following, stage of the amplifier the value of $C_{in}$ will closely approximate $C_{gk}$.

In the low-frequency region of operation for a given amplifier the parallel capacitance $C_g$ may be neglected as a very high reactance or open circuit, reducing the circuit of Fig. 5-3*a* to the low-frequency equivalent circuit of Fig. 5-3*b*. The voltage rise between cathode and plate is

$$\mathbf{E}_p = -\mathbf{I}_p\mathbf{Z}_L = \frac{-\mu\mathbf{E}_g\mathbf{Z}_L}{r_p + \mathbf{Z}_L} \tag{5-1}$$

where $\mathbf{Z}_L$ is the composite impedance of all a-c current-carrying elements in the load circuit, indicated inside the dashed box in the figure. The value of $\mathbf{Z}_L$, for the low-frequency range, is

$$\mathbf{Z}_L = \frac{R_a(R_g - j/\omega C)}{R_a + R_g - j/\omega C}$$

It can be seen that $\mathbf{E}_o$ is related to $\mathbf{E}_p$ as

$$\mathbf{E}_o = \frac{R_g}{R_g - j/\omega C}\,\mathbf{E}_p$$

so that the *low-frequency gain* is

$$\mathbf{A}_L = \frac{\mathbf{E}_o}{\mathbf{E}_s} = \frac{\mathbf{E}_o}{\mathbf{E}_g} = \frac{-\mu\mathbf{Z}_L}{r_p + \mathbf{Z}_L}\,\frac{R_g}{R_g - j/\omega C} \tag{5-2}$$

As the frequency of operation is raised, the reactance of $C$ drops until $1/\omega C \ll R_g$, at which frequency the *mid-frequency gain* becomes

$$\mathbf{A}_M = \frac{-\mu\mathbf{Z}_M}{r_p + \mathbf{Z}_M} \tag{5-3}$$

where the *mid-frequency* impedance $\mathbf{Z}_M$ is defined as

$$\mathbf{Z}_M = \frac{R_aR_g}{R_a + R_g}$$

The circuit of Fig. 5-3*c* is thus applicable in this frequency region, and the value of gain is independent of the frequency, since no reactances appear in the circuit. The phase angle is also constant and equal to 180°.

At some higher frequency the reactance of $C_g$ will no longer be large with respect to its paralleling resistors. At this frequency the component $R_{in}$ will usually still be large with respect to $R_a$ and may continue to be neglected, although at some higher frequency the value of $R_{in}$ must also be considered.

The equivalent circuit then becomes that of Fig. 5-3*d*, and the *high-frequency gain* is

$$\mathbf{A}_H = \frac{-\mu\mathbf{Z}_H}{r_p + \mathbf{Z}_H} \tag{5-4}$$

where the impedance of the *high-frequency* output circuit is defined as

$$\mathbf{Z}_H = \frac{R_a R_g}{R_a + R_g + j\omega C_g R_a R_g}$$

These expressions may be easily adapted for use with a pentode. Dividing the mid-frequency gain by $r_p$ gives

$$\mathbf{A}_M = \frac{-g_m \mathbf{Z}_M}{1 + \mathbf{Z}_M / r_p}$$

If $r_p \gg |Z_M|$, which is usually true for a pentode, then the gain becomes

$$\mathbf{A}_M = -g_m \mathbf{Z}_M \tag{5-5}$$

The expressions for the other frequency regions can likewise be modified for use with a high-$r_p$ pentode as

$$\mathbf{A}_L = -g_m \mathbf{Z}_L \frac{R_g}{R_g - j/\omega C} \tag{5-6}$$

$$\mathbf{A}_H = -g_m \mathbf{Z}_H \tag{5-7}$$

It should be noted that $C_{in}$ for a pentode is usually much less than for a triode since $C_{gp}$ is very small. Thus the point of demarcation between mid-frequency and high-frequency regions will be at a higher frequency for pentodes, and they will provide uniform, or "flat," amplification over a broader band than will triodes.

Throughout this section reference has been made to certain reactances being negligible in magnitude compared with other reactances or resistances. In practice it will usually be satisfactory to neglect one component with respect to another if a 10:1 ratio exists between their magnitudes.

The grid leak $R_g$ is ordinarily made large with respect to $R_a$ and so has little effect in the circuit. Its value may range from 100,000 ohms to 1 megohm. The coupling capacitor is ordinarily in the range of 0.0001 to 0.25 $\mu$f, depending on the desired low-frequency-range response. The capacity $C_g$ may run from 40 to 150 $\mu\mu$f for some triode amplifiers, being of the order of 5 to 20 $\mu\mu$f for well-arranged pentode stages. The load resistor is chosen as a compromise based on desired gain and frequency bandwidth to be covered, and this subject will be further discussed in later sections.

**5-2. Further Study of the Frequency and Phase Response of the *RC* Amplifier.** Analysis of an amplifier usually is presented in the form of gain-frequency and phase-frequency curves, such as Fig. 5-4. It is customary to plot the frequency axis on a logarithmic scale, in order to encompass more easily the required broad frequency ranges. The gain axis may be plotted on either a linear or a logarithmic scale, although

some advantages accrue to the use of the decibel scale for gain, as will be shown.

Referring to Fig. 5-4, the frequencies $f_1$ and $f_2$ represent frequencies at which the voltage gain has fallen to $1/\sqrt{2}$, or 0.707 of its mid-frequency value, or the output power is one-half of the mid-frequency value. On a decibel scale these frequencies would represent points at which the output is down 3.01 db from the mid-frequency value. The frequencies $f_1$ and $f_2$ are useful but arbitrary bounding values for the response, and it will be shown that they are simply related to the choice of circuit parameters for the amplifier. The frequency space between $f_1$ and $f_2$ is frequently referred to as the *bandwidth* of the amplifier.

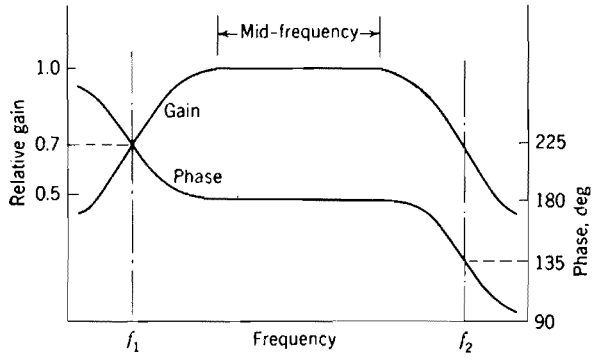



FIG. 5-4. Typical gain-frequency and phase-shift-frequency plot.

In design it is helpful to place data in generalized form so that they may be applied to any unit of a type. In the case of amplifiers the low- and high-frequency performance may be generalized to fit any *RC*-coupled amplifier by using the mid-frequency performance as a norm, or standard of comparison, since the mid-frequency response is constant in magnitude and phase over the mid-frequency band.

To demonstrate for the *RC*-coupled grounded-cathode amplifier, Eq. (5-3) gave the mid-frequency gain as

$$\mathbf{A}_M = \frac{-\mu \mathbf{Z}_M}{r_p + \mathbf{Z}_M} = -\frac{\mu R_a R_g}{r_p R_a + r_p R_g + R_a R_g}$$

The low-frequency gain from Eq. (5-2) can be reduced to

$$\mathbf{A}_L = -\frac{\mu R_a R_g}{r_p R_a + r_p R_g + R_a R_g - j(r_p + R_a)/\omega C}$$

Using $\mathbf{A}_M$ as the base of comparison, the variation of $\mathbf{A}_L$ with respect to the $\mathbf{A}_M$ value can be found by taking the ratio

$$\frac{\mathbf{A}_L}{\mathbf{A}_M} = \frac{1}{1 - \dfrac{j}{\omega C}\left(\dfrac{r_p + R_a}{r_p R_a + r_p R_g + R_a R_g}\right)} \tag{5-8}$$

The second term in the denominator is an inverse function of frequency and at some frequency will have a magnitude of unity, and the magnitude of the gain ratio then will become $|A_L/A_M| = 1/\sqrt{2}$. This gives physical significance to the frequency $f_1$ of Fig. 5-4. Then by definition

$$\begin{aligned} f_1 &= \frac{1}{2\pi C}\left(\frac{r_p + R_a}{r_p R_a + r_p R_g + R_a R_g}\right) \\ &= \frac{1}{2\pi C}\left[\frac{1}{R_g + r_p R_a/(r_p + R_a)}\right] \\ &= \frac{1}{2\pi C}\left(\frac{1}{R_g + R_x}\right) \end{aligned} \tag{5-9}$$

where $R_x$ is defined as the equivalent value for $r_p$ and $R_a$ in parallel.

It is apparent that frequency $f_1$ is inversely proportional to the time constant or $RC$ product of the series circuit of Fig. 5-5. This time constant is largely due to $C$ and $R_g$, since $R_x$ is usually much smaller than $R_g$.

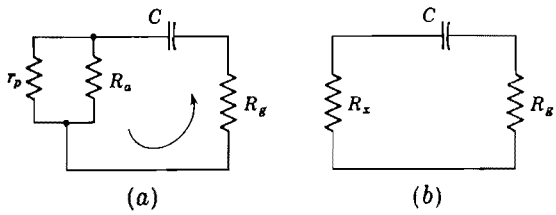



FIG. 5-5. Series circuit whose time constant determines the low-frequency performance.

If this time constant can be increased, as by increasing $C$ or $R_g$, then $f_1$ becomes a lower frequency and the amplifier range is extended. Because of the tiny grid current which may flow even with $e_c$ always negative, the value of $R_g$ does have a definite upper limit for a given tube type, if bias conditions are not to be upset. This limit is of the order of a megohm, and thus $C$ is the major variable available to the designer.

Inserting $f_1$ in Eq. (5-8) gives

$$\frac{\mathbf{A}_L}{\mathbf{A}_M} = \frac{1}{1 - jf_1/f} \tag{5-10}$$

wherein

$$\left|\frac{\mathbf{A}_L}{\mathbf{A}_M}\right| = \sqrt{\frac{1}{1 + (f_1/f)^2}} \tag{5-11}$$

and

$$\theta_L = \tan^{-1}\frac{f_1}{f} \tag{5-12}$$

in addition to the inherent 180° shift of the $\mathbf{A}_M$ gain. Choice of frequency $f_1$ thus fixes the low-frequency performance of the amplifier. Frequency $f_1$ is under the control of the designer through selection of coupling-circuit parameters. He may balance the low-frequency response of his circuit against cost, the latter being largely a function of the size of $C$.

A dimensionless plot of Eqs. (5-11) and (5-12) is presented in Fig. 5-6. These curves are actually plotted in terms of the reciprocal $f/f_1$, in

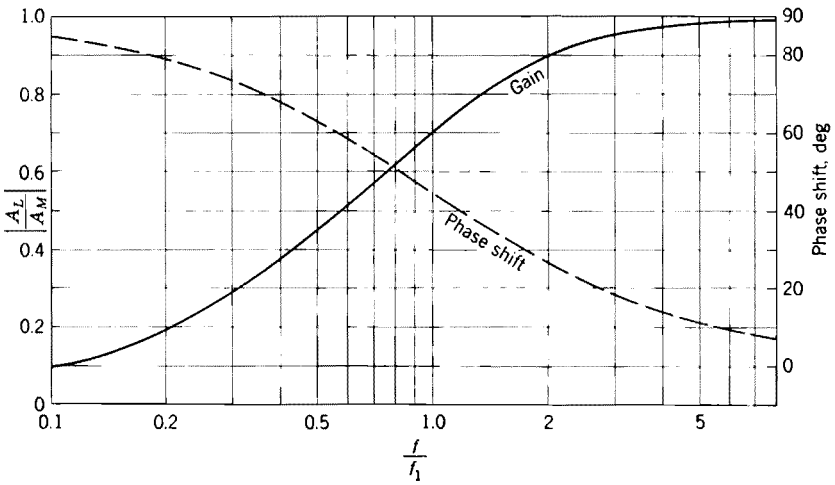



FIG. 5-6. Generalized gain and phase-shift curves for the $RC$ amplifier in the low-frequency range.

order that the ratio will increase with frequency. The gain asymptotically approaches zero, and the phase angle $+90°$, as $f/f_1$ decreases. These curves apply to any $RC$-coupled amplifier.

A similar use of Eq. (5-4) for $\mathbf{A}_H$ gives for the high-frequency-gain ratio

$$\frac{\mathbf{A}_H}{\mathbf{A}_M} = \frac{1}{1 + j\omega C_u \dfrac{1}{1/r_p + 1/R_a + 1/R_g}} \tag{5-13}$$

A frequency $f_2$ may be defined at which the second term in the denominator becomes unity, and the gain is $1/\sqrt{2}$ of $\mathbf{A}_M$ value. This confirms the use of $f_2$ in Fig. 5-4. Then

$$f_2 = \frac{1}{2\pi C_u}\left(\frac{1}{r_p} + \frac{1}{R_a} + \frac{1}{R_g}\right) = \frac{1}{2\pi C_u R_u} \tag{5-14}$$

where $R_y$ is the parallel equivalent resistance as shown in Fig. 5-7*b*. Substitution in Eq. (5-13) gives

$$\frac{\mathbf{A}_H}{\mathbf{A}_M} = \frac{1}{1 + jf/f_2} \tag{5-15}$$

from which

$$\left|\frac{\mathbf{A}_H}{\mathbf{A}_M}\right| = \frac{1}{\sqrt{1 + (f/f_2)^2}} \tag{5-16}$$

and

$$\theta_H = -\tan^{-1}\frac{f}{f_2} \tag{5-17}$$

in addition to the 180° shift of the $\mathbf{A}_M$ gain.

Frequency $f_2$ is determined by the time constant of the shunt-plate-circuit parameters as in Fig. 5-7. Reduction of this time constant by

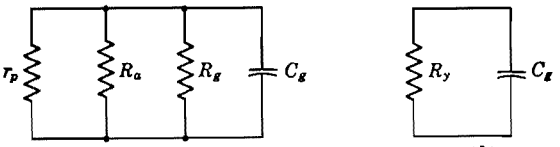



FIG. 5-7. Circuit configuration which limits the high-frequency response.

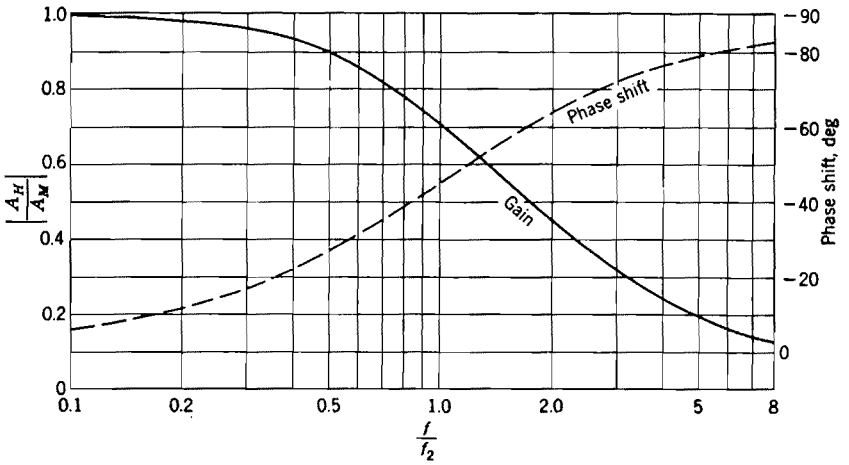



FIG. 5-8. Generalized gain and phase-shift curves for the *RC* amplifier in the high-frequency range.

reduction of $R_a$ or $C_g$ makes $f_2$ larger and extends the frequency response of the amplifier to higher frequencies. Figure 5-8 shows the variation of gain and phase shift with the ratio $f/f_2$. Limiting values are zero gain and a phase shift of −90° at infinite frequency. The response of the amplifier is fixed as soon as the designer selects frequency $f_2$.

**5-3. Steady-state Decibel-gain Plots.** In general, if an amplifier is made up of $n$ amplifier stages in cascade, the over-all gain $\mathbf{A}'$ can be found from

$$\begin{aligned}\mathbf{A}'/\underline{\theta^\circ} &= A_1/\underline{\theta_1^\circ} A_2/\underline{\theta_2^\circ} A_3/\underline{\theta_3^\circ} \cdots A_n/\underline{\theta_n^\circ} \\ &= A_1 A_2 A_3 \cdots A_n/\underline{\theta_1 + \theta_2 + \theta_3 + \cdots + \theta_n} \end{aligned} \quad (5\text{-}18)$$

The gain is the product of the individual gains, and the phase angle is the sum of the individual phase angles.

If the gain magnitudes are expressed in decibels,

$$\begin{aligned}20 \log A' &= 20(\log A_1 + \log A_2 + \log A_3 + \cdots + \log A_n) \\ \text{Over-all db gain} &= \text{db}_1 + \text{db}_2 + \text{db}_3 + \cdots + \text{db}_n \end{aligned} \quad (5\text{-}19)$$

and the over-all gain in decibels is simply expressed as the sum of the decibel gains of each stage.

If the gain-ratio curves of the preceding section are plotted in terms of decibels, it is possible to determine gain and phase shift of several dissimilar amplifier circuits in cascade by simple addition of their respective decibel gains and phase angles. Also, there is then available a simple method of plotting the respective gain curves, and some information concerning relative phase shift may be gleaned from the decibel curve. These advantages make consideration of the decibel-gain plot desirable.

For illustration the decibel value of the $|\mathbf{A}_H/\mathbf{A}_M|$ ratio leads to

$$\begin{aligned}\left|\frac{\mathbf{A}_H}{\mathbf{A}_M}\right| \text{ in db} &= 20 \log \left[\frac{1}{1 + (f/f_2)^2}\right]^{1/2} \\ &= 10 \log \frac{1}{1 + (f/f_2)^2} = -10 \log \left[1 + \left(\frac{f}{f_2}\right)^2\right] \end{aligned} \quad (5\text{-}20)$$

If the ratio $f/f_2$ is small with respect to unity, or the amplifier is operating in or near the mid-frequency range, then

$$\left|\frac{\mathbf{A}_H}{\mathbf{A}_M}\right|_{\text{db}} = 10 \log 1 = 0 \text{ db}$$

and this shows that the *reference level* is the mid-frequency gain level, or zero decibels.

If $f/f_2 = 1$, the equation becomes

$$\left|\frac{\mathbf{A}_H}{\mathbf{A}_M}\right|_{\text{db}} = 10 \log \frac{1}{2} = -10 \log 2 = -3.01 \text{ db}$$

and at this point on the curve the gain is *down* (minus) 3 db from the mid-frequency reference level. Other values of $f/f_2$ may be chosen and the gains of Table 5 computed. These values are plotted as the curve of Fig. 5-9.

When operation is at a high frequency where $(f/f_2)^2 \gg 1$, then

$$\left|\frac{\mathbf{A}_H}{\mathbf{A}_M}\right|_{\text{db}} = 10 \log \frac{1}{(f/f_2)^2} = -20 \log \frac{f}{f_2} \tag{5-21}$$

Equation (5-21) is the asymptote for the high-frequency gain of Table 5, at high $f/f_2$ values. It can be seen that Eq. (5-21) is that of a straight line on a semilog or decibel-frequency graph, and this line will pass through the point $f/f_2 = 1$, 0 db in Fig. 5-9, thus being the sloping dashed line of the figure. The frequency at which $f/f_2 = 1$ is occasionally called the *corner frequency*, by reason of the intersection of the 0-db line

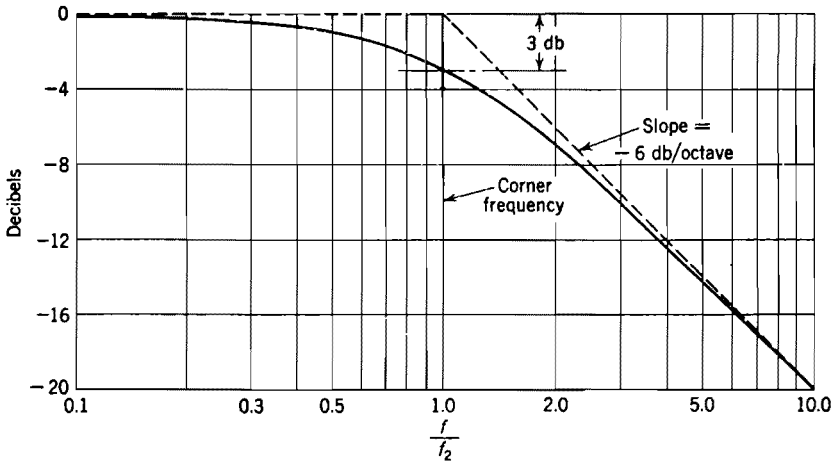



FIG. 5-9. Decibel gain vs. $f/f_2$ for the high-frequency response.

and the high-frequency asymptote in a corner at that frequency. Inspection of Eq. (5-21) also discloses that the asymptote has a slope of $-20$ db per frequency decade, which is equivalent to a slope of $-6$ db per octave of frequency.

TABLE 5. HIGH-FREQUENCY GAIN IN DECIBELS

| $f/f_2$ | *Db gain* | $f/f_2$ | *Db gain* |
|---|---|---|---|
| 0.1 | −0.04 | 1.5 | − 5.12 |
| 0.2 | −0.18 | 2.0 | − 7.0 |
| 0.3 | −0.38 | 3.0 | −10.0 |
| 0.5 | −0.96 | 5.0 | −14.1 |
| 0.7 | −1.73 | 8.0 | −18.1 |
| 1.0 | −3.01 | 10.0 | −20.0 |

Summarizing the properties of the decibel-frequency plot,

1. At $f/f_2 < 1$ the gain approaches the reference or zero-decibel line as an asymptote.

2. At $f/f_2 = 1$ the gain is $-3$ db.

3. At $f/f_2 > 1$ the gain falls, approaching a slope of $-6$ db per octave as an asymptote.

The plotting of a gain curve utilizing the above information is relatively simple, since the sloping asymptote passes through the $f/f_2 = 1$,

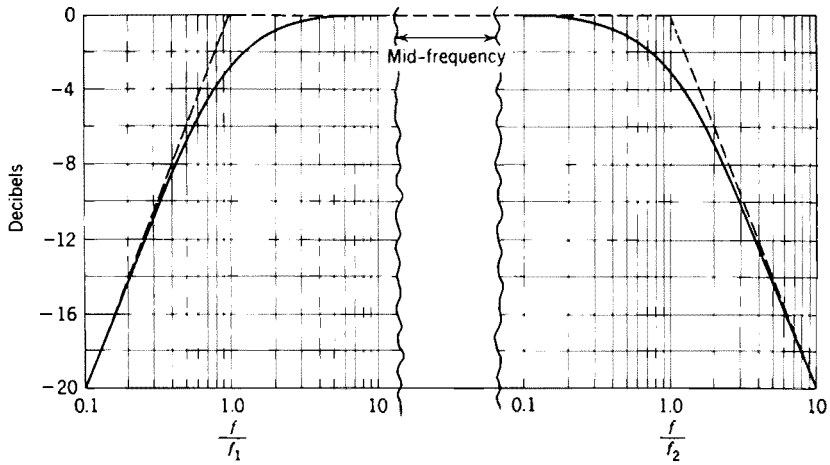



FIG. 5-10. Generalized gain curves for the $RC$ amplifier expressed in decibels.

0 point and falls off with a slope of $-20$ db per decade or $-6$ db per octave. The curve is quickly approximated by a freehand sketch.

TABLE 6. LOW-FREQUENCY GAIN IN DECIBELS

| $f/f_1$ | *Db gain* | $f/f_1$ | *Db gain* |
|---|---|---|---|
| 0.1 | −20.0 | 1.5 | −1.61 |
| 0.2 | −14.1 | 2.0 | −0.96 |
| 0.3 | −10.7 | 3.0 | −0.42 |
| 0.5 | − 7.0 | 5.0 | −0.18 |
| 0.7 | − 4.8 | 8.0 | −0.07 |
| 1.0 | − 3.0 | 10.0 | −0.04 |

Equation (5-11) for $|\mathbf{A}_L/\mathbf{A}_M|$ is of the same form as the high-frequency-gain ratio, and in decibels it is

$$\left|\frac{\mathbf{A}_L}{\mathbf{A}_M}\right|_{\text{db}} = 20 \log \left[\frac{1}{1 + (f_1/f)^2}\right]^{1/2} = -10 \log \left[1 + \left(\frac{f_1}{f}\right)^2\right] \quad (5\text{-}22)$$

Similar reasoning defines a corner frequency at $f_1/f = 1$, and a low-frequency asymptote which falls at $-6$ db per octave, or rises at $+6$ db per octave when considered in terms of an increasing $f/f_1$, as in Table 6. The method is further illustrated by the combined curves of Fig. 5-10 for the $RC$-coupled amplifier.

If a second amplifier is operated in cascade with the first, the over-all gain is the sum of the ordinates at each frequency. The curve for a two-stage amplifier of identical stages will be down −6 db at the stage-corner frequency, and the asymptote for the two stages will fall off at −12 db per octave. Figure 5-11 illustrates the changed response curves produced by cascading several identical stages.

In the range where $(f/f_2)^2$ is large, or where the gain curve approaches the asymptotic slope, the phase angle of a single-stage $RC$-coupled amplifier will approach 90° (−90° for the low-frequency response).

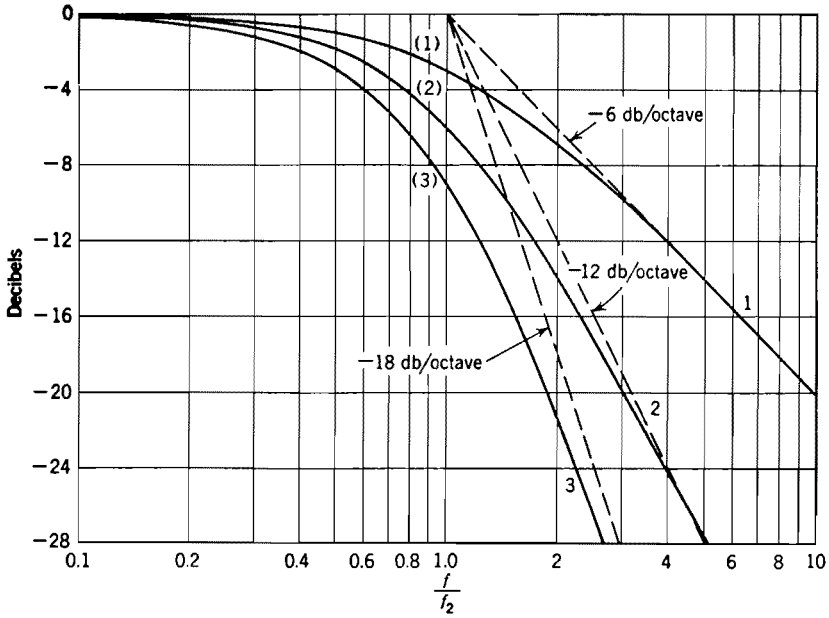



FIG. 5-11. Response for $n$-identical-stage $RC$ amplifier: one, two, three stages, respectively.

For a two-stage similar amplifier the asymptote has a slope of −12 db per octave, and when the gain curve approaches this slope, the phase angle will be nearing 270° (or −270° for the low-frequency case). Thus some knowledge of phase angle can be obtained from the steady-state frequency-response curve.

**5-4. $RC$ Circuits for Alteration of Bandwidth.** By use of simple $RC$ networks it is possible to reduce the response of an amplifier over either the low-frequency or the high-frequency end of an amplifier's range. Or, by sacrifice of gain in the wanted frequency range, the frequency band may be extended somewhat at either end. This is accomplished by matching the falling-gain characteristics of the amplifier with an equal rising characteristic from an $RC$ circuit.

The circuits of Fig. 5-12 are two of a possible wide variety. These may be analyzed by the corner-frequency slope of the asymptote methods of Sec. 5-3. The circuit of Fig. 5-12*a* cuts the high-frequency response,

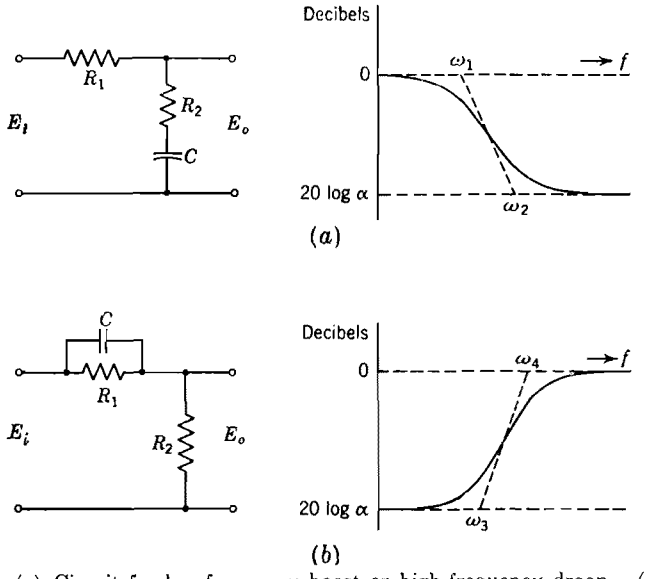



FIG. 5-12. (*a*) Circuit for low-frequency boost or high-frequency droop. (*b*) High-frequency boost or low-frequency droop.

thus appearing to boost the low-frequency response, and the voltage ratio is

$$\frac{\mathbf{E}_o}{\mathbf{E}_i} = \frac{1 + j\omega R_2 C}{1 + j\omega C(R_1 + R_2)} \tag{5-23}$$

For the gain limits at one extreme of frequency take $\omega$ *very small* and

$$\frac{\mathbf{E}_o}{\mathbf{E}_i} = 1$$

For $\omega$ *very large*

$$\frac{\mathbf{E}_o}{\mathbf{E}_i} = \frac{R_2}{R_1 + R_2} = \alpha$$

Taking twenty times the logarithm of the ratio gives 0 db as the low-frequency reference level and 20 log $\alpha$ as the high-frequency asymptote. Thus the low-frequency loss in the circuit (or the high-frequency boost) is a function of $\alpha$, the resistance ratio.

Rewriting Eq. (5-23) gives

$$\left|\frac{\mathbf{E}_o}{\mathbf{E}_i}\right| = \left[\frac{1 + (\omega/\omega_2)^2}{1 + (\omega/\omega_1)^2}\right]^{1/2} \tag{5-24}$$

where
$$\omega_1 = \frac{1}{C(R_1 + R_2)}$$
$$\omega_2 = \frac{1}{R_2 C}$$

Converting to decibels, the output-to-input ratio is

$$\left|\frac{\mathbf{E}_o}{\mathbf{E}_i}\right| \text{ db} = 10 \log \left[1 + \left(\frac{\omega}{\omega_2}\right)^2\right] - 10 \log \left[1 + \left(\frac{\omega}{\omega_1}\right)^2\right] \quad (5\text{-}25)$$

and it is to be noted that these terms are similar to those in Eq. (5-20). Thus a corner frequency, at which $|\mathbf{E}_o/\mathbf{E}_i|$ is down 3 db from the low-frequency level, occurs at $\omega = \omega_1$. This same term also provides an

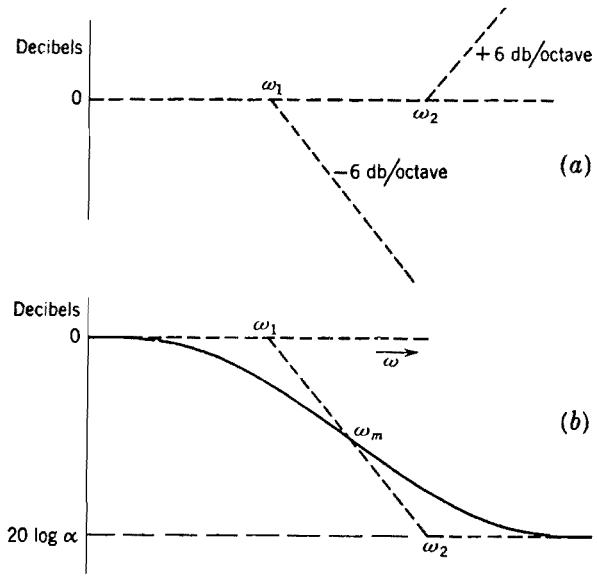



FIG. 5-13. (*a*) Corner frequencies and asymptote construction; (*b*) use of asymptotes to obtain response of Fig. 5-12*a*.

asymptote to the response curve which falls at 6 db per octave from the corner at $\omega = \omega_1$. This is shown dashed in Fig. 5-13*a*. A second corner frequency is contributed by the first term on the right at $\omega = \omega_2$, and because of the positive sign this asymptote *rises* at the rate of 6 db per octave.

The over-all response $|\mathbf{E}_o/\mathbf{E}_i|$ is then the sum of these two responses, as indicated in Fig. 5-13*b*. Beyond $\omega = \omega_2$ the two oppositely directed slopes annul each other, and the response asymptote becomes flat at a value of 20 log $\alpha$. Thus the high-frequency response is dropped 20 log $\alpha$ db with respect to the low-frequency end. Using these asymp-

totic curves, it is then readily possible to draw the *actual response curve* as the heavy line of the figure, using the data of Table 4 and Table 5 for intermediate points.

The circuit is a simple means of providing apparent low-frequency boost, at a sacrifice in high-frequency gain. In the field of servomechanisms it is known as a *lag* circuit, since the output phase angle lags the input.

The circuit of Fig. 5-12*b* operates in reverse fashion, reducing the low frequencies with respect to the high-frequency range. The response is

$$\frac{\mathbf{E}_o}{\mathbf{E}_i} = \frac{R_2}{R_1 + R_2} \frac{1 + j\omega CR_1}{1 + j\omega CR_1R_2/(R_1 + R_2)} \tag{5-26}$$

The gain asymptotes for the frequency extremes are, for $\omega$ *very small*,

$$\frac{\mathbf{E}_o}{\mathbf{E}_i} = \frac{R_2}{R_1 + R_2} = \alpha$$

and, for $\omega$ *very large*,

$$\frac{\mathbf{E}_o}{\mathbf{E}_i} = 1$$

Taking the logarithm of the latter expression to obtain decibels shows that the zero reference level occurs at high frequency. The decibel level at low frequency is 20 log $\alpha$, and since $\alpha$ is less than unity, the low-frequency gain asymptote is below zero or reference and the circuit reduces the low frequencies with respect to the high frequencies.

Equation (5-26) may be written as

$$\left|\frac{\mathbf{E}_o}{\mathbf{E}_i}\right| = \frac{R_2}{R_1 + R_2}\left[\frac{1 + (\omega/\omega_3)^2}{1 + (\omega/\omega_4)^2}\right]^{1/2} \tag{5-27}$$

where

$$\omega_3 = \frac{1}{CR_1}$$

$$\omega_4 = \frac{R_1 + R_2}{CR_1R_2} = \frac{\omega_3}{\alpha}$$

Converting to decibel units, the output-to-input ratio is

$$\left|\frac{\mathbf{E}_o}{\mathbf{E}_i}\right|_{\text{db}} = 20 \log \alpha + 10 \log \left[1 + \left(\frac{\omega}{\omega_3}\right)^2\right] - 10 \log \left[1 + \left(\frac{\omega}{\omega_4}\right)^2\right] \tag{5-28}$$

The similarity of this expression to that of Eq. (5-25) is apparent, and the asymptote sketch of Fig. 5-14*a* follows by similar reasoning.

The two asymptote slopes may be combined, the positive slope annulling the negative slope above the corner frequency at $\omega_4$, and the resultant is shown in Fig. 5-14*b*. The actual response curve is also sketched, passing through the $\pm 3$-db points at $\omega_3$ and $\omega_4$. The low frequencies are

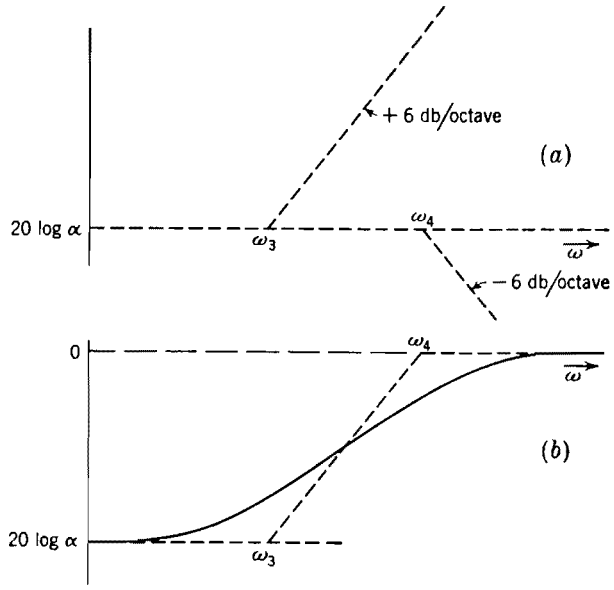



FIG. 5-14. (*a*) Corner frequencies and asymptotes for Fig. 5-12*b*. (*b*) Response of circuit.

dropped with respect to the high frequencies by an amount of 20 log $\alpha$ db. In the field of servomechanisms the circuit is known as a *lead* circuit, since the output voltage leads the input in phase.

These circuits introduce a phase shift which starts at zero, rises to a maximum lead or lag, and once more falls to zero as the frequency is progressively raised. The maximum shift occurs at

$$\omega_m = \sqrt{\omega_1\omega_2}$$

or $$\omega_m = \sqrt{\omega_3\omega_4}$$

with values plotted as functions of $\alpha$ in Fig. 5-15.

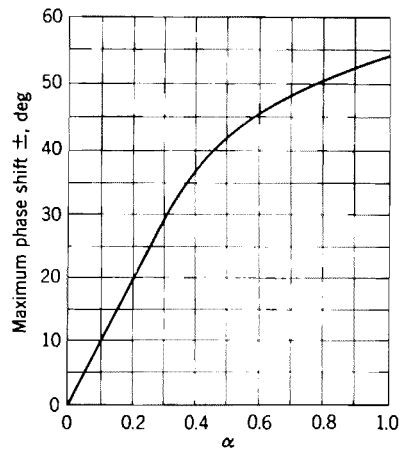



FIG. 5-15. Maximum phase shift ± for circuits of Fig. 5-12.

**5-5. Factors Affecting Amplifier Bandwidth.** The amplification of pulses and of rectangular, triangular, and other waveforms, with frequencies of many thousands of cycles, requires amplifiers with bandwidths considerably greater than would be required if the waveforms

were sinusoidal. The design of amplifiers of considerable bandwidth is of importance to those otherwise dealing with relatively low-frequency phenomena, and certain fundamentals should be understood.

Since the low-frequency range of an amplifier depends on $f_1$ and the coupling capacitor $C$, there may be an economic limit but there is no serious technical difficulty in extending $f_1$ to as low a frequency as desired. At the high-frequency end the shunt capacity $C_g = C_{pk} + C_{gk} + C_{\text{wiring}}$ is the limiting reactance, and it is not possible to reduce this below the inherent limit fixed by the sum of $C_{pk}$ and $C_{gk}$ of the associated tubes (pentodes assumed). Because of this, $RC$ amplifier bandwidth is usually discussed and measured in terms of the high-frequency limit alone.

Since pentodes with high $g_m$ provide high gain and have low input capacities approximating $C_{gk}$ only, pentodes are almost exclusively used when it is desired to extend frequency bands much above the 10,000- to 20,000-cycle audio-frequency range. Pentodes being assumed, it is possible to assume that $r_p \gg R_a$, where $R_a$ is the plate load. An assumption of $R_g \gg R_a$ is permissible because $R_a$ must usually be made small to raise $f_2$, whereas $R_g$ must be large to lower $f_1$.

*If these assumptions are employed*, the mid-frequency gain of Eq. (5-5) becomes

$$\mathbf{A}_M = -g_m R_a \tag{5-29}$$

and the high-frequency gain of Eq. (5-4) appears as

$$\mathbf{A}_H = \frac{-g_m R_a}{1 + j\omega C_g R_a} \tag{5-30}$$

The gain ratio for a pentode is then

$$\left| \frac{\mathbf{A}_H}{\mathbf{A}_M} \right| = \left[ \frac{1}{1 + (f/f_2)^2} \right]^{1/2} \tag{5-31}$$

where

$$f_2 = \frac{1}{2\pi C_g R_a} = \text{bandwidth in cycles} \tag{5-32}$$

Since $C_g$ has a minimum value which is fixed when the tube type is selected, the only variable present in the circuit by which bandwidth may be increased is reduction of $R_a$, and this also reduces gain.

Since gain is proportional to $R_a$ and bandwidth inversely proportional to the same quantity, their product is independent of $R_a$. If such a *gain-bandwidth* product be taken, there is obtained a quantity used as a *figure of merit* $M$, where

$$M = \frac{g_m R_a}{2\pi C_g R_a} = \frac{g_m}{2\pi C_g} \tag{5-33}$$

and *this is a constant for a given tube type.* If $C_g$ is given in micromicrofarads and $g_m$ in micromhos, the product becomes gain times bandwidth

in megacycles. Since the product is a constant, it is not possible to increase bandwidth without a commensurate sacrifice in gain, and vice versa. This poses somewhat of a fundamental limit on amplifiers of this type and points up the need for pentodes of high $g_m$ value.

Table 7 presents values for the bandwidth figure of merit for a few common tube types. For this purpose $C_g = C_{pk} + C_{gk}$, since this represents the irreducible minimum value for $C_g$ for pentodes.

TABLE 7. GAIN-BANDWIDTH FIGURE OF MERIT FOR A FEW PENTODES

| Tube type | $C_{gk}$, $\mu\mu$f | $C_{pk}$, $\mu\mu$f | $g_m$, $\mu$mhos | Figure of merit |
|---|---|---|---|---|
| 6AG5 | 6.5 | 1.8 | 5,000 | 98 |
| 6AK5 | 4.0 | 2.8 | 5,100 | 123 |
| 6BC5 | 6.6 | 2.6 | 5,700 | 98 |
| 6AH6 | 10 | 2.0 | 9,000 | 120 |
| 6AS6 | 4.0 | 3.0 | 3,200 | 73 |
| 6AU6 | 5.5 | 5.0 | 5,200 | 80 |
| 6AC7 | 11 | 5.0 | 9,000 | 90 |
| 6AG7 | 13 | 7.5 | 11,000 | 85 |
| 6CB6 | 6.3 | 1.9 | 6,200 | 120 |

If an amplifier is made up of $n$ cascaded $RC$-coupled stages, not necessarily identical, the over-all gain ratio, as indicated by the primes, is

$$\left|\frac{\mathbf{A}'_H}{\mathbf{A}'_M}\right| = \left[\frac{1}{1 + (f/f_a)^2}\right]^{1/2}\left[\frac{1}{1 + (f/f_b)^2}\right]^{1/2} \cdots \tag{5-34}$$

where the high-frequency expression is chosen only as an example. The over-all gain of an amplifier, expressed as a ratio to $\mathbf{A}_M$, is then given by a polynomial in $f/f_x$ ratios, where $f_x$ is defined in terms of the parameters of the individual stages.

If the amplifier stages are identical and $f_x = f_2$ for all, then

$$\left|\frac{\mathbf{A}'_H}{\mathbf{A}'_M}\right| = \left[\frac{1}{1 + (f/f_2)^2}\right]^{n/2} \tag{5-35}$$

It is possible to determine an important factor relating to bandwidth, which appears in Fig. 5-11. This figure shows that as $RC$ amplifiers are cascaded the bandwidth decreases, as defined by the half power or $-3$-db points on the response curve. If an over-all bandwidth is specified, then the individual stages must each have considerably greater bandwidth.

Again arbitrarily picking the high-frequency case, as expressed by Eq. (5-35), frequency $f_s$ may designate the $-3$-db point for the over-all ampli-

fier of $n$ identical stages, where the voltage-gain ratio is $1/\sqrt{2}$. Then

$$\frac{1}{\sqrt{2}} = \left[\frac{1}{1 + (f_s/f_2)^2}\right]^{n/2}$$

from which

$$\frac{f_s}{f_2} = \sqrt{2^{1/n} - 1} \tag{5-36}$$

is an expression for the half-power frequency point for a cascaded $n$-identical-stage amplifier. Table 8 gives calculated values for this expression which show that the bandwidth of each stage must be considerably increased, and the gain per stage correspondingly decreased, if an over-all wide band is to be obtained.

TABLE 8. RELATIVE VALUE OF HALF-POWER OR −3-DB POINTS FOR $n$ IDENTICAL CASCADED STAGES

| *Number of stages* $n$ | $f_s/f_2$ | *Number of stages* $n$ | $f_s/f_2$ |
|---|---|---|---|
| 1 | 1.00 | 5 | 0.39 |
| 2 | 0.64 | 6 | 0.35 |
| 3 | 0.51 | 7 | 0.32 |
| 4 | 0.44 | 8 | 0.30 |

The determination of number of stages, gain per stage, and stage half-power frequency is somewhat arbitrary. It will be illustrated by an example.

**Example.** A multistage amplifier is to be designed with over-all mid-frequency gain of 6000 (75 db) and with the gain only 5 per cent below mid-frequency value at 100,000 cycles.

Assume that two stages of pentodes will be sufficient, of $\sqrt{6000} = 78$ voltage gain each ($7\frac{5}{2} = 37.5$ db gain each). Then, from Eq. (5-35), it is possible to obtain the stage $f_2$ frequencies, from knowledge that the over-all gain is to be 0.95 of mid-range value at 100,000 cycles. Thus

$$\left|\frac{\mathbf{A}'_H}{\mathbf{A}'_M}\right| = 0.95 = \left[\frac{1}{1 + (f/f_2)^2}\right]^{3\frac{1}{2}}$$

$$\frac{f}{f_2} = 0.223$$

and since $f = 100{,}000$, then $f_2 = 450{,}000$ cycles for the half-power frequency for each stage. The choice of tube is arbitrary, but a common pentode, the 6AU6, may be tried. Using Table 7, $C_{gk} + C_{pk} = 10.5$ $\mu\mu$f, and allowing 10 $\mu\mu$f for stray capacity gives $C_g = 20.5$ $\mu\mu$f. Then since

$$f_2 = \frac{1}{2\pi C_g R_a}$$

$$R_a = \frac{1}{2\pi \times 20.5 \times 10^{-12} \times 450{,}000} = 17{,}200 \text{ ohms}$$

The gain per stage at mid-frequency then is

$$\mathbf{A}_M = -g_m R_a = -5200 \times 10^{-6} \times 17{,}200 = 89$$
$$\text{Over-all gain} = 89^2 = 7900$$
$$= 78 \text{ db}$$

The excess gain may be satisfactory, or $R_a$ may be reduced slightly to reduce the gain, thereby slightly improving the bandwidth.

**5-6. Low-frequency Compensation of *RC* Amplifiers.** It has been stated that it is technically possible to extend the low-frequency response of an *RC*-coupled amplifier to as low a frequency as required. This can, of course, be done by the direct-coupling techniques of Chap. 8, but employing a-c coupling methods may call for a coupling capacitor that is costly and physically too large.

The circuit of Fig. 5-16 provides a plate-load impedance that rises with decrease in frequency and can compensate for the reduction in gain caused by capacitor $C$. Assuming the use of a pentode tube, then $R_g$ and $r_p$ are large with respect to $R_a$, and the shunting effect of $C$ and $R_g$ on the load magnitude is negligible. However, $C$ and $R_g$ do introduce a voltage drop and a phase shift.

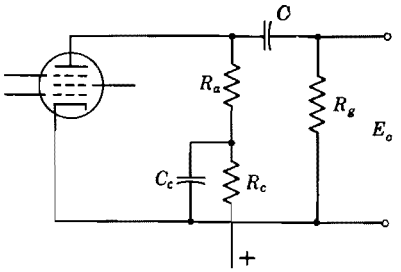



FIG. 5-16. Compensation network for improved low-frequency response.

In essence, the anode load is composed of $R_a$ and $C_c$ in series and will have a rising impedance with decrease in frequency. Resistor $R_c$ is merely present to furnish a d-c path for the plate current and should be large compared with the reactance of $C_c$. Since this cannot be true at some extremely low frequency, resistor $R_c$ causes a limit to be placed on the effectiveness of compensation at that extremely low frequency.

By use of the pentode equivalent circuit

$$\mathbf{A}_L = -g_m \left( R_a - \frac{jR_c/\omega C_c}{R_c - j/\omega C_c} \right) \frac{R_g}{R_g - j/\omega C}$$

Resistor $R_c \gg 1/\omega C$; so

$$\mathbf{A}_L = -g_m \left( \frac{\omega C_c R_a - j}{C_c} \right) \left( \frac{CR_g}{\omega C R_g - j} \right) \tag{5-37}$$

Within the assumptions it can be seen that if

$$C_c R_a = C R_g$$

the low-frequency gain will be independent of frequency. This is equivalent to saying that the time constant of the coupling circuit should be equal to the time constant $C_c R_a$.

The phase angle of the gain is

$$\phi = \tan^{-1} \frac{\omega(C_c R_a - C R_g)}{1 + \omega^2 C C_c R_a R_g} \tag{5-38}$$

and again, if

$$C_c R_a = C R_g$$

the phase angle will be zero, which is the desired condition. Proper equality of these two time constants will compensate for the effect of the coupling capacitor.

**5-7. High-frequency Compensation of the *RC* Amplifier.** It is possible considerably to extend the upper limit of frequency range in the

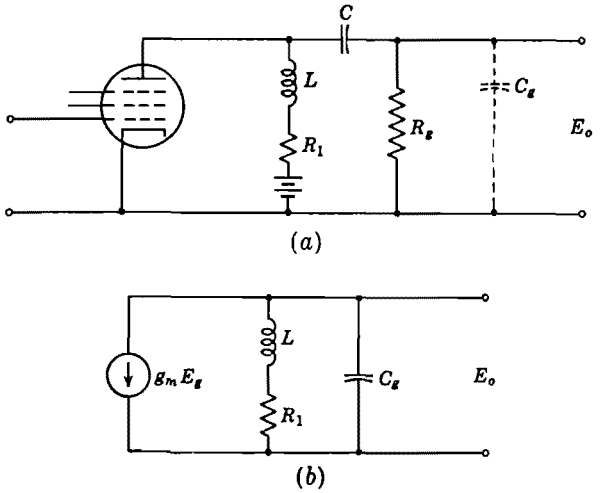



FIG. 5-17. (*a*) Shunt-compensated circuit. (*b*) Equivalent circuit for (*a*).

*RC*-coupled amplifier. A variety of plate-load circuits have been employed for this purpose, but only two of the simplest forms will be discussed here.

The *shunt-compensated* circuit of Fig. 5-17*a* is one form employed to raise the high-frequency response limit of an *RC* amplifier. In this circuit an inductor $L$ is placed in shunt with $C_g$ in the plate load so as to counteract the effect of $C_g$ in reducing the load impedance at high frequencies. Since pentodes are assumed and $R_g$ will be very large compared with the shunting effects of both the plate circuit and $C_g$, the gain may be written from the high-frequency equivalent circuit of Fig. 5-17*b* as

$$\mathbf{A}_H = -g_m \frac{(R_1 + j\omega L)(-j/\omega C_g)}{R_1 + j\omega L - j/\omega C_g}$$

$$= -g_m R_1 \frac{\omega L/R_1 - j}{\omega C_g R_1 + j(\omega^2 L C_g - 1)} \tag{5-39}$$

The magnitude of this gain is

$$|\mathbf{A}_H| = -g_m R_1 \sqrt{\frac{1 + \omega^2 L^2/R_1}{1 + \omega^2(C_g{}^2R_1{}^2 - 2LC_g) + \omega^4 L^2 C_g{}^2}}$$

For the uncompensated amplifier with $L = 0$ the upper half-power frequency is defined by Eq. (5-14) as

$$f_2 = \frac{1}{2\pi C_g R_1}$$

and, using this value, the above equation for $\mathbf{A}_H$ becomes

$$|\mathbf{A}_H| = -g_m R_1 \sqrt{\frac{1 + (L^2/C_g{}^2R_1{}^4)\,(f/f_2)^2}{1 + (1 - 2L/C_gR_1{}^2)(f/f_2)^2 + (L^2/C_g{}^2R_1{}^4)(f/f_2)^4}} \tag{5-40}$$

If the numerator and the denominator of the above were identical functions of frequency, there would be no variation of $\mathbf{A}_H$ with frequency. It is apparent from the form of the expression that this cannot be achieved. However, the coefficients of the second-power terms can be equated, leaving only the $(f/f_2)^4$ term to modify the gain-frequency curve at extreme values of frequency. Equating the coefficients of the second-power terms will indicate a desirable relation between $L$, $C_g$, and $R_1$. Thus

$$\frac{L^2}{C_g{}^2R_1{}^4} = 1 - \frac{2L}{C_gR_1{}^2} \tag{5-41}$$

from which

$$\frac{L}{C_gR_1{}^2} = -1 \pm \sqrt{2}$$

Choosing the positive sign since all quantities on the left are positive, a new parameter $a$ may be defined as

$$a = \frac{L}{C_gR_1{}^2} = 0.414 \tag{5-42}$$

as a condition of maximal flatness for the response curve. Since $C_g$ is fixed by choice of tube and $R_1$ determined by the required mid-frequency gain, then the required value of $L$ may be found from

$$L = 0.414C_gR_1{}^2 \tag{5-43}$$

A series of response curves is plotted in Fig. 5-18 for various values of the parameter $a$. While $a = 0.414$ gives a response with zero rise above the mid-frequency value, it can be seen that $a = 0.45$ provides an improvement in bandwidth if a very slight response rise can be tolerated. If $n$ such stages are used, it should be remembered that the rise will be increased by the factor $n$.

The phase angle of the gain can be obtained from Eq. (5-39) as

$$\phi = -\tan^{-1}\left(\frac{\omega^3 L^2 C_g}{R_1} + \omega C_g R_1 - \frac{\omega L}{R_1}\right) \tag{5-44}$$

and, after inserting values for $f_2$ and $a$,

$$\phi = -\tan^{-1}\frac{f}{f_2}\left[a^2\left(\frac{f}{f_2}\right)^2 - a + 1\right] \tag{5-45}$$

A criterion for desirable phase-response or time-delay properties is that $d\phi/df$ be constant, or that the phase shift be proportional to frequency. Equation (5-45) may be examined in this way by dividing $\phi$ by $f/f_2$ and plotting against $f/f_2$ for various values of $a$ as in Fig. 5-19.

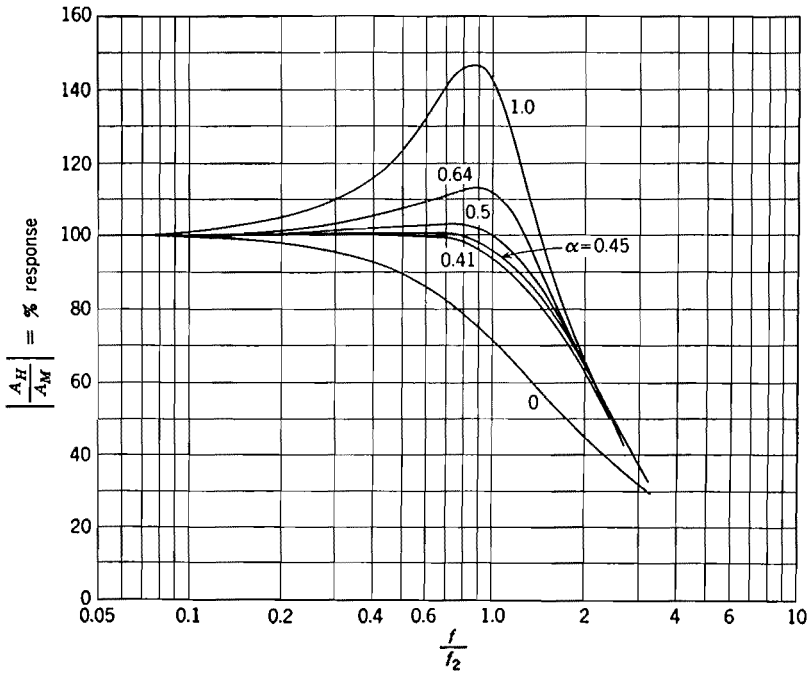



FIG. 5-18. Shunt-compensated amplifier response for various values of $a$.

It can be seen that there is the least change of $\phi$ with respect to $f/f_2$ for $a = 0.32$. Taking the derivative of Eq. (5-45) and setting to zero confirms this value of $a$ as required for constant time delay. It is apparent that a single value of $a$ is not satisfactory for simultaneous flat gain and constant time delay, and a compromise value of $a = 0.37$ is sometimes considered satisfactory.

The circuit of Fig. 5-20$a$ is formed by placing an inductor $L$ in series

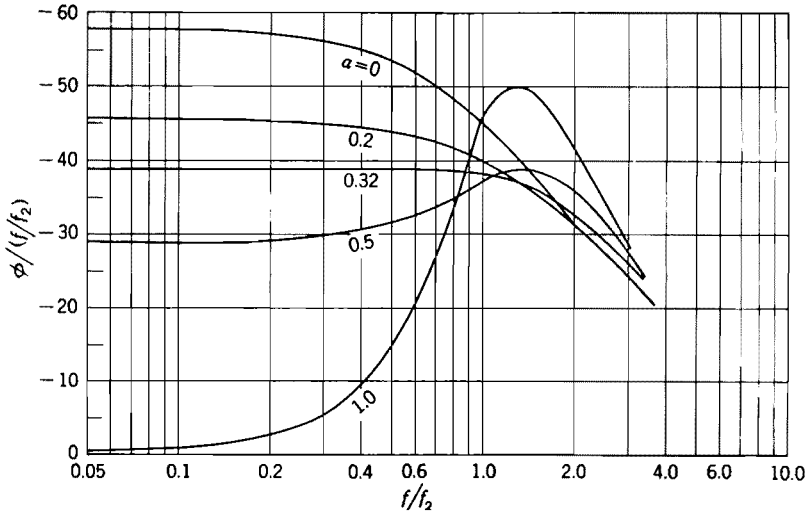



FIG. 5-19. Phase shift for the shunt-compensated amplifier.

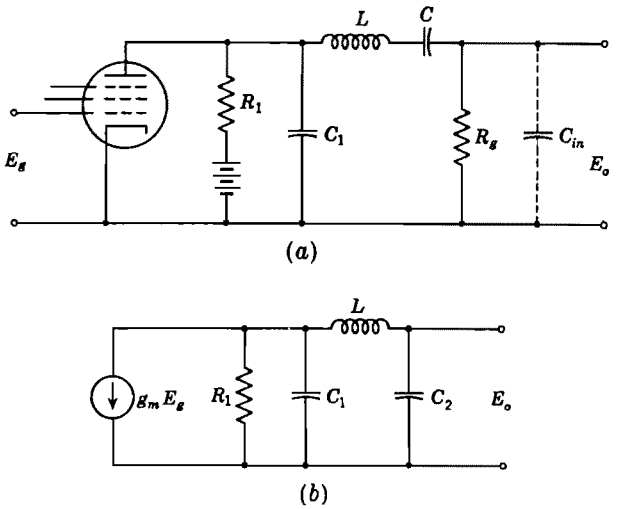



FIG. 5-20. (*a*) Series-compensated *RC* amplifier. (*b*) Equivalent circuit at high frequency.

with the limiting capacity $C_{in}$ of the following amplifier stage, and this is known as a *series-compensated* amplifier. By use of the high-frequency equivalent circuit of Fig. 5-20*b*, in which $C_2$ usually represents $C_{in}$ of the following stage, the gain may be found as

$$\mathbf{A}_H = -g_m \frac{[-jR_1/(\omega C_1 R_1 - j)](j\omega L - j/\omega C_2)}{[-jR_1/(\omega C_1 R_1 - j)] + j(\omega L - 1/\omega C_2)} \left(\frac{-1}{\omega^2 L C_2 - 1}\right)$$

This reduces to

$$\mathbf{A}_H = -g_m R_1 \frac{j}{\omega^3 LC_1C_2R_1 - \omega R_1(C_1 + C_2) - j(\omega^2 LC_2 - 1)} \tag{5-46}$$

After introduction of the definition

$$f_2 = \frac{1}{2\pi R_1(C_1 + C_2)}$$

which is the value of $f_2$ for the uncompensated stage, or for $L = 0$, and is analogous to the $f_2$ value used for the shunt-compensated circuit, the gain magnitude becomes

$$|A_H| = -g_m R_1 \left\{1 + \left[1 - \frac{2LC_2}{R_1^2(C_1 + C_2)^2}\right]\left(\frac{f}{f_2}\right)^2 + \left[\frac{L^2C_2^2}{R_1^4(C_1 + C_2)^4} - \frac{2LC_1C_2}{R_1^2(C_1 + C_2)^3}\right]\left(\frac{f}{f_2}\right)^4 + \left[\frac{L^2C_1^2C_2^2}{R_1^4(C_1 + C_2)^6}\right]\left(\frac{f}{f_2}\right)^6\right\}^{-1/2} \tag{5-47}$$

For the flattest possible gain it is desirable that

$$1 - \frac{2LC_2}{R_1^2(C_1 + C_2)^2} = 0 \tag{5-48}$$

$$\frac{L^2C_2^2}{R_1^4(C_1 + C_2)^4} - \frac{2LC_1C_2}{R_1^2(C_1 + C_2)^3} = 0 \tag{5-49}$$

by equating the appropriate coefficients. The $(f/f_2)^6$ term remains effective and limits the extreme frequency response.

The above two relations simplify to

$$2LC_2 = R_1^2(C_1 + C_2)^2$$
$$LC_2 = 2C_1R_1^2(C_1 + C_2)^2$$

By combining it is found that for maximum flatness of the response curve a parameter $m$ should have the value

$$m = \frac{C_1}{C_1 + C_2} = 0.25 \tag{5-50}$$

This implies that the capacities should have the ratio $C_1/C_2 = 0.33$ for the flattest response. A second design parameter $q$ may also be determined as

$$q = \frac{\omega_2 L}{R_1} = 2\frac{C_1}{C_2} \tag{5-51}$$

and $q$ has the value of 0.667 for the optimum value of $C_1/C_2 = 0.333$. The value of the series inductor can be computed from

$$L = qR_1^2(C_1 + C_2) \tag{5-52}$$

The phase angle for the gain in terms of $f_2$ becomes

$$\phi = -\tan^{-1}\left(\frac{f}{f_2}\right)\left[\frac{(f/f_2)^2LC_1C_2/(C_1+C_2) - R_1^2(C_1+C_2)^2}{R_1^2(C_1+C_2)^2 - (f/f_2)^2LC_2}\right] \quad (5\text{-}53)$$

If $d\phi/df$ is taken and the coefficients equated, it can be shown that relatively constant time delay is possible for parameter values of $m = 0.17$ and $q = 0.48$.

Since the capacitors $C_1$ and $C_2$ desirably represent only $C_{pk}$ and $C_{gk}$, respectively, and since they may not have the ratio of 1:3 required, the best conditions for this circuit cannot always be achieved. Of course,

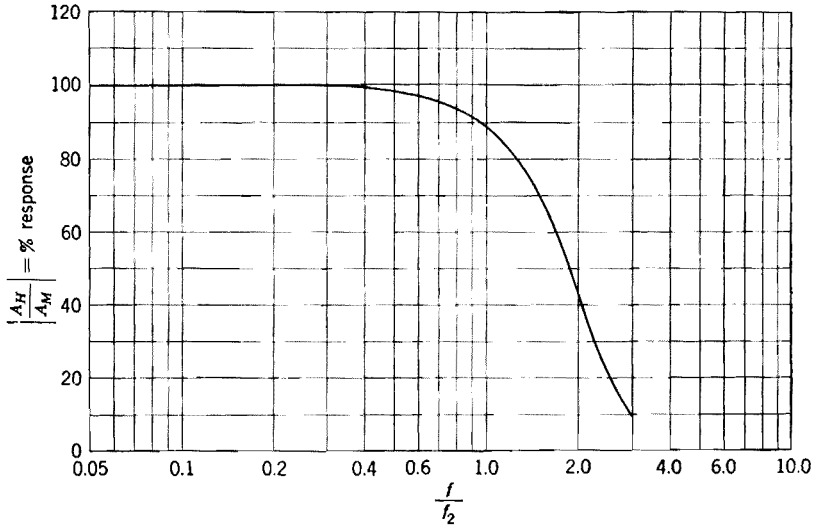



FIG. 5-21. Amplitude response for series-compensated circuit with $m = 0.20$, $q = 0.56$.

it is possible to parallel one or the other to reach the 1:3 ratio, but this decreases $f_2$ and reduces the possible bandwidth. A curve indicating the performance of the series-compensated circuit for compromise values of $m = 0.20$ and $q = 0.56$ is presented in Fig. 5-21. If $C_1$ and $C_2$ are found to present this 1:4 ratio, then the series-compensated circuit performance is approximately equal to that of the shunt-compensated circuit. Both represent a major improvement in bandwidth and phase response over the simple $RC$ amplifier.

Other more complex circuits utilizing additional circuit elements are also employed. As the number of circuit elements is increased, it becomes possible to meet additional boundary or design conditions, so that these more complex circuits may be expected to give improved performance over the two simple forms discussed above. Methods of analysis remain similar, however.

**5-8. Pulse Response of the *RC* Coupling Circuit.** Much electronic equipment employs waveforms other than sinusoids. Steady-state sinusoidal frequency and phase analysis which shows response adequate for the range between fundamental and harmonics of very high order (possibly to the hundredth) may ensure freedom from waveform distortion for nonsinusoidal waves, but greater assurance will be obtained by analysis of the response to an isolated voltage pulse or to a recurrent chain of such pulses. The direct treatment of pulses or square waves will indicate the nature and magnitude of design limitations which must be adhered to for preservation of pulse waveforms.

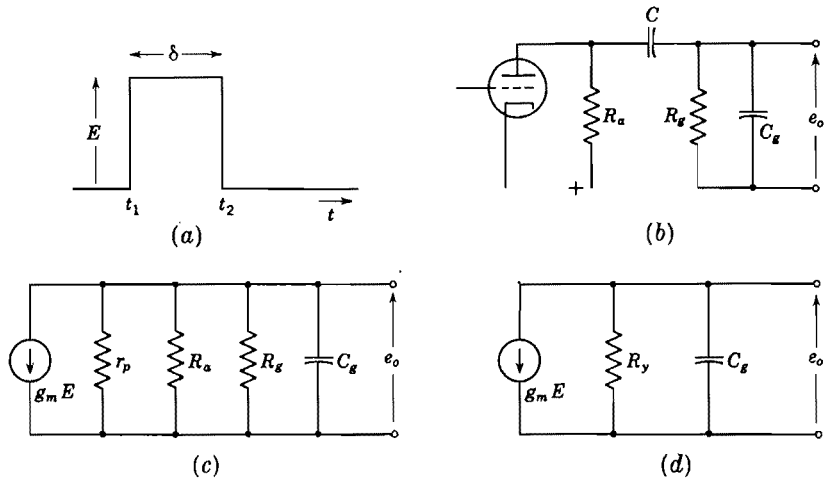



FIG. 5-22. (*a*) Pulse input. (*b*) *RC* amplifier. (*c*, *d*) Equivalent circuits for (*b*).

When a voltage pulse of the form of Fig. 5-22*a*, of amplitude $E$ and time interval $\delta = t_2 - t_1$, is applied as input to a tube in the circuit of Fig. 5-22*b*, the output can be determined in two steps, the first consisting in finding the rise in voltage due to the initial application of the pulse, and the second in finding the response during and after the constant-voltage portion of the pulse. It is found that the *rise time*, or time taken for $e_o$ to reach 90 per cent of its final value, is controlled by the high-frequency characteristics of the circuit, and the remainder of the pulse output depends on the low-frequency characteristics of the circuit.

For study of the rise time an equivalent circuit as at Fig. 5-22*c* is suitable, since the parallel capacity $C_g$ is the controlling factor. The capacity $C_g = C_{pk} + C_{gk}$ for pentodes, where $C_{pk}$ is that of the first tube, $C_{gk}$ that of the following tube. The rise time is independent of the usual values of coupling capacitor $C$, especially if $R_g \gg R_a$. Again defining $R_y$ as the equivalent to $r_p$, $R_a$, and $R_g$ in parallel, as was done in Eq. (5-14), the circuit can be further reduced to that of Fig. 5-22*d*.

Writing a current summation, using the notation $p = d/dt$, gives for the circuit of Fig. 5-22*d*

$$\left(\frac{1}{R_y} + pC_g\right) e_o = 0$$

for $t < t_1$, and

$$\left(\frac{1}{R_y} + pC_g\right) e_o = -g_m E \tag{5-54}$$

for $t > t_1$, for the applied grid-voltage pulse of Fig. 5-22*a*. Since

$$e_o = \frac{q_c}{C_g}$$

where $q_c$ is the charge on capacitor $C_g$ at any instant, then

$$pq_c = -g_m E - \frac{q_c}{R_y C_g} \tag{5-55}$$

After separation of variables the solution is

$$-g_m E - \frac{q_c}{R_y C_g} = \epsilon^{-(t-t_1+K_1)/R_y C_g} \tag{5-56}$$

where $t_1$ is the time of application of the pulse to the circuit and $K_1$ is a constant of integration.

If $C_g$ is initially uncharged, so that, at $t = t_1$, $q_c = 0$, then

$$-g_m E = \epsilon^{-K_1/R_y C_g}$$

so that

$$q_c = -g_m E R_y C_g (1 - \epsilon^{-(t-t_1)/R_y C_g})$$

Since $e_o = q_c/C_g$, the output-voltage waveform during the rise time is

$$e_o = -g_m E R_y (1 - \epsilon^{-(t-t_1)/R_y C_g}) \tag{5-57}$$

The rise time is thus dependent on the time constant $R_y C_g$, the gain or output being dependent on $R_y$. Since $C_g$ has a minimum value fixed by the tube capacities, the rise time can be reduced only by reduction of $R_y$ and sacrifice of gain. This bears out the reasoning of Sec. 5-5.

With 6AK5 pentodes having $C_{gk} + C_{pk} = 6.8$ $\mu\mu$f and assuming zero stray capacity, and with a load resistor of 10,000 ohms, $r_p$ and $R_g$ being large, the time constant is

$$R_y C_g = 10{,}000 \times 6.8 \times 10^{-12} = 0.068 \ \mu\text{sec}$$

The rise is exponential, according to the curve $1 - \epsilon^{-t/RC}$. Such a curve rises to 90 per cent of final value in 2.2 time constants, to 98 per cent in 4 time constants. The defined rise time for the 6AK5 pentodes will be 0.150 $\mu$sec, and 98 per cent of final value will be reached in 0.27 $\mu$sec. At the same time the gain will be $-g_m R_a = -5100 \times 10^{-6} \times 10{,}000 = -51$.

The time constant $R_yC_g$ is also the controlling factor in determining $f_2$, the high-frequency steady-state cutoff. Since this same time constant controls the rise time, the use of compensation to raise $f_2$ will reduce the rise time. However, the use of inductance makes the compensated circuit one of second order, and an overshoot or oscillation may occur. For values of $a$, $m$, and $q$ suggested in Sec. 5-7 serious overshoot will not usually be encountered, and reduced rise time may be obtained.

For usual amplifier constants the rise times will be assumed short with respect to pulse length, so that full pulse voltage may be assumed as reached, and will serve as the starting point for the study of the output voltage during and after the pulse. If full pulse voltage is not reached, the actual value reached may be used in place of $E$ in the following discussion.

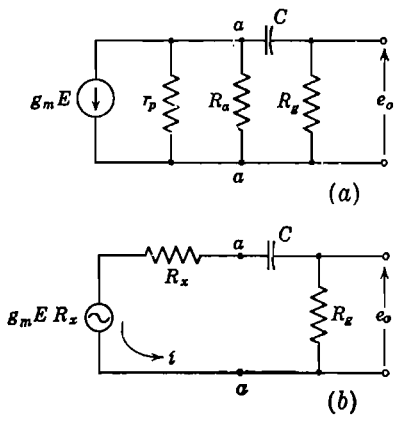



FIG. 5-23. ($a$) Equivalent circuit for study of pulse response during and after the pulse. ($b$) Thévenin modification of ($a$).

For study of response during and after the application of the pulse the low-frequency equivalent circuit of Fig. 5-23 is satisfactory. Using the symbol $R_x$ for the parallel combination of $r_p$ and $R_a$, as was done in Fig. 5-5, the circuit to the left of $a$, $a$ may be transformed by Thévenin's theorem, wherein the current source becomes a voltage-source generator of value $g_mER_x$ during the application of the pulse.

Writing a voltage summation for the circuit of Fig. 5-23$b$ gives

$$\frac{q_c}{C} + (R_g + R_x)i - g_mR_xE = 0 \tag{5-58}$$

After using $p = d/dt$,

$$pq_c = \frac{g_mR_xE}{R_g + R_x} - \frac{q_c}{C(R_g + R_x)} \tag{5-59}$$

This is an equation of the form of Eq. (5-55), and the solution may be written by analogy to Eq. (5-56) as

$$\frac{g_mR_xE}{R_g + R_x} - \frac{q_c}{C(R_g + R_x)} = \epsilon^{-(t-t_1+K_2)/C(R_g+R_x)}$$

At $t = t_1$, $q_c = 0$ by assumption, so that

$$q_c = g_mR_xEC(1 - \epsilon^{-(t-t_1)/C(R_g+R_x)}) \tag{5-60}$$

Since $i = pq$ and $e_o = -R_g i$, the output voltage during the time of application of the pulse voltage $E$ is

$$e_o = -R_g i = -\frac{g_m E R_x R_g}{R_x + R_g} \epsilon^{-(t-t_1)/C(R_x+R_g)} \qquad t_1 < t < t_2 \quad (5\text{-}61)$$

If $t = t_2$, then the above equation gives the voltage output at the end of the pulse application. Noting the exponential, it can be seen that, if the time constant $C(R_x + R_g) \gg \delta$, then the output voltage will be nearly constant during the pulse interval $\delta$ and the output wave will be flat-topped. If this inequality does not hold, the top of the pulse will slope off, or *droop*, during the pulse interval.

At $t_2$ the applied pulse voltage $E$ goes to zero, and the voltage summation of Eq. (5-58) becomes

$$\frac{q(t_2)}{C} + (R_g + R_x)i = 0$$

where $q(t_2)$ is the charge on $C$ at time $t_2$. Obtaining $q(t_2)$ from Eq. (5-60) leads to an expression for $i$ in the interval $t > t_2$, of the form

$$i = -\frac{g_m E R_x}{R_x + R_g}(1 - \epsilon^{-(t_2-t_1)/C(R_x+R_g)})\epsilon^{-(t-t_2)/C(R_x+R_g)} \qquad t > t_2$$

The output voltage $e_o$ after removal of the pulse voltage $E$ is then

$$e_o = -R_g i = \frac{g_m E R_x R_g}{R_x + R_g}(1 - \epsilon^{-(t_2-t_1)/C(R_x+R_g)})\epsilon^{-(t-t_2)/C(R_x+R_g)} \qquad t > t_2 \quad (5\text{-}62)$$

Equations (5-57), (5-61), and (5-62) are the expressions for the three portions of the output waveform of the $RC$-coupled amplifier when excited by a rectangular pulse. Their composite action is illustrated in Fig. 5-24. Computed waveforms for several ratios of time constant to pulse length $\delta$ are also presented in Fig. 5-25. The rise time is assumed negligible and is not indicated in these latter curves.

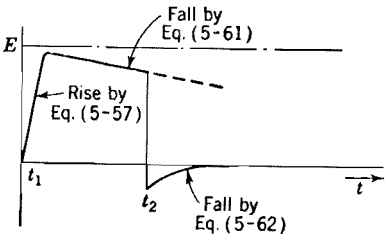



FIG. 5-24. Output-pulse shape.

Ideally the circuit should transmit the pulse without distortion and should have zero output after the end of the pulse at $t = t_2$. The pulse output will be flat-topped or have negligible droop if the time constant $C(R_x + R_g)$ is large with respect to $\delta$. The parentheses of Eq. (5-62) will be near zero for the same condition, and the output will then approximate zero for $t > t_2$, as it should.

The rise time of the output voltage is controlled by a time constant $R_yC_g$, and short rise time requires a value of this time constant small with respect to $\delta$. Droop of the wave top is controlled by the time constant $C(R_x + R_g)$, and from Fig. 5-25 it can be seen that this time constant should approximate ten to fifty times the pulse duration $\delta$ for reasonably good output waveform. Increase of this time constant is limited by the cost of capacitor $C$, since the maximum magnitude of $R_g$ is

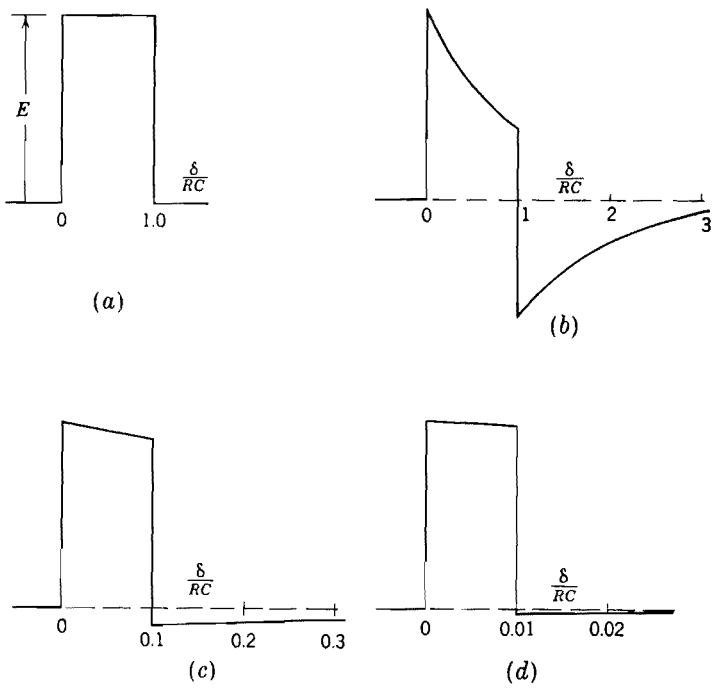



FIG. 5-25. (*a*) Input pulse. (*b*) Output pulse with $C(R_g + R_x)/\delta = 1.0$. (*c*) Output pulse with $C(R_g + R_y)/\delta = 10$. (*d*) Output pulse with $C(R_g + R_y)/\delta = 100$.

usually fixed by grid-current considerations and $R_g$ cannot usually exceed 1 or 2 megohms.

Thus the conditions for small distortion of a voltage pulse are:

1. Small shunt capacity $C_g$ and low load resistance $R_a$, for fast rise time.
2. The series time constant $C(R_x + R_g)$ must be large with respect to the pulse interval $\delta$.

The tube selected should have high $g_m$ value to obtain high gain under the usual conditions of small $R_a$ value.

**5-9. Relations between Sinusoidal and Pulse Response.** Sinusoidal response and pulse response may be readily correlated, so that a knowledge of either form allows the other type of response to be deduced.

In Eq. (5-9) the frequency $f_1$ was defined for an $RC$-coupled amplifier as

$$f_1 = \frac{1}{2\pi C(R_x + R_g)}$$

and in a discussion of Eq. (5-61) it was pointed out that this same time constant $C(R_x + R_g)$ controlled the tilt, or droop, of the top of the pulse-wave output. If $C(R_x + R_g) \gg \delta$, the pulse will be flat-topped and a large value of this time constant is also the condition for a small $f_1$ value. Thus the tilt of the pulse top is reduced if $f_1$ is small for sinusoidal steady-state response of an amplifier.

From Eq. (5-61) it can be seen that the droop is controlled by the exponential factor. That is, for the full pulse duration $t - t_1 = \delta$, the exponential term may be expressed by its series as

$$\epsilon^{-\delta/C(R_x+R_g)} = 1 - \frac{\delta}{C(R_x + R_g)} + \frac{1}{2}\left[\frac{\delta}{C(R_x + R_g)}\right]^2 \cdots$$

For values of $\delta/C(R_x + R_g) \leqq 0.15$ the exponential may be replaced with its first two series terms with small error. Equation (5-61) may then be written as

$$e_o \cong -\frac{g_m E R_x R_g}{R_x + R_g} + \frac{g_m E R_x R_g}{R_x + R_g}\left[\frac{\delta}{C(R_x + R_g)}\right]$$

The first term on the right is the output without droop or with $\delta \ll C(R_x + R_g)$. The second term on the right contributes droop, or reduces the output voltage. The per cent droop may then be found as

$$\text{Per cent droop} \cong \frac{[g_m E R_x R_g/(R_x + R_g)][\delta/C(R_x + R_g)]}{g_m E R_x R_g/(R_x + R_g)} \times 100\%$$

$$= \frac{\delta \times 100\%}{C(R_x + R_g)} \tag{5-63}$$

In view of the definition for $f_1$

$$\text{Per cent droop} \cong 2\pi\delta f_1 \times 100\% \tag{5-64}$$

and thus the droop, or tilt, of the pulse top is related to the $f_1$ value for the sinusoidal steady-state response.

In Eq. (5-14) the frequency $f_2$ was defined for sinusoidal steady-state response as

$$f_2 = \frac{1}{2\pi C_g R_g}$$

for an $RC$-coupled amplifier. By comparison with Eq. (5-57) it can be seen that the same time constant is related to the pulse rise time of the amplifier. Since the exponential term of Eq. (5-57) rises to 90 per cent

of final value in 2.2 time constants, then the *defined rise time* of an amplifier is

$$T_R = 2.2C_gR_g = \frac{2.2}{2\pi f_2} = \frac{0.35}{f_2} \tag{5-65}$$

and thus the rise time for pulse response is related to the $f_2$ frequency of sinusoidal response.

It may also be seen that

$$T_R f_2 = 0.35 \tag{5-66}$$

which states that the product of $f_2$ frequency (or bandwidth in most cases) and rise time is a constant. Thus a greater $f_2$ frequency leads to a shorter rise time.

It may also be shown that for a cascaded $RC$ amplifier the total rise time is given by

$$T_R = \sqrt{T_a^2 + T_b^2 + T_c^2 + \cdots}$$

and for $n$ identical stages

$$T_R = T_a\sqrt{n}$$

where $T_a$, $T_b$, etc., are the individual-stage rise times.

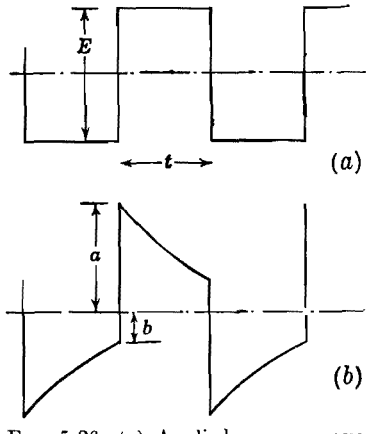



FIG. 5-26. (*a*) Applied square wave. (*b*) Output of an $RC$ circuit.

**5-10. Square-wave Response and Testing.** From the previous section it can be seen that amplifier performance around a voltage jump is an indication of high-frequency response, while performance during the flat pulse top is an indication of low-frequency response. Both such types of input are embodied on a repetitive basis in the square wave, and its use makes available a rapid means for amplifier-circuit study when the output is displayed on a cathode-ray oscilloscope.

For study of low-frequency response a square wave of frequency somewhat above $f_1$ is applied, giving an output which may appear as in Fig. 5-26*b*. The jump in voltage is composed of two parts, $a$ above the axis and $b$ below, so that

$$a + b = E$$

where $E$ is the pulse amplitude applied to the $RC$ coupling circuit. It

also appears that

$$b = a\epsilon^{-\delta/C(R_x+R_g)} = a\epsilon^{-1/C(R_x+R_g)2f}$$

where the nomenclature refers to Fig. 5-5, and where $\delta$ refers to the half period of the square wave or $f$ is the square-wave frequency. Then

$$a = \frac{E}{1 + \epsilon^{-1/C(R_x+R_g)2f}} \tag{5-67}$$

The response of an $RC$ amplifier to a 500-cycle square wave is shown in Fig. 5-27. Here ratios of $RC$ to the half period $\delta$ of 0.5, 2, and 25 were used, and it is apparent that for true reproduction of waveform the $RC$ time constant must be very long with respect to the half period of the

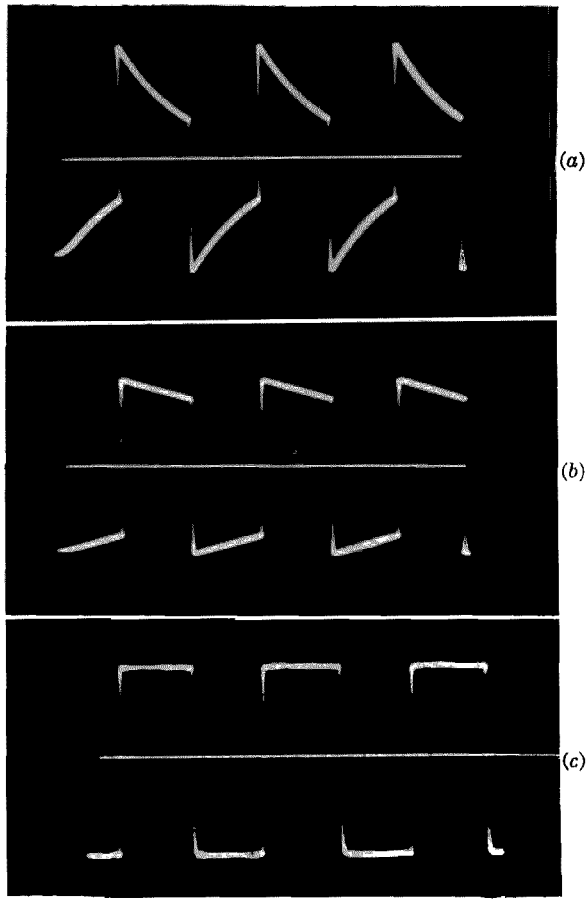



FIG. 5-27. Half period for all waves is 1000 $\mu$sec = $\delta$. (*a*) Response for $RC/\delta = 0.5$. (*b*) Response for $RC/\delta = 2$. (*c*) Response for $RC/\delta = 25$.

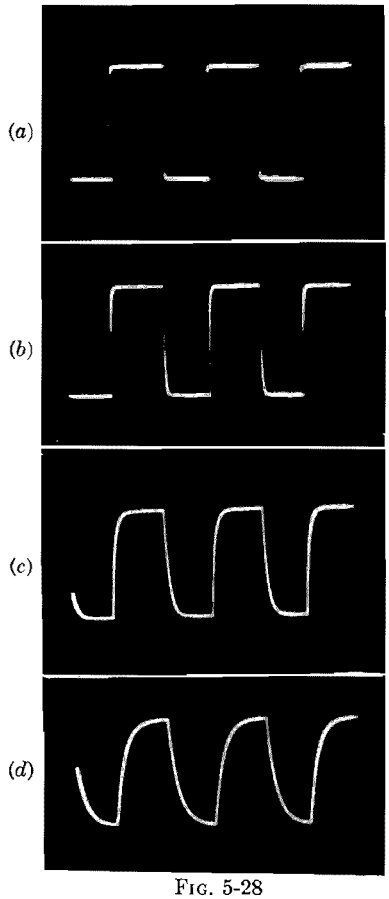



FIG. 5-28

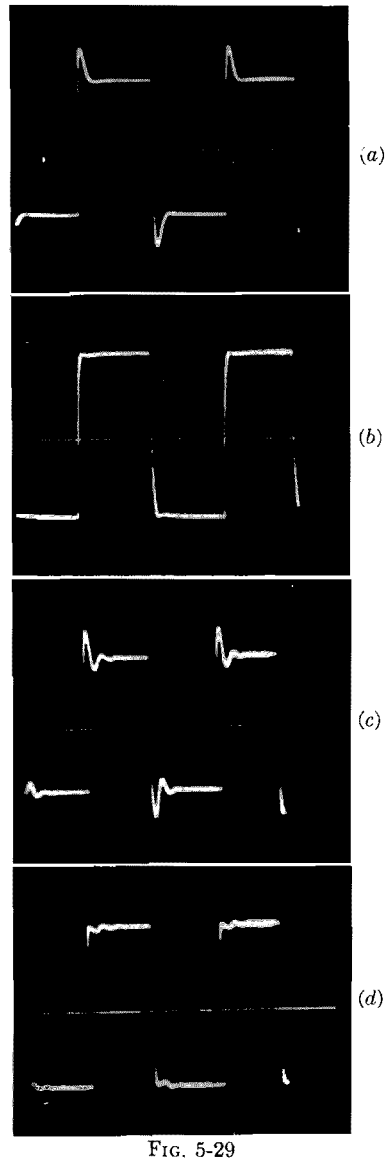



FIG. 5-29

FIG. 5-28. Rise time of $RC$ amplifier: ($a$) applied waveform, ($b$) $R_v C_g/\delta = 0.01$, ($c$) $R_v C_g/\delta = 0.04$, ($d$) $R_v C_g/\delta = 0.1$.

FIG. 5-29. ($a$, $b$) Compensated amplifier response to a square wave. ($c$, $d$) Response with more oscillatory conditions.

wave. These curves establish criteria for visual comparison with the low-frequency response of other amplifiers.

High-frequency response may be observed by application of a square wave of frequency approximating $f_2$. Figure 5-28 indicates the waveforms to be expected for various ratios of $\delta$ to the shunt $RC$ time constant. It can be seen that a very small time constant $C_g R_u$ is desirable to reduce rise time.

A second type of high-frequency response is also possible if high-frequency inductive compensation is used or if resonances occur within the amplifier. Depending on the damping or ratio $\omega L/R$ of the circuit, an overshoot or damped oscillation may occur, since the governing equation becomes of second order. Such performance will show on the oscilloscope as a shock-excited oscillation at the jump and is usually undesirable, although a very small overshoot may at times be allowed, to decrease rise time. Typical responses are shown in Fig. 5-29. An estimation of the frequency of the oscillation may be made by comparison with the period of the square wave, and this frequency estimate will aid in locating the circuit or combination of elements responsible for the oscillation.

**5-11. Fourier Series.** In the previous section the pulse response of the $RC$ amplifier was considered on a transient voltage or time basis (time domain). This seems the most useful method of directly predicting the waveforms to be encountered in the circuits. However, valuable additional information concerning amplifier response may also be obtained if the characteristics of a pulse are considered on a sinusoidal basis (frequency domain). Thus another tie between the time response and the frequency response may be developed.

If a chain of repeated pulses is considered as a periodic function, this function may be represented on a frequency basis by an infinite series, including a constant term and sine and cosine terms having angular velocities that are harmonics of a fundamental $\omega_r$. The lowest frequency or the fundamental angular velocity $\omega_r$ will be determined by the repetition rate, or the number of pulses per second. The function is then said to be represented by a *Fourier series.*

A Fourier series may be written as

$$f(t) = A_o + a_1 \cos \omega_r t + a_2 \cos 2\omega_r t + \cdots + a_k \cos k\omega_r t + \cdots + b_1 \sin \omega_r t + b_2 \sin 2\omega_r t + \cdots + b_k \sin k\omega_r t + \cdots \quad (5\text{-}68)$$

Such a series may be summed as

$$f(t) = A_o + \sum_{k=1}^{\infty} a_k \cos k\omega_r t + \sum_{k=1}^{\infty} b_k \sin k\omega_r t \quad (5\text{-}69)$$

The coefficients of the series may be evaluated through certain manipulations and trigonometric identities. In Fig. 5-30 the period is identified

as $2\pi/\omega_r$, equal to the repetition time $t_r$. If Eq. (5-68) is multiplied by $dt$ and integrated over the period $-\pi/\omega_r$ to $+\pi/\omega_r$, all sine and cosine terms on the right go to zero, leaving as a value for the constant term

$$A_0 = \frac{\omega_r}{2\pi} \int_{-\pi/\omega_r}^{+\pi/\omega_r} f(t)\, dt \tag{5-70}$$

This represents the average value of the function $f(t)$.

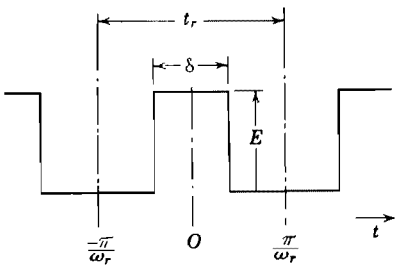



FIG. 5-30. Periodic interval for Fourier series.

If Eq. (5-68) is multiplied by $\cos k\omega_r t\, dt$ and integrated over the period $-\pi/\omega_r$ to $+\pi/\omega_r$ as before, it follows from trigonometric relations that all terms on the right side will vanish except the one having $a_k$ as its coefficient. This provides a method of determining the coefficients of the cosine terms as

$$a_k = \frac{\omega_r}{\pi} \int_{-\pi/\omega_r}^{+\pi/\omega_r} f(t) \cos k\omega_r t\, dt \tag{5-71}$$

Similar manipulation determines the coefficients of the sine terms as

$$b_k = \frac{\omega_r}{\pi} \int_{-\pi/\omega_r}^{+\pi/\omega_r} f(t) \sin k\omega_r t\, dt \tag{5-72}$$

An even function has $f(x) = f(-x)$, and an odd function has

$$f(-x) = -f(x)$$

Thus a cosine is an even function, and a sine is an odd function. If $f(t)$ is an odd function, then Eq. (5-71) includes the product of an odd function and an even function, and the result will be odd. The integral of an odd function taken through equal positive and negative limits is zero, since the positive contribution is nullified by the negative contribution. Therefore, if $f(t)$ is odd, its $a_k$ coefficients will be zero and a *Fourier series for an odd function includes only sine terms.*

By similar reasoning, if $f(t)$ is even, the $b_k$ coefficients will vanish and a *Fourier series for an even function will consist only of cosine terms.* The constant term, depending on the average value of $f(t)$ over the period, may appear in either form of series.

**Example.** Find the Fourier-series coefficients representing a time function given by a chain of square pulses, or the wave of Fig. 5-31, having amplitude $E$.

Choosing the time origin to make the wave an even function causes the resultant series to be one of cosine terms. Inspection shows that the wave has an average value or d-c term $A_0 = E/2$.

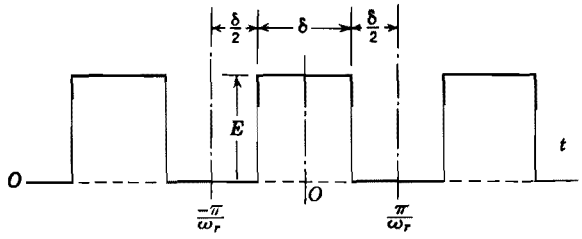



FIG. 5-31. A square wave.

The coefficients may be written as

$$a_k = \frac{\omega_r}{\pi} \int_{-\pi/2\omega_r}^{+\pi/2\omega_r} E \cos k\omega_r t \, dt$$

the limits being changed to the half period because $E = 0$ elsewhere. Then

$$a_k = \frac{\omega_r E}{\pi} \frac{\sin k\omega_r t}{k\omega_r} \Big]_{-\pi/2\omega_r}^{+\pi/2\omega_r}$$
$$= \frac{2E}{\pi k} \sin \frac{k\pi}{2} \tag{5-73}$$

The Fourier series is then given by

$$f(t) = \frac{E}{2} + \frac{2E}{\pi k} \sum_{k=1}^{\infty} \sin \frac{k\pi}{2} \cos k\omega_r t \tag{5-74}$$

and the first few terms may be written as

$$f(t) = \frac{E}{2} + \frac{2E}{\pi} \cos \omega_r t - \frac{2E}{3\pi} \cos 3\omega_r t + \frac{2E}{5\pi} \cos 5\omega_r t - \frac{2E}{7\pi} \cos 7\omega_r t + \cdots \tag{5-75}$$

The principle of building up a square wave such as discussed in the example is illustrated by superposition of harmonics in Fig. 5-32. Here the fundamental (or first harmonic) is added to the third, fifth, and

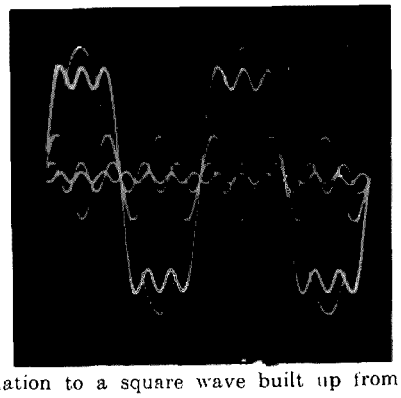


FIG. 5-32. Approximation to a square wave built up from first, third, fifth, and seventh harmonics.

seventh. Many higher harmonics would be needed for faithful reproduction of a square wave.

The coefficients for the various frequency terms present in a chain of square pulses or a square wave may be easily computed from the series of Eq. (5-75) and appear in Table 9. In the reproduction of a square wave the 55th harmonic still has an amplitude above 1 per cent of the fundamental and will contribute appreciably to building up the desired accurate waveform.

TABLE 9. RELATIVE HARMONIC AMPLITUDES IN A SQUARE WAVE
(Pulse amplitude $E = 1$)

| *Harmonic* | *Amplitude* | *Harmonic* | *Amplitude* |
|---|---|---|---|
| Fundamental | 0.636 | 25 | 0.025 |
| 3 | −0.212 | 35 | −0.018 |
| 5 | 0.127 | 45 | 0.014 |
| 7 | −0.091 | 55 | −0.011 |
| 9 | 0.070 | 65 | 0.0098 |
| 11 | −0.058 | 75 | −0.0085 |
| 13 | 0.049 | 85 | 0.0075 |
| 15 | −0.042 | 95 | −0.0067 |

A negative sign on a harmonic implies a phase reversal for that harmonic.

**5-12. Analysis of a Recurrent Rectangular Pulse.** The Fourier series allows sinusoidal interpretation of any periodic time function applied

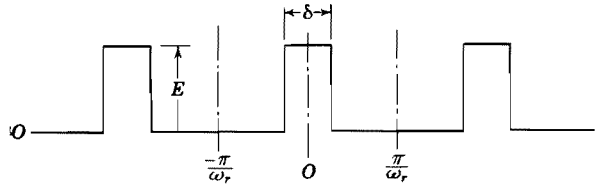



FIG. 5-33. Recurrent rectangular pulses.

to an amplifier. Each frequency component will be operated upon by the gain at that frequency, the result being a series giving the frequencies and their amplitudes and phase angles present in the amplifier output.

If the frequency response of the amplifier is such that some of the higher frequencies composing the input pulses are not amplified as much as the others, then the output pulses will not have the shape of the input pulses. The importance of each harmonic in building up the pulse shape is related to the relative harmonic amplitude as compared with the fundamental. Failure properly to amplify a large-amplitude harmonic will have more effect on the pulse waveform than would be caused by improper amplification of a small-amplitude harmonic.

For discussion consider the pulse chain of Fig. 5-33. Assuming even

symmetry, which can be achieved by proper choice of the time origin for the recurrent pulses, then the $b_k$ coefficients for the Fourier representation vanish, and

$$a_k = \frac{\omega_r}{\pi} \int_{-\pi/\omega_r}^{+\pi/\omega_r} f(t) \cos k\omega_r t \, dt$$

Since $f(t)$ is zero in the interval except between $t = -\delta/2$ and $t = +\delta/2$, the limits of integration may be changed,

$$a_k = 2f_r \int_{-\delta/2}^{+\delta/2} E \cos 2\pi f_r k t \, dt \tag{5-76}$$

Performing the integration and introducing the limits gives

$$a_k = 2Ef_r\delta \frac{\sin \pi k f_r \delta}{\pi k f_r \delta} \tag{5-77}$$

This expresses the amplitude of the $k$th harmonic for a rectangular pulse, in terms of duration $\delta$ and repetition frequency $f_r$.

The relative amplitude of the harmonics as predicted by Eq. (5-77) is illustrated in Fig. 5-34. The length of any vertical line under the

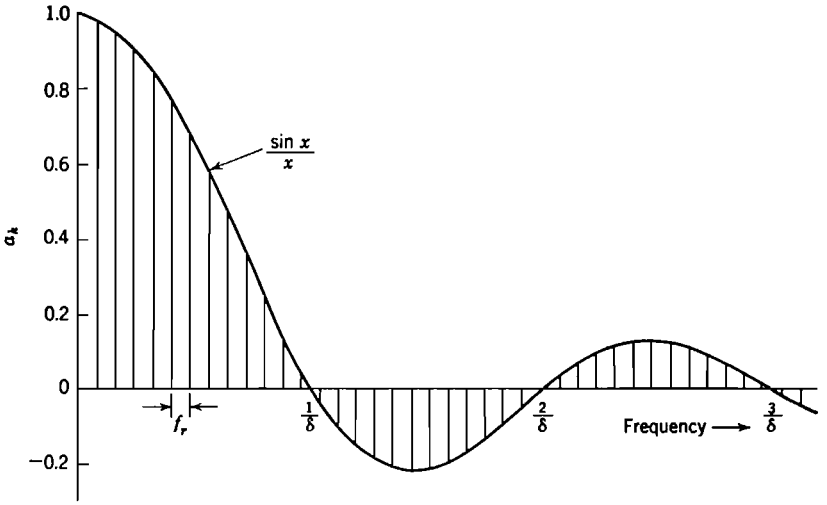



FIG. 5-34. Spectrum distribution of harmonics in recurrent pulse chain.

envelope indicates the amplitude $a_k$ of that particular cosine term in the Fourier-series representation of the pulses. Negative coefficients indicate a phase shift of 180° in that particular harmonic. The shape of the envelope indicates the relative importance of various groups of harmonics in adding to the shape of the pulse.

The crossover points or zero-amplitude points are functions of $\delta$ only, since at those points

$$\frac{\sin \pi k f_r \delta}{\pi k f_r \delta} = 0$$

$$\pi k f_r \delta = n\pi \qquad n = 1, 2, 3, \ldots$$

The particular harmonic frequency at which the amplitude is zero is then

$$k f_r = \frac{n}{\delta} \tag{5-78}$$

Therefore the frequency $kf_r$ at which an amplifier may be allowed to cut off is determined as a function of the pulse duration and is not a function of the repetition rate.

The number of harmonics in each loop is inversely proportional to the repetition rate. As a pulse occurs less and less often, more harmonics appear in each loop of the envelope and the frequency spectrum is filled in. *At the limit where a pulse occurs only once, or $f_r = 0$, a continuous spectrum of frequencies is obtained.**

To illustrate the point, assume a pulse having a duration $\delta$ of 1 $\mu$sec, so that the first zero in the spectrum occurs at $1/\delta = 1$ Mc, the second at 2 Mc, etc. If the repetition rate $f_r$ is 10,000 per second, there will be 100 harmonics in each loop of the envelope; if $f_r = 1000$, there will be 1000 harmonics in each loop; and if $f_r = 100$, then there will be 10,000 harmonics in the first and each other loop of the spectrum.

It can be seen that as further loops of Fig. 5-34 are included within an amplifier pass band, the fidelity of pulse reproduction will improve; or, conversely, for shorter pulse duration a greater amplifier bandwidth will be required for adequate reproduction. The decision as to how many harmonics or how many loops of the envelope are to be included is still arbitrary; it must be based on economics of design, but if the bandwidth of the amplifier extends out to the fourth or fifth zero, reproduction will usually be excellent. Actually, since most of the high-amplitude harmonics appear in the first loop of the envelope, reproduction with tolerable distortion may be obtained by transmitting frequencies to the first zero. Thus the amplifier bandwidth would range from the repetition frequency at the low-frequency end to a frequency of $1/\delta$ at the high-frequency end.

* In this case the frequency spectrum is obtained by means of the Fourier integral, giving for the amplitude distribution for an even function pulse

$$S(\omega) = \frac{E\delta}{2\pi} \frac{\sin (\omega\delta/2)}{\omega\delta/2}$$

and this is represented by the envelope of Fig. 5-34, being of the form $(\sin x)/x$. It may be found that the amplitude of the fundamental or differential frequency adjacent to zero is $E\delta/2\pi$.

This analysis may be applied to the square wave. For such a wave $\delta = 1/2f_r$, and Eq. (5-77) becomes

$$a_k = E\frac{\sin(k\pi/2)}{k\pi/2} = \frac{2E}{\pi k}\sin\frac{k\pi}{2}$$

and this result leads to a series identical with that of Eq. (5-75), as it should. A frequency-spectrum diagram is shown for this case in Fig. 5-35, along with the envelope predicted by Eq. (5-77). Since the wave is

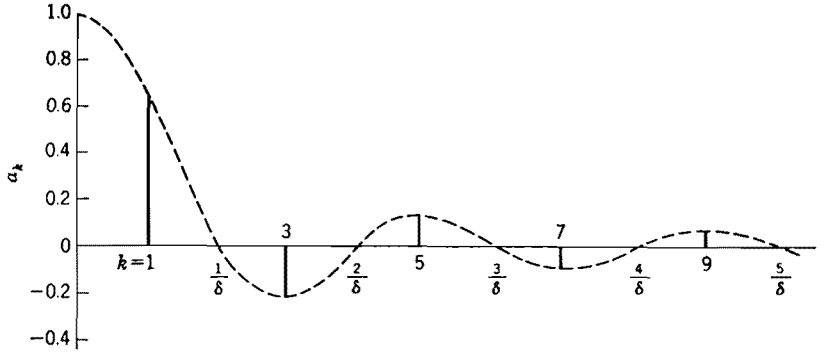



FIG. 5-35. Frequency spectrum of a square wave.

square, or $\delta = 1/2f_r$, it can be seen that $1/\delta = 2f_r$, or $k = 2$, in each loop. There will be only one odd harmonic in each loop of the envelope, the even harmonics falling at the zero points of the envelope, and these terms are absent from the series.

**Example.** For a chain of rectangular pulses having $\delta = 10$ μsec, and with a repetition rate of 4000 per second, find the amplitude of the first five harmonics. How many harmonics will fall in each loop of the frequency spectrum?

From Eq. (5-77)

$$a_k = 2Ef_r\delta\frac{\sin\pi kf_r\delta}{\pi kf_r\delta}$$

where $E$ is the pulse amplitude. Then

$$a_k = 2E \times 4 \times 10^3 \times 10^{-5}\frac{\sin\pi k \times 4 \times 10^3 \times 10^{-5}}{\pi k \times 4 \times 10^3 \times 10^{-5}}$$

$$= \frac{2E}{\pi k}\sin 0.04\pi k$$

The series (neglecting any d-c term) is

$$f(t) = 0.0795E\cos 2\pi \times 4000t + 0.0790E\cos 2\pi \times 8000t + 0.0780E\cos 2\pi \times 12{,}000t + 0.0764E\cos 2\pi \times 16{,}000t + 0.0745E\cos 2\pi \times 20{,}000t + \cdots$$

Since $kf_r = 1/\delta = 100{,}000$ cycles,

$$k = \frac{100{,}000}{4000} = 25 \text{ harmonics}$$

in each loop of the frequency spectrum.

**5-13. Noise in Amplifiers.** The maximum amplification which can be utilized by cascading successive stages of amplification is limited by electrical *noise*, or erratic currents generated in the tubes or in the circuits. Gain sufficient to bring the internal noise up to the normal signal output level is thus the usable limit. Any signal appreciably weaker than the noise will be covered up or masked.

Noise is introduced by a number of sources, some due to the inherent particle nature of electron flow in tubes, and some to circuit sources external to the tubes.

*Thermal Noise.* This source is created by the random motions of free charges present in any conductor. The noise energy is uniformly distributed over all frequencies and is believed due to impacts of electrons with molecules, each impact producing a very short pulse. The thermal activity in a resistor $R$ is a function of temperature, and the noise voltage of the resistor can be expressed as

$$E_{\text{noise}} = \sqrt{4kTR\,\Delta f} \tag{5-79}$$

where $k$ = Boltzmann's constant = $1.38 \times 10^{-23}$ joule/°K
$T$ = temperature, °K
$R$ = resistance, ohms
$\Delta f$ = bandwidth accepted by the circuit, cycles; this is the frequency differential between half-power points on a decibel-gain curve

A thermal-agitation noise source can be considered equivalent to a generator of voltage $E_{\text{noise}}$, in series with a noiseless internal resistance $R$. Because all circuit components have resistance, thermal noise is present in all circuit components.

*Shot Noise.* This is caused by the randomness of electron emission from a cathode. Space-charge operation of tubes reduces this noise considerably because the space-charge reservoir of electrons tends to smooth out the individual emission pulses. The noise appears to have a frequency spectrum uniformly distributed over the useful frequency region. It is frequently convenient to express this tube noise in terms of a resistor $R_{eq}$ which, if inserted in series with the grid of a noise-free tube, would generate the same noise current in the plate circuit as is caused by the actual noisy tube. An approximate expression for this resistance for *triodes* is

$$R_{eq} \cong \frac{2.5}{g_m} \quad \text{ohms}$$

Triodes have equivalent resistances ranging in value from 210 for a 6J4 to 1500 for a 6SL7.

*Partition Noise.* In pentodes the current divides irregularly between two or more positive electrodes. That is, it is quite arbitrary whether a particular electron moves to screen or anode, although the time average currents are constant. This arbitrary partition of the current from instant to instant adds to the shot noise and makes a pentode noisier than a triode. For *pentodes* the noise equivalent resistance becomes

$$R_{eq} \cong \frac{2.5}{g_m} + \frac{20 I_s}{g_m I_k}$$

where $I_s$ and $I_k$ are screen and cathode currents, respectively. Values of $R_{eq}$ for pentodes vary from 700 to over 13,000.

*Microphonic Noise.* This is an important noise source in equipment subject to vibration. The plate current is dependent on the space relation of anode to grid and cathode within the tube, and vibration of the tube elements will produce current variations, or noise. Tube types vary widely in their susceptibility to vibration, and special types are available for particularly difficult applications.

*Electric and Magnetic Pickup.* Shielding is useful in eliminating this source, but at 60 cycles shielding is either very expensive or ineffective. Induced voltages can be reduced by isolation or orientation of apparatus and by elimination of stray capacities to power circuits.

*Plate-supply Hum.* Adequate filtering can usually reduce the level of residual alternating current in rectified power supplies to a suitably low level, although reduction to levels below a few millivolts requires considerable care.

*Cathode Hum.* Tubes with a-c-operated heaters may have power-frequency hum from several sources. Emission of electrons from the heater may introduce hum, but this may usually be prevented by operating the heater circuit at 10 to 20 volts positive with respect to the cathode.

Capacitance between heater and cathode may produce hum in some circuits where the cathode operates above ground a-c potential. By grounding the heater circuit at a center tap some of this hum may be balanced out.

Certain types of tubes have a heater construction which sets up a 60-cycle magnetic field round the heater, and this field can influence electrons in transit, causing hum. Commercial tubes differ considerably from type to type in this respect. It is difficult completely to eliminate all 60-cycle hum when a-c heater operation is used, a residual of a few microvolts usually being present. When complete elimination is necessary, the heaters must be operated from rectified and filtered alternating current.

## PROBLEMS

**5-1.** In Fig. 5-36 $g_m$ = 2000 μmhos, $r_p$ = 30,000 ohms, $C_{pk}$ = 7.0 μμf, and $C_{in}$ = 70 μμf. Draw frequency-response and phase-angle curves for this amplifier, using a linear amplitude scale. Indicate frequencies $f_1$ and $f_2$, and calculate the mid-frequency gain in decibels.

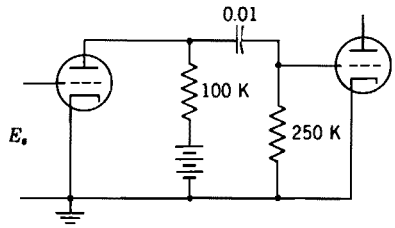
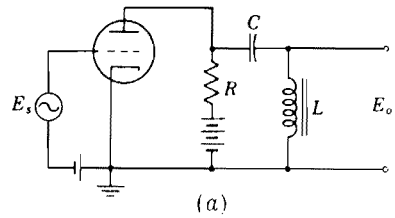


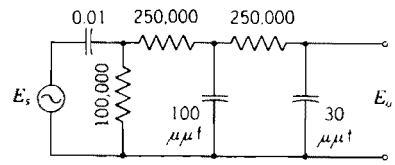

FIG. 5-36

**5-2.** The tube of Fig. 5-36 now has $\mu$ = 20, $r_p$ = 7700 ohms. (*a*) Compute the value of coupling capacitor $C$ for $f_1$ = 20 cycles. (*b*) To what value must $R_a$, the plate-load resistor, be changed to make $f_2$ = 250,000 cycles? What is the mid-frequency gain?

**5-3.** Using the gain-ratio methods of Sec. 5-2, find an expression for $|A_L/A_M|$ in terms of $f/f_1$ ratios and circuit parameters for the circuit of (*a*) Fig. 5-37. If $C$ and $L$




(*a*)




(*b*)

FIG. 5-37

are given, find an expression for $R$ giving flattest low-frequency gain. If $C$ = 0.01 μf and $L$ = 200 h, calculate $R$ for flattest gain, and plot gain in decibels and phase angle against frequency over the low-frequency region.

**5-4.** A pentode amplifier tube is to operate at $I_b$ = 2.5 ma, $E_{cc}$ = −4 volts, $E_{c2}$ = 100 volts. The load $R_a$ = 75,000 ohms, $g_m$ = 3200 μmhos, $r_p$ = 1 megohm, $C$ = 0.01 μfd, $R_g$ = 500,000 ohms, $C_{gk}$ = 10 μμf, $C_{gp}$ = 0.004 μμf, $C_{pk}$ = 9 μμf, stray wiring capacity = 15 μμf.

(*a*) Assuming screen and cathode resistors bypassed, calculate the mid-frequency gain and $f_1$ and $f_2$.

(*b*) Between what frequencies is the phase angle 180° ± 20°?

**5-5.** Write the expression for $E_o/E_s$ for the circuit of Fig. 5-37*b*. Also write $A_L/A_M$ and $A_H/A_M$ in terms of $f/f_x$ ratios.

**5-6.** For the circuit of Fig. 5-37*b* specify $f_1$ and $f_2$, and plot a decibel-gain curve; show the asymptotes for gain, and state the asymptote slopes in the several regions.

**5-7.** A three-stage pentode amplifier with *RC* coupling is to have constant gain within 0.5 db up to 50 kc. Compute the $f_2$ value for the identical individual stages.

**5-8.** What over-all gain is possible with a cascaded six-stage amplifier using 6AK5 pentodes, if the $f_2$ over-all value is to be 500 kc? What must $f_2$ be for the individual stages?

**5-9.** A pentode having $r_p = 0.8$ megohm, $g_m = 3000$ $\mu$mhos in an *RC* amplifier has $R_a = 25{,}000$ ohms, $R_g = 250{,}000$ ohms, $C = 0.01$ $\mu$fd, $C_{pk} = 8$ $\mu\mu$f, and $C_g = 27$ $\mu\mu$f.

(*a*) Find $f_1$ and $f_2$, and draw the frequency-response curve in decibel gain.

(*b*) Compensate the low-frequency response, specify the values of circuit components required, and draw the circuit.

(*c*) Provide shunt compensation for the high-frequency response to give a 5 per cent rise over $A_M$; specify circuit constants, and determine the new value of $f_2$.

**5-10.** Each stage of a three-stage pentode amplifier has a mid-frequency gain of 30 (voltage) and an $f_2$ value of 375 kc. The complete amplifier is to be modified by reducing $R_a$ values to give an over-all $f_2$ value of 600 kc. If $g_m = 3000$ $\mu$mhos, what value of $R_a$ should be used, and what stage gain and over-all gain will then be available?

**5-11.** The first stage of a two-stage *RC*-coupled amplifier has a voltage gain of 40 and $f_1$ of 60 cycles and $f_2$ of 28,000 cycles. The second stage has a voltage gain of 78 and $f_1$ of 82 cycles and $f_2$ of 37,000 cycles. Plot the over-all response curve in decibels.

**5-12.** A shunt-compensated wideband amplifier uses pentodes with $g_m = 4500$ $\mu$mhos to provide an over-all gain of 16,000 (voltage) with a 3.75-Mc bandwidth. If the value of plate-circuit shunt capacity $C_g$ is 27 $\mu\mu$f in each stage, find the number of stages required if $a = 0.41$.

**5-13.** A rectangular pulse having amplitude of 50 volts and duration of 100 $\mu$sec is applied to the pentode and circuit of Prob. 5-9, before compensation. Plot the waveform of the output voltage against time.

**5-14.** A rectangular pulse as in Fig. 5-38 is applied to the grid of an *RC*-coupled amplifier having a pentode with $g_m = 4500$ $\mu$mhos, $r_p = 0.8$ megohm. If $R_g = 1$ megohm and the total anode-circuit shunting capacity $C_g = 25$ $\mu\mu$f:

(*a*) Find the value of $R_a$ for a rise time of 0.20 $\mu$sec to 90 per cent final value.

(*b*) What peak output voltage will be obtained?

(*c*) Specify the value of coupling capacitor C to permit only 10 per cent droop in output voltage during the time of the pulse.

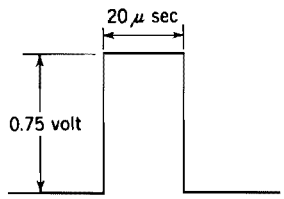



FIG. 5-38

**5-15.** Write the first six terms of the Fourier series for the periodic triangular wave of Fig. 5-39*a*. Include actual frequencies present if the period of the wave represents 0.00025 sec.

**5-16.** Write the first six terms of the Fourier series for the periodic wave of Fig. 5-39*b*. The frequency of repetition of the wave is 2000 per second.

**5-17.** The recurrent pulse chain of Fig. 5-39*c* has a pulse duration of 50 $\mu$sec, with a repetition rate of 4000 per second. Draw a frequency-spectrum diagram for all frequencies in the first three loops of the curve. What is the amplitude of the 67th harmonic?

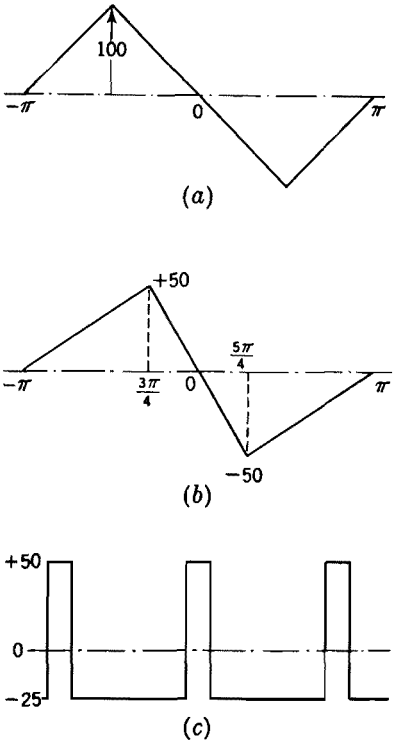



FIG. 5-39

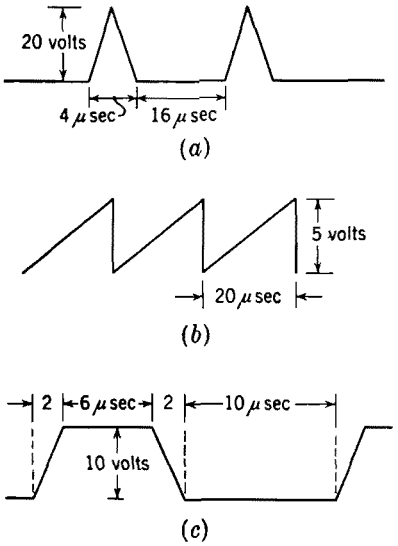



FIG. 5-40

**5-18.** Plot amplitude vs. frequency for the first 10 terms of the Fourier series representing the time functions of each of the waves of Fig. 5-40.

**5-19.** Plot frequency spectrums for the following pulse chains of rectangular pulses: (*a*) $f_r$ = 10,000 per second, $\delta$ = 50 $\mu$sec; (*b*) $f_r$ = 5000 per second, $\delta$ = 50 $\mu$sec; (*c*) $f_r$ = 1000 per second, $\delta$ = 50 $\mu$sec.

**5-20.** A certain *RC* two-stage amplifier has a mid-frequency voltage gain of 220 and an $f_2$ value of 16,000 cycles. If half the gain can be sacrificed, design a single *RC* network which will extend the high-frequency end of the band. One of the grid resistors is 250,000 ohms. What will be the new $f_2$ value?

## REFERENCES

1. Terman, F. E.: Universal Amplification Charts, *Electronics*, **10**: 34 (June, 1937).
2. Arguimbau, L. B.: "Vacuum Tube Circuits," John Wiley & Sons, Inc., New York, 1948.
3. Valley, G. E., Jr., and H. Wallman: "Vacuum Tube Amplifiers," McGraw-Hill Book Company, Inc., New York, 1948.

4. Freeman, R. L., and J. D. Schautz: Video-amplifier Design, *Electronics*, **10**: 22 (August, 1937).
5. Seely, S. W., and C. N. Kimball: Analysis and Design of Video Amplifiers, *RCA Rev.*, **3**: 290 (1939).
6. Herold, E. W.: High-frequency Correction in Resistance-coupled Amplifiers, *Communications (NY)*, **18**: 11 (August, 1938).
7. Larsen, M. J.: Low-frequency Compensation of Video Amplifiers, *Proc. IRE*, **33**: 666 (1945).
8. Bedford, A. V., and G. L. Fredendall: Transient Response of Multi-stage Video-frequency Amplifiers, *Proc. IRE*, **27**: 277 (1939).
9. MacLean, W. R.: Ultimate Bandwidth in High-gain Multistage Video Amplifiers, *Proc. IRE*, **32**: 12 (1944).
10. Frank, E.: "Pulsed Linear Networks," McGraw-Hill Book Company, Inc., New York, 1943.
11. Johnson, J. B.: Thermal Agitation of Electricity in Conductors, *Phys. Rev.*, **32**: 97 (1928).

CHAPTER 6

# LOW-FREQUENCY POWER AMPLIFIERS

In many applications involving vacuum tubes appreciable power output is required, and large input signals are applied to obtain that power. Graphical performance analysis must then be used, since the dynamic transfer characteristic will usually not be linear over the whole of large grid-voltage excursions.

This nonlinearity may create amplitude distortion, and operating conditions to prevent or reduce such distortion are available. In other cases, where the tube operates merely as an a-c or d-c power switch, distortion is not important and operating conditions are fixed by allowable currents and internal tube power loss. Since the vacuum tube is a loss-limited device, high power-handling efficiency is important.

Class A operation is normally employed when minimum distortion of waveform is desired. However, if some distortion can be tolerated, the operating efficiency and power output of a given tube can be increased by Class AB operation. Push-pull operation reduces distortion from either Class A or Class AB operation and makes possible high-power-output Class B operation.

Thus a designer must select the class of operation, a tube with sufficient plate-dissipation rating, and then determine the proper load resistance, the d-c supply voltages, the input signal required, and the resultant distortion if this is a critical matter.

**6-1. Output Circuits for Low-frequency Power Amplifiers.** The simple series-fed output circuit of Fig. 6-1*a* is almost never employed when appreciable a-c power output is desired. In the first place, the power loss due to the d-c component of anode current flowing through the load resistor is undesirable and reduces the power efficiency; second, the d-c anode voltage is lowered; third, the actual load rarely has the magnitude required for the desired power output and distortion conditions.

The circuit of Fig. 6-1*b* avoids the d-c power loss and reduction of anode voltage but still requires that the load be of a suitable magnitude, and this condition is not frequently met.

The circuit of Fig. 6-1*c* is most commonly employed for low-frequency power output. The transformer having a primary-to-secondary turns

ratio $n_1/n_2 = a$ is employed to present to the tube an *apparent a-c load* $R'$, while usually having a low d-c resistance. Ideally the transformer voltages and currents of Fig. 6-2*a* are related as

$$\frac{V_1}{V_2} = \frac{I_2}{I_1} = \frac{n_1}{n_2} = a$$

from which $V_1 = V_2 I_2 / I_1$. The impedance seen from the primary side is

$$Z_1 = \frac{V_1}{I_1} = \frac{V_2 I_2}{I_1^2} \frac{I_2}{I_2} = \frac{a^2 V_2}{I_2}$$

Since $V_2/I_2$ is the secondary circuit impedance, then

$$Z_1 = a^2 Z_2 \qquad (6\text{-}1)$$

is the theoretical load seen on the primary side due to any secondary load $Z_2$.

For instance, a given load of 500 ohms may have to be employed with a tube whose desirable load has been determined as 3000 ohms. A transformer to provide the proper value of $R'$ in Fig. 6-1*c* should have a ratio

$$a = \sqrt{\frac{R'}{R}} = \sqrt{\frac{3000}{500}} = \sqrt{6} = 2.45$$

Such an *ideal* transformer thus transforms magnitude but does not alter the angle of the load impedance. To be considered *ideal*, a transformer must have negligible losses, negligible magnetizing current, and negligible leakage reactance and internal capacities. While transformers can be built which approach these requirements over most of their operating frequency range, the cost will be high and many output transformers actually used depart considerably from the ideal.

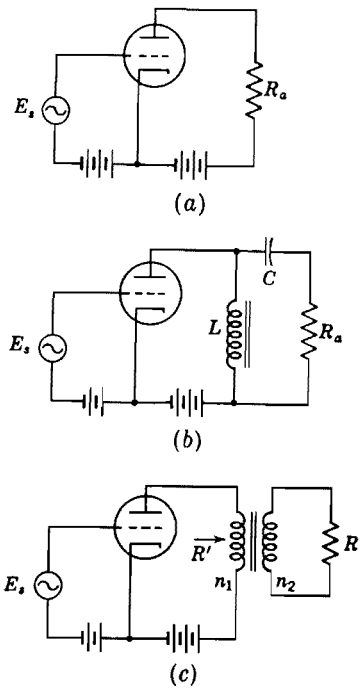



FIG. 6-1. Output circuits for power amplifiers.

In Fig. 6-2*a* the ideal transformer is illustrated. The actual equivalent circuit of an output transformer with a load $R$ is drawn in Fig. 6-2*b*, wherein $R_1$ and $R_2$ are winding resistances, $L_1$ and $L_2$ are leakage inductances, $R_c$ is an equivalent core loss resistance, $L_p$ is the primary inductance through which the magnetizing current flows, and the capacities are between the windings. In Fig. 6-2*c* all quantities have been referred to the primary circuit in accordance with Eq. (6-1). The terms $a^2R_2$

and $a^2L_2$ are the values of these secondary quantities referred to the primary side of the circuit.

The capacity $C'$ represents $C_1$ and $C_2/a^2$ plus an additional capacity representative of the effect of $C_m$ between the windings. This capacity is

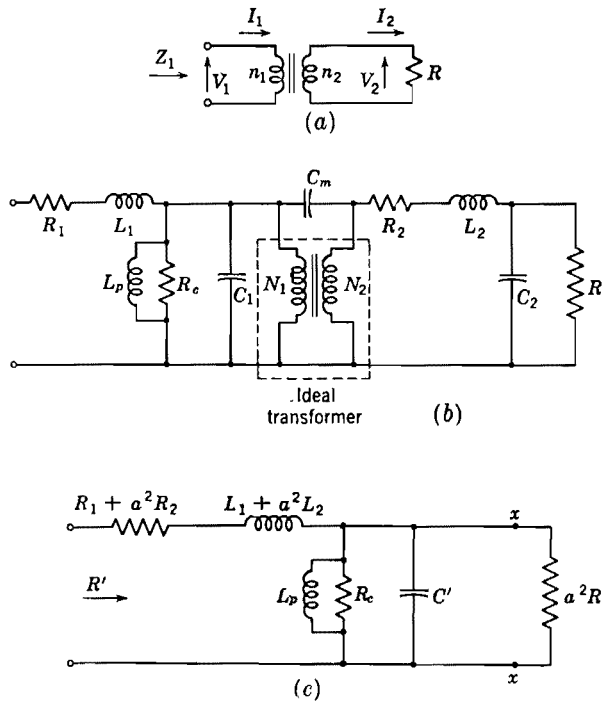



FIG. 6-2. (*a*) Ideal transformer. (*b*) Actual equivalent circuit. (*c*) Same, with all quantities referred to the primary.

connected, in effect, between points at potential $+E_1$ and $\pm E_2$ volts above reference. Since $E_1 = aE_2$, the potential across $C_m$ is

$$E_m = E_1 \pm E_2 = E_1\left(1 \pm \frac{1}{a}\right)$$

Capacitors store energy in proportion to the square of the applied potential so that

$$\text{Energy} = E_m{}^2C_m = E_1{}^2\left(1 \pm \frac{1}{a}\right)^2 C_m$$

which states that the energy actually stored in $C_m$ in Fig. 6-2*b* is the same as would be stored by the capacitor $\left(1 \pm \frac{1}{a}\right)^2 C_m$ at a potential $E_1$. Thus such a capacity may be connected across the primary or potential $E_1$ and

store the same energy or have the same circuit effect. Hence $C'$ of Fig. 6-2*c* is

$$C' = C_1 + \frac{C_2}{a^2} + \left(1 \pm \frac{1}{a}\right)^2 C_m$$

thereby placing all the capacities across the primary.

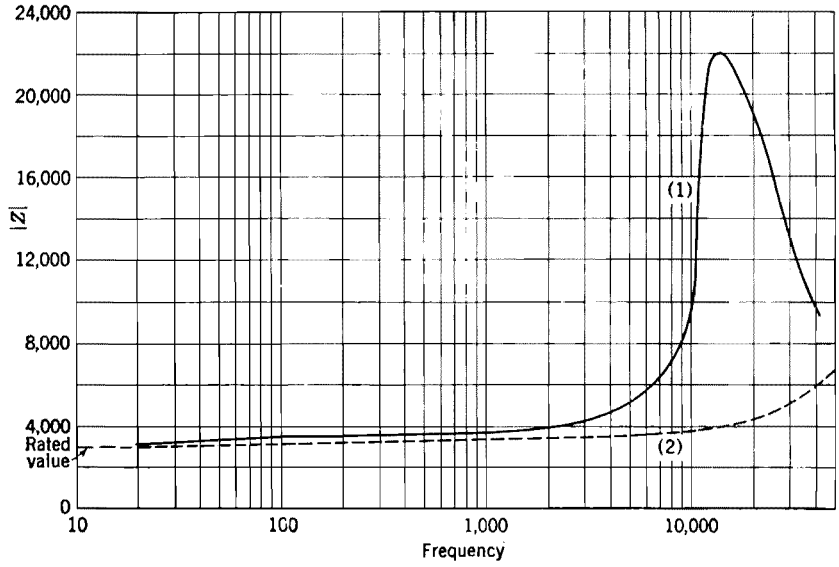
The input impedance as seen in Fig. 6-2*c* should be $R' = a^2R_2$ ideally but will differ owing to the presence of the series and shunt elements if they are appreciable in magnitude. The primary inductance $L_p$ will lower the effective impedance at low frequencies, and the presence of $L_p$ and $R_c$ will require that $a$ exceed the theoretical value at all frequencies. The resistances will also reduce the power efficiency. Series-resonant effects will be present at high frequencies owing to the leakage inductance $L_1 + a^2L_2$ and to $C'$.

Figure 6-3 illustrates the sort of departures from ideal theory which may be encountered in typical output transformers. This figure gives data on the performance at a 1-watt output level of (1) a small low-cost transformer with an 8-watt output rating and (2) a large high-cost transformer with 15-watt rating (dashed lines). It can be seen that the reflected impedances are approximately as predicted except at frequencies above 5000 cycles, where the low-cost unit has some extreme resonant effects. The effect of the leakage reactance begins to appear at about 2000 cycles for the small unit and at about 10,000 cycles for the large transformer, indicating greater care in design of the large unit in reducing leakages and capacitances.

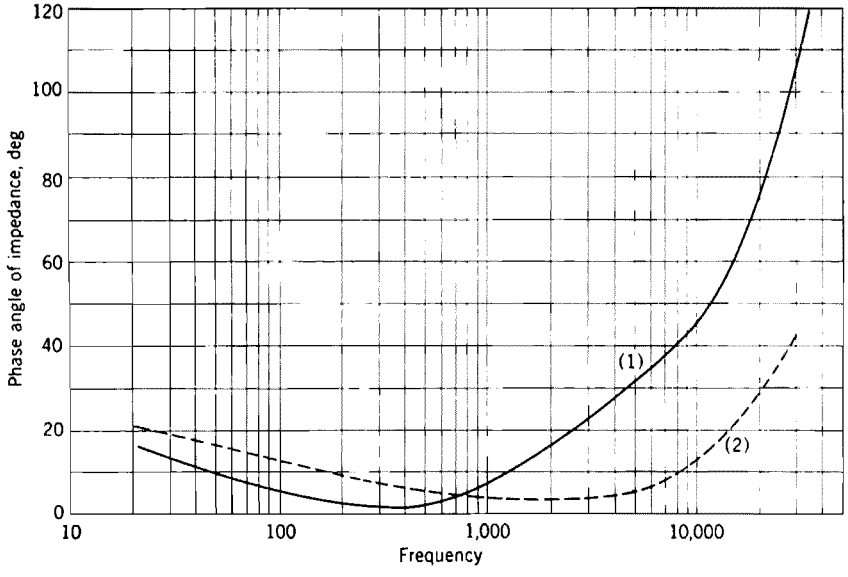
At the higher frequencies the load on the tube will not be resistive, and this may be a factor in distortion, since the load line then becomes an ellipse (see Sec. 6-3), and operation may be carried into the nonlinear portion of the tube characteristics. The power efficiency of neither transformer is particularly good, and this requires additional power from the tubes if rated transformer output is to be obtained. In the case of the small transformer the tube must supply 12 watts to place the rated 8 watts in the load.

While in amplifier design the output transformer may be frequently treated as ideal, the above data show this may be far from the truth. Transformers approaching ideal performance are available, however, but only at prices ten to thirty times those of the small unit discussed above. The difficulties caused by high-frequency resonance and phase angles will be further discussed in connection with feedback amplifiers.

The effect of the reduced transformer efficiency must be included in the load supplied by the tube. The transformer power losses can be represented by an equivalent parallel resistor $R_f$, as in Fig. 6-4. The actual load $R'$ presented to the tube is due to the parallel effect of $R_f$ and




(*a*)




(*b*)

FIG. 6-3. (*a*) Input impedance of (1) 8-watt rated low-cost transformer, rated 3000:8 ohms, operating at 1.0 watt output; (2) 15-watt high-cost output transformer, rated

$a^2R$, where

$$R_f = \frac{a^2R}{1 - \eta} \tag{6-2}$$

where $\eta$ is the transformer efficiency. The actual turns ratio required to present a given load $R'$ to the tube is then

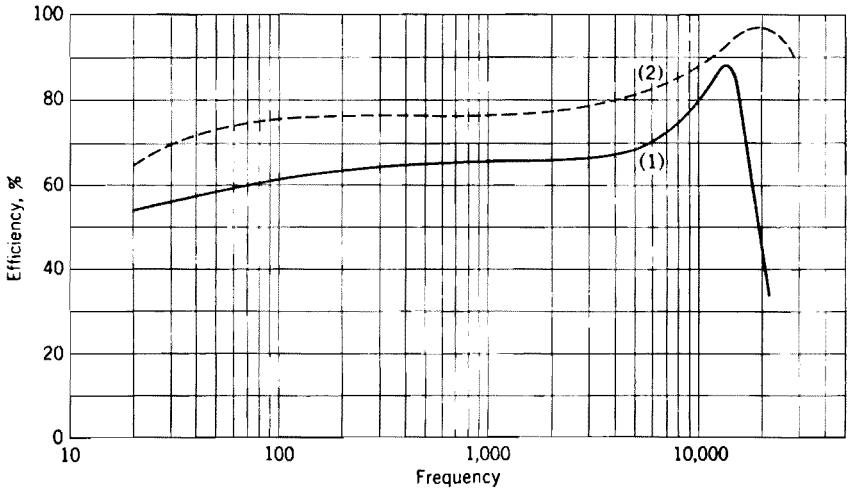
$$a = \sqrt{\frac{(2 - \eta)R'}{R}} \tag{6-3}$$

If the desired tube load is 7500 ohms = $R'$ and $R$ = 500 ohms with a transformer efficiency $\eta$ = 75 per cent then the turns ratio should be

$$a = \sqrt{\frac{(2 - 0.75)7500}{500}} = 4.33$$

whereas with an ideal transformer the ratio would be 3.87.

As with $RC$ amplifiers, square waves may be used to rapidly indicate the performance of an output transformer. Figure 6-5 shows the square-wave performance of the two transformers of Fig. 6-3, when operating at an output level of 1 watt. The distortion of the top of the low-frequency square wave indicates insufficient series time constant, or insufficient primary inductance in the transformer. The distortion indicated in Fig. 6-5*d* and *e* is due to the effects of the shunt capacity, the irregularities being caused by resonance between the leakage inductance




(*c*)

3000:600 ohms, operating at 1.0 watt output. (*b*) Phase angles of (*a*). (*c*) Efficiency curves for (*a*).

and the capacities of the transformer. The more expensive unit performs in a more satisfactory, although not perfect, manner.

**6-2. Load-line Construction.** In Fig. 6-6 are drawn the anode characteristics for a triode tube, employed in the output circuit of Fig. 6-7. Assuming $R = 200$ ohms, d-c resistance of transformer primary = 250 ohms, desired tube load $R' = 5000$ ohms, circuit operation can be graphically determined by the methods of Sec. 3-5. If the transformer is assumed ideal, then the turns ratio must be

$$a = \sqrt{\frac{5000}{200}} = 5 = \frac{n_1}{n_2}$$

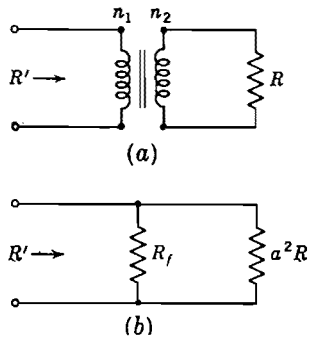



FIG. 6-4. (*a*) Ideal transformer. (*b*) Equivalent for a transformer of ratio $a$ and efficiency $\eta$.

A d-c load line can be drawn on the tube curves for $E_{bb} = 250$ volts, $R_{dc} = 250$ ohms, or with a slope of $-\frac{1}{250}$ from the $E_{bb}$ intercept. This line will pass through the point 225 volts, 100 ma, since at 100 ma there will be a 25-volt drop in the transformer. If the d-c resistance were zero, the d-c line would be vertical.

An a-c load line with a slope of $-\frac{1}{5000}$ must then be drawn. The proper slope would be that of a line having 200 volts and 40 ma as intercepts. Such a line is drawn dashed in Fig. 6-6 and serves as a construction line. The actual a-c load line can then be drawn through the $Q$ point on the $E_c = -20$ volts line, parallel to the construction line.

The use of the load line for linear-vacuum-tube analysis has already been covered in Sec. 3-5. Extension of this use into the region of nonlinear operation must be made with full consideration given to accompanying distortion, as is done in Sec. 6-7.

**6-3. Load Line for Reactive Loads.** While most tube circuits are designed to provide resistive tube loads, there will be occasions when a reactive load is presented to a tube, as in the use of an output transformer at its frequency extremes. Owing to the angle between sinusoidal current and voltage in reactive loads, the plate current will not have the same value for increasing plate voltage that it has for decreasing voltages, or the path of operation is not a straight line. In fact, the load line is in general some form of an ellipse, varying from a straight line for a resistive load to a circle for loads of pure reactance.

In a reactive-load circuit with a load of $Z/\underline{\phi}$ the load current and voltage expressions are

$$i_p = I_m \sin \omega t \tag{6-4}$$

$$e_o = E_m \sin (\omega t + \phi) \tag{6-5}$$

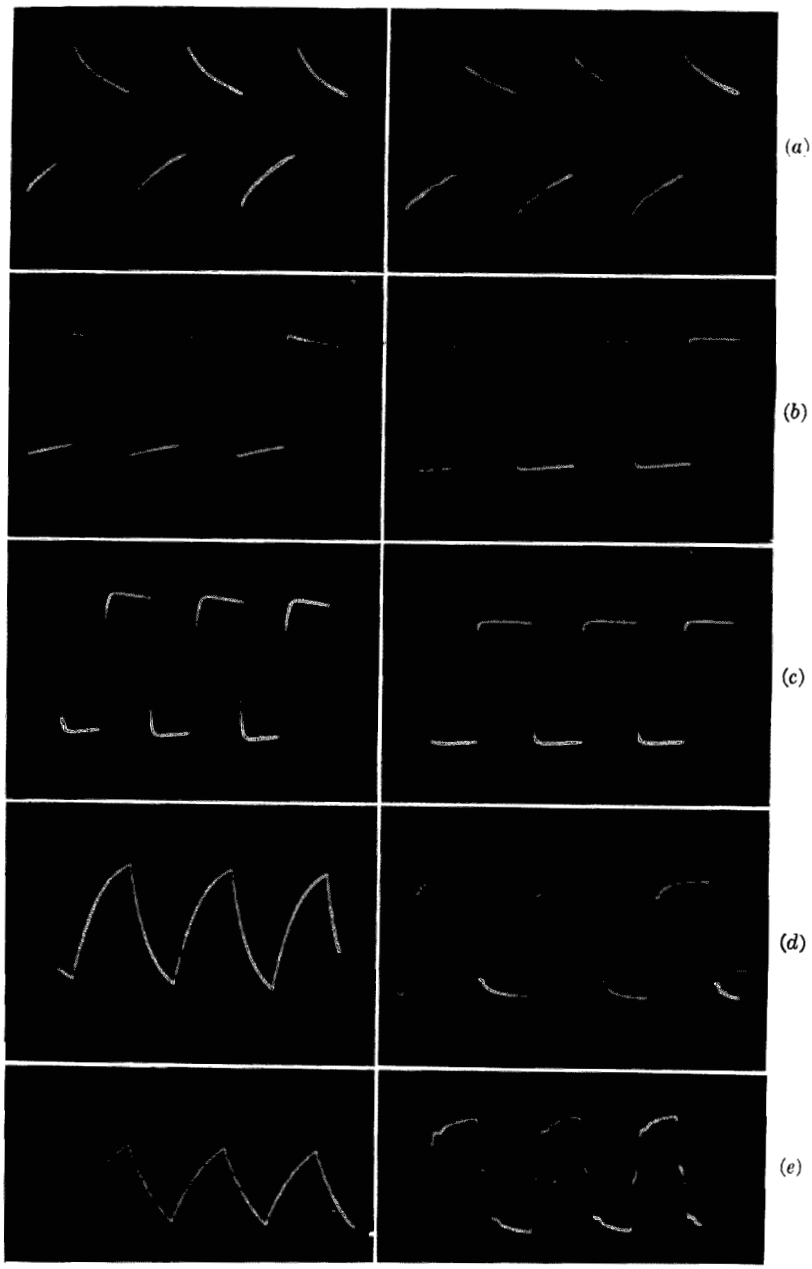



FIG. 6-5. Square waves applied to transformers (1) and (2) of Fig. 6-3. (*a*) 20 cycles; (*b*) 100 cycles; (*c*) 1000 cycles; (*d*) 10,000 cycles; (*e*) 20,000 cycles. Low-cost unit on left; large unit on right.

and these are the parametric equations of an ellipse. The elliptical load line for a 6J5 tube, operating with a load of 30,000/37° ohms, is shown in Fig. 6-8.

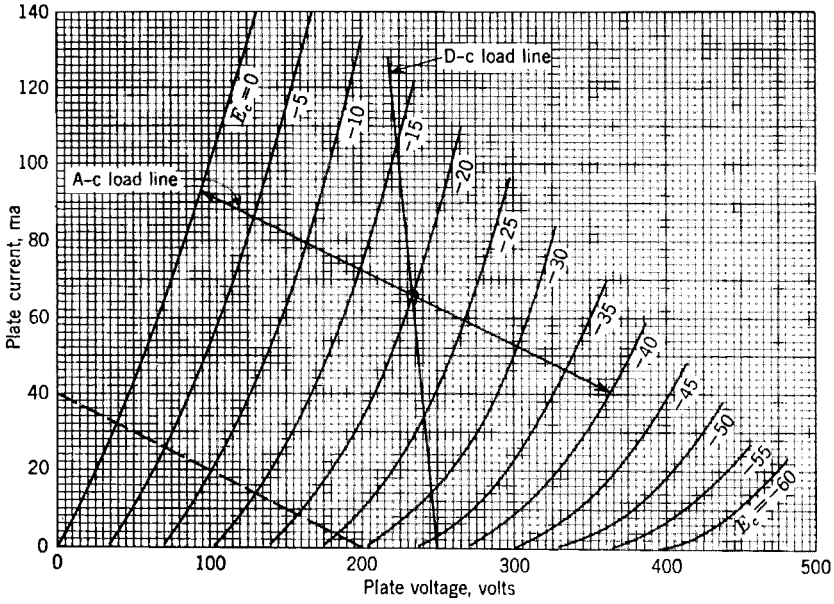



FIG. 6-6. Load-line construction.

This curve may be plotted through point-by-point calculations. For extreme swings the ellipse may pass into nonlinear regions of the tube characteristics and increased distortion result. Thus reactive loads are to be avoided for high-power-output amplifiers.

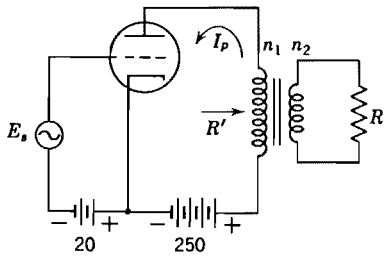



FIG. 6-7. Power-amplifier circuit.

**6-4. The Ideal Class A Amplifier.** Under an assumption of linearity which is an idealized condition for usual power amplifiers, the a-c anode current in the load circuit of Fig. 6-7 is

$$\mathbf{I}_p = \frac{\mu \mathbf{E}_g}{r_p + R'}$$

and the a-c power output in the load $R'$ is

$$P = I_p^2 R' = \left( \frac{\mu E_g}{r_p + R'} \right)^2 R'$$

The power output per input volt squared is

$$\frac{P}{E_g^2} = \mu g_m \frac{R'/r_p}{(1 + R'/r_p)^2} \tag{6-6}$$

This ratio is given the name *power sensitivity* since it measures the ability of an amplifier to produce power output per volt$^2$ of input. The power sensitivity is related to the tube chosen by the factor $\mu g_m$, and this product of tube parameters is considered a figure of merit for power tubes. Beam tubes and tetrodes have higher values of this figure than do triodes.

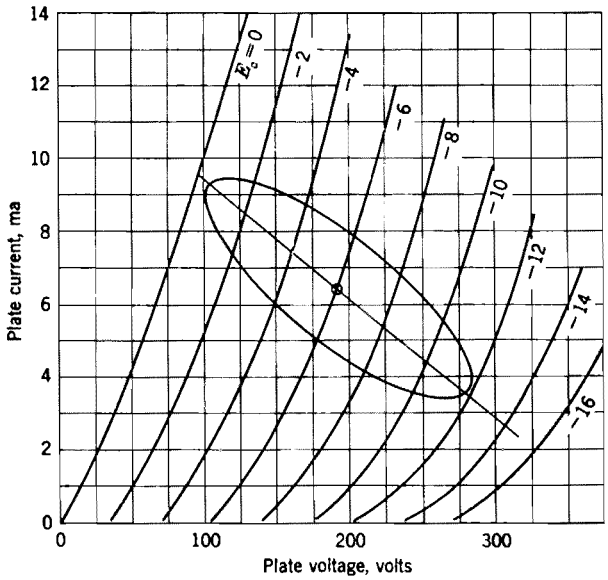



FIG. 6-8. Elliptical load line for 6J5 with load of 24,000 + $j$18,000 ohms.

Maximization of Eq. (6-6) with respect to $R'/r_p$ shows that maximum power output will occur for $R' = r_p$, the usual maximum-power-output condition. It should be noted that the actual value of power output is also proportional to $E_g^2$; thus a large grid swing is also necessary.

The above condition is correct if distortion is not considered, but if distortion is to be kept low, then a load of different value is required. For triodes this load will be larger than $r_p$, and for pentodes or beam tubes some one value of load will produce a minimum in the distortion. Selection of the proper load is arbitrary and depends on the particular tube application.

**6-5. Power Relations.** The power input supplied by the plate-voltage source must furnish the a-c signal-power output, the d-c power losses in the load or output transformer, and the power representative of the

energy required to accelerate the electrons in the tube. The latter energy appears as heat at the tube anode and is called the *plate loss*, or *plate dissipation*. Tubes must be designed to radiate or otherwise dissipate this heat. For tetrodes or pentodes additional heat losses will occur at the screen grid.

Since by the law of conservation of energy

$$\text{Power input} = \text{a-c power output} + \text{losses}$$

then

$$E_{bb}I_{bs} = I_p^2R' + \text{losses} \tag{6-7}$$

where $E_{bb}I_{bs}$ is the average power input to the circuit with signal present and $I_p^2R'$ is the output a-c power delivered to the a-c load $R'$. By definition

$$E_{bb} = E_b + I_{bs}R_{dc}$$

where $R_{dc}$ is the d-c resistance present in the $I_{bs}$ path, usually that of the transformer primary plus bias resistor if one is employed. Then, multiplying by $I_{bs}$,

$$E_{bb}I_{bs} = E_bI_{bs} + I_{bs}^2R_{dc} \tag{6-8}$$

and substitution of Eq. (6-7) in Eq. (6-8) allows the losses to be written as

$$\text{Losses} = E_bI_{bs} + I_{bs}^2R_{dc} - I_p^2R'$$

The term $I_{bs}^2R_{dc}$ is the d-c power loss in the load circuit. The remainder of the losses must occur in the tube; so the plate loss is given by

$$\text{Plate loss} = P_d = E_bI_{bs} - I_p^2R' \tag{6-9}$$

The average input to the tube is $E_bI_{bs}$; so the loss in the tube is the average input minus the a-c signal output.

In Class A the tube will operate with smaller loss with signal output than at the quiescent condition. Circuit design and tube selection must be based on the quiescent condition as the worst case, and the tube selected must be capable of dissipating $E_bI_{bs}$ watts, since the input signal may go to zero or fail.

The power efficiency of conversion of d-c input to a-c output is called the *plate-circuit efficiency* and is defined as

$$\eta_p = \frac{\text{a-c power output}}{\text{d-c power input}} \times 100\% \tag{6-10}$$

where average power is understood. The above may be written as

$$\eta_p = \frac{I_p^2R'}{E_{bb}I_{bs}} \times 100\% \tag{6-11}$$

where $I_{bs}$ is the average plate current with signal present.

Equation (6-10) may be otherwise written as

$$\text{Output} = \frac{\eta_p \times \text{loss}}{1 - \eta_p}$$

Tubes are rated in terms of allowable losses or maximum plate dissipation, and for a given rated loss the possible output is a function of plate-circuit efficiency. This shows an important difference between electron tubes and other electric machinery. In the case of tetrodes and pentodes the screen grid also has a rated dissipation. The screen loss $E_{c2}I_{c2}$ must be compared with the rating; otherwise screen damage may occur.

**6-6. Theoretical Limit on Class A Efficiency.** If the triode, tetrode, or pentode is considered ideal and linear, the characteristics might appear

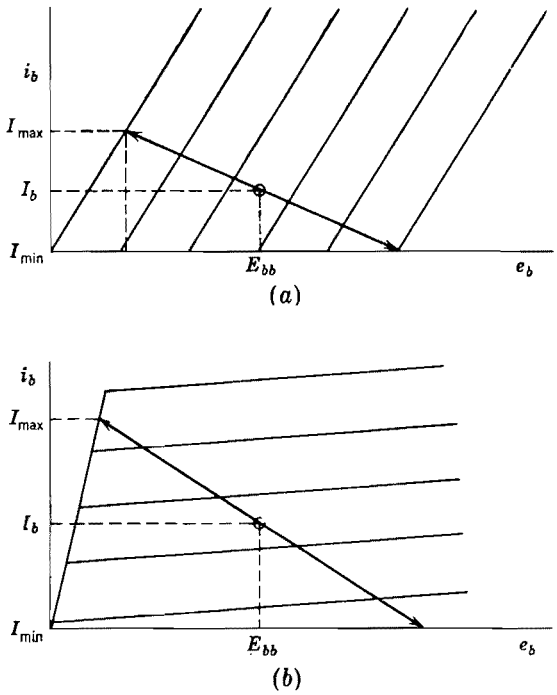



FIG. 6-9. (*a*) Theoretical linear-triode characteristics. (*b*) Same for pentode or beam tetrode.

as in Fig. 6-9. The output from the tube is $I_p^2R'$ and the input $E_{bb}I_b$. From the figure, assuming shunt feed of power,

$$\sqrt{2}\, I_p = \frac{I_{max} - I_{min}}{2} = I_m = I_b$$

The plate-circuit efficiency can then be written as

$$\text{Theoretical } \eta_p = \frac{I_p^2 R' \times 100\%}{E_{bb} I_b} = \frac{I_m^2 R' \times 100\%}{2E_{bb} I_m} = \frac{I_m R' \times 100\%}{2E_{bb}} \tag{6-12}$$

The term $I_m R'$ is the peak output-voltage swing. The greatest value this swing could have would occur for a very large load, making the load line nearly horizontal, and the peak voltage swing would approach $E_{bb}$ in value. Consequently the *maximum theoretical* plate-circuit efficiency for a shunt-fed or transformer-coupled amplifier would be

$$\text{Maximum theoretical } \eta_p = \tfrac{1}{2} \times 100\% = 50\%$$

To reach this condition, the load must be infinite, and zero output power would result. Practical efficiencies lie in the range of 5 to 25 per cent.

It can be seen that $I_m R'$ can more nearly approach $E_{bb}$ as a limit for a pentode or tetrode than for a triode. Practical operating efficiencies for pentodes or beam tubes exceed those usually possible with triodes.

For a series-fed circuit the load line has $E_{bb}$ as an intercept. The greatest value of output-voltage peak swing could then only be equal to $E_{bb}/2$. Substituting this in Eq. (6-12) gives 25 per cent as the maximum theoretical efficiency for the series-fed circuit. This lower value is the result of having to supply the d-c power losses in the load resistor.

**6-7. Determination of Amplitude Distortion.** The presence of harmonic or amplitude distortion in a power amplifier may be undesirable. The output-current or -voltage waveforms may be plotted from the load line and analyzed by conventional methods for harmonic content. However, it is possible to avoid the labor of plotting the waveforms by obtaining the magnitude of the necessary wave ordinates directly from the load line.

An equation for the dynamic grid characteristic of a tube may be written around a given $E_b$, $I_b$ operating point as a Taylor's series. In terms of variation of grid voltage from the operating point such an equation may be

$$i_b = a_o + a_1 e_g + a_2 e_g^2 + a_3 e_g^3 + a_4 e_g^4 + \cdots \tag{6-13}$$

where $a_o = I_b$, since this is the current value at the operating point for $e_g = 0$.

It can be shown by successive differentiations that the coefficients have the values

$$a_1 = \frac{\partial i_b}{\partial e_g} = g_d$$

$$a_2 = \frac{\partial^2 i_b}{\partial e_g^2} = \frac{1}{2!}\frac{\partial g_d}{\partial e_g}$$
$$a_3 = \frac{\partial^3 i_b}{\partial e_g^3} = \frac{1}{3!}\frac{\partial^2 g_d}{\partial e_g^2}$$

and so on. Thus $a_1$ is equal to the slope $g_d$ of the *dynamic curve*, and $a_2$ is proportional to the rate of change of slope. Retention of the first two terms of the series gives a linear approximation; retention of the first three terms approximates the dynamic characteristic by a straight line plus a parabola. An approximation as close as desired may be obtained by retaining additional higher-order terms.

If a grid voltage of form

$$e_g = E_m \cos \omega t$$

is applied to the tube, then

$$i_b = I_b + a_1 E_m \cos \omega t + a_2 E_m^2 \cos^2 \omega t + a_3 E_m^3 \cos^3 \omega t + a_4 E_m^4 \cos^4 \omega t + \cdots \quad (6\text{-}14)$$

By use of trigonometric identities for multiple angles, and combining the coefficients of similar terms, there results a series

$$i_b = I_b + A_o + A_1 \cos \omega t + A_2 \cos 2\omega t + A_3 \cos 3\omega t + A_4 \cos 4\omega t + \cdots \quad (6\text{-}15)$$

If the analysis leading to Eq. (6-15) is followed through in detail, it will be seen that the nonlinear terms of the series for the dynamic curve have added a d-c term $A_o$ and components of harmonic frequencies. It may be reasoned that the higher-order harmonics will be of small amplitude because the higher-order coefficients of the power series would be small for the usual near-parabolic vacuum-tube dynamic curves. Harmonics above the fourth usually may be dropped with small error.

If the amplitudes of the five constants of Eq. (6-15) are to be determined, five equations must be used, written for $i_b$ at five instants in time. With the load line drawn as in Fig. 6-10, five values of current can be selected from the load line, at five instants of time equally spaced over the positive and negative current swings. These are indicated by the numerals 1 to 5 on the load line and are at $\omega t = 0$, with a current value of $I_{max}$ when the a-c grid voltage is at positive maximum; $\omega t = \pi/3$, with a current $I_x$ when the grid voltage is at half the maximum; $\omega t = \pi/2$, with a current $I_b$ when the a-c grid voltage is zero; $\omega t = 2\pi/3$, with $I_b = I_y$ when the grid voltage is at half the negative maximum; and $\omega t = \pi$, with a current $I_{min}$ when the a-c grid voltage is at the negative maximum.

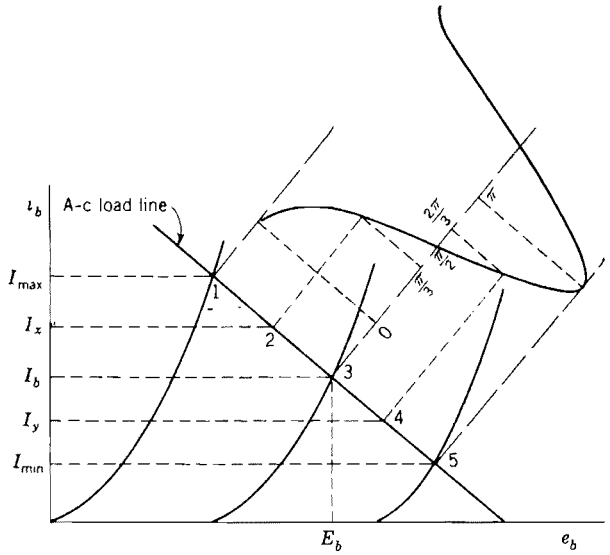



FIG. 6-10. Calculation of distortion from the load line.

Substituting these values into Eq. (6-15) gives five equations:

$$\omega t = 0: \qquad I_{max} = I_b + A_o + A_1 + A_2 + A_3 + A_4$$

$$\omega t = \frac{\pi}{3}: \qquad I_x = I_b + A_o + \frac{A_1}{2} - \frac{A_2}{2} - A_3 - \frac{A_4}{2}$$

$$\omega t = \frac{\pi}{2}: \qquad I_b = I_b + A_o - A_2 + A_4$$

$$\omega t = \frac{2\pi}{3}: \qquad I_y = I_b + A_o - \frac{A_1}{2} - \frac{A_2}{2} + A_3 - \frac{A_4}{2}$$

$$\omega t = \pi: \qquad I_{min} = I_b + A_o - A_1 + A_2 - A_3 + A_4 \tag{6-16}$$

Solution of these five equations gives the following expressions for the amplitudes of the various harmonic components:

$$A_o = \frac{I_{max} + I_{min}}{6} + \frac{I_x + I_y}{3} - I_b \tag{6-17}$$

$$A_1 = \frac{I_{max} - I_{min}}{3} + \frac{I_x - I_y}{3} \tag{6-18}$$

$$A_2 = \frac{I_{max} + I_{min}}{4} - \frac{I_b}{2} \tag{6-19}$$

$$A_3 = \frac{I_{max} - I_{min}}{6} - \frac{I_x - I_y}{3} \tag{6-20}$$

$$A_4 = \frac{I_{max} + I_{min}}{12} - \frac{I_x + I_y}{3} + \frac{I_b}{2} \tag{6-21}$$

The sum of $A_o$ and $I_b$ represents the average current $I_{bs}$ flowing when a signal is present. That is,

$$I_{bs} = \frac{I_{max} + I_{min}}{6} + \frac{I_x + I_y}{3} \tag{6-22}$$

This is a necessary value when total plate input with signal present is to be computed.

Harmonic percentage may be defined in terms of the amplitudes of the harmonic components with respect to the fundamental as

$$D_2 = \frac{A_2}{A_1} \times 100\% \qquad D_3 = \frac{A_3}{A_1} \times 100\%$$

The total harmonic distortion is the ratio of the effective value of all the harmonics to the effective value of the fundamental, or

$$D = \sqrt{D_2^2 + D_3^2 + D_4^2 + \cdots} \tag{6-23}$$

A triode usually has predominant second-harmonic distortion, and if all higher harmonics are considered small, then Eq. (6-15) becomes

$$i_b = I_b + C_o + C_1 \cos \omega t + C_2 \cos 2\omega t$$

If a three-point voltage-current-time table is used, the values of the amplitude coefficients can be found as

$$C_o = \frac{I_{max} + I_{min}}{4} - \frac{I_b}{2} \tag{6-24}$$

$$C_1 = \frac{I_{max} - I_{min}}{2} \tag{6-25}$$

$$C_2 = \frac{I_{max} + I_{min}}{4} - \frac{I_b}{2} \tag{6-26}$$

These equations permit some savings in time over Eqs. (6-17) to (6-21) if it is known that harmonics above the second are negligible. This is not a reasonable assumption for tetrodes or pentodes where the third harmonic is usually appreciable. Equations (6-17) to (6-21) must then be used.

A schedule for any number of harmonics can be set up through choice of a sufficient number of current amplitudes from the load line.

**Example.** Compute the fundamental and harmonic amplitudes for the beam tetrode of Fig. 6-11, with $E_{bb} = 250$ volts, $E_{cc} = -10$ volts, reflected transformer load $R' = 2500$ ohms, $R_{dc} = 0$, for a sinusoidal grid voltage of 10 volts peak.

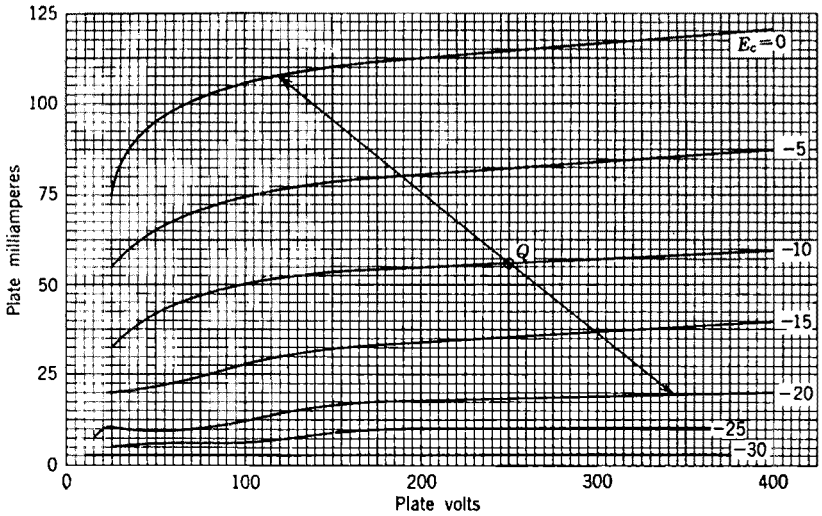



FIG. 6-11. Beam-tetrode curves and load line.

The d-c load line will be vertical for $E_{bb} = 250$ volts, and the $Q$ point is at the intersection with $E_c = -10$ volts. Readings from the load line are:

$$I_{max} = 107.5 \text{ ma} \qquad I_x = 80.0 \text{ ma}$$
$$I_{min} = 19.0 \text{ ma} \qquad I_y = 37.0 \text{ ma}$$
$$I_b = 56 \text{ ma}$$

Then

$$A_o = \frac{107.5 + 19.0}{6} + \frac{80.0 + 37.0}{3} - 56 = 4.1 \text{ ma}$$

$$A_1 = \frac{107.5 - 19.0}{3} + \frac{80.0 - 37.0}{3} = 43.8 \text{ ma}$$

$$A_2 = \frac{107.5 + 19.0}{4} - \frac{56}{2} = 3.6 \text{ ma.}$$

$$A_3 = \frac{107.5 - 19.0}{6} - \frac{80.0 - 37.0}{3} = 0.45 \text{ ma}$$

$$A_4 = \frac{107.5 + 19.0}{12} - \frac{80.0 + 37.0}{3} + \frac{56}{2} = -0.5 \text{ ma}$$

Percentage harmonic distortion then is

$$D_2 = \frac{3.6}{43.8} \times 100\% = 8.2\% \qquad D_3 = \frac{0.45}{43.8} = 1.0\%$$

$$D_4 = \frac{0.5}{43.8} = 1.1\%$$

**6-8. Intermodulation Distortion.** Waveforms having multiple sinusoidal components are frequently encountered as inputs to power amplifiers and will produce another form of distortion known as *intermodulation* if nonlinearity of the dynamic curve exists. This form of distortion produces additional frequencies in the output which are sums and differences of all frequencies present.

For simplicity assume that an input consisting of two frequencies

$$e_g = E_a \cos \omega_a t + E_b \cos \omega_b t \tag{6-27}$$

is applied to an amplifier whose dynamic transfer characteristic may be assumed as the first three terms of the series of Eq. (6-13), thus:

$$i_b = I_b + a_1 e_g + a_2 e_g^2$$

The anode current will be

$$i_b = I_b + a_1(E_a \cos \omega_a t + E_b \cos \omega_b t) + a_2(E_a \cos \omega_a t + E_b \cos \omega_b t)^2 \tag{6-28}$$

After expansion and use of

$$\cos^2 x = \frac{1}{2} + \frac{\cos 2x}{2}$$

$$\cos x \cos y = \tfrac{1}{2}[\cos (x + y) + \cos (x - y)]$$

then Eq. (6-28) may be written as

$$i_b = I_b + \frac{a_2}{2}(E_a^2 + E_b^2) + a_1 E_a \cos \omega_a t + a_1 E_b \cos \omega_b t + a_2 E_a E_b \cos (\omega_a + \omega_b)t + a_2 E_a E_b \cos (\omega_a - \omega_b)t + \frac{a_2 E_a^2}{2} \cos 2\omega_a t + \frac{a_2 E_b^2}{2} \cos 2\omega_b t \tag{6-29}$$

Thus the original input frequencies are present, as well as terms representing the sum and difference of the input frequencies, and the second harmonics of the input frequencies.

A more complex set of input frequencies, and higher-order terms in the series for the tube characteristic, would produce a more involved set of output frequencies, but the above analysis is indicative of the situation. It should be noted that all the distortion terms contain the coefficient $a_2$ and are present only because of the nonlinearity of the dynamic curve of the tube. The distortion terms are usually small but are physically present and can be found by examination of the output with a wave analyzer. They appear as background noise in a sound amplifier.

**6-9. Choice of Load for Optimum Power Output.** Proper operating conditions for power-output tubes are dependent on available anode supply voltage, anode dissipation rating, available grid input voltage, and allowable amplitude distortion in the output. The condition in which $R' = r_p$, or the load resistance equals the tube resistance, will give maximum power output, but distortion will also be high. If the distortion limit is set low, then choice of a load for a triode is arbitrary and will usually exceed $r_p$. For tetrodes or pentodes the load must be much less than $r_p$ to keep distortion within bounds.

The amount of distortion to be allowed is a subject of decision for each application. Five per cent is considered limiting if the ear is to be the

final receiver of the signal however, precision equipment will usually have somewhat less than this figure.

As a first step in selection of the desirable operating conditions for a tube, the rated dissipation line should be drawn on the plate characteristics. This line will be a hyperbola drawn for

$$E_b I_b = K \tag{6-30}$$

where $K$ is the rated plate dissipation of the tube. If the output transformer or shunt-feed inductor has negligible d-c resistance, Eq. (6-30) becomes $E_{bb} I_b = K$. The $Q$ point must lie on or below this hyperbola to ensure that the rated plate dissipation will not be exceeded under no-signal conditions. The $Q$ point is usually placed at the rated maximum $E_{bb}$ value of the selected tube.

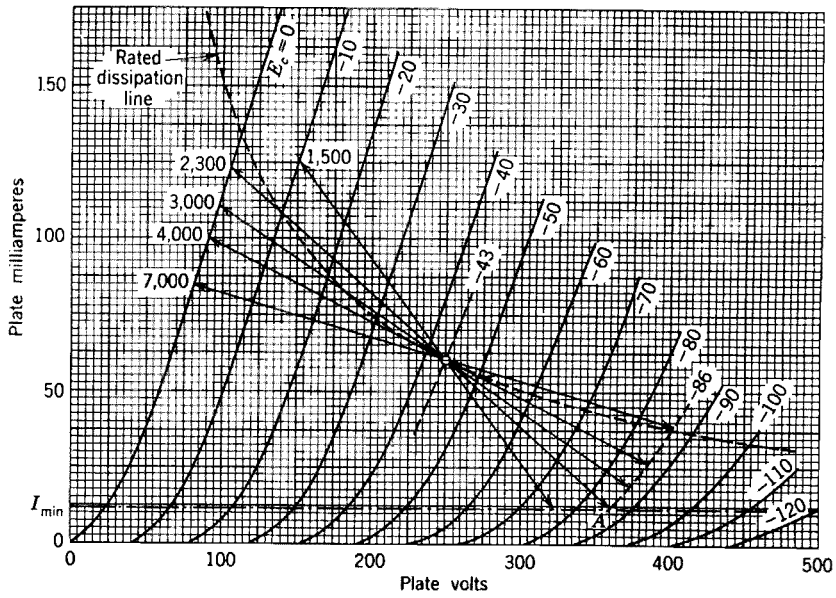
Having chosen the $Q$ point, the proper output load may be selected by experiment in the laboratory, supplementing output-power measurements with harmonic-amplitude measurements. Graphical methods can also be employed in an experimental manner to predict performance; these methods will be discussed here as applied to the Class A amplifier.

Following determination of the dissipation line and the $Q$ point, a value for $I_{min}$ must be chosen so that operation will not carry into low values of anode current, where the characteristics become seriously nonlinear. A horizontal line drawn on the plate family will serve as the $I_{min}$ limit of operation. It is assumed for Class A service that the other limit will be the $e_c = 0$ line. Both the rated dissipation line for $E_{bb} I_b = 15$ watts and a line for $I_{min} = 12$ ma are drawn in Fig. 6-12*a* for a triode tube.

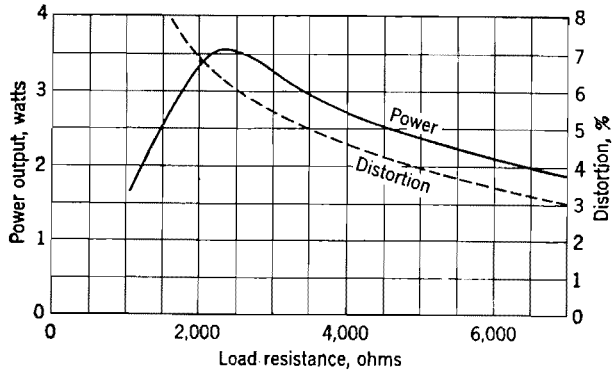
A point on the $I_{min}$ line is selected at which $e_c$ has a value equal to twice the $e_c$ value at the $Q$ point. This is indicated as point $A$ in the figure. Through this point and the $Q$ point draw a load line, carrying up to $e_c = 0$. This load will be one giving maximum or near maximum power output, but distortion may or may not be within allowable limits. Such a load line is drawn in Fig. 6-12*a* and is found to represent a load of 2300 ohms.

TABLE 10. DETERMINATION OF THE OPTIMUM LOAD FOR THE TRIODE, FIG. 6-12*a*

| Load, ohms | $I_{max}$, ma | $I_{min}$, ma | $I_x$, ma | $I_y$, ma | $A_1$ | $I_p$, ma | $P$, watts | $D_2$, % | $D_3$, % | $D_4$, % | Total $D$, % |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1500 | 126 | 12 | 93 | 33 | 58.0 | 41.0 | 2.5 | 7.8 | −1.7 | −0.9 | 8.0 |
| 2300 | 123 | 12 | 90 | 33 | 56.0 | 39.6 | 3.6 | 6.7 | −0.9 | −0.5 | 6.8 |
| 3000 | 111 | 19 | 85 | 38 | 46.7 | 33.0 | 3.3 | 5.3 | −0.7 | ..... | 5.3 |
| 4000 | 100 | 27 | 80 | 43 | 36.7 | 25.9 | 2.7 | 4.8 | ..... | −0.9 | 4.9 |
| 7000 | 85 | 38 | 72 | 49 | 23.3 | 16.5 | 1.9 | 3.0 | ..... | ..... | 3.0 |




(*a*)




(*b*)

FIG. 6-12. (*a*) Experimental load lines for determination of power output. Type 2A3 tube, 15 watts rated dissipation. (*b*) Power output and distortion vs. load.

Load lines should be drawn for other loads with values above and below that for the line first drawn, and additional lines appear in the figure. For each load the power-output and distortion components are calculated using a sinusoidal grid voltage sufficient to swing either to $e_c = 0$ at the positive peak or to $I_{min}$ at the negative peak, as determined by which of these limits is first reached. The results are tabulated in Table 10.

AU CENTRAL LIBRARY
Acc: No: 12469
Date: 13.8.85

Power output and total distortion are plotted against load resistance in Fig. 6-12*b* for the triode. An arbitrary choice of load can then be made to yield as near maximum output as possible with a desirable value of distortion. For the triode of Fig. 6-12*a* such a load might be 3500 ohms, giving distortion of about 5 per cent, mostly second-harmonic. Higher loads would reduce this distortion figure.

If distortion is of no importance, then the load may be chosen as 2300 ohms, giving maximum power output for the given conditions and amount of grid drive. It should be noted that larger power output can be had if higher anode voltage and larger grid-signal voltage are available. For instance, with the triode of this illustration, with $E_{cc} = -60$ volts, $E_{bb} = 315$ volts, and a load of 5200 ohms, the power output will be 4.1 watts with a distortion of 5 per cent.

Figure 6-13*a* illustrates the above graphical process for a beam tetrode. For such a tube, after locating the $Q$ point in a manner similar to that employed for the triode, a first load line may be drawn, and it will probably pass through the $e_c = 0$ line near the knee of the curve. Additional load lines should then be drawn to intersect the $E_c = 0$ curve at both sides of the knee. Computations of power output and distortion should be made for each load line. For the beam tetrode of Fig. 6-13*a* these results are tabulated in Table 11 and plotted in Fig. 6-13*b*.

TABLE 11. DETERMINATION OF OPTIMUM LOAD FOR THE TETRODE, FIG. 6-13*a*

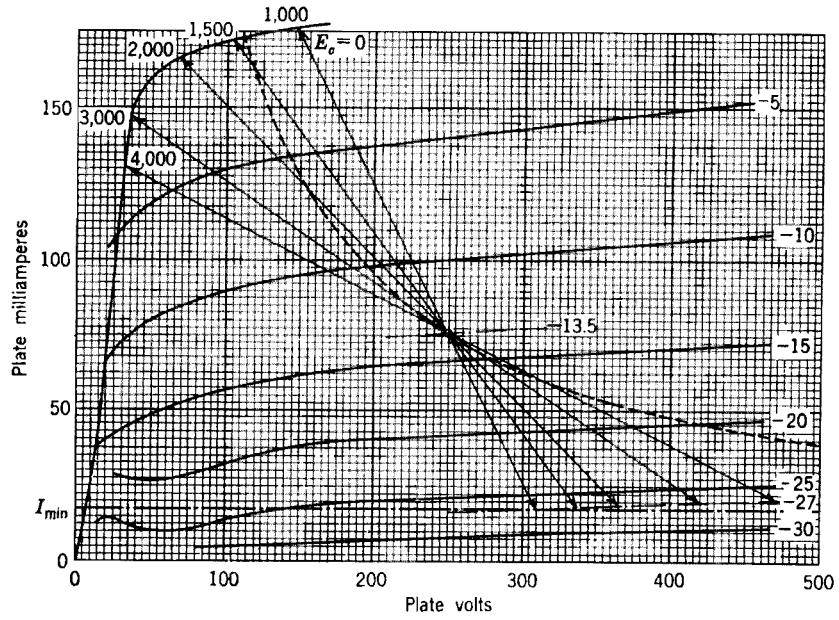
| Load, ohms | $I_{max}$, ma | $I_{min}$, ma | $I_x$, ma | $I_y$, ma | $A_1$ | $I_p$, ma | $P$, watts | $D_2$, % | $D_3$, % | $D_4$, % | Total $D$, % |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1000 | 176 | 17.5 | 125 | 42 | 80.5 | 57.0 | 3.3 | 13.0 | −1.1 | −2.0 | 13.2 |
| 1500 | 172 | 17.5 | 122 | 42 | 78.2 | 55.3 | 4.6 | 12.2 | −1.2 | −1.1 | 12 3 |
| 2000 | 166 | 18 | 120 | 43 | 75.0 | 53.0 | 5.6 | 10.7 | −1.4 | −1.3 | 10.9 |
| 3000 | 147 | 19 | 117 | 43 | 67.4 | 47.6 | 6.8 | 5.2 | −5.0 | −2.2 | 7.6 |
| 4000 | 131 | 20 | 113 | 44 | 60.0 | 42.4 | 7.2 | −0.4 | −7.5 | −2.8 | 8.0 |

The tetrode results show a minimum in the total distortion curve so that choice of a load for such a tube on the minimum-distortion basis is usually definite and for the tube of the figure would approximate 3800 ohms, with 7 watts output and 7 per cent total distortion.

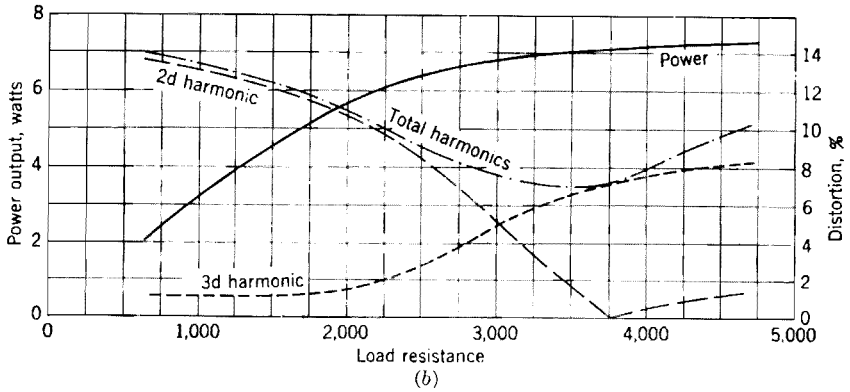
The minimum in the total distortion curve is due to the second harmonic going to zero at a particular load. This occurs for the load at which

$$I_{max} + I_{min} = 2I_b$$

or for a load line drawn in such a way that the current swings are symmetrical on peaks, or for a $Q$ point with $I_b$ value exactly halfway between $I_{max}$ and $I_{min}$. Experimental calculations indicate that this would occur




(*a*)




(*b*)

FIG. 6-13. (*a*) Experimental load lines for the 6L6 tetrode, rated 19 watts plate dissipation. (*b*) Power output vs. load.

for a load line giving $I_{max} = 132$ ma, $I_{min} = 20$ ma for the tube of Fig. 6-13*a*, or for a load of about 3800 ohms.

Choice of a load for maximum power output from a tetrode without regard to distortion is somewhat arbitrary, but little power would be gained by going above 4000 ohms for the tube under discussion, at the assigned $Q$ point. Under the assigned conditions the 4000-ohm load

produces a 440-volt peak-to-peak swing across the load, and this could increase to only 500 peak-to-peak volts as a limit for higher loads.

For low anode voltages the screen current of a tetrode is very large. Thus the use of high load resistances may overload the screen during the low-anode-voltage portion of the cycle and cause the screen to exceed its rated dissipation. For the same reason anode voltage should never be reduced to low values or removed without first removing the screen voltage. It is quite easy for the screen input to reach values which will cause the screen wires to melt.

**6-10. The Distortion Rule.** When it is known that the second harmonic is predominant, as is usual for a triode, then a faster method may be used. The load line which will give a specified amount of second-harmonic distortion for a given $Q$ point and driving voltage can be drawn by use of the so-called *distortion rule.*

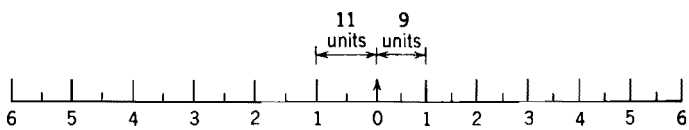



FIG. 6-14. The distortion rule.

If $X$ is the stated per cent of allowable second-harmonic distortion, then Eqs. (6-25) and (6-26) may be used to write

$$\frac{X}{100} = \frac{I_{max} + I_{min} - 2I_b}{2(I_{max} - I_{min})}$$

from which it is possible to write

$$\frac{I_{max} - I_b}{I_b - I_{min}} = \frac{1 + 0.02X}{1 - 0.02X}$$

If it is desired to limit distortion to 5 per cent maximum, which will occur at a grid drive sufficient to swing up to $e_c = 0$, then the right side above becomes $11/9$ and the plate-current swing from $I_b$ up to $I_{max}$ should be $11/9$ times the swing from $I_b$ down to $I_{min}$.

A rule may be laid out as in Fig. 6-14, wherein spaces to the left of zero are laid out 11 units long and those to the right of zero are 9 units long. If this rule is laid on the plate characteristics with 0 at the $Q$ point, and turned until the number on the left portion falling on $e_c = 0$ is the same as the number on the right portion falling on the $e_c = 2E_{cc}$ curve, then the load line is divided 11:9 and the distortion is 5 per cent.

**6-11. The Push-pull Amplifier.** Situations frequently arise where more output power is desired than can be supplied from a given tube. An obvious solution to the problem would be to parallel two or more output tubes, and this is occasionally done. However, by use of a connection

known as the *push-pull* circuit, two tubes may be so used as to provide more than twice the power output of one tube, and with much less even-harmonic distortion.

In the case of the triode the distortion is almost entirely second-harmonic, and since the push-pull amplifier cancels even-harmonic distortion, then higher power output is obtainable from the push-pull connection at a given distortion. For the pentode or beam tetrode an increase in load value will increase the power output, but with a penalty of increased harmonic distortion. A portion of this increase is due to the second harmonic, and this the push-pull circuit will eliminate, making higher power output available at a given distortion.

An additional advantage of the push-pull circuit is that the d-c components of magnetomotive force acting on the transformer core cancel and transformer-core saturation is eliminated. This simplifies output-transformer design.

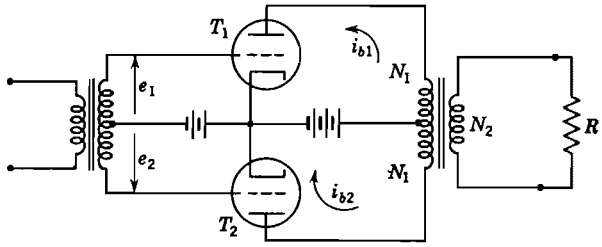



FIG. 6-15. The push-pull amplifier.

The circuit of a push-pull amplifier is illustrated in Fig. 6-15. For pentodes or beam tetrodes no change other than provision of a screen voltage supply and bypass capacitor are needed. The input grid-cathode voltages are required to have 180° phase relation, and this is shown as provided by a center-tapped transformer. Any circuit providing this phase relation over the desired frequency band is suitable, and several are discussed in Sec. 6-19. Owing to cost and limited frequency range a transformer is rarely employed unless appreciable power must be handled.

Since the grid-cathode voltages are out of phase, the plate current of one tube will reach its maximum at the instant the other tube reaches its minimum. This is illustrated in Fig. 6-16, in which the dynamic characteristic of tube 2 has been reversed to express the grid-voltage phase relation and aligned on the $E_{cc}$ value of tube 1. Both dynamic characteristics are curves, and harmonic generation in $i_{b1}$ and $i_{b2}$ is severe. The figure illustrates the even-harmonic cancellation properties of the circuit, giving an $i_{b1} - i_{b2}$ output with even harmonics canceled.

An analysis of the circuit may be made to further demonstrate its operation. If the dynamic characteristics of the two tubes are assumed

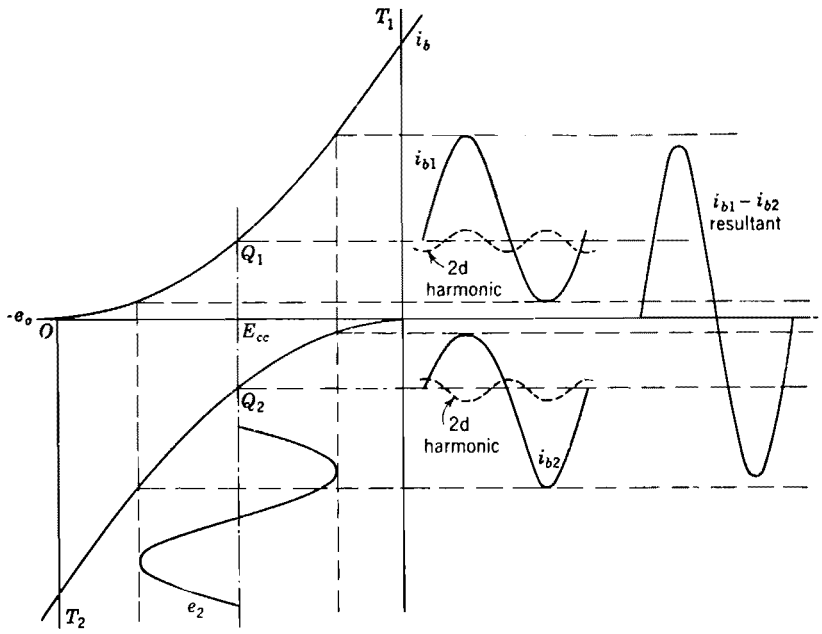



FIG. 6-16. Harmonic cancellation in the push-pull amplifier.

identical, then, from Sec. 6-7, the plate current in either tube may be written as

$$i_b = I_b + a_1 e_g + a_2 e_g^2 + a_3 e_g^3 + a_4 e_g^4 + \cdots \tag{6-31}$$

If a sinusoidal voltage is applied to the primary of the input transformer of Fig. 6-15, the grid voltages will be

$$e_1 = E_2 \sin \omega t \tag{6-32}$$

$$e_2 = E_2 \sin (\omega t + \pi) \tag{6-33}$$

The plate current of tube $T_1$ then is

$$i_{b1} = I_b + a_1 E_2 \sin \omega t + a_2 E_2^2 \sin^2 \omega t + a_3 E_2^3 \sin^3 \omega t + \cdots$$

and that of tube $T_2$ is

$$i_{b2} = I_b + a_1 E_2 \sin (\omega t + \pi) + a_2 E_2^2 \sin^2 (\omega t + \pi) + a_3 E_2^3 \sin^3 (\omega t + \pi) + \cdots$$

Use of various trigonometric identities allows reduction of the above to expressions involving multiple angles, or harmonic frequencies, as

$$i_{b1} = I_b + B_o + B_1 \sin \omega t - B_2 \cos 2\omega t + B_3 \sin 3\omega t - B_4 \cos 4\omega t + \cdots \tag{6-34}$$

$$i_{b2} = I_b + B_o + B_1 \sin (\omega t + \pi) - B_2 \cos 2(\omega t + \pi) + B_3 \sin 3(\omega t + \pi) - B_4 \cos 4(\omega t + \pi) + \cdots \tag{6-35}$$

However, Eq. (6-35) can be further simplified because

$$\sin (\omega t + \pi) = -\sin \omega t$$
$$\cos 2(\omega t + \pi) = \cos 2\omega t$$

and similarly for all odd and even harmonics; so the current in $T_2$ becomes

$$i_{b2} = I_b + B_o - B_1 \sin \omega t - B_2 \cos 2\omega t - B_3 \sin 3\omega t - B_4 \cos 4\omega t - \cdots \quad (6\text{-}36)$$

The currents $i_{b1}$ and $i_{b2}$ have directions as shown in the output-transformer primary winding and subtract in their effect on the magnetic circuit. The primary magnetomotive force is then proportional to

$$NI = N_1 i_{b1} - N_1 i_{b2}$$

The net ampere turns are then

$$NI = N_1(2B_1 \sin \omega t + 2B_3 \sin 3\omega t + 2B_5 \sin 5\omega t + \cdots) \quad (6\text{-}37)$$

If the transformer operates as an ideal device, the secondary voltage $E_s$ will be proportional to the ampere turns, or

$$E_s = k(2B_1 \sin \omega t + 2B_3 \sin 3\omega t + 2B_5 \sin 5\omega t + \cdots) \quad (6\text{-}38)$$

This shows that the push-pull connection, with tubes having matched dynamic characteristics, will eliminate all even-harmonic terms from the output. Tubes may be operated at loads and $Q$ points which give increased power output, and the resulting distortion in even harmonics will be canceled by the push-pull action. By this means, outputs approximating three times that possible from one tube can be obtained with distortion less than from one tube operating normally.

As was previously mentioned, the d-c components of plate current cancel and produce no d-c component of magnetizing force on the output-transformer core. Saturation of this core is avoided, and higher primary inductances are possible with given amounts of iron, giving better transformer performance. Power-supply ripple remaining in the $E_{bb}$ supply will also produce inphase components which cancel. Thus a push-pull amplifier may tolerate a power supply with less filtering than that required for a single-tube output, or *single-ended* amplifier, as it is often called.

**6-12. Bias Considerations in Push-pull Amplifiers.** A common cathode resistor is usually employed for grid bias in push-pull amplifiers, in place of the battery of Fig. 6-15. When connected as in Fig. 6-17 and used without a bypass capacitor $C_k$, the current flowing through $R_k$ will be $i_k = i_{b1} + i_{b2}$, and the cathode potential then is

$$e_k = R_k(2I_b + 2B_o - 2B_2 \cos 2\omega t - 2B_4 \cos 4\omega t - \cdots) \quad (6\text{-}39)$$

where $e_k$ is expressed as a rise from ground to cathode.

The even-harmonic voltages present across $R_k$ are of such phase as to produce a component of grid-cathode voltage opposing the even harmonics generated in the plate circuit. However, if the tubes are not dynamically matched, there may be odd-harmonic components remaining in $i_k$, due to inequality of the coefficients $a_3$, $a_5$, etc. These odd harmonics are of such phase as to produce a component of grid-cathode voltage which will increase the magnitude of the odd-harmonic terms. If there is a chance of the tubes not being dynamically matched, then the cathode bypass capacitor should be used. This is of especial importance for beam tetrodes and pentodes, where the odd harmonics are large and where tube mismatches are common owing to variations in screen voltage and internal construction.

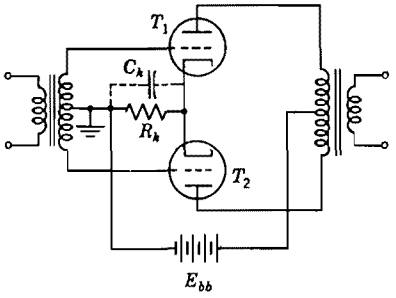


FIG. 6-17. Cathode bias in the push-pull amplifier.

Matching of the $I_b$ components, which is the usual practice, does not ensure equality of the other coefficients of the series or a dynamic match. Actually a dynamic match is rare, and it is best to utilize the bypass capacitor.

With the bypass capacitor in place the bias voltage is

$$E_{cc} = 2I_bR_k + 2B_oR_k \tag{6-40}$$

This voltage will be constant only if $B_o$ is small. Since $B_o \cong B_2$, then the $Q$ point is fixed only if individual-tube second-harmonic distortion is low. This condition is normally satisfactory in Class A amplifiers, and the bias voltage is then

$$E_{cc} = 2I_bR_k$$

In push-pull Class AB operation, second-harmonic distortion is large, and the size of grid swing must be limited if cathode bias is used, to limit distortion due to shifting of the $Q$ point. Larger grid-voltage swings and greater power output are possible if fixed bias is used.

In Class B operation, quiescent operation is near cutoff, and $I_b$ is small. The second term of Eq. (6-40) is large, and cathode bias cannot be utilized.

**6-13. Equivalent Circuit for the Push-pull Amplifier.** If a push-pull Class A amplifier is assumed as operating in linear fashion, the equivalent circuits may be drawn as in Fig. 6-18*a*. Polarities are indicated for the half cycle in which grid $G_1$ is positive. Since operation is linear, no harmonics are generated and the connection between $K$ and the transformer center tap carries zero net fundamental current. Thus for

alternating current this connection can be eliminated as in Fig. 6-18*b*, and the load when referred to the transformer primary has the value $(2N_1/N_2)^2R$.

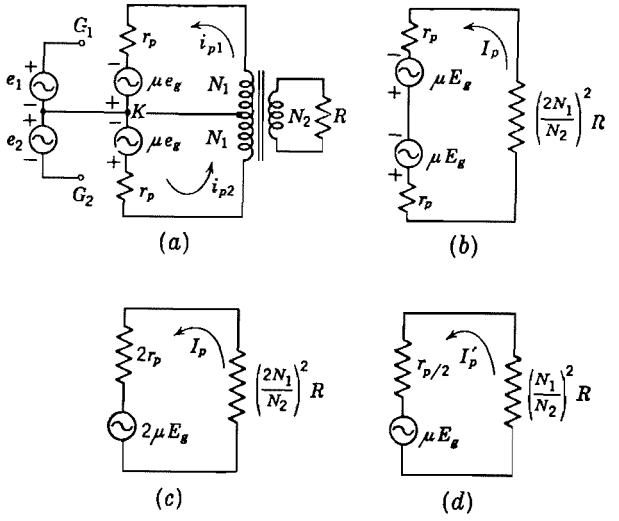



FIG. 6-18. Reduction of the equivalent circuit for the linear push-pull amplifier.

The two plate-circuit generators are then effectively in series and the circuit in Fig. 6-18*c* may be drawn. The plate current then is

$$\mathbf{I}_p = \frac{2\mu\mathbf{E}_g}{2r_p + (2N_1/N_2)^2R} \tag{6-41}$$

and the power output is

$$P = \left[\frac{2\mu E_g}{2r_p + (2N_1/N_2)^2R}\right]^2 \left(\frac{2N_1}{N_2}\right)^2 R \tag{6-42}$$

It can be easily determined that maximum power output will be obtained when

$$2r_p = \left(\frac{2N_1}{N_2}\right)^2 R \tag{6-43}$$

which states that the plate-to-plate load impedance should equal twice the plate resistance of one tube, or that of two tubes in series.

Equation (6-42) may be rewritten as

$$P = \left[\frac{\mu E_g}{r_p/2 + (N_1/N_2)^2R}\right]^2 \left(\frac{N_1}{N_2}\right)^2 R \tag{6-44}$$

and this would be supplied by a circuit as in Fig. 6-18*d*. Such a circuit would represent a single *composite tube*, equivalent in operation to that of

the complete push-pull amplifier. The reflected impedance $(N_1/N_2)^2R$ is the *per tube* load. The composite tube is that tube which, if placed in one socket of a push-pull amplifier, the other socket being empty, would give an output exactly equivalent to that of the complete amplifier. Obviously such a tube would have to operate with linear characteristics.

The reduction of distortion by the push-pull circuit allows use of a load equal to the internal resistance of the amplifier. This load value permits greater power output from the push-pull amplifier.

**6-14. Graphical Analysis Using the Composite-tube Characteristics.** A method for graphical analysis for push-pull amplifiers is available. Since the total plate currents are subtractive in the output transformer, plate characteristics may be assembled, with the currents for one tube plotted as negative values, and with its plate voltages increasing from right to left. The two curve families are adjusted on the zero-current axis so that the $E_b$ values are aligned. The construction is illustrated in Fig. 6-19 for two type 2A3 triodes operating at $E_b = 250$ volts and $E_{cc} = -50$ volts. The plate-voltage axes are adjusted so that the 250-volt line is the same for both tubes. The grid-voltage lines are marked with actual $e_c$ values, and also with signal voltages above and below the $Q$ point in parentheses.

With no signal the tubes will be at their respective $Q$ points, each at 32 ma. Since the effects of the tube currents are subtractive in the output transformer, the net output, or *composite*, *current* at the $Q$ point is zero. For any other values of anode voltage, net or composite currents can be found as the difference of the individual current magnitudes (or as the sum of the positive and negative plotted values). Points so determined as the graphical difference of the two $e_c = -50$ volts curves are found to lie on a straight line, the line being dashed in the figure. This straight line is the $e_g = 0$ curve of the composite tube representing the push-pull amplifier.

For a signal of 25 volts the upper tube will operate on its $e_c = -25$ volts line and the lower tube on its $e_c = -75$ volts line. That is, the curve for $n$ volts less than the $Q$ bias on the upper tube is combined with the curve for $n$ volts more than the $Q$ bias on the lower tube, in each case. The resultant composite curves are shown by long dashes and for matched tubes will be straight lines in the operating region.

These curves are the characteristics for the composite tube, and their linearity indicates very low distortion, as was pointed out in the preceding section.

The composite characteristics may be plotted on only one set of characteristics if desired. The curve for $e_c = -25$ volts can be combined with current values chosen from the $e_c = -75$ volts curve to yield the upper half of the $e_g = +25$ volts curve (for $Q$ at $e_c = -50$ volts). Since the

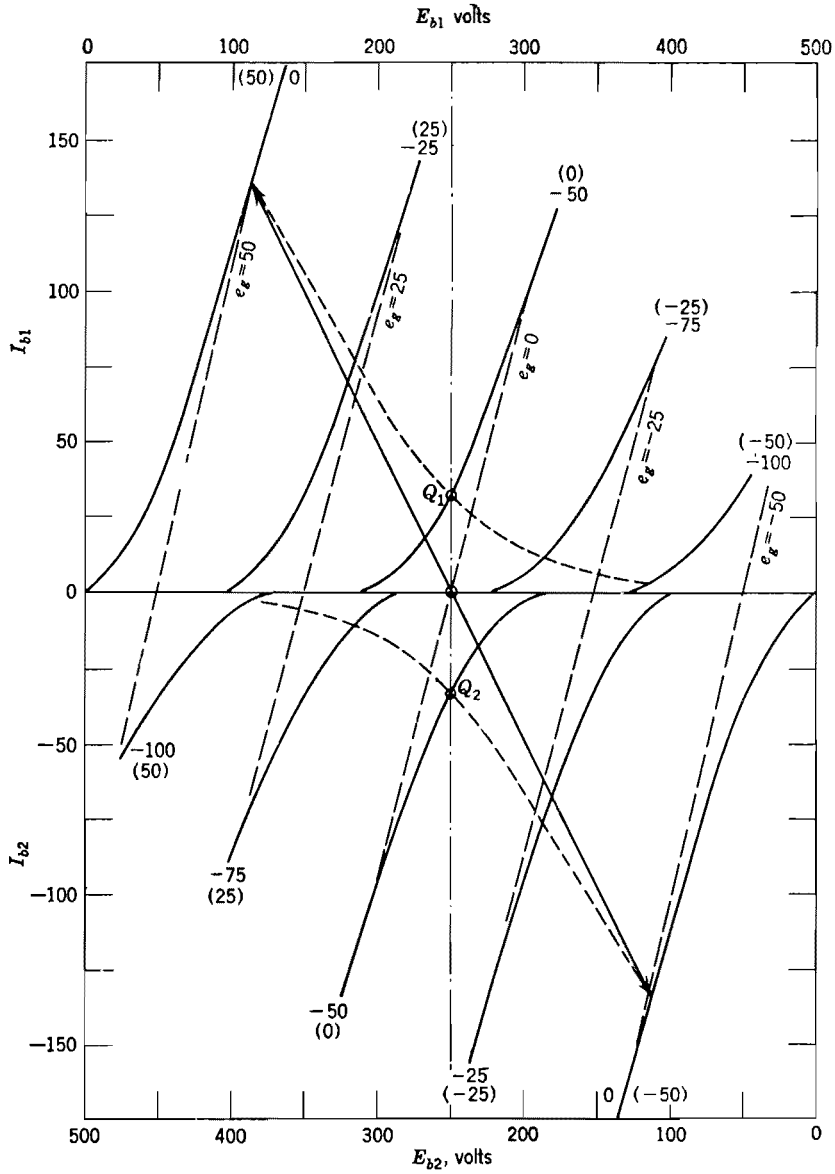



FIG. 6-19. Composite push-pull characteristics; load line for 1000 ohms.

diagram will be symmetrical about $i_b = 0$, $E_b$, the other half of the diagram is not needed.

A load line for the composite tube may be drawn on the composite characteristics. Such a line is shown in Fig. 6-19 for a composite-tube load of 1000 ohms, or a plate-to-plate load four times as great, or 4000 ohms. Previously discussed methods of calculation are available for determination of distortion and power output. The load line of the figure gives the values in the accompanying table. Comparison of this

| $I_{max}$, ma | $I_{min}$, ma | $I_x$, ma | $I_y$, ma | $A_1$ | $I_p$, ma | $P$, watts | $D_2$, % | $D_3$, % |
|---|---|---|---|---|---|---|---|---|
| 135 | −135 | 64 | −64 | 133 | 95.4 | 9.1 | 0 | 1.7 |

result with Table 10 proves the statement that power outputs up to three times that of one tube are obtainable, with low distortion, from the push-pull amplifier.

Because of the linearity of the composite characteristics, it becomes possible to choose the load almost solely on the basis of power output. Measurement of the slope of the composite curves gives the internal resistance $r_p/2$ of the composite tube. For the figure this has a value of 500 ohms. If this is matched for maximum power output, then the plate-to-plate load of the amplifier should be four times the composite-tube load, or 2000 ohms, and this value also checks that from Eq. (6-43). In all the above, $r_p$ retains its meaning as the plate resistance of one of the push-pull tubes.

The actual paths of operation for each of the individual tubes are indicated by the short-dash curves passing through each of the $Q$ points. Intersections of the load line and the composite characteristics are projected upward and downward to the actual pair of tube characteristics which generated the composite line. For example, the intersection of the load line and the +25-volt composite characteristic means that the upper tube is operating on its −25-volt curve and the lower tube on its −75-volt curve, at the common anode voltage fixed by the intersection of the load line and the composite characteristic. Thus the upper tube carries 76 ma, the lower tube 11 ma at that instant.

The operating curves are definitely nonlinear, indicating considerable distortion in each tube which is canceled by the push-pull connection. From the operating line of Fig. 6-19 the currents in milliamperes are

$$I_{max} = 135 \qquad I_x = 77$$
$$I_{min} = 3 \qquad I_y = 11$$

and the individual-tube second-harmonic distortion can be found from Eqs. (6-18) and (6-19) as 28 per cent.

Because of the high distortion present in each tube, the average value of plate current may differ considerably from the $Q$-point value, and the input and efficiency will differ from those calculated from $Q$-point readings. Using the data above, the average current with signal $I_{bs}$ may be found from Eq. (6-22) as 52.3 ma. The $Q$-point value is shown as 32 ma. The plate input per tube is then

$$P_{in} = 250 \times 0.0523 = 13.1 \text{ watts}$$

The power output was determined as 9.1 watts so that the plate-circuit efficiency is

$$\eta_p = \frac{9.1 \times 100\%}{26.2} = 34.8\%$$

Since the power output per tube was $9.1/2 = 4.55$ watts, the actual anode dissipation is very much below the rated 15 watts, but this would not always be the case. If the dissipation rating is exceeded, it is usually advisable to increase the load resistance, reducing the nonlinearity of the individual tube and making $I_{bs}$ approach $I_b$.

It should be noted that output is calculated from the composite load line, but individual tube input and tube loss are calculated from the tube operation line.

Pentodes and beam tetrodes may be handled in a similar manner. However, permissible distortion must be considered in the selection of a load, since the odd harmonics produced by these tubes are not canceled by the push-pull connection. The desirable load is a compromise between power output and distortion.

**6-15. The Class AB Push-pull Amplifier.** By raising the anode-supply voltage, keeping the plate power dissipation constant, and increasing the $Q$-point bias, the operating point and the lower set of characteristics of Fig. 6-19 move to the right. The load line then becomes longer, and larger output may be obtained. Each tube may be driven to cutoff during part of its negative cycle, but the even harmonics produced are eliminated by the push-pull connection. Such operation into cutoff of the individual tubes is called Class AB, wherein the tubes are cut off for less than 180°. The graphical analysis methods of the previous section are still applicable for quantitative information.

The subscript 1 is used, as in Class $AB_1$, if the grids are not driven positive and no grid current flows. For greater output it is possible to use still larger input voltages, and the subscript 2, as in Class $AB_2$, indicates that the grids are being driven positive on peaks and that grid current flows at some time during the positive intervals.

Amplifiers supplying the input circuit must then furnish power to the grids, and must have good voltage regulation or low impedance if the

grid-voltage waveforms are not to be distorted. The impedance of the source can be reduced if a low-$r_p$ tube is used with a step-down ratio transformer to the grid circuit. The transformer reduces the voltage by its turns ratio, but it lowers the impedance reflected into the grid circuit by the square of the turns ratio.

Class $AB_1$ and $AB_2$ amplifiers are capable of greater power output than Class A amplifiers, but distortion will be greater. The d-c component of plate current $I_o$ will also be large and will vary with signal amplitude. Plate-voltage sources should provide good voltage regulation, and fixed bias should preferably be used.

**6-16. The Class B Push-pull Amplifier.** The reasoning of the previous section may be extended to full Class B operation of a push-pull amplifier,

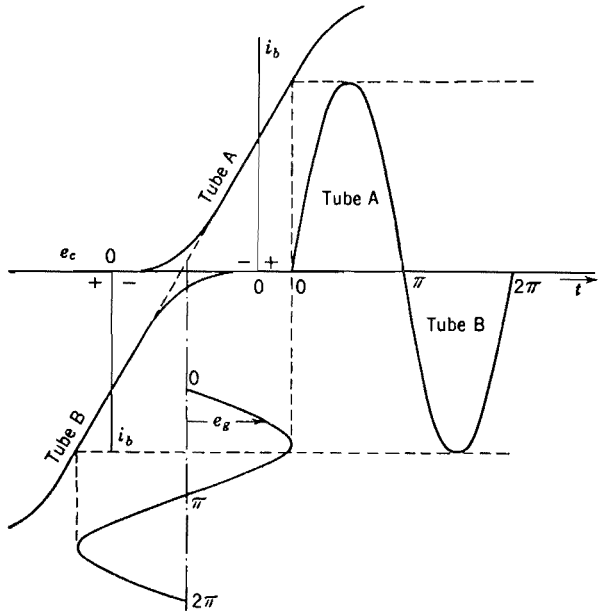



FIG. 6-20. Analysis of the Class B push-pull amplifier.

wherein the tubes are biased to cutoff at the $Q$ point, and a greater length of load line is available for the positive grid-voltage swing. The negative swing goes into cutoff, but the even-harmonic distortion is canceled by the push-pull action.

Class B push-pull operation may be studied by reference to Fig. 6-20. The dynamic grid characteristic of tube $B$ has been plotted upside down and reversed to that of tube $A$, and with the grid-voltage scales adjusted so that $E_{cc}$ is aligned for both tubes. The value of $E_{cc}$ is determined by extending the linear portion of the characteristic to the zero-current axis.

This value of bias, which is actually slightly above cutoff, is called the *extended-cutoff* voltage. Because the amplifier is push-pull, a composite grid characteristic can be drawn by adding algebraically the plate-current values for each $e_c$ value. The result will be nearly linear for matched tubes, except at high current values. The use of extended cutoff avoids a break in this curve which would occur if the characteristics were aligned at true cutoff.

It can be seen that, with a sinusoidal voltage applied out of phase to the grids of tubes $A$ and $B$, tube $A$ carries current in approximately one half cycle and tube $B$ carries current approximately in the other half cycle, the push-pull output transformer combining the outputs of the two tubes into a composite result as a good image of the applied sine wave.

The dynamic characteristic of the individual tubes is usually designed to be linear well into the positive-grid region, and increased output can be obtained by driving the grids to positive voltage values. Grid-circuit distortion must be avoided, and the driving amplifier must be capable of supplying the necessary power when the grids go positive.

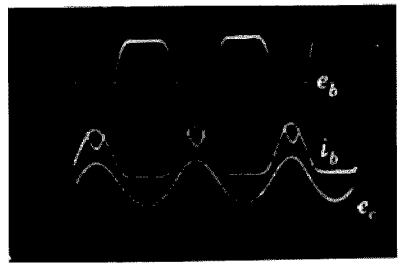



FIG. 6-21. Waveforms in a one-tube Class B amplifier.

If the tubes are biased as in Fig. 6-20, grid current will flow only near the positive peaks and this will represent a varying load on the driving amplifier. For this reason many Class B tubes are designed which operate at extended cutoff with zero-bias voltage. The signal drives one or the other grid positive during the whole of the cycle, and the grid circuit appears as a much more constant load. Driver amplifier distortion is reduced by use of these zero-bias tubes. The zero-bias tube also makes fixed-bias voltage supplies unnecessary.

Linearity of the dynamic characteristic is desired if distortion is to be low, since, with only one tube operating at a time, nonlinear distortion produced by one tube cannot be canceled by nonlinear distortion in the other tube, except near cutoff.

Figure 6-21 shows the voltage and current waveforms in one tube operating in Class B with $E_{bb} = 150$ volts. The zero axis shown applies for all waves so that it can be estimated that the grid was driven about 20 volts positive, with the minimum plate voltage about 35 volts. The fact that the cutoff is not sharp is shown by the rounded corners on the plate-current pulses.

There is little difference between extended-cutoff operation and actual cutoff when high-$\mu$ tubes are used. A *theoretical* analysis of Class B

amplifier performance is then possible if the composite dynamic characteristic is assumed linear *with the tubes biased to actual cutoff*. The output-current waveform will be sinusoidal for sinusoidal input. Each tube is *assumed* to carry half-sine-wave pulses of current as illustrated in Fig. 6-20. By use of the properties of sine waves the average value of a half sine wave gives the d-c current per tube as

$$I_{dc} = \frac{I_m}{\pi} \qquad \text{amp/tube} \tag{6-45}$$

and the total d-c plate current is twice this value, where $I_m$ is the peak current value in either tube.

Since the output current is a full sine wave, as in Fig. 6-20, the effective a-c load current is

$$I_{rms} = \frac{I_m}{\sqrt{2}} \qquad \text{amp} \tag{6-46}$$

and the total a-c power output is

$$P_o = \frac{I_m^2 R}{2} \qquad \text{watts} \tag{6-47}$$

where $R$ is the per tube load of the amplifier (one-fourth the plate-to-plate value). The d-c power input is easily found as

$$P_{dc} = \frac{2I_m E_{bb}}{\pi} \tag{6-48}$$

The plate-circuit efficiency of a Class B push-pull amplifier is then given, under the assumptions, as

$$\eta_p = \frac{(I_m^2 R/2) \times 100\%}{2I_m E_{bb}/\pi} = \frac{\pi}{4} \frac{I_m R}{E_{bb}} \times 100\%$$

The term $I_m R$ is the peak departure of plate voltage from $E_{bb}$, or is equal to $E_{bb} - E_{min}$, so that

$$\eta_p = \frac{\pi}{4}\left(1 - \frac{E_{min}}{E_{bb}}\right) \times 100\% \tag{6-49}$$

In theory the smallest possible value for $E_{min}$ is zero so that the *maximum theoretical* plate-circuit efficiency of a Class B push-pull amplifier is

$$\text{Maximum } \eta_p = \frac{\pi}{4} \times 100\% = 78.5\% \tag{6-50}$$

This figure shows a considerable improvement over the possible efficiency for Class A operation and makes possible more power output from a given tube.

From the current relations of Eqs. (6-45) and (6-46) for assumed half sinusoids of current in each tube, it is possible to calculate the a-c power output as

$$P_o = 2\left(\frac{\pi I_{dc}}{4}\right)^2 R = \frac{\pi^2 I_{dc}^2 R}{8} \tag{6-51}$$

where $I_{dc}$ is here the total current taken by the two tubes. In Class B operation the plate input is large only for large signals, and this is the reason for the high power efficiency. The plate dissipation of the tubes must be determined for the condition of peak plate loss, which does not occur at zero signal as in a Class A amplifier.

The plate loss $P_d$ is

$$P_d = \text{input} - \text{output} = E_{bb} I_{dc} - \frac{\pi^2 I_{dc}^2 R}{8}$$

where $I_{dc}$ is the value for both tubes.

Taking the derivative with respect to $I_{dc}$ and setting it to zero shows that the d-c plate current of the amplifier will have a value of $I_{dc} = 4E_{bb}/\pi^2 R$ when the maximum loss occurs. The tubes must then have a dissipation rating greater than

$$\text{Maximum } P_d = \frac{2E_{bb}^2}{\pi^2 R} \tag{6-52}$$

The power output at the time of greatest plate loss is $2E_{bb}^2/\pi^2 R$, with an efficiency of 50 per cent.

Class B amplifiers are ordinarily employed only where large power output is required and costs dictate economy of input power and tubes.

**6-17. Design of Class B Amplifiers for Large Power Output.** The use of the Class B push-pull amplifier becomes a necessity when hundreds of watts or kilowatts of low-frequency power output are required, since it is the only circuit capable of generating such powers with reasonable distortion, at an economical cost in tube investment and d-c power input.

The design of Class B amplifiers involves selection of tube type on the basis of expected power output and dissipation, specification of load resistance, and a choice of the peak driving signal $e_{g,max}$ to be employed. Power output, plate loss, efficiency, and distortion are the criteria used in selection of the optimum operating conditions.

Any tube with a sufficient plate-loss rating may be biased to near cutoff for Class B operation, but, when available, the high-$\mu$ zero-bias types are preferred. Bias to total cutoff creates a nonlinear composite dynamic curve, and this causes zero-point distortion. Extended cutoff bias ensures smooth transition from one tube to the other. Somewhat greater plate dissipation is the price paid for this lowered distortion.

For high power output in Class B it is necessary to drive the grids

considerably positive. The question of how positive the grid should be driven is indefinite and can be settled in only an approximate manner. Figure 6-22 shows the plate-current (solid line) and the grid-current curves (short dashes) for low anode voltages for the 805 tube, a typical zero-bias triode. The *diode line* indicated thereon is the plate-current value for equal instantaneous anode and grid voltages. Operation to the left of this line, or with $e_b < e_c$, is not desirable because of excessive grid current and reduced plate current.

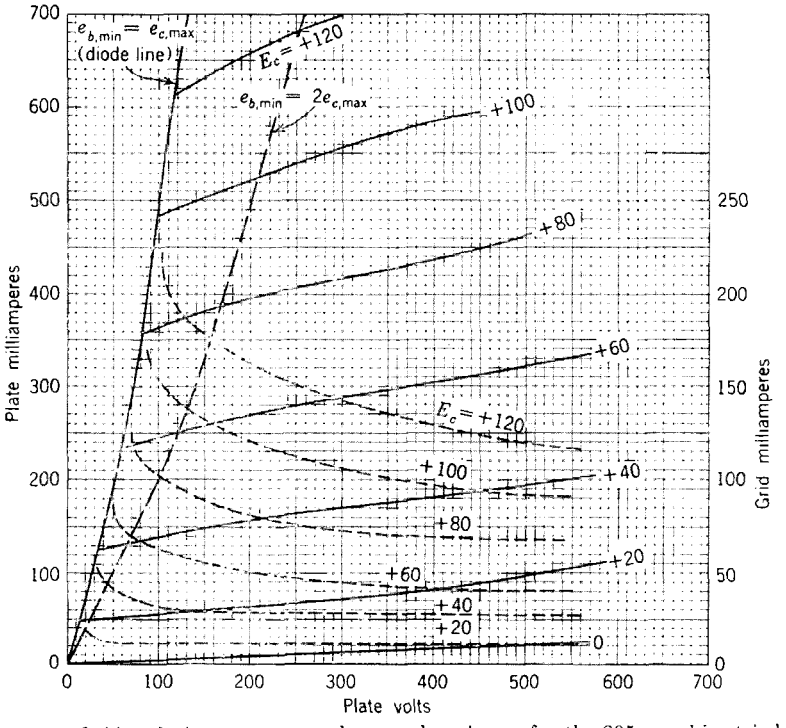



FIG. 6-22. Grid and plate currents at low anode voltages for the 805 zero-bias triode.

In Chap. 11 it will be shown that operation with a peak grid signal sufficient to make $e_{b,min} = e_{c,max}$, or to drive the tube to the diode line on positive grid voltage peaks, is optimum for the Class C amplifier. However, in Class B service it is not desirable owing to high grid-circuit distortion, since the grid current rises abruptly and the changing load on the driver amplifier increases its distortion. It is usually considered good Class B design to limit the instantaneous value of $e_{b,min}$ to values of 1½ to 2 times those of the diode line, and a limit of twice the diode-line value is drawn as the long-dashed line of Fig. 6-22. It can be seen

that such a limit avoids the sharply rising grid-current region which is reached for lower instantaneous $e_{b,min}$ values.

Using the above limit on positive grid swing, the characteristics of Fig. 6-23 may be drawn as those of one tube of a Class B push-pull pair, with the dashed line indicating $e_{c,max}$ at values of $e_b$ twice those of the diode line. The desirable load for such an amplifier may be determined by drawing several load lines for the composite tube. Starting at the $E_{bb}$ point on the zero-current axis, lines are drawn for several loads and

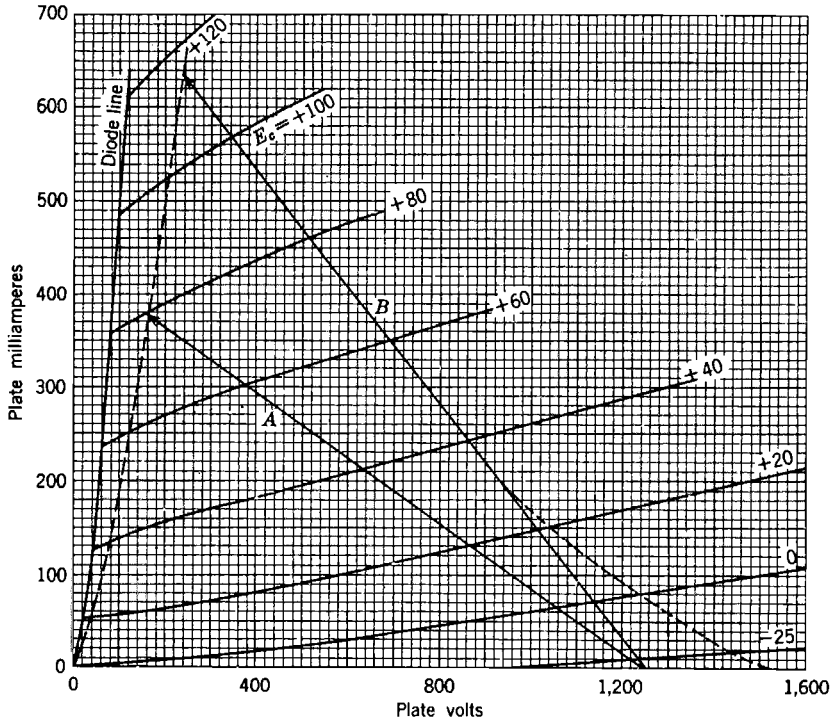



FIG. 6-23. Composite curves for one tube of a Class B amplifier using an 805 tube.

several values of positive peak grid drive $e_{c,max}$, as shown in Fig. 6-23 for the 805 tube, operating at zero grid volts and 1250 plate-supply volts. The other half of the composite characteristic is not shown, although it has been used to determine the portion of the actual tube operating characteristic passing through the actual tube $Q$ point. Above their intersection the actual tube operating line and the composite load line are identical since the other tube is cut off.

The next step in design after choice of $e_{c,max}$ is to determine which load gives a satisfactory output and at the same time holds the plate loss and average plate current within ratings. For the 805 tube, as an example,

the limiting ratings are

$E_{bb}$ = 1250 volts  Plate dissipation = 125 watts
$E_{cc}$ = 0 volts  Maximum $I_{dc}$ value = 210 ma

In the figure, load line $A$ is drawn for a composite-tube load of 2900 ohms (11,600 ohms plate to plate in the actual amplifier). Load line $B$ is drawn for a composite load of 1600 ohms (6400 ohms plate to plate). Using load line $A$ as an example, the design may be checked as follows:

Readings from the composite load line $A$ give

$I_{max}$ = 380 ma  $e_{c,max}$ = +80 volts
$I_x$ = 212 ma (at $e_c$ = +40 volts)  $e_{b,min}$ = 160 volts
$I_y$ = −212 ma  $I_b$ = 80 ma (from actual $Q$ point)
$I_{min}$ = −380 ma

Using Eq. (6-22), the average plate current for one tube with signal is

$$I_{bs} = \frac{I_{max} + I_{min}}{6} + \frac{I_x + I_y}{3}$$
$$= \frac{0.380 + 0}{6} + \frac{0.212 + 0}{3} = 0.134 \text{ amp/tube}$$

The plate input is then given by

$$P_{in} = 1250 \times 0.134 \times 2 = 336 \text{ watts total}$$

The power output can be computed by use of Eq. (6-18) for the fundamental plate-current component,

$$I_p = \frac{A_1}{\sqrt{2}} = \frac{\frac{0.380 - (-0.380)}{3} + \frac{0.212 - (-0.212)}{3}}{\sqrt{2}} = 0.278 \text{ amp}$$

$P_o = 0.278^2 \times 2900 = 224$ watts total output

Plate loss = 336 − 224 = 112 watts total, or 56 watts/tube

Since this loss figure is considerably below the rated value, it seems desirable to choose a lower load and a larger signal input voltage to increase the power output and the loss.

Choosing load line $B$ representing a composite-tube load of 1600 ohms, the operating conditions are read from the line as

$I_{max}$ = 630 ma
$I_x$ = 335 ma (at $e_c$ = +57.5 volts)
$I_y$ = −335 ma
$I_{min}$ = −630 ma
$e_{c,max}$ = +115 volts

The plate-current average per tube with signal will be

$$I_{bs} = \frac{0.630 + 0}{6} + \frac{0.335 + 0}{3} = 0.216 \text{ amp}$$

by use of Eq. (6-22). This value slightly exceeds the rating but could be reduced if necessary by a slight reduction in peak grid volts. The power input is

$$P_{in} = 1250 \times 0.216 \times 2 = 540 \text{ watts total}$$

The fundamental plate-current component is obtained from the composite load line as

$$I_p = \frac{A_1}{\sqrt{2}} = \frac{\frac{0.630 - (-0.630)}{3} + \frac{0.335 - (-0.335)}{3}}{\sqrt{2}} = 0.455 \text{ amp}$$

from which the power output for the composite tube is

$$P_o = 0.455^2 \times 1600 = 332 \text{ watts}$$
$$\text{Plate loss} = 540 - 332 = 208 \text{ watts total, or } 104 \text{ watts/tube}$$

The design will be considered satisfactory, although a slightly higher load might be used to lower the average plate current to rating. Amplifier operation will then be at

$$E_{bb} = 1250 \text{ volts} \qquad e_{c,\max} = +115 \text{ volts}$$
$$E_{cc} = 0 \qquad R = 1600 \text{ ohms}$$
$$= 6400 \text{ ohms plate to plate}$$

The plate-circuit efficiency is

$$\eta_p = \frac{332 \times 100\%}{540} = 61.2\%$$

As a further check on the design the third-harmonic distortion may be obtained for the composite tube through the use of Eq. (6-20),

$$A_3 = \frac{I_{max} - I_{min}}{6} - \frac{I_x - I_y}{3}$$
$$= \frac{0.630 - (-0.630)}{6} - \frac{0.335 - (-0.335)}{3} = -0.013$$

As computed above, $A_1 = \sqrt{2} \times 0.455 = 0.644$ peak amp. The third harmonic distortion is then

$$D_3 = \frac{A_3}{A_1} \times 100\% = \frac{0.013}{0.644} \times 100\%$$
$$= 2.0\%$$

value would usually be considered satisfactory.

It is of interest to compare these results using extended cutoff conditions with those which would be obtained from the theoretical cutoff conditions of Sec. 6-16. Then

| | *Theoretical cutoff* | *Extended cutoff* |
|---|---|---|
| $I_{dc} = \frac{I_m}{\pi}$ | = 205 ma | 216 ma |
| $P_o = \frac{I_m^2 R}{2}$ | = 317 watts | 332 watts |
| $P_{in} = \frac{2I_m E_{bb}}{\pi}$ | = 500 watts | 540 watts |
| $P_{loss}$ | = 500 − 317 = 183 watts | 208 watts |
| $\eta_p$ | $= \frac{317}{500} \times 100\% = 64\%$ | 61.2% |

The theoretical equations of the preceding section serve as quick approximations to indicate general operating values within about 10 per cent, although the plate-loss error is more serious. The disagreement is due to the theoretical assumption that the individual tube plate currents are each half sinusoids, a situation not fully met with extended-cutoff operation, as is demonstrated by the oscillograms of Fig. 6-21.

**6-18. Grid-circuit Power Requirements.** The waveform of the grid current will be distorted, owing to the nonlinearity of the grid-current curves of a tube. However, the exact waveform can be obtained by a point-by-point plot and the power input calculated by numerical integration. The instantaneous grid voltage is $e_c = E_{cc} + e_{g,max} \cos \omega t$, and with zero-bias triodes $E_{cc} = 0$. The instantaneous anode voltage will be given by

$$e_b = (E_{bb} - e_{b,min}) \cos \omega t$$

where $e_{b,min}$ is the minimum value of anode voltage during the plate-current cycle, occurring at the time of $e_{g,max}$. Values of $e_c$ and $e_b$ may be computed for evenly spaced values of $\omega t$, and the accompanying $i_c$ grid current read from the tube grid curves. For the tube and example of the preceding section these values may be read from the curves of Fig.

TABLE 12. CALCULATED GRID INPUT POWER

| | $\omega t$, deg | | | | | | |
|---|---|---|---|---|---|---|---|
| | 0 | 15 | 30 | 45 | 60 | 75 | 90 |
| cos $\omega t$ | 1.000 | 0.966 | 0.866 | 0.707 | 0.500 | 0.259 | 0.000 |
| $e_c$ | 115 | 111 | 99.6 | 81.4 | 57.5 | 29.7 | 0 |
| $e_b$ | 230 | 265 | 366 | 528 | 740 | 986 | 1250 |
| $i_c$, amp | 0.142 | 0.130 | 0.099 | 0.071 | 0.037 | 0.018 | 0.0 |
| $e_c i_c$, watts | 16.3 | 14.4 | 9.9 | 5.8 | 2.1 | 0.5 | 0.0 |
| $R_{in}$, ohms | 810 | 850 | 1010 | 1150 | 1560 | 1650 | |

6-22, after using the equations

$$e_c = 0 + 115 \cos \omega t \qquad e_b = (1250 - 230) \cos \omega t$$

The results are tabulated in Table 12, and the grid-current waveform is drawn in Fig. 6-24.

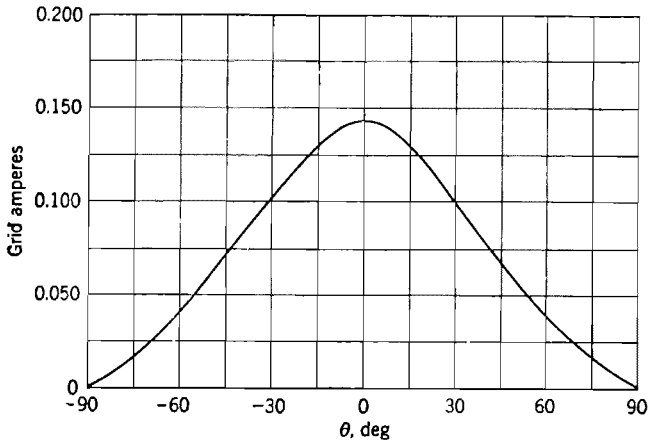



FIG. 6-24. Grid-current waveform for the 805 tube; $E_{bb}$ = 1250 volts, $e_{c,\max}$ = 115 volts, $e_{b,\min}$ = 230 volts.

The instantaneous grid input watts are also computed as $e_c i_c$, and the average power input to the grid may then be found as

$$P_g = \frac{1}{2\pi} \int_{-\pi}^{\pi} e_c i_c \, d\theta \tag{6-53}$$

The integral may be evaluated by use of the trapezoidal rule wherein, if the interval between 0 and $2\pi$ is divided into $n$ intervals of uniform length $\Delta\theta$, then

$$\int_{-\pi}^{\pi} y \, d\theta = \Delta\theta(y_o + 2y_1 + 2y_2 + \cdots + 2y_{n-1} + y_n) \tag{6-54}$$

where $y_o$ is taken for $\theta = 0$ and $y_n = 0$, cosine symmetry being assumed.

For the particular example under consideration with 15° intervals or $\Delta\theta = \pi/12$, there results

$$P_g = \frac{1}{2\pi} \frac{\pi}{12} (16.3 + 28.8 + 19.8 + \cdots + 1.0 + 0) \tag{6-55}$$

using the values of Table 12, and the average grid power input is found to be 3.4 watts per tube, or 6.8 watts for both tubes.

The driver amplifier must, however, deliver a peak of 16.3 watts, and a d rule of thumb calls for its average power rating to be half of this

peak value, even though it is called on for only 6.8 watts average. This ensures better waveform of the input voltage to the grids.

The grid circuits represent a load of about 1000 ohms average to the driver circuit, which means that a step-down transformer should be used to match the driver plate-load circuits to the Class B grid circuit. Since this load is not steady, it is usually desirable to operate the driver with a higher than normal load, to reduce distortion with Class A triode driver amplifiers. Theoretically, if the driver amplifier employed a pair of 2A3 tubes in push-pull, capable of 9 to 10 watts output at a load of 2000 ohms, then the Class B input transformer should have a ratio of $a = \sqrt{2000/1000} = 1.4:1$ step-down to the grids. Practical applications may employ a ratio of about 1.7:1 to raise the driver output-circuit load.

**6-19. Signal Sources for Push-pull Amplifiers.** Two equal voltages, in phase opposition, are required as input for push-pull amplifiers. High-power Class B amplifiers are usually driven by lower-power Class A

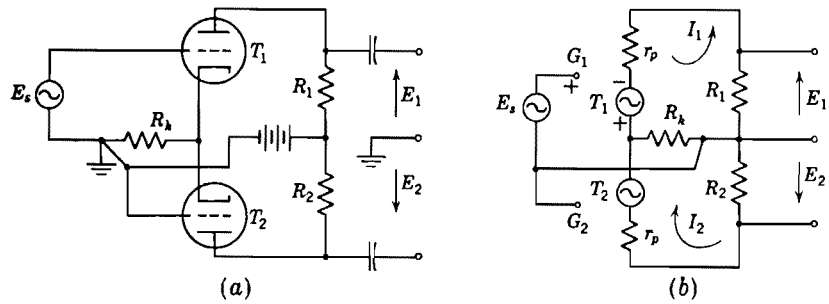



FIG. 6-25. Cathode-coupled phase inverter and the equivalent circuit.

circuits through center-tapped transformers to give the phase opposition, to provide impedance matching, and to handle the required power efficiently. Class A push-pull amplifiers, which require no exciting power, are usually supplied with input voltages from special circuits known as *phase inverters*.

A phase inverter is required to change a single grounded potential into two equal voltages, with a 180° phase relation between them, at all frequencies of interest. Unfortunately the simple circuits usually employed do not completely fulfill this specification, owing to the need for special resistor adjustment and to the consideration of capacity effects at the higher frequencies. While suitable for much radio equipment, the balance achieved may not be equally satisfactory for all applications and at all frequencies.

One simple circuit is illustrated in Fig. 6-25 and is known as a *cathode-coupled* phase inverter. The input signal $\mathbf{E}_s$ is supplied to the

circuit of $T_1$, the second tube $T_2$ obtaining its input signal from the common cathode resistor $R_k$. The grid voltages can be written as

$$\mathbf{E}_{g1} = \mathbf{E}_s - (\mathbf{I}_1 + \mathbf{I}_2)R_k \tag{6-56}$$

$$\mathbf{E}_{g2} = -(\mathbf{I}_1 + \mathbf{I}_2)R_k \tag{6-57}$$

The two plate-circuit equations are

$$\mu\mathbf{E}_{g1} = \mathbf{I}_1(R_1 + r_p + R_k) + \mathbf{I}_2R_k \tag{6-58}$$

$$\mu\mathbf{E}_{g2} = \mathbf{I}_2(R_2 + r_p + R_k) + \mathbf{I}_1R_k \tag{6-59}$$

Tubes $T_1$ and $T_2$ are usually halves of a double triode and may be assumed to have identical parameters. Because of the unbalance of grid voltages, the output voltages will not be equal in magnitude unless the circuit is properly designed. To study this point further, the grid voltages may be substituted into Eqs. (6-58) and (6-59), leading to

$$\mathbf{I}_1 = \frac{\mu\mathbf{E}_s[R_2 + r_p + (\mu + 1)R_k]}{[(R_1 + r_p) + (R_2 + r_p)][1 + (\mu + 1)R_k]} \tag{6-60}$$

$$\mathbf{I}_2 = \frac{-\mu(\mu + 1)\mathbf{E}_sR_k}{[(R_1 + r_p) + (R_2 + r_p)][1 + (\mu + 1)R_k]} \tag{6-61}$$

Since the output voltages are $\mathbf{E}_1 = -\mathbf{I}_1R_1$ and $\mathbf{E}_2 = -\mathbf{I}_2R_2$, the ratio of the two output-voltage magnitudes will be

$$\left|\frac{E_1}{E_2}\right| = \frac{R_1[R_2 + r_p + (\mu + 1)R_k]}{(\mu + 1)R_kR_2}$$

As ordinarily used with equal load resistors $R_1 = R_2 = R_a$, the output voltages will be

$$\left|\frac{E_1}{E_2}\right| = 1 + \frac{R_a + r_p}{(\mu + 1)R_k} \tag{6-62}$$

and the output voltages will be in balance only for $R_k$ infinite in value. Actually it is not difficult to make $(\mu + 1)R_k \gg R_a + r_p$, and reasonable balance is achieved. For the 6SN7 or 12AU7 triodes, commonly used in this application, with $\mu = 20$ and $r_p = 7700$ ohms, the voltages will only be 5 per cent out of balance if $R_1 = R_2 = R_k = 50{,}000$ ohms. Should exactly balanced voltages be required, then the load resistors must meet the requirement

$$\frac{R_2}{R_1} = \frac{1 + r_p/(\mu + 1)R_k}{1 - R_1/(\mu + 1)R_k} \tag{6-63}$$

Another form of phase inverter is illustrated in Fig. 6-26. The tubes are assumed identical, with $T_1$ acting as an ordinary voltage amplifier. A portion of its output voltage is selected and again amplified, introducing an additional 180° phase shift. The two output voltages then have the desired out-of-phase relation. Ordinarily $R_1 + R_2 = R_3$, and if $R_1$ and

$R_2$ are so chosen that

$$\frac{R_2}{R_1 + R_2} = \frac{1}{A_2}$$

where $A_2$ is the gain of the second tube $T_2$, then the output voltage $E_2$ will equal $E_1$, as

$$E_1 \frac{R_2}{R_1 + R_2} A_2 = E_2$$

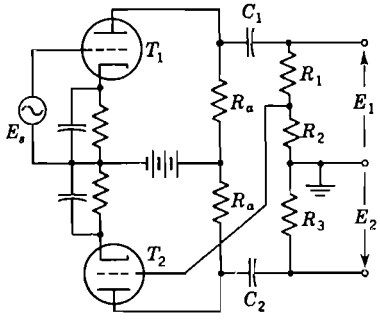



FIG. 6-26. Common phase-inverting circuit.

The success of the circuit operation is dependent on a knowledge of the gain $A_2$ and constancy of this gain throughout tube life. Some unbalance of the two voltages is to be expected, but because of its high gain and simplicity the circuit is widely used.

Capacitors $C_1$ and $C_2$ will introduce some phase shift at low frequencies, and the 180° relation will not be maintained, since the output $\mathbf{E}_1$ is affected by only one capacitor-resistor combination, whereas output $\mathbf{E}_2$ has a phase angle affected by two such circuits.

A third inverter circuit is drawn in Fig. 6-27, and this is usually referred to as a split-load inverter, since the load is split between anode and

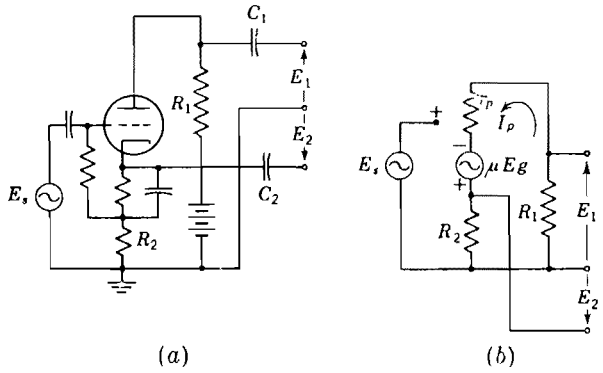



FIG. 6-27. The split load or cathode-anode form of phase inverter and its equivalent circuit.

cathode circuits. The circuit employs only one tube and the same current flows through both load resistors, so that the circuit is balanced and independent of changes in tube characteristics. The gain will be small, however, because of the relation of the circuit to the cathode follower, and this is the chief disadvantage

The circuit equations may be formed as

$$\mathbf{E}_g = \mathbf{E}_s - \mathbf{I}_p R_2 \tag{6-64}$$

$$\mu \mathbf{E}_g = \mathbf{I}_p(r_p + R_1 + R_2) \tag{6-65}$$

and if $R_1 = R_2 = R$, the gain to either output may be calculated as

$$\mathbf{A} = \frac{\mathbf{E}_1}{\mathbf{E}_s} = \frac{\mu R/(\mu + 2)}{r_p/(\mu + 2) + R}$$

If $\mu \gg 2$, this reduces to

$$\mathbf{A} = \frac{1}{1 + 1/g_m R} \tag{6-66}$$

and the gain is less than unity to either output.

This circuit is stable and constant in gain. Hum may be introduced because of the cathode operating at an appreciable potential above ground, as discussed in Chap. 4, and this limits the values of $R_1$ and $R_2$ to 25,000 to 50,000 ohms.

In all phase inverters, if precise balance is to be obtained, it is necessary to select and match resistors and capacitors in the balanced output circuits. Usual tolerances of 10 to 20 per cent introduce considerable phase and magnitude unbalance.

## PROBLEMS

(Characteristics of tube types mentioned appear in the usual tube handbooks.)

**6-1.** A certain triode has a plate family expressed by

$$i_b = 0.0012(3.5e_c + e_b) \quad \text{amp}$$

Calculate the power sensitivity, and for Class A operation at $E_{bb} = 250$ volts, $E_{cc} = -40$ volts find the power output, efficiency, and d-c power input for a signal of 40 volts peak, with a load of 2000 ohms.

**6-2.** A 2A3 triode operates in the circuit of Fig. 6-1*b* with $E_{bb} = 250$ volts, $I_b = 60$ ma. If the inductor $L$ has a very high inductance and d-c resistance of 400 ohms, draw a load line for $R_a = 2000$ ohms, neglecting the reactance of $C$. For $E_s = 30$ volts rms compute the power output, plate loss, and plate-circuit efficiency.

**6-3.** A 6V6 tetrode is operated at $E_{c2} = E_{bb} = 250$ volts, $E_{c1} = -12$ volts, with a peak sinusoidal grid signal of 10 volts. If the tube is connected in the circuit of Fig. 6-1*a*, find the output a-c power, the plate-circuit efficiency, the plate loss, and the total plate input, if $R_a = 2000$ ohms.

**6-4.** Repeat Prob. 6-3 if the tube is operated in the circuit of Fig. 6-1*b*, with $R_a = 2000$ ohms, d-c resistance of $L = 250$ ohms, inductance of $L$ very large, and reactance of $C$ negligible.

**6-5.** A Class A power amplifier using a 6V6 tetrode is coupled to a load of 500 ohms by an ideal transformer with turns ratio $a = 3$. If $E_{bb} = 300$ volts, $E_{c2} = 250$ volts, $E_{cc} = -12.5$ volts, and the sinusoidal grid signal has a 12.5-volt peak, find (*a*) output power, (*b*) plate dissipation, (*c*) plate-circuit efficiency, (*d*) total harmonic distortion (second, third, fourth harmonics).

**6-6.** Find the power output, plate-circuit efficiency, and per cent second and third harmonics for a 6L6 tetrode for loads of 2500, 6000, and 10,000 ohms coupled into the plate circuit through an ideal 1:1 transformer. The $Q$ point is $E_{bb} = 300$ volts, $E_{c2} = 250$ volts, $E_{c1} = -15$ volts, with $I_{b,\min} = 10$ ma.

**6-7.** A Class A power amplifier using a 6V6 tetrode is coupled to a load of 500 ohms by a transformer having a turns ratio $a = 3$ and power efficiency of 78 per cent. If $E_{bb} = 300$ volts, $E_{c2} = 250$ volts, $E_{c1} = -10$ volts, determine the plate dissipation, and output power in the load in the transformer secondary circuit, if $E_s = 7$ volts rms.

**6-8.** For the amplifier of Prob. 6-7, choose the transformer ratio giving the minimum total harmonic distortion (second and third) by plotting distortion and power output against turns ratio values $a = 1.5$, 2.0, 2.5, 3.4, and 5. The transformer efficiency is constant at 78 per cent.

**6-9.** A 6L6 beam tetrode is operated at $E_b = E_{c2} = 250$ volts, $E_{c1} = -15$ volts and is coupled to a load of 16 ohms through a transformer of negligible losses. Find the transformer turns ratio which will reduce the second-harmonic distortion to zero, for a sinusoidal signal of 15 volts peak. Compute also the fundamental power output, per cent third-harmonic distortion, total d-c plate current, and the anode loss with and without signal.

**6-10.** A 6K6 pentode is operated from a transformer having secondary taps giving ratios of $a = 10$, 15, 22, 30, with a secondary load of 16 ohms. $E_{bb} = E_{c2} = 250$ volts, and $E_{c1} = -17.5$ volts. Neglecting transformer losses and with $I_{b,\min} = 5$ ma, determine power output and total second- and third-harmonic distortion for the transformer tap giving greatest power output.

**6-11.** Two 2A3 triodes are operated in Class A push-pull at $E_b = 300$ volts, $E_{cc} = -50$ volts. Find the value of load resistance giving greatest power output, irrespective of distortion, and compute the no-signal plate loss and efficiency, for a sinusoidal signal of 50 volts peak to each tube grid.

**6-12.** Two 6V6 tetrodes are operated in push-pull in Class AB with $E_b = 300$ volts, $E_{c2} = 250$ volts, $E_{c1} = -20$ volts.

(*a*) Draw the composite characteristics.

(*b*) Determine the plate resistance of the composite tube.

(*c*) For a plate-to-plate load of 8000 ohms determine the power output, efficiency, and plate dissipation for an input signal of 14.1 rms volts, grid to ground.

**6-13.** A 6K6, with a desired load of 10,000 ohms, is to be transformer-coupled to a load of 200 ohms resistance. The transformer is expected to have an efficiency of 82 per cent, and $E_b = E_{c2} = 250$ volts, $E_{c1} = -15$ volts, $I_{b,\min} = 5$ ma, and $E_s = 15$ volts peak. Find (*a*) transformer ratio needed, (*b*) fundamental power output in the 200-ohm load, (*c*) total harmonic distortion in per cent.

**6-14.** Using a 2A3 tube with an output transformer of ratio 24:1 and secondary load of 8 ohms, with a transformer efficiency of 88 per cent, and assuming the grid is never positive, $E_b = 275$ volts, and $i_b$ is always greater than 10 ma, find (*a*) $Q$ point for rated plate dissipation of 15 watts, (*b*) maximum possible power output in the load under the assumptions, (*c*) value of $E_g$ required for (*b*), (*d*) fundamental plate current, (*e*) total harmonic distortion.

**6-15.** The two triodes of a 6AS7G tube are operated in push-pull at $E_b = 135$ volts, $E_{cc} = -60$ volts.

(*a*) Draw the composite characteristics, and determine the plate-to-plate load for maximum power output.

(*b*) Calculate the Class A power output for this load.

(*c*) Construct the path of operation for one triode.

**6-16.** Two 6L6 tubes are used in push-pull with $E_{bb} = 300$ volts, $E_{c2} = 250$ volts,

and $E_{c1} = -16$ volts, under Class A conditions. Plot as a function of plate-to-plate load (*a*) power output, (*b*) third-harmonic distortion, (*c*) plate-circuit efficiency. Try loads from 1000 to 10,000 ohms, plate to plate.

**6-17.** A 6SN7 and two 6L6 tubes are available, with a 250-volt power supply. For an operating frequency range of 50 to 7500 cycles, and with a push-pull output transformer of 85 per cent efficiency, design a complete circuit to supply 7 watts to an 8-ohm load with minimum distortion. The input signal available is 1.5 volts, ground reference. Specify transformer ratio, all resistors and capacitors, and all resistor wattage ratings.

**6-18.** Two 805 zero-bias Class B tubes are used in push-pull at $E_b = 1250$ volts, $E_{cc}$ = cutoff volts, with a plate-to-plate load of 7000 ohms. For a steady sinusoidal input signal a d-c milliammeter in the common anode lead reads 400 ma. What are the a-c power output, plate-circuit efficiency, and plate loss? What maximum plate dissipation rating should these tubes have if the maximum power output expected is 300 watts?

**6-19.** A push-pull output stage may serve as its own phase inverter. Draw a possible circuit, and discuss its limitations, particularly with respect to distortion.

**6-20.** Two type 810 tubes are used in a Class B audio-frequency amplifier, with $E_b = 2000$ volts. Determine the extended cutoff bias, and use this value. Assume $e_{b,\min} = 1.5e_{c,\max}$, and determine a load for an output of at least 500 watts per pair of tubes, with operation inside ratings, of $I_{dc}$ maximum per tube = 250 ma, maximum plate dissipation = 125 watts per tube. For specified load determine $P_o$, $P_{in}$, $\eta_p$, plate loss, and third-harmonic distortion for the amplifier. Determine average grid-circuit power input.

## REFERENCES


1. Espley, D. C.: The Calculation of Harmonic Production in Thermionic Valves with Resistive Load, *Proc. IRE*, **21**: 1439 (1933).
2. Preisman, A.: "Graphical Constructions for Vacuum-tube Circuits," McGraw-Hill Book Company, Inc., New York, 1943.
3. Thompson, B. J.: Graphical Determination of Performance of Push-pull Amplifiers, *Proc. IRE*, **21**: 591 (1933).
4. Barton, L. E.: High Audio Power from Relatively Small Tubes, *Proc. IRE*, **19**: 1131 (1931).
5. Sah, A. P-T.: Quasi Transients in Class B Audio-frequency Push-pull Amplifiers, *Proc. IRE*, **24**: 1522 (1936).
6. Nottingham, W. B.: Optimum Conditions for Maximum Power in Class A Amplifiers, *Proc. IRE*, **29**: 620 (1941).
7. Ryder, J. D.: "Electronic Fundamentals and Applications," Prentice-Hall, Inc., Englewood Cliffs, N.J., 1950.

CHAPTER 7

# FEEDBACK IN AMPLIFIERS

While the principle of *feedback*, or the comparison of the output of a device with its input to improve the accuracy of the output-input relation, had been previously applied in the process-control and other industries, a real understanding of its effects and in particular its application to electronic amplifiers was not available until H. S. Black published an important paper in 1934. This served to mathematically formalize the field of feedback, and the applications of this mathematical analysis have now made possible a real understanding of the control of physical processes in many fields, as well as creating usable and accurate servomechanisms, process controls, and vacuum-tube amplifiers.

The vacuum tube, alone, is an uncertain and variable electronic device. Combined with negative feedback, it becomes a precise and dependable engineering tool.

**7-1. Feedback Principles.** The response, or the ratio of output to input of an amplifier, is the gain **A**, which has been defined as

$$\mathbf{A} = \frac{\mathbf{E}_o}{\mathbf{E}_s}$$

If feedback is introduced into an amplifier, a simple block diagram may appear as in Fig. 7-1. Here a circuit has been added to introduce the output of the amplifier, possibly altered in magnitude and phase by a complex term $\beta$, back into the input. The term $\beta$ may be defined as

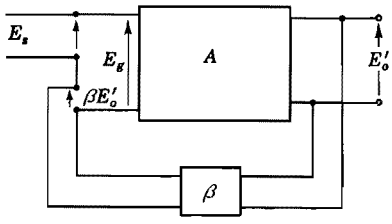



FIG. 7-1. Block diagram of a feedback system.

$$\beta = \frac{\text{voltage feedback } \underline{/\theta_1}}{\text{output voltage } \underline{/\theta_2}} \qquad (7\text{-}1)$$

The total input to the amplifier is given as

$$\mathbf{E}_g = \mathbf{E}_s + \beta\mathbf{E}'_o \qquad (7\text{-}2)$$

where the output of the amplifier, *with feedback*, is indicated as $\mathbf{E}'_o$. The amplifier has an internal gain **A**, so that $\mathbf{AE}_g = \mathbf{E}'_o$; it is then possible to

eliminate $\mathbf{E}_g$ in Eq. (7-2), and

$$\mathbf{E}'_o = (\mathbf{E}_s + \beta\mathbf{E}'_o)\mathbf{A}$$

from which the gain $A'$ with feedback is

$$\mathbf{A}' = \frac{\mathbf{E}'_o}{\mathbf{E}_s} = \frac{\mathbf{A}}{1 - \mathbf{A}\beta} \tag{7-3}$$

The gain of the amplifier is changed by the factor $1 - \mathbf{A}\beta$ when feedback is introduced. In general both $\mathbf{A}$ and $\beta$ will be complex. Three cases arise, depending on the magnitude of the denominator in Eq. (7-3):

1. If $|1 - A\beta|$ *is less than unity*, then the gain $\mathbf{A}'$ exceeds $\mathbf{A}$. The voltage feedback $\beta\mathbf{E}'_o$ adds to $\mathbf{E}_s$ and increases $\mathbf{E}_g$. The feedback is then said to be *positive*, and the circuit is *regenerative*. Positive feedback increases gain, increases distortion, reduces stability, and is usually avoided.

2. If $|1 - A\beta| = 0$, then the gain becomes infinite and output can be obtained with no external input, or the circuit becomes an oscillator. This situation will be studied in Chap. 11.

3. If $|1 - A\beta|$ *is greater than unity*, the gain $\mathbf{A}'$ is less than $\mathbf{A}$. The voltage feedback $\beta\mathbf{E}'_o$ subtracts from $\mathbf{E}_s$ and reduces $\mathbf{E}_g$. The feedback is then said to be *negative*, or *inverse*, and the circuit is *degenerative*. Inverse feedback reduces gain, reduces distortion and noise, increases stability, and is the subject of study in this chapter.

The sign in the denominator of Eq. (7-3) is obviously dependent on the manner in which $\beta\mathbf{E}'_o$ is introduced into the amplifier input, as well as its phase angle. The sign as given is a matter of convention, 1, 2, and 3 above serving to define positive and negative feedback.

It is now necessary to ascertain the effects of negative feedback on amplifier performance, to prove the statements made above. Requirements on $1 - \mathbf{A}\beta$, especially as it approaches zero, must also be studied.

**7-2. Negative Feedback and Gain Stability.** Equation (7-2) may be rearranged to

$$\mathbf{E}_s = \mathbf{E}_g - \beta\mathbf{E}'_o \tag{7-4}$$

This equation states that $\mathbf{E}_s$ and $-\beta\mathbf{E}'_o$ would be equal except for a voltage $\mathbf{E}_g$. It implies comparison of $\mathbf{E}_s$ and a portion of the output $-\beta\mathbf{E}'_o$ such that, if they are near equality, then $\mathbf{E}_g$ must be small. This comparison can be thought of as on an instantaneous basis, and if input and a portion of the output are near equality, then any internal changes of waveform or distortion must be small. Thus it is helpful to determine the conditions which will make the actual amplifier input $\mathbf{E}_g$ small.

Since $\mathbf{E}_o' = \mathbf{A}\mathbf{E}_g$, then

$$\mathbf{E}_s = \mathbf{E}_g - \mathbf{A}\boldsymbol{\beta}\mathbf{E}_g$$
$$\frac{\mathbf{E}_g}{\mathbf{E}_s} = \frac{1}{1 - \mathbf{A}\boldsymbol{\beta}} \tag{7-5}$$

Thus, for $\mathbf{E}_g \ll \mathbf{E}_s$, it appears that $|1 - A\beta|$ should be large with respect to unity. The term $\mathbf{A}\boldsymbol{\beta}$ is called the *feedback factor*, and the above implies that $|A\beta| \gg 1$. The feedback factor is in general a function of frequency, but in circuit design an attempt is usually made to bring its angle to 0° or 180° over the operating frequency range.

If $|A\beta| \gg 1$, then Eq. (7-3) for feedback gain becomes

$$\mathbf{A}' = -\frac{1}{\boldsymbol{\beta}} \tag{7-6}$$

showing that the gain can be made independent of tube parameters, supply voltages, and similar circuit variations. Since $\boldsymbol{\beta}$ is usually determined by a network of fixed circuit parameters, the amplifier gain $\mathbf{A}'$ can be stable as these parameters.

It may not always be possible to make $|A\beta|$ extremely large with respect to unity. To illustrate the improved gain stability achieved with feedback, it is necessary to be able to compute the resultant change in $A'$ for a given per cent change in $A$. For a small incremental change in $A$, Eq. (7-3) gives

$$dA' = \frac{(1 - A\beta)\,dA + A\beta\,dA}{(1 - A\beta)^2}$$

from which

$$\frac{dA'}{A'} = \left(\frac{1}{A} + \frac{\beta}{1 - A\beta}\right) dA$$
$$= \frac{1}{1 - A\beta}\frac{dA}{A} \tag{7-7}$$

Thus the change in gain with feedback is $1/(1 - A\beta)$ times as large as the change in internal gain of the amplifier. If $|1 - A\beta|$ is very large, the gain change with feedback can be made very small. This illustrates the stabilizing effect with feedback. By use of negative feedback many electronic instruments can be made precise and constant in accuracy.

**Example.** A given amplifier has a gain $A$ of $-20$. Feedback with $\beta = 0.25$ is added. If the internal gain changes by 10 per cent, find the change in over-all gain.

Here $dA/A = 0.1$, and $A\beta = -5$. Then

$$\frac{dA'}{A'} = \frac{1}{1 + 5} \times 0.1 = \frac{0.1}{6}$$
$$= 0.0166$$

so that, for a 10 per cent change in internal gain, the change in gain with feedback is only 1.66 per cent.

If the internal gain had been $-200$, so that $A\beta = -50$, the change in gain with feedback would have been only 0.00196, or 0.196 per cent.

**7-3. Negative Feedback and Distortion.** If $\mathbf{A}\beta$ is large, so that the gain $\mathbf{A}'$ depends largely on $\beta$, and if $\beta$ does not vary with frequency, then frequency distortion which may be present in the gain $\mathbf{A}$ will be reduced by feedback. Phase distortion will also be reduced. Conversely, by

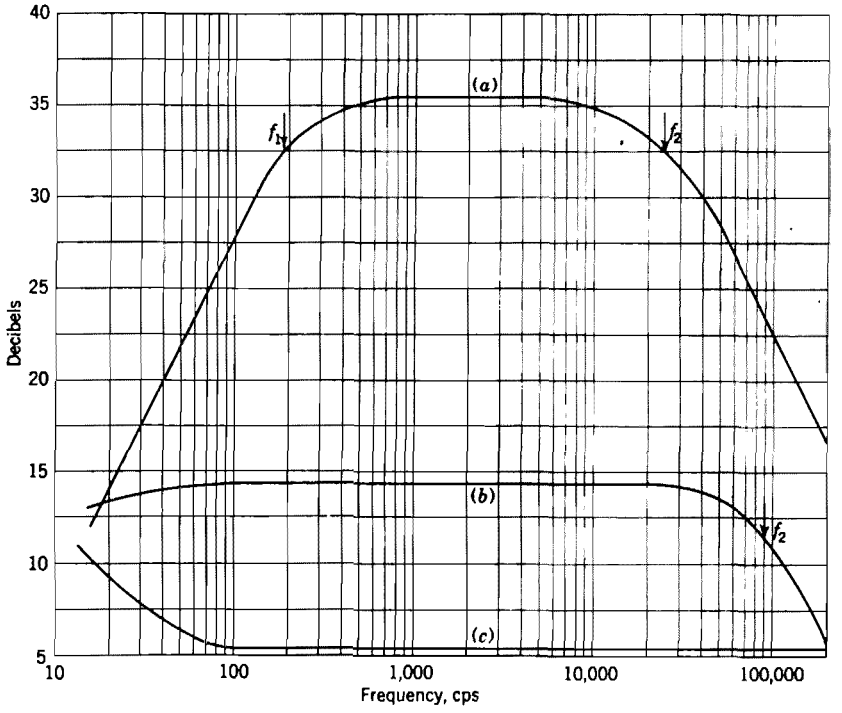



FIG. 7-2. Response of $RC$ amplifier. ($a$) No feedback; ($b$) with $\beta \cong 0.1$; ($c$) with $\beta \cong 0.55$.

building the $\beta$ circuit to vary as some function of frequency, Eq. (7-6) points out that the gain will then vary as an inverse function of the $\beta$-circuit frequency characteristic.

In Fig. 7-2 is an experimental demonstration of the ability of feedback to reduce frequency distortion when applied to a two-stage $RC$ coupled amplifier (using one $RC$ network). Without feedback curve (1) shows that the mid-frequency gain is 35.5 db at 1000 cycles, and the bandwidth frequencies are $f_1 = 190$ cycles, $f_2 = 25{,}000$ cycles. When approximately 10 per cent negative feedback ($\beta = 0.1$) is applied from the output to the first-stage cathode, the curve appears as at (2), with gain reduced to 14.3 db, but with $f_1$ below 15 cycles and $f_2$ extended to 90,000 cycles.

A further increase in feedback to a value of $\beta = 0.55$ produces curve (3). The gain is only 5.3 db, but the high-frequency range is extended above the limit of measurement at 200,000 cycles. The low-frequency end is interesting in that it shows the effect of regeneration. The d-c anode voltage was blocked from the feedback by a capacitor $C$, in series with a resistor. At frequencies below 100 cycles the phase shift introduced in the $\beta$ circuit by the capacitor reduced the magnitude of $1 - \mathbf{A}\beta$ until it was less than unity, and the gain increased. With the small value of $\beta$ used in curve (2) the resistor in series with $C$ was large, the $\beta$-circuit phase shift was small, and regeneration was not so evident. Actually a small amount of regeneration was probably present, as indicated by the small droop of curve (2) at the low-frequency end.

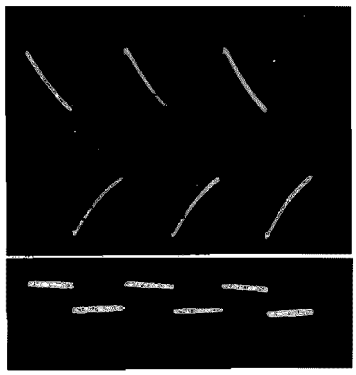



FIG. 7-3. Reduction of distortion by feedback. (*a*) Output with square-wave input and no feedback; (*b*) same with feedback.

Feedback will also reduce amplitude distortion. If $D$ is the amplitude distortion in an amplifier output at some frequency, it can be considered that $D$ is a function of the output voltage as

$$D = f(E_o)$$

since the distortion increases in some manner as the tube is driven over a wider range. With feedback added

$$D' = f(\mathbf{E}'_o) + \mathbf{A}\beta D'$$

where higher-order terms in $\beta$ are neglected. Then

$$D' = \frac{f(\mathbf{E}'_o)}{1 - \mathbf{A}\beta}$$

and if the output voltage with feedback is required to equal that without feedback, or $\mathbf{E}_o = \mathbf{E}'_o$, then

$$D' = \frac{D}{1 - \mathbf{A}\beta} \tag{7-8}$$

Thus the distortion component at a given frequency is reduced in the same magnitude as the gain. Other frequencies must be considered individually, each treated by the proper value of $\beta$ for that frequency.

The ability of negative feedback to reduce amplitude distortion is illustrated by Fig. 7-3. Curve (*a*) represents the output of an amplifier with square-wave input, but without feedback. Curve (*b*) shows the output of the same amplifier with negative feedback, the waveform being well restored to square form, although the gain has been reduced.

An output amplifier where amplitude distortion often occurs is usually

not required to provide much voltage gain, most of the over-all amplifier gain coming in the early low-amplitude stages. Thus a considerable reduction in distortion can be achieved by feedback across the output stage, and with the loss in gain made up in the earlier stages.

**7-4. Negative Feedback and Noise.** Noise, or hum, components introduced into an amplifier inside the feedback loop are also reduced by feedback. Suppose that an a-c hum source, possibly due to pickup from a-c heater wiring, exists in an amplifier as shown in Fig. 7-4, with gain $\mathbf{A}_1$ ahead and gain $\mathbf{A}_2$ following the source $N$, such that $\mathbf{A}_1\mathbf{A}_2 = \mathbf{A}$ as the over-all internal gain. The output voltage with feedback can be written as

$$\begin{aligned}\mathbf{E}_o' &= \mathbf{A}_1\mathbf{A}_2\mathbf{E}_s + \mathbf{A}_1\mathbf{A}_2\beta\mathbf{E}_o' + \mathbf{A}_2 N \\ &= \frac{\mathbf{A}_1\mathbf{A}_2}{1 - \mathbf{A}_1\mathbf{A}_2\beta}\left(\mathbf{E}_s + \frac{N}{\mathbf{A}_1}\right) \\ &= \frac{\mathbf{A}}{1 - \mathbf{A}\beta}\left(\mathbf{E}_s + \frac{N}{\mathbf{A}_1}\right) \qquad (7\text{-}9)\end{aligned}$$

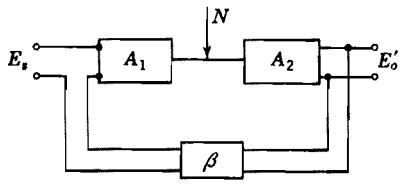



FIG. 7-4. Feedback amplifier with internally introduced noise.

It is apparent that the over-all gain has been reduced by $1 - \mathbf{A}\beta$ as was expected. In addition the noise output has been reduced by the factor $\mathbf{A}_1$, or the noise is additionally reduced in proportion to the *gain which precedes the noise source.*

**7-5. Feedback Measured in Decibels; an Example.** The ratio of power output with constant input, expressed in decibels, with and without feedback, is a measure of the amount of feedback introduced into an amplifier. For constant input

$$\text{db feedback} = 20 \log \frac{E_o'}{E_o} \qquad (7\text{-}10)$$

The previously developed relations will now be used in an example to illustrate the operation of a feedback amplifier.

**Example.** An amplifier has an internal gain of $-400$, distortion of 8 per cent, and a normal input of 0.5 volt. Find the effects of adding negative feedback with $\beta = 0.05$.

The feedback factor will be $\mathbf{A}\beta = -20$. The gain is reduced to

$$\mathbf{A}' = \frac{\mathbf{A}}{1 - \mathbf{A}\beta} = \frac{-400}{1 + 20} = -19$$

and the amplifier stability to gain changes is improved by a factor of 21. The distortion is reduced to

$$D' = \frac{D}{1 - \mathbf{A}\beta} = \frac{8}{21} = 0.38\%$$

If the amplifier is required to produce the same output with feedback as without, or $\mathbf{E}_o' = \mathbf{E}_o = \mathbf{A}'\mathbf{E}_s' = \mathbf{A}\mathbf{E}_s$, then

$$\mathbf{E}_s' = \mathbf{E}_s(1 - \mathbf{A}\beta)$$

so that the input-signal voltage required with feedback is

$$\mathbf{E}_s' = 0.5 \times 21 = 10.5 \text{ volts}$$

The output, by either method, is

$$\mathbf{E}_o = \mathbf{A}\mathbf{E}_s = -400 \times 0.5 = -200 \text{ volts}$$
$$\mathbf{E}_o' = \mathbf{A}'\mathbf{E}_s' = -19 \times 10.5 = -200 \text{ volts}$$

The actual input voltage with feedback, $\mathbf{E}_g'$, is

$$\mathbf{E}_g' = \mathbf{E}_s' + \beta\mathbf{E}_o' = 10.5 - 0.05 \times 200 = 0.5 \text{ volt}$$

While the required external input signal is increased by feedback from 0.5 to 10.5 volts, yet the amplifier is operating *internally* under a net input of 0.5 volt in both cases. This reasoning should not be carried too far, however, because if distortion is present, the input voltage $\mathbf{E}_g$ may differ in waveform when feedback is added, as in Fig. 7-5. In that figure the phase of the grid-cathode voltage with feedback (Fig. 7-5*c*) is reversed to illustrate the fact that Fig. 7-5*c* is the difference of the waves in Fig. 7-5*b* and *d*. With Fig. 7-5*c* oriented as it is, it can be seen that the grid-cathode voltage *with feedback* has developed a leading-edge spike, to reinforce the output of the amplifier at this point, since Fig. 7-5*a* shows that without feedback the amplifier suffers from leading-edge distortion or long rise time.

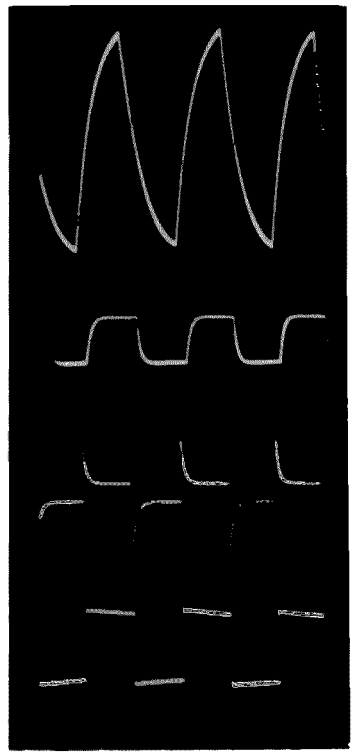



FIG. 7-5. (*a*) Output without feedback. (*b*) Output with feedback. (*c*) Grid-cathode voltage with feedback. (*d*) Square-wave input.

If the input voltage of the amplifier in the example is held constant at, say, 0.5 volt, the output voltages would be

Without feedback:

$$E_o = 0.5 \times (-400) = -200 \text{ volts}$$

With feedback:

$$E_o' = 0.5 \times (-19) = -9.5 \text{ volts}$$

The reduction in output power due to feedback, expressed in decibels, is

$$\text{db} = 20 \log \frac{9.5}{200} = -20 \times 1.324 = -26.5$$

Feedback has reduced the amplifier output by 26.5 db, or it is said that 26.5 db of feedback has been applied to the amplifier. It may also be stated that a feedback factor of 20 has reduced the output by 26.5 db.

**7-6. Basic Negative-feedback Circuits.** A great variety of circuits may be used to introduce negative feedback into an amplifier. However, they all fall into two general classes known as *voltage* feedback and *current* feedback. In voltage feedback the voltage fed back, or introduced into, the grid circuit is proportional to the output load voltage. In current feedback the voltage fed back is proportional to the current through the load. The two types of circuit produce somewhat different results.

A circuit of the voltage-feedback type is drawn in Fig. 7-6. The voltage introduced into the grid circuit is the drop across $R_2$, and this

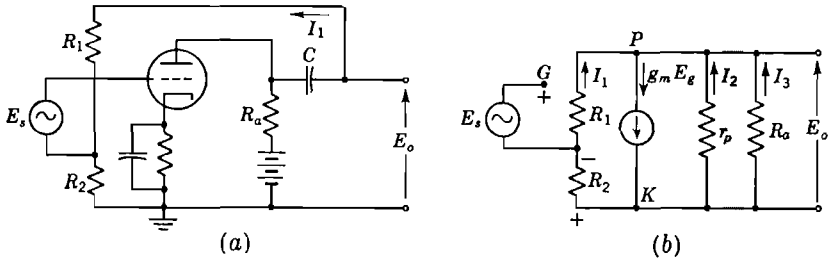



FIG. 7-6. Voltage-feedback circuit and its equivalent.

voltage drop is proportional to the output voltage, since $R_1$ and $R_2$ form a voltage divider across $\mathbf{E}_o$. The current $\mathbf{I}_1$ is given by $\mathbf{E}_o/(R_1 + R_2)$ and the cathode-to-grid voltage rise is

$$\mathbf{E}_g = \mathbf{E}_s + \mathbf{I}_1 R_2 = \mathbf{E}_s + \frac{R_2}{R_1 + R_2}\mathbf{E}_o \tag{7-11}$$

By comparison with Eq. (7-2) the value of $\beta$ is recognized as

$$\beta = \frac{R_2}{R_1 + R_2} \tag{7-12}$$

and the feedback $\mathbf{A}\beta$ is negative, since the gain $A$ is negative.

Usually $R_1 + R_2$ is large with respect to $r_p$ and $R_a$, and the reactance of $C$ is neglected. If the latter is not true, the value of $\beta$ will be altered at some frequencies. The circuit will be analyzed by ordinary circuit methods and by the feedback equations, as an illustration. Neglecting $I_1$, a current summation yields

$$g_m \mathbf{E}_g = \mathbf{I}_2 + \mathbf{I}_3 = -\mathbf{E}_o\left(\frac{1}{r_p} + \frac{1}{R_a}\right) = -\mathbf{E}_o \frac{r_p + R_a}{r_p R_a}$$

Bringing in Eq. (7-11), there results

$$\mathbf{A}' = \frac{\mathbf{E}_o}{\mathbf{E}_s} = \frac{-g_m}{(r_p + R_a)/r_p R_a + g_m\beta} = \frac{-\mu R_a}{r_p + (1 + \mu\beta)R_a} \tag{7-13}$$

This expression for gain with feedback may be altered to

$$\mathbf{A}' = \frac{-[\mu/(1 + \mu\beta)]R_a}{[r_p/(1 + \mu\beta)] + R_a} \tag{7-14}$$

Upon comparison of the above expression with the gain expression for a nonfeedback amplifier as

$$\mathbf{A} = \frac{-\mu R_a}{r_p + R_a}$$

it can be seen that the factors in brackets in Eq. (7-14) might be considered as $\mu'$ and $r_p'$ parameters of a hypothetical equivalent nonfeedback tube. Both parameters being reduced by the same factor shows the hypothetical equivalent tube to have lower $\mu$ and plate resistance but with an unchanged $g_m$ value. Thus *voltage feedback reduces the internal resistance of an amplifier* and tends to maintain constant voltage with changes in load resistance.

If the feedback expressions are used, it is necessary to compute $\mathbf{A}$, which is readily done by considering $R_2 = 0$, so that

$$\mathbf{A} = \frac{-\mu R_a}{r_p + R_a}$$

Then

$$\mathbf{A}' = \frac{\mathbf{A}}{1 - \mathbf{A}\beta} = \frac{-\mu R_a/(r_p + R_a)}{1 + [\mu R_a/(r_p + R_a)]\beta} = \frac{-\mu R_a}{r_p + (1 + \mu\beta)R_a} \tag{7-15}$$

which is the same result as in Eq. (7-13).

Figure 7-7 shows a simple form of current-feedback circuit. Here the voltage fed back into the grid circuit is $\mathbf{I}_p R_k$ and is proportional to the load current, and only incidentally to the load voltage. The grid voltage is

$$\mathbf{E}_g = \mathbf{E}_s - \mathbf{I}_p R_k \tag{7-16}$$

The output voltage is $-\mathbf{I}_p R_a$, and the feedback fraction is

$$\beta = \frac{R_k}{R_a} \tag{7-17}$$

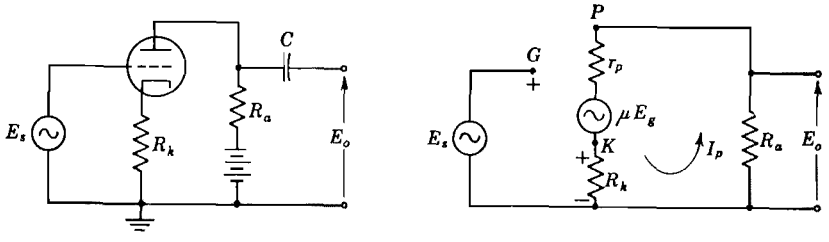
The plate current can be written from the circuit as

$$\mathbf{I}_p = \frac{\mu \mathbf{E}_g}{r_p + R_k + R_a} = \frac{\mu \mathbf{E}_s}{r_p + (1 + \mu)R_k + R_a} \tag{7-18}$$

and the gain is

$$\mathbf{A}' = \frac{\mathbf{E}_o}{\mathbf{E}_s} = \frac{-\mu R_a}{r_p + (1 + \mu)R_k + R_a} \tag{7-19}$$

for the current-feedback amplifier of Fig. 7-7.




(a) (b)

FIG. 7-7. Current-feedback circuit and its equivalent.

To use the feedback equation, the gain **A** is obtained by returning the grid source to $K$, so that

$$\mathbf{A} = \frac{-\mu R_a}{r_p + R_k + R_a} \tag{7-20}$$

The gain with feedback then is

$$\mathbf{A}' = \frac{\mathbf{A}}{1 - \mathbf{A}\beta} = \frac{-\mu R_a/(r_p + R_k + R_a)}{1 - (R_k/R_a)[-\mu R_a/(r_p + R_k + R_a)]} = \frac{-\mu R_a}{r_p + (1 + \mu)R_k + R_a} \tag{7-21}$$

Comparison of this expression with a nonfeedback amplifier shows that *current feedback raises the internal resistance* to $r_p + (1 + \mu)R_k$. If this is high with respect to $R_a$, the output current tends to remain constant or is independent of the load.

It should be noted that the cathode follower is a form of voltage-feedback amplifier. Since the voltage fed back is $\mathbf{E}_o = \mathbf{I}_p R_k$, the value of grid input is

$$\mathbf{E}_g = \mathbf{E}_s - \mathbf{E}_o$$

and $\beta = -1$. This high value of $\beta$ makes the cathode follower a very stable circuit.

The output impedance of the cathode follower was found to be

$$Z_{out} = \frac{[r_p/(1 + \mu)]R_k}{r_p/(1 + \mu) + R_k}$$

and this is reduced below that of the tube alone, as predicted by the statements above on voltage feedback.

With voltage feedback the output impedance of the amplifier is reduced, and the amplifier tends to become a constant-voltage source for a constant input. With current feedback the output impedance of the amplifier is increased, and the amplifier tends to become a constant-current source for constant input.

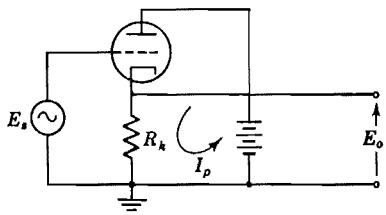



FIG. 7-8. Cathode-follower circuit.

**7-7. Other Feedback Circuits.** Feedback may be applied across more than one amplifier stage if attention be given to obtaining the proper phase relations. In the mid-frequency range of circuits having an odd number of grid-input–plate-output stages, the phase angle of the gain is an odd multiple of 180° or the gain is negative. The $\beta$ circuit then introduces a portion of the output directly into the grid circuit, and the feedback factor is $-\mathbf{A}\beta$. Such a circuit was the one shown in Fig. 7-6.

For circuits having an even number of grid-input–plate-output stages the phase angle of the gain is an even multiple of 180° and the gain is positive. The $\beta$ circuit must produce another 180° shift in introducing the feedback voltage into the input, and this is most easily done by feeding back into the cathode circuit of the first stage. This is shown for two stages in Fig. 7-9*a*. Some current feedback is also produced by $R_k$ in this circuit.

The circuits of Fig. 7-9*b* and *c* show methods of feeding back from the secondary of an output transformer. This places the transformer in the feedback loop, and frequency and waveform distortion generated by the transformer will be reduced. The circuit of Fig. 7-9*b* shows current feedback, and that of Fig. 7-9*c* gives voltage feedback. Positive or negative feedback is available, the proper sign of $\beta$ being ensured by choice of polarity of the transformer windings. Transformer phase angle may introduce difficulties, as discussed in Sec. 7-11.

Figure 7-9*d* shows a compound form of current and voltage feedback. The value of $\beta$ can be found to be

$$\beta = \frac{R_2}{R_1 + R_2} + \frac{R_k}{R_a} = \beta_v + \beta_c$$

thus combining the feedback percentages of the two methods. The gain

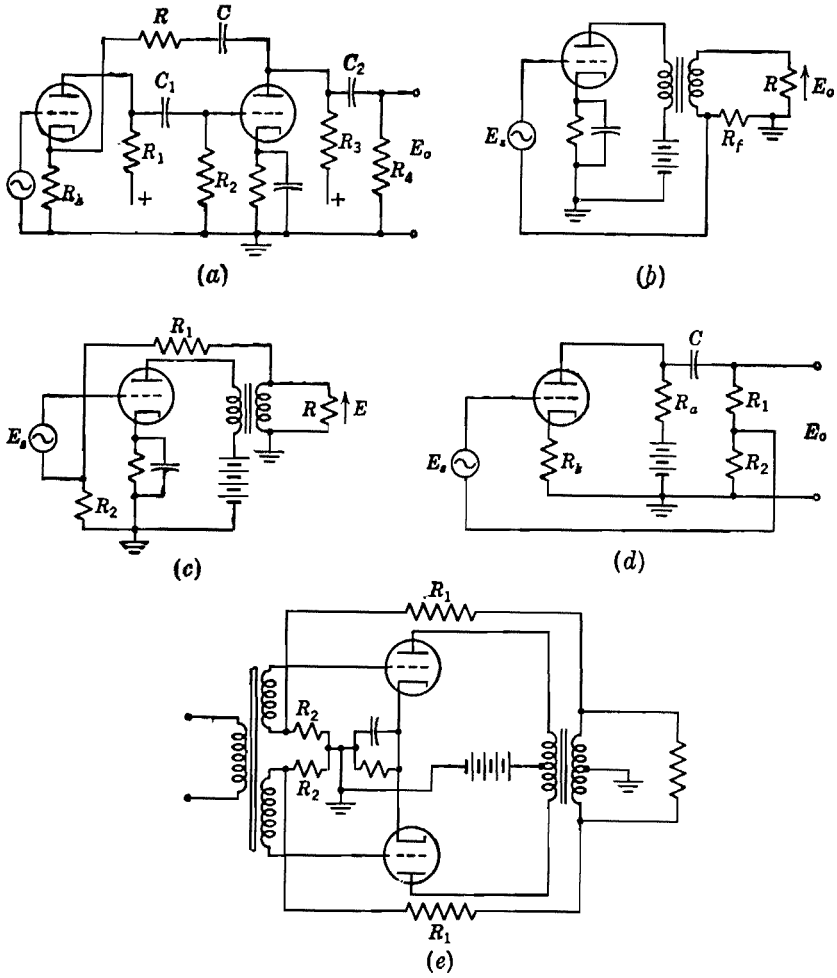



FIG. 7-9. Forms of feedback circuits.

can be written as

$$\mathbf{A}' = \frac{-\mu R_a}{r_p + (1 + \mu)R_k + (1 + \mu\beta_v)R_a} = \frac{-[\mu/(1 + \mu\beta_v)]R_a}{[r_p + (1 + \mu)R_k]/(1 + \mu\beta_v) + R_a} \tag{7-22}$$

The $\mu$ of the tube is reduced by the voltage feedback, and the effective plate resistance can be made greater, or less than, that of the tube alone by adjustment of $R_k$ and $\beta_v$. Thus such a circuit permits considerable control of internal impedance.

Figure 7-9*e* illustrates a method of adding voltage feedback to a push-pull amplifier. Feedback may be applied by several of the other methods as well.

**7-8. Positive Feedback; Decoupling.** Positive feedback is ordinarily avoided because of its usually undesirable effects. While it has an adverse effect on stability, there are occasions on which it may be added inside a negative-feedback loop, the negative feedback being sufficient to maintain the over-all amplifier stable and degenerative. In such cases the positive feedback will raise the gain and lower the output impedance since positive feedback introduces a feedback of such phase as to represent a negative resistance. When positive current feedback is placed inside a negative-voltage-feedback loop, it is possible to adjust the two types so that the output impedance becomes zero.

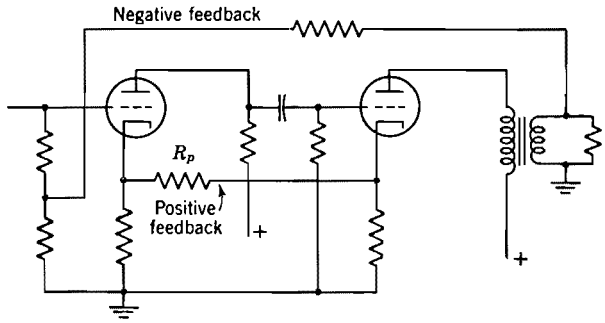



FIG. 7-10. Introduction of simultaneous positive and negative feedback.

A simple means of placing positive current feedback inside a negative-voltage-feedback loop is shown in Fig. 7-10, wherein resistor $R_p$ couples the cathodes of two successive stages. Overriding negative feedback is taken from the secondary of the output transformer to the grid circuit of the first tube.

Accidental sources of either positive or negative feedback exist in electric and magnetic couplings between input and output circuits, couplings through common wiring, and couplings through the internal impedance of common plate-voltage supplies. These accidental positive or negative feedbacks are responsible for amplifiers with unstable gain, for unexpectedly high or low gain, or for low- or high-frequency oscillation. A low-frequency oscillation of the order of a few cycles, which frequently occurs in audio amplifiers, has the descriptive name of "motorboating."

Shielding, isolation, and careful wiring location will usually eliminate sources of trouble other than that due to power-supply coupling. *Decoupling circuits*, as indicated by the *RC* filters of Fig. 7-11, are usually required in most plate and grid return leads to prevent the various signal currents from flowing through the common power-supply imped-

ance, the signal currents in each stage then being effectively returned directly to the cathodes. Values of $R$ are chosen to be as high as possible without serious plate-voltage drop. The time constant of the combination should be ten or more times the period of the lowest frequency to be passed.

The amount and kind of instability produced by these stray feedbacks will depend on the number of amplifier stages and the points at which the feedback voltage originates and at which it is fed back. For reasons already discussed, feedback through power supply from output to input grid circuit will usually be degenerative for an odd number of stages, regenerative for an even number.

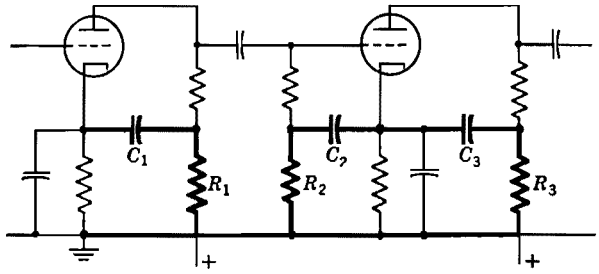



FIG. 7-11. Decoupling circuits in an amplifier.

**7-9. Stability of Feedback Amplifiers.** It was previously pointed out that $|1 - A\beta| > 1$ if the feedback was to be negative and the amplifier stable. Since $\mathbf{A}\beta$ is complex and usually a frequency function, it is necessary to study its magnitude and angle throughout the frequency range to ascertain that $|1 - A\beta| > 1$ for all frequencies. It is quite possible that over some frequency range $|1 - A\beta| < 1$, or may even be equal to $1 + j0$. If the former situation applies, then the amplifier will be regenerative, whereas if the latter condition is true, the amplifier will oscillate.

It is possible to study the action of $\mathbf{A}\beta$ most readily by a complex-plane plot. The magnitude and phase of $\mathbf{A}\beta$ are calculated at all frequencies from $f = 0$ to $f = \infty$ and the results plotted as a polar curve on the complex plane with real and $j$ axes.

As an illustration Fig. 7-12 is drawn for a 6J5 tube in an $RC$-coupled amplifier with a mid-frequency gain of $-14$ and $\beta = 0.25$ for all frequencies. Data are available in Figs. 5-6 and 5-8, to prepare the accompanying table of $A\beta$ values. It is found that the polar plot for an $RC$

| *Frequency* | $A\beta$ | *Frequency* | $A\beta$ |
|---|---|---|---|
| $0.1f_1$ | $0.35/264°$ | $0.5f_2$ | $3.11/153°$ |
| $0.5f_1$ | $1.57/243°$ | $f_2$ | $2.47/135°$ |
| $f_1$ | $2.47/225°$ | $2f_2$ | $1.57/117°$ |
| $2f_1$ | $3.11/207°$ | $8f_2$ | $0.35/96°$ |
| Mid-range | $3.50/180°$ | | |

amplifier is a circle, with frequency progressing clockwise. Various frequency points are indicated along the circle, and the entire mid-frequency range appears at the 180° point. The circle represents the locus of $\mathbf{A}\beta$ for all frequencies. Also drawn is a vector $1/0°$, so that the vector difference $1 - \mathbf{A}\beta$ can be determined graphically, and this difference is also indicated.

A circle of unit radius is shown dashed around the 1, $j0$ point, and it can be reasoned that the value of $|1 - \mathbf{A}\beta|$ can be less than unity only if

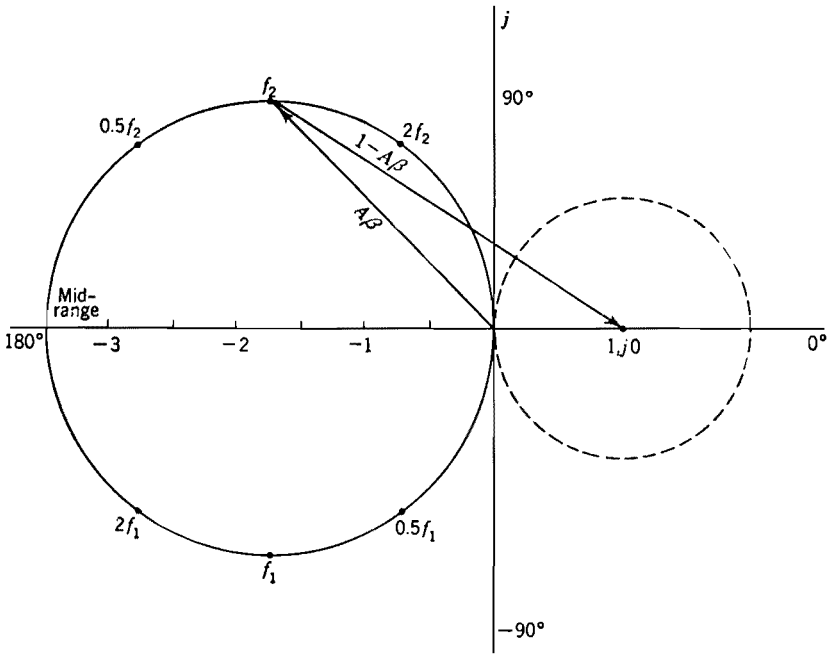



FIG. 7-12. Polar plot of $A\beta$ for an $RC$ amplifier.

the $1 - \mathbf{A}\beta$ line ends inside this unit circle. The only way in which $1 - \mathbf{A}\beta$ may end inside the unit circle is for the polar $\mathbf{A}\beta$ locus to pass inside the unit circle. Thus a condition for stability without regeneration is that the $\mathbf{A}\beta$ locus not pass through the unit circle around the 1, $j0$ point. Obviously the feedback amplifier for which Fig. 7-12 is drawn has degenerative feedback at all frequencies.

If the $\mathbf{A}\beta$ locus passes through the unit circle, the amplifier will be regenerative. If the $\mathbf{A}\beta$ locus passes through the 1, $j0$ point, the amplifier will oscillate. In fact, Nyquist has shown *that if the* $\mathbf{A}\beta$ *locus envelops the* 1, $j0$ *point the amplifier will oscillate.*

To illustrate the point further, Fig. 7-13 is drawn for a feedback amplifier including an output transformer in the feedback loop. At frequencies

below the mid-range the behavior is similar to that of the *RC* amplifier. As the frequency rises, the phase and gain both increase because of the effects of transformer resonance, and at still higher frequencies the phase angle may near 0°, although at small gain. The $\mathbf{A\beta}$ locus there passes through the unit circle, and at such frequencies the amplifier will be regenerative.

In general, oscillation will not occur if the absolute value of $\mathbf{A\beta}$ becomes less than unity when the phase angle reaches 360°. This requires control of amplifier gain over frequency ranges much wider than the expected operating range.

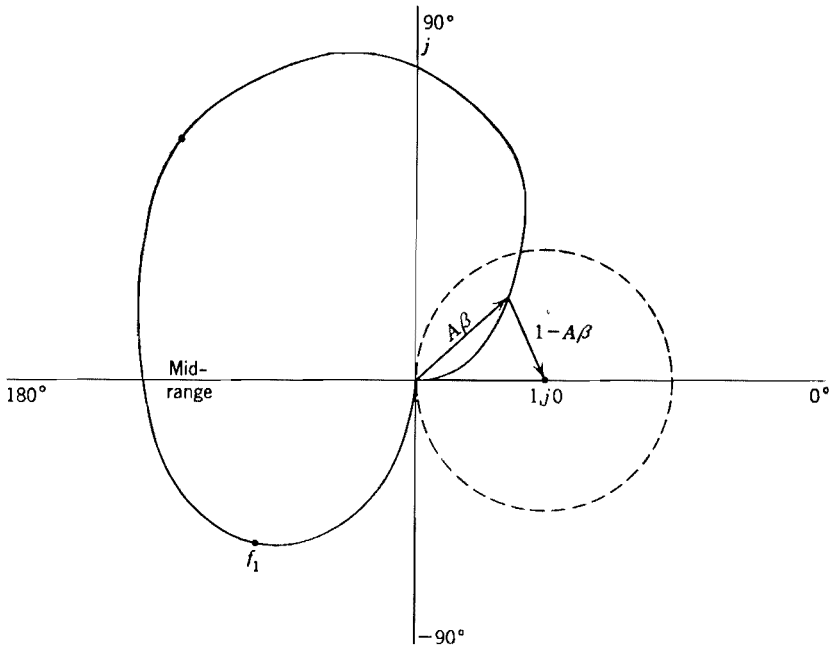



FIG. 7-13. Polar plot of $A\beta$ of an amplifier including a transformer.

**7-10. Further Discussion of Stability Conditions.** The plotting of the complete $\mathbf{A\beta}$ locus for an amplifier may be quite tedious. Other methods which avoid this, and also introduce some new concepts, will be presented here.

A network of pure reactances shows a succession of resonances and antiresonances as the sinusoidal frequency is varied. The resonances and antiresonances are called zeros and poles, respectively, as they appear along the frequency axis. The reactance function describing the network occurs as a quotient of two polynomials in angular frequency $\omega$, and each of these is factorable into its roots, the zeros occurring at the

root values of the numerator, and the poles occurring at the roots of the denominator. Such a function may have the form

$$F(\omega) = M \frac{(\omega - \omega_1)(\omega - \omega_3) \cdots}{(\omega - \omega_2)(\omega - \omega_4) \cdots} \tag{7-23}$$

where $M$ is a constant multiplier known as the scale factor. This function would have zeros at $\omega_1$ and $\omega_3$, poles at $\omega_2$ and $\omega_4$.

The transfer ratio or gain of an amplifier is a polynomial ratio of this form. If it is generalized further by using the notation of the Laplace transform, substituting $s$ for $j\omega$, the resulting mathematical form is analogous to that above. For instance, the high-frequency response of an $RC$ amplifier is available as

$$\mathbf{A}_H = -g_m R \frac{1}{1 + j\omega RC}$$

If $j\omega$ is replaced by $s$, this becomes

$$\mathbf{A}_H = \frac{-g_m}{C} \frac{1}{s + 1/RC} \tag{7-24}$$

$$= \frac{-g_m}{C} \frac{1}{s + s_1}$$

The denominator polynomial contributes a pole at $s = -1/RC$. The scale factor, or constant multiplier, is $g_m/C$. If a complete amplifier includes several such stages, the zeros and poles from the several stages are superposed to form the complex plane pattern of the whole design.

As a further illustration of the determination of zeros and poles, consider the gain expression for the shunt-compensated wide band amplifier of Sec. 5-7. The gain is, by Eq. (5-39),

$$\mathbf{A}(\omega) = -g_m \frac{(R + j\omega L)(1/j\omega C)}{R + j\omega L + 1/j\omega C} \tag{7-25}$$

where $C = C_a$, the shunt capacity of the stage. Rewriting in terms of $s$,

$$A(s) = -g_m \frac{(R + sL)(1/sC)}{R + sL + 1/sC}$$

which can be reduced to

$$A(s) = \frac{-g_m}{C} \frac{s + R/L}{s^2 + sR/L + 1/LC} \tag{7-26}$$

$$= \frac{-g_m}{C} \frac{s + s_o}{(s + s_1)(s + s_2)} \tag{7-27}$$

where $s_o$ is the zero at $s = -R/L$ and the two denominator poles are real or complex conjugate, depending on the value of $R$.

While $s$ is analogous to frequency, Laplace-transform theory indicates

that in general $s$ is a complex number, $s = \sigma + j\omega$. The concept of $\omega$ alone is as a positive real frequency, and $s$ is spoken of as a complex frequency. This is a valuable generalizing concept, since if we consider the exponential

$$\epsilon^{st} = \epsilon^{(\sigma+j\omega)t}$$

it is seen as a growing or decaying sinusoid, and thus all transient and steady-state network response is included in the complex-frequency concept. The real component $\sigma$ represents the damping which controls the rate of growth or decay, and the imaginary component supplies the real frequency. Values of $s$ can be plotted on a complex plane, with $\sigma$ on the real axis and $\omega$ along the imaginary axis. Zeros and poles of $s$ can be marked by circles and crosses, respectively, to indicate the function

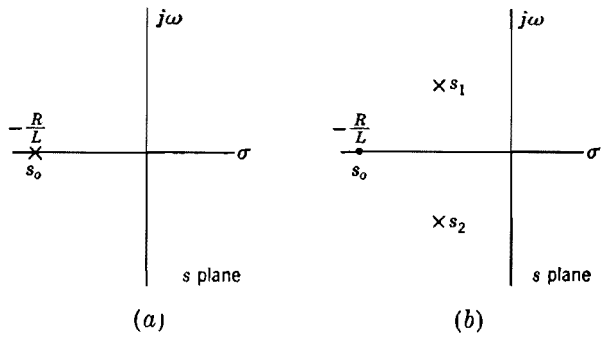



FIG. 7-14. Zero and pole location for (*a*) Eq. (7-24), (*b*) Eq. (7-27).

responses. Plots of the functions of Eqs. (7-24) and (7-27) are shown in Fig. 7-14.

The location of the poles and zeros becomes a matter of factoring the polynomials in the gain expression as a function of $s$. It should be noted that zeros and poles must be real and lie on the $\sigma$ axis, or they must occur in complex conjugate pairs. Any zero or pole on one side of the real axis must have a paired conjugate. Conjugate complex poles represent growing or decaying sinusoids, depending on whether $\sigma$ is positive or negative. Conjugate imaginary poles represent constant-amplitude sinusoidal response and appear on the $j$ axis.

A positive value for $\sigma$, or a pole location in the right half plane, indicates a sinusoid growing without limit, and the amplifier will be unstable. In turn a negative value of $\sigma$ indicates a damped sinusoid which will eventually decay to zero. Since only a theoretical system with no losses will be represented by a pole with $\sigma = 0$, this situation is of no concern in physically realizable systems.

Thus it can be reasoned that the necessary and sufficient condition

that *a feedback amplifier be stable is that its gain or transfer ratio, as a function of s, shall have no poles in the right-hand half of the s plane.*

If this reasoning is carried one step further to the general expression for the gain of a feedback amplifier expressed as an $s$ function,

$$G(s) = \frac{\mathbf{A}}{1 - \mathbf{A}\beta} \tag{7-28}$$

it can be seen that the critical matter of stability revolves around the *zeros* of the denominator $1 - A\beta$ which occur in the right half plane. If the denominator has zeros there, then the full fraction $G(s)$ will have

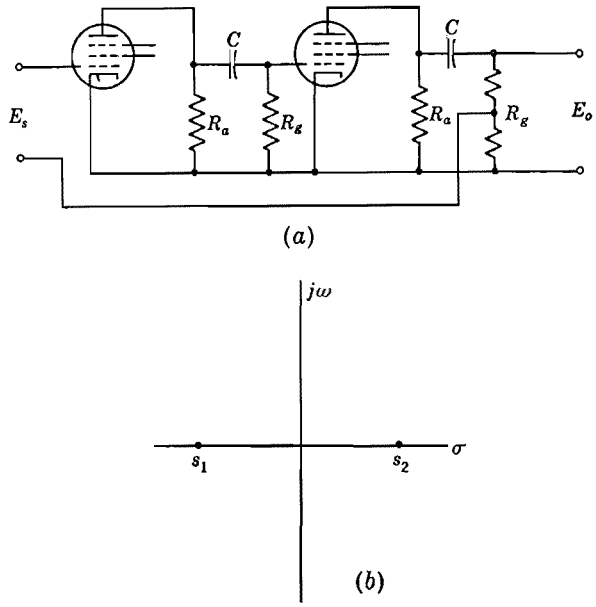



FIG. 7-15. Illustration of Eq. (7-30).

poles in the right half plane and instability follows. If there are no zeros of the denominator in the right half plane, the amplifier is stable. This result is one form of the Nyquist criterion.

To illustrate the method, consider a two-stage $RC$-coupled amplifier of identical stages, with feedback introduced as in Fig. 7-15*a*. The value of $\beta$ is fixed, and its angle is zero for all frequencies. Using the high-frequency gain, with $R_g \gg R_a$,

$$\mathbf{A}(\omega) = \left(\frac{-g_m R_a}{1 + jf/f_2}\right)\left(\frac{-g_m R_a}{1 + jf/f_2}\right)$$

$$A(s) = \sqrt{\frac{(-g_m R_a)^2}{(1 + s/\omega_2)^2}}$$

Writing $G(s)$, by use of Eq. (7-28),

$$G(s) = \frac{g_m^2 R_a^2}{(1 + s/\omega_2)^2 - g_m^2 R_a^2 \beta} \tag{7-29}$$

which becomes

$$G(s) = \frac{g_m^2 R_a^2 \omega_2^2}{s^2 + 2\omega_2 s + \omega_2^2(1 - g_m^2 R_a^2 \beta)} \tag{7-30}$$

It is the denominator of Eq. (7-30) which is to be examined for zeros in the right half plane. The roots of the denominator may be found to be

$$s_1 = -\omega_2(1 + g_m R_a \sqrt{\beta})$$
$$s_2 = -\omega_2(1 - g_m R_a \sqrt{\beta})$$

For $g_m R_a \sqrt{\beta} = 1$ the root $s_2$ will represent a zero in the right half plane, and the zeros will appear in the $s$ plane as in Fig. 7-15*b*. Since $g_m R_a$ is the mid-frequency gain of the amplifier, it follows that, for a mid-frequency gain of $1/\sqrt{\beta}$ or greater, the amplifier with feedback over two stages will be unstable, as expected from previous theory.

Since $g_m^2 R_a^2$ will also be unity or greater for instability, it follows that any over-all gain greater than unity, within the feedback loop, will lead to instability. Thus the result above applies for nonidentical as well as identical stage amplifiers.

## PROBLEMS

**7-1.** An amplifier with negative feedback has a gain of 41. It is found that without feedback 0.15 volt signal is required to produce a given output, whereas the input must be 3.2 volts with feedback, for the same output. What is the value of $\beta$ and of the gain before feedback is added?

**7-2.** A single-stage amplifier has a gain of $-20$ and an output of 105 volts with 8 per cent distortion. If feedback is added to reduce the distortion to 2 per cent, what value of $\beta$ should be used, and what output voltage will be obtained for the same input that gave 27 volts without feedback?

**7-3.** (*a*) Draw a feedback $\beta$ network which will allow the amplifier output to rise at low frequencies.

(*b*) Repeat with a $\beta$ circuit which will raise the output of the high frequencies with respect to the low frequencies.

**7-4.** An amplifier has an internally introduced hum output of 1 volt when the signal output is 125 volts. Determine the value of $\beta$ required to improve the signal-to-noise ratio in the output to 1000:1, for the same signal output. The gain with feedback is $+220$.

**7-5.** An *RC*-coupled amplifier has $f_1 = 60$ cycles, $f_2 = 15{,}500$ cycles. If the mid-frequency gain is $-127$, find the values of $f_1$ and $f_2$ if feedback with $\beta = 0.1$ is added.

**7-6.** The circuit of Fig. 7-6*a* is used with a tube having $\mu = 70$, $r_p = 60{,}000$ ohms. If $R_a = 100{,}000$ ohms, $R_1 + R_2 = 1$ megohm, specify $R_1$ and $R_2$ if the gain is to be $-6$, the reactance of $C$ being negligible.

**7-7.** (*a*) Find the value of $R_k$ needed in Fig. 7-7*a* to provide 10 per cent feedback,

with $\mu = 20$, $r_p = 6000$ ohms, $R_a = 100{,}000$ ohms. (*b*) A capacitor of 0.5 $\mu$f is connected across $R_k$. Compute the gain and phase shift over the low-frequency range. Would this be a satisfactory method of reducing 60-cycle response compared with that at 600 cycles?

**7-8.** (*a*) For the circuit of Fig. 7-6*a*, with $\mu = 70$, $r_p = 77{,}000$ ohms, $R_a = 100{,}000$ ohms, $R_1 = 150{,}000$ ohms, $R_2 = 10{,}000$ ohms, find the gain. (*b*) Remove the feedback by returning the grid to ground, and determine the gain again.

**7-9.** The circuit of Fig. 7-9*a* is used with 6J5 tubes having $\mu = 20$, $r_p = 7700$ ohms, and $R_k = 2000$ ohms, $R_1 = R_3 = 100{,}000$ ohms, $R_2 = R_4 = 250{,}000$ ohms, $C_1 = C_2 = 0.1$ $\mu$f. (*a*) If $R = 40{,}000$ ohms, $C = 1.0$ $\mu$f, find the gain and phase shift at 100 cycles. (*b*) Let $R = \infty$, and find the gain and phase shift at 100 cycles.

**7-10.** (*a*) Draw the current-source equivalent circuit for Fig. 7-9*d*. (*b*) Prove that the gain is given by Eq. (7-22). Neglect the reactance of $C$. (*c*) Also show that $\beta$ is equal to the sum of the voltage and current values for $\beta$.

**7-11.** If the tube of Fig. 7-9*c* is a 6V6 tetrode, in linear Class A operation, and if the transformer ratio is $a = 4.5$, $R = 500$ ohms, find the value of $R_2$ for $E_s = 10$ volts, $R_1 = 50{,}000$ ohms, and the secondary voltage is to be 20 volts. Give careful consideration to computing $\beta$.

**7-12.** In Fig. 7-9*b* $R = 8$ ohms. If the transformer has a ratio of $a = 23$, and if $\mu = 19$, $r_p = 5000$ ohms, find the value for $R_f$ if 10 db of negative feedback is to be added.

**7-13.** In Fig. 7-9*c* the tube is a tetrode with $g_m = 5000$ $\mu$mhos and $r_p = 46{,}000$ ohms. The load $R$ is 10 ohms, and the transformer ratio is 20:1 step-down to the load. (*a*) Specify values for $R_1$ and $R_2$ if 12 db of negative feedback is to be added. (*b*) If $E_s = 1.0$ volt, find the output power supplied to $R$.

**7-14.** An $RC$-coupled amplifier of three identical stages has $f_1 = 23$ cycles, $f_2 = 92$ kc for each stage. If mid-range gain is 230, plot the complex voltage-gain locus for the amplifier, indicating actual frequencies thereon.

**7-15.** A feedback amplifier with $\beta = 0.1$ has a mid-frequency gain of $14\underline{/180°}$. Included in the feedback loop is a transformer with the frequency characteristics shown in the accompanying tabulation. Plot the amplifier $A\beta$ locus, and discuss performance, assuming changes in $A\beta$ are due only to the transformer.

| *f, cycles* | *Ratio a* | *f, cycles* | *Ratio a* |
|---|---|---|---|
| 20 | $0.6\underline{/80°}$ | 12,000 | $4.5\underline{/-20°}$ |
| 50 | $1.35\underline{/63°}$ | 16,000 | $6.8\underline{/-40°}$ |
| 100 | $2.1\underline{/45°}$ | 20,000 | $9.0\underline{/-90°}$ |
| 800 | $3.0\underline{/10°}$ | 30,000 | $2.1\underline{/-156°}$ |
| 1000–7000 | $3.1\underline{/0°}$ | 40,000 | $1.2\underline{/-164°}$ |
| 8000 | $3.6\underline{/-8°}$ | | |

**7-16.** Reduce Eq. (5-46), for the gain of a series-compensated amplifier, to an expression in $s$, and draw an $s$-plane map showing poles and zeros.

**7-17.** A given feedback amplifier has a feedback factor given by

$$A\beta = \frac{-40(1 + 0.02s)}{1 + 0.25s + 0.01s^2}$$

Plot poles and zeros in the complex plane, and state whether the amplifier is stable.

**7-18.** Determine stability of each of the following feedback amplifiers, by plotting poles and zeros:

(*a*) $$A\beta = \frac{45}{s^2 + 2s + 1}$$

(*b*) $$A\beta = \frac{12}{s^2(0.4s^2 + 1.3s + 1)}$$

(*c*) $$A\beta = \frac{1 + 3s}{-s(1 + 0.3s)}$$

## REFERENCES


1. Black, H. S.: Stabilized Feedback Amplifiers, *Elec. Eng.*, **53**: 114 (1934).
2. Nyquist, H.: Regeneration Theory, *Bell System Tech. J.*, **11**: 126 (July, 1932).
3. Bode, H. W.: "Network Analysis and Feedback Amplifier Design," D. Van Nostrand Company, Inc., Princeton, N.J., 1945.
4. Terman, F. E.: Feedback Amplifiers, *Electronics*, **10**: 15 (January, 1937).
5. Mayer, H. F.: Control of the Effective Internal Impedance of Amplifiers by Means of Feedback, *Proc. IRE*, **27**: 213 (1939).
6. Learned, V.: Corrective Networks for Feedback Circuits, *Proc. IRE*, **32**: 403 (1944).
7. Terman, F. E., R. R. Buss, W. R. Hewlett, and F. C. Cahill: Some Applications of Negative Feedback with Particular Reference to Laboratory Equipment, *Proc. IRE*, **27**: 649 (1939).
8. Scott, H. H.: A New Type of Selective Network and Some of Its Applications, *Proc. IRE*, **26**: 226 (1938).
9. Lynch, W. A.: The Stability Problem in Feedback Amplifiers, *Proc. IRE*, **39**: 1000 (1951).
10. Nixon, F. E.: "Principles of Automatic Control," Prentice-Hall, Inc., Englewood Cliffs, N.J., 1953.

CHAPTER 8

# DIRECT-COUPLED AMPLIFIERS; COMPUTING AMPLIFIERS

In much measurement work it is desirable to retain the d-c components, initial displacements, and the like, that may accompany the a-c components or varying signals. Amplifiers which will amplify down to d-c or zero frequency as one limit are then required. Since such amplifiers cannot employ reactive coupling elements, phase shift and frequency distortion are absent up to those frequencies at which the internal tube and stray shunt capacities become effective.

A tube in a direct-coupled amplifier, however, is unable to distinguish between an input d-c signal and a shift in its bias voltage, so that direct-coupled amplifiers introduce special problems in stability, and methods of overcoming these difficulties will be discussed here.

Certain forms of direct-coupled amplifiers are also employed for performing mathematical operations in electronic computing machines of the continuous variable or analog type, and these circuits will also be introduced.

While the direct-coupled amplifier is able to amplify direct voltages, the term as abbreviated, "d-c" amplifier, means a direct-coupled amplifier and not necessarily a direct-current amplifier, since it is quite possible that a direct-coupled amplifier may be used to amplify alternating current only.

**8-1. Early Direct-coupled Amplifiers.** If the blocking capacitors are removed from an *RC*-coupled amplifier, it is necessary to restore the tubes to the proper *Q* points by inserting grid-bias batteries, as shown in Fig. 8-1. This circuit will amplify a steady or slowly varying input voltage and will have a gain of

$$\mathbf{A} = \frac{-\mu R_1}{r_p + R_1} \tag{8-1}$$

for each stage. Considering the interstage network, the voltage $E_{cc}$ can be determined if $I_{b1}$ is known. If $E_{bb} = 250$ volts, $I_{b1} = 1.5$ ma, and $R_1 = 100{,}000$ ohms, then the drop in $R_1$ is 150 volts. The potential at point $A$ is then $+100$ volts to cathode, and this would be the grid

voltage without the bias battery. To place the tube $T_2$ at a $Q$ point of $E_c = -10$ volts requires that $E_{cc} = -110$ volts, which is a large battery.

The battery requirement is not a desirable feature and would make the circuit also suffer from high shunt capacity, due to the capacity of the

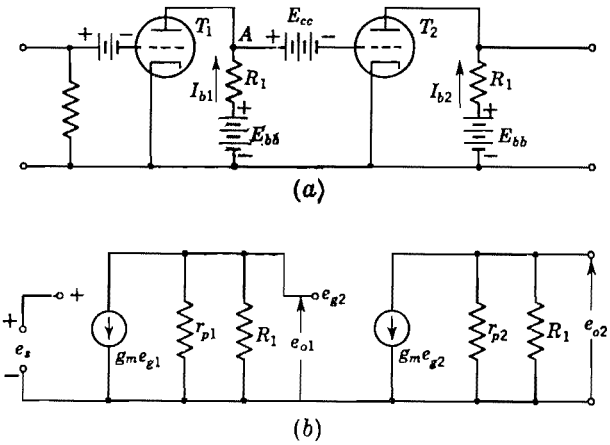



FIG. 8-1. Two-stage direct-coupled amplifier.

grid-bias batteries to ground. This would seriously limit the transient and high-frequency response.

Another version, more suited to operation from a rectified a-c supply, is the Loftin-White circuit of Fig. 8-2. A single d-c supply furnishes all grid and anode voltages by means of a tapped resistor. The current $I$ bled through this resistor should be several times the total current demand of the amplifier, in order to make the tapped resistor low in value and to

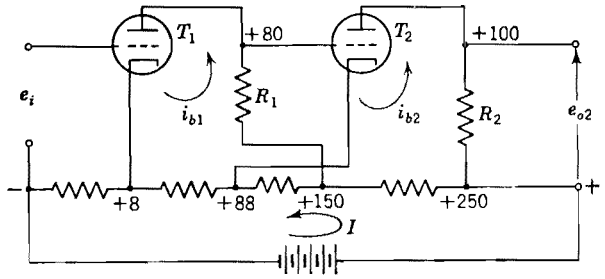



FIG. 8-2. The Loftin-White form of d-c amplifier.

establish the supply voltages as somewhat independent of the tube currents.

A typical set of voltages is indicated in the figure. The cathode of $T_1$ is at +8 volts, determining the grid bias of this tube. The anode supply

voltage of $T_1$ is $150 - 8 = +142$ volts. The $I_{b1}R_1$ drop is 62 volts, making the potential $E_{b1} = +80$ volts, and this is also the potential of the grid of $T_2$ above negative reference. In order to obtain $-8$ volts bias on $T_2$, its cathode is returned to a tap on the resistor at $+88$ volts, 8 volts above its grid. The anode of $T_2$ is supplied at $+250$ volts, or $250 - 88 = 162$ volts above its cathode. If the $I_{b2}R_2$ drop is 62 volts, then potential $E_{b2} = +100$ volts, and

$$e_{o2} = +100 - (+250) = -150 \text{ volts output}$$

Such circuits will amplify extremely low frequencies and introduce no frequency and phase distortion. However, owing to the high supply voltage required for multistage amplifiers, the high operating voltage between cathodes and heaters, and the waste of power in the dividing resistor, the circuit has been superseded by other types.

These circuits also suffer from the inherent instability of all d-c amplifiers, due to tube-characteristic drift, supply-voltage change, and heater-voltage change with cathode-type tubes. Changes in supply voltage move the operating point in just the same way as does a signal, so that the amplifier cannot tell the difference between a change in bias and a change in signal, and the output drifts. Zero adjustment features are usually required to restore the output voltage to zero for zero input, or to compensate for internal drift.

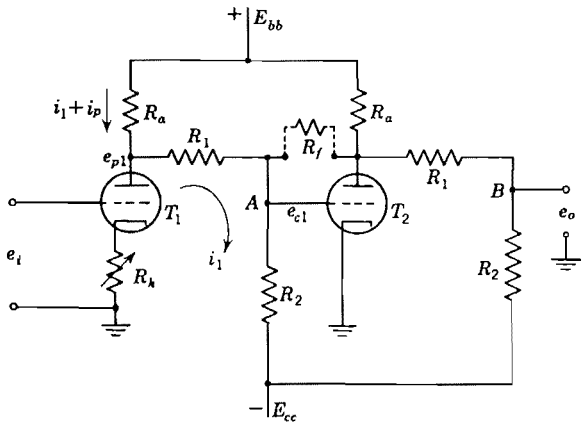



FIG. 8-3. Resistance coupling with positive and negative power supplies.

**8-2. The Resistance-coupled Amplifier.** In modern d-c amplifiers many basic circuit principles are combined, and it is the intent of this and following sections to discuss a few of the most common.

To overcome the common coupling and feedback possible through the tapped supply resistor of the Loftin-White circuit, it is possible to use the resistance-coupled circuit of Fig. 8-3. This requires two power

supplies, one positive to ground, the other negative to ground. By adjustment of the bias of $T_1$ with a variable cathode resistor, it is possible to shift the potential level at point $B$, so that $B$ is brought to ground or midway between the two power-supply voltages, for zero input signal. The output voltage will then also be zero. If regulated power supplies are used and the tubes well aged, the output may drift only a fraction of a volt in a few hours. Occasional manual adjustment will then cancel the drift.

Obviously, there is a loss in gain through introducing the voltage divider $R_1$, $R_2$, but this and the dual power supply are the price to be paid for gaining some control of the internal d-c potential levels. The gain of each stage can readily be found as

$$\begin{aligned} A &= \frac{-\mu R_a(R_1 + R_2)/(R_a + R_1 + R_2)}{r_p + [R_a(R_1 + R_2)/(R_a + R_1 + R_2)]} \frac{R_2}{R_1 + R_2} \\ &= \frac{-\mu R_a R_2/(R_a + R_1 + R_2)}{r_p + [R_a(R_1 + R_2)/(R_a + R_1 + R_2)]} \end{aligned} \tag{8-2}$$

and for high gain $R_2 \gg R_1$. Thus d-c amplification can be obtained with this sort of circuit with a common anode-voltage supply, without introducing biasing complications, and with provision for manually balancing the internal drift by use of a variable bias in the first stage.

Referring to Fig. 8-3, if $e_{p1}$ and $e_{c1}$ are the potential levels above ground reference for the plate of the first stage and the grid of the second, respectively, then suitable values for $R_1$ and $R_2$ may be computed from the equations

$$\begin{aligned} (i_1 + i_p)R_a &= E_{bb} - e_{p1} \\ e_{p1} - e_{c1} &= i_1 R_1 \\ e_{c1} - E_{cc} &= i_1 R_2 \end{aligned}$$

from which

$$\begin{aligned} \frac{R_1}{R_a} &= \frac{e_{p1} - e_{c1}}{E_{bb} - e_{p1} - i_p R_a} \\ \frac{R_2}{R_a} &= \frac{e_{c1} - E_{cc}}{E_{bb} - e_{p1} - i_p R_a} \end{aligned}$$

Negative feedback may be added as indicated by the dashed resistor $R_f$ in Fig. 8-3, to improve stability. Because of large values of $\mu$ available in many tubes, the gain of the circuit can be large even though part of this gain is sacrificed to place the desired grid and plate voltages on the tube. In a circuit such as shown in the figure the $i_p R_a$ drop might be equal to one-half of $E_{bb}$. Suitable values of $R_1$ and $R_2$ then result to place point $B$ at ground potential with zero signal. The circuit is frequently employed.

Figure 8-4 shows a variation of the previous circuit, replacing $R_1$ with a gas-discharge tube of the voltage-regulator type. When sufficient cur-

rent is flowing through such a tube, it tends to maintain a constant voltage between its electrodes. Thus the grid of $T_2$ may be dropped to the desired d-c level by the d-c drop in the neon tube. Since the voltage drop across the neon bulb is constant, any *changes* in potential at the anode of $T_1$ will also appear across $R_2$ and thus there is no loss in gain due to the voltage-divider action of $R_2/(R_1 + R_2)$ as in the previous circuit. Hence the gain is greater, but this is obtained at the expense of some noise due to random and minute voltage fluctuations in the voltage-regulator tube. Therefore the circuit is not well suited to the amplification of very small input voltages.

**8-3. Cathode Coupling in D-C Amplifiers.** Another type of circuit often found as an element of d-c amplifiers is the cathode-coupled amplifier of Fig. 8-5. Essentially, it is a cathode follower placed ahead of a grounded-grid amplifier and introduces d-c coupling and considerable gain without serious circuit complications. Assuming the use of a double triode of identical parameters, from Fig. 8-5 the circuit equations are

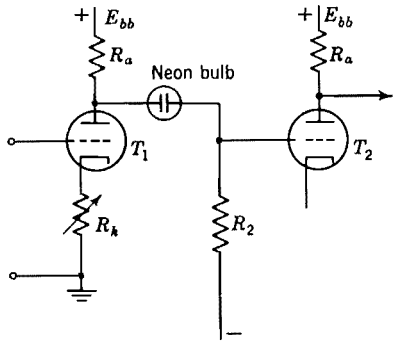



FIG. 8-4. Use of a gas-discharge tube in a resistance-coupled d-c amplifier.

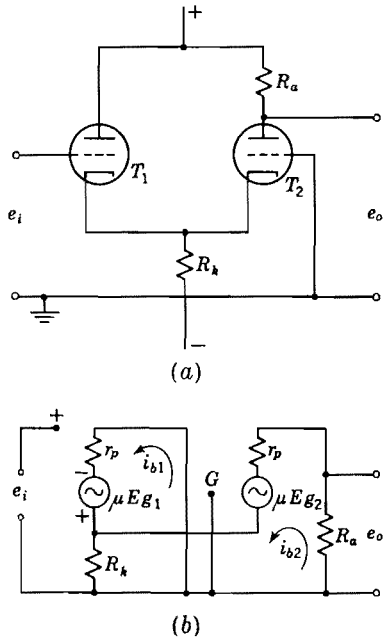



FIG. 8-5. The cathode-coupled circuit and its equivalent.

$$e_{g1} = e_i - (i_{b1} + i_{b2})R_k \tag{8-3}$$

$$e_{g2} = -(i_{b1} + i_{b2})R_k \tag{8-4}$$

$$\mu e_{g1} = i_{b1}(r_p + R_k) + i_{b2}R_k \tag{8-5}$$

$$\mu e_{g2} = i_{b1}R_k + i_{b2}(r_p + R_a + R_k) \tag{8-6}$$

from which

$$i_{b1} = -i_{b2}\frac{r_p + R_a + (\mu + 1)R_k}{(\mu + 1)R_k}$$

and $\mu(\mu + 1)R_k e_i = -i_{b2}\{[r_p + (\mu + 1)R_k][r_p + R_a + (\mu + 1)R_k] - (\mu + 1)^2 R_k^2\}$

The gain of the circuit is

$$A = \frac{e_o}{e_i} = \frac{-i_{b2}R_a}{e_i} = \frac{\mu R_a}{[(r_p + R_a)/(\mu + 1)](r_p/R_k) + 2r_p + R_a} \tag{8-7}$$

For usual circuit and tube parameters the first term in the denominator is small and may be dropped, giving

$$A = \frac{\mu R_a}{2r_p + R_a} \tag{8-8}$$

Amplification is achieved, and again it is possible to use common anode-voltage supplies. Two triodes are required, but the circuit is simple and frequently employed in instruments where the stability, introduced by the negative current feedback from the cathode resistor, is desirable.

The principle is further demonstrated in Fig. 8-6.

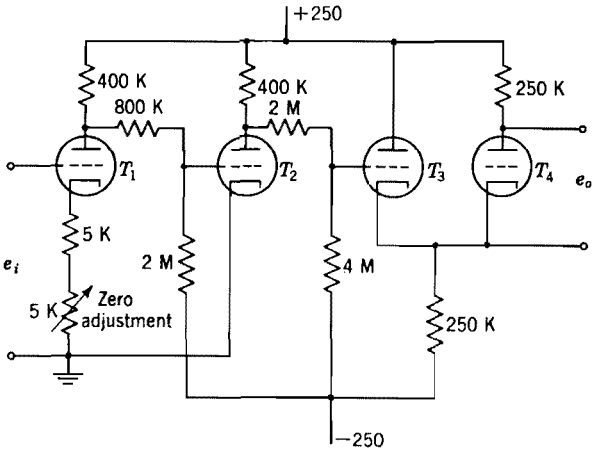



FIG. 8-6. A complete d-c amplifier employing resistance coupling and cathode coupling.

**8-4. The Balanced Push-pull Amplifier.** A balanced or push-pull type of operation is sometimes employed in an attempt to overcome zero drifts in d-c amplifiers due to changing anode-supply and heater voltages. In the circuit of Fig. 8-7 tubes $T_1$ and $T_2$ constitute a phase inverter to invert the single-ended, or grounded, input voltage into two voltages, $e_1$ and $e_2$, balanced to ground. Tubes $T_3$ and $T_4$ form a resistance-coupled and balanced d-c amplifier stage. By use of the zero adjustment in the first stage the output voltage $e_o$ can be made zero for zero input $e_i$.

The circuit has the disadvantage that both output terminals are above

ground potential. If further stages are to follow, or if the output is to be supplied to an indicating instrument or the plates of a cathode-ray oscillograph, this is not important.

If it be assumed that all tubes have matched characteristics, then it may be reasoned that the output will be independent of supply-voltage changes. Actually this is a condition difficult to achieve because of tube dissimilarity, but some reduction of drift is realized. The circuit does not give protection from individual tube parameter changes.

A simpler version of a balanced d-c amplifier, sometimes called a *difference amplifier*, is derived from the cathode-coupled phase inverter of Sec. 6-19 and is related to the circuit of Fig. 8-5. This is shown in Fig. 8-8, where $M$ is an indicating meter for showing amplifier output,

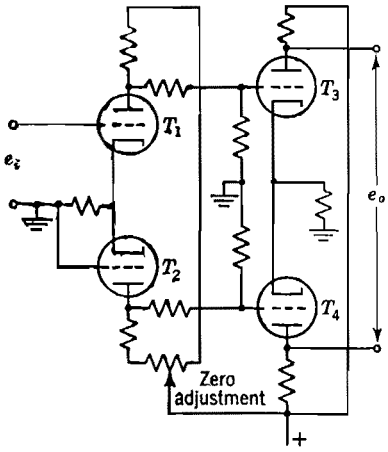



FIG. 8-7. A push-pull d-c amplifier.

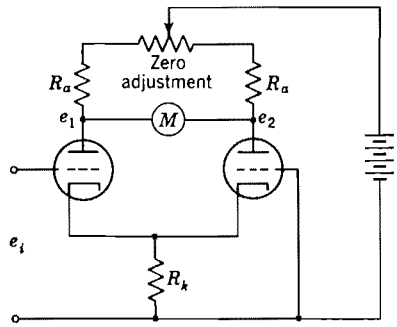



FIG. 8-8. A balanced circuit employing an indicating instrument.

since the circuit is commonly employed in vacuum-tube voltmeters. By use of the adjustment the instrument can be brought to zero for zero input voltage.

The analysis leading to Eqs. (6-60) and (6-61) provides expressions for the tube currents. If the two plate-load resistors are made equal to $R_a$, then the currents may be written as

$$i_1 = \frac{\mu e_s[R_a + r_p + (\mu + 1)R_k]}{2(R_a + r_p)[1 + (\mu + 1)R_k]} \tag{8-9}$$

$$i_2 = \frac{-\mu(\mu + 1)e_s R_k}{2(R_a + r_p)[1 + (\mu + 1)R_k]} \tag{8-10}$$

Referring to Fig. 8-8, the potentials at $e_2$ and $e_1$ are

$$e_2 = E_{bb} - i_2 R_a$$
$$e_1 = E_{bb} - i_1 R_a$$

and the potential applied to $M$ is

$$e_M = e_2 - e_1 = (i_1 - i_2)R_a \tag{8-11}$$

and use of the currents leads to

$$e_M = \frac{\mu e_s R_a[R_a + r_p + 2(\mu + 1)R_k]}{2(R_a + r_p)[1 + (\mu + 1)R_k]}$$

Circuit design and choice of a reasonable value of $\mu$ will lead to the condition that $(R_a + r_p) \ll (\mu + 1)R_k$, and the above reduces to

$$e_M = \frac{\mu e_s R_a}{R_a + r_p} \tag{8-12}$$

as a simple result, neglecting the loading effect of the instrument current. With matched tubes, changes in plate-supply potential have no effect on the instrument readings. Since a double triode is used, changes in heater potential are also reasonably balanced and the circuit is frequently employed in indicating instruments.

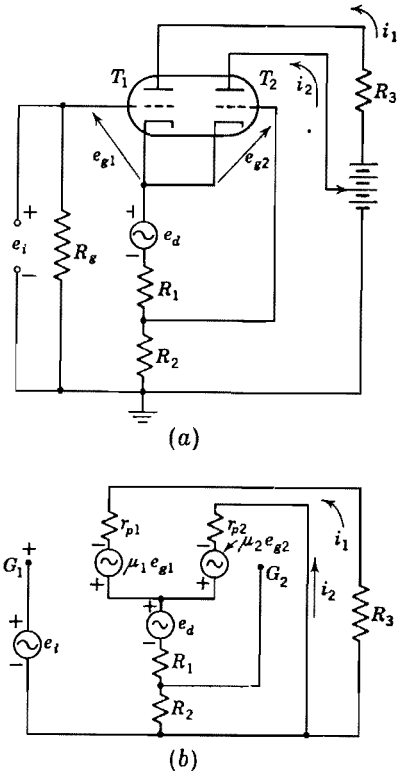



FIG. 8-9. Analysis of Miller circuit for cathode-drift compensation.

**8-5. Cathode-drift Compensation.** One of the problems associated with the operation of sensitive d-c amplifiers is the erratic drift in output voltage caused by random changes in cathode emission. The latter changes may be due to small heater-voltage changes or to actual changes in the internal emission mechanisms. The effect is usually considered as located in the first tube, since at that point any drift is multiplied by the whole gain of the amplifier and is therefore more serious than when it occurs in following tubes.

The Miller compensation circuit of Fig. 8-9 provides a means of overcoming cathode drift in a d-c amplifier by use of an additional triode. Usually a double triode is employed, and it is necessary to assume that both cathodes have identical drift characteristics and that the drift can be simulated, for analysis purposes, by an equivalent random-drift gener-

ator $e_d$ in the common cathode lead of the two tubes. It may then be considered that the second tube amplifies the drift voltage and feeds it back in opposition to the drift.

Using the equivalent circuit of Fig. 8-9*b*, an analysis may be made to determine the conditions which must be met for drift cancellation. Writing mesh equations with no external signal applied,

$$e_{g1} = -(i_1 + i_2)(R_1 + R_2) - e_d + e_i \tag{8-13}$$

$$e_{g2} = -(i_1 + i_2)R_1 - e_d \tag{8-14}$$

$$i_1(r_{p1} + R_1 + R_2 + R_3) + i_2(R_1 + R_2) + e_d = \mu_1 e_{g1}$$

$$i_1(R_1 + R_2) + i_2(r_{p2} + R_1 + R_2) + e_d = \mu_2 e_{g2}$$

Solving for $i_1$,

$$i_1 = \frac{\mu_1 e_i[r_{p2} + R_2 + (\mu_2 + 1)R_1] + (\mu_1 + 1)e_d(\mu_2 R_2 - r_{p2})}{(r_{p1} + R_3)[r_{p2} + R_2 + (\mu_2 + 1)R_1] + (\mu_1 + 1)r_{p2}(R_1 + R_2)} \tag{8-15}$$

It is desired that the current $i_1$ be independent of changes in the drift voltage $e_d$, so that $\partial i_1/\partial e_d = 0$ is obtained. Therefore

$$\frac{\partial i_1}{\partial e_d} = \frac{(\mu_1 + 1)(\mu_2 R_2 - r_{p2})}{\text{denominator}} = 0$$

and this leads to the condition that

$$R_2 = \frac{1}{g_{m2}} \tag{8-16}$$

for cancellation of the cathode drift, where $g_{m2}$ is the transconductance of the second, or compensation, triode. This condition is easily complied with.

If the tube is a double triode and $\mu_1 = \mu_2$, $r_{p1} = r_{p2}$, and if Eq. (8-16) is satisfied, or $g_{m2}R_2 = 1$, the gain expression simplifies to

$$A = \frac{-i_1 R_3}{e_i} = \frac{-\mu R_3}{2r_p + R_3} \tag{8-17}$$

showing the presence of current feedback, by reason of the increase of $r_p$ over that of one tube.

For d-c amplifiers having extremely small input signals, this method of compensation is helpful. By its use heater-voltage variations can be compensated, the residual change in $i_1$ being less than 0.5 per cent. This residue appears because $g_m$ varies somewhat with available electron emission, and this may make the compensating resistor $R_2$ incorrect for an extreme range of heater voltages.

**8-6. Series-cathode Compensation; the Bridge-balanced Amplifier.** The Miller compensator may be considered as a parallel method of compensation for heater and emission variations. Compensation can also be

obtained through use of an identical tube in series, resulting in a *bridge-balanced* d-c amplifier, as shown in Fig. 8-10. Here $T_1$ is the amplifier and $T_2$ the compensator. The circuit appears as a bridge, considering $T_1$, $T_2$, and the two resistors $R$ as the bridge arms.

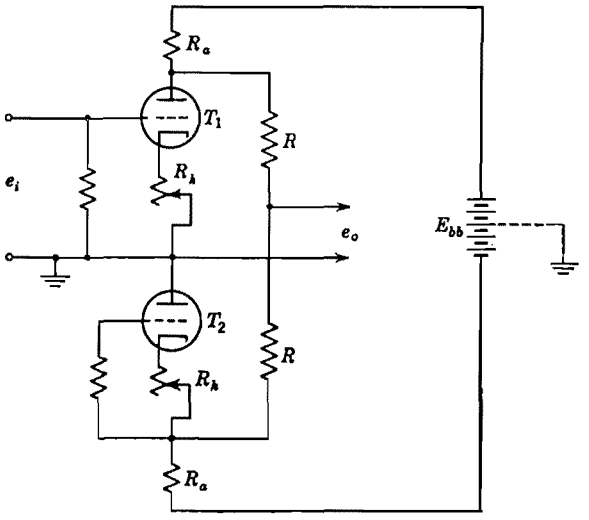



FIG. 8-10. Bridge amplifier using series compensation for cathode drift.

The output can be brought to zero volts for zero input by use of the variable cathode resistors. The output is at ground potential and may swing either positive or negative depending on the input $e_i$. If it be assumed that heater changes occur similarly in both tubes, the bridge will remain balanced or compensated for heater changes. For this reason double triodes are frequently used. Voltage-supply changes are likewise compensated if the positive and negative supplies do not have a grounded center tap or are floating. However, if more than one stage is to be operated in cascade from the same supply, a fixed center tap is required and then each supply should be well regulated.

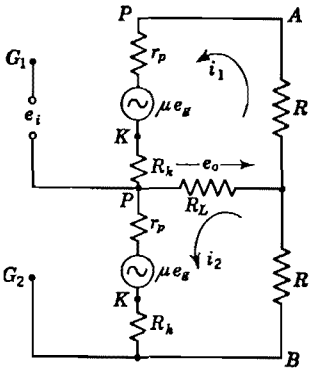



FIG. 8-11. Equivalent circuit for the bridge amplifier.

If a load $R_L$ is connected across the output, the equivalent circuit becomes that of Fig. 8-11, with the anode supply floating. Because of the balanced nature of the circuit the potential difference across $A$-$B$ is constant, and resistors $R$ do not enter into

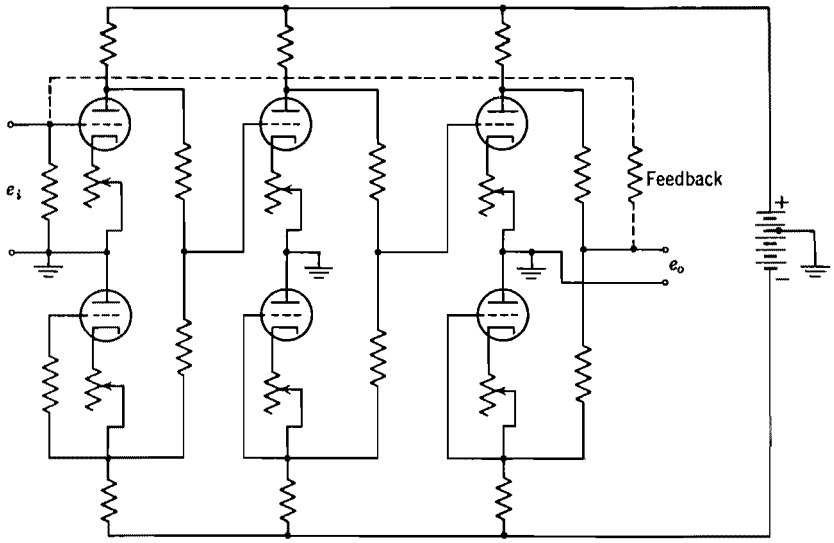



FIG. 8-12. Multistage bridge amplifier with negative feedback.

the varying current equivalent circuit. Writing mesh equations for identical tubes gives

$$e_{g1} = e_i - i_1 R_k \tag{8-18}$$

$$e_{g2} = -i_2 R_k \tag{8-19}$$

$$\mu e_{g1} = i_1(r_p + R_k + R + R_L) - i_2 R_L \tag{8-20}$$

$$\mu e_{g2} = i_2(r_p + R_k + R + R_L) - i_1 R_L \tag{8-21}$$

Solving for $i_1$ and $i_2$ and noting that

$$e_o = (i_2 - i_1)R_L$$

then

$$e_o = \frac{\mu e_i R_L\{R_L - [r_p + (\mu + 1)R_k + R + R_L]\}}{[r_p + (\mu + 1)R_k + R + R_L]^2 - R_L^2}$$

and the gain is

$$A = \frac{e_o}{e_i} = \frac{-\mu R_L}{r_p + (\mu + 1)R_k + R + 2R_L} \tag{8-22}$$

If the amplifier is employed to excite a following stage, as in Fig. 8-12, $R_L$ may be considered very large and the gain then reduces to

$$A \cong -\frac{\mu}{2} \tag{8-23}$$

When the anode supply is operated with grounded center tap, the resistors $R$ become part of the equivalent circuit and it is possible to

obtain the gain as

$$A = \frac{-\mu R}{2[r_p + (\mu + 1)R_k + R]} \tag{8-24}$$

for $R_L$ large. This is just one-half of the gain of a conventional cathode-feedback amplifier. The circuit is easy to adjust to zero and has considerable value for compensating slow cathode drifts.

When pentodes are employed in the circuit of Fig. 8-10, with stabilized screen voltages, Eq. (8-23) shows that the gain may be very high because of the high $\mu$ of the pentode, and values of 4000 can be reached.

The circuit of Fig. 8-12 illustrates three cascaded stages using bridge stabilization of zero. In addition the dashed connection may be employed for negative feedback to stabilize the gain. Such circuits are very insensitive to supply-voltage changes but are expensive because of the large number of tubes and components involved.

**8-7. Vacuum-tube Electrometers.** Some of the first applications of d-c amplifiers were as electrometers, to measure electrostatic potentials or currents of the order of $10^{-13}$ amp or less. The vacuum tube is a voltage-operated device; to produce an appreciable potential from such small currents requires that they pass through an extremely high resistance, and values up to $10^{14}$ ohms are used. It was quickly found that the grid currents in the usual tubes, even with quite negative grids, might greatly exceed the magnitude of the current to be measured, owing to gas ions formed by the tube high voltages, photoelectric emission from the grid, and leakage currents across the insulating surfaces of the tube.

Special electrometer tubes were developed, employing low electrode potentials and filamentary cathodes which operated at low temperatures and gave off almost no light to excite photoelectric emission, and with external leakage minimized by surface treatments of the glass.

Electrometer tubes usually are employed as tetrodes, with the grid nearest the cathode operated at a low positive potential, forming a virtual cathode or electron source in its region. The positive potential also repels positive ions emitted by the cathode and prevents them from reaching the second, or control, grid. The control grid is operated negatively and the plate at a positive potential of a few volts, below the ionizing potential of any residual gas. The amplification factor is about unity, the tube serving not as a voltage amplifier but as an impedance transformer, to step down the very high impedance of the grid circuit to that of the indicating instrument. By these methods grid currents as indicated in Fig. 8-13 may be obtained.

A typical circuit for an electrometer tube is shown in Fig. 8-14, all potentials being supplied by one battery source. Resistor $R$ is the high-value input resistor, which must be carefully protected from dirt, mois-

ture, and other contamination if surface leakage is not to reduce its value. The indicating microammeter $M$ is in a *bucking circuit*, whereby its current may be brought to zero for zero voltage input. Resistors $R_3$ and $R_4$, $R_2$, and $R_1$ and $R_x$ constitute three arms of a bridge, with the tube as a fourth arm. Adjustment of $R_x$ for zero input voltage balances the bridge and brings the meter $M$ to zero. Thus an instrument having a full-scale rating much less than the static plate current may be used, the meter reading only the change in current and increasing the circuit sensitivity.

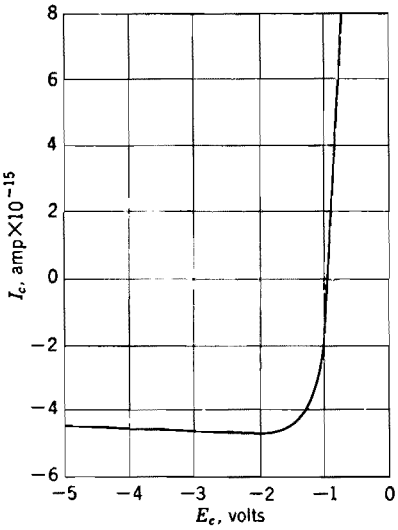



FIG. 8-13. Control-grid current in an electrometer tube; $E_b = 6$ volts.

Since the output of such a circuit is current, as indicated on an instrument, the over-all circuit transconductance, or milliamperes output change per input grid volts change, is of importance. For an amplifier with an indicating instrument of resistance $R_M$ in its anode circuit, the plate current as obtained from the

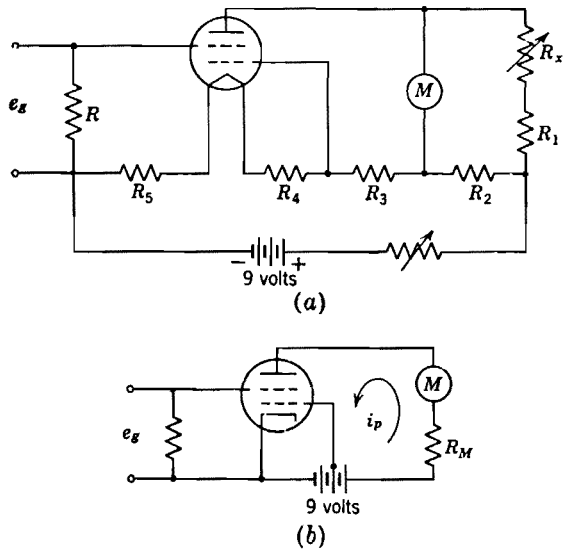



FIG. 8-14. (*a*) Electrometer tube circuit. (*b*) Equivalent for (*a*).

equivalent circuit of Fig. 8-14*b* is

$$i_p = \frac{\mu e_g}{r_p + R_M}$$

The circuit transconductance is then

$$g_c = \frac{i_p}{e_g} = \frac{\mu}{r_p + R_M} \tag{8-25}$$

and if $R_M$ is small, the circuit transconductance is just that of the grid-plate transconductance of the tube. Large values of this quantity, along with very minute grid currents, are tube requirements for this service.

**8-8. Residual Drift in Direct-coupled Amplifiers.** Direct-coupled amplifiers can be reasonably well stabilized against the types of output drift due to varying supply voltages, either by the methods of the preceding sections or by use of well-regulated power supplies. There remains, however, the effect of any internal tube changes of emission or cathode characteristics, contact potentials, or tube parameters, which may produce a shift in the $Q$ point of an amplifier, as well as a change in its voltage-current characteristics. Such changes frequently accompany the general aging of a tube and are very low-frequency phenomena, far below the low-frequency cutoff of a resistance-capacitance-coupled amplifier, and thus are not noted in the output of such a non-direct-coupled amplifier. However, a change in $Q$ point is immediately reflected as a change in the d-c output values of a direct-coupled amplifier having a pass band to zero frequency, and the general phenomenon is known as *residual drift.*

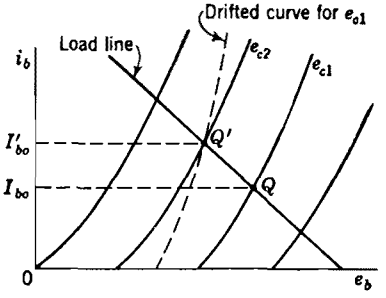



FIG. 8-15. Change in $Q$ point due to tube aging.

Consider Fig. 8-15, wherein the solid characteristics are those of the new tube, the dashed curve being one for bias $e_{c1}$ after an aging process has taken place. Initial operation, with a bias $e_{c1}$, was at point $Q$ on the solid curve. At a later time in the life of the tube this solid curve becomes altered to the dashed curve by reason of aging of some internal tube element or part. With bias still at $e_{c1}$ the operating point is now on the load line at $Q'$ with output current $I'_{bo}$.

It can be assumed, as far as static change in output current is concerned, that the tube did not change but instead the bias was merely changed from $e_{c1}$ to $e_{c2}$. Thus amplifier drift is equivalent to insertion of an extra voltage source in the grid circuit, and it is possible to express a drift figure for an amplifier in terms of *equivalent grid volts change per*

*hour of operation.* The drift figure is that grid voltage required to restore the output current to its initial value.

In feedback direct-coupled amplifiers this drift may be likened to noise or a perturbation produced within the feedback loop. The analysis of Sec. 7-4 may be used, wherein Eq. (7-9) showed

$$\mathbf{E}_o' = \frac{\mathbf{A}}{1 - \mathbf{A}\beta}\left(\mathbf{E}_s + \frac{N}{\mathbf{A}_1}\right)$$

where $\mathbf{A} = \mathbf{A}_1\mathbf{A}_2$, the gain before and after the point of noise introduction, respectively. Since the drift or noise equivalent voltage is assumed as introduced at the input, the gain $A_1$ of the previous analysis becomes unity and the output voltage with feedback is

$$e_o = \frac{A}{1 - A\beta}(e_s + e_d) \tag{8-26}$$

where $e_s$ is the input signal and $e_d$ is the equivalent drift voltage. Thus both these inputs are treated alike, and feedback does not discriminate in favor of the desired signal.

For some years output drift due to this residual cause was compensated by adjustment of manual controls on input bias. However, several more automatic systems are available and will be discussed in succeeding sections.

**8-9. Chopper-modulated Amplifiers.** Mechanical *choppers*, or synchronous switches, have been developed to interrupt or modulate a low-level direct current or very low-frequency input, giving an alternating voltage of amplitude proportional to the input direct current or slowly varying signal. The alternating voltage is then amplified by a conventional a-c signal amplifier and its output used either in a-c form or rectified back to d-c again. For amplification of slowly changing signals (thus requiring narrow bandwidth) such a system is satisfactory and will reduce drift problems to zero.

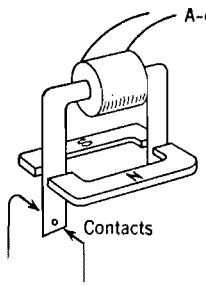



FIG. 8-16. Elements of a chopper.

A diagrammatic sketch of a chopper is shown in Fig. 8-16. A magnetic circuit is polarized by a permanent magnet and an a-c flux superimposed. This latter flux saturates the magnetic circuit, so that change in flux occurs only near the cycle zero points. The vibrating contact alternates between contacts on each half cycle of the a-c excitation, and owing to the saturation its switching operation occurs close to the cycle

zero. While 60 cycles is the usual excitation frequency, some choppers are made to operate at frequencies of 400 to 1000 cycles.

Figure 8-17*a* shows a series form of chopper amplifier which supplies an a-c input to an a-c amplifier by alternately connecting the amplifier

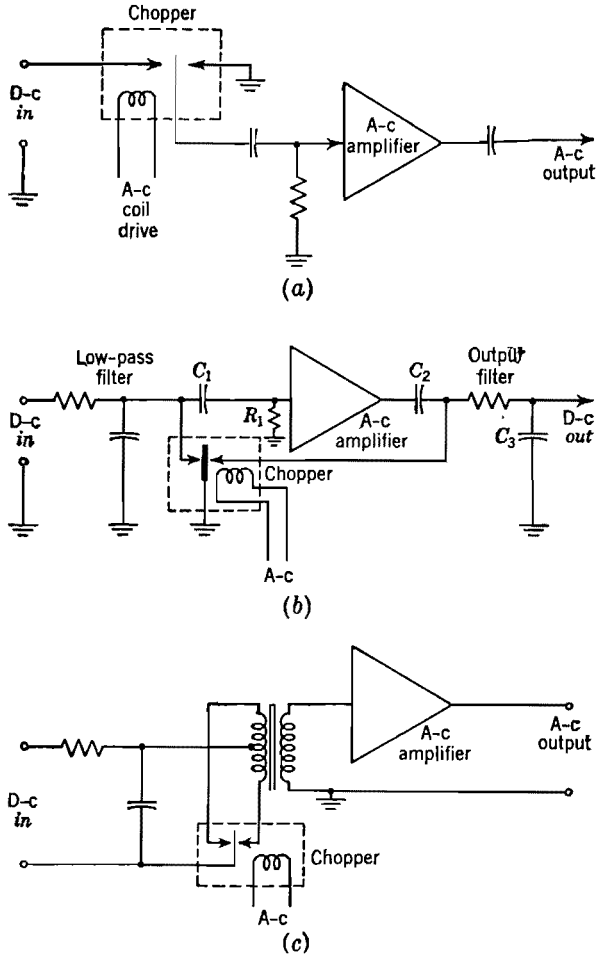



FIG. 8-17. (*a*) Simple chopper amplifier with a-c output. (*b*) Synchronous chopper with d-c output. (*c*) Chopper amplifier with input transformer and a-c output.

input to the d-c source and to ground. The amplifier output would be in a-c form unless rectified by a diode or other device.

The circuit of Fig. 8-17*b* removes the need for the additional rectifier diode by use of a shunt chopper which synchronously rectifies the amplifier output back to direct current. For analysis of operation assume that

the chopper reed has been on the right contact. Capacitor $C_1$ has then been charged through $R_1$ to the potential of the d-c source. When the reed returns to the left, capacitor $C_1$ is grounded and discharged, causing a current flow through $R_1$ and applying a voltage as input to the amplifier. At the same time capacitor $C_2$ is discharging from the amplifier output through the output filter.

When the reed returns to the right contact, capacitor $C_1$ charges from the d-c signal source, producing an input of opposite polarity on the amplifier. At the same time the output capacitor $C_2$ is grounded, so that no charge is contributed to $C_3$. Thus an a-c input is supplied to the amplifier, but pulses of unidirectional charge are supplied to the output capacitor $C_3$ on alternate half cycles, and the charge on $C_3$ does not reverse and the output appears as a rectified direct current.

If this output direct current is to be distortionless, or is to have a value proportional to the input voltage at all times, all higher-frequency components introduced at the switching frequency or its harmonics must be eliminated. This is the purpose of the output filter. To ensure proper filtering, its time constant must be long with respect to a period of the switch, and its high-frequency cutoff may need to be of the order of one-tenth to one-fifth of the switching rate. This imposes a fundamental limit on the rate of variation which may appear in the input to the amplifier, or on the allowable bandwidth which may be handled. Sixty cycles is a common frequency of chopper operation, and this basic filter requirement places the passband width in the range of 0 to 6 cycles. Thus in obtaining a solution to the problem of drift it has been necessary to sacrifice bandwidth.

In some applications this is not a serious limit, as in the amplification of d-c voltages from thermocouples measuring furnace temperatures, and many such chopper amplifiers are in use in industrial temperature recorders.

Another major difficulty in the operation of choppers at low d-c inputs is pickup from the a-c operating coil, this pickup voltage being amplified like a signal. If considerable gain is built into the amplifier, the hum voltages may overload the later stages and cause faulty operation, even though the output filter will remove the hum from the amplifier output. Narrow band response is an aid in eliminating this interference, but isolation of the a-c operating coil and leads from the low-level input circuit is very important.

Another variation of the input circuit appears in Fig. 8-17*c*, where the chopper operates in the primary of an input transformer and rectification is either not used or can be supplied by diodes. Input-transformer isolation and shielding are very important if hum and noise pickup are not to be serious with low-level inputs. Choice of circuits in Fig. 8-17*a* to *c*

is partly dependent on the source impedance, those in Fig. 8-17*a* and *b* being better adapted to high-impedance sources, whereas that in Fig. 8-17*c* can be used to match a low-impedance source to the high impedance of the amplifier input.

Where a wider band of frequencies must be amplified, it is possible to use a higher chopping frequency, but 400 to 1000 cycles represents about the limit of mechanical choppers. For ranges of zero to still higher frequencies, up to several hundred thousand cycles, a balanced bridge of diode rectifiers, called a *varistor*, may be used as in Fig. 8-18. In theory, when alternating current is applied through transformer $T$ to the bridge diodes, all four diodes conduct on one half cycle, thus shorting the grid to ground. All four diodes appear as an open circuit on the other half cycle, thus applying full signal input to the grid of the tube.

If the diode internal resistances went to zero when conducting, the triode would be driven by a voltage varying from the signal instantaneous value to zero. Since the diodes do not represent zero resistance, resistor $R$ is inserted in series with the d-c input and the ratio of internal bridge resistance to $R$ fixes the fraction of the input voltage which is transmitted during the conducting interval. The output has the same carrier or center frequency as that of the alternating current applied to the bridge through $T$, and because this may be several hundred thousand cycles, the input bandwidth handled may approximate zero to 10,000 cycles.

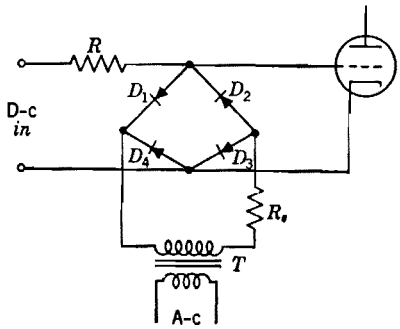


FIG. 8-18. Diode chopper or modulator.

Since the diodes may drift in internal-resistance value, this method is not drift-free and the drift problem is merely transferred from the amplifier input tube to the varistor chopping device.

**8-10. Chopper-stabilized Direct-coupled Amplifiers.** Possibilities of further improvement in drift compensation are suggested by two of the previous discussions. First, Eq. (7-9) shows that noise in a feedback amplifier is reduced in proportion to the gain preceding the point of noise introduction. Considering drift as a "noise" voltage, placement of driftless gain ahead of the point in the amplifier at which drift originates will reduce drift. Second, if it be remembered that drift is a *very* low-frequency phenomenon, then low-frequency gain may be added ahead of the first tube of a direct-coupled amplifier by use of a chopper modulator, with its accompanying narrow bandwidth, the high-frequency components being permitted to pass directly through $C_3$ into the direct-coupled

amplifier with its wider pass band as in the *chopper-stabilized amplifier* of Fig. 8-19.

As shown in this figure, any output voltage due to drift in the d-c amplifier is fed back through $R_2$ and being of very low frequency is

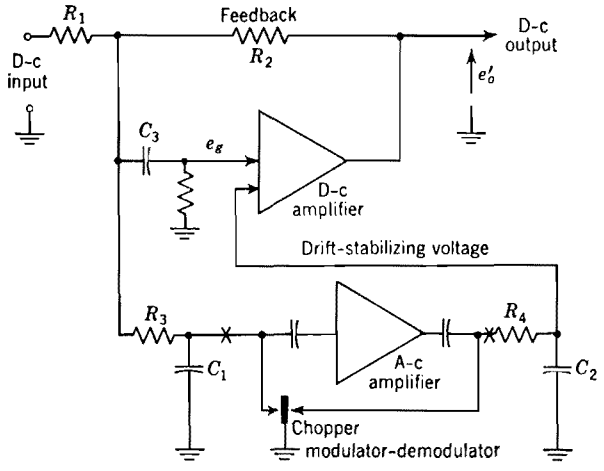



FIG. 8-19. Chopper-stabilized d-c amplifier.

effective in charging $C_1$ through the input filter network $R_3$, $C_1$, of time constant long with respect to the time of a cycle of the chopper frequency. Voltage appearing on $C_1$ is chopped, amplified, and synchronously rectified by the chopper. The output is filtered and the chopper frequency components removed by $R_4$, $C_2$, of long time constant, and applied to the input of the d-c amplifier as a *drift-stabilizing* voltage. This voltage cancels the equivalent grid drift voltage and annuls the drift output voltage to zero. A method of introducing the annulling voltage into the d-c amplifier input is shown in Fig. 8-20, wherein cathode coupling is employed in the first stage of the d-c amplifier to add the two inputs.

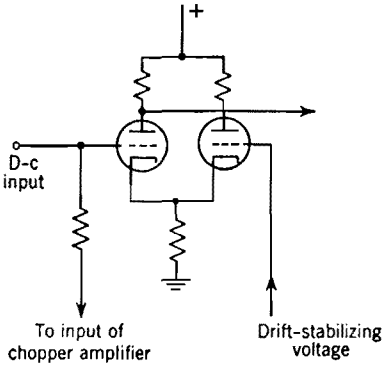



FIG. 8-20. Method of introducing low-frequency drift components into a d-c amplifier.

The capacitor $C_3$ is not always required, because even though the drift low-frequency voltage also appears in $e_g$ along with the higher-frequency components it is insignificant in its effect. If the gain of the a-c amplifier is $A_2$ and large, the drift-stabilizing voltage is $A_2$ times the original drift

voltage; thus the drift component directly introduced is negligible with respect to the drift voltage times $A_2$.

This may be further explained. If $e_y$ is the input representative of a certain amount of drift in the d-c amplifier, then the output voltage $e_o$, without the stabilizing amplifier present, is

$$e_o = -A_1 e_y$$

where $A_1$ is the gain of the d-c amplifier. However, with an input circuit as in Fig. 8-19 with an a-c amplifier gain of $A_2$, the output voltage with the a-c amplifier operating is

$$e_o' = -A_1 A_2 e_g' - A_1 e_g'$$

where $e_g'$ is the input with the stabilizing amplifier operating. To understand this statement, it must be remembered that the two gains $A_1$ and

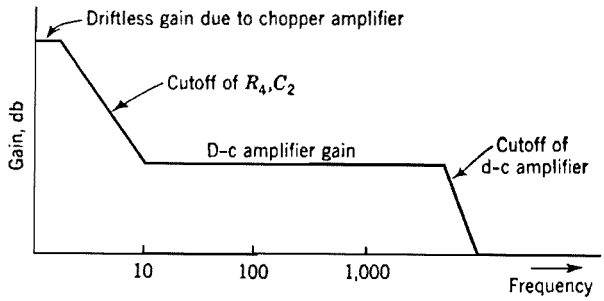



FIG. 8-21. Gain-frequency plot for a chopper-stabilized amplifier.

$A_2$ have different frequency-response characteristics due to chopper limitations, and that a low frequency, such as a drift signal, is amplified by the two amplifiers in series giving the $-A_1A_2$ gain, while a higher-frequency term is amplified by the d-c amplifier only, giving only a gain of $A_1$.

In effect this circuit places the a-c amplifier in series with the d-c amplifier for low frequencies. A gain-frequency plot for such an amplifier is presented in Fig. 8-21 and illustrates the extra gain provided for the low frequencies by $A_2$ of the a-c amplifier.

If the output drift is to be limited to the same amount with or without the stabilizing amplifier, then $e_o = e_o'$, and from the above equations

$$e_g' = \frac{e_g}{A_2 + 1} \tag{8-27}$$

and the input-voltage component to give the same drift is reduced by a factor of $A_2 + 1$ from its value without the chopper and a-c amplifier. That is, the equivalent drift voltage is reduced by this amount, thus indicating improved stability.

The long time constant associated with the output of the drift-stabilizing amplifier allows one such amplifier to be used to stabilize a number of d-c amplifiers. Motor-driven commutators are used to switch the a-c amplifier and chopper at points $x$, $x$ of Fig. 8-19 from one to another of several d-c amplifiers. Because of the long time constants the filters maintain their voltages over the switching cycle, and since the drift is such a slow-changing phenomenon, correction of the filter output voltage on $C_2$ is needed only at intervals. Thus cost of such stabilization for large groups of amplifiers is reduced, and amplifiers of this nature are now quite frequently employed.

**8-11. The Operational Amplifier.** Many d-c amplifiers are employed in performing mathematical operations such as summing, differentiation, and integration. When so applied, they are generally referred to as

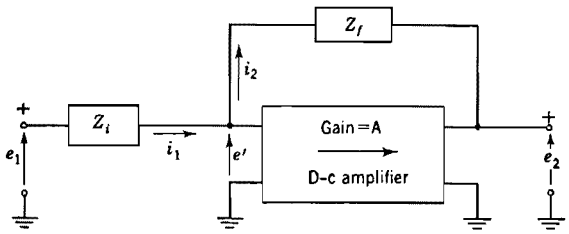



FIG. 8-22. Basic operational amplifier.

*operational amplifiers.* These operations may involve steady or slowly varying components so that d-c amplifiers are required. Almost any combination of the principles previously discussed, which leads to a linear, stable, direct-coupled amplifier having an odd number of stages or equivalent phase shift, combined with over-all feedback of the proper nature, may be employed.

Such a combination is represented by the amplifier symbol of Fig. 8-22, with gain

$$A = -\frac{e_2}{e'} \tag{8-28}$$

and having an inherent odd number of phase reversals.

The voltages and currents represented in Fig. 8-22 are instantaneous time functions and are thus related by the usual electrical laws,

$$e = iR \tag{8-29}$$

$$e = L\frac{di}{dt} = Lpi \tag{8-30}$$

$$e = \frac{1}{C}\int_{t_1}^{t_2} i\,dt = \frac{i}{Cp} \tag{8-31}$$

where the operator-$p$ notation is used once again, with $p = d/dt$. It can be readily shown that the notation $1/p$ implies the operation of integration in Eq. (8-31). Voltage-current ratios for the three basic circuit elements then become

$$\frac{e}{i} = R \qquad \frac{e}{i} = Lp \qquad \frac{e}{i} = \frac{1}{Cp}$$

depending on the type of circuit element employed. In general it is possible to refer to any of these relations as

$$\frac{e}{i} = Z(p) \tag{8-32}$$

where $Z(p)$ is called an *operational impedance*.

Analysis of performance of the circuit of Fig. 8-22 can be most general if initially the type of circuit elements contemplated for $Z_i$ and $Z_f$ not be specified, but these elements simply referred to as operational impedances $Z_i(p)$ and $Z_f(p)$ having some $e/i$ relation. After the performance analysis is completed, it will be possible to substitute the $e/i$ relation for any circuit elements or combination actually used and obtain a prediction of circuit performance.

The input voltage $e_i$ is applied through $Z_i(p)$, with over-all feedback introduced through the $Z_f(p)$ path. If current into the first grid circuit of the amplifier is negligible, then

$$i_1 = i_2$$

and, by use of Eq. (8-32), the current relations can be written as

$$\frac{e_1 - e'}{Z_i(p)} = \frac{e' - e_2}{Z_f(p)} \tag{8-33}$$

In view of the gain definition of Eq. (8-28), the input $e'$ can be eliminated and

$$-e_2\left[Z_i(p) + \frac{Z_i(p)}{A} + \frac{Z_f(p)}{A}\right] = Z_f(p)\, e_1$$

from which

$$-e_2 = \frac{Z_f(p)}{Z_i(p)} \frac{1}{1 + (1/A)[1 + Z_f(p)/Z_i(p)]} e_1 \tag{8-34}$$

If the amplifier gain $A$ is made large so that

$$A \gg 1 + \frac{Z_f(p)}{Z_i(p)}$$

then Eq. (8-34) reduces to

$$e_2 = -\frac{Z_f(p)}{Z_i(p)} e_1 \tag{8-35}$$

This is the basic equation for the operational form of amplifier. The accuracy with which the operations indicated by $Z_f(p)/Z_i(p)$ are performed is dependent on the requirement of very large amplifier gain being fulfilled. This is usually not a difficult specification to meet.

By definition and assumption the polarity of $e_2$ is always opposite to that of $e_1$, so that if $Z_f(p) = Z_i(p) = R$ (possibly 1 megohm each) then the circuit will produce an output of magnitude equal to its input, but of opposite sign, or $A = -1$. The circuit is thus always a sign changer. If $Z_f(p)$ and $Z_i(p)$ are resistive but not equal, then

$$e_2 = -\frac{R_f}{R_i} e_1 \tag{8-36}$$

and it becomes possible to obtain an output equal to the negative of the input multiplied by a constant, independent of amplifier characteristics. This usage is represented in Fig. 8-23*a*.

If $Z_f(p)$ is a capacitor, having an operational impedance $1/Cp$, and $Z_i(p)$ is a resistance, then Eq. (8-35) becomes

$$e_2 = -\frac{1}{RC}\frac{1}{p} e_1 \tag{8-37}$$

The amplifier is then an accurate integrator with respect to time, and a reversal of sign is still present. The result is multiplied by a constant $1/RC$, but if $R = 1$ megohm, $C = 1$ $\mu$f, the multiplying factor can be made unity. This application is illustrated in Fig. 8-23*b*.

The input voltage $e'$ is only $1/A$ of the output voltage $e_2$, and with $A$ very large, as required, the voltage $e'$ is negligible, and the potential at point $X$ at the amplifier input may be considered essentially zero or at ground reference. Therefore the potential across the integrating capacitor $C$ is almost equal to the output $e_2$. This is important in introducing the initial conditions into the integration process, since an initial charge on $C$ applied between output terminal and ground will represent an initial condition.

As another alternative let $Z_i(p)$ be a capacitor and $Z_f(p)$ a resistor as in Fig. 8-23*c*. Then Eq. (8-35) becomes

$$e_2 = -RCpe_1 \tag{8-38}$$

and the output of the amplifier is the negative of the derivative of the input voltage.

Another operation, that of summing several independent voltages, is introduced in Fig. 8-23*d*. As before

$$i_a + i_b + i_c = i_2$$

and if $A$ is large, making $e'$ very small with respect to either input or

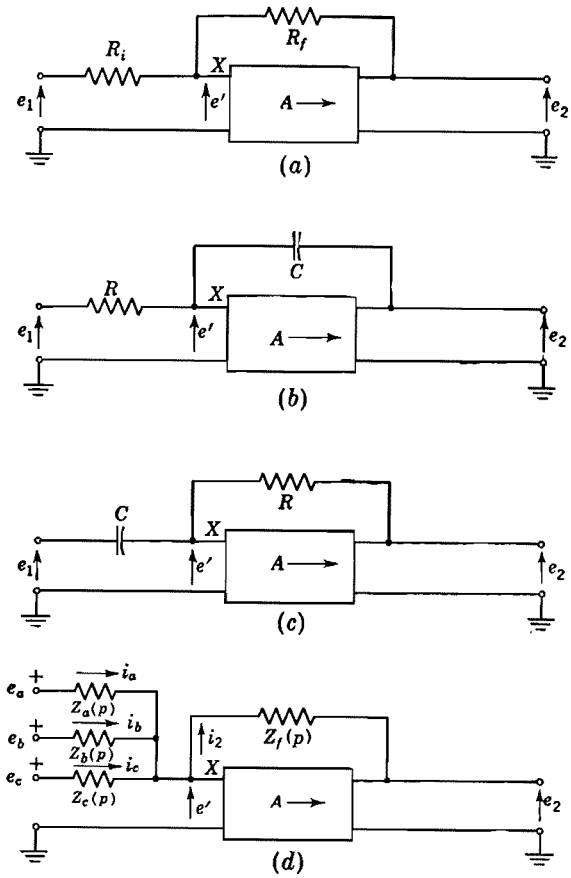



FIG. 8-23. Uses of the operational amplifier. (*a*) Sign changing and multiplication by a constant; (*b*) integration; (*c*) differentiation; (*d*) summation.

output voltages, then

$$\frac{e_a}{Z_a(p)} + \frac{e_b}{Z_b(p)} + \frac{e_c}{Z_c(p)} = \frac{-e_2}{Z_f(p)}$$

from which

$$-e_2 = \frac{Z_f(p)}{Z_a(p)} e_a + \frac{Z_f(p)}{Z_b(p)} e_b + \frac{Z_f(p)}{Z_c(p)} e_c \tag{8-39}$$

If all the impedances are equal resistors, then

$$-e_2 = e_a + e_b + e_c \tag{8-40}$$

and the output represents the summation of the inputs. This may be done for any reasonable number of input voltages.

If some negative voltages are introduced (possibly having passed through a sign-changing amplifier), then algebraic summation is possible.

In addition, if the ratios $R_f/R_a$, $R_f/R_b$, $R_f/R_c$ are not chosen equal, then multiplication by various constant factors may be done. By making $Z_f(p)$ a capacitor three integrals may be simultaneously added, or by making $Z_a(p)$ a capacitor several input variables may be added to the derivative of another variable.

If the input voltage to the integrator circuit were sinusoidal,

$$e_1 = E_m \sin \omega t$$

the output voltage would be

$$e_2 = -\frac{1}{RC} \int E_m \sin \omega t \, dt = \frac{E_m}{\omega RC} \cos \omega t$$

and for effective values

$$\mathbf{E}_2 = \frac{-\mathbf{E}_1}{j\omega RC} \tag{8-41}$$

which is most easily obtained by direct substitution of $p = j\omega$ in Eq. (8-37). It is apparent that the output of the integrating circuit falls off as $1/\omega$, or at the rate of 6 db per octave, and this offers a simple test of integrator accuracy. If at some frequency this relation fails to hold, then it is known that integration is accurate only up to this frequency.

In the case of the differentiator, if $e_1 = E_m \sin \omega t$, the output voltage would be

$$e_2 = -RC \frac{d}{dt} (E_m \sin \omega t) = -\omega RCE_m \cos \omega t$$

and for effective values

$$\mathbf{E}_2 = -j\omega RCE_1$$

which is Eq. (8-38) with $p = j\omega$. The output of a differentiator increases with $\omega$, at a rate of 6 db per octave. If $RC$ were equal to unity and $\omega = 377$ for 60 cycles, the output voltage $\mathbf{E}_2$ would be 377 times the input voltage $\mathbf{E}_1$. If there is undesired 60-cycle pickup or hum introduced into the input, the hum would appear in the output with a magnitude $\omega$ times greater. Because of the increasing gain with frequency, and the problems of noise and hum which then arise, the use of the operational amplifier as a differentiator is usually avoided. Mathematically this is not difficult to do.

It is possible to use parallel or series combinations of $R$ and $C$ as $Z_f$ or $Z_i$, and performance will be predicted by the form of the $Z_f(p)/Z_i(p)$ relation. The operational impedance of any $RC$ combination may be easily written if it be remembered that the $Z(p)$ form is obtainable from the steady-state impedance $Z$ by replacing $j\omega$ by $p$ wherever it appears. A few such circuit forms appear in Fig. 8-24 as examples.

The operational amplifier is basic to a great variety of electronic computation equipment. In application it is necessary that resistors be accurate, and capacitors of very low leakage and low retentivity of charge must be used. These usually employ polystyrene as a dielectric.

**Example.** Write the expression for the operation performed on the input voltage of the circuit in Fig. 8-24*c*. First write

$$Z_f = R_2 + \frac{1}{j\omega C}$$
$$Z_i = R_1$$

Then, replacing $j\omega$ by $p$, the operational impedances are

$$Z_f(p) = R_2 + \frac{1}{pC}$$
$$Z_i(p) = R_1$$

and the relation between $e_2$ and $e_1$ is

$$e_2 = \frac{R_2 + 1/pC}{R_1} e_1 = \frac{R_2Cp + 1}{R_1Cp} e_1$$

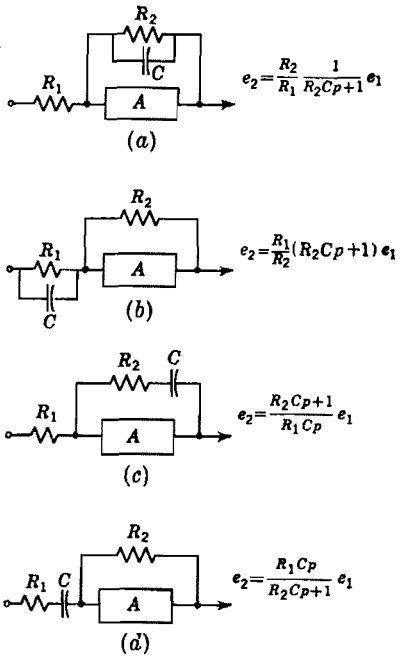



FIG. 8-24. A few operational amplifier circuits.

**8-12. Use of Positive Feedback.** The accuracy of the desired output-input relation in an operational amplifier is dependent on the assumption of very high gain inside the feedback loop. It has been found that such high gain may be easily obtained by use of positive feedback or regeneration in the amplifier, inside the overriding external negative feedback. The positive feedback may be very large; yet the overall negative feedback will maintain the stability of the amplifier.

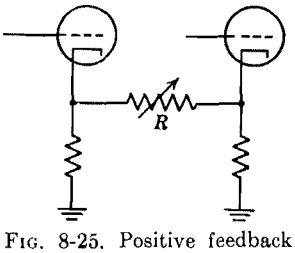



FIG. 8-25. Positive feedback inside an operational amplifier.

A circuit utilizing positive feedback between the cathodes of two successive stages, as in Fig. 8-25, is a very simple means of introducing the desired regeneration. Internal gains in the range of $10^6$ to infinity may be obtained in simple amplifier circuits and the accuracy of the functional operation made very great.

**8-13. Analog Computation.** The *analog* type of electronic computer uses electronic amplifiers and provides solutions to mathematical prob-

lems in terms of a physical quantity, such as the angular position of a shaft, a voltage or current, or the position of the index on a scale. As a result the accuracy may be limited by the precision with which these physical quantities can be read, or by the ability of a human read-out operator, if one is employed.

In contrast, the digital computer to be discussed in Chap. 9 provides solutions directly in numerical digits, as in the case of a desk calculator. The degree of accuracy is limited only by the capacity of the machine in terms of significant digits, and this is fixed largely by the cost and complexity to be tolerated. A digital computer is usually more expensive than the analog form, but when the accuracy requirements of the latter are pushed below about 0.1 per cent, the analog costs may exceed those of the digital form. Combinations of the two forms exist, in which digital output is provided for an analog computer and analog indication or curve plotting is also available for digital computers.

The applications of the analog computer lie largely in the field of engineering design or in the solution of systems of linear or nonlinear differential equations. Solutions are usually in the form of traced curves or oscillographic records of varying output voltages and currents. The basic operations in the application of such computers, or differential analyzers as they have been called, is the writing down of the differential equation or equations describing the particular physical system of interest. Of great aid in this respect is *Lagrange's equation* relating the interchange of system energies, and this will be considered in the next section.

Given the differential equations, the analog result may be obtained either by the *direct* method, in which equivalent electrical circuits are directly used to simulate the time variations of the dependent variables of the physical system, or the *functional* method, in which electronic circuits perform the mathematical operations indicated by the terms of the differential equation. The former method is well suited to design problems, especially where many degrees of freedom are involved, and the latter is best suited to a wide variety of engineering problems arising in control design and system operation.

**8-14. Lagrange's Equation for Dynamical Systems.** The starting points in most analog computation are the differential equations which describe the action of a dynamical system. In many complex physical systems, however, it is not easy to write down these differential equations. Often in engineering problems there are so many variables present that it is difficult to write the equations of the system, and these difficulties are increased by the frequent presence of both mechanical and electrical (or other) effects in the system.

Lagrange's equation provides a systematic means whereby the process is simplified. Its advantage lies in the usual ease with which expressions

for the kinetic and potential energies of a complex system may be written, the equations of motion following readily by application of Lagrange's equation. This equation, which will not be derived here, is based on the physical fact that nature, in moving from one condition of equilibrium to another, seems to do so in the manner requiring the least interchange, or transfer, of energy.

The equations of motion must employ some frame of reference or a coordinate system, and in a complicated problem there may be a number of these coordinates or independent variables, which will be referred to in general as $q_1, q_2, \ldots$, although they may bear other symbolism in actual problems. For instance, in a three-mesh electric network these $q$'s may actually appear as the mesh currents, $i_1, i_2, i_3, \ldots$, or may represent some set of node voltages. The first and second derivatives with respect to time are designated $\dot{q}$ and $\ddot{q}$ in the usual manner. The number of such independent coordinates is fixed by $n$, the number of degrees of freedom of the system, and this is equal to the number of independent equations required to describe the action of the system.

If $T$ be allowed to represent the total kinetic energy of the system, $V$ the total potential energy in terms of the selected coordinates, $D$ the dissipation of energy per second, and $f$ the applied force if present, then Lagrange's equation states the relation between these quantities as

$$\frac{d}{dt}\left[\frac{\partial(T - V)}{\partial \dot{q}_i}\right] - \frac{\partial(T - V)}{\partial q_i} + \frac{1}{2}\frac{\partial D}{\partial \dot{q}_i} = f \tag{8-42}$$

The term $q_i$ represents any of the coordinates $q_1, \ldots, q_n$, and thus there will be as many equations of the form of the above as there are independent quantities of reference.

As a simple illustration of the method consider a projectile of mass $m$ whose motion is defined by two coordinates, $x$, the distance from the point of reference, and $y$, the height above earth. The energies may be written as

$$T = \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2) \qquad V = mgy \tag{8-43}$$

Preparing the terms needed for substitution in Eq. (8-42),

$$\frac{\partial T}{\partial \dot{x}} = m\dot{x} \qquad \frac{\partial(T - V)}{\partial x} = 0$$

$$\frac{\partial T}{\partial \dot{y}} = m\dot{y} \qquad \frac{\partial(T - V)}{\partial y} = -mg$$

The dissipation $D$ and applied force $f$ are zero. Writing Lagrange's equation for each of the coordinates $x$ and $y$ gives

$$\frac{d}{dt}(m\dot{x}) - 0 = 0 \qquad \frac{d}{dt}(m\dot{y}) + mg = 0$$

from which $m\ddot{x} = 0 \qquad m\ddot{y} = -mg$

$\ddot{x} = 0 \qquad \ddot{y} = -g$

which are the usual results.

Lagrange's equation shows a greater advantage in more complicated problems, such as the condenser microphone of Fig. 8-26. Such a device consists of a stretched metallic diaphragm closely spaced in front of a solid metal backplate. The sound pressure applies a force on the diaphragm, varying the spacing and the capacity between diaphragm and

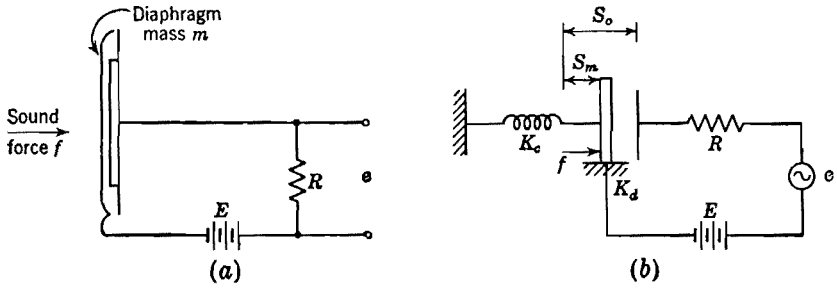



FIG. 8-26. Condenser microphone.

plate. Since the applied voltage $E$ is constant, the varying capacity requires a varying charge in accordance with

$$q = CE$$

and a current therefore flows through $R$, producing a varying output voltage $e$.

The coordinates of reference are taken as the diaphragm deflection $s$ and the charge $q$ on the capacitor. The kinetic energy of the system is then

$$T = \frac{m\dot{s}^2}{2} \tag{8-44}$$

to which would be added a term $L\dot{q}^2/2$ if inductance were present in the electric circuit. The potential energy is

$$V = \frac{k_C s^2}{2} + \frac{q^2}{2C} \tag{8-45}$$

The instantaneous value of capacity $C$ is

$$C = \frac{s_o}{s_o - s} C'$$

where $C'$ is the capacity when the movable diaphragm is at $s = 0$. Then

$$V = \frac{k_C s^2}{2} + \frac{q^2}{2C'} \frac{s_o - s}{s_o} \tag{8-46}$$

The dissipation is due to the resistance $R$ and to friction $k_d$ in bending the metal of the diaphragm. Then $D$ is

$$D = k_d\dot{s}^2 + R\dot{i}^2 = k_d\dot{s}^2 + R\dot{q}^2 \tag{8-47}$$

Taking the indicated derivatives for Lagrange's equations,

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{s}}\right) = m\ddot{s} \qquad \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}}\right) = 0$$

$$\frac{\partial(T - V)}{\partial s} = -k_C s + \frac{q^2}{2C'}\frac{s_o - s}{s_o} \qquad \frac{\partial(T - V)}{\partial q} = -\frac{q}{C'}\frac{s_o - s}{s_o}$$

$$\frac{\partial D}{\partial \dot{s}} = 2k_d\dot{s} \qquad \frac{\partial D}{\partial \dot{q}} = 2R\dot{q}$$

and combining by Eq. (8-42) leads to

$$m\ddot{s} + k_C s - \frac{q^2}{2C'}\frac{s_o - s}{s_o} + k_d\dot{s} = f \tag{8-48}$$

$$\frac{q}{C'}\frac{s_o - s}{s_o} + R\dot{q} = e \tag{8-49}$$

as the differential equations describing the action of the condenser microphone. It is seen that one equation is obtained for each of the coordinates of reference.

Lagrange's equation provides a simple tool by which system differential equations can be written, preparatory to analog solution.

**8-15. The Direct-analog Computer.** The so-called *direct analog* of a problem is arrived at by replacing the physical system under study, as described by a set of differential equations, with an electrical circuit whose performance is described by differential equations of *identical mathematical form.* It is well known that two sets of equations of the same form have identical forms of solutions, and therefore the performance of the electrical circuit is analogous to the operation of the physical system replaced. The solution, as obtained in the electrical circuit in terms of easily measured currents and voltages, is analogous to the original physical system in terms of its appropriate independent and dependent variables.

For instance, the common mass-spring-damper mechanical system of Fig. 8-27*a* may be expressed by a differential equation of the form

$$m\frac{d^2x}{dt^2} + k_d\frac{dx}{dt} + k_C x = F(t)$$

which may be otherwise written, by use of $v = dx/dt$, as

$$m\frac{dv}{dt} + k_d v + k_C\int v\,dt = F(t) \tag{8-50}$$

The system of rotation of Fig. 8-27*b* yields a similar equation

$$I\frac{d^2\theta}{dt^2} + k_d\frac{d\theta}{dt} + k_C\theta = T(t)$$

and by using $\omega = d\theta/dt$,

$$I\frac{d\omega}{dt} + k_d\omega + k_C\int \omega\, dt = T(t) \tag{8-51}$$

Both Eqs. (8-50) and (8-51) have been written under an assumption of viscous friction in the damping device, or that friction force is proportional to the first power of the velocity. This is satisfactory for a great

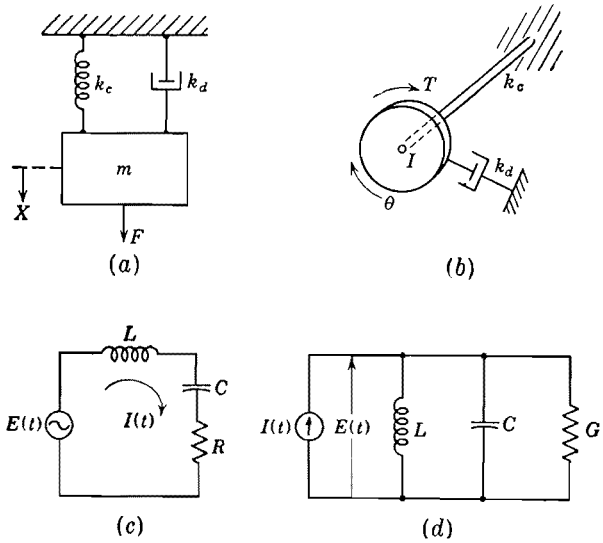



FIG. 8-27. Analogous systems.

variety of problems, and methods exist whereby this limitation on performance may be removed if necessary.

A mere change in the choice of alphabetical symbolism may lead to

$$L\frac{di}{dt} + Ri + \frac{1}{C}\int i\, dt = E(t) \tag{8-52}$$

or to

$$C\frac{de}{dt} + Ge + \frac{1}{L}\int e\, dt = I(t) \tag{8-53}$$

It is apparent that Eqs. (8-52) and (8-53) are of identical form to either Eq. (8-50) or Eq. (8-51), and to one versed in electrical terminology it is likewise apparent that Eqs. (8-52) and (8-53) describe the performance of the circuits of Fig. 8-27*c* and *d*, respectively. Thus these circuits

become the *direct analogs* of the translational or rotational systems of Fig. 8-27*a* and *b*.

The systems of Fig. 8-27*a* and *b* may be studied by use of the circuits of Fig. 8-27*c* or *d*, the choice being arbitrary and determined largely by the availability of suitably sized electric-circuit elements. In Fig. 8-27*c* the independent variable is $E(t)$, the dependent variable $I(t)$, analogous to the force $F(t)$ or torque $T(t)$ and to velocities $v$ or $\omega$, respectively. Inductance is analogous to mass, and this analog is sometimes called the *force-voltage*, or the *mass-inductance*, analog. In Fig. 8-27*d* the independent variable is $I(t)$ and the dependent variable $E(t)$, analogous to $F(t)$ or $T(t)$ and to $v$ or $\omega$, respectively. Capacitance is analogous to mass, and this form is sometimes referred to as the *mass-capacitance*, or the *force-current*, analog.

Study of $I(t)$ in Fig. 8-27*c* or $E(t)$ in Fig. 8-27*d* by oscillographic or other methods, when the circuits are excited by a voltage or current time function of form analogous to the mechanical force function, yields results which may be interpreted in terms of velocities, accelerations, and displacements in the analogous mechanical, thermal, or other physical system under study.

By use of electronic circuits it is relatively easy to produce sinusoids, pulses, saw-toothed waves, decaying exponentials, and the like, as driving sources or time functions analogous to the mechanical or other force functions which may occur in an actual system. In this connection, constant-current sources may be simulated by voltage sources of very high internal resistance or by pentodes operated in a constant-anode-current region.

TABLE 13. SUMMARY OF ANALOGOUS QUANTITIES

| Translational | Rotational | Electrical | |
|---|---|---|---|
| | | $M - I$ | $M - C$ |
| Force $f$ | Torque $T$ | $E$ | $I$ |
| Mass $m$ | Inertia $I$ | $L$ | $C$ |
| Damping $k_d$ | Damping $k_d$ | $R$ | $G$ |
| Velocity $v$ | Velocity $\omega$ | $I$ | $E$ |
| Compliance $k_C$ | Compliance $k_C$ | $1/C$ | $1/L$ |
| Displacement $x$ | Displacement $\theta$ | $q$ | $e\,dt$ |

To illustrate further the direct-analogy viewpoint, Table 13 develops the analogous translational, rotational, and electric-circuit fundamental quantities. Similar analogs in other energy systems can also be set up if desired.

Table 14 also shows how the mathematical analogs may be set up to convert system elements to circuit elements.

One difficulty is occasionally encountered in these electrical analog circuits. Some resistance is always associated with inductors and friction with any mechanical system. It should be noted, however, that the ratios of mass or inertia to unavoidable friction may be very much greater in mechanical devices than the achievable ratios of $L/R$ in electrical inductors. This leads to a difficulty in simulation and, where the situation becomes limiting in application of the method, may make advisable the employment of the mass-capacitance analog since capacitors have much lower internal losses than inductors.

TABLE 14. ANALOGOUS RELATIONS

| | | |
|---|---|---|
| Mass to inductance | $f = m\frac{dv}{dt}$ | $e = L\frac{di}{dt}$ |
| | $v = \frac{1}{m}\int f\,dt$ | $i = \frac{1}{L}\int e\,dt$ |
| | $W = \frac{mv^2}{2}$ | $W = \frac{Li^2}{2}$ |
| Spring to capacitance | $f = k_C\int v\,dt$ | $e = \frac{1}{C}\int i\,dt$ |
| | $v = \frac{1}{k_C}\frac{df}{dt}$ | $i = C\frac{de}{dt}$ |
| | $W = \frac{fx}{2} = \frac{f^2}{2k_C}$ | $W = \frac{eq}{2} = \frac{Ce^2}{2}$ |
| Damping to resistance | $f = k_d v$ | $e = Ri$ |
| | $v = \frac{f}{k_d}$ | $i = \frac{e}{R}$ |
| | $W = \int fv\,dt$ | $W = \int ei\,dt$ |
| | $P = fv$ | $P = ei$ |

The direct analog has its greatest application in design studies, wherein mass, elastance, or damping parameters may be represented directly by $L$, $C$, or $R$ components and the effects of a wide range of parameter values studied in a short time. The similarity in structure of the physical system and its electrical analog is often an aid in the use of the method, and to an experienced worker in the field will usually eliminate the necessity for writing the system equations for simple systems.

**8-16. Scale Factors.** Any system of units which is consistent within itself may be employed in the opposing analogous forms, that is, in a mechanical system pound-foot-second or meter-kilogram-second (mks) units may be employed, while in the electrical analog the mks units will probably be chosen because of general agreement on their use within this field. In actual application, then, 1 g mass may be analogous to 1 h

or 1 lb mass might be chosen as analogous to one mh if all other elements are appropriately selected.

Scale factors relating or modeling the relative magnitudes of the electrical-circuit elements must be properly selected. In mechanical systems there are usually three dimensions, mass $M$, length $L$, and time $T$, which are arbitrarily chosen as basic, or fundamental. Thus three scaling factors may be chosen as a matter of convenience, all others being derived therefrom.

As an example, the mass-inductance analogy may be scaled by setting up three scale factors $S_1$, $S_2$, and $S_3$, to yield satisfactory and reasonable electric element sizes. Then*

$$\mathcal{L} \doteq S_1 M \qquad Q \doteq S_2 L \qquad t \doteq S_3 T$$

where $\mathcal{L}$, $Q$, and $t$ are the inductance, charge, and time referring to the electric analog and $M$, $L$, and $T$ are those of the mechanical system. Employing the usual dimensional relations, the other analog scaling factors may be derived.

For instance, it will not be convenient to choose $S_2$ directly in many cases because we ordinarily do not work directly with charge. However, if current be taken as dimensionally $Q/t$ and its analog velocity as $L/T$, then

$$1 \text{ unit of current} = \frac{Q}{t} \doteq \frac{S_2 L}{S_3 T} = \frac{S_2}{S_3} \times 1 \text{ unit of velocity}$$

The factor $S_3$ may be arbitrarily chosen to provide a convenient time relationship and the above used to determine $S_2$ to give a desirable value of current in the analog.

The scale factor for voltage may be derived by use of the relation $e = \mathcal{L}\, di/dt$, giving

$$1 \text{ unit of voltage} = \frac{Q}{t^2} \doteq \frac{S_1 M S_2 L}{(S_3 T)^2} = \frac{S_1 S_2}{S_3^2} \times 1 \text{ unit of force}$$

since force has dimensions $MLT^{-2}$. Other common relations follow as in Table 15.

TABLE 15. DERIVED SCALE FACTORS FOR MECHANICAL SYSTEMS

| | |
|---|---|
| Current | $\doteq S_2 S_3^{-1} \times$ velocity |
| Voltage | $\doteq S_1 S_2 S_3^{-1} \times$ force |
| Resistance | $\doteq S_1 S_3^{-1} \times$ damping (force/velocity) |
| Capacity | $\doteq S_1^{-1} S_3^{2} \times$ compliance (displacement/force) |
| Power | $\doteq S_1 S_2^{2} S_3^{-3} \times$ mechanical power |

* The symbol $\mathcal{L}$ will be employed to refer to inductance here to avoid conflict with $L$, denoting length. Also the symbol $\doteq$ will be used to denote "analogous to."

**8-17. Functional or Operational Analog Computers.** The operational analog (differential analyzer) employs electrical circuits to perform the mathematical operations called for by the integrodifferential equations of the dynamical system. Operational amplifiers of the form described

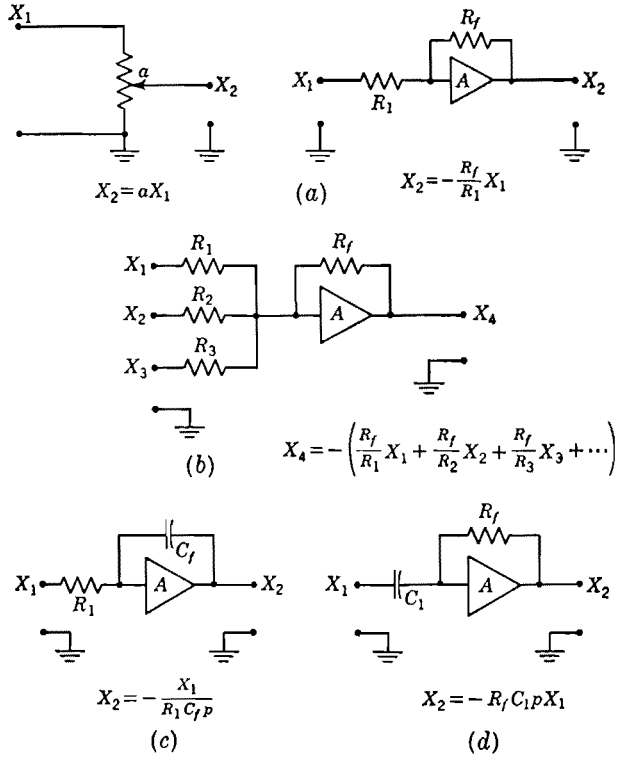



FIG. 8-28. Methods of using operational amplifiers in computation.

previously in this chapter are employed in the manner indicated in Fig. 8-28, to:

1. Form the sum of two or more variables.
2. Multiply a variable by a constant.
3. Integrate a variable against time.
4. Take the time derivative of a variable.

Units performing these operations are connected by cords and plugs in the sequence required by the particular problem. In the *slow* form of computer the process of obtaining the solution may take several minutes to accommodate servomechanism operation, and the solution appears as a plotted curve on a graph sheet on a plotting table. In the *fast*, or repetitive, computer the time scale is speeded, and the input

function and initial conditions applied 30 to 60 times per second, with the output appearing as a trace on a cathode-ray oscillograph.

These two types differ largely in accuracy, since a cathode-ray display is only 5 to 10 per cent accurate unless calibrated at the time of use. It is customary to draw the coordinate axes for the slow-computer plot directly with computer output, thus eliminating calibration errors that may exist.

The operation of differentiation is usually avoided by prior manipulation of the equations, since any hum or other noise present is increased in importance by a factor of angular frequency. Integration does not so suffer, noise being in effect reduced by $\omega$.

To illustrate only the principle of operation, consider a typical second-order differential equation as

$$A\ddot{x} + B\dot{x} + Cx + D = f(t) \tag{8-54}$$

Division by $A$ and equating the highest-order term to the remainder leads to

$$\ddot{x} = -\frac{B}{A}\dot{x} - \frac{C}{A}x - \frac{D}{A} + \frac{1}{A}f(t) \tag{8-55}$$

This equation indicates that the second derivative of $x$ is equal to the sum of the other $x$ functions plus $f(t)/A$. Assume a summing amplifier be arranged with four inputs as in Fig. 8-29. The output of this amplifier must be equal to $\ddot{x}$ by Eq. (8-55) (actually $-\ddot{x}$ because of the sign-changing properties of the operational amplifier). A complete solution for $x$ is then possible by two successive integrations on $\ddot{x}$. At the same time the other input terms proportional to $\dot{x}$, and $x$, may be formed by additional amplifiers as in Fig. 8-30. A solution for $x$ is then obtainable and for a repetitive computer will appear as a trace on a cathode-ray screen. For a slow computer the output voltage to ground available at the $-x$ terminal may be used to actuate a recorder to give a record as a function of time.

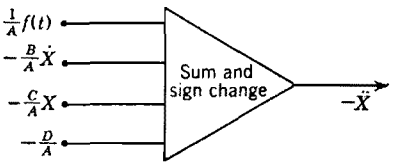



FIG. 8-29. Obtaining $\ddot{x}$ in the analog computer.

Switches $S_1$ and $S_2$ may be relay-operated at the instant the computer time scale equals zero, will place desired magnitudes of initial voltage on the integrating capacitors to represent initial conditions, and will also trigger the generator for $f(t)$. For a repetitive, or fast, computer these relays must be operated each time that $t = 0$ in the $f(t)$ driving function. Such a time scale requires that the complete transient solution of the problem be scaled in time so that the transient will die out in somewhat

less time than that of one period. Long-persistence cathode-ray screens are often helpful.

The preparation of the computer may be systematized as follows:

1. Write the differential equations of the system.
2. Solve for the highest-order derivative in each equation.
3. Form the block diagram for each equation.

   *a.* Derive lower-order terms by integration.
   *b.* Generate the coefficients of the lower-order terms as required, by use of available apparatus.
   *c.* Using summing amplifiers, add the lower-order terms after multiplication by their coefficients.

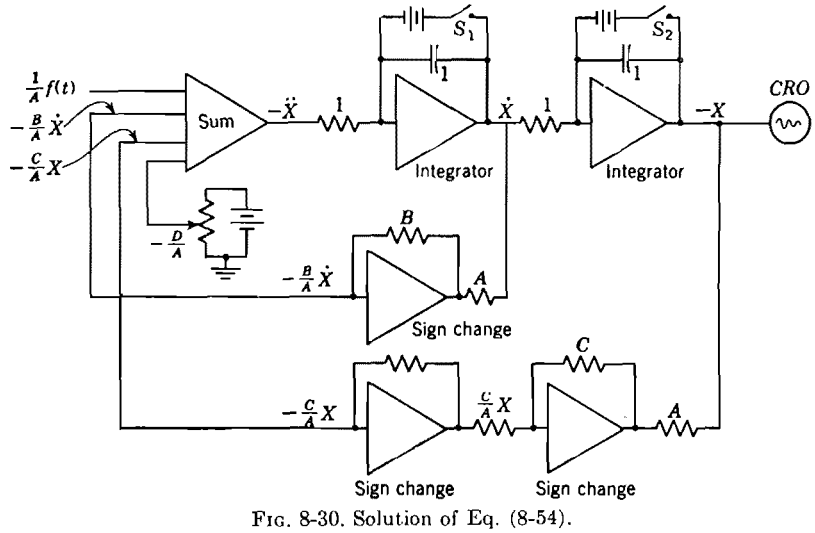



FIG. 8-30. Solution of Eq. (8-54).

**Example.** A body of mass $m$ falls under gravity through a viscous liquid. Design the computer circuit to plot displacement $y$, and velocity $\dot{y}$ against time, with initial displacement = 50 ft upward and initial velocity = 95 fps upward.

In preparation for the use of Lagrange's equation the energies may be written as

$$T = m\dot{y}^2/2 \qquad V = mgy \qquad D = k_d\dot{y}^2$$

Application of Lagrange's equation then yields the equation of motion as

$$m\ddot{y} - mg + k_d\dot{y} = 0$$

Rearranging,

$$\ddot{y} = g - \frac{k_d}{m}\dot{y} \tag{8-56}$$

The initial conditions call for voltages to be applied to the capacitors of the integrating amplifiers at $t = 0$, by use of the relays. These voltages are adjusted manually by potentiometers to bring the integrator outputs to the desired initial value with all

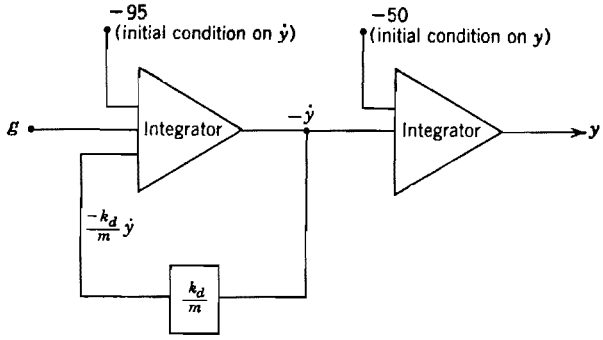



FIG. 8-31. Computer set up to solve Eq. (8-56).

other inputs removed. The circuit of Fig. 8-31 will then perform the desired mathematical operation.

**8-18. Function Generators.** Since analog computers are called upon to simulate and solve a wide variety of dynamic problems, it is necessary to be able to generate many driving-voltage waveforms, either as time functions or as functions of other variables.

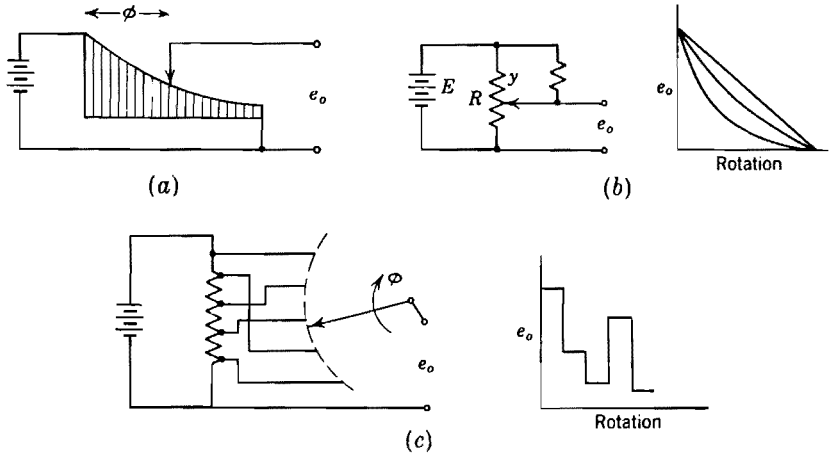



FIG. 8-32. Function generation by potentiometers.

Tapered or nonlinear wire-wound circular potentiometers, as in Fig. 8-32, may be driven by a constant-speed motor if a time function $e_o = f(t)$ is desired, or by a servomotor controlled by another variable $y$ if $e_o = f(y)$ is the desired output. The method is limited to monotonic functions (continuously increasing or decreasing) for which a suitable winding can be manufactured on a tapered card, of which the sine and cosine are examples.

The effect of a current load on the output terminals of a tapered

potentiometer must be considered. If the loading effect is appreciable, the output may depart from the desired function of shaft position. In fact, a linear potentiometer may be deliberately loaded,a s in Fig. 8-32*b*, to achieve quite a variety of nonlinear functions relating angle of rotation $\phi$ and output $e_o$. For instance, assuming zero external loading on $e_o$, the output as a function of angle $\phi$ for Fig. 8-32*b* is

$$e_o = E \frac{1 - \phi}{1 - \phi^2 R/(\phi R + y)}$$

where $\phi$ is measured from the $e_o = E$ position. For linear unloaded output $e_o = E(1 - \phi)$.

Photoelectric curve followers are available which will rotate a potentiometer, or provide an output voltage, proportional to the ordinate of a plotted curve. The output may also be a time function, if the coordinates are moved at a linear rate, or a function of some other variable $y$, if the coordinates are moved by a servomotor in proportion to the variable $y$.

While all the above devices are slow and only suited to slow-speed computers giving plotted outputs, the accuracy is very high, often of the order of a fraction of 1 per cent.

Early forms of computers used rotating switches to introduce potentials tapped from a resistive potentiometer as in Fig. 8-32*c*. This provides a step approximation to the desired time function, which can be made as close as desired by increase in the number of taps and commutator segments. Other mechanical refinements can be used if linear variation between tap points is desired.

The rotating switch is the only one of the above devices suited to the repetitive, or high-speed, analog computer. Two other electronic devices, the *photoformer* and the *monoformer*, have come into use to generate $e_o$ as a function of time or of another variable $y$. The photoformer is illustrated by the block diagram of Fig. 8-33. The input voltage $e_i$ drives the cathode-ray beam in the horizontal direction. At the same time the positioning bias applied to the d-c amplifier tends to deflect the beam upward, so that the screen spot rises above the mask. However, as soon as the spot appears above the mask edge, its illumination on the photocell develops an additional input to the d-c amplifier so polarized as to deflect the beam and spot downward. Equilibrium is reached with the spot just riding along the top edge of the mask.

Optical nonlinear feedback is thus provided for the d-c amplifier in the vertical-deflection circuit of the cathode-ray tube, the feedback path comprising the cathode-ray beam, the nonlinear mask, and the photocell. Since the output of a feedback amplifier is strongly dependent on the

characteristics of the feedback path if the gain is high, the output $e_o$ is dependent on, or is a function of, the mask shape.

It may also be seen that, as the spot is swept horizontally by the input $e_i$, the output voltage $e_o$ required to maintain the spot at the mask edge is made to vary as some function of $e_i$, this function being determined by the shape of the mask boundary. It is usually best to determine the exact mask shape experimentally, as variations in cathode-ray-tube sensitivity and in parallax induced by photocell size and location make calculations difficult. Accuracy to 1 or 2 per cent can usually be obtained if the mask is close to the tube surface and the spot intensity is kept low enough to prevent visual broadening of the spot. Speed of response is not a limiting factor with the usual oscillograph amplifiers

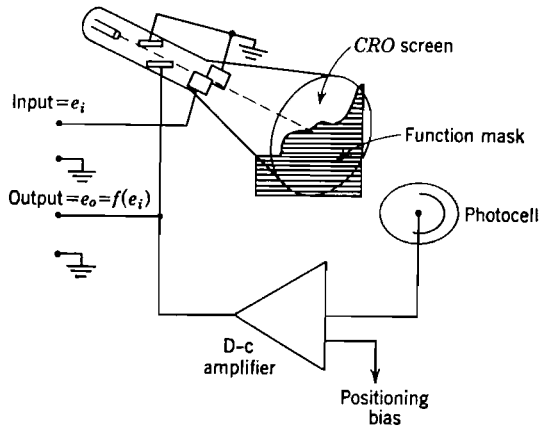



FIG. 8-33. Photoformer for generating nonlinear function.

and short-persistence screen tubes, although some difficulty may be expected in gaining good reproduction of steps or discontinuities in the mask pattern.

The monoformer places the mask inside the tube, reducing parallax error. Instead of a fluorescent screen the cathode-ray tube end has a mask made of a poor secondary-electron emitter such as graphite, the unmasked area being a good secondary emitter, such as aluminum. A collector electrode is positioned to receive the secondary electrons. With the d-c amplifier biased as for the photoformer, when the beam rises above the masked area the resultant shower of secondary electrons reaching the collector introduces an amplifier input polarized to drive the beam down. Equilibrium is again reached with the beam on the edge of the masked area.

The disadvantage of difficult and expensive mask change is somewhat balanced by lack of parallax and greater accuracy. Different tubes must be stocked for the expected variety of functions.

**8-19. Nonlinear-circuit Elements.** While vacuum diodes, triodes, crystal diodes, and some special materials such as *thyrite* will yield nonlinear volt-ampere characteristics, stability and reproducibility are somewhat lacking. Building up of nonlinear characteristics by addition of straight-line segments is more accurate and stable, and such a process is possible with reasonable numbers of crystal diodes, employed only as switches. The process is illustrated in Fig. 8-34, where three straight-line segments are utilized to approximate a volt-ampere curve. The conductances required for each segment are

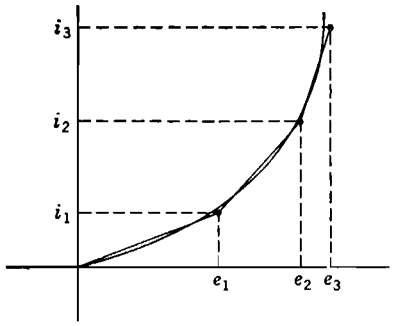



FIG. 8-34. Approximation of a curve by three straight lines.

$$g_1 = \frac{i_1}{e_1} \qquad g_2 = \frac{i_2 - i_1}{e_2 - e_1}$$

$$g_3 = \frac{i_3 - i_2}{e_3 - e_2}$$

if suitable switching be provided at $o$, $e_1$, and $e_2$ to place the proper conductances in the circuit.

Crystal diodes are well suited to operate as the needed switches. Since they are inherently rectifiers, the diode represents an open circuit when the cathode is more positive than the anode, whereas when the anode is more positive than the cathode the diode has an internal resistance in the range of a few ohms to a few hundred ohms. The diode thus is a potential-actuated switch in a circuit such as that of Fig. 8-35*a*. The volt-ampere curve of the circuit is that of Fig. 8-35*b*, showing that when the applied voltage makes terminal $A$ more than $E$ volts positive then the circuit appears as a resistance $R$ or conductance $1/R$ ($R \gg r_b$ of the diode). Modification of the circuit to that of Fig. 8-35*c* provides ready adjustment of the switching potential in either positive or negative direction and variation of the resistances appearing in the circuit.

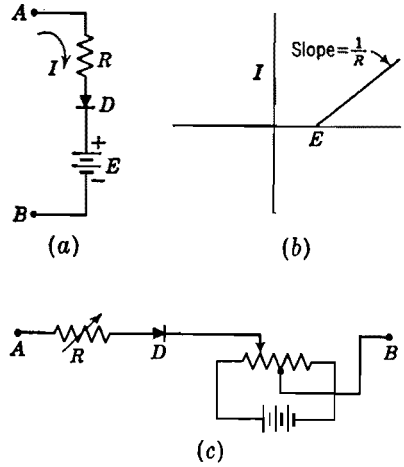



FIG. 8-35. Simple nonlinear-diode elements.

By combining a number of such diode circuits it is possible to generate

nonlinear-circuit elements with almost any desired relation between input and output voltages. A combination is shown in Fig. 8-36*a* along with a resultant volt-ampere curve yielding a nonlinear resistance decreasing with applied voltage $E$. Resistance increase with voltage is obtained by the modification in Fig. 8-36*b*. By the use of straight-line segments in this manner it is possible to approximate many mathematical functions. For instance, a curve of $E = KI^2$ can be represented to an

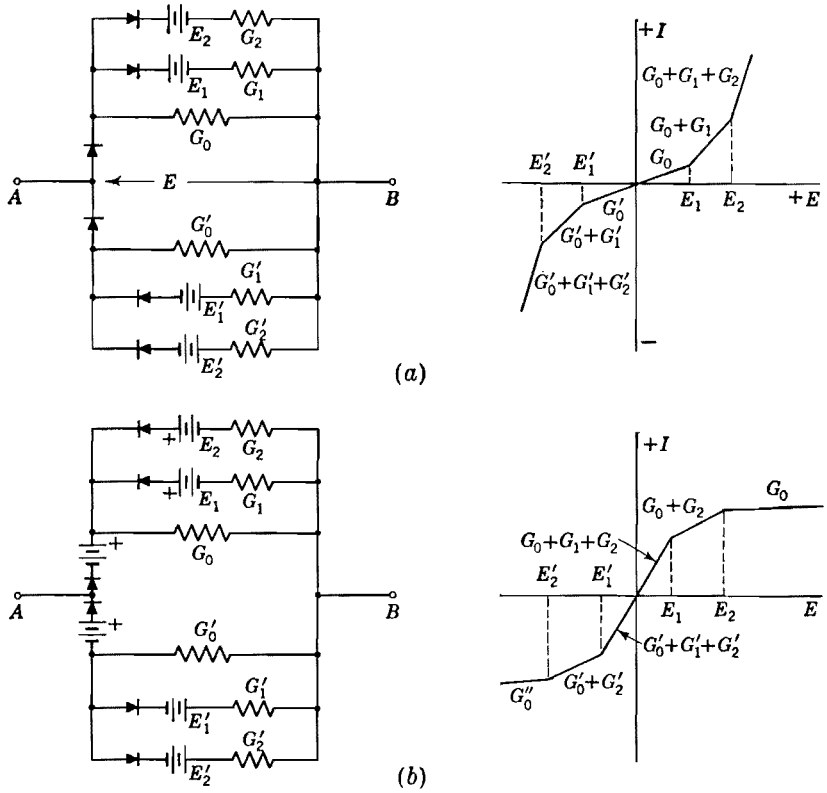



FIG. 8-36. Nonlinear resistors formed by diode switching.

accuracy of 1 per cent with five straight-line segments. Due to the fact the diode changes from open circuit to low resistance over a small range of voltage, some curvature is introduced at each intersection, and this improves the fit to a desired curve. By this means an electronic instrument for measuring rms input quantities has been designed.

These combinations may also be used in the feedback paths of operational amplifiers to yield additional interesting properties.

**8-20. Limiting Circuits.** In the study of some systems, particularly those encountered in automatic control, it is necessary to introduce into

the computer solution the idea of limiting or overload. This might occur in a control system if a valve reached the limit of its travel and could allow no more flow regardless of increasing demand on its actuator.

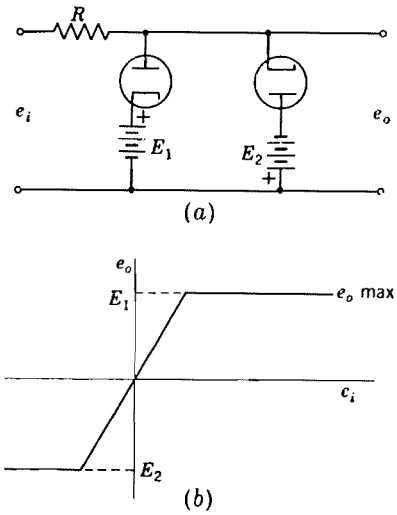



FIG. 8-37. Limiter circuit and characteristics.

This limiting action is readily obtained through simplifying the nonlinear diode circuits of the preceding section to two diodes, as in Fig. 8-37*a*, which is also the clipper of Chap. 9. In this usage $R \gg r_b$ of the diodes. The bias voltage should be taken from a power supply of low internal resistance; otherwise the flat limiting portion of the output characteristic will be somewhat curved. Voltages $E_1$ and $E_2$ should be adjustable to provide choice of the limiting value.

**8-21. Multiplying Devices and Circuits.** It is easily possible to multiply a variable by a constant of value less than unity by use of the potentiometer of Fig. 8-38*a* or by use of the operational amplifier. Multiplication of a variable $e_1$ by a second variable $e_2$ is more difficult.

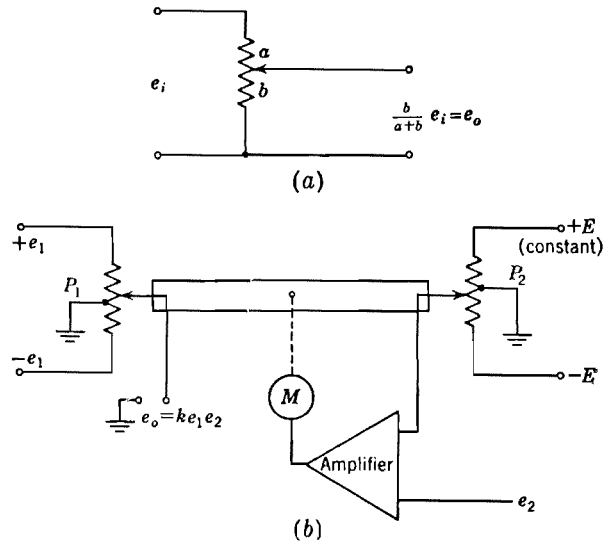



FIG. 8-38. (*a*) Multiplication of a variable by a constant. (*b*) Multiplication of two variables.

For the slow-speed class of computer this operation is usually performed by a potentiometer $P_1$ with an applied potential proportional to $e_1$ and with the tap positioned by a servomotor in proportion to $e_2$, as given below. A second potentiometer $P_2$ mounted on the same shaft with $P_1$ has applied to it a fixed and constant potential $E$. An amplifier compares the second potential $e_2$ and the tapped voltage on $P_2$ and drives the motor to adjust $P_2$ to bring the difference of potentials on the amplifier input to zero. Since $P_1$ and $P_2$ are on the same shaft, the tap on $P_1$ is being adjusted in accordance with $e_2$ and its over-all voltage is $e_1$.

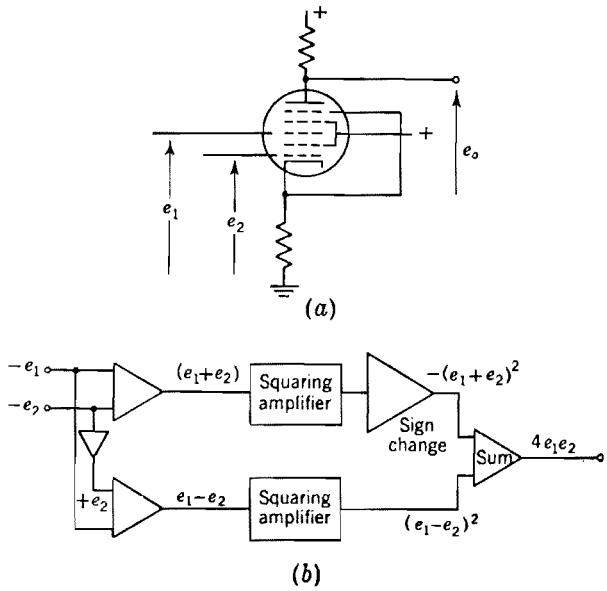



FIG. 8-39. Electronic forms of multipliers.

Thus the output voltage from the tap on $P_1$ will be proportional to $e_1$ times $e_2$. As shown, the multiplier has ranges from $+e_2$ to $-e_2$. Such devices are highly accurate but limited in speed of response to that of the servomechanism, usually below 2 cycles.

The problem of a multiplier for the high-speed repetitive computer has not yet been completely solved. The multigrid vacuum tube (6L7, for example) of Fig. 8-39*a* will have an output proportional to the product of $e_1$ and $e_2$ over a limited input range but is dependent on selection and maintenance of certain tube characteristics.

It is possible to select tubes to have grid-transfer characteristics which show pronounced quadratic form over limited ranges, that is,

$$i_b = Ae_c^2$$

to a reasonable approximation. Use of such tubes in the circuit of Fig. 8-39*b*, as squaring elements, gives an output of

$$e_1e_2 = \frac{(e_1 + e_2)^2}{4} - \frac{(e_1 - e_2)^2}{4} \tag{8-57}$$

which for obvious reasons is referred to as a *quarter-square* multiplier. While suited to analog multiplication by reason of speed, its accuracy is dependent on the maintenance of critical shapes in the transfer characteristics of several tubes, and these cannot be maintained to a high degree.

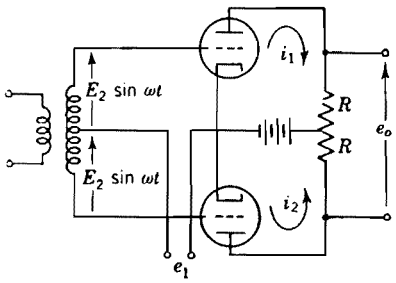



FIG. 8-40. Balanced modulator multiplier.

Another all-electronic method is based on the principle of amplitude modulation as in Fig. 8-40. Let a slowly varying voltage $e_1$ and an a-c carrier voltage $e_2 = E_2 \sin \omega t$ be applied to a pair of tubes having assumed square-law grid-transfer characteristics given by

$$i_b = ae_c^2$$

For tube $T_1$ the plate current will be due to an input voltage $e_1 + E_2 \sin \omega t$ so that

$$\begin{aligned} i_1 &= a(e_1 + E_2 \sin \omega t)^2 \\ &= ae_1^2 + 2ae_1E_2 \sin \omega t + aE_2^2 \sin^2 \omega t \end{aligned}$$

and for tube $T_2$

$$\begin{aligned} i_2 &= a(e_1 - E_2 \sin \omega t)^2 \\ &= ae_1^2 - 2ae_1E_2 \sin \omega t + aE_2^2 \sin^2 \omega t \end{aligned}$$

The output voltage $e_o$ is given by

$$\begin{aligned} e_o &= (i_1 - i_2)R \\ &= 4ae_1E_2 \sin \omega t \end{aligned}$$

This is the desired product term. It may be rectified back to direct-current for reintroduction into the computer. The carrier frequency may be initially produced by a chopper and must be high with respect to the highest input frequencies present in $e_1$. Since the output is proportional to $a$, the accuracy is dependent on tube characteristics and is not likely to be much better than 5 to 10 per cent.

The photoformer and monoformer may be employed as multipliers, through use of logarithmic and antilogarithmic masks, in the circuit of Fig. 8-41. Division could be carried out as easily. The operation is

restricted to positive values of $e_1$ and $e_2$ since logarithms of negative numbers are not defined.

Several other multiplier systems are based on the principle of sampling voltage $e_1$ periodically, the length of the sample pulses varying linearly with the other variable $e_2$. Thus the amplitude of the sample pulse is proportional to the value of $e_1$ at a particular instant, the length of the pulse is proportional to $e_2$ at that instant, and the area (or average) of the pulse is then proportional to the desired product $e_1e_2$. The only requirement on sampling frequency is that it be high with respect to the highest input frequency. With sampling frequencies of the order of 40 to 50 kc this condition is met for most applications.

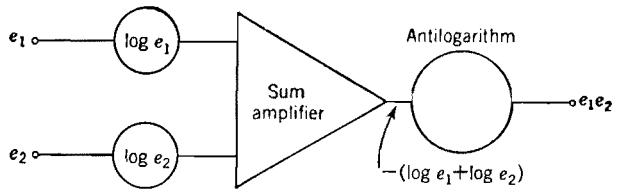



FIG. 8-41. Use of three photoformers with logarithmic masks to produce multiplication.

The output is the averaged value of the pulses, and it is possible to design multipliers for four-quadrant operation, or for values of $e_1$ and $e_2$ between plus and minus limits, and for $-100$- to $+100$-volt inputs.

## PROBLEMS

**8-1.** In Fig. 8-1, specify all resistors and battery voltages if the tubes have $\mu = 70$, $r_p = 40{,}000$ ohms and are to operate at $E_b = 100$ volts, $E_c = -2$ volts. The over-all gain is to be 2250, with first-stage gain 150 per cent of the second-stage gain, and $I_{b1} = 1.25$ ma, $I_{b2} = 1.50$ ma.

**8-2.** In Fig. 8-6, $T_1 = T_2$ and $\mu = 70$, $r_p = 45{,}000$ ohms; $T_3 = T_4$ and $\mu = 20$, $r_p = 7700$ ohms. Compute the over-all gain, with the zero adjuster set for 2000 ohms.

**8-3.** The circuit of Fig. 8-2 is to be redesigned. The tubes are halves of a 6SN7 and are to operate at

$$E_{b1} = 200 \text{ volts} \qquad E_{c1} = -4 \text{ volts}$$
$$E_{b2} = 250 \text{ volts} \qquad E_{c2} = -6 \text{ volts}$$

If $R_1 = 25{,}000$ ohms, $R_2 = 15{,}000$ ohms and the battery is to supply a total current of 20 ma, find the bleeder current and the required power-supply voltage, and specify the potentials at all bleeder taps and at the grids.

**8-4.** In Fig. 8-8 the meter $M$ has a resistance of 1500 ohms. If $R_a = 15{,}000$ ohms, $R_k = 1000$ ohms, $\mu = 20$, $r_p = 7700$ ohms and the circuit is balanced for $e_i = 0$, find the current through the instrument per volt input.

**8-5.** Prove that Eq. (8-17) is correct for the gain of the Miller cathode-drift compensator.

**8-6.** Derive an expression for the gain of the d-c amplifier of Fig. 8-42, where $e_1$ is the grid-to-grid input and the tubes are identical.

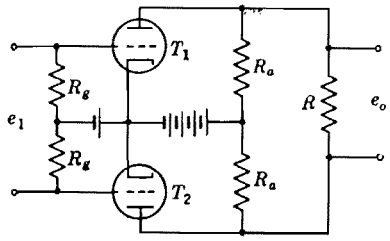
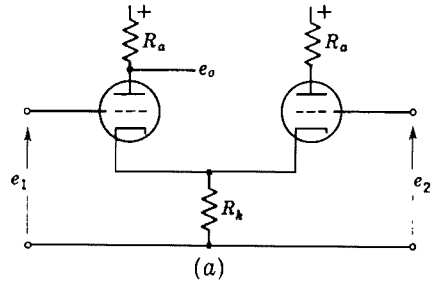


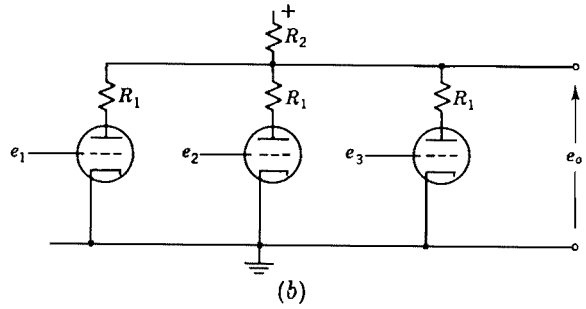

FIG. 8-42

**8-7.** Show that Eq. (8-24) is correct for the bridge-balanced amplifier operated with grounded center tap on the power supply and with $R_L$ large.

**8-8.** Show that the circuit of Fig. 8-43*a* will subtract or that $e_o = k(e_1 - e_2)$.




(*a*)




(*b*)

FIG. 8-43

**8-9.** Demonstrate that the circuit of Fig. 8-43*b* will add, or that $e_o = K(e_1 + e_2 + e_3)$.

**8-10.** For each operational circuit of Fig. 8-44 determine the relation of $e_o$ to $e_i$ in terms of operator $p$.

**8-11.** By use of Lagrange's equation prove that Eq. (8-50) is correct for the system of Fig. 8-27*a*.

**8-12.** Using Lagrange's equation, develop the equation of motion of a pendulum, assuming all mass concentrated in the bob and the angle of swing small.

**8-13.** (*a*) Draw a block diagram for a functional analog computer to determine $y$ for the equation

$$\ddot{y} - f(t) = 0$$

(*b*) If $f(t)$ is a step function [$f(t) = 0$ for $t < 0, f(t) = 1$ for $t > 0$], draw the value of $y$ and $\dot{y}$ as functions of time.

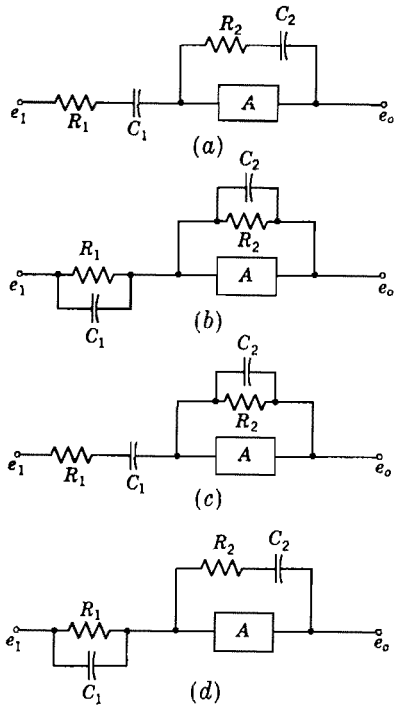



FIG. 8-44

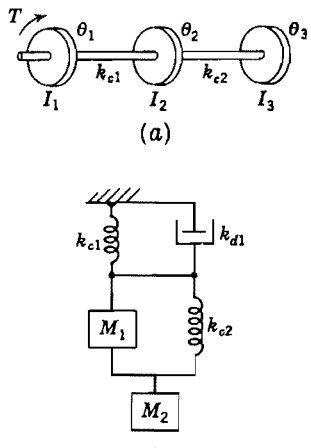



FIG. 8-45

**8-14.** Indicate by a block diagram how an operational amplifier may be made to integrate the sum of three variables.

**8-15.** Set up a block diagram for an operational solution of the equation

$$\ddot{y} + a\dot{y} + y = 0 \qquad \text{for } 0 < a < 1$$

**8-16.** Draw an electric circuit for direct-analog solution of the equation of Prob. 8-15, and list the analogous variables.

**8-17.** Set up an electric circuit for the mechanical system of Fig. 8-45*a*, and list the analogous quantities. For what mechanical conditions will $\theta_2$ be maximum?

**8-18.** By Lagrange's equations set up the equations of motion for inertia $I_3$ of Fig. 8-45*a*.

**8-19.** Set up the direct-analog circuit for the mechanical system of Fig. 8-45*b*.

## REFERENCES


1. Loftin, E. H., and S. Y. White: Cascaded Direct-coupled Tube Systems Operated from Alternating Current, *Proc. IRE*, **18**: 669 (1930).
2. Valley, G. E., Jr., and H. Wallman: "Vacuum Tube Amplifiers," McGraw-Hill Book Company, Inc., New York, 1948.
3. Miller, S. E.: Sensitive D-C Amplifier with A-C Operation, *Electronics*, **14**: 27 (November, 1941).
4. Yu, Y. P.: Cathode-follower Coupling, *Electronics*, **19**: 99 (1946).
5. Rittenhouse, J. W.: Cathode Drift Compensation in D-C Amplifiers, *Elec. Eng.*, **72**: 299 (1953).
6. Artzt, M.: Survey of D-C Amplifiers, *Electronics*, **18**: 112 (August, 1945).
7. Korn, G. A., and T. M. Korn: "Electronic Analog Computers," McGraw-Hill Book Company, Inc., New York, 1952.
8. Williams, A. J., R. E. Tarpley, and W R. Clark: D-C Amplifier Stabilized for Zero and Gain, *Trans. AIEE*, **67**: 47 (1948).
9. Goldberg, E. A.: Stabilization of D-C Amplifiers, *RCA Rev.*, **11**: 296 (1950).
10. Dubridge, L. A.: The Amplification of Small Direct Currents, *Phys. Rev.*, **37**: 392 (1931).
11. Collins, D. L.: Electrometer Tubes, *Instruments*, **26**: 1708 (1953).
12. Victoreen, J. A.: Electrometer Tubes for the Measurement of Small Currents, *Proc. IRE*, **37**: 432 (1949).
13. Ragazzini, J. R., R. H. Randall, and F. A. Russell: Analysis of Problems in Dynamics by Electronic Circuits, *Proc. IRE*, **35**: 444 (1947).
14. Chance, F. B., D. Sayre, V. Hughes, E. F. MacNichol, Jr., and F. C. Williams: "Waveforms," McGraw-Hill Book Company, Inc., New York, 1949.
15. Ginzton, E. L.: D-C Amplifier Techniques, *Electronics*, **17**: 98 (March, 1944).
16. Bush, V.: The Differential Analyzer, A New Machine for Solving Differential Equations, *J. Franklin Inst.*, **212**: 447 (1931).
17. Harder, E. L., and G. D. McCann: A Large-scale General Purpose Electric Analog Computer, *Trans. AIEE*, **67**: 664 (1948).
18. Sunstein, D. E.: Photoelectric Wave Form Generator, *Electronics*, **22**: 100 (1949).
19. MacNee, A. B.: An Electronic Differential Analyzer, *Proc. IRE*, **37**: 1315 (1949).
20. Soroka, W. W.: "Analog Methods in Computation and Simulation," McGraw-Hill Book Company, Inc., New York, 1954.
21. Korn, G. A.: Design and Construction of Universal Function Generating Potentiometers, *Rev. Sci. Instr.*, **21**: 77 (1950).
22. Murphy, Glenn: "Similitude in Engineering," The Ronald Press Company, New York, 1950.

CHAPTER 9

# SWITCHING CIRCUITS; DIGITAL COMPUTATION

In previous chapters the ability of the vacuum tube faithfully to amplify a given input voltage has been studied and emphasized. It is now necessary to introduce another large class of vacuum-tube circuits in which emphasis is on the ability of the tube to operate as a switch. In this latter class the output voltage may bear no relation to the input magnitude or waveform, except to be present or absent in accordance with certain levels of input.

The vacuum tube thus used is an inertialess and nearly instantaneous switch. Having speed of operation as its main advantage, its major difficulty is that when closed it still represents a small but appreciable resistance. When open, of course, it has a resistance which is essentially infinite. The low value of closed resistance must be allowed for in circuit design but can be neglected in many cases.

Circuits employing vacuum tubes as switches are frequently used to generate pulses and other nonsinusoidal wave shapes or to operate with or by such pulses or waveforms. Radar equipment forms one such class of circuit, with television, digital computers, and various pulse-modulated communication systems also included.

**9-1. Circuit Elements.** As mentioned in Sec. 3-1, resistive circuit elements may have linear or nonlinear volt-ampere characteristics. A nonlinear resistor is represented by a curve in Fig. 9-1*a* and obeys a law

$$e = R(i)i \tag{9-1}$$

where the proportionality coefficient $R(i)$ is a function of current. A linear resistor follows the usual law

$$e = Ri \tag{9-2}$$

Some resistance elements, particularly crystal or vacuum diodes, may be unilateral (and therefore nonlinear) as in Fig. 9-1*b*, where the volt-ampere curve is dissimilar for positive and negative voltages.

In circuits having nonlinear resistive elements it is necessary to employ the concepts of *static* and *dynamic* resistance as in Sec. 3-2. These are

illustrated in Fig. 9-2. As defined, the static resistance is

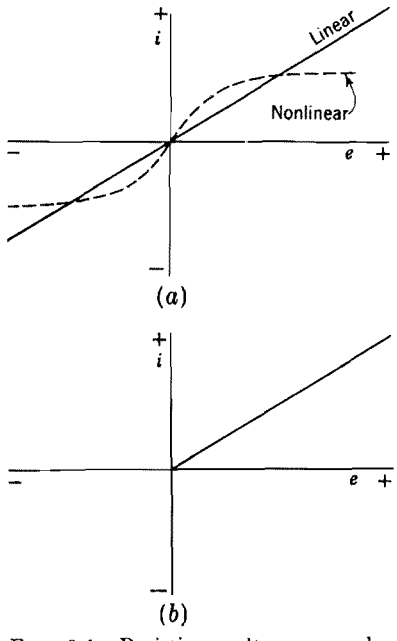



FIG. 9-1. Resistive volt-ampere characteristics.

$$r_b = \frac{e_A}{i_A} \tag{9-3}$$

at point $A$, and the dynamic resistance is

$$r_p = \frac{de_A}{di_A} \tag{9-4}$$

or the reciprocal of the slope at the point $A$. For a linear resistor these values are identical, but in Fig. 9-2 the values of $r_b$ and $r_p$ will vary from point to point, and near $A$ the value of $r_p$ is negative.

For capacitive-circuit elements

$$e_C = E_o + \frac{1}{C}\int_o^t i_C\, dt \tag{9-5}$$

$$i_C = C\frac{de_C}{di_C} \tag{9-6}$$

and for inductive-circuit elements

$$e_L = L\frac{di_L}{dt} \tag{9-7}$$

$$i_L = I_o + \frac{1}{L}\int_o^t e\, dt \tag{9-8}$$

It is apparent that the potential across a capacitor or the current through an inductor cannot change instantaneously, without an infinite applied emf.

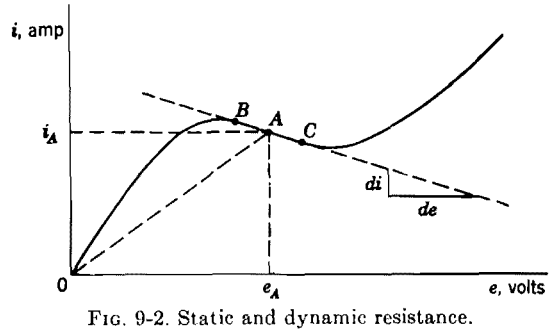
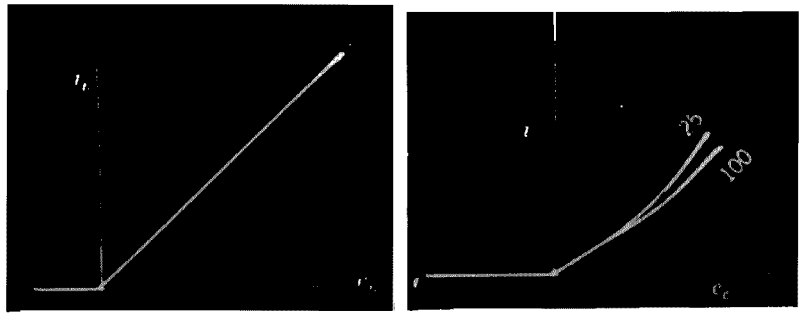



FIG. 9-2. Static and dynamic resistance.

**9-2. Tubes in Switching Circuits.** While diode-tube characteristics are nonlinear, the degree of nonlinearity is frequently not great enough to warrant use of graphical analysis in switching circuits. Sufficient

accuracy may usually be obtained by assuming a constant average value of static resistance, which may be as low as 1000 to 2000 ohms for many vacuum diodes, and this may often be completely neglected when the diode is in series with a much larger resistor. The near approach to linearity of such a characteristic is shown in Fig. 9-3*a*.

Similar treatment may be used for the plate circuits of triodes, although recourse must usually be had to the actual characteristics for the dynamic plate resistance and actual current values. Triodes may often be operated with grid voltages swinging from cutoff to zero or positive values. With a positive grid voltage the static resistance of the grid-cathode circuit of the usual triode may be taken as 1000 ohms as a good approximation. The linearity of the grid resistance for several positive grid voltages is shown in Fig. 9-3*b*, being good to +50 volts.




(*a*) (*b*)

FIG. 9-3. (*a*) Dynamic characteristic of 6H6 diode with load of 20,000 ohms. (*b*) Grid volt-ampere curves for a 6J5 triode with +100 volts $e_{c,\max}$, $E_b$ = 25 and 100 volts.

Crystal diodes are frequently employed and offer lower static resistance than vacuum diodes, and with no current at zero applied emf as is the case of the vacuum diode. Static resistances as low as a few hundred ohms are possible, although the back resistance (with anode negative) will not be so high as with thermionic diodes. A typical volt-ampere curve for a crystal diode is given in Fig. 9-4*a*.

Gas diodes and triodes may be considered as having an idealized volt-ampere characteristic as in Fig. 9-4*b*. The constant voltage drop is assumed independent of current value.

**9-3. Response of *RC* and *RL* Series Circuits.** In many applications desired waveforms are built up from segments of *RC* or *RL* transient voltages, with triodes or diodes employed as switches. A discussion of *RC* and *RL* transients will be introduced here for reference.

Consider the circuit of Fig. 9-5*a*, in which switch *s* is thrown to position 2 at $t = 0$. The potential on *C* at $t = 0$ is $E_o$, which may or may not be equal to $E_1$, dependent on the time *s* has been in position 1. At $t = 0$

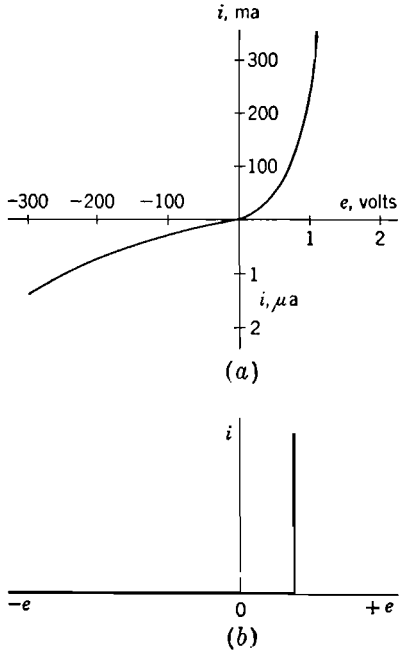



FIG. 9-4. (*a*) Germanium crystal-diode volt-ampere curve. (*b*) Ideal-gas-diode volt-ampere curve.

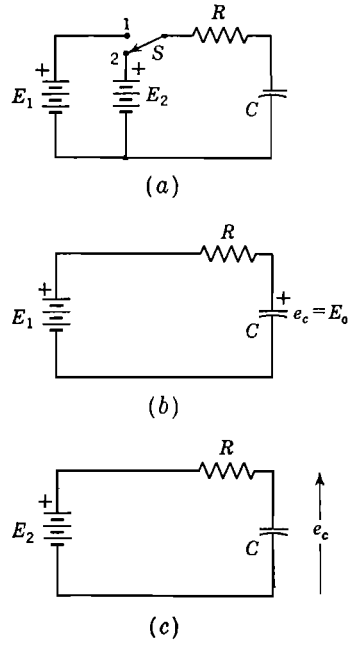



FIG. 9-5. Switching in an *RC* circuit.

the circuit changes from that of Fig. 9-5*b* to that of Fig. 9-5*c*, and the switching action might be obtained by proper operation of a triode. Writing an equation for the circuit of Fig. 9-5*c* at $t = 0$ gives

$$iR + \frac{1}{C}\int_o^t i\,dt + E_o - E_2 = 0 \tag{9-9}$$

Differentiating and dividing by $R$ gives

$$\frac{di}{dt} + \frac{i}{RC} = 0$$

or

$$\left(p + \frac{1}{RC}\right) i = 0 \tag{9-10}$$

The solution of Eq. (9-10) is

$$i = A\epsilon^{-t/RC}$$

and from Eq. (9-9), at $t = 0$, where $i = i(0)$,

$$i(0) = \frac{E_2 - E_o}{R}$$

It can be seen that $A = (E_2 - E_o)/R$ and

$$i = \frac{E_2 - E_o}{R} \epsilon^{-t/RC} \tag{9-11}$$

The resistor voltage is then

$$e_R = iR = (E_2 - E_o)\epsilon^{-t/RC} \tag{9-12}$$

and the capacitor voltage is

$$\begin{aligned} e_C = E_o + \frac{1}{C}\int_o^t i\,dt &= E_o + \frac{1}{C}\int_o^t \frac{E_2 - E_o}{R} \epsilon^{-t/RC}\,dt \\ &= E_o + (E_2 - E_o)(1 - \epsilon^{-t/RC}) \end{aligned} \tag{9-13}$$

Equations (9-12) and (9-13) serve to introduce the exponential terms $\epsilon^{-t/RC}$ and $1 - \epsilon^{-t/RC}$, which will be discussed later.

In a similar manner an $RL$ circuit with a steady current $I_o$ flowing may have switch $s$ thrown from 1 to 2 at $t = 0$, as in Fig. 9-6. A circuit equation for $t > 0$ may be written as

$$L\frac{di}{dt} + Ri - E_2 = 0$$

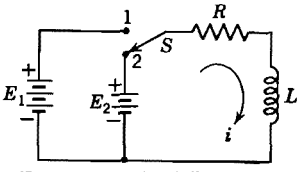



FIG. 9-6. Series $RL$ circuit.

Separation of variables and integration gives

$$E_2 - Ri = B\epsilon^{-Rt/L} \tag{9-14}$$

Since $i = I_o$ at $t = 0$,

$$A = E_2 - RI_o$$

and then

$$i = I_o + \left(\frac{E_2}{R} - I_o\right)(1 - \epsilon^{-Rt/L}) \tag{9-15}$$

The voltage across the resistor is

$$e_R = iR = I_oR + (E_2 - I_oR)(1 - \epsilon^{-Rt/L}) \tag{9-16}$$

and that across the inductor (assumed resistanceless) is

$$e_L = L\frac{di}{dt} = (E_2 - I_oR)\epsilon^{-Rt/L} \tag{9-17}$$

It can be seen by comparison of Eqs. (9-13) and (9-16) that $e_C$ for the $RC$ circuit and $e_R$ for the $RL$ circuit are of similar form, likewise $e_R$ for the $RC$ circuit and $e_L$ for the $RL$ circuit. The $RL$-circuit equations also involve the terms $\epsilon^{-Rt/L}$ and $1 - \epsilon^{-Rt/L}$, of similar form to those for the $RC$ circuit.

**9-4. The Time Constant.** All single-capacitor or single-inductor $RC$ and $RL$ circuits involve the terms $e^{-t/RC}$ and $\epsilon^{-Rt/L}$. The reciprocals of

the exponent coefficients $RC$ and $L/R$ have dimensions of time and are known as the *time constants* of the circuits. These terms determine the rate of change of the exponential term and are defined as the time in seconds for the output transient to change $1/\epsilon$, or 64 per cent of its infinite time value. It is convenient to measure time of response of a circuit in units of time constant, since universal curves as in Fig. 9-7 may be plotted

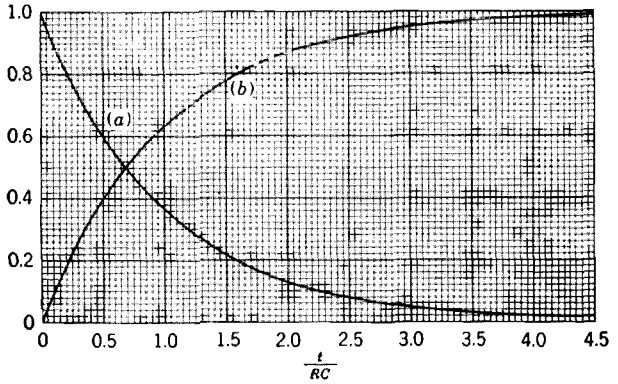



FIG. 9-7. (*a*) $\epsilon^{-t/RC}$. (*b*) $1 - \epsilon^{-t/RC}$.

and used to determine values of the exponential terms in various transient switching problems. The per cent of final value attained in lengths of time equal to multiples of circuit time constant is indicated in Table 16.

TABLE 16. TIME-CONSTANT VALUES

| *Time constant* | *Per cent of complete transient* | *Time constant* | *Per cent of complete transient* |
|---|---|---|---|
| 0 | 0.0 | 3.0 | 95.5 |
| 0.5 | 39.4 | 3.5 | 97.0 |
| 1.0 | 63.2 | 4.0 | 98.2 |
| 1.5 | 77 7 | 4.5 | 98.9 |
| 2.0 | 86.5 | 5.0 | 99.4 |
| 2.5 | 91.8 | 10.0 | 99.999 |

It is frequently sufficiently accurate to assume the transient completed after four time constants, when the value of the function is within 1.8 per cent of the final value.

**Example.** The switch in Fig. 9-8 has been closed on position 1 for a long time. It is then switched to position 2 for 750 $\mu$sec, then back to position 1. Sketch the voltage waveform for $e_o$ during the complete transient.

Since the switch $s$ has been closed on 1 for a long time, the value of $e_o$ at $t = 0$ is taken as 250 volts. At $t = 0$ the switch is thrown to 2, and the capacitor starts to discharge. From Eq. (9-11) with $E_2 = 0$, $E_o = 250$ volts, and $RC = 0.01 \times 10^{-6} \times$

$50 \times 10^3 = 500 \times 10^{-6}$ sec,

$$i = \frac{250}{50 \times 10^3} \epsilon^{-t/(500 \times 10^{-6})}$$

Voltage $e_o$ is given by use of Eq. (9-13) as

$$e_o = -Ri + e_C = -10^4 i + 250 + (-250)(1 - \epsilon^{-t/(500 \times 10^{-6})})$$
$$= 200\epsilon^{-t/(500 \times 10^{-6})}$$

At 750 $\mu$sec, which is equal to 1.5 time constants, the value of the exponential term is 0.22 from Fig. 9-7, and at this instant the switch $s$ returns to position 1. The voltage $e_o$ at the switching instant is

$$e_o = 200 \times 0.22 = 44.0 \text{ volts}$$

Intermediate values to draw the waveform may be obtained from Fig. 9-7.

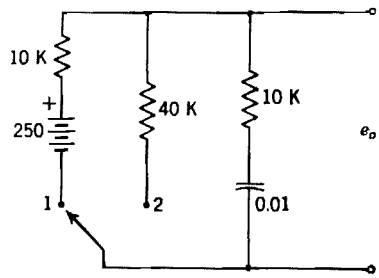



FIG. 9-8

At the switching instant the capacitor potential $e_C$ is given by Eq. (9-13) as

$$e_C = 250 + (-250)(1 - \epsilon^{-1.5}) = 55.0 \text{ volts}$$

and this is the new value for $E_o$ during the second time interval. The time constant for the circuit with $s$ on 1 or for $t + 750 \times 10^{-6}$ sec is

$$RC = 0.01 \times 10^{-6} \times 20 \times 10^3 = 200 \times 10^{-6} \text{ sec}$$

and the current will be given by

$$i = \frac{250 - 55}{2 \times 10^4} \epsilon^{-t/(200 \times 10^{-6})}$$
$$= 97.5 \times 10^{-4} \epsilon^{-t/(200 \times 10^{-6})}$$

and voltage $e_o$ will be given by Eq. (9-13) as

$$e_o = 10^4 i + 55 + (250 - 55)(1 - \epsilon^{-t/(200 \times 10^{-6})})$$
$$= 250 - 97.5\epsilon^{-t/(200 \times 10^{-6})}$$

At the instant of switching ($t = 0$ for the above equation) the value of $e_o$ jumps to 152.5 volts and then rises exponentially toward its final value of 250 volts. Intermediate values may be found from Fig. 9-7. The complete waveform is plotted in Fig. 9-9.

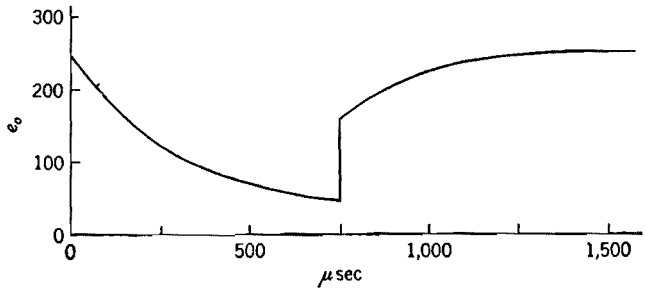



FIG. 9-9. Waveform for example.

**9-5. The Differentiating Circuit.** Referring to a circuit of the form of Fig. 9-10*a*, with the initial value of $e_C = 0$, then

$$e_i = \frac{1}{C}\int_0^t i\,dt + iR \tag{9-18}$$

By making the *RC time constant small with respect to the interval of any applied voltage*, the capacitor $C$ will charge and discharge rapidly, and the integral in the equation above will be large or important in relation to the $iR$ term. The capacitor will essentially determine the value of current, or

$$i \cong C\frac{de_i}{dt} \tag{9-19}$$

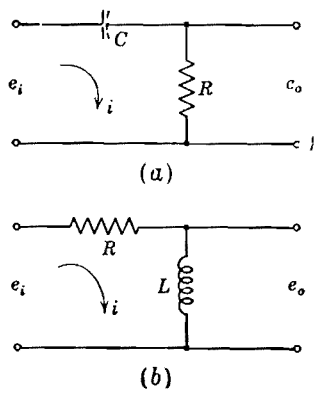


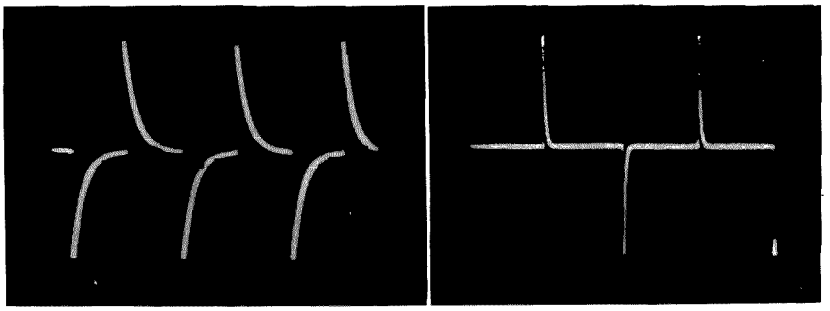

FIG. 9-10. Differentiating circuits.

The output voltage will then be $iR$, and under the requirement of the $RC$ assumption this gives

$$e_o = iR \cong RC\frac{de_i}{dt} \tag{9-20}$$

Thus, with a *small* value of time constant $RC$, the output voltage will be proportional to the derivative of the input voltage, and the circuit is known as a *differentiating circuit.*

Theoretically the output for an applied square wave of voltage should be a series of alternate positive and negative pulses of zero width in time.



(a) (b)

FIG. 9-11. Output of an $RC$ differentiator with square-wave period $T = 2 \times 10^{-4}$ sec. (*a*) With $RC/T = 0.05$; (*b*) with $RC/T = 0.0005$.

As the time constant is made smaller, this condition is approached. The amplitude of the pulses will be equal to the instantaneous change of input voltage of the input wave. In Fig. 9-11 are shown the output-voltage waveforms for several ratios of the $RC$ value to the period of

the applied wave. As $RC$ is made smaller, the $e_o$ voltage approaches more closely the spike form to be expected for a derivative of a square input wave.

This circuit is frequently used to sharpen pulse waveforms or to generate pips, or sharp pulses, for triggering other electronic circuits.

The circuit of Fig. 9-10*b* also has differentiating properties. Writing

$$e_i = L\frac{di}{dt} + iR \tag{9-21}$$

then if $L/R$ is small, the applied potential will largely appear across $R$, or $R$ will determine the current $i$, so that

$$i \cong \frac{e_i}{R}$$

and then

$$e_o \cong \frac{L}{R}\frac{de_i}{dt} \tag{9-22}$$

and thus the inductive form of circuit will also differentiate the input waveform. Since physical inductors have some resistance, the output voltage will have an $iR$ component and will not be so accurate a derivative as given by the $RC$ circuit.

The $RC$ form of Fig. 9-10*a* is also that of the $RC$ coupling for amplifiers. In such applications, to avoid differentiation and distortion of the waveform, it is desirable that the time constant $RC$ be very long with respect to the lowest period of the applied voltage. This circuit has been previously discussed in Sec. 5-8.

**Example.** A square wave of peak value +100 volts is applied to an $RC$ differentiator of time constant 0.05 of the period of the wave. Assuming initial value of the wave equal to zero and initial value of $e_C = 0$, plot the output waveform.

Let $t = 0$ at the time the input wave jumps from 0 to +100 volts. Since the potential across the capacitor cannot change instantaneously, $e_C$ will remain zero for an instant and the input jump will appear across $R$, resulting in a jump of +100 volts in $e_o$ and establishing the initial current $I_o = 100/R$. The capacitor will then charge toward an ultimate +100 volts, the current decreasing exponentially with a value given by

$$i = \frac{100}{R}\epsilon^{-t/RC}$$

and the output voltage $e_o = iR$ will also fall. The variation may be plotted using values from Fig. 9-7. After four time constants = 0.2 of the period the output voltage will be essentially zero once more. At the half period the input voltage returns to zero, or falls 100 volts. The output voltage falls a like amount from zero to −100 volts. The capacitor once again starts to charge until after four time constants its potential is once more 100 volts and the current and output voltage are again zero. This is illustrated in Fig. 9-12.

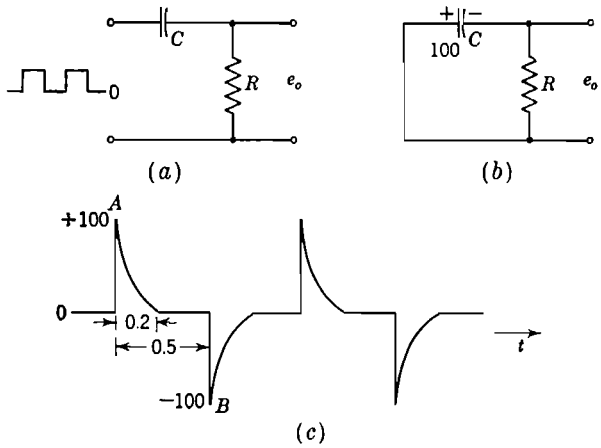



FIG. 9-12. (*a*) Initial conditions at *A*. (*b*) Conditions in circuit at *B*. (*c*) Output wave.

**9-6. Integrating Circuits.** The circuits of Fig. 9-13 may be shown to perform integration of input waveforms. Taking the *RC* circuit of Fig. 9-13*a*, with $e_C$ initially zero, the circuit equation is

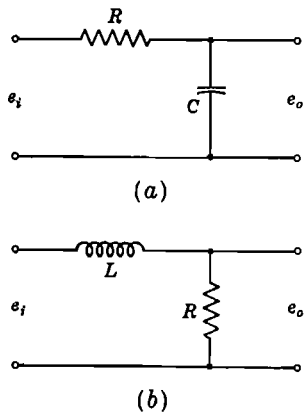



FIG. 9-13. Simple integrating circuits.

$$e_i = iR + \frac{1}{C}\int_o^t i\,dt$$

or

$$i = \frac{e_i}{R} - \frac{1}{RC}\int_o^t i\,dt \qquad (9\text{-}23)$$

If *RC is made very large* with respect to the longest period of the input signal, then the second term will be very small with respect to the first and

$$i \cong \frac{e_i}{R} \qquad (9\text{-}24)$$

Since

$$e_o = \frac{1}{C}\int_o^t i\,dt$$

then

$$e_o \cong \frac{1}{RC}\int_o^t e_i\,dt \qquad (9\text{-}25)$$

and the output voltage is proportional to the integral of the input voltage under the assumption concerning *RC*. The output of such an *RC* circuit, with long time constant, is illustrated by the oscillogram of Fig. 9-14, when driven by a square wave. The circuit provides a simple means of obtaining an approximate integral. For high accuracy the circuits of Chap. 8 should be employed.

The $RL$ circuit of Fig. 9-13*b* will also result in integration. For a large value of $L/R$ the input voltage will appear largely across $L$ so that

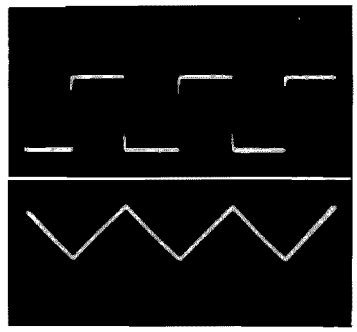
$$e_i \cong L\frac{di}{dt}$$

or
$$i \cong \frac{1}{L}\int_o^t e_i\,dt + i_o \quad (9\text{-}26)$$

If the initial inductor current is zero, then

$$e_o = iR \cong \frac{1}{L/R}\int_o^t e_i\,dt \quad (9\text{-}27)$$

and the output voltage appears as the integral of the input voltage. As for the inductive differentiator, this circuit also loses precision if the resistance of the inductor is appreciable.



(*a*)

(*b*)

FIG. 9-14. Input square wave and its integral taken by an $RC$ circuit where $RC/T = 200$.

Circuits of this nature may be used for separation of pulses having different lengths, and one such application is in the separation of the horizontal and vertical synchronizing pulses in television.

**9-7. Clipping Circuits.** A frequent application of the electron tube as a switching device is the *clipper circuit*, in which waveforms are intentionally distorted by removal, or clipping off, of a portion of the applied wave. The clipping action may be controlled as a function of potential level or polarity.

Figure 9-15*a* illustrates the essential elements of a clipper. For an assumed case the switch is on point 1 for all applied negative potentials, on point 2 for all potentials $0 < e_i < E_1$, and on point 3 for all $e_i > E_1$. As a result the output potential $e_o$ (with output terminals open) will be

| | |
|---|---|
| For $e$ negative: | $e_o = 0$ |
| $0 < e_i < E_1$: | $e_o = e_i$ |
| $e_i > E_1$: | $e_o = E_1$ |

and the resultant output characteristic is drawn in Fig. 9-15*c*. An applied sine wave of emf is shown clipped to a series of semitrapezoidal positive pulses, of amplitude fixed by $E_1$.

A diode in series with a switching potential is a suitable device to perform the action of switch $s$. As long as the anode of a diode is negative to cathode, the diode is open-circuited. When the anode becomes positive to cathode, the diode conducts with a relatively low internal resistance.

The source $e_i$ should have a high internal resistance or be supplemented by the series resistor $R$, so that the effective source resistance is made very large with respect to the diode forward resistance. Referring to

Fig. 9-15*c*, and neglecting the diode resistance, it can be seen that, for all negative values of $e_i$, diode $D_1$ short circuits the output and $e_o = 0$. For positive values of $e_i < E_1$ both diodes $D_1$ and $D_2$ are open, or have positive cathodes, and the input $e_i$ will appear at the output, assuming no load. When $e_i > E_1$, diode $D_2$ will conduct and voltage $E_1$ will appear in the output, the excess of $e_i$ over $E_1$ appearing as a drop in $R$.

Adjustment of value and polarity of $E_1$ allows clipping at any desired level, reversal of diode and battery resulting in clipping at negative levels. Several applications and their results on sine waves are shown in Figs. 9-16 and 9-17.

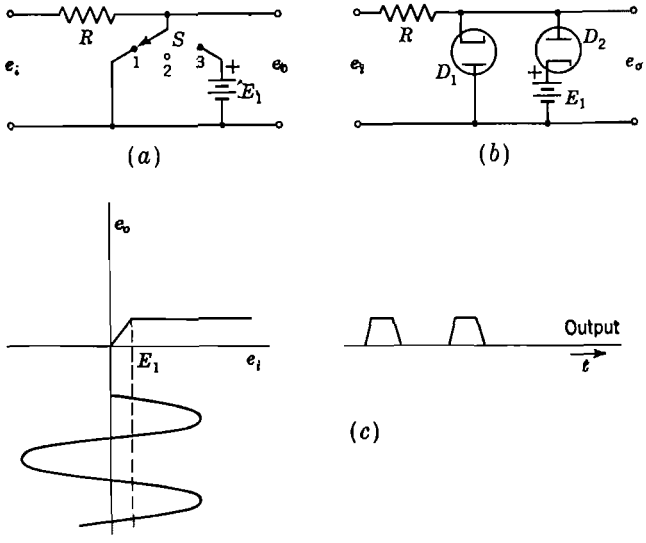



FIG. 9-15. (*a*) Basis of a clipping circuit. (*b*) Diode equivalent for (*a*). (*c*) Input-output characteristics for (*a*).

If the diode internal resistance $r_b$ is not negligible with respect to $R$, or if the excess of $e_i$ over $E_1$ is not great, then the diode voltage drop will not be small in comparison with $E_1$ and the output $e_o$ will not be flat over the clipped portion, since diode drop is a function of current. This situation appears in Fig. 9-17*b*.

Either crystal or vacuum diodes may be employed. Crystal diodes provide lower forward resistances and more perfect clipping under the conditions of the paragraph above. At the same time they have a finite back resistance, which may reduce the output voltage available during the time of diode open circuit. Vacuum-diode internal capacitance may cause trouble at high applied frequencies so that crystals may be preferred, especially in situations where the vacuum-diode cathode would be

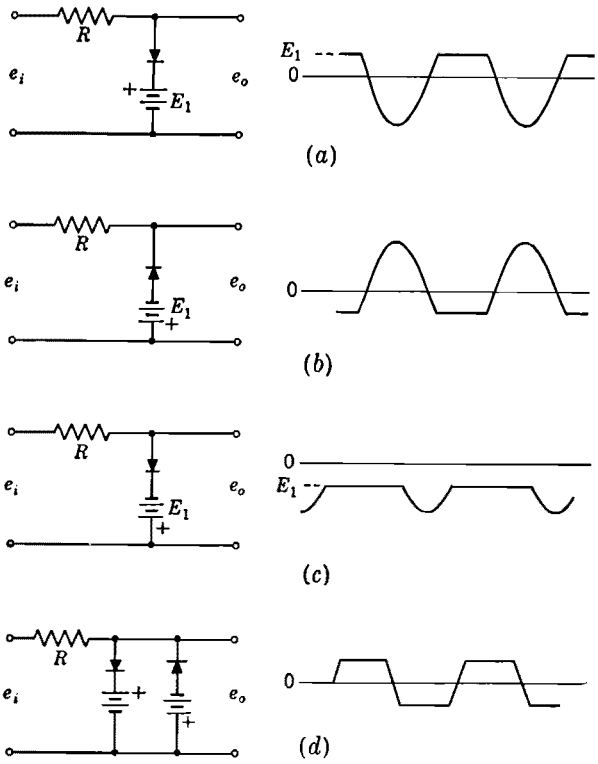
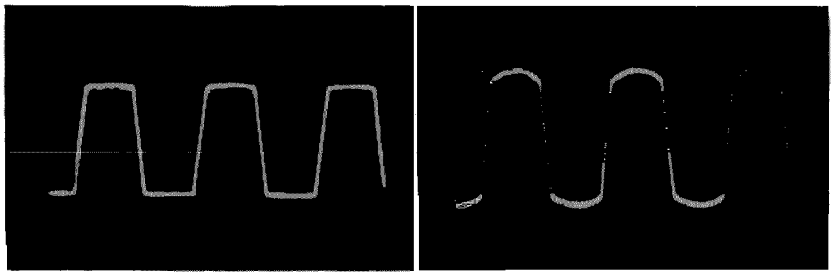



FIG. 9-16. Basic shunt-diode clipping circuits.



(a) (b)

FIG. 9-17. Clipping of a sine wave by diodes. (a) $R/r_b$ large; (b) $R/r_b$ small or diode resistance large.

operated above ground potential, with the cathode-to-heater capacitance then shunting the load.

A series form of clipper is possible, as in Fig. 9-18*a*. Since the rectifier conducts only in the direction of the arrow, only the positive portion of the input will pass, the rectifier appearing as an open circuit for nega-

tive portions. A bias may be placed as at $E$ in Fig. 9-18*b* to vary the potential at which the switch or clipper operates, as desired. In this form the series circuit is occasionally used to cut off excessive noise peaks rising above signal level, or as a *noise limiter*.

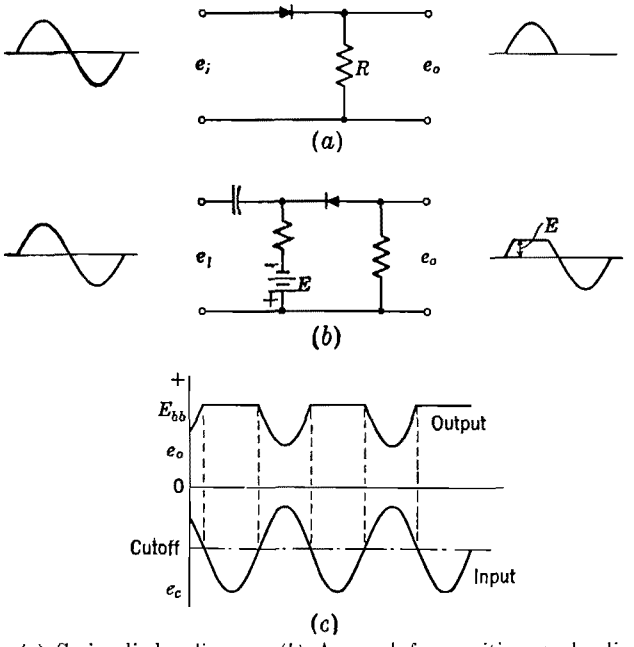



FIG. 9-18. (*a*) Series-diode clipper. (*b*) As used for positive peak clipping; (*c*) Action of biased-triode clipper.

Triode circuits are also employed as clippers. The triode is biased near cutoff, the voltage at which the applied wave is clipped being determined by the difference between $E_{cc}$ and the cutoff voltage. One-half of an input wave can be clipped in this manner, the other half being clipped in a second stage after the 180° amplifier inversion. This is known as *cutoff clipping* and may occur inadvertently in amplifiers which are accidentally overdriven. The waveforms are illustrated in Fig. 9-18*c*.

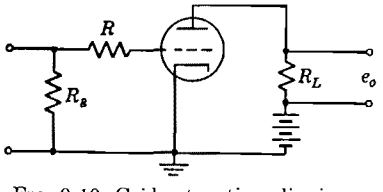



FIG. 9-19. Grid-saturation clipping.

Clipping may also be accomplished by *grid saturation* in a circuit of the form of Fig. 9-19. On positive inputs the grid and cathode act as a diode clipper, with $R$ large with respect to the grid-cathode resistance $r_c$. Clipping occurs at the zero level, the grid current on positive input producing an $i_c R$ drop equal to the positive input value. Bias may also be

inserted in series with $R_g$, so that the bias (the zero axis of the input wave) is intermediate between zero grid volts and cutoff. By driving the tube with an amplitude considerably larger than the bias value, simultaneous grid-saturation and cutoff clipping may be obtained. The results of such clipping with an applied sine wave are shown in Fig. 9-20, the effect of $r_c$ not being entirely negligible with respect to $R$ giving a slightly rounded wave on the positive input excursion. By repeating the process in a following stage after the usual 180° phase shift, both wave tops will be flat, and good square waves can be generated from an initial sine form.

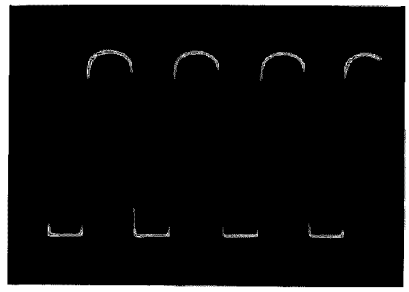


FIG. 9-20. Grid saturation and cutoff clipping of a sine wave.

Clipping circuits find an important application in television, being used to separate the synchronizing signals from the composite picture signal. The synchronizing pulses are sent between the groups of frequencies representing the scanning of a picture line and ride above these signals in amplitude. Clippers then separate the synchronizing pulses and send them to the sweep oscillator

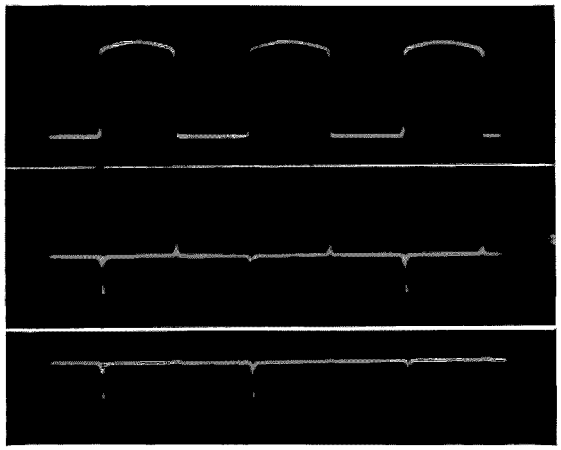


FIG. 9-21. Formation of negative pulses by grid saturation and cutoff clipping of a sine wave, followed by differentiation and rectification.

circuits, the picture signals below clipping level going to the picture circuits.

Another application in pulse forming is shown in Fig. 9-21, wherein a sine wave is clipped by grid saturation and cutoff. The resulting waveform is differentiated, giving a chain of positive and negative pulses.

These are rectified by a diode, eliminating the positive and leaving a chain of negative pulses.

**9-8. Clamping Circuits.** A circuit which will introduce a desired d-c level in an amplifier output is called a *clamper*, or *d-c restorer*. A simple form is shown in Fig. 9-22*a*, where the $RC$ time constant should be long with respect to a period of the input wave and the source should be of low resistance. The diode conducts whenever output terminal $A$ is

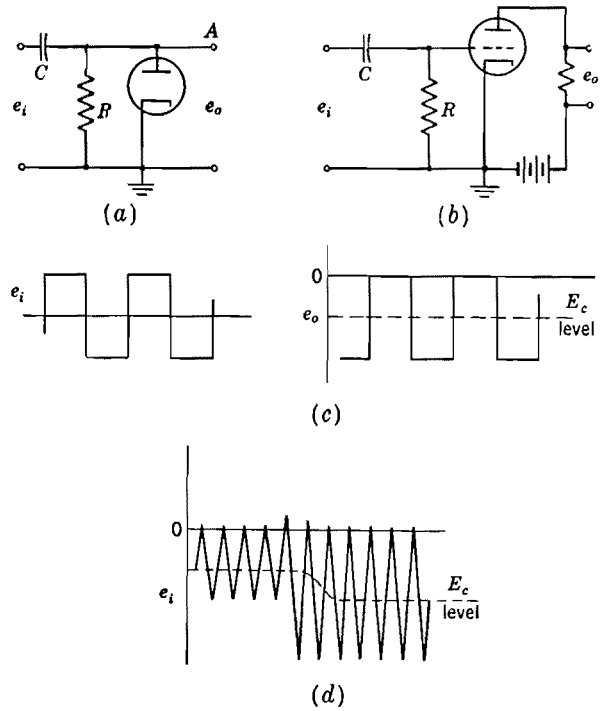



FIG. 9-22. (*a*, *b*) Diode and triode clamping circuits. (*c*) Action in clamping a square wave. (*d*) Clamping a sine wave.

positive, and connects that terminal to ground, giving essentially zero output. The full applied value of $e_i$ then appears across $C$, and $C$ charges rapidly through the low-resistance source and diode.

When the input falls to a zero or negative value, the right side of $C$ must instantaneously fall by the same amount, taking point $A$ to a negative potential. Since the diode is then open and the controlling time constant of $RC$ is long, $C$ cannot discharge appreciably in the negative interval so that $A$ remains negative. When the input $e_i$ rises with the square wave of Fig. 9-22*c*, the output once more goes up to zero. Therefore, under the assumption of $RC$ large and source resistance small, the

positive peaks of $e_i$ will be clamped at zero and a d-c axis other than zero supplied to the wave.

Sine-wave or other inputs would be likewise treated, as shown in Fig. 9-22 or 9-23. Upon a change in input amplitude a few peaks will rise to a more positive value until additional negative charge is gained by $C$ from the grid current. This additional charge causes the average grid potential $E_c$ to be shifted negative until a new equilibrium is reached at which each peak goes just enough above zero to attract a charge for $C$ sufficient to replace that which leaked off between cycles. If the input peaks fall, a few cycles will fail to reach zero, allowing charge to leak off $C$ until a new equilibrium is established at which the positive peaks again just slightly exceed, or are clamped close to, zero volts.

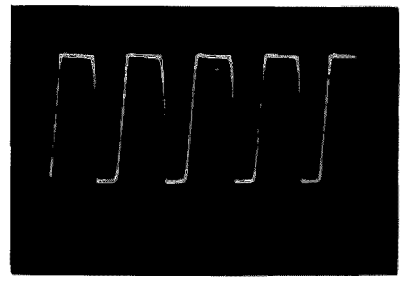


FIG. 9-23. Sine wave, clipped and clamped at zero. Note tiny pip representing charging of $C$.

By reversing the diode negative peaks may be clamped, and by inserting bias in series with the diode other d-c levels may be maintained in the output, as in Fig. 9-24.

Since the grid-cathode circuit of a triode provides the elements of a diode, it is possible to introduce the $RC$ elements into the grid circuit as in Fig. 9-22*b* and obtain clamping as well as amplifier gain.

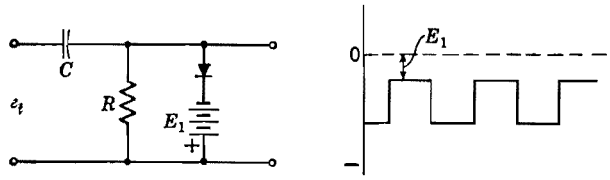



FIG. 9-24. Clamping of a wave at a value other than zero.

**9-9. Clamping in the Television System.** Television signal standards provide that, if the blanking-pulse level be taken as the base, then white is represented by a large voltage above base, gray a medium value, black occurring at the blanking level. The blanking pulse is provided to cut off the cathode-ray-tube grid during the time the beam returns from right to left so that retrace lines are not seen in the picture. Thus the blanking level must be fixed at a negative voltage sufficient to cut off the cathode beam, and white should provide a grid voltage to turn the beam on strongly.

However, upon passing a signal, such as that of Fig. 9-25*a*, through a capacitive-coupled amplifier the reference or blanking-level axis is lost

and the signals appear as in Fig. 9-25*b*, centered on an axis determined by the average value. This would give a picture in which the average illumination would always be the same and in which the blanking level could eliminate the retrace lines for some scenes and not for others.

Clamping is then resorted to, fixing the d-c level of the blanking pulses, and resulting in the waveforms of Fig. 9-25*c*, giving sufficient negative blanking voltage at all times and placing the general picture illumination at the proper level as well.

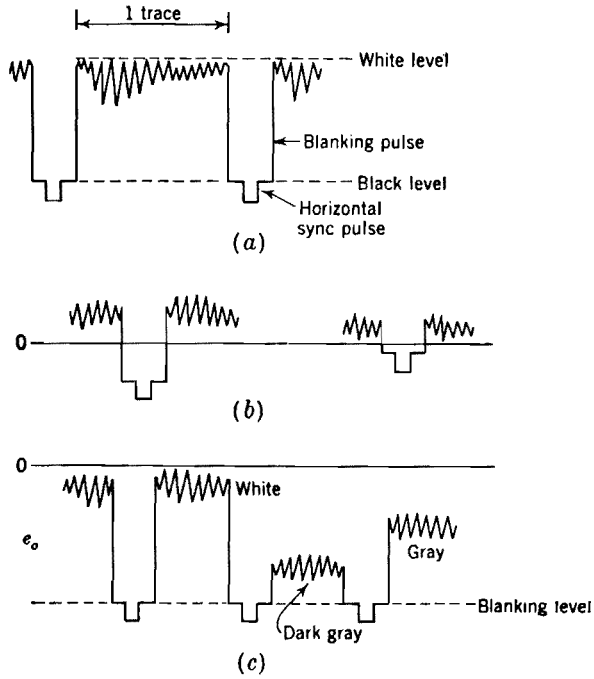



FIG. 9-25. (*a*) Video signal in television. (*b*) Television signals after passing through *RC* amplifier. (*c*) Television signals after clamping.

**9-10. The Ringing Circuit.** Another use of a triode as a switch is the *ringing circuit*, in which a tube alternately connects and disconnects an energy source to an oscillating *RLC* circuit. It is also useful as another means of generating sharp pulse chains. Analysis of the circuit illustrates a difficulty occasionally encountered in other forms of amplifier circuits, namely, the ability of a circuit to generate a damped oscillation when excited by a transient impulse.

Figure 9-26 illustrates a triode with clamping action in the grid circuit and a plate load consisting of a parallel *RLC* circuit, in which *R* is any shunted value, including the equivalent resistance of the inductor. If

this circuit is driven by a high-amplitude square wave, the positive peaks will be clamped close to zero volts by automatic adjustment of the derived $i_c R_g$ bias. The negative half cycles will be of an amplitude great enough to drive the triode beyond cutoff.

During the time the grid is at zero, the $RLC$ plate load is shunted by the rather low $r_b$ value of the triode, the load circuit being heavily damped by this tube load. A steady current $i_o$ flows through $L$ and a component $e/R$ through $R$, where $e$ is the instantaneous potential across the $RLC$ circuit.

When the grid is suddenly driven negative past cutoff, the value of $r_b$ goes to infinity and the circuit damping is much reduced, this circuit in effect being left isolated and free to oscillate. The phenomena then occurring may be investigated.

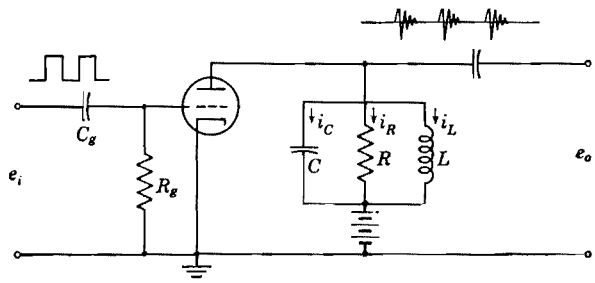



FIG. 9-26. The ringing oscillator.

Writing a current summation for the load circuit at $t = 0$, when the grid goes negative, leads to

$$i_o + \frac{1}{L}\int_o^t e\,dt + C\frac{de}{dt} + \frac{e}{R} = 0 \tag{9-28}$$

Differentiating and introducing the operator $p$ leads to

$$p^2 e + \frac{pe}{RC} + \frac{e}{LC} = 0$$

The solution proceeds as

$$\left(p^2 + \frac{p}{RC} + \frac{1}{LC}\right) e = 0 \tag{9-29}$$

from which the roots are given by

$$p = -\frac{1}{2RC} \pm \sqrt{\frac{1}{4R^2C^2} - \frac{1}{LC}} \tag{9-30}$$

Three forms are possible for the solution of Eq. (9-28). These are

obtained when

$A$: $\dfrac{1}{4R^2C^2} > \dfrac{1}{LC}$ and the circuit is *overdamped*

$B$: $\dfrac{1}{4R^2C^2} = \dfrac{1}{LC}$ and the circuit is *critically damped*

$C$: $\dfrac{1}{4R^2C^2} < \dfrac{1}{LC}$ and the circuit is *underdamped*

Cases $A$, $B$, and $C$ may occur depending on circuit constants. However, case $C$ is of greatest interest, and the radical of Eq. (9-30) then becomes

$$\sqrt{\frac{1}{4R^2C^2} - \frac{1}{LC}} = j\sqrt{\frac{1}{LC} - \frac{1}{4R^2C^2}} = j\omega$$

For case $C$ the solution of the differential equation is

$$e = (A_1\epsilon^{j\omega t} + A_2\epsilon^{-j\omega t})\epsilon^{-t/2RC}$$

which may be converted to

$$e = (A_3 \cos \omega t + A_4 \sin \omega t)\epsilon^{-t/2RC} \tag{9-31}$$

Since $L$ has been made resistanceless by inclusion of its losses in $R$, it is known that, at $t = 0$, $i_L = i_o$, a constant and thus, at $t = 0$, $e = L\, di/dt = 0$. Applying this condition to Eq. (9-31) shows that $A_3 = 0$, and Eq. (9-31) may then be written as

$$e = (A_4 \sin \omega t)\epsilon^{-t/2RC} \tag{9-32}$$

which may be differentiated to produce $de/dt$ as

$$\frac{de}{dt} = \left(-\frac{A_4}{2RC} \sin \omega t + A_4\omega \cos \omega t\right)\epsilon^{-t/2RC}$$

Evaluated at $t = 0$, this becomes

$$\left.\frac{de}{dt}\right]_{t=0} = A_4\omega$$

With $e = 0$ at $t = 0$ the original equation (9-28) reduces to

$$\left.\frac{de}{dt}\right]_{t=0} = -\frac{i_o}{C} = A_4\omega$$

and so the coefficient is evaluated as

$$A_4 = -\frac{i_o}{\omega C} \tag{9-33}$$

The solution to the differential equation is then

$$e = \left(-\frac{i_o}{\omega C} \sin \omega t\right)\epsilon^{-t/2RC} \tag{9-34}$$

This result is oscillatory in nature, being damped by the exponential term. The frequency of oscillation may be determined from $\omega$ as

$$f = \frac{1}{2\pi}\sqrt{\frac{1}{LC} - \frac{1}{4R^2C^2}} \tag{9-35}$$

This is the frequency of free oscillation of the circuit, which approaches the unity-power-factor resonant frequency as the losses are reduced. This frequency is nearly independent of changes in tube characteristics and supply voltages, except for the output capacitance of the tube.

The output voltage, while oscillatory, is damped by the factor $\epsilon^{-t/2RC}$, and an output is shown in Fig. 9-27. As a source of single pulses rather

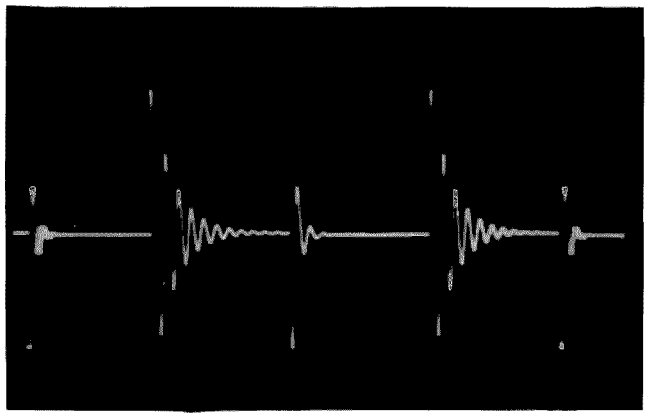


FIG. 9-27. Ringing-oscillator output.

high damping may be used and the output waveform clipped to provide only a pulse due to the peak of the first oscillation. This may be differentiated, giving sharp pulses, which always lag the negative-going portion of the input square wave by a fixed time. Thus the circuit may supply a fixed time delay as well.

If the length of the input square wave is sufficient for steady-state conditions to be reached during the positive or heavily damped period, then the initial inductor current $i_o$ may be found from the tube curves for zero grid voltage. The maximum circuit output voltage, obtained during the first oscillatory cycle, can be calculated by noting that for high $Q$ or small losses, as for the oscillatory case, the maximum energy stored during the first oscillation in the inductance is substantially transferred during the next quarter cycle to the capacitor, or that

$$\frac{Li_o^2}{2} = \frac{Ce_{\max}^2}{2}$$

from which

$$e_{\max} = -i_o\sqrt{\frac{L}{C}} \tag{9-36}$$

It is necessary that the input square wave have amplitude sufficient to drive the tube beyond cutoff during even the peak of the first oscillation. At this instant the voltage of Eq. (9-36) adds to $E_{bb}$, so that the input peak-to-peak amplitude must be given by

$$E_{p\text{-}p} > \frac{E_{bb} + i_o\sqrt{L/C}}{\mu} \tag{9-37}$$

In the discussion of square-wave-testing applications in Chap. 5 it was noted that the transients produced by such wave application might be oscillatory at times. It can be seen that the presence of an *RLC* circuit in an amplifier may cause an oscillation when impulse-excited by a transient and produce spurious output and distortion.

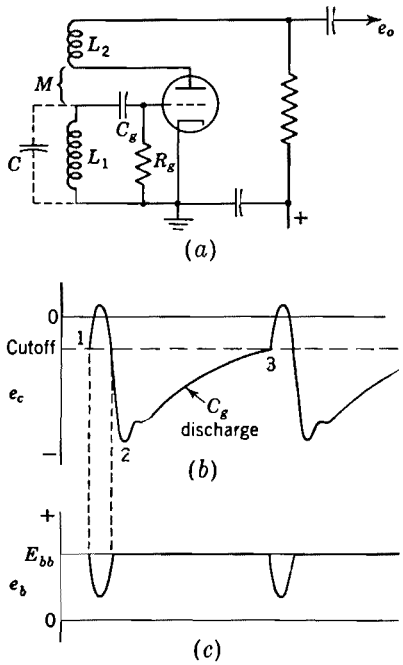



FIG. 9-28. (*a*) Blocking-oscillator circuit. (*b*) Grid-voltage waveform. (*c*) Plate-voltage waveform.

From Fig. 9-27 it may also be seen that the positive-going portion of the square wave also produces a ringing transient, but this is much more heavily damped since at that instant the *RLC* circuit is shunted by $r_b$ of the tube.

**9-11. The Blocking Oscillator.** Another circuit form in which the switch properties of a triode are employed to produce accurate narrow pulses is the *blocking oscillator*. The basic circuit is usually one employing magnetic feedback from plate to grid circuit, with a low $Q$ value of the *RLC* tuned circuit at the grid, so that highly damped oscillations are produced.

The circuit is shown in Fig. 9-28, wherein the $C_gR_g$ time constant must be long with respect to the time of a cycle of the resonant frequency of the grid circuit. It is usual to employ the internal distributed capacitance $C$ of the coil for resonant tuning. The mutual coupling or magnetic feedback between $L_1$ and $L_2$ must be large and so connected as to give positive feedback. Any change in plate current will induce a voltage on the grid to aid the plate-current change. An increasing plate current will tend to make the grid positive, further increasing the plate current, and at the same time charging capacitor $C_g$ from the current flowing to the positive grid. When the plate current becomes constant

at saturation value, the $L\,di/dt$ voltage in $L_1$ falls to zero. Because there is no longer an induced positive grid potential, $C_g$ starts to discharge through $R_g$. The grid becomes less positive, reducing plate current, and a falling plate current induces a negative voltage in $L_1$, driving the grid negative and beyond cutoff. Oscillations tend to occur in the tuned grid winding as a ringing circuit but are heavily damped by the low $Q$ value. All this is illustrated in the portion of the curve from 1 to 2 in Fig. 9-28*b*.

Since with the grid negative the only discharge path for $C_g$ is through $R_g$, the discharge may be slow and the tube will remain cut off for a considerable period. When the $C_g$ charge leaks off to the point where conduction once more starts at point 3, a positive voltage will again be induced by the feedback into the grid circuit and the action will be repeated.

The interval between pulses of Fig. 9-28*b* will be largely determined by the $C_gR_g$ time constant and will be independent of resonant-circuit and tube characteristics. The pulses may be synchronized with external phenomena by introduction of pulse signals in the grid circuit, as illustrated in Fig. 9-29, where the oscillator output is synchronized with every second input pulse. This gives accurate frequency control, or timing, of the output pulses and also provides a means of reshaping input pulses which have been distorted in transmission.

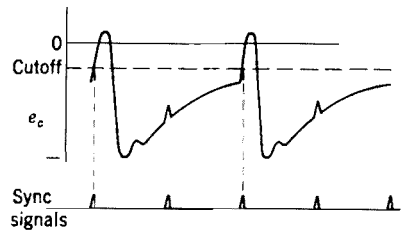


FIG. 9-29. Synchronizing the blocking oscillator with pulses.

**9-12. Time Base or Sweep Generation for Cathode-ray-tube Deflection.** The deflection of the electron spot for electrostatic cathode-ray tubes is given by Eq. (3-32), which may be written as

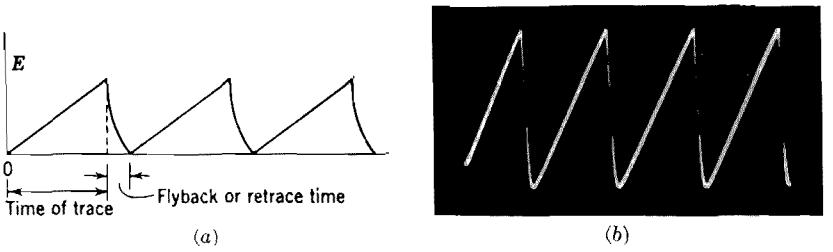
$$D = K_1E_d$$

all quantities in the equation other than $E_d$ being included in $K_1$. If $E_d$ is made a linear function of time, or $E_d = K_2t$, then

$$D = K_1K_2t \tag{9-38}$$

and the deflection will be a linear function of time. This is the requirement for the sweep wave for many cathode-ray applications and as actually employed becomes the repetitive saw-tooth wave of Fig. 9-30. This wave will cause voltages applied to the $y$-axis deflection plates to be plotted as functions of time, the repetitive saw tooth being adjusted in frequency such that $n$ cycles of the $y$-axis voltage appear on the screen for each saw tooth. With the form shown, the electron spot is moved in the $x$ direction across the screen at a uniform velocity, and with a speed

depending on the slope of the saw tooth. The spot is caused to return very rapidly to the starting point, and the return trace usually is not seen, or is seen at much reduced brilliance due to its high writing speed during the retrace time.

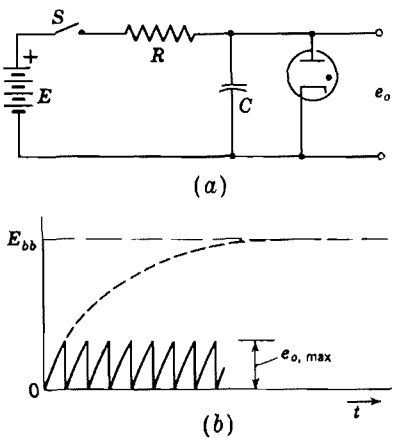



(a) (b)

FIG. 9-30. (a) Saw-tooth-sweep voltage wave. (b) Saw-tooth sweep from an actual oscillograph.

The ideal saw tooth is not easily obtained by direct electronic means, and various methods have been employed to achieve near linearity. Several of these methods will be discussed in the following sections.

**9-13. Simple *RC* Sweep Generation.** It is possible to employ the potential across a capacitor $C$, charged through a resistor $R$ from a constant-potential source, as an approximate but simple sweep source, one example appearing in Fig. 9-31. Suppose that $C$ is initially uncharged when $s$ is closed. The potential across $C$ will rise according to




(a) (b)

FIG. 9-31. (a) *RC* charging circuit with gas diode for discharge. (b) Achieving near-linear saw tooth from *RC* circuit.

$$e_o = e_C = E(1 - \epsilon^{-t/RC}) \quad (9\text{-}39)$$

When $e_o$ reaches the breakdown potential of the gas diode, current will flow through the tube. The discharge is accompanied by a rapidly falling capacitor voltage, because of the near short circuit imposed by the diode. When $e_o$ falls below the extinction potential of the diode, the discharge ceases and $e_o$ begins to rise as $C$ charges, and the operation is repeated. The discharge current is limited only by the resistance and inductance of the diode wiring.

The process is periodic and results in saw-tooth waves of period $T$, equal to the time required for the capacitor to charge from zero (considering the diode extinguishing potential as small) to $e_{o,\max}$. From Eq.

voltage without the bias battery. To place the tube $T_2$ at a $Q$ point of $E_c = -10$ volts requires that $E_{cc} = -110$ volts, which is a large battery.

The battery requirement is not a desirable feature and would make the circuit also suffer from high shunt capacity, due to the capacity of the

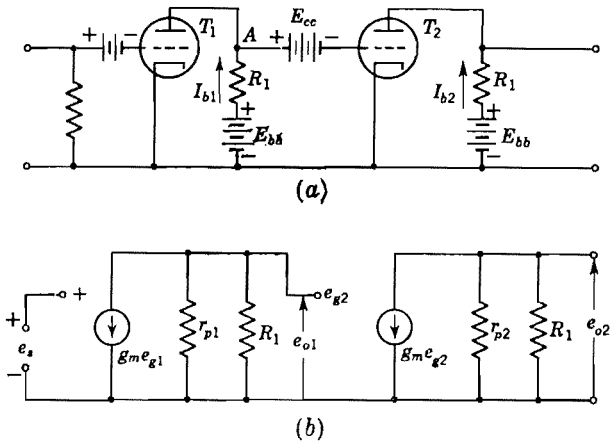



FIG. 8-1. Two-stage direct-coupled amplifier.

grid-bias batteries to ground. This would seriously limit the transient and high-frequency response.

Another version, more suited to operation from a rectified a-c supply, is the Loftin-White circuit of Fig. 8-2. A single d-c supply furnishes all grid and anode voltages by means of a tapped resistor. The current $I$ bled through this resistor should be several times the total current demand of the amplifier, in order to make the tapped resistor low in value and to

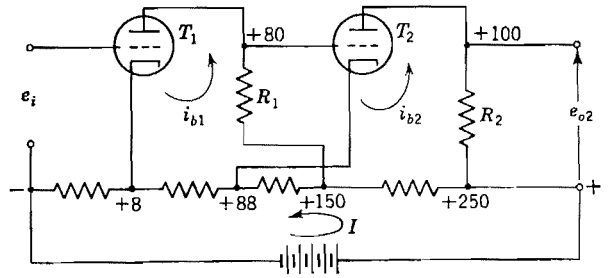



FIG. 8-2. The Loftin-White form of d-c amplifier.

establish the supply voltages as somewhat independent of the tube currents.

A typical set of voltages is indicated in the figure. The cathode of $T_1$ is at +8 volts, determining the grid bias of this tube. The anode supply

voltage of $T_1$ is $150 - 8 = +142$ volts. The $I_{b1}R_1$ drop is 62 volts, making the potential $E_{b1} = +80$ volts, and this is also the potential of the grid of $T_2$ above negative reference. In order to obtain $-8$ volts bias on $T_2$, its cathode is returned to a tap on the resistor at $+88$ volts, 8 volts above its grid. The anode of $T_2$ is supplied at $+250$ volts, or $250 - 88 = 162$ volts above its cathode. If the $I_{b2}R_2$ drop is 62 volts, then potential $E_{b2} = +100$ volts, and

$$e_{o2} = +100 - (+250) = -150 \text{ volts output}$$

Such circuits will amplify extremely low frequencies and introduce no frequency and phase distortion. However, owing to the high supply voltage required for multistage amplifiers, the high operating voltage between cathodes and heaters, and the waste of power in the dividing resistor, the circuit has been superseded by other types.

These circuits also suffer from the inherent instability of all d-c amplifiers, due to tube-characteristic drift, supply-voltage change, and heater-voltage change with cathode-type tubes. Changes in supply voltage move the operating point in just the same way as does a signal, so that the amplifier cannot tell the difference between a change in bias and a change in signal, and the output drifts. Zero adjustment features are usually required to restore the output voltage to zero for zero input, or to compensate for internal drift.

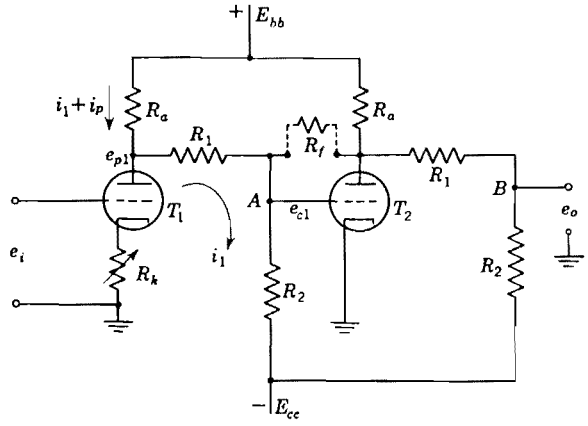



FIG. 8-3. Resistance coupling with positive and negative power supplies.

**8-2. The Resistance-coupled Amplifier.** In modern d-c amplifiers many basic circuit principles are combined, and it is the intent of this and following sections to discuss a few of the most common.

To overcome the common coupling and feedback possible through the tapped supply resistor of the Loftin-White circuit, it is possible to use the resistance-coupled circuit of Fig. 8-3. This requires two power

supplies, one positive to ground, the other negative to ground. By adjustment of the bias of $T_1$ with a variable cathode resistor, it is possible to shift the potential level at point $B$, so that $B$ is brought to ground or midway between the two power-supply voltages, for zero input signal. The output voltage will then also be zero. If regulated power supplies are used and the tubes well aged, the output may drift only a fraction of a volt in a few hours. Occasional manual adjustment will then cancel the drift.

Obviously, there is a loss in gain through introducing the voltage divider $R_1$, $R_2$, but this and the dual power supply are the price to be paid for gaining some control of the internal d-c potential levels. The gain of each stage can readily be found as

$$\begin{aligned} A &= \frac{-\mu R_a(R_1 + R_2)/(R_a + R_1 + R_2)}{r_p + [R_a(R_1 + R_2)/(R_a + R_1 + R_2)]} \frac{R_2}{R_1 + R_2} \\ &= \frac{-\mu R_a R_2/(R_a + R_1 + R_2)}{r_p + [R_a(R_1 + R_2)/(R_a + R_1 + R_2)]} \end{aligned} \tag{8-2}$$

and for high gain $R_2 \gg R_1$. Thus d-c amplification can be obtained with this sort of circuit with a common anode-voltage supply, without introducing biasing complications, and with provision for manually balancing the internal drift by use of a variable bias in the first stage.

Referring to Fig. 8-3, if $e_{p1}$ and $e_{c1}$ are the potential levels above ground reference for the plate of the first stage and the grid of the second, respectively, then suitable values for $R_1$ and $R_2$ may be computed from the equations

$$\begin{aligned} (i_1 + i_p)R_a &= E_{bb} - e_{p1} \\ e_{p1} - e_{c1} &= i_1 R_1 \\ e_{c1} - E_{cc} &= i_1 R_2 \end{aligned}$$

from which

$$\begin{aligned} \frac{R_1}{R_a} &= \frac{e_{p1} - e_{c1}}{E_{bb} - e_{p1} - i_p R_a} \\ \frac{R_2}{R_a} &= \frac{e_{c1} - E_{cc}}{E_{bb} - e_{p1} - i_p R_a} \end{aligned}$$

Negative feedback may be added as indicated by the dashed resistor $R_f$ in Fig. 8-3, to improve stability. Because of large values of $\mu$ available in many tubes, the gain of the circuit can be large even though part of this gain is sacrificed to place the desired grid and plate voltages on the tube. In a circuit such as shown in the figure the $i_p R_a$ drop might be equal to one-half of $E_{bb}$. Suitable values of $R_1$ and $R_2$ then result to place point $B$ at ground potential with zero signal. The circuit is frequently employed.

Figure 8-4 shows a variation of the previous circuit, replacing $R_1$ with a gas-discharge tube of the voltage-regulator type. When sufficient cur-

rent is flowing through such a tube, it tends to maintain a constant voltage between its electrodes. Thus the grid of $T_2$ may be dropped to the desired d-c level by the d-c drop in the neon tube. Since the voltage drop across the neon bulb is constant, any *changes* in potential at the anode of $T_1$ will also appear across $R_2$ and thus there is no loss in gain due to the voltage-divider action of $R_2/(R_1 + R_2)$ as in the previous circuit. Hence the gain is greater, but this is obtained at the expense of some noise due to random and minute voltage fluctuations in the voltage-regulator tube. Therefore the circuit is not well suited to the amplification of very small input voltages.

**8-3. Cathode Coupling in D-C Amplifiers.** Another type of circuit often found as an element of d-c amplifiers is the cathode-coupled amplifier of Fig. 8-5. Essentially, it is a cathode follower placed ahead of a grounded-grid amplifier and introduces d-c coupling and considerable gain without serious circuit complications. Assuming the use of a double triode of identical parameters, from Fig. 8-5 the circuit equations are

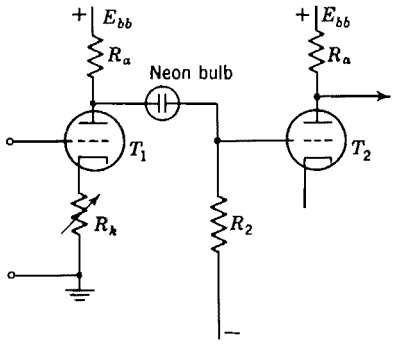



FIG. 8-4. Use of a gas-discharge tube in a resistance-coupled d-c amplifier.

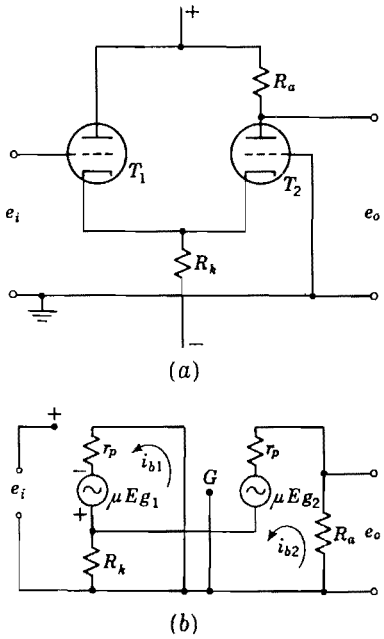



FIG. 8-5. The cathode-coupled circuit and its equivalent.

$$e_{g1} = e_i - (i_{b1} + i_{b2})R_k \tag{8-3}$$

$$e_{g2} = -(i_{b1} + i_{b2})R_k \tag{8-4}$$

$$\mu e_{g1} = i_{b1}(r_p + R_k) + i_{b2}R_k \tag{8-5}$$

$$\mu e_{g2} = i_{b1}R_k + i_{b2}(r_p + R_a + R_k) \tag{8-6}$$

from which

$$i_{b1} = -i_{b2}\frac{r_p + R_a + (\mu + 1)R_k}{(\mu + 1)R_k}$$

and $$\mu(\mu + 1)R_k e_i = -i_{b2}\{[r_p + (\mu + 1)R_k][r_p + R_a + (\mu + 1)R_k] - (\mu + 1)^2 R_k^2\}$$

The gain of the circuit is

$$A = \frac{e_o}{e_i} = \frac{-i_{b2}R_a}{e_i} = \frac{\mu R_a}{[(r_p + R_a)/(\mu + 1)](r_p/R_k) + 2r_p + R_a} \tag{8-7}$$

For usual circuit and tube parameters the first term in the denominator is small and may be dropped, giving

$$A = \frac{\mu R_a}{2r_p + R_a} \tag{8-8}$$

Amplification is achieved, and again it is possible to use common anode-voltage supplies. Two triodes are required, but the circuit is simple and frequently employed in instruments where the stability, introduced by the negative current feedback from the cathode resistor, is desirable.

The principle is further demonstrated in Fig. 8-6.

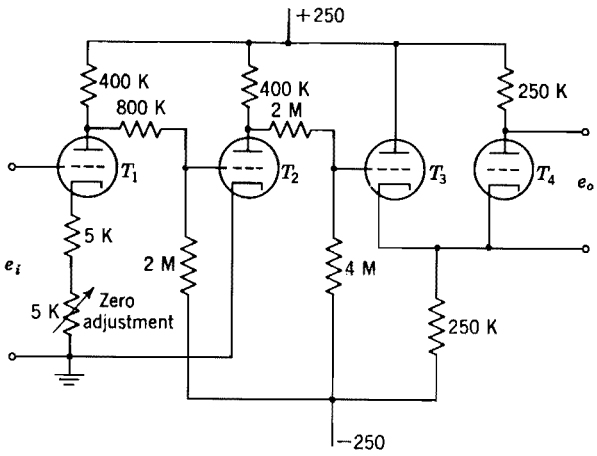



FIG. 8-6. A complete d-c amplifier employing resistance coupling and cathode coupling.

**8-4. The Balanced Push-pull Amplifier.** A balanced or push-pull type of operation is sometimes employed in an attempt to overcome zero drifts in d-c amplifiers due to changing anode-supply and heater voltages. In the circuit of Fig. 8-7 tubes $T_1$ and $T_2$ constitute a phase inverter to invert the single-ended, or grounded, input voltage into two voltages, $e_1$ and $e_2$, balanced to ground. Tubes $T_3$ and $T_4$ form a resistance-coupled and balanced d-c amplifier stage. By use of the zero adjustment in the first stage the output voltage $e_o$ can be made zero for zero input $e_i$.

The circuit has the disadvantage that both output terminals are above

ground potential. If further stages are to follow, or if the output is to be supplied to an indicating instrument or the plates of a cathode-ray oscillograph, this is not important.

If it be assumed that all tubes have matched characteristics, then it may be reasoned that the output will be independent of supply-voltage changes. Actually this is a condition difficult to achieve because of tube dissimilarity, but some reduction of drift is realized. The circuit does not give protection from individual tube parameter changes.

A simpler version of a balanced d-c amplifier, sometimes called a *difference amplifier*, is derived from the cathode-coupled phase inverter of Sec. 6-19 and is related to the circuit of Fig. 8-5. This is shown in Fig. 8-8, where $M$ is an indicating meter for showing amplifier output,

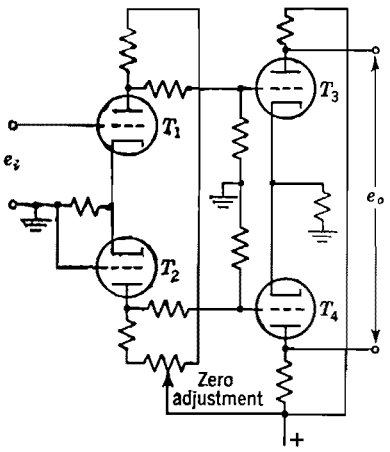



FIG. 8-7. A push-pull d-c amplifier.

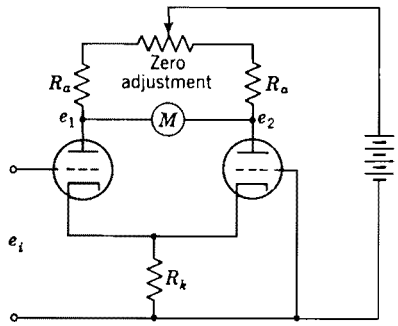



FIG. 8-8. A balanced circuit employing an indicating instrument.

since the circuit is commonly employed in vacuum-tube voltmeters. By use of the adjustment the instrument can be brought to zero for zero input voltage.

The analysis leading to Eqs. (6-60) and (6-61) provides expressions for the tube currents. If the two plate-load resistors are made equal to $R_a$, then the currents may be written as

$$i_1 = \frac{\mu e_s[R_a + r_p + (\mu + 1)R_k]}{2(R_a + r_p)[1 + (\mu + 1)R_k]} \tag{8-9}$$

$$i_2 = \frac{-\mu(\mu + 1)e_s R_k}{2(R_a + r_p)[1 + (\mu + 1)R_k]} \tag{8-10}$$

Referring to Fig. 8-8, the potentials at $e_2$ and $e_1$ are

$$e_2 = E_{bb} - i_2 R_a$$
$$e_1 = E_{bb} - i_1 R_a$$

and the potential applied to $M$ is

$$e_M = e_2 - e_1 = (i_1 - i_2)R_a \tag{8-11}$$

and use of the currents leads to

$$e_M = \frac{\mu e_s R_a[R_a + r_p + 2(\mu + 1)R_k]}{2(R_a + r_p)[1 + (\mu + 1)R_k]}$$

Circuit design and choice of a reasonable value of $\mu$ will lead to the condition that $(R_a + r_p) \ll (\mu + 1)R_k$, and the above reduces to

$$e_M = \frac{\mu e_s R_a}{R_a + r_p} \tag{8-12}$$

as a simple result, neglecting the loading effect of the instrument current. With matched tubes, changes in plate-supply potential have no effect on the instrument readings. Since a double triode is used, changes in heater potential are also reasonably balanced and the circuit is frequently employed in indicating instruments.

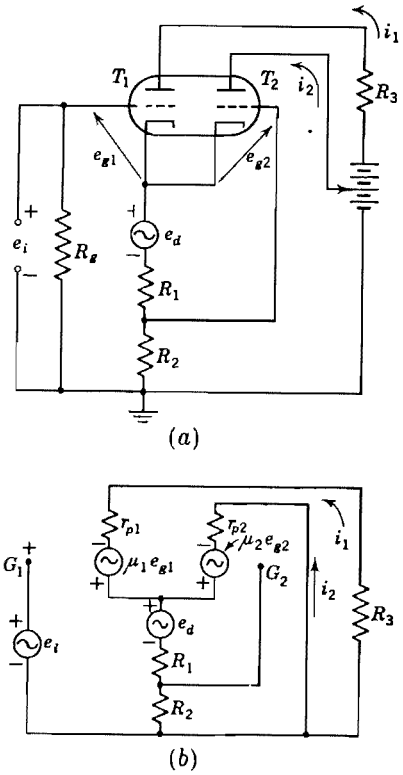



FIG. 8-9. Analysis of Miller circuit for cathode-drift compensation.

**8-5. Cathode-drift Compensation.** One of the problems associated with the operation of sensitive d-c amplifiers is the erratic drift in output voltage caused by random changes in cathode emission. The latter changes may be due to small heater-voltage changes or to actual changes in the internal emission mechanisms. The effect is usually considered as located in the first tube, since at that point any drift is multiplied by the whole gain of the amplifier and is therefore more serious than when it occurs in following tubes.

The Miller compensation circuit of Fig. 8-9 provides a means of overcoming cathode drift in a d-c amplifier by use of an additional triode. Usually a double triode is employed, and it is necessary to assume that both cathodes have identical drift characteristics and that the drift can be simulated, for analysis purposes, by an equivalent random-drift gener-

ator $e_d$ in the common cathode lead of the two tubes. It may then be considered that the second tube amplifies the drift voltage and feeds it back in opposition to the drift.

Using the equivalent circuit of Fig. 8-9*b*, an analysis may be made to determine the conditions which must be met for drift cancellation. Writing mesh equations with no external signal applied,

$$e_{g1} = -(i_1 + i_2)(R_1 + R_2) - e_d + e_i \tag{8-13}$$

$$e_{g2} = -(i_1 + i_2)R_1 - e_d \tag{8-14}$$

$$i_1(r_{p1} + R_1 + R_2 + R_3) + i_2(R_1 + R_2) + e_d = \mu_1 e_{g1}$$

$$i_1(R_1 + R_2) + i_2(r_{p2} + R_1 + R_2) + e_d = \mu_2 e_{g2}$$

Solving for $i_1$,

$$i_1 = \frac{\mu_1 e_i[r_{p2} + R_2 + (\mu_2 + 1)R_1] + (\mu_1 + 1)e_d(\mu_2 R_2 - r_{p2})}{(r_{p1} + R_3)[r_{p2} + R_2 + (\mu_2 + 1)R_1] + (\mu_1 + 1)r_{p2}(R_1 + R_2)} \tag{8-15}$$

It is desired that the current $i_1$ be independent of changes in the drift voltage $e_d$, so that $\partial i_1/\partial e_d = 0$ is obtained. Therefore

$$\frac{\partial i_1}{\partial e_d} = \frac{(\mu_1 + 1)(\mu_2 R_2 - r_{p2})}{\text{denominator}} = 0$$

and this leads to the condition that

$$R_2 = \frac{1}{g_{m2}} \tag{8-16}$$

for cancellation of the cathode drift, where $g_{m2}$ is the transconductance of the second, or compensation, triode. This condition is easily complied with.

If the tube is a double triode and $\mu_1 = \mu_2$, $r_{p1} = r_{p2}$, and if Eq. (8-16) is satisfied, or $g_{m2}R_2 = 1$, the gain expression simplifies to

$$A = \frac{-i_1 R_3}{e_i} = \frac{-\mu R_3}{2r_p + R_3} \tag{8-17}$$

showing the presence of current feedback, by reason of the increase of $r_p$ over that of one tube.

For d-c amplifiers having extremely small input signals, this method of compensation is helpful. By its use heater-voltage variations can be compensated, the residual change in $i_1$ being less than 0.5 per cent. This residue appears because $g_m$ varies somewhat with available electron emission, and this may make the compensating resistor $R_2$ incorrect for an extreme range of heater voltages.

**8-6. Series-cathode Compensation; the Bridge-balanced Amplifier.** The Miller compensator may be considered as a parallel method of compensation for heater and emission variations. Compensation can also be

obtained through use of an identical tube in series, resulting in a *bridge-balanced* d-c amplifier, as shown in Fig. 8-10. Here $T_1$ is the amplifier and $T_2$ the compensator. The circuit appears as a bridge, considering $T_1$, $T_2$, and the two resistors $R$ as the bridge arms.

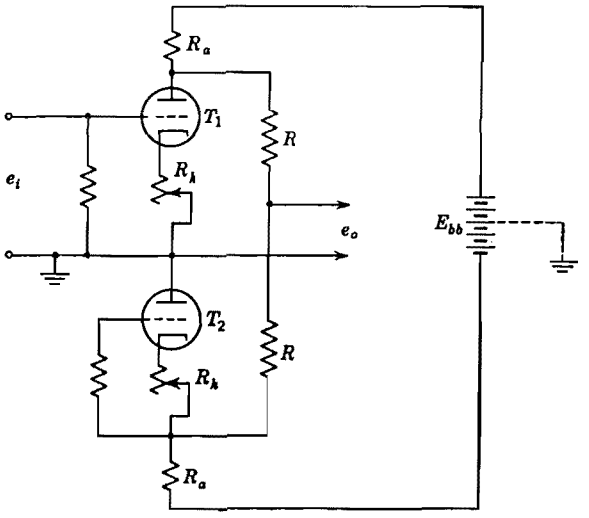



FIG. 8-10. Bridge amplifier using series compensation for cathode drift.

The output can be brought to zero volts for zero input by use of the variable cathode resistors. The output is at ground potential and may swing either positive or negative depending on the input $e_i$. If it be assumed that heater changes occur similarly in both tubes, the bridge will remain balanced or compensated for heater changes. For this reason double triodes are frequently used. Voltage-supply changes are likewise compensated if the positive and negative supplies do not have a grounded center tap or are floating. However, if more than one stage is to be operated in cascade from the same supply, a fixed center tap is required and then each supply should be well regulated.

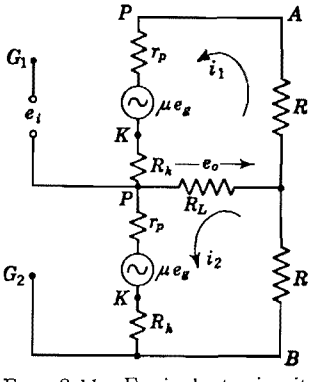



FIG. 8-11. Equivalent circuit for the bridge amplifier.

If a load $R_L$ is connected across the output, the equivalent circuit becomes that of Fig. 8-11, with the anode supply floating. Because of the balanced nature of the circuit the potential difference across $A$-$B$ is constant, and resistors $R$ do not enter into

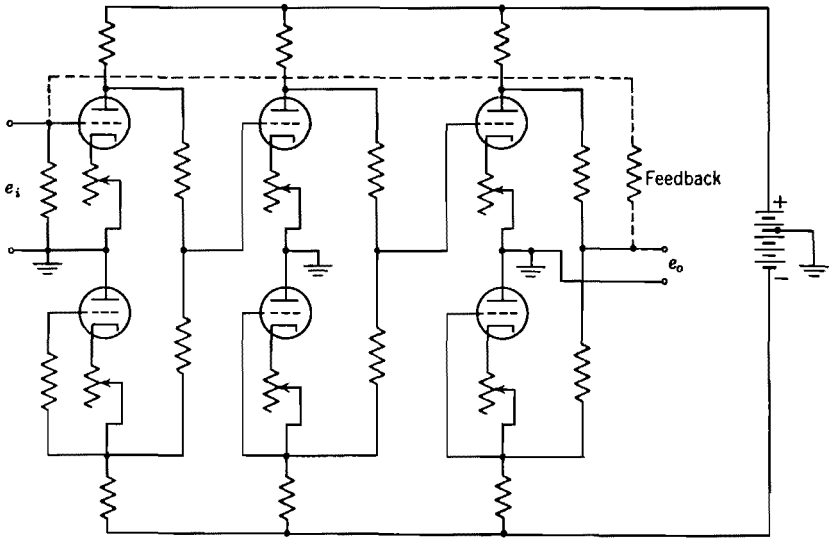



FIG. 8-12. Multistage bridge amplifier with negative feedback.

the varying current equivalent circuit. Writing mesh equations for identical tubes gives

$$e_{g1} = e_i - i_1 R_k \tag{8-18}$$

$$e_{g2} = -i_2 R_k \tag{8-19}$$

$$\mu e_{g1} = i_1(r_p + R_k + R + R_L) - i_2 R_L \tag{8-20}$$

$$\mu e_{g2} = i_2(r_p + R_k + R + R_L) - i_1 R_L \tag{8-21}$$

Solving for $i_1$ and $i_2$ and noting that

$$e_o = (i_2 - i_1) R_L$$

then

$$e_o = \frac{\mu e_i R_L \{R_L - [r_p + (\mu + 1)R_k + R + R_L]\}}{[r_p + (\mu + 1)R_k + R + R_L]^2 - R_L^2}$$

and the gain is

$$A = \frac{e_o}{e_i} = \frac{-\mu R_L}{r_p + (\mu + 1)R_k + R + 2R_L} \tag{8-22}$$

If the amplifier is employed to excite a following stage, as in Fig. 8-12, $R_L$ may be considered very large and the gain then reduces to

$$A \cong -\frac{\mu}{2} \tag{8-23}$$

When the anode supply is operated with grounded center tap, the resistors $R$ become part of the equivalent circuit and it is possible to

obtain the gain as

$$A = \frac{-\mu R}{2[r_p + (\mu + 1)R_k + R]} \tag{8-24}$$

for $R_L$ large. This is just one-half of the gain of a conventional cathode-feedback amplifier. The circuit is easy to adjust to zero and has considerable value for compensating slow cathode drifts.

When pentodes are employed in the circuit of Fig. 8-10, with stabilized screen voltages, Eq. (8-23) shows that the gain may be very high because of the high $\mu$ of the pentode, and values of 4000 can be reached.

The circuit of Fig. 8-12 illustrates three cascaded stages using bridge stabilization of zero. In addition the dashed connection may be employed for negative feedback to stabilize the gain. Such circuits are very insensitive to supply-voltage changes but are expensive because of the large number of tubes and components involved.

**8-7. Vacuum-tube Electrometers.** Some of the first applications of d-c amplifiers were as electrometers, to measure electrostatic potentials or currents of the order of $10^{-13}$ amp or less. The vacuum tube is a voltage-operated device; to produce an appreciable potential from such small currents requires that they pass through an extremely high resistance, and values up to $10^{14}$ ohms are used. It was quickly found that the grid currents in the usual tubes, even with quite negative grids, might greatly exceed the magnitude of the current to be measured, owing to gas ions formed by the tube high voltages, photoelectric emission from the grid, and leakage currents across the insulating surfaces of the tube.

Special electrometer tubes were developed, employing low electrode potentials and filamentary cathodes which operated at low temperatures and gave off almost no light to excite photoelectric emission, and with external leakage minimized by surface treatments of the glass.

Electrometer tubes usually are employed as tetrodes, with the grid nearest the cathode operated at a low positive potential, forming a virtual cathode or electron source in its region. The positive potential also repels positive ions emitted by the cathode and prevents them from reaching the second, or control, grid. The control grid is operated negatively and the plate at a positive potential of a few volts, below the ionizing potential of any residual gas. The amplification factor is about unity, the tube serving not as a voltage amplifier but as an impedance transformer, to step down the very high impedance of the grid circuit to that of the indicating instrument. By these methods grid currents as indicated in Fig. 8-13 may be obtained.

A typical circuit for an electrometer tube is shown in Fig. 8-14, all potentials being supplied by one battery source. Resistor $R$ is the high-value input resistor, which must be carefully protected from dirt, mois-

ture, and other contamination if surface leakage is not to reduce its value. The indicating microammeter $M$ is in a *bucking circuit*, whereby its current may be brought to zero for zero voltage input. Resistors $R_3$ and $R_4$, $R_2$, and $R_1$ and $R_x$ constitute three arms of a bridge, with the tube as a fourth arm. Adjustment of $R_x$ for zero input voltage balances the bridge and brings the meter $M$ to zero. Thus an instrument having a full-scale rating much less than the static plate current may be used, the meter reading only the change in current and increasing the circuit sensitivity.

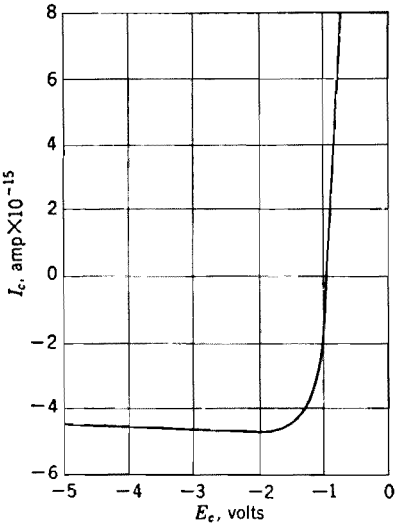



FIG. 8-13. Control-grid current in an electrometer tube; $E_b$ = 6 volts.

Since the output of such a circuit is current, as indicated on an instrument, the over-all circuit transconductance, or milliamperes output change per input grid volts change, is of importance. For an amplifier with an indicating instrument of resistance $R_M$ in its anode circuit, the plate current as obtained from the

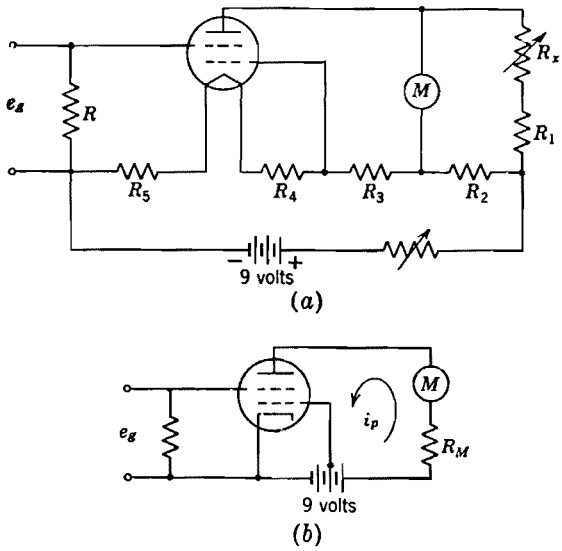



FIG. 8-14. (*a*) Electrometer tube circuit. (*b*) Equivalent for (*a*).

equivalent circuit of Fig. 8-14*b* is

$$i_p = \frac{\mu e_g}{r_p + R_M}$$

The circuit transconductance is then

$$g_c = \frac{i_p}{e_g} = \frac{\mu}{r_p + R_M} \tag{8-25}$$

and if $R_M$ is small, the circuit transconductance is just that of the grid-plate transconductance of the tube. Large values of this quantity, along with very minute grid currents, are tube requirements for this service.

**8-8. Residual Drift in Direct-coupled Amplifiers.** Direct-coupled amplifiers can be reasonably well stabilized against the types of output drift due to varying supply voltages, either by the methods of the preceding sections or by use of well-regulated power supplies. There remains, however, the effect of any internal tube changes of emission or cathode characteristics, contact potentials, or tube parameters, which may produce a shift in the $Q$ point of an amplifier, as well as a change in its voltage-current characteristics. Such changes frequently accompany the general aging of a tube and are very low-frequency phenomena, far below the low-frequency cutoff of a resistance-capacitance-coupled amplifier, and thus are not noted in the output of such a non-direct-coupled amplifier. However, a change in $Q$ point is immediately reflected as a change in the d-c output values of a direct-coupled amplifier having a pass band to zero frequency, and the general phenomenon is known as *residual drift.*

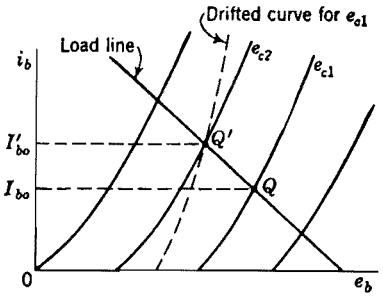



FIG. 8-15. Change in $Q$ point due to tube aging.

Consider Fig. 8-15, wherein the solid characteristics are those of the new tube, the dashed curve being one for bias $e_{c1}$ after an aging process has taken place. Initial operation, with a bias $e_{c1}$, was at point $Q$ on the solid curve. At a later time in the life of the tube this solid curve becomes altered to the dashed curve by reason of aging of some internal tube element or part. With bias still at $e_{c1}$ the operating point is now on the load line at $Q'$ with output current $I'_{bo}$.

It can be assumed, as far as static change in output current is concerned, that the tube did not change but instead the bias was merely changed from $e_{c1}$ to $e_{c2}$. Thus amplifier drift is equivalent to insertion of an extra voltage source in the grid circuit, and it is possible to express a drift figure for an amplifier in terms of *equivalent grid volts change per*

*hour of operation.* The drift figure is that grid voltage required to restore the output current to its initial value.

In feedback direct-coupled amplifiers this drift may be likened to noise or a perturbation produced within the feedback loop. The analysis of Sec. 7-4 may be used, wherein Eq. (7-9) showed

$$\mathbf{E}_o' = \frac{\mathbf{A}}{1 - \mathbf{A}\beta}\left(\mathbf{E}_s + \frac{N}{\mathbf{A}_1}\right)$$

where $\mathbf{A} = \mathbf{A}_1\mathbf{A}_2$, the gain before and after the point of noise introduction, respectively. Since the drift or noise equivalent voltage is assumed as introduced at the input, the gain $A_1$ of the previous analysis becomes unity and the output voltage with feedback is

$$e_o = \frac{A}{1 - A\beta}(e_s + e_d) \tag{8-26}$$

where $e_s$ is the input signal and $e_d$ is the equivalent drift voltage. Thus both these inputs are treated alike, and feedback does not discriminate in favor of the desired signal.

For some years output drift due to this residual cause was compensated by adjustment of manual controls on input bias. However, several more automatic systems are available and will be discussed in succeeding sections.

**8-9. Chopper-modulated Amplifiers.** Mechanical *choppers,* or synchronous switches, have been developed to interrupt or modulate a low-level direct current or very low-frequency input, giving an alternating voltage of amplitude proportional to the input direct current or slowly varying signal. The alternating voltage is then amplified by a conventional a-c signal amplifier and its output used either in a-c form or rectified back to d-c again. For amplification of slowly changing signals (thus requiring narrow bandwidth) such a system is satisfactory and will reduce drift problems to zero.

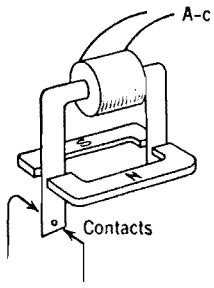


FIG. 8-16. Elements of a chopper.

A diagrammatic sketch of a chopper is shown in Fig. 8-16. A magnetic circuit is polarized by a permanent magnet and an a-c flux superimposed. This latter flux saturates the magnetic circuit, so that change in flux occurs only near the cycle zero points. The vibrating contact alternates between contacts on each half cycle of the a-c excitation, and owing to the saturation its switching operation occurs close to the cycle

zero. While 60 cycles is the usual excitation frequency, some choppers are made to operate at frequencies of 400 to 1000 cycles.

Figure 8-17*a* shows a series form of chopper amplifier which supplies an a-c input to an a-c amplifier by alternately connecting the amplifier

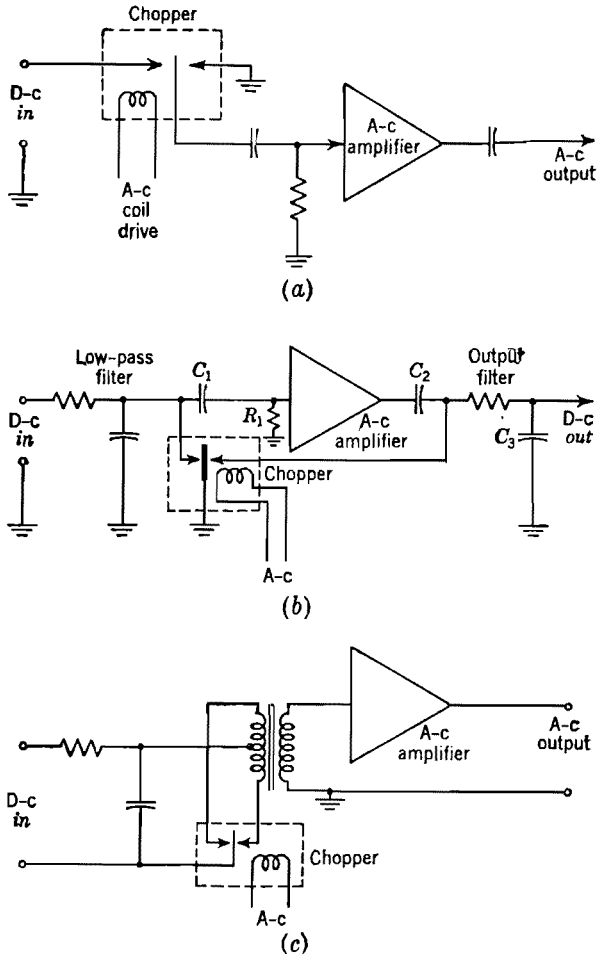



FIG. 8-17. (*a*) Simple chopper amplifier with a-c output. (*b*) Synchronous chopper with d-c output. (*c*) Chopper amplifier with input transformer and a-c output.

input to the d-c source and to ground. The amplifier output would be in a-c form unless rectified by a diode or other device.

The circuit of Fig. 8-17*b* removes the need for the additional rectifier diode by use of a shunt chopper which synchronously rectifies the amplifier output back to direct current. For analysis of operation assume that

the chopper reed has been on the right contact. Capacitor $C_1$ has then been charged through $R_1$ to the potential of the d-c source. When the reed returns to the left, capacitor $C_1$ is grounded and discharged, causing a current flow through $R_1$ and applying a voltage as input to the amplifier. At the same time capacitor $C_2$ is discharging from the amplifier output through the output filter.

When the reed returns to the right contact, capacitor $C_1$ charges from the d-c signal source, producing an input of opposite polarity on the amplifier. At the same time the output capacitor $C_2$ is grounded, so that no charge is contributed to $C_3$. Thus an a-c input is supplied to the amplifier, but pulses of unidirectional charge are supplied to the output capacitor $C_3$ on alternate half cycles, and the charge on $C_3$ does not reverse and the output appears as a rectified direct current.

If this output direct current is to be distortionless, or is to have a value proportional to the input voltage at all times, all higher-frequency components introduced at the switching frequency or its harmonics must be eliminated. This is the purpose of the output filter. To ensure proper filtering, its time constant must be long with respect to a period of the switch, and its high-frequency cutoff may need to be of the order of one-tenth to one-fifth of the switching rate. This imposes a fundamental limit on the rate of variation which may appear in the input to the amplifier, or on the allowable bandwidth which may be handled. Sixty cycles is a common frequency of chopper operation, and this basic filter requirement places the passband width in the range of 0 to 6 cycles. Thus in obtaining a solution to the problem of drift it has been necessary to sacrifice bandwidth.

In some applications this is not a serious limit, as in the amplification of d-c voltages from thermocouples measuring furnace temperatures, and many such chopper amplifiers are in use in industrial temperature recorders.

Another major difficulty in the operation of choppers at low d-c inputs is pickup from the a-c operating coil, this pickup voltage being amplified like a signal. If considerable gain is built into the amplifier, the hum voltages may overload the later stages and cause faulty operation, even though the output filter will remove the hum from the amplifier output. Narrow band response is an aid in eliminating this interference, but isolation of the a-c operating coil and leads from the low-level input circuit is very important.

Another variation of the input circuit appears in Fig. 8-17*c*, where the chopper operates in the primary of an input transformer and rectification is either not used or can be supplied by diodes. Input-transformer isolation and shielding are very important if hum and noise pickup are not to be serious with low-level inputs. Choice of circuits in Fig. 8-17*a* to *c*

is partly dependent on the source impedance, those in Fig. 8-17*a* and *b* being better adapted to high-impedance sources, whereas that in Fig. 8-17*c* can be used to match a low-impedance source to the high impedance of the amplifier input.

Where a wider band of frequencies must be amplified, it is possible to use a higher chopping frequency, but 400 to 1000 cycles represents about the limit of mechanical choppers. For ranges of zero to still higher frequencies, up to several hundred thousand cycles, a balanced bridge of diode rectifiers, called a *varistor*, may be used as in Fig. 8-18. In theory, when alternating current is applied through transformer $T$ to the bridge diodes, all four diodes conduct on one half cycle, thus shorting the grid to ground. All four diodes appear as an open circuit on the other half cycle, thus applying full signal input to the grid of the tube.

If the diode internal resistances went to zero when conducting, the triode would be driven by a voltage varying from the signal instantaneous value to zero. Since the diodes do not represent zero resistance, resistor $R$ is inserted in series with the d-c input and the ratio of internal bridge resistance to $R$ fixes the fraction of the input voltage which is transmitted during the conducting interval. The output has the same carrier or center frequency as that of the alternating current applied to the bridge through $T$, and because this may be several hundred thousand cycles, the input bandwidth handled may approximate zero to 10,000 cycles.

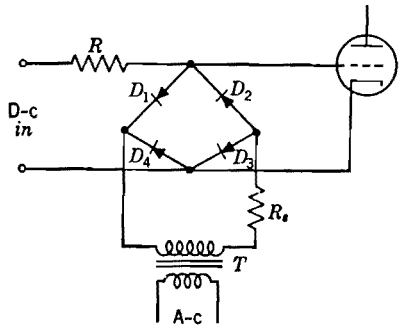


FIG. 8-18. Diode chopper or modulator.

Since the diodes may drift in internal-resistance value, this method is not drift-free and the drift problem is merely transferred from the amplifier input tube to the varistor chopping device.

**8-10. Chopper-stabilized Direct-coupled Amplifiers.** Possibilities of further improvement in drift compensation are suggested by two of the previous discussions. First, Eq. (7-9) shows that noise in a feedback amplifier is reduced in proportion to the gain preceding the point of noise introduction. Considering drift as a "noise" voltage, placement of driftless gain ahead of the point in the amplifier at which drift originates will reduce drift. Second, if it be remembered that drift is a *very* low-frequency phenomenon, then low-frequency gain may be added ahead of the first tube of a direct-coupled amplifier by use of a chopper modulator, with its accompanying narrow bandwidth, the high-frequency components being permitted to pass directly through $C_3$ into the direct-coupled

amplifier with its wider pass band as in the *chopper-stabilized amplifier* of Fig. 8-19.

As shown in this figure, any output voltage due to drift in the d-c amplifier is fed back through $R_2$ and being of very low frequency is

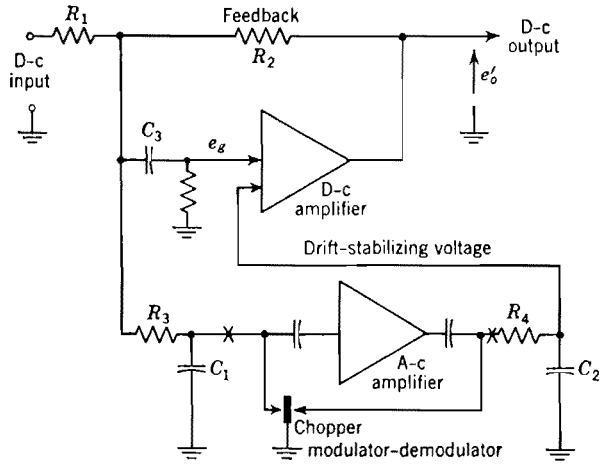



FIG. 8-19. Chopper-stabilized d-c amplifier.

effective in charging $C_1$ through the input filter network $R_3$, $C_1$, of time constant long with respect to the time of a cycle of the chopper frequency. Voltage appearing on $C_1$ is chopped, amplified, and synchronously rectified by the chopper. The output is filtered and the chopper frequency components removed by $R_4$, $C_2$, of long time constant, and applied to the input of the d-c amplifier as a *drift-stabilizing* voltage. This voltage cancels the equivalent grid drift voltage and annuls the drift output voltage to zero. A method of introducing the annulling voltage into the d-c amplifier input is shown in Fig. 8-20, wherein cathode coupling is employed in the first stage of the d-c amplifier to add the two inputs.

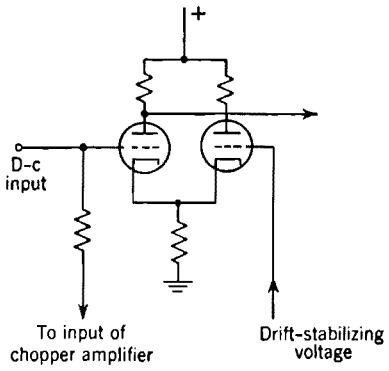



FIG. 8-20. Method of introducing low-frequency drift components into a d-c amplifier.

The capacitor $C_3$ is not always required, because even though the drift low-frequency voltage also appears in $e_g$ along with the higher-frequency components it is insignificant in its effect. If the gain of the a-c amplifier is $A_2$ and large, the drift-stabilizing voltage is $A_2$ times the original drift

voltage; thus the drift component directly introduced is negligible with respect to the drift voltage times $A_2$.

This may be further explained. If $e_g$ is the input representative of a certain amount of drift in the d-c amplifier, then the output voltage $e_o$, without the stabilizing amplifier present, is

$$e_o = -A_1 e_g$$

where $A_1$ is the gain of the d-c amplifier. However, with an input circuit as in Fig. 8-19 with an a-c amplifier gain of $A_2$, the output voltage with the a-c amplifier operating is

$$e_o' = -A_1 A_2 e_g' - A_1 e_g'$$

where $e_g'$ is the input with the stabilizing amplifier operating. To understand this statement, it must be remembered that the two gains $A_1$ and

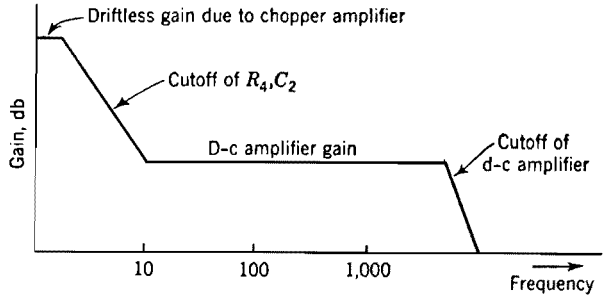



FIG. 8-21. Gain-frequency plot for a chopper-stabilized amplifier.

$A_2$ have different frequency-response characteristics due to chopper limitations, and that a low frequency, such as a drift signal, is amplified by the two amplifiers in series giving the $-A_1A_2$ gain, while a higher-frequency term is amplified by the d-c amplifier only, giving only a gain of $A_1$.

In effect this circuit places the a-c amplifier in series with the d-c amplifier for low frequencies. A gain-frequency plot for such an amplifier is presented in Fig. 8-21 and illustrates the extra gain provided for the low frequencies by $A_2$ of the a-c amplifier.

If the output drift is to be limited to the same amount with or without the stabilizing amplifier, then $e_o = e_o'$, and from the above equations

$$e_g' = \frac{e_g}{A_2 + 1} \tag{8-27}$$

and the input-voltage component to give the same drift is reduced by a factor of $A_2 + 1$ from its value without the chopper and a-c amplifier. That is, the equivalent drift voltage is reduced by this amount, thus indicating improved stability.

The long time constant associated with the output of the drift-stabilizing amplifier allows one such amplifier to be used to stabilize a number of d-c amplifiers. Motor-driven commutators are used to switch the a-c amplifier and chopper at points $x$, $x$ of Fig. 8-19 from one to another of several d-c amplifiers. Because of the long time constants the filters maintain their voltages over the switching cycle, and since the drift is such a slow-changing phenomenon, correction of the filter output voltage on $C_2$ is needed only at intervals. Thus cost of such stabilization for large groups of amplifiers is reduced, and amplifiers of this nature are now quite frequently employed.

**8-11. The Operational Amplifier.** Many d-c amplifiers are employed in performing mathematical operations such as summing, differentiation, and integration. When so applied, they are generally referred to as

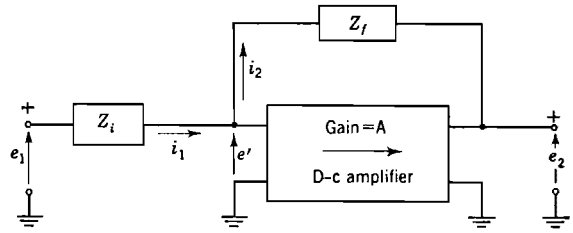



FIG. 8-22. Basic operational amplifier.

*operational amplifiers.* These operations may involve steady or slowly varying components so that d-c amplifiers are required. Almost any combination of the principles previously discussed, which leads to a linear, stable, direct-coupled amplifier having an odd number of stages or equivalent phase shift, combined with over-all feedback of the proper nature, may be employed.

Such a combination is represented by the amplifier symbol of Fig. 8-22, with gain

$$A = -\frac{e_2}{e'} \tag{8-28}$$

and having an inherent odd number of phase reversals.

The voltages and currents represented in Fig. 8-22 are instantaneous time functions and are thus related by the usual electrical laws,

$$e = iR \tag{8-29}$$

$$e = L\frac{di}{dt} = Lpi \tag{8-30}$$

$$e = \frac{1}{C}\int_{t_1}^{t_2} i\,dt = \frac{i}{Cp} \tag{8-31}$$

where the operator-$p$ notation is used once again, with $p = d/dt$. It can be readily shown that the notation $1/p$ implies the operation of integration in Eq. (8-31). Voltage-current ratios for the three basic circuit elements then become

$$\frac{e}{i} = R \qquad \frac{e}{i} = Lp \qquad \frac{e}{i} = \frac{1}{Cp}$$

depending on the type of circuit element employed. In general it is possible to refer to any of these relations as

$$\frac{e}{i} = Z(p) \tag{8-32}$$

where $Z(p)$ is called an *operational impedance*.

Analysis of performance of the circuit of Fig. 8-22 can be most general if initially the type of circuit elements contemplated for $Z_i$ and $Z_f$ not be specified, but these elements simply referred to as operational impedances $Z_i(p)$ and $Z_f(p)$ having some $e/i$ relation. After the performance analysis is completed, it will be possible to substitute the $e/i$ relation for any circuit elements or combination actually used and obtain a prediction of circuit performance.

The input voltage $e_i$ is applied through $Z_i(p)$, with over-all feedback introduced through the $Z_f(p)$ path. If current into the first grid circuit of the amplifier is negligible, then

$$i_1 = i_2$$

and, by use of Eq. (8-32), the current relations can be written as

$$\frac{e_1 - e'}{Z_i(p)} = \frac{e' - e_2}{Z_f(p)} \tag{8-33}$$

In view of the gain definition of Eq. (8-28), the input $e'$ can be eliminated and

$$-e_2\left[Z_i(p) + \frac{Z_i(p)}{A} + \frac{Z_f(p)}{A}\right] = Z_f(p)\, e_1$$

from which

$$-e_2 = \frac{Z_f(p)}{Z_i(p)} \frac{1}{1 + (1/A)[1 + Z_f(p)/Z_i(p)]} e_1 \tag{8-34}$$

If the amplifier gain $A$ is made large so that

$$A \gg 1 + \frac{Z_f(p)}{Z_i(p)}$$

then Eq. (8-34) reduces to

$$e_2 = -\frac{Z_f(p)}{Z_i(p)} e_1 \tag{8-35}$$

This is the basic equation for the operational form of amplifier. The accuracy with which the operations indicated by $Z_f(p)/Z_i(p)$ are performed is dependent on the requirement of very large amplifier gain being fulfilled. This is usually not a difficult specification to meet.

By definition and assumption the polarity of $e_2$ is always opposite to that of $e_1$, so that if $Z_f(p) = Z_i(p) = R$ (possibly 1 megohm each) then the circuit will produce an output of magnitude equal to its input, but of opposite sign, or $A = -1$. The circuit is thus always a sign changer. If $Z_f(p)$ and $Z_i(p)$ are resistive but not equal, then

$$e_2 = -\frac{R_f}{R_i} e_1 \tag{8-36}$$

and it becomes possible to obtain an output equal to the negative of the input multiplied by a constant, independent of amplifier characteristics. This usage is represented in Fig. 8-23*a*.

If $Z_f(p)$ is a capacitor, having an operational impedance $1/Cp$, and $Z_i(p)$ is a resistance, then Eq. (8-35) becomes

$$e_2 = -\frac{1}{RC}\frac{1}{p} e_1 \tag{8-37}$$

The amplifier is then an accurate integrator with respect to time, and a reversal of sign is still present. The result is multiplied by a constant $1/RC$, but if $R = 1$ megohm, $C = 1$ $\mu$f, the multiplying factor can be made unity. This application is illustrated in Fig. 8-23*b*.

The input voltage $e'$ is only $1/A$ of the output voltage $e_2$, and with $A$ very large, as required, the voltage $e'$ is negligible, and the potential at point $X$ at the amplifier input may be considered essentially zero or at ground reference. Therefore the potential across the integrating capacitor $C$ is almost equal to the output $e_2$. This is important in introducing the initial conditions into the integration process, since an initial charge on $C$ applied between output terminal and ground will represent an initial condition.

As another alternative let $Z_i(p)$ be a capacitor and $Z_f(p)$ a resistor as in Fig. 8-23*c*. Then Eq. (8-35) becomes

$$e_2 = -RCpe_1 \tag{8-38}$$

and the output of the amplifier is the negative of the derivative of the input voltage.

Another operation, that of summing several independent voltages, is introduced in Fig. 8-23*d*. As before

$$i_a + i_b + i_c = i_2$$

and if $A$ is large, making $e'$ very small with respect to either input or

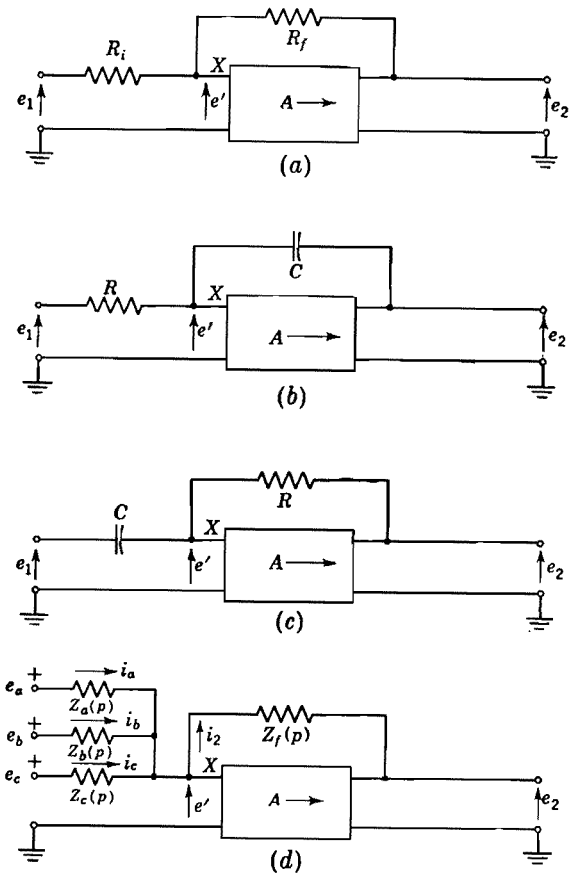



FIG. 8-23. Uses of the operational amplifier. (*a*) Sign changing and multiplication by a constant; (*b*) integration; (*c*) differentiation; (*d*) summation.

output voltages, then

$$\frac{e_a}{Z_a(p)} + \frac{e_b}{Z_b(p)} + \frac{e_c}{Z_c(p)} = \frac{-e_2}{Z_f(p)}$$

from which

$$-e_2 = \frac{Z_f(p)}{Z_a(p)}\, e_a + \frac{Z_f(p)}{Z_b(p)}\, e_b + \frac{Z_f(p)}{Z_c(p)}\, e_c \tag{8-39}$$

If all the impedances are equal resistors, then

$$-e_2 = e_a + e_b + e_c \tag{8-40}$$

and the output represents the summation of the inputs. This may be done for any reasonable number of input voltages.

If some negative voltages are introduced (possibly having passed through a sign-changing amplifier), then algebraic summation is possible.

In addition, if the ratios $R_f/R_a$, $R_f/R_b$, $R_f/R_c$ are not chosen equal, then multiplication by various constant factors may be done. By making $Z_f(p)$ a capacitor three integrals may be simultaneously added, or by making $Z_a(p)$ a capacitor several input variables may be added to the derivative of another variable.

If the input voltage to the integrator circuit were sinusoidal,

$$e_1 = E_m \sin \omega t$$

the output voltage would be

$$e_2 = -\frac{1}{RC}\int E_m \sin \omega t \, dt = \frac{E_m}{\omega RC} \cos \omega t$$

and for effective values

$$\mathbf{E}_2 = \frac{-\mathbf{E}_1}{j\omega RC} \tag{8-41}$$

which is most easily obtained by direct substitution of $p = j\omega$ in Eq. (8-37). It is apparent that the output of the integrating circuit falls off as $1/\omega$, or at the rate of 6 db per octave, and this offers a simple test of integrator accuracy. If at some frequency this relation fails to hold, then it is known that integration is accurate only up to this frequency.

In the case of the differentiator, if $e_1 = E_m \sin \omega t$, the output voltage would be

$$e_2 = -RC\frac{d}{dt}(E_m \sin \omega t) = -\omega RCE_m \cos \omega t$$

and for effective values

$$\mathbf{E}_2 = -j\omega RCE_1$$

which is Eq. (8-38) with $p = j\omega$. The output of a differentiator increases with $\omega$, at a rate of 6 db per octave. If $RC$ were equal to unity and $\omega = 377$ for 60 cycles, the output voltage $\mathbf{E}_2$ would be 377 times the input voltage $\mathbf{E}_1$. If there is undesired 60-cycle pickup or hum introduced into the input, the hum would appear in the output with a magnitude $\omega$ times greater. Because of the increasing gain with frequency, and the problems of noise and hum which then arise, the use of the operational amplifier as a differentiator is usually avoided. Mathematically this is not difficult to do.

It is possible to use parallel or series combinations of $R$ and $C$ as $Z_f$ or $Z_i$, and performance will be predicted by the form of the $Z_f(p)/Z_i(p)$ relation. The operational impedance of any $RC$ combination may be easily written if it be remembered that the $Z(p)$ form is obtainable from the steady-state impedance $Z$ by replacing $j\omega$ by $p$ wherever it appears. A few such circuit forms appear in Fig. 8-24 as examples.

The operational amplifier is basic to a great variety of electronic computation equipment. In application it is necessary that resistors be accurate, and capacitors of very low leakage and low retentivity of charge must be used. These usually employ polystyrene as a dielectric.

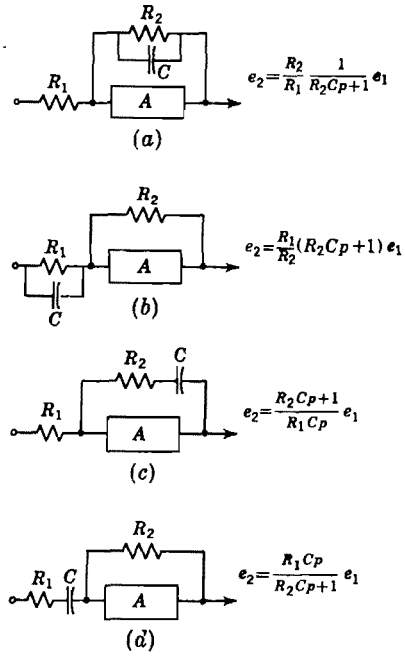



FIG. 8-24. A few operational amplifier circuits.

**Example.** Write the expression for the operation performed on the input voltage of the circuit in Fig. 8-24*c*. First write

$$Z_f = R_2 + \frac{1}{j\omega C}$$
$$Z_i = R_1$$

Then, replacing $j\omega$ by $p$, the operational impedances are

$$Z_f(p) = R_2 + \frac{1}{pC}$$
$$Z_i(p) = R_1$$

and the relation between $e_2$ and $e_1$ is

$$e_2 = \frac{R_2 + 1/pC}{R_1} e_1 = \frac{R_2Cp + 1}{R_1Cp} e_1$$

**8-12. Use of Positive Feedback.** The accuracy of the desired output-input relation in an operational amplifier is dependent on the assumption of very high gain inside the feedback loop. It has been found that such high gain may be easily obtained by use of positive feedback or regeneration in the amplifier, inside the overriding external negative feedback. The positive feedback may be very large; yet the overall negative feedback will maintain the stability of the amplifier.

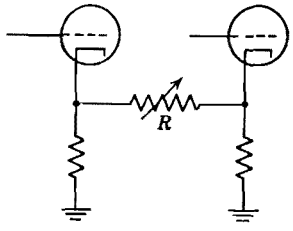



FIG. 8-25. Positive feedback inside an operational amplifier.

A circuit utilizing positive feedback between the cathodes of two successive stages, as in Fig. 8-25, is a very simple means of introducing the desired regeneration. Internal gains in the range of $10^6$ to infinity may be obtained in simple amplifier circuits and the accuracy of the functional operation made very great.

**8-13. Analog Computation.** The *analog* type of electronic computer uses electronic amplifiers and provides solutions to mathematical prob-

lems in terms of a physical quantity, such as the angular position of a shaft, a voltage or current, or the position of the index on a scale. As a result the accuracy may be limited by the precision with which these physical quantities can be read, or by the ability of a human read-out operator, if one is employed.

In contrast, the digital computer to be discussed in Chap. 9 provides solutions directly in numerical digits, as in the case of a desk calculator. The degree of accuracy is limited only by the capacity of the machine in terms of significant digits, and this is fixed largely by the cost and complexity to be tolerated. A digital computer is usually more expensive than the analog form, but when the accuracy requirements of the latter are pushed below about 0.1 per cent, the analog costs may exceed those of the digital form. Combinations of the two forms exist, in which digital output is provided for an analog computer and analog indication or curve plotting is also available for digital computers.

The applications of the analog computer lie largely in the field of engineering design or in the solution of systems of linear or nonlinear differential equations. Solutions are usually in the form of traced curves or oscillographic records of varying output voltages and currents. The basic operations in the application of such computers, or differential analyzers as they have been called, is the writing down of the differential equation or equations describing the particular physical system of interest. Of great aid in this respect is *Lagrange's equation* relating the interchange of system energies, and this will be considered in the next section.

Given the differential equations, the analog result may be obtained either by the *direct* method, in which equivalent electrical circuits are directly used to simulate the time variations of the dependent variables of the physical system, or the *functional* method, in which electronic circuits perform the mathematical operations indicated by the terms of the differential equation. The former method is well suited to design problems, especially where many degrees of freedom are involved, and the latter is best suited to a wide variety of engineering problems arising in control design and system operation.

**8-14. Lagrange's Equation for Dynamical Systems.** The starting points in most analog computation are the differential equations which describe the action of a dynamical system. In many complex physical systems, however, it is not easy to write down these differential equations. Often in engineering problems there are so many variables present that it is difficult to write the equations of the system, and these difficulties are increased by the frequent presence of both mechanical and electrical (or other) effects in the system.

Lagrange's equation provides a systematic means whereby the process is simplified. Its advantage lies in the usual ease with which expressions

for the kinetic and potential energies of a complex system may be written, the equations of motion following readily by application of Lagrange's equation. This equation, which will not be derived here, is based on the physical fact that nature, in moving from one condition of equilibrium to another, seems to do so in the manner requiring the least interchange, or transfer, of energy.

The equations of motion must employ some frame of reference or a coordinate system, and in a complicated problem there may be a number of these coordinates or independent variables, which will be referred to in general as $q_1, q_2, \ldots$, although they may bear other symbolism in actual problems. For instance, in a three-mesh electric network these $q$'s may actually appear as the mesh currents, $i_1, i_2, i_3, \ldots$, or may represent some set of node voltages. The first and second derivatives with respect to time are designated $\dot{q}$ and $\ddot{q}$ in the usual manner. The number of such independent coordinates is fixed by $n$, the number of degrees of freedom of the system, and this is equal to the number of independent equations required to describe the action of the system.

If $T$ be allowed to represent the total kinetic energy of the system, $V$ the total potential energy in terms of the selected coordinates, $D$ the dissipation of energy per second, and $f$ the applied force if present, then Lagrange's equation states the relation between these quantities as

$$\frac{d}{dt}\left[\frac{\partial(T - V)}{\partial \dot{q}_i}\right] - \frac{\partial(T - V)}{\partial q_i} + \frac{1}{2}\frac{\partial D}{\partial \dot{q}_i} = f \tag{8-42}$$

The term $q_i$ represents any of the coordinates $q_1, \ldots, q_n$, and thus there will be as many equations of the form of the above as there are independent quantities of reference.

As a simple illustration of the method consider a projectile of mass $m$ whose motion is defined by two coordinates, $x$, the distance from the point of reference, and $y$, the height above earth. The energies may be written as

$$T = \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2) \qquad V = mgy \tag{8-43}$$

Preparing the terms needed for substitution in Eq. (8-42),

$$\frac{\partial T}{\partial \dot{x}} = m\dot{x} \qquad \frac{\partial(T - V)}{\partial x} = 0$$

$$\frac{\partial T}{\partial \dot{y}} = m\dot{y} \qquad \frac{\partial(T - V)}{\partial y} = -mg$$

The dissipation $D$ and applied force $f$ are zero. Writing Lagrange's equation for each of the coordinates $x$ and $y$ gives

$$\frac{d}{dt}(m\dot{x}) - 0 = 0 \qquad \frac{d}{dt}(m\dot{y}) + mg = 0$$

from which $m\ddot{x} = 0 \quad m\ddot{y} = -mg$

$\ddot{x} = 0 \quad \ddot{y} = -g$

which are the usual results.

Lagrange's equation shows a greater advantage in more complicated problems, such as the condenser microphone of Fig. 8-26. Such a device consists of a stretched metallic diaphragm closely spaced in front of a solid metal backplate. The sound pressure applies a force on the diaphragm, varying the spacing and the capacity between diaphragm and

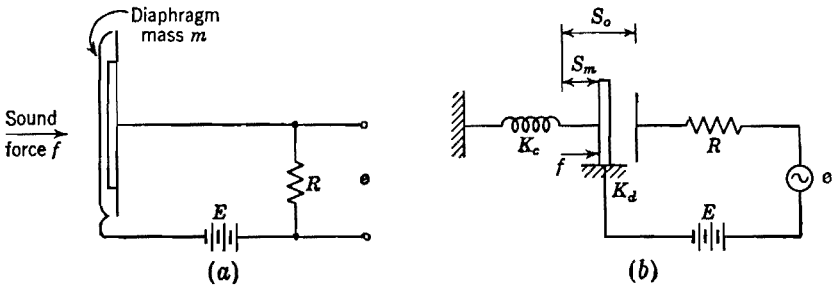



FIG. 8-26. Condenser microphone.

plate. Since the applied voltage $E$ is constant, the varying capacity requires a varying charge in accordance with

$$q = CE$$

and a current therefore flows through $R$, producing a varying output voltage $e$.

The coordinates of reference are taken as the diaphragm deflection $s$ and the charge $q$ on the capacitor. The kinetic energy of the system is then

$$T = \frac{m\dot{s}^2}{2} \tag{8-44}$$

to which would be added a term $L\dot{q}^2/2$ if inductance were present in the electric circuit. The potential energy is

$$V = \frac{k_C s^2}{2} + \frac{q^2}{2C} \tag{8-45}$$

The instantaneous value of capacity $C$ is

$$C = \frac{s_o}{s_o - s} C'$$

where $C'$ is the capacity when the movable diaphragm is at $s = 0$. Then

$$V = \frac{k_C s^2}{2} + \frac{q^2}{2C'} \frac{s_o - s}{s_o} \tag{8-46}$$

The dissipation is due to the resistance $R$ and to friction $k_d$ in bending the metal of the diaphragm. Then $D$ is

$$D = k_d\dot{s}^2 + R\dot{i}^2 = k_d\dot{s}^2 + R\dot{q}^2 \tag{8-47}$$

Taking the indicated derivatives for Lagrange's equations,

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{s}}\right) = m\ddot{s} \qquad \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}}\right) = 0$$

$$\frac{\partial(T - V)}{\partial s} = -k_C s + \frac{q^2}{2C'}\frac{s_o - s}{s_o} \qquad \frac{\partial(T - V)}{\partial q} = -\frac{q}{C'}\frac{s_o - s}{s_o}$$

$$\frac{\partial D}{\partial \dot{s}} = 2k_d\dot{s} \qquad \frac{\partial D}{\partial \dot{q}} = 2R\dot{q}$$

and combining by Eq. (8-42) leads to

$$m\ddot{s} + k_C s - \frac{q^2}{2C'}\frac{s_o - s}{s_o} + k_d\dot{s} = f \tag{8-48}$$

$$\frac{q}{C'}\frac{s_o - s}{s_o} + R\dot{q} = e \tag{8-49}$$

as the differential equations describing the action of the condenser microphone. It is seen that one equation is obtained for each of the coordinates of reference.

Lagrange's equation provides a simple tool by which system differential equations can be written, preparatory to analog solution.

**8-15. The Direct-analog Computer.** The so-called *direct analog* of a problem is arrived at by replacing the physical system under study, as described by a set of differential equations, with an electrical circuit whose performance is described by differential equations of *identical mathematical form.* It is well known that two sets of equations of the same form have identical forms of solutions, and therefore the performance of the electrical circuit is analogous to the operation of the physical system replaced. The solution, as obtained in the electrical circuit in terms of easily measured currents and voltages, is analogous to the original physical system in terms of its appropriate independent and dependent variables.

For instance, the common mass-spring-damper mechanical system of Fig. 8-27*a* may be expressed by a differential equation of the form

$$m\frac{d^2x}{dt^2} + k_d\frac{dx}{dt} + k_C x = F(t)$$

which may be otherwise written, by use of $v = dx/dt$, as

$$m\frac{dv}{dt} + k_d v + k_C\int v\,dt = F(t) \tag{8-50}$$

The system of rotation of Fig. 8-27*b* yields a similar equation

$$I\frac{d^2\theta}{dt^2} + k_d\frac{d\theta}{dt} + k_C\theta = T(t)$$

and by using $\omega = d\theta/dt$,

$$I\frac{d\omega}{dt} + k_d\omega + k_C\int \omega\, dt = T(t) \tag{8-51}$$

Both Eqs. (8-50) and (8-51) have been written under an assumption of viscous friction in the damping device, or that friction force is proportional to the first power of the velocity. This is satisfactory for a great

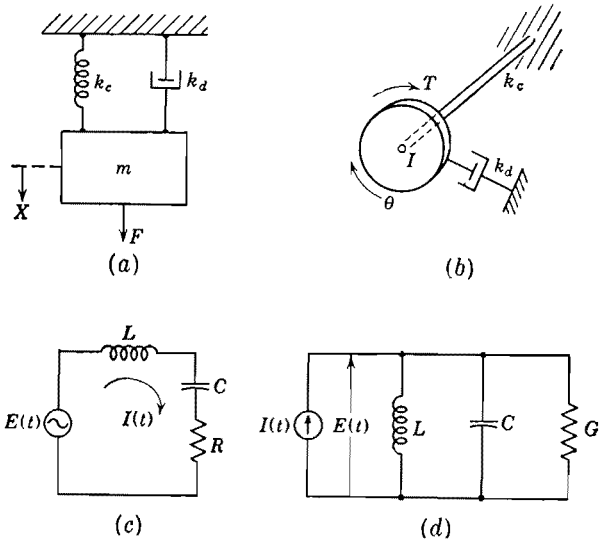



FIG. 8-27. Analogous systems.

variety of problems, and methods exist whereby this limitation on performance may be removed if necessary.

A mere change in the choice of alphabetical symbolism may lead to

$$L\frac{di}{dt} + Ri + \frac{1}{C}\int i\, dt = E(t) \tag{8-52}$$

or to

$$C\frac{de}{dt} + Ge + \frac{1}{L}\int e\, dt = I(t) \tag{8-53}$$

It is apparent that Eqs. (8-52) and (8-53) are of identical form to either Eq. (8-50) or Eq. (8-51), and to one versed in electrical terminology it is likewise apparent that Eqs. (8-52) and (8-53) describe the performance of the circuits of Fig. 8-27*c* and *d*, respectively. Thus these circuits

become the *direct analogs* of the translational or rotational systems of Fig. 8-27*a* and *b*.

The systems of Fig. 8-27*a* and *b* may be studied by use of the circuits of Fig. 8-27*c* or *d*, the choice being arbitrary and determined largely by the availability of suitably sized electric-circuit elements. In Fig. 8-27*c* the independent variable is $E(t)$, the dependent variable $I(t)$, analogous to the force $F(t)$ or torque $T(t)$ and to velocities $v$ or $\omega$, respectively. Inductance is analogous to mass, and this analog is sometimes called the *force-voltage*, or the *mass-inductance*, analog. In Fig. 8-27*d* the independent variable is $I(t)$ and the dependent variable $E(t)$, analogous to $F(t)$ or $T(t)$ and to $v$ or $\omega$, respectively. Capacitance is analogous to mass, and this form is sometimes referred to as the *mass-capacitance*, or the *force-current*, analog.

Study of $I(t)$ in Fig. 8-27*c* or $E(t)$ in Fig. 8-27*d* by oscillographic or other methods, when the circuits are excited by a voltage or current time function of form analogous to the mechanical force function, yields results which may be interpreted in terms of velocities, accelerations, and displacements in the analogous mechanical, thermal, or other physical system under study.

By use of electronic circuits it is relatively easy to produce sinusoids, pulses, saw-toothed waves, decaying exponentials, and the like, as driving sources or time functions analogous to the mechanical or other force functions which may occur in an actual system. In this connection, constant-current sources may be simulated by voltage sources of very high internal resistance or by pentodes operated in a constant-anode-current region.

TABLE 13. SUMMARY OF ANALOGOUS QUANTITIES

| Translational | Rotational | Electrical | |
|---|---|---|---|
| | | $M-I$ | $M-C$ |
| Force $f$ | Torque $T$ | $E$ | $I$ |
| Mass $m$ | Inertia $I$ | $L$ | $C$ |
| Damping $k_d$ | Damping $k_d$ | $R$ | $G$ |
| Velocity $v$ | Velocity $\omega$ | $I$ | $E$ |
| Compliance $k_C$ | Compliance $k_C$ | $1/C$ | $1/L$ |
| Displacement $x$ | Displacement $\theta$ | $q$ | $e\,dt$ |

To illustrate further the direct-analogy viewpoint, Table 13 develops the analogous translational, rotational, and electric-circuit fundamental quantities. Similar analogs in other energy systems can also be set up if desired.

Table 14 also shows how the mathematical analogs may be set up to convert system elements to circuit elements.

One difficulty is occasionally encountered in these electrical analog circuits. Some resistance is always associated with inductors and friction with any mechanical system. It should be noted, however, that the ratios of mass or inertia to unavoidable friction may be very much greater in mechanical devices than the achievable ratios of $L/R$ in electrical inductors. This leads to a difficulty in simulation and, where the situation becomes limiting in application of the method, may make advisable the employment of the mass-capacitance analog since capacitors have much lower internal losses than inductors.

TABLE 14. ANALOGOUS RELATIONS

| | | |
|---|---|---|
| Mass to inductance | $f = m\frac{dv}{dt}$ | $e = L\frac{di}{dt}$ |
| | $v = \frac{1}{m}\int f\,dt$ | $i = \frac{1}{L}\int e\,dt$ |
| | $W = \frac{mv^2}{2}$ | $W = \frac{Li^2}{2}$ |
| Spring to capacitance | $f = k_C\int v\,dt$ | $e = \frac{1}{C}\int i\,dt$ |
| | $v = \frac{1}{k_C}\frac{df}{dt}$ | $i = C\frac{de}{dt}$ |
| | $W = \frac{fx}{2} = \frac{f^2}{2k_C}$ | $W = \frac{eq}{2} = \frac{Ce^2}{2}$ |
| Damping to resistance | $f = k_d v$ | $e = Ri$ |
| | $v = \frac{f}{k_d}$ | $i = \frac{e}{R}$ |
| | $W = \int fv\,dt$ | $W = \int ei\,dt$ |
| | $P = fv$ | $P = ei$ |

The direct analog has its greatest application in design studies, wherein mass, elastance, or damping parameters may be represented directly by $L$, $C$, or $R$ components and the effects of a wide range of parameter values studied in a short time. The similarity in structure of the physical system and its electrical analog is often an aid in the use of the method, and to an experienced worker in the field will usually eliminate the necessity for writing the system equations for simple systems.

**8-16. Scale Factors.** Any system of units which is consistent within itself may be employed in the opposing analogous forms, that is, in a mechanical system pound-foot-second or meter-kilogram-second (mks) units may be employed, while in the electrical analog the mks units will probably be chosen because of general agreement on their use within this field. In actual application, then, 1 g mass may be analogous to 1 h

or 1 lb mass might be chosen as analogous to one mh if all other elements are appropriately selected.

Scale factors relating or modeling the relative magnitudes of the electrical-circuit elements must be properly selected. In mechanical systems there are usually three dimensions, mass $M$, length $L$, and time $T$, which are arbitrarily chosen as basic, or fundamental. Thus three scaling factors may be chosen as a matter of convenience, all others being derived therefrom.

As an example, the mass-inductance analogy may be scaled by setting up three scale factors $S_1$, $S_2$, and $S_3$, to yield satisfactory and reasonable electric element sizes. Then*

$$\mathfrak{L} \doteq S_1 M \qquad Q \doteq S_2 L \qquad t \doteq S_3 T$$

where $\mathfrak{L}$, $Q$, and $t$ are the inductance, charge, and time referring to the electric analog and $M$, $L$, and $T$ are those of the mechanical system. Employing the usual dimensional relations, the other analog scaling factors may be derived.

For instance, it will not be convenient to choose $S_2$ directly in many cases because we ordinarily do not work directly with charge. However, if current be taken as dimensionally $Q/t$ and its analog velocity as $L/T$, then

$$1 \text{ unit of current} = \frac{Q}{t} \doteq \frac{S_2 L}{S_3 T} = \frac{S_2}{S_3} \times 1 \text{ unit of velocity}$$

The factor $S_3$ may be arbitrarily chosen to provide a convenient time relationship and the above used to determine $S_2$ to give a desirable value of current in the analog.

The scale factor for voltage may be derived by use of the relation $e = \mathfrak{L}\, di/dt$, giving

$$1 \text{ unit of voltage} = \frac{Q}{t^2} \doteq \frac{S_1 M S_2 L}{(S_3 T)^2} = \frac{S_1 S_2}{S_3^2} \times 1 \text{ unit of force}$$

since force has dimensions $MLT^{-2}$. Other common relations follow as in Table 15.

TABLE 15. DERIVED SCALE FACTORS FOR MECHANICAL SYSTEMS

| | |
|---|---|
| Current | $\doteq S_2 S_3^{-1} \times$ velocity |
| Voltage | $\doteq S_1 S_2 S_3^{-1} \times$ force |
| Resistance | $\doteq S_1 S_3^{-1} \times$ damping (force/velocity) |
| Capacity | $\doteq S_1^{-1} S_3^2 \times$ compliance (displacement/force) |
| Power | $\doteq S_1 S_2^2 S_3^{-3} \times$ mechanical power |

* The symbol $\mathfrak{L}$ will be employed to refer to inductance here to avoid conflict with $L$, denoting length. Also the symbol $\doteq$ will be used to denote "analogous to."

**8-17. Functional or Operational Analog Computers.** The operational analog (differential analyzer) employs electrical circuits to perform the mathematical operations called for by the integrodifferential equations of the dynamical system. Operational amplifiers of the form described

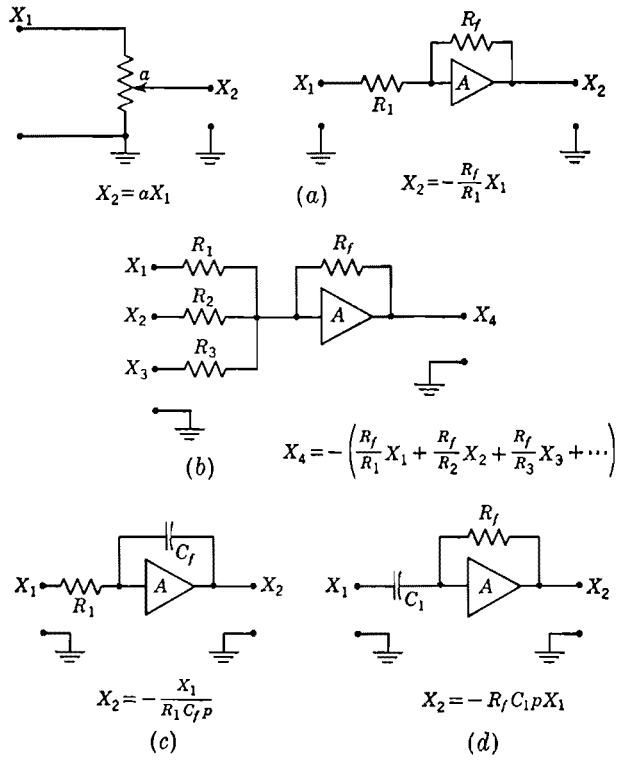



FIG. 8-28. Methods of using operational amplifiers in computation.

previously in this chapter are employed in the manner indicated in Fig. 8-28, to:

1. Form the sum of two or more variables.
2. Multiply a variable by a constant.
3. Integrate a variable against time.
4. Take the time derivative of a variable.

Units performing these operations are connected by cords and plugs in the sequence required by the particular problem. In the *slow* form of computer the process of obtaining the solution may take several minutes to accommodate servomechanism operation, and the solution appears as a plotted curve on a graph sheet on a plotting table. In the *fast*, or repetitive, computer the time scale is speeded, and the input

function and initial conditions applied 30 to 60 times per second, with the output appearing as a trace on a cathode-ray oscillograph.

These two types differ largely in accuracy, since a cathode-ray display is only 5 to 10 per cent accurate unless calibrated at the time of use. It is customary to draw the coordinate axes for the slow-computer plot directly with computer output, thus eliminating calibration errors that may exist.

The operation of differentiation is usually avoided by prior manipulation of the equations, since any hum or other noise present is increased in importance by a factor of angular frequency. Integration does not so suffer, noise being in effect reduced by $\omega$.

To illustrate only the principle of operation, consider a typical second-order differential equation as

$$A\ddot{x} + B\dot{x} + Cx + D = f(t) \tag{8-54}$$

Division by $A$ and equating the highest-order term to the remainder leads to

$$\ddot{x} = -\frac{B}{A}\dot{x} - \frac{C}{A}x - \frac{D}{A} + \frac{1}{A}f(t) \tag{8-55}$$

This equation indicates that the second derivative of $x$ is equal to the sum of the other $x$ functions plus $f(t)/A$. Assume a summing amplifier be arranged with four inputs as in Fig. 8-29. The output of this amplifier must be equal to $\ddot{x}$ by Eq. (8-55) (actually $-\ddot{x}$ because of the sign-changing properties of the operational amplifier). A complete solution for $x$ is then possible by two successive integrations on $\ddot{x}$ At the same time the other input terms proportional to $\dot{x}$, and $x$, may be formed by additional amplifiers as in Fig. 8-30. A solution for $x$ is then obtainable and for a repetitive computer will appear as a trace on a cathode-ray screen. For a slow computer the output voltage to ground available at the $-x$ terminal may be used to actuate a recorder to give a record as a function of time.

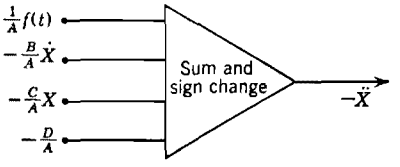



FIG. 8-29. Obtaining $\ddot{x}$ in the analog computer.

Switches $S_1$ and $S_2$ may be relay-operated at the instant the computer time scale equals zero, will place desired magnitudes of initial voltage on the integrating capacitors to represent initial conditions, and will also trigger the generator for $f(t)$. For a repetitive, or fast, computer these relays must be operated each time that $t = 0$ in the $f(t)$ driving function. Such a time scale requires that the complete transient solution of the problem be scaled in time so that the transient will die out in somewhat

less time than that of one period. Long-persistence cathode-ray screens are often helpful.

The preparation of the computer may be systematized as follows:

1. Write the differential equations of the system.
2. Solve for the highest-order derivative in each equation.
3. Form the block diagram for each equation.

*a.* Derive lower-order terms by integration.
*b.* Generate the coefficients of the lower-order terms as required, by use of available apparatus.
*c.* Using summing amplifiers, add the lower-order terms after multiplication by their coefficients.

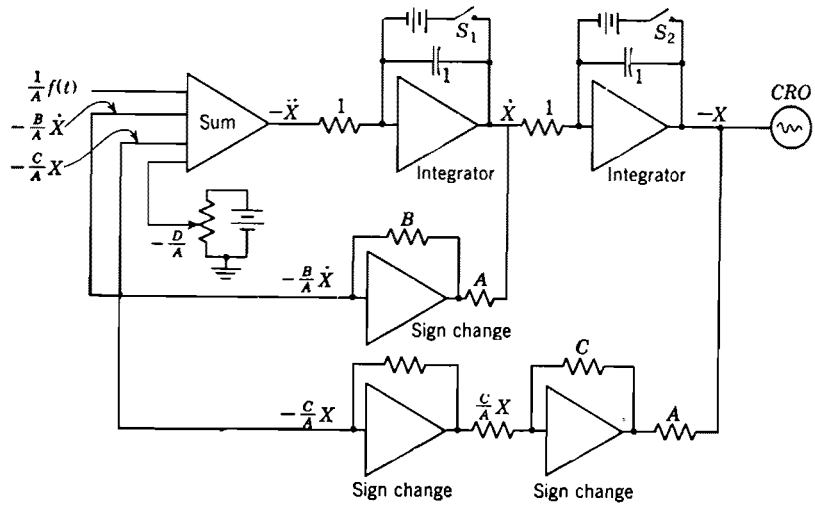



FIG. 8-30. Solution of Eq. (8-54).

**Example.** A body of mass $m$ falls under gravity through a viscous liquid. Design the computer circuit to plot displacement $y$, and velocity $\dot{y}$ against time, with initial displacement = 50 ft upward and initial velocity = 95 fps upward.

In preparation for the use of Lagrange's equation the energies may be written as

$$T = m\dot{y}^2/2 \qquad V = mgy \qquad D = k_d\dot{y}^2$$

Application of Lagrange's equation then yields the equation of motion as

$$m\ddot{y} - mg + k_d\dot{y} = 0$$

Rearranging,

$$\ddot{y} = g - \frac{k_d}{m}\dot{y} \tag{8-56}$$

The initial conditions call for voltages to be applied to the capacitors of the integrating amplifiers at $t = 0$, by use of the relays. These voltages are adjusted manually by potentiometers to bring the integrator outputs to the desired initial value with all

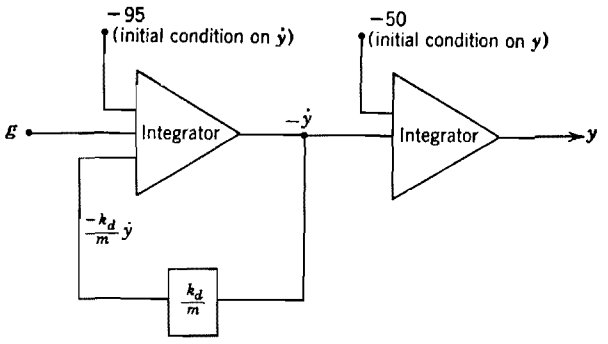



FIG. 8-31. Computer set up to solve Eq. (8-56).

other inputs removed. The circuit of Fig. 8-31 will then perform the desired mathematical operation.

**8-18. Function Generators.** Since analog computers are called upon to simulate and solve a wide variety of dynamic problems, it is necessary to be able to generate many driving-voltage waveforms, either as time functions or as functions of other variables.

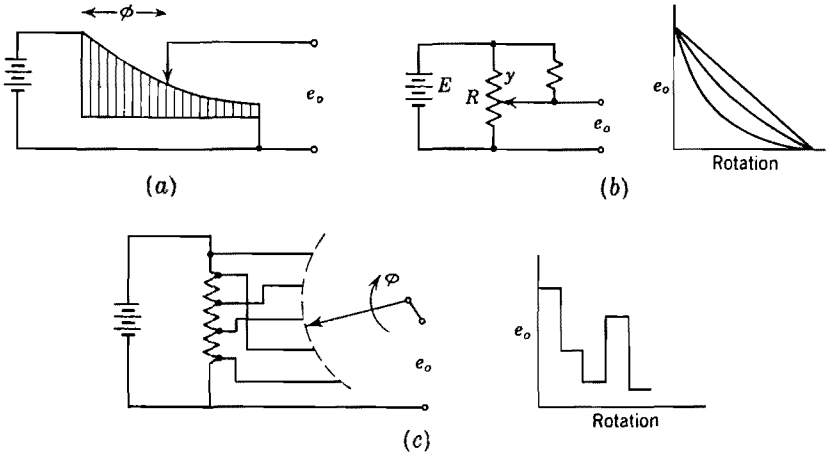



FIG. 8-32. Function generation by potentiometers.

Tapered or nonlinear wire-wound circular potentiometers, as in Fig. 8-32, may be driven by a constant-speed motor if a time function $e_o = f(t)$ is desired, or by a servomotor controlled by another variable $y$ if $e_o = f(y)$ is the desired output. The method is limited to monotonic functions (continuously increasing or decreasing) for which a suitable winding can be manufactured on a tapered card, of which the sine and cosine are examples.

The effect of a current load on the output terminals of a tapered

potentiometer must be considered. If the loading effect is appreciable, the output may depart from the desired function of shaft position. In fact, a linear potentiometer may be deliberately loaded,a s in Fig. 8-32*b*, to achieve quite a variety of nonlinear functions relating angle of rotation $\phi$ and output $e_o$. For instance, assuming zero external loading on $e_o$, the output as a function of angle $\phi$ for Fig. 8-32*b* is

$$e_o = E \frac{1 - \phi}{1 - \phi^2 R/(\phi R + y)}$$

where $\phi$ is measured from the $e_o = E$ position. For linear unloaded output $e_o = E(1 - \phi)$.

Photoelectric curve followers are available which will rotate a potentiometer, or provide an output voltage, proportional to the ordinate of a plotted curve. The output may also be a time function, if the coordinates are moved at a linear rate, or a function of some other variable $y$, if the coordinates are moved by a servomotor in proportion to the variable $y$.

While all the above devices are slow and only suited to slow-speed computers giving plotted outputs, the accuracy is very high, often of the order of a fraction of 1 per cent.

Early forms of computers used rotating switches to introduce potentials tapped from a resistive potentiometer as in Fig. 8-32*c*. This provides a step approximation to the desired time function, which can be made as close as desired by increase in the number of taps and commutator segments. Other mechanical refinements can be used if linear variation between tap points is desired.

The rotating switch is the only one of the above devices suited to the repetitive, or high-speed, analog computer. Two other electronic devices, the *photoformer* and the *monoformer*, have come into use to generate $e_o$ as a function of time or of another variable $y$. The photoformer is illustrated by the block diagram of Fig. 8-33. The input voltage $e_i$ drives the cathode-ray beam in the horizontal direction. At the same time the positioning bias applied to the d-c amplifier tends to deflect the beam upward, so that the screen spot rises above the mask. However, as soon as the spot appears above the mask edge, its illumination on the photocell develops an additional input to the d-c amplifier so polarized as to deflect the beam and spot downward. Equilibrium is reached with the spot just riding along the top edge of the mask.

Optical nonlinear feedback is thus provided for the d-c amplifier in the vertical-deflection circuit of the cathode-ray tube, the feedback path comprising the cathode-ray beam, the nonlinear mask, and the photocell. Since the output of a feedback amplifier is strongly dependent on the

characteristics of the feedback path if the gain is high, the output $e_o$ is dependent on, or is a function of, the mask shape.

It may also be seen that, as the spot is swept horizontally by the input $e_i$, the output voltage $e_o$ required to maintain the spot at the mask edge is made to vary as some function of $e_i$, this function being determined by the shape of the mask boundary. It is usually best to determine the exact mask shape experimentally, as variations in cathode-ray-tube sensitivity and in parallax induced by photocell size and location make calculations difficult. Accuracy to 1 or 2 per cent can usually be obtained if the mask is close to the tube surface and the spot intensity is kept low enough to prevent visual broadening of the spot. Speed of response is not a limiting factor with the usual oscillograph amplifiers

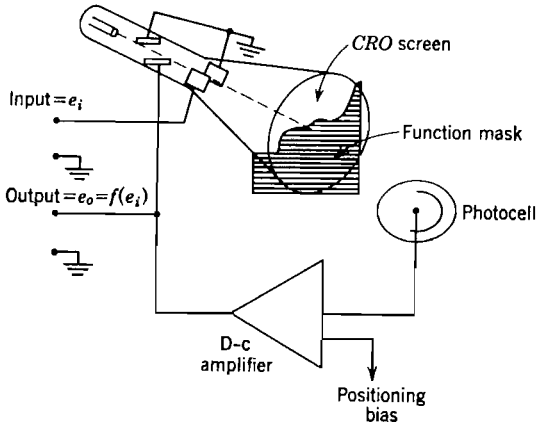



FIG. 8-33. Photoformer for generating nonlinear function.

and short-persistence screen tubes, although some difficulty may be expected in gaining good reproduction of steps or discontinuities in the mask pattern.

The monoformer places the mask inside the tube, reducing parallax error. Instead of a fluorescent screen the cathode-ray tube end has a mask made of a poor secondary-electron emitter such as graphite, the unmasked area being a good secondary emitter, such as aluminum. A collector electrode is positioned to receive the secondary electrons. With the d-c amplifier biased as for the photoformer, when the beam rises above the masked area the resultant shower of secondary electrons reaching the collector introduces an amplifier input polarized to drive the beam down. Equilibrium is again reached with the beam on the edge of the masked area.

The disadvantage of difficult and expensive mask change is somewhat balanced by lack of parallax and greater accuracy. Different tubes must be stocked for the expected variety of functions.

**8-19. Nonlinear-circuit Elements.** While vacuum diodes, triodes, crystal diodes, and some special materials such as *thyrite* will yield nonlinear volt-ampere characteristics, stability and reproducibility are somewhat lacking. Building up of nonlinear characteristics by addition of straight-line segments is more accurate and stable, and such a process is possible with reasonable numbers of crystal diodes, employed only as switches. The process is illustrated in Fig. 8-34, where three straight-line segments are utilized to approximate a volt-ampere curve. The conductances required for each segment are

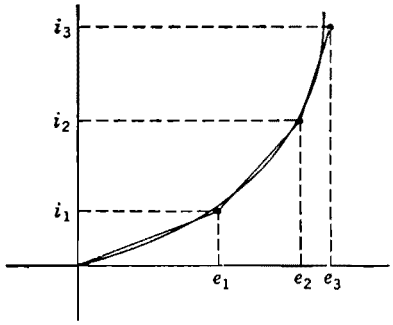



FIG. 8-34. Approximation of a curve by three straight lines.

$$g_1 = \frac{i_1}{e_1} \qquad g_2 = \frac{i_2 - i_1}{e_2 - e_1}$$

$$g_3 = \frac{i_3 - i_2}{e_3 - e_2}$$

if suitable switching be provided at $o$, $e_1$, and $e_2$ to place the proper conductances in the circuit.

Crystal diodes are well suited to operate as the needed switches. Since they are inherently rectifiers, the diode represents an open circuit when the cathode is more positive than the anode, whereas when the anode is more positive than the cathode the diode has an internal resistance in the range of a few ohms to a few hundred ohms. The diode thus is a potential-actuated switch in a circuit such as that of Fig. 8-35*a*. The volt-ampere curve of the circuit is that of Fig. 8-35*b*, showing that when the applied voltage makes terminal $A$ more than $E$ volts positive then the circuit appears as a resistance $R$ or conductance $1/R$ ($R \gg r_b$ of the diode). Modification of the circuit to that of Fig. 8-35*c* provides ready adjustment of the switching potential in either positive or negative direction and variation of the resistances appearing in the circuit.

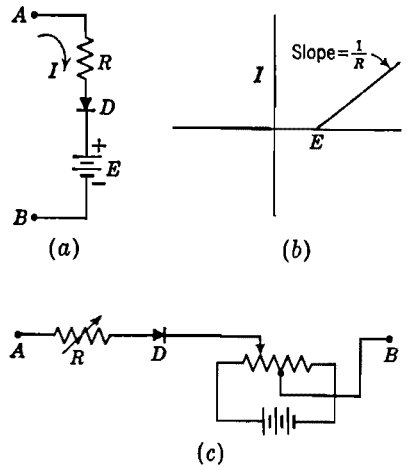



FIG. 8-35. Simple nonlinear-diode elements.

By combining a number of such diode circuits it is possible to generate

nonlinear-circuit elements with almost any desired relation between input and output voltages. A combination is shown in Fig. 8-36*a* along with a resultant volt-ampere curve yielding a nonlinear resistance decreasing with applied voltage $E$. Resistance increase with voltage is obtained by the modification in Fig. 8-36*b*. By the use of straight-line segments in this manner it is possible to approximate many mathematical functions. For instance, a curve of $E = KI^2$ can be represented to an

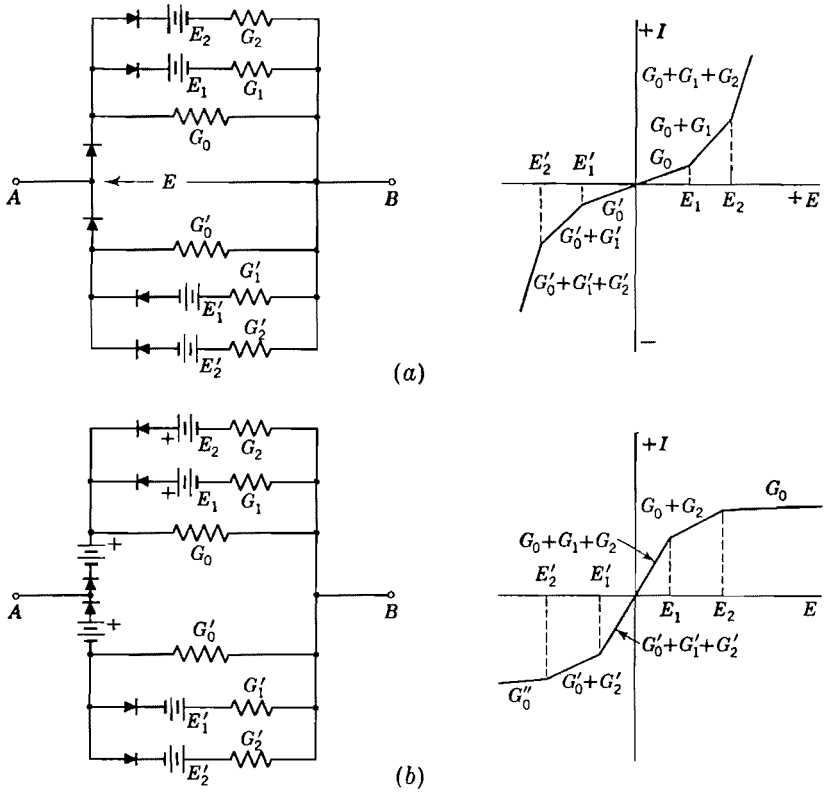



FIG. 8-36. Nonlinear resistors formed by diode switching.

accuracy of 1 per cent with five straight-line segments. Due to the fact the diode changes from open circuit to low resistance over a small range of voltage, some curvature is introduced at each intersection, and this improves the fit to a desired curve. By this means an electronic instrument for measuring rms input quantities has been designed.

These combinations may also be used in the feedback paths of operational amplifiers to yield additional interesting properties.

**8-20. Limiting Circuits.** In the study of some systems, particularly those encountered in automatic control, it is necessary to introduce into

the computer solution the idea of limiting or overload. This might occur in a control system if a valve reached the limit of its travel and could allow no more flow regardless of increasing demand on its actuator.

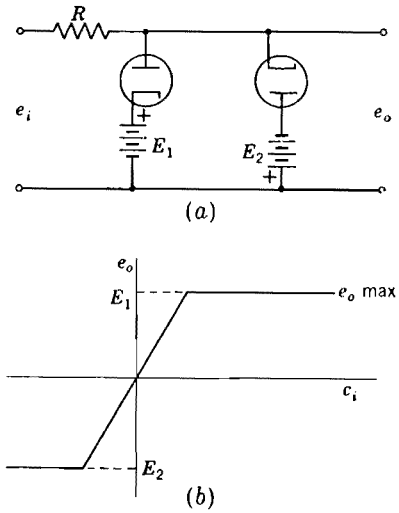



FIG. 8-37. Limiter circuit and characteristics.

This limiting action is readily obtained through simplifying the nonlinear diode circuits of the preceding section to two diodes, as in Fig. 8-37*a*, which is also the clipper of Chap. 9. In this usage $R \gg r_b$ of the diodes. The bias voltage should be taken from a power supply of low internal resistance; otherwise the flat limiting portion of the output characteristic will be somewhat curved. Voltages $E_1$ and $E_2$ should be adjustable to provide choice of the limiting value.

**8-21. Multiplying Devices and Circuits.** It is easily possible to multiply a variable by a constant of value less than unity by use of the potentiometer of Fig. 8-38*a* or by use of the operational amplifier. Multiplication of a variable $e_1$ by a second variable $e_2$ is more difficult.

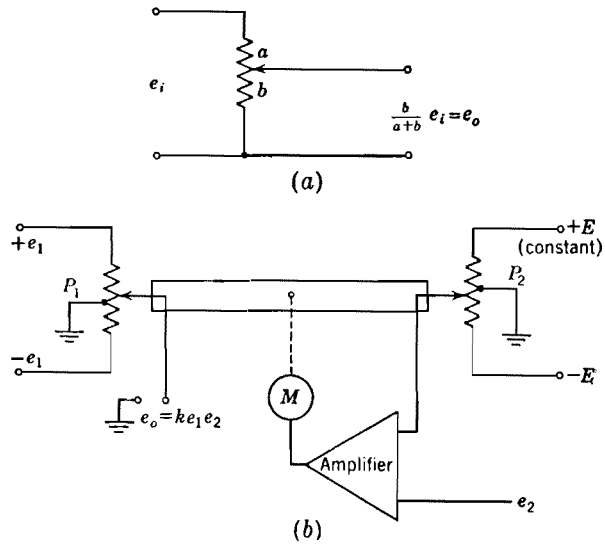



FIG. 8-38. (*a*) Multiplication of a variable by a constant. (*b*) Multiplication of two variables.

For the slow-speed class of computer this operation is usually performed by a potentiometer $P_1$ with an applied potential proportional to $e_1$ and with the tap positioned by a servomotor in proportion to $e_2$, as given below. A second potentiometer $P_2$ mounted on the same shaft with $P_1$ has applied to it a fixed and constant potential $E$. An amplifier compares the second potential $e_2$ and the tapped voltage on $P_2$ and drives the motor to adjust $P_2$ to bring the difference of potentials on the amplifier input to zero. Since $P_1$ and $P_2$ are on the same shaft, the tap on $P_1$ is being adjusted in accordance with $e_2$ and its over-all voltage is $e_1$.

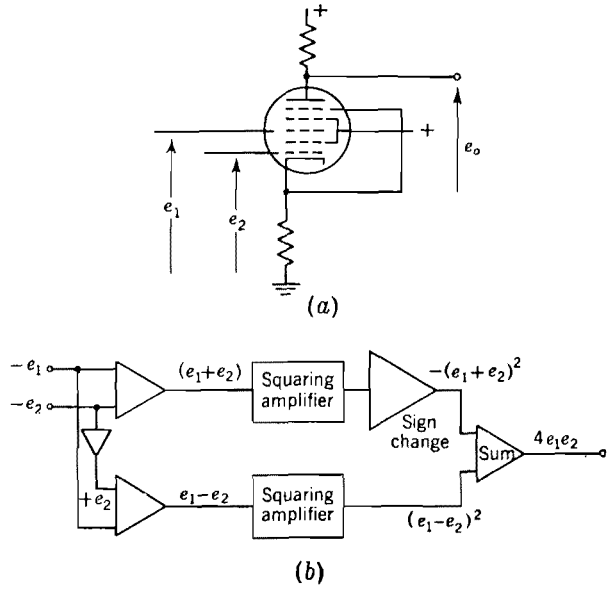



FIG. 8-39. Electronic forms of multipliers.

Thus the output voltage from the tap on $P_1$ will be proportional to $e_1$ times $e_2$. As shown, the multiplier has ranges from $+e_2$ to $-e_2$. Such devices are highly accurate but limited in speed of response to that of the servomechanism, usually below 2 cycles.

The problem of a multiplier for the high-speed repetitive computer has not yet been completely solved. The multigrid vacuum tube (6L7, for example) of Fig. 8-39*a* will have an output proportional to the product of $e_1$ and $e_2$ over a limited input range but is dependent on selection and maintenance of certain tube characteristics.

It is possible to select tubes to have grid-transfer characteristics which show pronounced quadratic form over limited ranges, that is,

$$i_b = A e_c^2$$

to a reasonable approximation. Use of such tubes in the circuit of Fig. 8-39*b*, as squaring elements, gives an output of

$$e_1 e_2 = \frac{(e_1 + e_2)^2}{4} - \frac{(e_1 - e_2)^2}{4} \tag{8-57}$$

which for obvious reasons is referred to as a *quarter-square* multiplier. While suited to analog multiplication by reason of speed, its accuracy is dependent on the maintenance of critical shapes in the transfer characteristics of several tubes, and these cannot be maintained to a high degree.

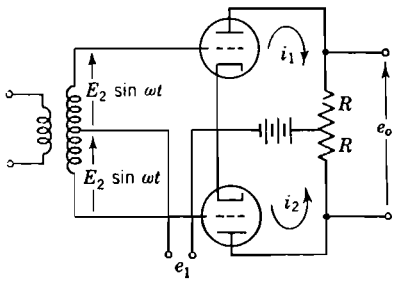



FIG. 8-40. Balanced modulator multiplier.

Another all-electronic method is based on the principle of amplitude modulation as in Fig. 8-40. Let a slowly varying voltage $e_1$ and an a-c carrier voltage $e_2 = E_2 \sin \omega t$ be applied to a pair of tubes having assumed square-law grid-transfer characteristics given by

$$i_b = a e_c^2$$

For tube $T_1$ the plate current will be due to an input voltage $e_1 + E_2 \sin \omega t$ so that

$$\begin{aligned} i_1 &= a(e_1 + E_2 \sin \omega t)^2 \\ &= a e_1^2 + 2a e_1 E_2 \sin \omega t + a E_2^2 \sin^2 \omega t \end{aligned}$$

and for tube $T_2$

$$\begin{aligned} i_2 &= a(e_1 - E_2 \sin \omega t)^2 \\ &= a e_1^2 - 2a e_1 E_2 \sin \omega t + a E_2^2 \sin^2 \omega t \end{aligned}$$

The output voltage $e_o$ is given by

$$\begin{aligned} e_o &= (i_1 - i_2)R \\ &= 4a e_1 E_2 \sin \omega t \end{aligned}$$

This is the desired product term. It may be rectified back to direct-current for reintroduction into the computer. The carrier frequency may be initially produced by a chopper and must be high with respect to the highest input frequencies present in $e_1$. Since the output is proportional to $a$, the accuracy is dependent on tube characteristics and is not likely to be much better than 5 to 10 per cent.

The photoformer and monoformer may be employed as multipliers, through use of logarithmic and antilogarithmic masks, in the circuit of Fig. 8-41. Division could be carried out as easily. The operation is

restricted to positive values of $e_1$ and $e_2$ since logarithms of negative numbers are not defined.

Several other multiplier systems are based on the principle of sampling voltage $e_1$ periodically, the length of the sample pulses varying linearly with the other variable $e_2$. Thus the amplitude of the sample pulse is proportional to the value of $e_1$ at a particular instant, the length of the pulse is proportional to $e_2$ at that instant, and the area (or average) of the pulse is then proportional to the desired product $e_1e_2$. The only requirement on sampling frequency is that it be high with respect to the highest input frequency. With sampling frequencies of the order of 40 to 50 kc this condition is met for most applications.

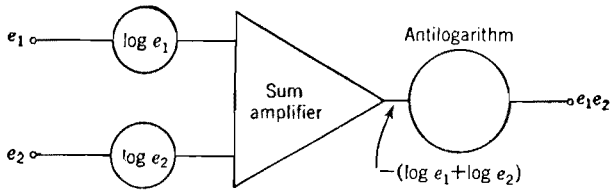



FIG. 8-41. Use of three photoformers with logarithmic masks to produce multiplication.

The output is the averaged value of the pulses, and it is possible to design multipliers for four-quadrant operation, or for values of $e_1$ and $e_2$ between plus and minus limits, and for $-100$- to $+100$-volt inputs.

## PROBLEMS

**8-1.** In Fig. 8-1, specify all resistors and battery voltages if the tubes have $\mu = 70$, $r_p = 40{,}000$ ohms and are to operate at $E_b = 100$ volts, $E_c = -2$ volts. The over-all gain is to be 2250, with first-stage gain 150 per cent of the second-stage gain, and $I_{b1} = 1.25$ ma, $I_{b2} = 1.50$ ma.

**8-2.** In Fig. 8-6, $T_1 = T_2$ and $\mu = 70$, $r_p = 45{,}000$ ohms; $T_3 = T_4$ and $\mu = 20$, $r_p = 7700$ ohms. Compute the over-all gain, with the zero adjuster set for 2000 ohms.

**8-3.** The circuit of Fig. 8-2 is to be redesigned. The tubes are halves of a 6SN7 and are to operate at

$$E_{b1} = 200 \text{ volts} \qquad E_{c1} = -4 \text{ volts}$$
$$E_{b2} = 250 \text{ volts} \qquad E_{c2} = -6 \text{ volts}$$

If $R_1 = 25{,}000$ ohms, $R_2 = 15{,}000$ ohms and the battery is to supply a total current of 20 ma, find the bleeder current and the required power-supply voltage, and specify the potentials at all bleeder taps and at the grids.

**8-4.** In Fig. 8-8 the meter $M$ has a resistance of 1500 ohms. If $R_a = 15{,}000$ ohms, $R_k = 1000$ ohms, $\mu = 20$, $r_p = 7700$ ohms and the circuit is balanced for $e_i = 0$, find the current through the instrument per volt input.

**8-5.** Prove that Eq. (8-17) is correct for the gain of the Miller cathode-drift compensator.

**8-6.** Derive an expression for the gain of the d-c amplifier of Fig. 8-42, where $e_1$ is the grid-to-grid input and the tubes are identical.

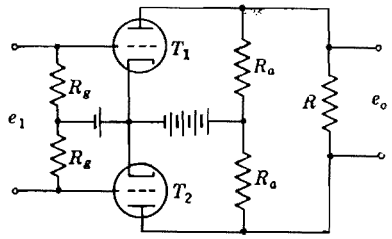
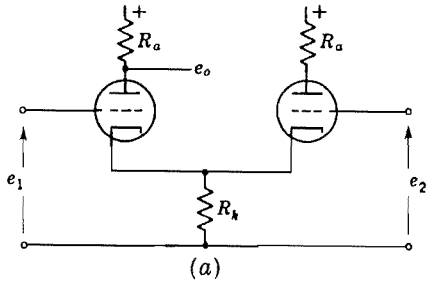


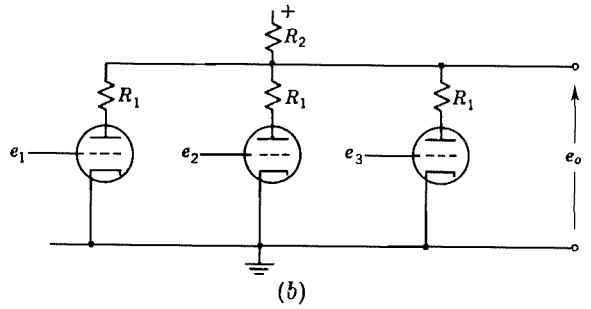

FIG. 8-42

**8-7.** Show that Eq. (8-24) is correct for the bridge-balanced amplifier operated with grounded center tap on the power supply and with $R_L$ large.

**8-8.** Show that the circuit of Fig. 8-43*a* will subtract or that $e_o = k(e_1 - e_2)$.




(*a*)




(*b*)

FIG. 8-43

**8-9.** Demonstrate that the circuit of Fig. 8-43*b* will add, or that $e_o = K(e_1 + e_2 + e_3)$.

**8-10.** For each operational circuit of Fig. 8-44 determine the relation of $e_o$ to $e_i$ in terms of operator $p$.

**8-11.** By use of Lagrange's equation prove that Eq. (8-50) is correct for the system of Fig. 8-27*a*.

**8-12.** Using Lagrange's equation, develop the equation of motion of a pendulum, assuming all mass concentrated in the bob and the angle of swing small.

**8-13.** (*a*) Draw a block diagram for a functional analog computer to determine $y$ for the equation

$$\ddot{y} - f(t) = 0$$

(*b*) If $f(t)$ is a step function [$f(t) = 0$ for $t < 0$, $f(t) = 1$ for $t > 0$], draw the value of $y$ and $\dot{y}$ as functions of time.

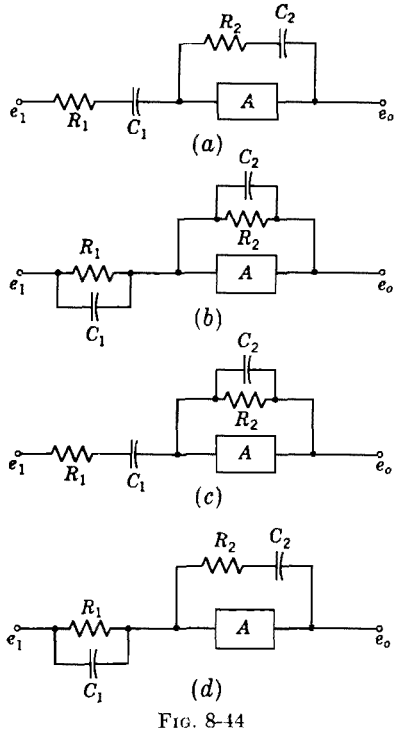



FIG. 8-44

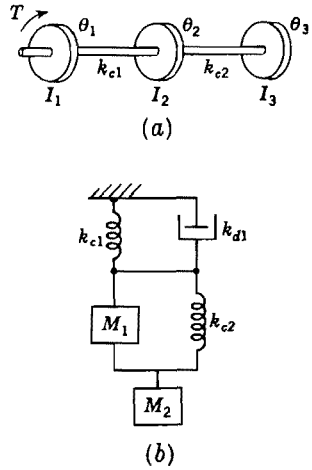



FIG. 8-45

**8-14.** Indicate by a block diagram how an operational amplifier may be made to integrate the sum of three variables.

**8-15.** Set up a block diagram for an operational solution of the equation

$$\ddot{y} + a\dot{y} + y = 0 \qquad \text{for } 0 < a < 1$$

**8-16.** Draw an electric circuit for direct-analog solution of the equation of Prob. 8-15, and list the analogous variables.

**8-17.** Set up an electric circuit for the mechanical system of Fig. 8-45*a*, and list the analogous quantities. For what mechanical conditions will $\theta_2$ be maximum?

**8-18.** By Lagrange's equations set up the equations of motion for inertia $I_3$ of Fig. 8-45*a*.

**8-19.** Set up the direct-analog circuit for the mechanical system of Fig. 8-45*b*.

## REFERENCES


1. Loftin, E. H., and S. Y. White: Cascaded Direct-coupled Tube Systems Operated from Alternating Current, *Proc. IRE*, **18**: 669 (1930).
2. Valley, G. E., Jr., and H. Wallman: "Vacuum Tube Amplifiers," McGraw-Hill Book Company, Inc., New York, 1948.
3. Miller, S. E.: Sensitive D-C Amplifier with A-C Operation, *Electronics*, **14**: 27 (November, 1941).
4. Yu, Y. P.: Cathode-follower Coupling, *Electronics*, **19**: 99 (1946).
5. Rittenhouse, J. W.: Cathode Drift Compensation in D-C Amplifiers, *Elec. Eng.*, **72**: 299 (1953).
6. Artzt, M.: Survey of D-C Amplifiers, *Electronics*, **18**: 112 (August, 1945).
7. Korn, G. A., and T. M. Korn: "Electronic Analog Computers," McGraw-Hill Book Company, Inc., New York, 1952.
8. Williams, A. J., R. E. Tarpley, and W R. Clark: D-C Amplifier Stabilized for Zero and Gain, *Trans. AIEE*, **67**: 47 (1948).
9. Goldberg, E. A.: Stabilization of D-C Amplifiers, *RCA Rev.*, **11**: 296 (1950).
10. Dubridge, L. A.: The Amplification of Small Direct Currents, *Phys. Rev.*, **37**: 392 (1931).
11. Collins, D. L.: Electrometer Tubes, *Instruments*, **26**: 1708 (1953).
12. Victoreen, J. A.: Electrometer Tubes for the Measurement of Small Currents, *Proc. IRE*, **37**: 432 (1949).
13. Ragazzini, J. R., R. H. Randall, and F. A. Russell: Analysis of Problems in Dynamics by Electronic Circuits, *Proc. IRE*, **35**: 444 (1947).
14. Chance, F. B., D. Sayre, V. Hughes, E. F. MacNichol, Jr., and F. C. Williams: "Waveforms," McGraw-Hill Book Company, Inc., New York, 1949.
15. Ginzton, E. L.: D-C Amplifier Techniques, *Electronics*, **17**: 98 (March, 1944).
16. Bush, V.: The Differential Analyzer, A New Machine for Solving Differential Equations, *J. Franklin Inst.*, **212**: 447 (1931).
17. Harder, E. L., and G. D. McCann: A Large-scale General Purpose Electric Analog Computer, *Trans. AIEE*, **67**: 664 (1948).
18. Sunstein, D. E.: Photoelectric Wave Form Generator, *Electronics*, **22**: 100 (1949).
19. MacNee, A. B.: An Electronic Differential Analyzer, *Proc. IRE*, **37**: 1315 (1949).
20. Soroka, W. W.: "Analog Methods in Computation and Simulation," McGraw-Hill Book Company, Inc., New York, 1954.
21. Korn, G. A.: Design and Construction of Universal Function Generating Potentiometers, *Rev. Sci. Instr.*, **21**: 77 (1950).
22. Murphy, Glenn: "Similitude in Engineering," The Ronald Press Company, New York, 1950.

CHAPTER 9

# SWITCHING CIRCUITS; DIGITAL COMPUTATION

In previous chapters the ability of the vacuum tube faithfully to amplify a given input voltage has been studied and emphasized. It is now necessary to introduce another large class of vacuum-tube circuits in which emphasis is on the ability of the tube to operate as a switch. In this latter class the output voltage may bear no relation to the input magnitude or waveform, except to be present or absent in accordance with certain levels of input.

The vacuum tube thus used is an inertialess and nearly instantaneous switch. Having speed of operation as its main advantage, its major difficulty is that when closed it still represents a small but appreciable resistance. When open, of course, it has a resistance which is essentially infinite. The low value of closed resistance must be allowed for in circuit design but can be neglected in many cases.

Circuits employing vacuum tubes as switches are frequently used to generate pulses and other nonsinusoidal wave shapes or to operate with or by such pulses or waveforms. Radar equipment forms one such class of circuit, with television, digital computers, and various pulse-modulated communication systems also included.

**9-1. Circuit Elements.** As mentioned in Sec. 3-1, resistive circuit elements may have linear or nonlinear volt-ampere characteristics. A nonlinear resistor is represented by a curve in Fig. 9-1*a* and obeys a law

$$e = R(i)i \tag{9-1}$$

where the proportionality coefficient $R(i)$ is a function of current. A linear resistor follows the usual law

$$e = Ri \tag{9-2}$$

Some resistance elements, particularly crystal or vacuum diodes, may be unilateral (and therefore nonlinear) as in Fig. 9-1*b*, where the volt-ampere curve is dissimilar for positive and negative voltages.

In circuits having nonlinear resistive elements it is necessary to employ the concepts of *static* and *dynamic* resistance as in Sec. 3-2. These are

illustrated in Fig. 9-2. As defined, the static resistance is

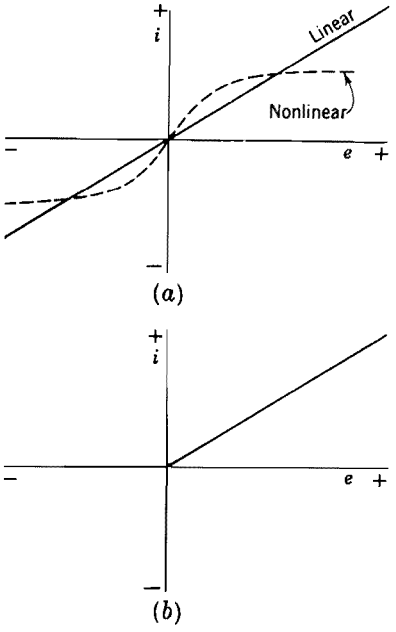



FIG. 9-1. Resistive volt-ampere characteristics.

$$r_b = \frac{e_A}{i_A} \tag{9-3}$$

at point $A$, and the dynamic resistance is

$$r_p = \frac{de_A}{di_A} \tag{9-4}$$

or the reciprocal of the slope at the point $A$. For a linear resistor these values are identical, but in Fig. 9-2 the values of $r_b$ and $r_p$ will vary from point to point, and near $A$ the value of $r_p$ is negative.

For capacitive-circuit elements

$$e_C = E_o + \frac{1}{C}\int_o^t i_C\, dt \tag{9-5}$$

$$i_C = C\frac{de_C}{di_C} \tag{9-6}$$

and for inductive-circuit elements

$$e_L = L\frac{di_L}{dt} \tag{9-7}$$

$$i_L = I_o + \frac{1}{L}\int_o^t e\, dt \tag{9-8}$$

It is apparent that the potential across a capacitor or the current through an inductor cannot change instantaneously, without an infinite applied emf.

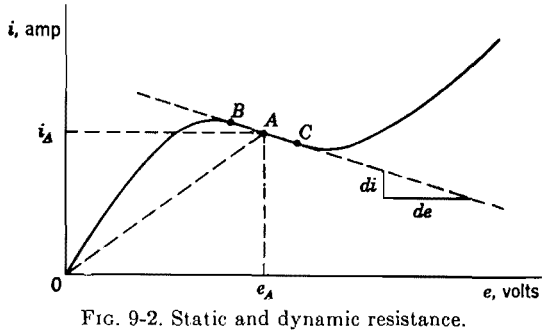
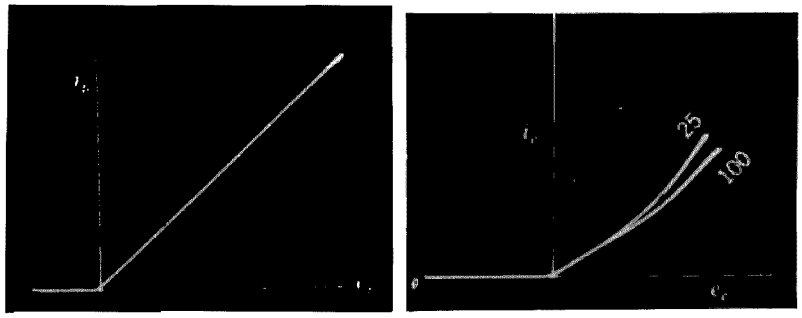



FIG. 9-2. Static and dynamic resistance.

**9-2. Tubes in Switching Circuits.** While diode-tube characteristics are nonlinear, the degree of nonlinearity is frequently not great enough to warrant use of graphical analysis in switching circuits. Sufficient

accuracy may usually be obtained by assuming a constant average value of static resistance, which may be as low as 1000 to 2000 ohms for many vacuum diodes, and this may often be completely neglected when the diode is in series with a much larger resistor. The near approach to linearity of such a characteristic is shown in Fig. 9-3*a*.

Similar treatment may be used for the plate circuits of triodes, although recourse must usually be had to the actual characteristics for the dynamic plate resistance and actual current values. Triodes may often be operated with grid voltages swinging from cutoff to zero or positive values. With a positive grid voltage the static resistance of the grid-cathode circuit of the usual triode may be taken as 1000 ohms as a good approximation. The linearity of the grid resistance for several positive grid voltages is shown in Fig. 9-3*b*, being good to +50 volts.




(*a*) (*b*)

FIG. 9-3. (*a*) Dynamic characteristic of 6H6 diode with load of 20,000 ohms. (*b*) Grid volt-ampere curves for a 6J5 triode with +100 volts $e_{c,max}$, $E_b$ = 25 and 100 volts.

Crystal diodes are frequently employed and offer lower static resistance than vacuum diodes, and with no current at zero applied emf as is the case of the vacuum diode. Static resistances as low as a few hundred ohms are possible, although the back resistance (with anode negative) will not be so high as with thermionic diodes. A typical volt-ampere curve for a crystal diode is given in Fig. 9-4*a*.

Gas diodes and triodes may be considered as having an idealized volt-ampere characteristic as in Fig. 9-4*b*. The constant voltage drop is assumed independent of current value.

**9-3. Response of *RC* and *RL* Series Circuits.** In many applications desired waveforms are built up from segments of *RC* or *RL* transient voltages, with triodes or diodes employed as switches. A discussion of *RC* and *RL* transients will be introduced here for reference.

Consider the circuit of Fig. 9-5*a*, in which switch *s* is thrown to position 2 at $t = 0$. The potential on *C* at $t = 0$ is $E_o$, which may or may not be equal to $E_1$, dependent on the time *s* has been in position 1. At $t = 0$

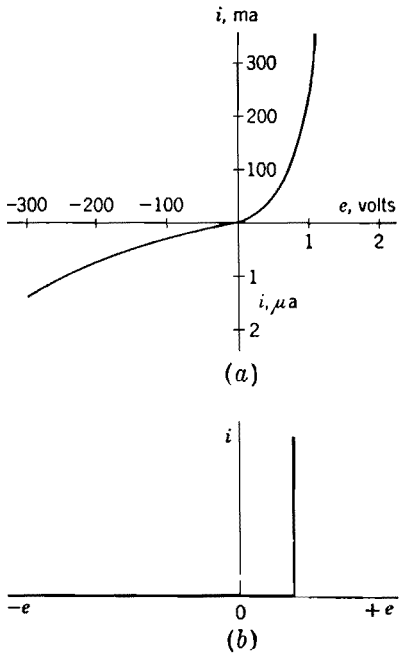



FIG. 9-4. (*a*) Germanium crystal-diode volt-ampere curve. (*b*) Ideal-gas-diode volt-ampere curve.

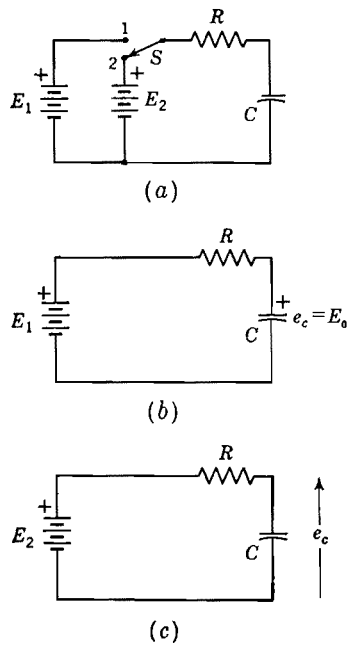



FIG. 9-5. Switching in an *RC* circuit.

the circuit changes from that of Fig. 9-5*b* to that of Fig. 9-5*c*, and the switching action might be obtained by proper operation of a triode. Writing an equation for the circuit of Fig. 9-5*c* at $t = 0$ gives

$$iR + \frac{1}{C}\int_o^t i\,dt + E_o - E_2 = 0 \tag{9-9}$$

Differentiating and dividing by $R$ gives

$$\frac{di}{dt} + \frac{i}{RC} = 0$$

or

$$\left(p + \frac{1}{RC}\right) i = 0 \tag{9-10}$$

The solution of Eq. (9-10) is

$$i = Ae^{-t/RC}$$

and from Eq. (9-9), at $t = 0$, where $i = i(0)$,

$$i(0) = \frac{E_2 - E_o}{R}$$

It can be seen that $A = (E_2 - E_o)/R$ and

$$i = \frac{E_2 - E_o}{R} \epsilon^{-t/RC} \tag{9-11}$$

The resistor voltage is then

$$e_R = iR = (E_2 - E_o)\epsilon^{-t/RC} \tag{9-12}$$

and the capacitor voltage is

$$\begin{aligned} e_C = E_o + \frac{1}{C}\int_o^t i\,dt &= E_o + \frac{1}{C}\int_o^t \frac{E_2 - E_o}{R} \epsilon^{-t/RC}\,dt \\ &= E_o + (E_2 - E_o)(1 - \epsilon^{-t/RC}) \end{aligned} \tag{9-13}$$

Equations (9-12) and (9-13) serve to introduce the exponential terms $\epsilon^{-t/RC}$ and $1 - \epsilon^{-t/RC}$, which will be discussed later.

In a similar manner an $RL$ circuit with a steady current $I_o$ flowing may have switch $s$ thrown from 1 to 2 at $t = 0$, as in Fig. 9-6. A circuit equation for $t > 0$ may be written as

$$L\frac{di}{dt} + Ri - E_2 = 0$$

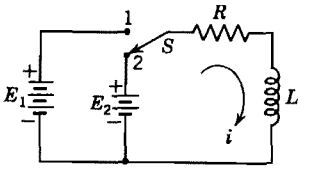



FIG. 9-6. Series $RL$ circuit.

Separation of variables and integration gives

$$E_2 - Ri = B\epsilon^{-Rt/L} \tag{9-14}$$

Since $i = I_o$ at $t = 0$,

$$A = E_2 - RI_o$$

and then

$$i = I_o + \left(\frac{E_2}{R} - I_o\right)(1 - \epsilon^{-Rt/L}) \tag{9-15}$$

The voltage across the resistor is

$$e_R = iR = I_oR + (E_2 - I_oR)(1 - \epsilon^{-Rt/L}) \tag{9-16}$$

and that across the inductor (assumed resistanceless) is

$$e_L = L\frac{di}{dt} = (E_2 - I_oR)\epsilon^{-Rt/L} \tag{9-17}$$

It can be seen by comparison of Eqs. (9-13) and (9-16) that $e_C$ for the $RC$ circuit and $e_R$ for the $RL$ circuit are of similar form, likewise $e_R$ for the $RC$ circuit and $e_L$ for the $RL$ circuit. The $RL$-circuit equations also involve the terms $\epsilon^{-Rt/L}$ and $1 - \epsilon^{-Rt/L}$, of similar form to those for the $RC$ circuit.

**9-4. The Time Constant.** All single-capacitor or single-inductor $RC$ and $RL$ circuits involve the terms $\epsilon^{-t/RC}$ and $\epsilon^{-Rt/L}$. The reciprocals of

the exponent coefficients $RC$ and $L/R$ have dimensions of time and are known as the *time constants* of the circuits. These terms determine the rate of change of the exponential term and are defined as the time in seconds for the output transient to change $1/\epsilon$, or 64 per cent of its infinite time value. It is convenient to measure time of response of a circuit in units of time constant, since universal curves as in Fig. 9-7 may be plotted

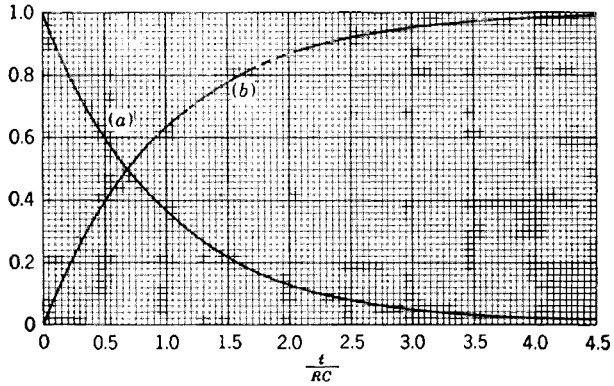



FIG. 9-7. (*a*) $\epsilon^{-t/RC}$. (*b*) $1 - \epsilon^{-t/RC}$.

and used to determine values of the exponential terms in various transient switching problems. The per cent of final value attained in lengths of time equal to multiples of circuit time constant is indicated in Table 16.

TABLE 16. TIME-CONSTANT VALUES

| *Time constant* | *Per cent of complete transient* | *Time constant* | *Per cent of complete transient* |
|---|---|---|---|
| 0 | 0.0 | 3.0 | 95.5 |
| 0.5 | 39.4 | 3.5 | 97.0 |
| 1.0 | 63.2 | 4.0 | 98.2 |
| 1.5 | 77.7 | 4.5 | 98.9 |
| 2.0 | 86.5 | 5.0 | 99.4 |
| 2.5 | 91.8 | 10.0 | 99.999 |

It is frequently sufficiently accurate to assume the transient completed after four time constants, when the value of the function is within 1.8 per cent of the final value.

**Example.** The switch in Fig. 9-8 has been closed on position 1 for a long time. It is then switched to position 2 for 750 $\mu$sec, then back to position 1. Sketch the voltage waveform for $e_o$ during the complete transient.

Since the switch $s$ has been closed on 1 for a long time, the value of $e_o$ at $t = 0$ is taken as 250 volts. At $t = 0$ the switch is thrown to 2, and the capacitor starts to discharge. From Eq. (9-11) with $E_2 = 0$, $E_o = 250$ volts, and $RC = 0.01 \times 10^{-6} \times$

$50 \times 10^3 = 500 \times 10^{-6}$ sec,

$$i = \frac{250}{50 \times 10^3} \epsilon^{-t/(500\times10^{-6})}$$

Voltage $e_o$ is given by use of Eq. (9-13) as

$$e_o = -Ri + e_C = -10^4 i + 250 + (-250)(1 - \epsilon^{-t/(500\times10^{-6})})$$
$$= 200\epsilon^{-t/(500\times10^{-6})}$$

At 750 $\mu$sec, which is equal to 1.5 time constants, the value of the exponential term is 0.22 from Fig. 9-7, and at this instant the switch $s$ returns to position 1. The voltage $e_o$ at the switching instant is

$$e_o = 200 \times 0.22 = 44.0 \text{ volts}$$

Intermediate values to draw the waveform may be obtained from Fig. 9-7.

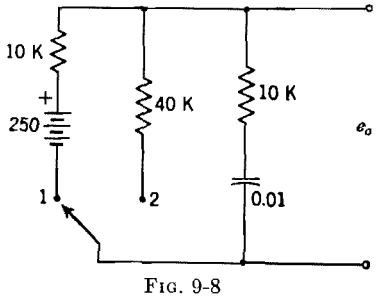


FIG. 9-8

At the switching instant the capacitor potential $e_C$ is given by Eq. (9-13) as

$$e_C = 250 + (-250)(1 - \epsilon^{-1.5}) = 55.0 \text{ volts}$$

and this is the new value for $E_o$ during the second time interval. The time constant for the circuit with $s$ on 1 or for $t + 750 \times 10^{-6}$ sec is

$$RC = 0.01 \times 10^{-6} \times 20 \times 10^3 = 200 \times 10^{-6} \text{ sec}$$

and the current will be given by

$$i = \frac{250 - 55}{2 \times 10^4} \epsilon^{-t/(200\times10^{-6})}$$
$$= 97.5 \times 10^{-4}\epsilon^{-t/(200\times10^{-6})}$$

and voltage $e_o$ will be given by Eq. (9-13) as

$$e_o = 10^4 i + 55 + (250 - 55)(1 - \epsilon^{-t/(200\times10^{-6})})$$
$$= 250 - 97.5\epsilon^{-t/(200\times10^{-6})}$$

At the instant of switching ($t = 0$ for the above equation) the value of $e_o$ jumps to 152.5 volts and then rises exponentially toward its final value of 250 volts. Intermediate values may be found from Fig. 9-7. The complete waveform is plotted in Fig. 9-9.

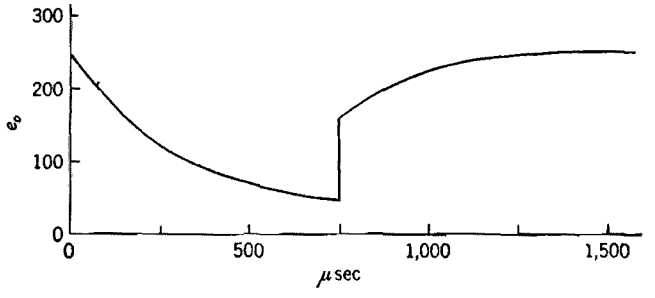


FIG. 9-9. Waveform for example.

**9-5. The Differentiating Circuit.** Referring to a circuit of the form of Fig. 9-10*a*, with the initial value of $e_C = 0$, then

$$e_i = \frac{1}{C}\int_o^t i\,dt + iR \tag{9-18}$$

By making the *RC time constant small with respect to the interval of any applied voltage*, the capacitor $C$ will charge and discharge rapidly, and the integral in the equation above will be large or important in relation to the $iR$ term. The capacitor will essentially determine the value of current, or

$$i \cong C\frac{de_i}{dt} \tag{9-19}$$

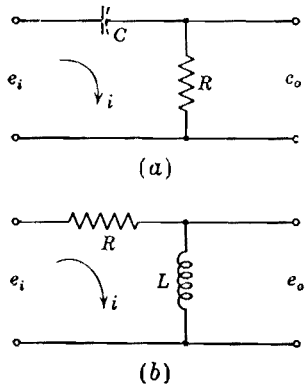


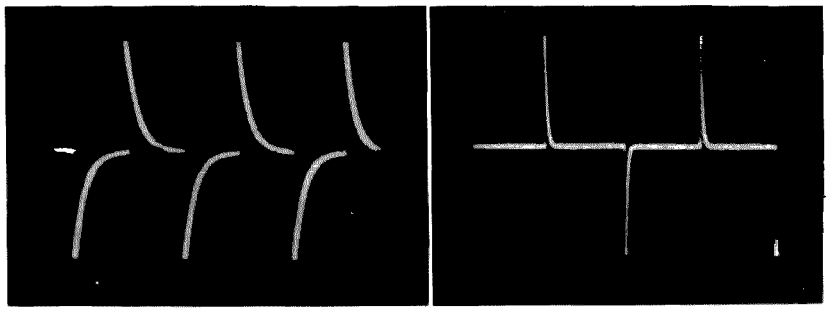

FIG. 9-10. Differentiating circuits.

The output voltage will then be $iR$, and under the requirement of the $RC$ assumption this gives

$$e_o = iR \cong RC\frac{de_i}{dt} \tag{9-20}$$

Thus, with a *small* value of time constant $RC$, the output voltage will be proportional to the derivative of the input voltage, and the circuit is known as a *differentiating circuit.*

Theoretically the output for an applied square wave of voltage should be a series of alternate positive and negative pulses of zero width in time.



(a) (b)

FIG. 9-11. Output of an $RC$ differentiator with square-wave period $T = 2 \times 10^{-4}$ sec. (*a*) With $RC/T = 0.05$; (*b*) with $RC/T = 0.0005$.

As the time constant is made smaller, this condition is approached. The amplitude of the pulses will be equal to the instantaneous change of input voltage of the input wave. In Fig. 9-11 are shown the output-voltage waveforms for several ratios of the $RC$ value to the period of

the applied wave. As $RC$ is made smaller, the $e_o$ voltage approaches more closely the spike form to be expected for a derivative of a square input wave.

This circuit is frequently used to sharpen pulse waveforms or to generate pips, or sharp pulses, for triggering other electronic circuits.

The circuit of Fig. 9-10*b* also has differentiating properties. Writing

$$e_i = L\frac{di}{dt} + iR \tag{9-21}$$

then if $L/R$ is small, the applied potential will largely appear across $R$, or $R$ will determine the current $i$, so that

$$i \cong \frac{e_i}{R}$$

and then

$$e_o \cong \frac{L}{R}\frac{de_i}{dt} \tag{9-22}$$

and thus the inductive form of circuit will also differentiate the input waveform. Since physical inductors have some resistance, the output voltage will have an $iR$ component and will not be so accurate a derivative as given by the $RC$ circuit.

The $RC$ form of Fig. 9-10*a* is also that of the $RC$ coupling for amplifiers. In such applications, to avoid differentiation and distortion of the waveform, it is desirable that the time constant $RC$ be very long with respect to the lowest period of the applied voltage. This circuit has been previously discussed in Sec. 5-8.

**Example.** A square wave of peak value +100 volts is applied to an $RC$ differentiator of time constant 0.05 of the period of the wave. Assuming initial value of the wave equal to zero and initial value of $e_C = 0$, plot the output waveform.

Let $t = 0$ at the time the input wave jumps from 0 to +100 volts. Since the potential across the capacitor cannot change instantaneously, $e_C$ will remain zero for an instant and the input jump will appear across $R$, resulting in a jump of +100 volts in $e_o$ and establishing the initial current $I_o = 100/R$. The capacitor will then charge toward an ultimate +100 volts, the current decreasing exponentially with a value given by

$$i = \frac{100}{R}\epsilon^{-t/RC}$$

and the output voltage $e_o = iR$ will also fall. The variation may be plotted using values from Fig. 9-7. After four time constants = 0.2 of the period the output voltage will be essentially zero once more. At the half period the input voltage returns to zero, or falls 100 volts. The output voltage falls a like amount from zero to −100 volts. The capacitor once again starts to charge until after four time constants its potential is once more 100 volts and the current and output voltage are again zero. This is illustrated in Fig. 9-12.

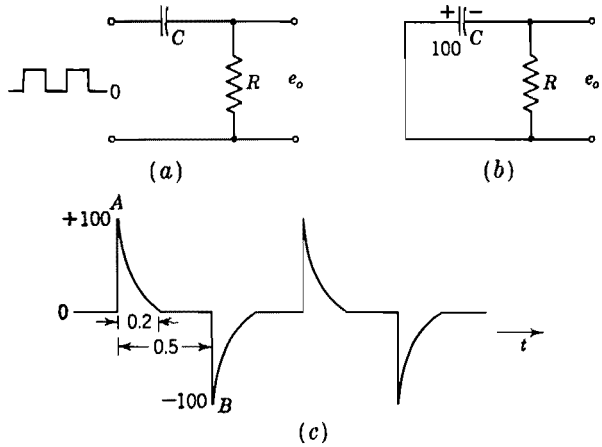



FIG. 9-12. (*a*) Initial conditions at *A*. (*b*) Conditions in circuit at *B*. (*c*) Output wave.

**9-6. Integrating Circuits.** The circuits of Fig. 9-13 may be shown to perform integration of input waveforms. Taking the *RC* circuit of Fig. 9-13*a*, with $e_C$ initially zero, the circuit equation is

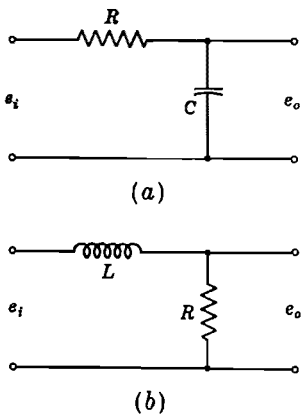



FIG. 9-13. Simple integrating circuits.

$$e_i = iR + \frac{1}{C}\int_o^t i\,dt$$

or

$$i = \frac{e_i}{R} - \frac{1}{RC}\int_o^t i\,dt \qquad (9\text{-}23)$$

If *RC is made very large* with respect to the longest period of the input signal, then the second term will be very small with respect to the first and

$$i \cong \frac{e_i}{R} \qquad (9\text{-}24)$$

Since

$$e_o = \frac{1}{C}\int_o^t i\,dt$$

then

$$e_o \cong \frac{1}{RC}\int_o^t e_i\,dt \qquad (9\text{-}25)$$

and the output voltage is proportional to the integral of the input voltage under the assumption concerning *RC*. The output of such an *RC* circuit, with long time constant, is illustrated by the oscillogram of Fig. 9-14, when driven by a square wave. The circuit provides a simple means of obtaining an approximate integral. For high accuracy the circuits of Chap. 8 should be employed.

The $RL$ circuit of Fig. 9-13*b* will also result in integration. For a large value of $L/R$ the input voltage will appear largely across $L$ so that

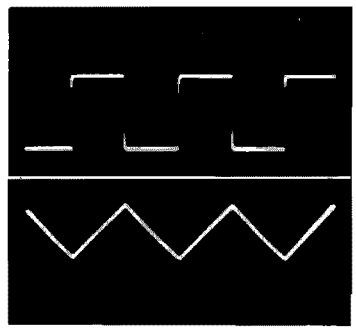
$$e_i \cong L\frac{di}{dt}$$

or
$$i \cong \frac{1}{L}\int_o^t e_i\,dt + i_o \quad (9\text{-}26)$$

If the initial inductor current is zero, then

$$e_o = iR \cong \frac{1}{L/R}\int_o^t e_i\,dt \quad (9\text{-}27)$$

and the output voltage appears as the integral of the input voltage. As for the inductive differentiator, this circuit also loses precision if the resistance of the inductor is appreciable.



(*a*)

(*b*)

FIG. 9-14. Input square wave and its integral taken by an $RC$ circuit where $RC/T = 200$.

Circuits of this nature may be used for separation of pulses having different lengths, and one such application is in the separation of the horizontal and vertical synchronizing pulses in television.

**9-7. Clipping Circuits.** A frequent application of the electron tube as a switching device is the *clipper circuit*, in which waveforms are intentionally distorted by removal, or clipping off, of a portion of the applied wave. The clipping action may be controlled as a function of potential level or polarity.

Figure 9-15*a* illustrates the essential elements of a clipper. For an assumed case the switch is on point 1 for all applied negative potentials, on point 2 for all potentials $0 < e_i < E_1$, and on point 3 for all $e_i > E_1$. As a result the output potential $e_o$ (with output terminals open) will be

| | |
|---|---|
| For $e$ negative: | $e_o = 0$ |
| $0 < e_i < E_1$: | $e_o = e_i$ |
| $e_i > E_1$: | $e_o = E_1$ |

and the resultant output characteristic is drawn in Fig. 9-15*c*. An applied sine wave of emf is shown clipped to a series of semitrapezoidal positive pulses, of amplitude fixed by $E_1$.

A diode in series with a switching potential is a suitable device to perform the action of switch $s$. As long as the anode of a diode is negative to cathode, the diode is open-circuited. When the anode becomes positive to cathode, the diode conducts with a relatively low internal resistance.

The source $e_i$ should have a high internal resistance or be supplemented by the series resistor $R$, so that the effective source resistance is made very large with respect to the diode forward resistance. Referring to

Fig. 9-15*c*, and neglecting the diode resistance, it can be seen that, for all negative values of $e_i$, diode $D_1$ short circuits the output and $e_o = 0$. For positive values of $e_i < E_1$ both diodes $D_1$ and $D_2$ are open, or have positive cathodes, and the input $e_i$ will appear at the output, assuming no load. When $e_i > E_1$, diode $D_2$ will conduct and voltage $E_1$ will appear in the output, the excess of $e_i$ over $E_1$ appearing as a drop in $R$.

Adjustment of value and polarity of $E_1$ allows clipping at any desired level, reversal of diode and battery resulting in clipping at negative levels. Several applications and their results on sine waves are shown in Figs. 9-16 and 9-17.

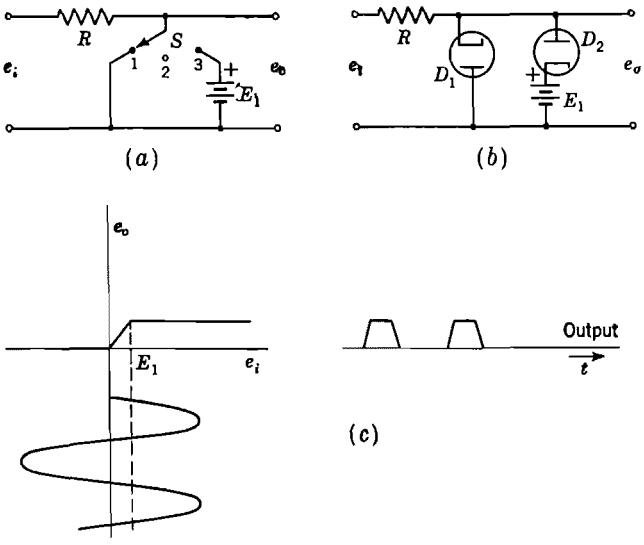



FIG. 9-15. (*a*) Basis of a clipping circuit. (*b*) Diode equivalent for (*a*). (*c*) Input-output characteristics for (*a*).

If the diode internal resistance $r_b$ is not negligible with respect to $R$, or if the excess of $e_i$ over $E_1$ is not great, then the diode voltage drop will not be small in comparison with $E_1$ and the output $e_o$ will not be flat over the clipped portion, since diode drop is a function of current. This situation appears in Fig. 9-17*b*.

Either crystal or vacuum diodes may be employed. Crystal diodes provide lower forward resistances and more perfect clipping under the conditions of the paragraph above. At the same time they have a finite back resistance, which may reduce the output voltage available during the time of diode open circuit. Vacuum-diode internal capacitance may cause trouble at high applied frequencies so that crystals may be preferred, especially in situations where the vacuum-diode cathode would be

(9-39) this is

$$T = RC \ln \frac{E}{E - e_{o,\max}} \tag{9-40}$$

The amplitude is $e_{o,\max}$. The time of the sweep is controllable by adjustment of $R$ or $C$, or both.

The output potential $e_o$ can be written from Eq. (9-39), using the series for the exponential as

$$\begin{aligned} e_o &= E\left[1 - \left(1 - \frac{t}{RC} + \frac{t^2}{2R^2C^2} - \frac{t^3}{6R^3C^3} + \cdots\right)\right] \\ &= E\left(\frac{t}{RC} - \frac{t^2}{2R^2C^2} + \frac{t^3}{6R^3C^3} - \cdots\right) \end{aligned} \tag{9-41}$$

If $t/RC$ is sufficiently small,

$$e_o \cong \frac{Et}{RC} = kt \tag{9-42}$$

and the desired linear sweep voltage is approximately obtained. It is also apparent that

$$\frac{e_o}{E} = \frac{t}{RC}$$

and if $t/RC$ is small to achieve linearity, then the ratio of output voltage to charging potential must also be small. Actually an amplifier is usually provided to furnish sufficient output from $e_o$ to deflect a cathode-ray beam.

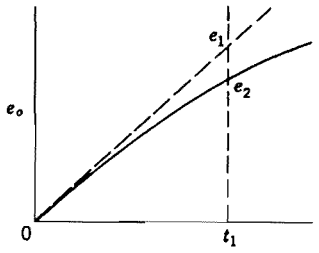



FIG. 9-32. Departure from linearity of the $RC$ charging circuit.

It is of interest to determine the magnitude of the departure from linearity occasioned by the approximation of Eq. (9-42). Taking as the ideal linear waveform

$$e_o = \frac{E}{RC} t$$

resulting from the first term of the exponential series, then the potential $e_1$ at time $t = t_1$, as indicated in Fig. 9-32, will be

$$e_1 = \frac{E}{RC} t_1 \tag{9-43}$$

and

$$\frac{e_1}{E} = \frac{t_1}{RC} \tag{9-44}$$

The potential $e_2$ on the actual charging curve at the time $t_1$ will be

$$e_2 = E(1 - \epsilon^{-t_1/RC})$$

from which

$$\frac{e_2}{E} = 1 - \epsilon^{-e_1/E} \tag{9-45}$$

The difference between potentials $e_1$ and $e_2$ is

$$e_1 - e_2 = E\left(\frac{t_1}{RC} - 1 + \epsilon^{-t_1/RC}\right)$$

$$\frac{e_1 - e_2}{e_1} = 1 - \frac{e_2}{e_1} = 1 - \frac{E}{e_1}(1 - \epsilon^{-e_1/E})$$

from which the fraction of linearity obtained in the sweep is

$$\frac{e_2}{e_1} = \frac{E}{e_1}(1 - \epsilon^{-e_1/E}) \tag{9-46}$$

as a function of the parameter $e_1/E$. This result is best expressed by the curve of Fig. 9-33, which shows that use of a ratio $e_1/E = 0.1$ gives a voltage 95 per cent linear, or having an error of 5 per cent.

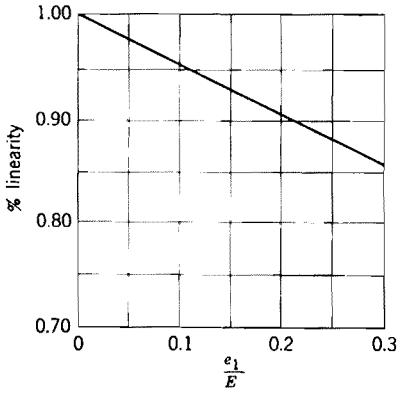



FIG. 9-33. Per cent of linearity of $RC$ sweep as a function of $e_1/E$.

**9-14. Thyratron Sweep Generators.** The $RC$ circuit with a gas-diode discharge device lacks stability of frequency and amplitude because of uncertainties in the breakdown potential of the diode. This makes a stable time plot difficult to obtain on the cathode-ray screen. Use of a thyratron as the discharge control still employs the $RC$ linearity approximation but permits greater flexibility and synchronization of the sweep frequency with the input signal. Such a thyratron-controlled sweep generator is shown in Fig. 9-34, wherein $R$ and $C$ are as before and $R_g$ is a current-limiting grid resistor. The resistor $R$ is shown variable as a means of frequency control.

For a given value of grid bias $E_{cc}$ and anode voltage $e_o$ the thyratron will break down, ionize, and discharge the capacitor $C$. If $E_{cc}$ is made more negative, a new and higher value of breakdown potential is arrived at. If a sine-wave synchronizing signal is introduced in the grid circuit, swinging the bias alternately more positive and more negative, the breakdown or firing potential will likewise swing, as indicated in Fig. 9-35. The natural frequency of the generator is adjusted slightly lower than that of the synchronizing signal, and without the synchronizing sine wave present the thyratron would fire at point $A$ when $e_o$ reaches the critical

breakdown value. On the next cycle the firing point would be $C$, indicating no fixed relation to the desired signal frequency. With the synchronizing signal present the firing-potential curve becomes the indicated sinusoid, and the firing point occurs at $B$ in the first cycle and $D$ in the

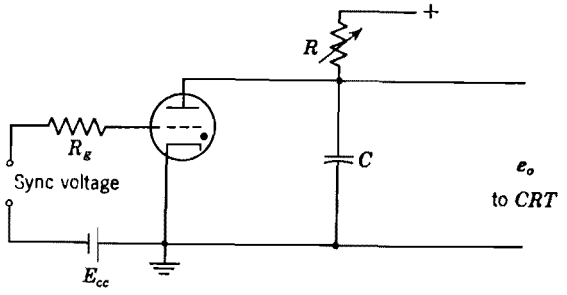



FIG. 9-34. Thyratron-controlled saw-tooth-sweep generator.

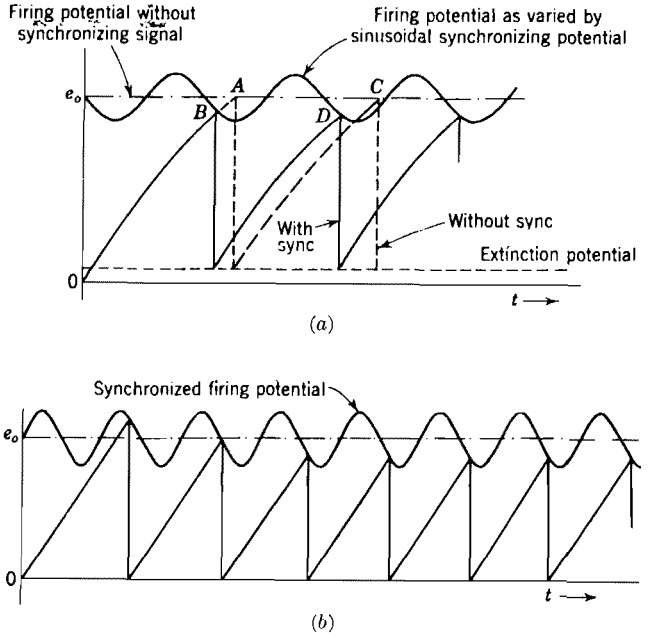



FIG. 9-35. (*a*) Synchronizing a thyratron saw-tooth generator. (*b*) The transient on starting.

second. The period of the saw tooth is reduced, and the firing point will occur at the same point in each cycle, or the generator is locked to the frequency of the injected signal. Figure 9-35*b* illustrates the transient in the generator period during its starting operation. Further discussion of synchronization appears in Sec. 9-21.

The thyratron saw-tooth generator may also be locked or synchronized to a multiple or submultiple of the synchronizing frequency. All thyratron sweep generators are limited to relatively low frequencies by the time required for deionization of the gas between sweep operations. Such a limit usually appears at about 30,000 to 50,000 cycles.

**9-15. Triggered Vacuum-tube Sweep Generator.** In some applications it is not desired that the sweep be synchronized to a repetitive signal, but that a sweep wave be produced at the time of incidence of another event. The vacuum tube used as a switch lends itself to such

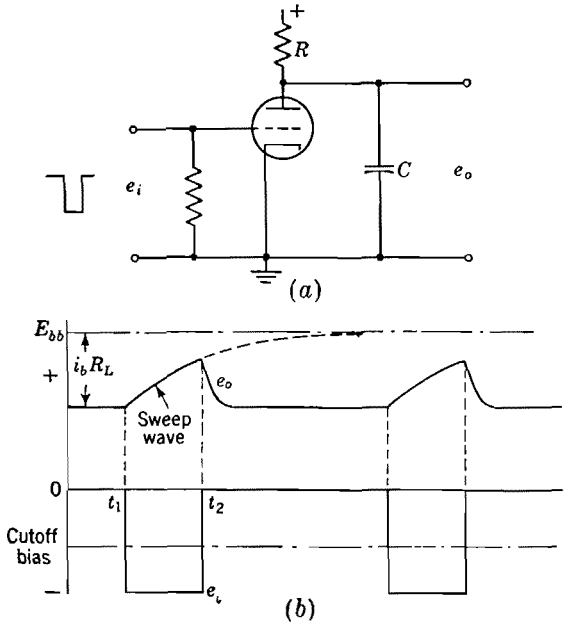



FIG. 9-36. Vacuum-tube sweep using an $RC$ charging circuit.

an application and also provides a source of higher-frequency sweep voltages, with very accurate timing of the start of the sweep wave.

In the simple $RC$ circuit of Fig. 9-36 the grid is normally at zero potential when no input signal is applied. The potential across $C$ is then equal to $E_{bb} - i_bR_L = e_b = e_o$, the plate voltage across the tube. With the application of an input rectangular pulse of desired duration the grid is driven beyond cutoff, the capacitor $C$ charges through $R_L$, and the voltage $e_o$ across $C$ begins to increase toward $E_{bb}$. This increase continues during the length of the pulse.

At the end of the pulse the grid is once more at zero, the tube is conducting, capacitor $C$ quickly discharges through the low-resistance path provided by the tube, and the circuit returns to the initial condition.

Thus a sweep waveform accurate as to starting time and duration is generated.

**Example.** Design the plate circuit for a vacuum-tube sweep circuit, using the 6C4 triode, to give a sweep period of 100 $\mu$sec, and with nonlinearity of 2 per cent.

For the nonlinearity limitation Figure 9-33 shows $e_i/E = 0.05$. If $E_{bb} = 250$ volts and $R$ is assumed as 50,000 ohms, then the tube curves show a zero-bias plate current of 4.3 ma and a plate voltage of $250 - 0.0043 \times 50{,}000 = 35$ volts. The plate voltage then changes from 35 to 250 volts, and the potential available for charging $C$ is $250 - 35 = 215$ volts. Then from Eq. (9-45), interpreting the $E$ value as 215,

$$e_2 = 215(1 - \epsilon^{-0.05}) = 20.4 \text{ volts}$$

as the peak sweep output at time $t_2$. Assuming $e_1 = e_2$, which is reasonable for small nonlinearity, Eq. (9-43) gives

$$\frac{e_1}{E} = \frac{t_1}{RC} = 0.05$$

$$RC = \frac{100 \times 10^{-6}}{0.05} = 0.002 \text{ sec}$$

and since $R = 50{,}000$ ohms, then

$$C = \frac{0.002}{50{,}000} = 0.04 \times 10^{-6} = 0.04 \ \mu\text{f}$$

During the zero grid-bias interval the value of $r_b$ of the tube is

$$r_b = \frac{35}{0.0043} = 8100 \text{ ohms}$$

The capacitor will discharge back down to 35 volts through the tube resistance in parallel with $R$, or approximately 7000 ohms. The $r_bC$ time constant is

$$r_bC = 7000 \times 0.04 \times 10^{-6} = 0.00028 \text{ sec}$$

The complete discharge transient will be approximated in four time constants, or

$$T = 4 \times 0.00028 = 720 \ \mu\text{sec}$$

This indicates a top limit for the pulse-repetition rate of about 1000 pulses per second, unless the design is further modified to decrease the $r_bC$ time constant.

**9-16. Linearization of the $RC$ Sweep.** The potential across the capacitor of an $RC$ sweep circuit is given by

$$e_C = \frac{1}{C} \int_o^t i \, dt$$

and if a constant current be supplied for charging, rather than a constant potential,

$$e_C = \frac{I_o}{C} t + V_o$$

and the desired voltage will be obtained as linear with time.

Since a pentode approximates a constant-current device, or the anode current of a pentode is nearly independent of anode potential with fixed screen and control-grid potentials, then a pentode may serve as a constant-current source. Figure 9-37 illustrates pentodes supplying gas-triode and vacuum-triode forms of sweep generators.

With careful adjustment of the pentode operating point, it is possible to obtain quite linear sweep potentials. Some difficulty in circuit design may be encountered, however, because of the necessity for various pentode supply potentials to operate above ground reference.

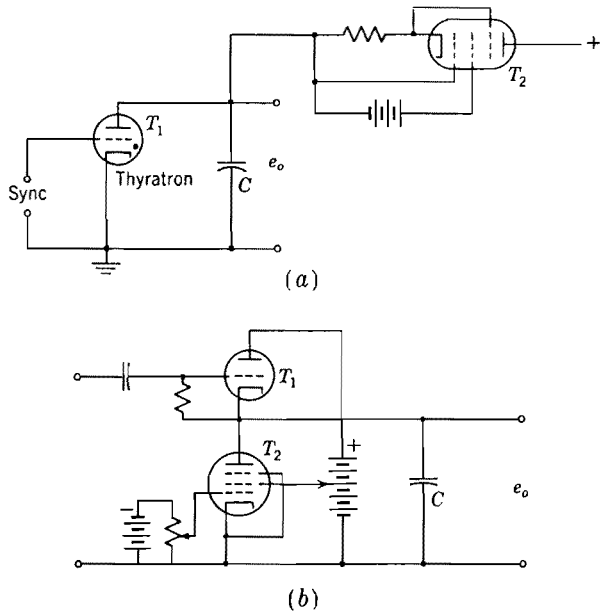



FIG. 9-37. (*a*) Constant-current supply with a gas triode for switching. (*b*) Constant-current pentode in vacuum-tube sweep.

Another circuit known as the *bootstrap sweep* uses a form of feedback to supply a constant charging current for linearizing the $RC$ sweep voltage and is shown in Fig. 9-38*a*. Therein tube $T_1$ is driven by the rectangular switching pulse, or *gate*, so called because it turns on, or operates, the circuit. In the output of $T_1$ is capacitor $C_1$, charged through $R_1$ and $R_2$ in the plate circuit, the output $e_o$ being the potential across $C_1$. A cathode follower $T_2$ provides feedback from the output to the plate circuit of $T_1$, thus giving a linearizing current feedback.

The circuit action can be studied by use of Fig. 9-38*b*, which represents the charging circuit of tube $T_1$, with the three currents indicated to the node between $R_1$ and $R_2$. Writing emf equations to this node

with $s$ closed gives

$$e_1 = E - i_1 R_2$$

$$e_1 = i_3 R_1 + \frac{1}{C}\int_o^{t_1} i_3\, dt \tag{9-47}$$

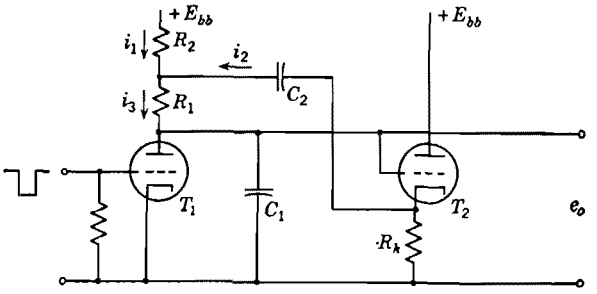
Equating,

$$E - i_1 R_2 = i_3 R_1 + \frac{1}{C}\int_o^{t_1} i_3\, dt \tag{9-48}$$

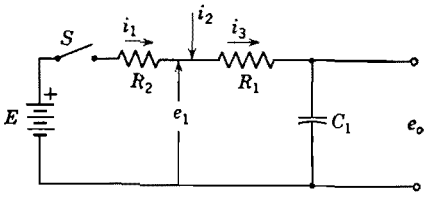
At $t = 0$ it can be seen that the value of $i_3$ must be

$$i_3 = \frac{E}{R_1 + R_2} \qquad t = 0 \tag{9-49}$$

assuming $i_2 = 0$ at that time, a condition which will later be found physically correct. However, if $i_3$ is to be constant, or constant-current




(a)




(b)

FIG. 9-38. (*a*) "Bootstrap" sweep employing feedback through $C_2$ for linearization. (*b*) Equivalent $RC$ charging circuit.

charging of $C_1$ is to be achieved, then $i_3$ must have the value of Eq. (9-49) at $t = 0$, and at all other times as well, during the switching cycle.

Noting that

$$i_1 = i_3 - i_2$$

and using this in Eq. (9-48) gives

$$i_3(R_1 + R_2) - i_2 R_2 = E - \frac{1}{C}\int_o^t i_3\, dt$$

Using the condition of Eq. (9-49) to force constant-current charging, then

$$i_2 R_2 = \frac{1}{C}\int_0^t i_3\, dt$$

$$i_2 = \frac{1}{R_2} e_o \tag{9-50}$$

Thus the current $i_2$ fed back to the plate circuit must be proportional to the output potential $e_o$ across $C_1$.

In the circuit of Fig. 9-38*a* tube $T_1$ is normally conducting but is cut off by the application of the negative gate pulse, at which time capacitor $C_1$ begins to charge toward the potential $E_{bb}$, in exactly similar manner to that discussed in Sec. 9-15. With the cathode follower $T_2$ connected the potential on $R_k$ rises, and if the gain of the cathode follower were unity, the potential $e_o$ would be applied across $R_2$, as required by Eq. (9-50) for constant-current charging of $C_1$.

In effect the circuit acts such that capacitor $C_1$, instead of charging toward a potential $E_{bb}$, charges toward a constantly increasing potential, $E_{bb} + (1/C)\int i\, dt$, the increase in potential compensating for the $C_1$ voltage, thereby placing a constant potential across $R_1$ and causing a constant current therein.

The extent to which the resulting saw-tooth wave is linear is dependent on how closely the gain of the cathode follower approaches unity. It is also assumed that $C_2/C_1$ is large, so that the potential of $C_2$ remains constant during a sweep, and changes in $e_o$ are passed entirely to $R_1$. Since the gain of $T_2$ may be 0.95 or more, the nonlinearity error may be reduced by a factor of 2 or 3 over the simple $RC$ exponential sweep.

It is, of course, possible to provide amplification inside the current-feedback loop, so that almost any degree of compensation for nonlinearity is possible. Other variations of this basic circuit are also employed.

**9-17. Other Forms of Time Bases.** While the linear time base is the most common form of presentation of information on a cathode-ray screen, there are other forms of occasional advantage. One is the circular sweep of Fig. 9-39*a*, with signals appearing as radial "pips," or changes in the circle radius. Sine waves are employed as sweep voltages, with the two voltages written as

$$E_v = A \sin \omega t$$

$$E_h = A \sin\left(\omega t - \frac{\pi}{2}\right)$$

provided through an appropriate phase shifter and amplifiers. A signal may be applied to the screen of pentode tubes in these amplifiers to increase gain and radius on an instantaneous basis, thus forming a pip.

A sweep axis three times as long as that possible on a linear basis is the result.

Figure 9-39*d* is a simple circuit for supplying potentials for a circular sweep. The potentiometer controls the size of circle, and the variable resistor $R$ controls the degree of ellipticity. A waveform free of harmonics is required for a good circle, since the variation of $R$ is to provide a point where $R = 1/\omega C$, and this can be true for only one frequency.

In another form of sweep for radar the beam is caused to start at the center and move radially outward, and then the radial sweep is rotated or scanned around the center of the screen in synchronism with the rotation of a radar antenna around its base. The signal circuits trigger

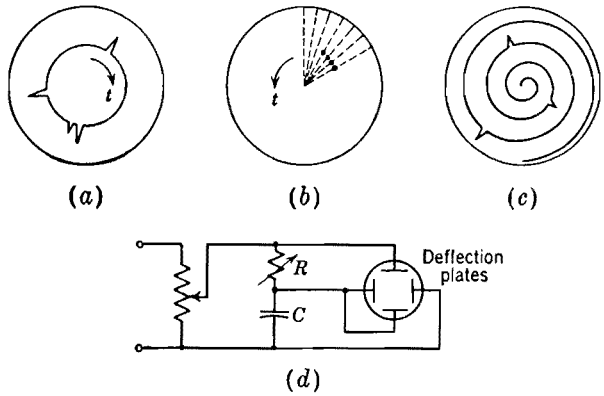



FIG. 9-39. (*a*) Circular-sweep presentation. (*b*) PPI indication. (*c*) Spiral sweep. (*d*) Circuit for simple circular sweep.

the sweep for each radial, and distance to an echo-creating point is indicated as radial distance on the screen. An echo from a target is received and indicated by increase of cathode-ray-tube grid voltage, intensifying the beam and causing a bright spot to appear on the screen. A radar on a ship near a coast line will present a line of bright spots indicating signal reflections from the coast and providing a map on the screen; thus the name plan-position indicator (PPI). The sweep voltage for this system is usually derived from a rotating wire-wound potentiometer of desired nonlinear relation.

The spiral sweep is a modified and lengthened form of circular sweep, for greater precision. Its deflecting voltages are obtained as for the PPI indication, but the speed of rotation is made much greater than the time of scanning of one radial line, thus widening that radial into a spiral.

**9-18. Sweep Circuits for Magnetically Deflected Cathode-ray Tubes.** Since flux density $B$ is proportional to $i$ in air-cored deflecting coils, then a current linear with time is needed for a linear sweep in a magnetically

deflected cathode-ray tube, and this current will vary as

$$i = kt$$

It is then necessary to determine the applied voltage waveform that will cause a current to vary as $kt$ in a deflection-coil system, having resistance, inductance, and distributed capacity. To simplify analysis, the latter parameter will be neglected here as small.

The circuit of Fig. 9-40 employs a small portion of an $RL$ transient current as a first approximation to linearity. The circuit is driven by a

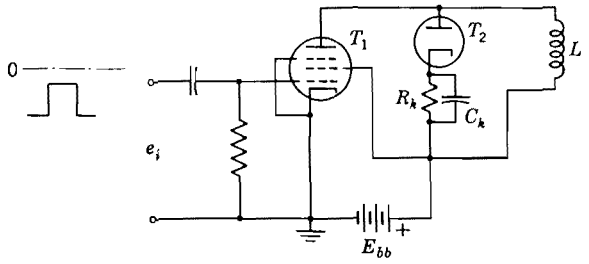



FIG. 9-40. Magnetic sweep generator.

positive-going rectangular pulse, which drives the grid to near zero during the sweep interval. At the instant the grid is driven up from cutoff, the plate current is zero but it will rise as

$$i = \frac{E_{bb}}{R + r_b}(1 - \epsilon^{-(R+r_b)t/L}) \tag{9-51}$$

where $R$ and $L$ are the resistance and inductance of the deflection coil. If the time constant $L/(R + r_b)$ be made large with respect to the time of the rectangular pulse, then the sweep current will approach linearity and the linearity-error information previously developed for the $RC$ circuit is applicable, using equivalent fractions of time constant as sweep time.

During the conduction interval the polarity of the $L\ di/dt$ voltage existing across the sweep coil is in such direction that the damper diode $T_2$ is nonconducting. When $T_1$ returns to cutoff condition, the $L\ di/dt$ voltage reverses polarity and $T_2$ conducts, allowing the energy stored in $L$ to be dissipated in the diode and $R_k$. Owing to distributed capacity in the deflecting coil it is possible for the circuit to become oscillatory, and $R_k$ is usually chosen small enough to damp the discharge transient quickly to zero.

It is also possible to find a voltage waveform which, when applied to the inductive load, will yield a coil current which is linear with time. For a linear current having the form $i = Kt$ the deflection coil, of induct-

ance $L$ and resistance $R$, will have an applied voltage

$$\begin{aligned} e &= L\frac{di}{dt} + Ri \\ &= LK + RKt \end{aligned} \tag{9-52}$$

and this will be of the form shown in Fig. 9-41*b*.

The circuit of Fig. 9-42 will generate such a wave as input to $T_2$, through the action of $R_2$, $C_2$ on the rectangular positive pulse applied

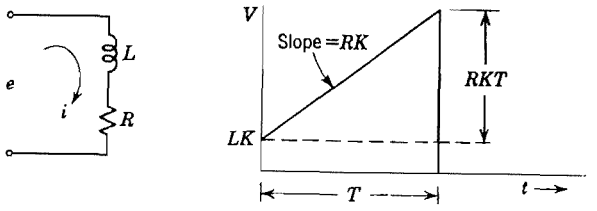



FIG. 9-41. (*a*) Deflection-coil circuit. (*b*) Voltage wave required to produce a saw-tooth current in an $RL$ circuit.

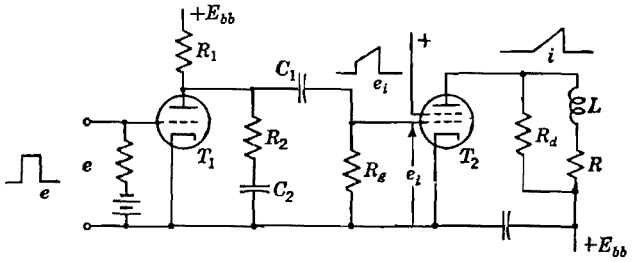



FIG. 9-42. Generation of a linear saw-tooth current.

to $T_1$, and will then apply this waveform to the Class A amplifier $T_2$. The voltage $e_i$ must supply the excitation indicated by Eq. (9-52), so that

$$AE_{bb}\left(1 - \frac{R_1}{R_1 + R_2}\epsilon^{-t/(R_1+R_2)C_2}\right) = LK + RKt \tag{9-53}$$

where $A$ is the gain of $T_2$. Hence the initial jump of the wave of Fig. 9-41*b* must be

$$AE_{bb}\left(1 - \frac{R_1}{R_1 + R_2}\right) = LK \tag{9-54}$$

and the initial slope must be

$$\frac{AE_{bb}R_1}{(R_1 + R_2)^2C_2} = RK \tag{9-55}$$

A waveform of $e_i$ produced by such a circuit is shown in Fig. 9-43.

If the sweep-coil distributed capacity is appreciable, it is necessary to add a narrow spike of voltage at the start of $e_i$ to aid in supplying initial charge to the distributed capacitance. Other forms of magnetic sweep circuit are also possible.

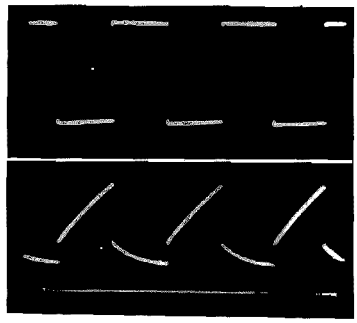



FIG. 9-43. (*a*) Square-wave input. (*b*) Trapezoidal voltage $e_i$ of Fig. 9-42.

**9-19. The Free-running Multivibrator.** The term *multivibrator* is generally employed for the circuit of Fig. 9-44*a*, developed by Abraham and Bloch in 1918, because of the abundance of harmonics present in the output waveform. Its major applications make use of this feature to provide pulses or square, trapezoidal, or saw-tooth waveforms. Variations in circuit are possible which lead to three forms of operation:

1. A circuit having no stable conditions of operation, which yields a *free-running oscillator.*

2. A circuit having one stable operating point, but which can be driven into a region of instability, eventually returning to the initial stable state. This type of multivibrator is referred to as a *one-shot* or *monostable* form.

3. A circuit having two stable operating points with an unstable region between, such that the circuit may be driven alternately from one to the other stable condition. This is known as a *trigger circuit.*

The three types are illustrated in Fig. 9-44. The free-running multivibrator will be discussed here, the other types being covered in later sections. Consider Fig. 9-44*a*, with plate currents momentarily equal, a condition of unstable equilibrium. Some small disturbance may cause the plate current of $T_1$ to increase momentarily, lowering the potential of point $A$. Since $C_1$ and $C_2$ are large, their potentials cannot change rapidly, so that the drop in potential of $A$ is coupled through $C_2$ to the grid of $T_2$, reducing the current in $T_2$. This raises the potential of point $B$ and, by coupling through $C_1$, raises the grid potential of $T_1$ and further increases its plate current. The effect is cumulative, the current in $T_1$ reaching a maximum, and as point $A$ goes down in potential, the grid of $T_2$ is driven so far negatively, almost instantaneously, as to cut off $T_2$.

The charges on $C_1$ and $C_2$ then leak away through $r_c$ and $R_g$, and at some later time the grid potential of $T_2$ rises above cutoff, and $T_2$ conducts. This causes a negative signal to reach the grid of $T_1$ as point $B$ falls. The current in $T_1$ now decreases, starting the cumulative chain again such that $T_1$ becomes cut off and $T_2$ conducting, after which the cycle repeats.

There are two unstable limiting conditions, in one of which $T_1$ is cut off

and $T_2$ conducting, and in the other the situation is reversed. The triggering action occurs at the point of plate-current equality, and this condition is thus also one of instability.

Since grid current flows in the tube having a positive grid, and its $r_c$ resistance may then approximate 1000 ohms, the associated capacity discharges more quickly than the capacity in the circuit of the negative-grid tube. The latter tube is thus the one that determines the delay time, its associated capacitor having to discharge through $R_g$, and the time constant determining the delay time between triggering actions is

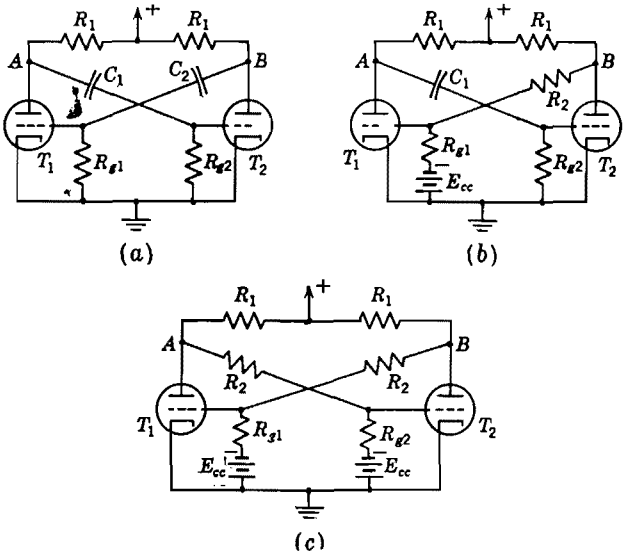



FIG. 9-44. (*a*) Free-running multivibrator. (*b*) One-shot multivibrator. (*c*) Trigger circuit.

then $R_{g1}C_1$ or $R_{g2}C_2$, assuming $R_g \gg R_1$. If these time constants are unequal, then unequal delays, or pulses, will result.

Since $r_c$ will be small with respect to both $R_g$ and $R_1$, then the time of charge of $C_1$ or $C_2$ is fixed by the $R_1C_1$ time constant, usually much smaller than that dependent on $R_g$.

The grid- and plate-voltage waveforms are illustrated in Figs. 9-45 and 9-46, which represent theoretical and actual cases. Attention may be directed to the small positive pips on the grid-voltage waves, these occurring during the time that $C_1$ or $C_2$ charges through the $r_c$, $R_1$ path. The long time constant of the discharge through $R_g$ is also shown as limiting the pulse delay, the grid voltage rising exponentially until the triggering grid voltage is reached at the cutoff level.

The cutoff potential $V_o$, at which the rapid rise in grid potential begins,

is given by

$$V_o = (E_{bb} - e_b)\epsilon^{-t_1/R_{g1}C_1} \tag{9-56}$$

where $e_b$ is the anode potential of the conducting tube and $E_{bb}$ is the supply potential to which $C_1$ was charged. Then $E_{bb} - e_b$ is the initial potential acting to discharge the capacitor $C_1$ through the circuit composed of $R_1$, $C_1$, and $R_{g1}$, as shown in Fig. 9-47.

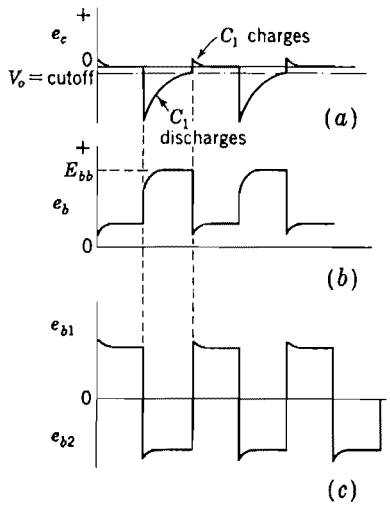



FIG. 9-45. (*a*) Grid-voltage waveform $T_1$. (*b*) Plate-voltage waveform $T_1$. (*c*) Plate-to-plate voltage.

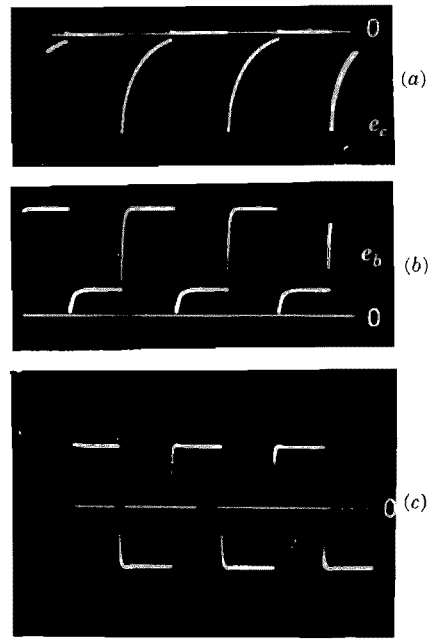



FIG. 9-46. (*a*) Grid voltage $T_1$. (*b*) Plate voltage $T_1$. (*c*) Plate-to-plate voltage.

If the circuit is symmetrical, or the pulses are equal, then the time $T$, or the period of oscillation of the multivibrator, is

$$T = 2t_1 = 2R_gC_1 \ln \frac{E_{bb} - e_b}{V_o} \tag{9-57}$$

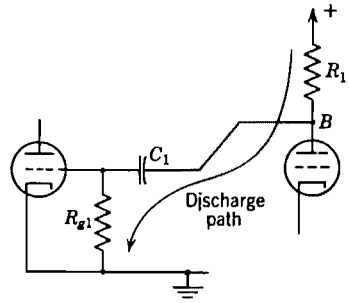
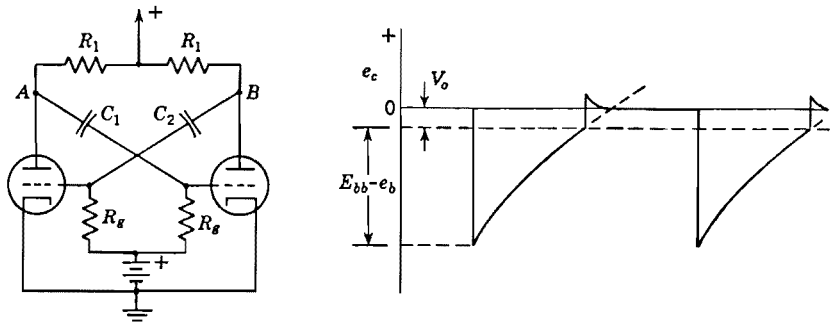



FIG. 9-47. Path for discharge of $C_1$.

This expression does not include the effect of stray and tube capacities.

The multivibrator is an excellent source of square waves, the plate-to-plate wave of Fig. 9-46*c* needing only clipping to improve the form. If an extremely rapid rise of voltage is needed, then careful circuit design in terms of low plate resistance, low-capacitance tubes, and broad-band amplifier techniques must be used. Rise times of fractions of microseconds may be obtained.

**9-20. Effect of Bias.** The time at which the trailing edge of a multivibrator output pulse is formed is fixed by the intersection of the exponentially rising grid-voltage line and the cutoff line of Fig. 9-45*a*. However, the cutoff voltage is dependent on tube constants and voltages, and the time of trailing-edge formation and length of the pulse may not be precise. With the exponential grid voltage rising toward a low value of




(*a*) (*b*)

FIG. 9-48. (*a*) Biased multivibrator. (*b*) Grid waveform for (*a*).

potential, the slope of the exponential curve at time of reaching cutoff may be quite flat and the triggering time therefore somewhat indefinite.

To assure a rapidly rising exponential voltage at time of reaching cutoff, the positive-biased circuit of Fig. 9-48*a* may be employed, leading to grid-voltage waveforms as in Fig. 9-48*b* or 9-49. The steeper angle at which the grid-voltage wave crosses the cutoff line gives greater precision in the triggering time. Even with this precaution stability of frequency is not great.

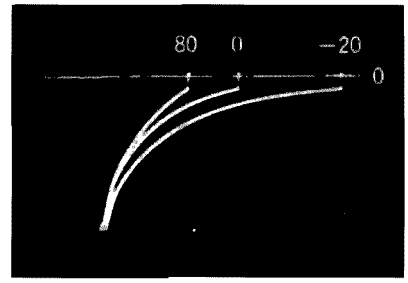



FIG. 9-49. Multivibrator grid-voltage waves for biases of +80, 0, and −20 volts.

**9-21. Synchronization of the Multivibrator.** Further to improve the frequency stability of the free-running multivibrator, the system may be driven or synchronized with an external pulse or sinusoidal voltage. This forces the period of the multivibrator to be exactly the same as a multiple or a submultiple of the synchronizing frequency.

One method of introducing a synchronizing signal into a multivibrator is shown in Fig. 9-50*a*. Other methods are possible, the only requirement being that the input be symmetrical to both tubes. If positive pulses are introduced on the grids, they obviously have no effect on the

conducting tube whose grid is already positive. The pulses add to the potential on the negative but rising grid, producing pips along the rising wave as in Fig. 9-50*b*. If these pips have sufficient amplitude, then some pip adding to the exponential voltage will cause the cutoff voltage to be exceeded sooner than would be the case without the synchronizing signal. The multivibrator will then lock onto the latter signal, triggering each time on arrival of the same relative pulse. Even if the driving signal changes frequency over a small range, the multivibrator will likewise shift frequency and stay locked. If the signal drift is greater, then the circuit may be triggered by the earlier or a following pulse, and the circuit will trigger every $n - 1$ or $n + 1$ pulses, instead of every $n$ pulses.

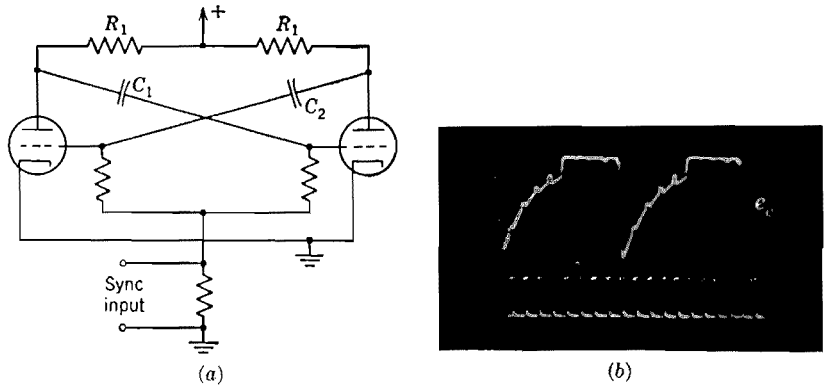



FIG. 9-50. (*a*) One input method for synchronization. (*b*) Grid voltage of pulse-synchronized multivibrator.

*Frequency division* is usually the object of such synchronization, the multivibrator being caused to trigger on every $n$th pulse, the factor $n = 10$ being convenient. The frequency of a highly stable 100-kc oscillator is frequently subdivided in this way, giving accurate outputs of 10 kc, 1 kc, and finally 100 cycles. This last frequency may be made to drive a clock whose time may be compared with the rotation of the earth, the whole then becoming a very accurate *primary standard* of frequency and time.

**9-22. The One-shot Multivibrator.** The *one-shot multivibrator* operates with one tube normally biased beyond cutoff, so that the circuit is stable and inactive until a synchronizing or triggering pulse is applied. The pulse causes the circuit to change from its stable limiting condition to an unstable limiting condition, and, after an interval of time determined by the circuit constants, the circuit returns to its initial stable state, where it remains until another triggering pulse arrives.

Such a circuit is shown in Fig. 9-51*a*, as a modification of the free-running multivibrator obtained by adding resistor $R_2$ and greater than cutoff bias for $T_1$. In this circuit the stable limiting condition finds $T_1$

cut off and $T_2$ conducting. Application of a positive-going pulse to the grid of $T_1$, of sufficient magnitude to make $T_1$ conduct, abruptly lowers the potential of point $X$. Because the potential of $C_1$ cannot change abruptly, point $Z$ or the grid of $T_2$ follows point $X$ down in potential

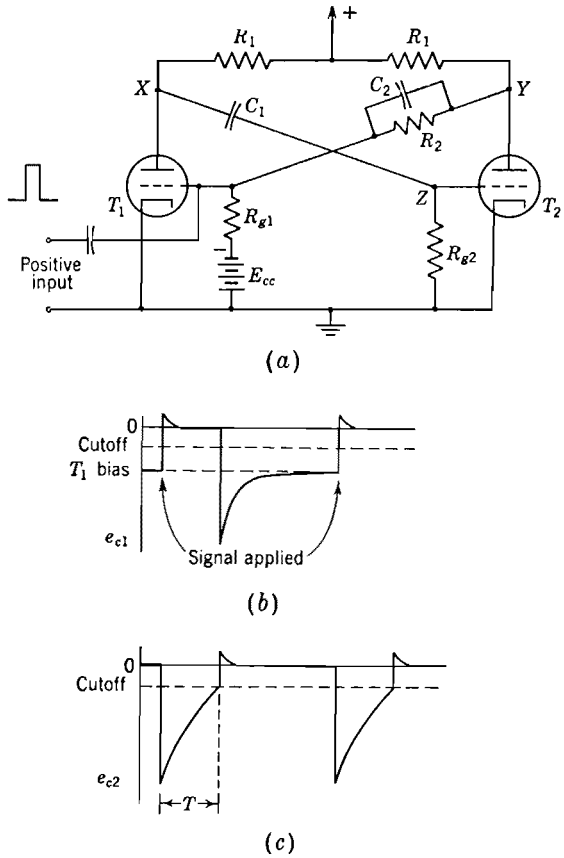



FIG. 9-51. (*a*) One-shot multivibrator. (*b*) Grid voltage on biased tube $T_1$. (*c*) Grid voltage on unbiased tube $T_2$.

to a value below cutoff. This is indicated in Fig. 9-51*c*. As $C_1$ discharges through $R_{g2}$, the grid potential of $T_2$ rises until at cutoff value current again flows in $T_2$, lowering the potential of point $Y$ and forcing down the grid potential of $T_1$. The action is then cumulative and rapid, $T_2$ going to full conduction and $T_1$ being cut off as in the initial state. The length of time for the triggering operation is determined by the $R_{g2}$, $C_1$ time constant.

By analogy with the discussion of Sec. 9-19 the time duration of the

triggering operation is given by

$$T = R_{g2}C_1 \ln \frac{E_{bb} - e_b}{V_o} \tag{9-58}$$

The grid resistor of $T_2$ may be returned to a positive bias, as with the free-running circuit, giving somewhat more precise operation, due to the increased slope of the rising exponential grid voltage at the instant of reaching the cutoff value. Variation of this positive bias also becomes a simple means of changing the pulse or gate width, the time $T$ being proportional to grid bias over a wide range. A circuit modification to provide this feature is shown in Fig. 9-52.

Capacitor $C_2$ serves to decrease the time of switching conduction from $T_2$ back to $T_1$. Referring to Fig. 9-51*a*, the voltage of point $Y$ reaches

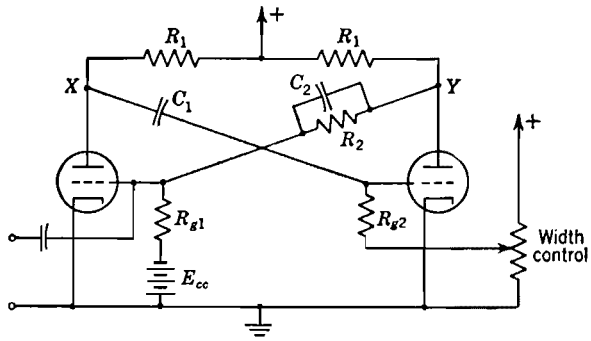



FIG. 9-52. Pulse-width control.

the grid of $T_1$ only through a voltage divider formed of $R_2$ and $R_{g1}$; thus only a fraction of the d-c voltage at $Y$ appears on the $T_1$ grid. The capacitor, however, bypasses $R_2$ for rapidly changing voltages, and, at the instant of switching back to $T_1$, the whole potential at $Y$ is transferred to the $T_1$ grid. Also, without $C_2$ the internal tube capacities, shunted across $R_g$, would have to be charged through $R_2$. With $C_2$ appreciably larger than the tube capacities, it is possible quickly to charge the internal capacities and make the potential of the grid of $T_1$ closely follow that of point $Y$ during switching.

The circuit is quite generally employed to generate short rectangular pulses, with the output being taken from the plate of $T_2$. It may also be used to reshape pulses, the input pulses of distorted form being used to trigger the circuit, giving uniform output pulses to actuate controls or pulse counters.

**9-23. The Trigger, or Flip-Flop, Circuit.** The *trigger* circuit of Fig. 9-53, first described by Eccles and Jordan in 1919, is a further development of the multivibrator circuit, having two stable conditions, separated

by an unstable region. One stable condition is with $T_1$ conducting and $T_2$ cut off, the other stable condition representing a reverse condition. Pulses or signal inputs are required to cause the circuit to shift from one condition to the other, and the potentials at $A$ and $B$ may be utilized to control appropriate switching operations.

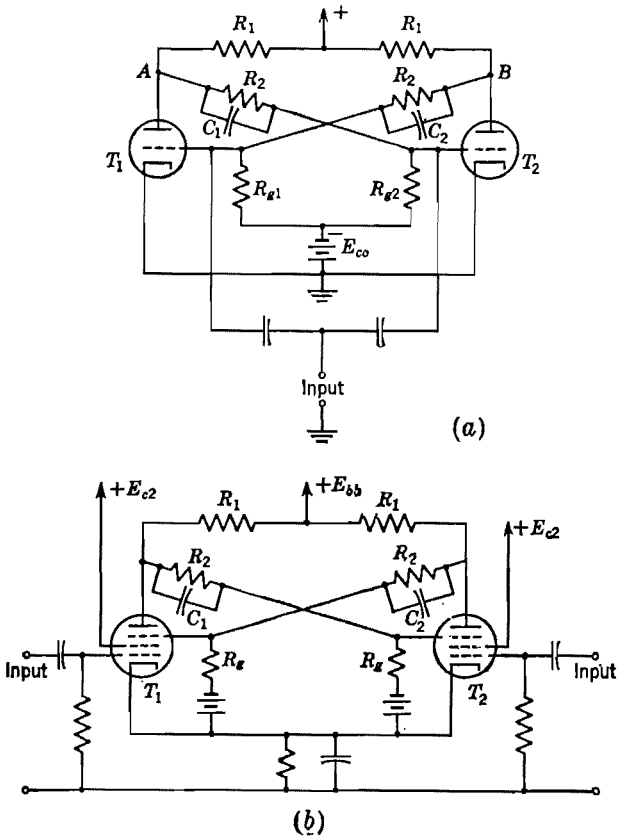



FIG. 9-53. (*a*) Eccles-Jordan trigger circuit. (*b*) Pentode trigger circuit.

Assume that $T_1$ is initially conducting and $T_2$ is cut off by the combined action of $R_2$, $R_{g2}$ as a voltage divider and the negative bias. If a positive pulse be applied to both grids simultaneously, there will be no reaction in $T_1$, already conducting, but $T_2$ will begin conduction. This lowers the potential at $B$ and reduces the grid potential on $T_1$, which reduces the tube current, raising the potential at $A$ and the grid potential of $T_2$. A regenerative action thus follows until $T_1$ is cut off and $T_2$ is fully conducting, in a manner similar to that already discussed for the other multivibrator circuits. Since no $RC$ transients are involved, the

circuit will remain in the last stable condition until another positive pulse is applied, at which instant it returns to the original condition.

Negative trigger pulses may also be used, with no reaction in the cutoff tube, but reducing the current in the conducting tube and starting the regenerative action. The only requirement on such input pulses is that they be of sufficient magnitude and duration to drive the tube to a condition slightly past the equal-current point, after which the cumulative action completes the current transfer.

Input pulses may be applied to the common plate connection as well as to the common grid connection, negative pulses being used. As the anode potentials go down, the grid potentials follow through the voltage-divider action and thus the action is much as described above.

Unsymmetrical triggering, in which a negative pulse is applied to the "on" tube or a positive pulse to the "off" tube, may be used, but this then requires alternation of the input-pulse polarities as the tubes switch. In addition the magnitude of the positive pulse needed for triggering will be greater than that of a negative pulse, since the negative pulse need merely drive the grid down slightly beyond the current-equality point, whereas the positive pulse will have to overcome a considerable negative bias before conduction even starts.

Capacitors $C_1$ and $C_2$ are added to aid the circuit operation on short pulses, as discussed for the one-shot circuit. Pentodes may also be employed for trigger operation, the suppressor grids replacing the control grids in the circuit of Fig. 9-53*b*, in so far as cross coupling is concerned. The normal control grids are then available for signal-pulse triggering by short negative pulses, which drive both tube currents to zero. However, the potential across $C_1$ or $C_2$ of the tube that has been conducting makes the suppressor of that tube more negative than that of the other tube, and the circuit triggers when the input pulse is removed. This requires that $C_1$ and $C_2$ be large enough so that their potentials do not appreciably change during the pulse interval.

The time that it takes to transfer conduction from $T_1$ to $T_2$, or vice versa, is called the *resolving time*. During this interval the circuit is insensitive to the arrival of a second pulse, and thus the resolving time is the minimum interval between pulses which can be resolved, or both pulses indicated. The time constants of $R_1C_1$ and $R_2C_2$ should be large with respect to the time of an individual pulse, but short with respect to the expected time between pulses, in order to obtain good resolution.

**9-24. Scaling Circuits.** It is apparent that, in the circuit of Fig. 9-53*a*, tube $T_2$ conducts once for every two input pulses, and the number of negative pulses available from its anode is one-half the number of triggering pulses. This gives it the name of a *scale-of-2* circuit, and these are extensively used in scaling down the number of pulses produced in

experiments with radioactivity and other areas of nuclear physics. It is possible to connect such circuits in cascade, so that *scale-of-4*, *scale-of-8*, and other ratios may be produced. Thus a high initial pulse rate may be counted on a low-speed mechanical counter. The time between pulses need not be constant, the circuit simply waiting in its last condition until the next pulse arrives.

It is, of course, necessary to know which tube of a given pair is conducting, and this can most easily be determined by connection of a small neon-gas diode across the plate-load resistor $R_1$ of one tube. If this diode is not glowing, it indicates that the *other* tube of a pair is conducting,

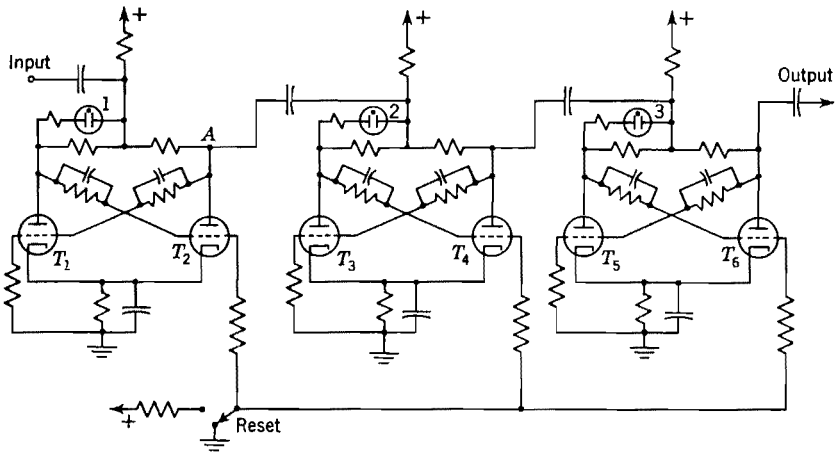



FIG. 9-54. Scale-of-8 circuit, including plate coupling and reset.

and only one diode per trigger pair is needed. These connections are shown in Fig. 9-54.

If five flip-flops are used, they will count down by a factor of $2^5 = 32$, and in general a scaling circuit will count down by a factor of $2^n$, where $n$ indicates the number of flip-flops. The circuit of Fig. 9-54 illustrates a scale-of-8 counter with coupling of the output pulse from the plates of $T_2$ and $T_4$ to the common anode lead of the following pair. Arrangements are also shown through a *reset* switch, to ensure that $T_2$, $T_4$, and $T_6$ are all initially conducting, so that the starting condition of each trigger pair is known and all indicator diodes are off.

Then the first negative input pulse transfers conduction from $T_2$ back to $T_1$, but since point $A$ rises in potential, the second pair is not triggered, a negative pulse being required for anode triggering. However, indicator 1 is turned on. The second input pulse transfers conduction from $T_1$ back to $T_2$, and at this time as point $A$ falls in potential a negative trigger is transmitted to the second pair, causing the transfer of con-

duction from $T_4$ to $T_3$. Indicator light 1 goes out, and 2 is turned on. A third input pulse transfers conduction back to $T_1$, lighting indicator 1. A fourth pulse returns conduction to $T_2$ and transmits another negative trigger to the second pair, causing $T_4$ to conduct and further transmitting a trigger to the third pair, causing $T_5$ to conduct.

After four pulses it can be seen that tubes $T_2$, $T_4$, and $T_5$ are conducting and indicator 3 is on. It can be further reasoned that after eight input pulses tubes $T_2$, $T_4$, and $T_6$ are conducting and an output pulse has been transmitted to the recording counter. The sequence is summarized in the accompanying tabulation, where the indicator "off"

| Number of pulses | Indicator light | | | Output signal |
|---|---|---|---|---|
| | 1 | 2 | 3 | |
| 0 | 0 | 0 | 0 | |
| 1 | 1 | 0 | 0 | |
| 2 | 0 | 1 | 0 | |
| 3 | 1 | 1 | 0 | |
| 4 | 0 | 0 | 1 | |
| 5 | 1 | 0 | 1 | |
| 6 | 0 | 1 | 1 | |
| 7 | 1 | 1 | 1 | |
| 8 | 0 | 0 | 0 | $X$ |

condition is shown by 0, the indicator "on" condition by 1. If the indicator lights are given the values 1, 2, 4, . . . , it can be seen that the number of pulses received is given by the properly weighted sum of the "on" indicator lights, seven input pulses being given by $1 + 2 + 4 = 7$.

**9-25. Decimal Scaling Circuits.** In general a set of trigger pairs will count down by the ratio $2^n$, where $n$ is an integer, and thus count-down ratios of 2, 4, 8, 16, . . . are possible directly. However, feedback can be applied in many different ways to cause a circuit to reset itself to a previous condition after arrival of the $k$th pulse, thus essentially subtracting a few pulses of number $r$, and providing a scale factor equal to $k + r$. In particular it is desirable to be able to scale by the decimal factor 10, or to count in the decades of our usual number system. The circuit of Fig. 9-55, to be described as an example, provides such a decimal counting system with four trigger pairs in which the count-down ratio would normally be 16. The circuit provides decimal counting by making $k = 9$ and $r = 1$, so that $k + r = 10$; the result is accomplished by forward as well as backward coupling or feedback between stages.

The circuit of Fig. 9-55 employs coupling between trigger pairs by means of the common anode connection, only negative pulses on this

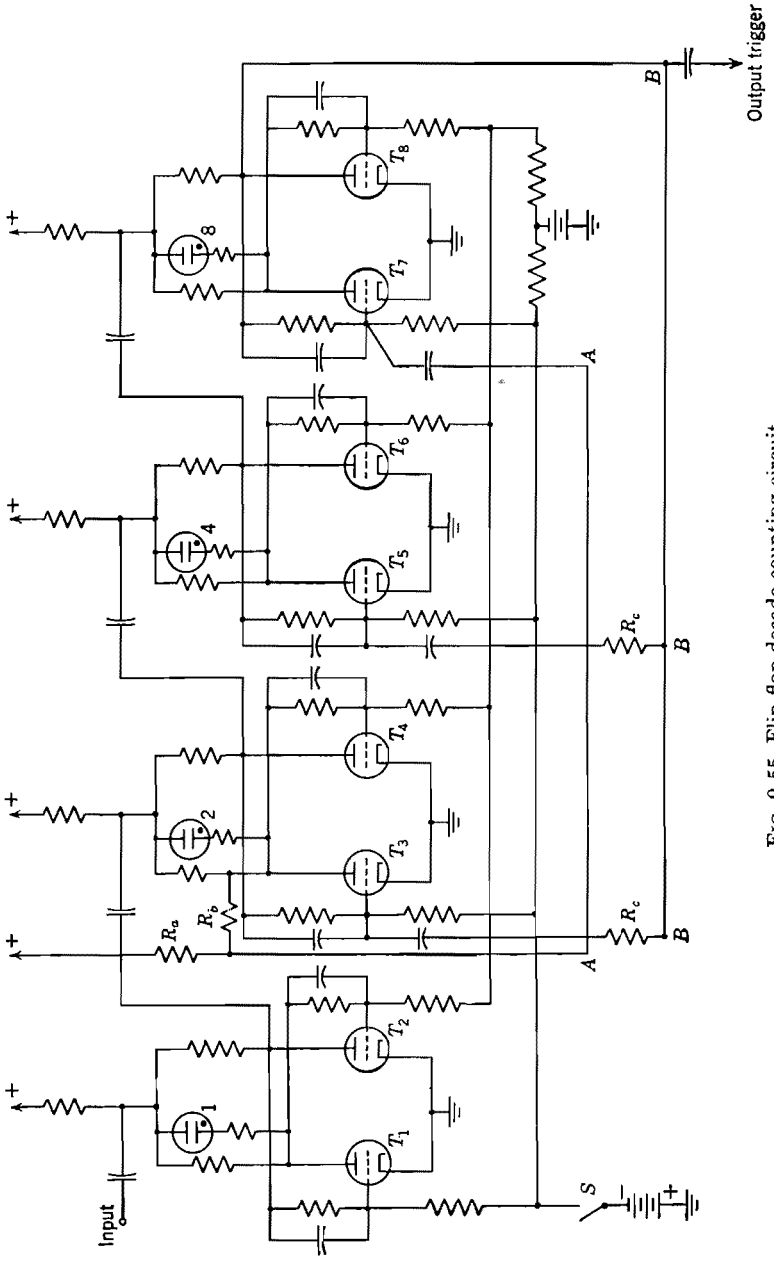



FIG. 9-55. Flip-flop decade-counting circuit.

lead then being effective to produce switching. Reset to the desired initial condition of $T_1$, $T_3$, $T_5$, $T_7$ nonconducting is possible by use of switch $s$, which biases the corresponding grids in negative fashion. After this initial setting the circuit will count input pulses and give an output pulse for every tenth input pulse. This output might be used as input to further decade circuits, counting 100, 1000, etc., respectively. Indicating lamps are associated with $T_1$, $T_3$, $T_5$, and $T_7$ which glow if their respective tube is conducting.

Forward coupling from the anode of $T_3$ is provided to the grid of $T_7$ through line $A$-$A$, and further feedback is also applied from the output or anode $T_8$ to the grids of $T_3$ and $T_5$ through line $B$-$B$. The actions induced by these feedback couplings will be explained below.

The switch $s$ initially resets the odd-numbered tubes to "off," or nonconducting, this action giving the first line of the accompanying tabulation. The tube conditions are indicated as "on" or "off," the signal lights being indicated as 0 for off and 1 for on. The number of pulses received at any instant may be counted by weighting the lamps by the factors 1, 2, 4, 8 and adding. To this must be added the reading of the output counter or higher decades.

ACTIONS OF DECADE SCALER

| Number of pulses | $T_1$ | $T_2$ | Light count 1 | $T_3$ | $T_4$ | Light count 2 | $T_5$ | $T_6$ | Light count 4 | $T_7$ | $T_8$ | Light count 8 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | Off | On | 0 | Off | On | 0 | Off | On | 0 | Off | On | 0 |
| 1 | On | Off | 1 | Off | On | 0 | Off | On | 0 | Off | On | 0 |
| 2 | Off | On | 0 | On | Off | 1 | Off | On | 0 | Off | On | 0 |
| 3 | On | Off | 1 | On | Off | 1 | Off | On | 0 | Off | On | 0 |
| 4 | Off | On | 0 | Off | On | 0 | On | Off | 1 | Off | On | 0 |
| 5 | On | Off | 1 | Off | On | 0 | On | Off | 1 | Off | On | 0 |
| 6 | Off | On | 0 | On | Off | 1 | On | Off | 1 | Off | On | 0 |
| 7 | On | Off | 1 | On | Off | 1 | On | Off | 1 | Off | On | 0 |
| 8 | Off | On | 0 | Off | On | 0 | Off | On | 0 | On | Off | 1 |
| 9 | On | Off | 1 | Off | On | 0 | Off | On | 0 | On | Off | 1 |

The first pulse turns on $T_1$ and lights the units lamp. The turning off of $T_2$ transmitted a positive pulse to $T_3$ which is ineffective since anode switching occurs only with negative pulses. Thus the conditions of the table for a single pulse are obtained. The second pulse turns off $T_1$ and turns $T_2$ on, extinguishing lamp 1. A negative pulse is transmitted to $T_3$ by the turning on of $T_2$, as its anode potential drops, and $T_3$ conducts and lights lamp 2, giving the conditions of the table.

It may be noted that the turning on of $T_3$ transmitted a negative pulse

through $A$-$A$ to the grid of $T_7$. However, $T_7$ is already cut off, and a negative pulse *to the grid* has no effect. Later, at the fourth pulse, a positive pulse will be transmitted to the grid of $T_7$ through $A$-$A$, but this is adjusted in magnitude by resistors $R_a$ and $R_b$ to be too small to trigger $T_7$, in accordance with the discussion of grid triggering in Sec. 9-23. The reception of the eighth pulse similarly has no effect on $T_3$ and $T_5$ through the $B$-$B$ feedback because the turning off of $T_8$ transmits a positive pulse which also is too small to turn on $T_3$ and $T_5$. Thus the action follows in normal fashion until the ninth pulse is recorded in the table (remember that only a negative pulse causes a tube to switch).

The tenth pulse causes further triggering actions, leading to the changes from pulse 9 to pulse 10 given in the accompanying tabulation. The

| Number of pulse | $T_1$ | $T_2$ | Light count 1 | $T_3$ | $T_4$ | Light count 2 | $T_5$ | $T_6$ | Light count 4 | $T_7$ | $T_8$ | Light count 8 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 9 | On | Off | 1 | Off | On | 0 | Off | On | 0 | On | Off | 1 |
| 10 | Off | On | 0 | On | Off | 1 | Off | On | 1 | Off | On | 0 |
| Plus an output pulse, then by feedback changes to: | | | | | | | | | | | | |
| 10 | Off | On | 0 | Off | On | 0 | Off | On | 0 | Off | On | 0 |

action of $T_4$ in going from on to off at pulse 10 transmits a positive pulse to the next pair which cannot operate the $T_5$, $T_6$ combination. However, the action of turning on $T_3$ transmits a negative pulse through the $A$-$A$ coupling to $T_7$, turning it off and turning on $T_8$. When $T_8$ turns on, a negative pulse is fed back through $B$-$B$ to the grids of $T_3$ and $T_5$, causing $T_3$ to turn off and preventing $T_5$ from turning on at the instant $T_4$ goes on and transmits a negative pulse to $T_5$, $T_6$. The result is the second step above for pulse 10. At the same time a negative output pulse is transmitted from $T_8$ to a counter or to the next decade of the scaler. The condition of the decade after 10 counts is identical with its starting condition; so the circuit stands ready to count a second group of 10 pulses.

It can be seen that the tenth pulse causes three sets of operations to occur, the first of which switches $T_3$ on but does not operate $T_5$. Then there is transmitted a negative pulse through the $A$-$A$ coupling to switch $T_7$ off and $T_8$ on. This latter action transmits a negative pulse through the $B$-$B$ feedback, causing $T_3$ to be switched off or reset and $T_5$ to be held off. To allow time for these events to occur in sequence, the feedback pulses are delayed slightly through the action of $R_b$ and associated capacity, and the resistor $R_c$ and the coupling capacitors; otherwise the circuit might appear to be attempting to make the triggers change both ways at once.

Many other feedback paths may be provided to achieve different countdown ratios. In addition, since a set of voltages exists in the circuit for every stable state which is unique to that state, it is possible to incorporate control circuits which will indicate or control only after a preset number of pulses has been counted. For instance, the number of pills going into a bottle may be counted photoelectrically and the filling operation stopped after the desired number have been supplied. These scaling and counting methods provide a means of measurement of number of events to an accuracy and speed unparalleled by other means. In particular is this true in measurement of frequency, where such counters can be designed to count cycles to an accuracy of six, seven, or more significant figures.

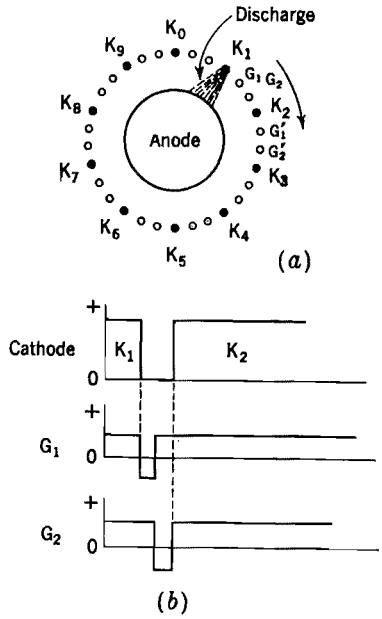



FIG. 9-56. (*a*) Decade-counter tube. (*b*) Potential relations during switching.

**9-26. The Decade-counter Tube.** The complexity of the typical circuit of the foregoing section has led to the development of single tubes, usually gas-filled, to perform the decade-counting operation. They are suited to medium-speed operation and are read by noting the position of the ionization glow on one of 10 numbered cathodes.

One such tube operates on the principle that the breakdown voltage for a gas discharge is lowered if ions are already present in the space between cathode and anode. If a series of cathodes are placed about a common anode, the gas discharge can be made to move from cathode to cathode by a succession of negative voltage pulses, if initial ionization is provided in the space between anode and the succeeding cathode.

Figure 9-56*a* illustrates a top view of such a tube, having a central anode surrounded by 30 cathodes in groups of three. The main cathodes are designated $K_o$, . . . , $K_9$ and are all connected in parallel. Also in each group of three are two guide cathodes, $G_1$ and $G_2$, or $G_1'$ and $G_2'$, and similarly around the ring. All $G_1$ cathodes are in parallel, and all $G_2$ cathodes are likewise connected together.

Assuming a gaseous-glow discharge present between the anode and $K_1$, a negative pulse on all $G_1$ guide cathodes will cause the glow to move to $G_1$ nearest $K_1$, as the most negative electrode in the region. It is easy

for the discharge to make this transfer since ions from the discharge to $K_1$ are in and near the path to this $G_1$ cathode. Upon removal of the negative pulse to $G_1$ a second negative pulse is applied to all $G_2$ cathodes, and the discharge transfers to the $G_2$ nearest $G_1$, since this guide cathode is now the nearest and most negative electrode. At the end of the negative pulse on $G_2$ the glow transfers to $K_2$, as nearby and more negative than the normal $G_1$ or $G_2$ potential.

All transfers are preferential to any electrode because of the presence of ions from the adjacent glow. Although the negative transfer pulses are applied to all 10 of the $G_1$ or $G_2$ cathodes, the priming influence of the nearby discharge moves the glow to the guide cathode closest to the discharge. The discharge is not likely to back up at any time because of the shielding effect of the intervening cathode. Thus the discharge moves clockwise three electrodes for every input pulse. The associated circuits are called upon to generate two pulses for guide-cathode operation for every input pulse, and this may be done by circuits already described.

Internal light shields are incorporated so that the glow may illuminate the appropriate digit on a scale plate, the device then being direct-reading. It is possible to make the cathode potential of $K_o$ transmit a pulse to a second tube so that successive decade tubes may be operated, to extend the counting to hundreds, thousands, or more. Because of the time required to deionize the gas, the input frequency is limited to about 4000 pulses per second.

A vacuum tube of another form employs a ribbon-shaped electron beam which in passing between a pair of deflection plates is successively shifted from an initial stable position to each of 10 stable positions, in each of which it passes through an anode slot and strikes a fluorescent screen. This places a luminescent mark opposite a corresponding numeral on a mask fixed outside the tube envelope. The stepping action is obtained by interaction between the currents intercepted by a slotted screen grid and the slotted anode, causing the beam to lock into each position in turn, after being deflected by a positive-going triangular pulse. Input pulse rates up to 30,000 per second are possible.

**9-27. The Binary-number System.** In the decimal system of numbers 10 states, or marks, 0, 1, 2, 3, 4, 5, 6, 7, 8, and 9 are recognized. A number is represented as the sum of these marks times powers of 10; for example, the decimal number 1074 is actually formed as

$$1074 = 1 \times 10^3 + 0 \times 10^2 + 7 \times 10^1 + 4 \times 10^0$$

but by convention we write only the coefficients in proper position as 1074. In general a decimal number is formed as

$$S = k_n \times 10^n + k_{n-1} \times 10^{n-1} + \cdots + k_1 \times 10^1 + k_o \times 10^0 \quad (9\text{-}59)$$

An ordinary electrical switch has two recognizable states, "open" and "closed," or "on" and "off." This is also the case with electrical relays and with the trigger, or flip-flop, circuit previously described. The frequent use of on-off switching circuits has led to the application of a two-state, or *binary*, number system in telephone switching-circuit design, and this has now been further applied in electronic digital computers because of the properties of the trigger, or flip-flop, circuit.

The binary system makes it necessary only to recognize that a tube is on or off, or conducting or nonconducting. To represent more states than two with vacuum tubes would require recognition of discrete plate-current values, and for this purpose the vacuum tube lacks sufficient precision.

TABLE 17. DECIMAL-BINARY CONVERSION

| *Decimal* | *Binary* | *Decimal* | *Binary* | *Decimal* | *Binary* |
|---|---|---|---|---|---|
| 0 | 0 | 11 | 1011 | 22 | 10110 |
| 1 | 1 | 12 | 1100 | 23 | 10111 |
| 2 | 10 | 13 | 1101 | 24 | 11000 |
| 3 | 11 | 14 | 1110 | 25 | 11001 |
| 4 | 100 | 15 | 1111 | 26 | 11010 |
| 5 | 101 | 16 | 10000 | 27 | 11011 |
| 6 | 110 | 17 | 10001 | 28 | 11100 |
| 7 | 111 | 18 | 10010 | 29 | 11101 |
| 8 | 1000 | 19 | 10011 | 30 | 11110 |
| 9 | 1001 | 20 | 10100 | 31 | 11111 |
| 10 | 1010 | 21 | 10101 | 32 | 100000 |

The two recognizable states in the binary system are taken as off and on, or false and true, and are usually represented as 0 and 1. The writing of a number in binary notation employs powers of 2, because of the 2 states, similar to the use of powers of 10 and the 10 states of the decimal system. Writing the decimal number 1074 in the binary system gives

$$S = 1 \times 2^{10} + 0 \times 2^9 + 0 \times 2^8 + 0 \times 2^7 + 0 \times 2^6 + 1 \times 2^5 + 0 \times 2^4 + 1 \times 2^3 + 0 \times 2^2 + 1 \times 2^1 + 0 \times 2^0 \quad (9\text{-}60)$$

the successive terms representing decimal equivalents of

$$S = 1032 + 0 + 0 + 0 + 0 + 32 + 0 + 8 + 0 + 2 + 0 = 1074$$

Writing the coefficients of Eq. (9-60) in order, however, gives the binary representation of the decimal number 1074 as 10000101010.

Arithmetic operations may be carried out with the binary notation. Addition follows from the simple rules

$$0 + 0 = 0$$
$$1 + 0 = 1$$
$$1 + 1 = 10$$

the latter step following from the fact that the binary equivalent for 2 is 10.

Subtraction follows from the rules

$$0 - 0 = 0$$
$$1 - 0 = 1$$
$$1 - 1 = 0$$

A multiplication table may also be written from the rules

$$0 \times 0 = 0$$
$$1 \times 0 = 0$$
$$1 \times 1 = 1$$

Division follows as an inverse operation.

**Example.** Add binary 11010 to binary 10010.

$$\begin{array}{rll} & 11010 & = 26 \\ & 10010 & = 18 \\ \hline & 01000 & \\ \text{Carry} & 1\ \ 1 & \\ \hline & 101100 & = 44 \end{array}$$

**Example.** Multiply 13 by 6 in binary manner. Binary 13 = 1101; binary 6 = 110.

$$\begin{array}{rll} & 1101 & 13 \\ & 110 & 6 \\ \hline & 0000 & 78 \\ & 1101 & \\ & 1101 & \\ \hline & 101110 & \\ \text{Carry} & 1 & \\ \hline & 1001110 & \end{array}$$

$$1001110 \doteq 1 \times 2^6 + 0 \times 2^5 + 0 \times 2^4 + 1 \times 2^3 + 1 \times 2^2 + 1 \times 2^1 + 0 \times 2^0$$
$$\doteq 64 + 8 + 4 + 2$$
$$1001110 \doteq 78 \quad \textit{Check}$$

The possibility of computation with binary numbers and the high potential speed of operation of electron tubes have led to the development of electronic digital computers, to be discussed in following sections. Therein the simplicity of the binary system, in terms of only two recognizable states, is purchased at the expense of carrying increased numbers of digits, decimal 31 requiring 5 binary digits, with a 10-place decimal number requiring up to 40 binary digits.

It can be seen that four digits are required to cover one decade, since binary 10 = 1010. These same four digits, and the necessary equipment, can be used to extend the system to binary 15, and thus a sexidecimal, or 16-mark, system is available, and occasionally employed in computer

design. The numbers above 9 in this system may be represented as letters to avoid ambiguity. An octal system employing only eight digits is also possible, decimal digits 8 and 9 being represented by sums such as $8 = 3 + 5$, $9 = 2 + 7$.

In information theory each basic piece of information is spoken of as a *bit;* thus in binary notation each digit is a bit, and the number 10011 is composed of five bits.

**9-28. Digital Computers.** The desk calculator is an example of a class of devices operating as *digital computers*, where the input and output are numerical digits in contrast to the analog computer, in which the input and output are physical quantities. The desk calculator performs various arithmetical operations largely through an ability to add or subtract, the human operator directing, or *programming*, the machine in such a way that through a sequence of additions or subtractions many more complex arithmetical operations are performed.

Assuming proper machine operation, the accuracy of a digital computer is dependent only on the number of significant digits carried through the computation, and this number is largely fixed by questions of allowable complexity and cost of the machine. Human reading error and component calibration are largely eliminated by the digital form of the result.

Modern science has many problems which are incapable of solution in reasonable time by the mechanical form of computer. During the period 1944 to 1947 several computers were built which sought to improve the speed of operation through the use of electrical relays as operating elements with the binary system of arithmetic, to which the relays lent themselves through their two-state open-or-closed form of operation. Numbers were handled in binary form as chains of electrical pulses representing 1s and 0s in proper sequence.

The first relay machine, the Harvard Mark I calculator, completed in 1944, was capable of adding two 23-digit (decimal equivalent) numbers in 0.3 sec, thus materially exceeding the speed of a desk calculator, which is of the order of two 10-digit additions per minute, including time to record the result.

Further increases in operating speed were obtained by application of the electron tube to replace the relay, it being realized that the flip-flop circuit was well suited to such numerical computation in binary numbers. With the flip-flop and certain other circuits to be described it is possible to build an electronic digital computer capable of the usual arithmetic operations at high speeds. The first such large-scale electronic digital computer was the ENIAC (Electronic Numerical Integrator And Computer) completed at the University of Pennsylvania in 1947. This increased the basic speed of computers to the addition of two 10-digit (decimal equivalent) numbers in about 200 $\mu$sec. Other computers have

been built since that time which exceed this speed or have greater flexibility in operation. The ENIAC employed about 18,000 vacuum tubes, later types reducing this to 2000 to 4000.

The various forms of electronic digital computers operate, in general, by principles which parallel the usual numerical methods, and essentially by processes of addition or subtraction. They employ well-known methods of successive approximation or iteration to achieve interpolation, function values, integration, and the solution of systems of differential and algebraic equations. The pulse lengths employed with the binary numbers vary from a few microseconds to fractions of a microsecond.

The parallelism in operation may be illustrated by the following. In the usual manual process of addition of two numbers, say 618 to 723, it is customary to proceed as follows:

| *Manual operation* | | *Computer function* |
|---|---|---|
| Write down | 618 | Input and store |
| Write down | 723 | Input and store |
| Add | 331 | Add |
| Carry | 1 1 | Carry |
| Write result | 1341 | Output or store |

It can be seen that actually, in writing down the number, one performs an *input* function, or introduces the numerical data to the process. Writing the numbers also places them where they will be available later, or the numbers may be considered as *stored* on the paper until required. Then follows the *arithmetical operation* of addition, including carries. Finally, writing down of the sum represents an *output* function, and also one of *storage* of the answer.

In the above there was a mental *control* of the problem, the brain telling the hand what to write and in what order. Thus a digital computer must in some measure include the functions of:

Input
Storage or memory
Control
Arithmetical operation
Output

and these may be interrelated somewhat as in Fig. 9-57.

To solve a problem requires that the input supply both numerical data and control orders to the memory. Orders are then issued from the memory to the control, which in turn actuates the arithmetic unit to perform the ordered calculations on numbers withdrawn from the memory. The results of the calculation are then returned to the memory or are sent to the output unit.

Electronic digital computers are being built to handle 10 or more decimal digits, which requires the handling of 40 or more binary digits. Internally twice this number may be required in multiplication operations. As was pointed out, the binary system gains simplicity in recognizable states at the price of handling additional digits, which adds to the cost.

It might appear, in fact, that, while the 10 states of the decimal system may be too many in terms of accurate electronic resolution, yet the 2 states of the binary system may be too few in terms of equipment complexity and cost. Following this further, if $r$ be the number of recognizable states ($r = 2$ for binary) and $n$ the number of machine digits to be employed, then the decimal numbers expressible in the system are $r^n$.

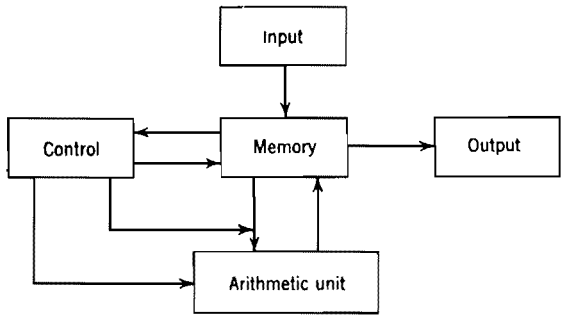



FIG. 9-57. Functions of a digital computer.

If it be assumed that one triode tube or equivalent is required for each digit and each recognizable state, then the complexity or quantity of apparatus and cost may be measured by $rn$. The largest decimal number to be handled is usually specified, and thus $r^n = k$ is a fixed quantity. Therefore

$$\text{Cost} = rn \qquad \text{but} \qquad r^n = k$$

Then

$$n \ln r = k$$

$$\text{Cost} = \frac{rk}{\ln r}$$

Taking the derivative with respect to $r$, it is possible to show that the cost is a minimum when

$$r = \epsilon = 2.718$$

Consequently values of $r$ of 2 or 3 are near minimum, and the value of 2 for $r$ is chosen because of the properties of the flip-flop circuit and the easing of accuracy requirements on the associated electron tubes.

Computers may be of several types. A *synchronous* computer includes a *clock*, or pulse-frequency generator, and operations proceed at the clock-pulse rate, with definite numbers of clock pulses assigned for the completion of various operations. An *asynchronous* computer has no fixed

rate, the completion of each operation providing an enabling signal for the computer to proceed to the next operation.

A *series* computer operates on each binary digit in order or serially; thus the time to complete addition of two 15-digit numbers is 15 times that of handling 1 digit. A *parallel* computer operates on all digits simultaneously, after which time must be allowed for the carry digits to propagate through the system. The parallel form requires more complex circuits but usually allows a saving in time for an operation.

Transfer from series to parallel representation is possible. If 15 digits are being handled, then the first digit of the serial representation must be delayed 14 digit intervals, the second 13 digit intervals, etc., the last digit then being transmitted in parallel with the others, each simultaneously into its own channel.

In most digital computers the addition operation is carried out much as indicated in Sec. 9-27. Circuit complexity is increased by the necessity to provide for handling the carries, and this is discussed in Sec. 9-34. Three registers or accumulators for transient number storage may be needed, two to receive the numbers to be added and the third to receive the result before transfer to the memory or the output.

Multiplication follows as a process in addition, carried out by repeated shifts of the multiplicand to the left (or right for numbers less than unity) and adding of the multiplicand when called for by a 1 in the multiplier. Thus 1011 × 1010 (decimal 11 × 10) gives:

| | |
|---|---|
| | 1011 |
| | 1010 |
| Shift, nonadd | 0 |
| Shift, add | 1011 |
| Shift, nonadd | 0 |
| Shift, add | 1011 |
| | 1101110 |

The result is given by one left shift to represent multiplication by 0, a second left shift and addition to represent multiplication by 1, another left shift without addition for the next 0, and a final left shift and addition for the multiplication by the final 1.

Division is carried out by right shifts and subtractions. Other operations such as square root or integration are performed by numerical processes based primarily on addition methods.

**9-29. Computer Input and Output.** Input and output equipment for digital computers must be fast enough to handle data and orders without delaying computer operation significantly. Information is usually supplied in the form of punched cards, punched teletypewriter tape, or magnetic tape. All these methods have the advantage of being capable of

advance preparation and checking; thus computer time is not used until the problem is actually run. Data are usually transformed to binary, octal, sexidecimal, or a special code by equipment similar to a teletypewriter, a sample punched tape appearing in Fig. 9-58*a*. Magnetic tape is prepared in equipment typical of the sound-recording field except for its higher speed of operation.

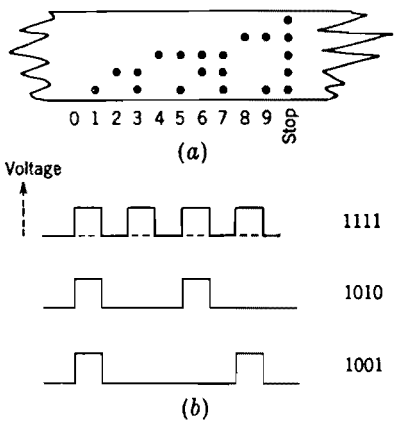



FIG. 9-58. (*a*) Manner of punching binary digits in input tape. (*b*) Voltage pulses as introduced into the computer.

Tape readers may be mechanical or photoelectric, and, for a tape as of the figure, will read each decimal digit as a group of four pulses as in Fig. 9-58*b*. These may be transmitted either in parallel or in series to flip-flops in a register chain of flip-flops. However, these numbers are present on the tape in a binary-coded decimal system and must be converted by the machine to true binary numbers before further operations can be carried on.

In one type of operation the conversion may be performed by reading the highest-order decimal digit into the first four flip-flops to the left of the decimal location. This quantity is then multiplied by 1010, or decimal ten, and the tape code of the next-lowest-order decimal digit added to the product. If another decimal digit follows, then the process is repeated. The multiplication by 1010 is not required for the least significant decimal digit, since it is found in the units place. For instance, assume that decimal 132 is punched into the tape as 0001, 0011, 0010. The first four binary digits are read into the register and operated on as follows:

| | |
|---|---|
| Add 0001(1) | 0001 |
| Multiply by 1010 (10) | 1010 |
| | 1010 |
| Add 0011 (3) | 0011 |
| | 1101 |
| Multiply by 1010 (10) | 1010 |
| | 11010 |
| | 1101 |
| | 10000010 |
| Add 0010 (2) | 0010 |
| | 10000100 |

and the result is the binary equivalent for 132. The multiplication by ten (1010) is easily accomplished, as previously illustrated in Sec. 9-28.

Fractions may be handled starting with the least significant decimal digit and dividing by 1010 to reach the binary equivalent. Some computers are capable of handling this operation at rates of several hundred decimal digits per second.

The reverse process may be employed in converting from the binary numbers to decimals for the punching of output tape. Repeated division by 1010 (decimal ten) can be performed, the remainder after each division indicating the binary value of a decimal digit, this four-bit binary number then being punched in the tape, later to operate an electric typewriter suited to the code.

Output devices may consist of tape punches, for providing teletypewriter tape which may be stored and later run through the teletypewriter for reproduction, or data may go directly from the memory to an electric typewriter. These machines must, of course, be altered to operate on the code being employed. Since the output typewriter is the slowest element in the chain, at about 10 typed digits per second, active development continues on speeded methods, including typewriters capable of 200 characters per second, and various methods employing the cathode-ray tube for recording the data, which are then photographed from the screen, and electrostatic printing methods with speeds of up to several thousand characters per second.

**9-30. Memory, or Storage, Facilities.** The memory is the heart of a digital computer. It must be of capacity sufficient, in terms of digits, to store numerical data and orders for a complex operation without need for frequent reference to slowly available external orders, and the information in storage must be available in a time short enough as not to delay the computation unduly. The study of various types of storage systems requires definition of certain terms.

*Cyclic systems* are those in which the stored information is being continously moved, or cycled, in time or space, such that a given digit can be reached only once per cycle. *Noncyclic* systems are those in which digits are stored statically, and all are equally accessible to the read-out circuit. Switching systems such as those described in Sec. 9-32 may be required to select the desired digits. Time provides this switching for cyclic systems.

Some systems are *erasable;* others are *nonerasable.* For instance, of the types of input equipment discussed, punched cards or paper tape are nonerasable, having to be repunched to remove errors. Magnetic tape is erasable immediately for reuse.

Memory systems must also be compared on the basis of *volatility* and *nonvolatility.* A volatile system is one in which the information will dis-

appear, or "evaporate," upon failure of power or of some component. A nonvolatile system is one in which the data are permanently recorded even though machine power may fail.

*Access time* is the interval between ordering a quantity out of storage and the instant at which the storage actually supplies that information. This may involve merely the time for the circuits to respond, measured in microseconds, or the time necessary to search a long loop of magnetic tape for the requested data.

A number of basically different storage systems are in use or contemplated, including the following:

*Electrical Relays.* Relays are arranged to receive digits as open or closed operations. They are slow in operation and not perfectly reliable on contact closure, and each relay can store only one digit. The information is volatile unless mechanical lockup of the relays is provided, and access time may be slow because of the 10- to 30-msec operating time of relays.

*Acoustic Delay Lines.* Pulsed sound waves are transmitted through a material such as mercury or fused quartz, the information or pulsed digits being kept recirculating from the output end of the line to input by use of amplifiers. At a desired instant the signal may be picked off the circuit and sent into the computer. Quartz or other piezoelectric crystals are used as transducers to transmit and receive the sound waves.

For mercury the velocity of sound is about $1.45 \times 10^5$ cm per sec. The time of storage, for one traverse of an acoustic path 40 cm long, is

$$\frac{40}{1.45 \times 10^5} = 276 \ \mu\text{sec}$$

If the digits are sent into storage at a rate of four digits per microsecond, the number of digits stored per path in the mercury is

$$4 \times 276 = 1104$$

This number can be increased by the use of parallel channels in each mercury cell; however, the number of amplifier tubes required for regeneration and recirculation of the signals may be excessive.

Greater rates of pulse input, or shorter physical digit lengths, are not possible in mercury because of the effect of temperature variation on sound velocity, which is about 1 part in 3000 per degree centigrade. If as many as 3000 digits were to be stored in one traverse, the mercury temperature would have to be held to much better than 1°C; otherwise, it would not be possible to tell which digit was being read at a given time. Fused quartz has a temperature dependence about half that of mercury and has found some application.

*Magnetic Tape.* This method employs variation of intensity of resid-

ual magnetization along a cellophane tape coated with a thin layer of magnetic material. The tape passes across an air gap in a magnetic circuit, as in Fig. 9-59, being exposed to the varying magnetic flux at speeds of several feet per second. Spots of strong and zero residual magnetization are left as evidence of 1s and 0s recorded. Continuous cycling of a ring of tape allows storage of considerable data and ability to read off the recorded information as desired. This leads to an access time equal to the time of rotation of the ring of tape, and the storage is nonvolatile.

*Magnetic Drum.* Numerous magnetic tracks may be recorded side by side on the surface of a rotating magnetically coated drum. With a recording-reading head as in Fig. 9-59 available for each track many digits can be recorded in parallel and stored until required. It is necessary to wait up to one revolution time to obtain a given number; so high rotational speeds are employed to reduce this access time. Such storage is nonvolatile.

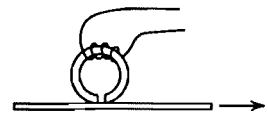


FIG. 9-59. Magnetic recording head.

*Electrostatic Cathode-ray Storage.* This method is occasionally referred to as "Williams-tube storage" after its inventor. It is possible to store electric charge on the dielectric screen of a cathode-ray tube. If the electron beam is positioned and turned on for an instant, an electric charge is imparted to the screen dielectric at the point of impact. It is found that if the beam is moved, or "twitched," during its "on" time, forming a dash on the screen, the pattern of charge will be distinctly different from what it would be if the beam were not moved. Thus a raster of dot positions is chosen on a screen, possibly a 32- by 32-dot pattern, having 1024 storage locations. If at each position the beam is turned on but not moved, a 0 is recorded. If the beam is turned on and twitched sideways, a 1 is recorded. The locations of the spots are referred to by *addresses* in binary code.

For read-out the beam is returned to a given address and twitched. This discharges the spot, and the output-current waveform to a metal electrode covering the outside of the tube screen differs, depending on whether a 0 or a 1 had been recorded on that spot.

The charge tends to leak off the screen gradually, and circuits must be provided to cause the beam to scan all spots and re-record the data thereon at a rate of a few hundred times per second. This operation is sandwiched between read-in and read-out operations and can be interrupted at any time that read-in or read-out of data is called for. Any spot can be read as quickly as any other spot, so that a high rate of access to the memory is a major advantage. A disadvantage is that, if power or circuits fail, all stored data are lost.

*Magnetic-core Matrix.* The availability of magnetic materials with substantially rectangular hysteresis loops, as in Fig. 9-60*a*, and high residual-flux densities makes possible a magnetic-core storage. In this system a 0 state represents a negatively magnetized core in saturation and a 1 state a core in positive saturation with a flux-density value remaining at point $X$ of Fig. 9-60*a*.

Numbers of very small ring cores of these materials are strung together in a matrix of horizontal and vertical wires. At every point where two of these insulated wires cross, they run through the center of a core, Fig. 9-60*b* being a section of the matrix holding four such cores. A current flowing through two wires is required to drive a core up to saturation, and, by applying these currents simultaneously to one horizontal and one vertical wire, only one core at a time receives the required value of magnetizing intensity. Thus a digit is stored, a 0 or 1 depending on the polarity of the currents. The material holds the saturation state indefinitely until reverse ampere turns are supplied.

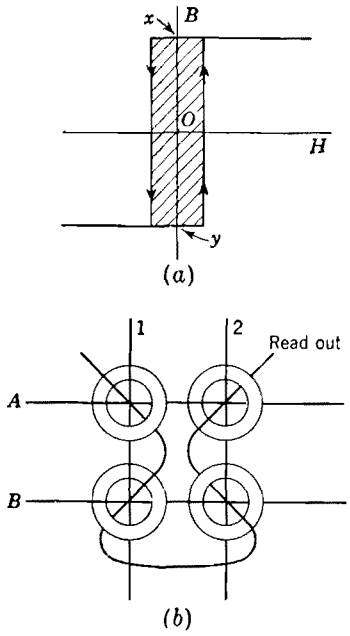



FIG. 9-60. Magnetic-core matrix memory. (*a*) Hysteresis loop of material; (*b*) wiring of portion of matrix.

If a 1 is stored by the condition of $X$, read-out is possible by sending a test current through the two wires meeting in the desired core. If the polarity is that required to write a 1, and a 1 state exists, there will be no flux change and no emf induced in the read-out wire. If a 0 state had existed in the core, the magnetic conditions would have reversed from $Y$ to $X$ and a read-out emf obtained. Read-out destroys the stored information, but associated pulse circuits are designed to restore the core to its initial state, and no information is lost.

Such a memory is erasable but nonvolatile, and access time is only that required to operate the associated circuits.

A comparison of volume requirements and the relative cost of several memory systems are given in Table 18. These data are subject to change with time and further development.

**9-31. Vacuum-tube Gating Circuits.** Vacuum-tube *gates* are circuits designed to require coincidence of two or more input signals to produce an output. Thus a signal from the control might be applied, and if a

TABLE 18. APPROXIMATE STORAGE VOLUME AND COST OF VARIOUS MEMORY SYSTEMS

| Type | Digits per cu ft | Digits per dollar cost | Comment |
|---|---|---|---|
| Relay | 2–4 | 0.2–1 | Obsolete |
| Punched tape | $3 \times 10^7$ | $3 \times 10^6$ | Very slow access |
| Magnetic tape | $3 \times 10^9$ | $10^7$ | Slow access |
| Mercury line | $2 \times 10^4$ | 10 | Fast access |
| Electrostatic cathode-ray tube | $10^3$ | ..... | Very fast access |
| Magnetic-core matrix | $10^6$ | ..... | Very fast access |

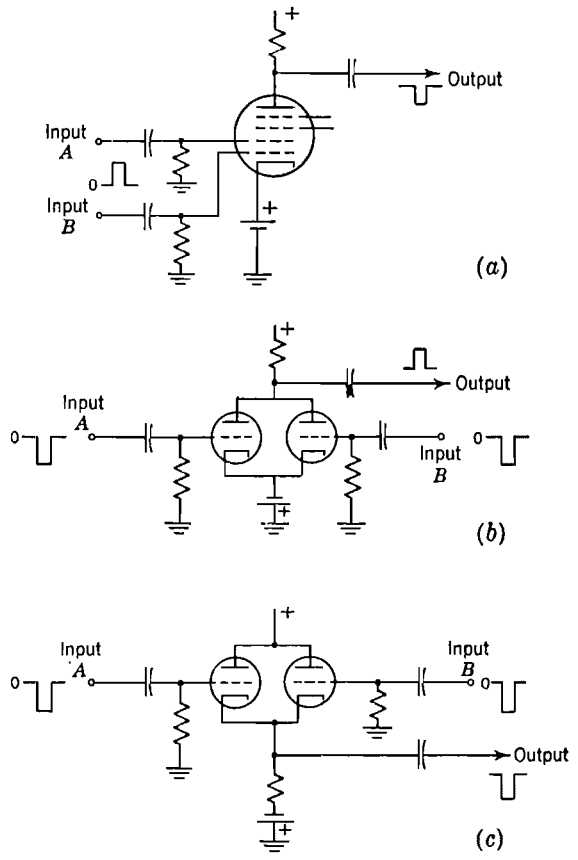



FIG. 9-61. Gating circuits. (*a*) Multigrid gate; (*b*) parallel gate; (*c*) cathode-coupled gate.

second signal then arrives, an output is produced. If the control signal had not been present to open the gate, the second signal would not have been passed on. Such circuits can be used for switching or for selection of signal paths in a computer as actuated by signals from the control unit.

The multigrid circuit employing a 6SA7 tube or similar type as in Fig. 9-61*a* operates normally with negative bias and at cutoff on both grids. A positive signal on one grid is insufficient to yield current flow, but if both input grids are simultaneously positive, the tube conducts and a negative output is obtained. The two grids normally do not have equal $\mu$, and required signal amplitudes may differ, although it is usually intended to drive both grids positive to saturation.

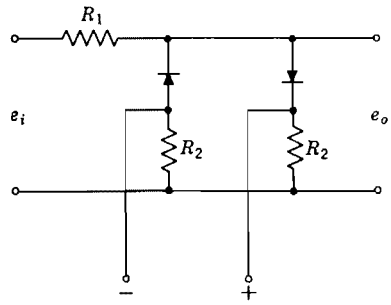



FIG. 9-62. Diode gate.

The circuit of Fig. 9-61*b* employs two triodes, each normally conducting at saturation level. Cutoff of one triode does not appreciably alter the current and plate potential, but if both tubes receive a negative cutoff signal simultaneously, the plate potential rises to $E_{bb}$ and a positive output pulse is obtained.

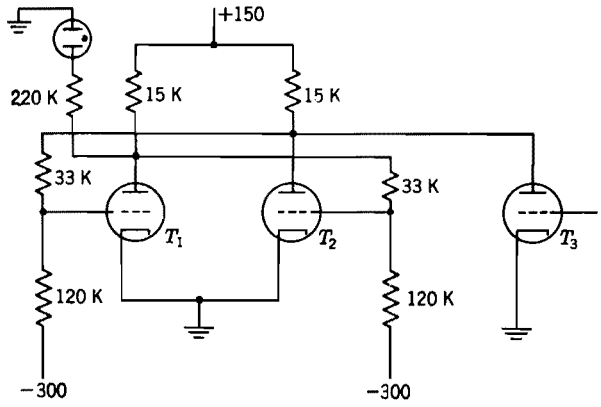



FIG 9-63. Gating or switching tube to restore flip-flop to initial condition.

The circuit in Fig. 9-61*c* is similar, with both triodes normally conducting saturation current. Cutoff of one tube by a negative signal has little influence on the cathode potential, but if simultaneous negative signals arrive, the tubes both cut off and an output negative signal is obtained. As a cathode output stage the circuit has low output impedance and no reversal of polarity of output over input signal.

A diode form of gate is shown in Fig. 9-62. Resistor $R_1$ is made much

greater than $R_2$, so that the circuit is a clipper at essentially zero level for either polarity of input, or gives zero output. However, application to $R_2$ of a gate voltage of the polarity indicated places a bias on the diode such that input pulses are not bypassed to ground and appear in the output. The two diodes provide selective gating according to the desired output-pulse polarity, the left diode passing negative pulses, the right passing positive pulses.

Another use for a switching tube is indicated in Fig. 9-63, where a triode restores, or *clears*, a flip-flop to a desired initial condition. When the grid of $T_3$ is driven from cutoff up to zero, the plate potential of $T_2$ is pulled down. This forces $T_1$ off, if it were conducting, turns $T_2$ on, and turns off the indicator light. Such a circuit might be employed with each flip-flop to clear an arithmetic register of flip-flops after completion of an operation.

**9-32. Diode Switching Matrices.** It is frequently necessary, in computer or counting operations or in the magnetic-core matrix, to be able to energize a single output line for each definite input count of pulses. This is not difficult with relay-operated stepping switches, but they are slow in operation, and electronic methods are needed to keep pace with high-speed electronic decade scalers and computer operations. A suitable circuit is the diode matrix of Fig. 9-64. Controlled by flip-flops to count the input pulses, the circuit as shown will handle a count of four, but may be extended by adding additional flip-flops in binary or decade form and providing additional horizontal and vertical switching leads and diodes.

The matrix is composed of horizontal and vertical leads interconnected by diodes having low forward resistance (direction of arrow in diode symbol) and an extremely high back resistance. Crystal diodes are usually employed because of small space requirements and lack of heater power.

A tube in the flip-flop will be considered in the 0 state when nonconducting, with its plate voltage equal to the supply voltage shown as +300 volts. In the 1 state the tube is conducting, and the plate voltage will be low and should be made less than the +100 volts of the matrix supply bus. This bus voltage must be chosen less than $E_{bb}$ also. Thus a diode connected to horizontal line $H_1$ will not conduct if $T_1$ is 0 or nonconducting, the diode being back-biased, or the potential at $A$ is more positive than the vertical bus value of +100 volts. However, if $T_1$ is conducting or in the 1 state, its anode potential will be less than that on the vertical lead and the diodes connected to $H_1$ will conduct, since they point into a lower potential at $A$.

Assume the flip-flops initially in the 00 state, $T_1$ and $T_3$ nonconducting, $T_2$ and $T_4$ conducting. Points $B$ and $D$ will be at low potential, and the four diodes connected to lines $H_2$ and $H_4$ will be conducting.

These conducting diodes effectively shunt output resistors $R_1$, $R_2$, and $R_4$ and cause the voltage on lines $V_1$, $V_2$, and $V_3$ to be low. At the same time points $A$ and $C$ will be at the $E_{bb}$ value of 300 volts, and the diodes connected to $H_1$ and $H_3$ will be nonconducting. Line $V_0$ is the only vertical line not connected to a conducting diode. Examination of the diagram then shows that the bus supply voltage of +100 appears across $R_0$ and on line $V_0$ only when the flip-flops are in the 00 state.

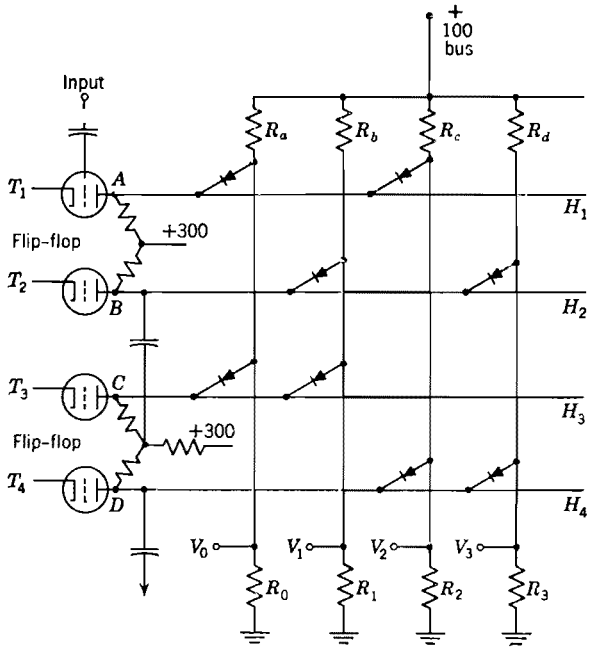



FIG. 9-64 Diode matrix for switching, controlled by flip-flops.

A single pulse may now be inserted at the input, setting the flip-flops to the 01 state, with $T_1$ and $T_4$ conducting. Diodes connected to $H_1$ and $H_4$ are conducting; diodes connected to $H_2$ and $H_3$ are nonconducting. The only vertical lead not connected by conducting diodes to $H_1$ or $H_4$ is $V_1$; thus the matrix bus voltage appears across $R_1$ and on $V_1$ only, for the 01 state.

Apply a second pulse setting the flip-flops to the 10 state with $T_2$ and $T_3$ conducting. Diodes connected to $H_2$ and $H_3$ are conducting, shorting $R_0$, $R_1$, and $R_3$, and the only vertical lead not connected to such a diode is $V_2$; thus the matrix bus voltage appears on $V_2$. It can be reasoned that after a third pulse to the 11 state the output will appear only on $V_3$. A fourth pulse will restore the 00 state and transmit an output pulse to further flip-flops if desired.

It is thus possible to distribute the bus voltage to any of four lines in this example, one at a time, by setting up the proper binary number in the flip-flops. This can be extended to higher numbers; for example. with eight flip-flops and proper diode connections the bus voltage might be switched to any one of 256 lines simply by inserting the proper number of pulses into the counters.

**9-33. And, Or, Not Circuits.** As will be seen in the next section, in order properly to combine the data entered in the flip-flop registers of a

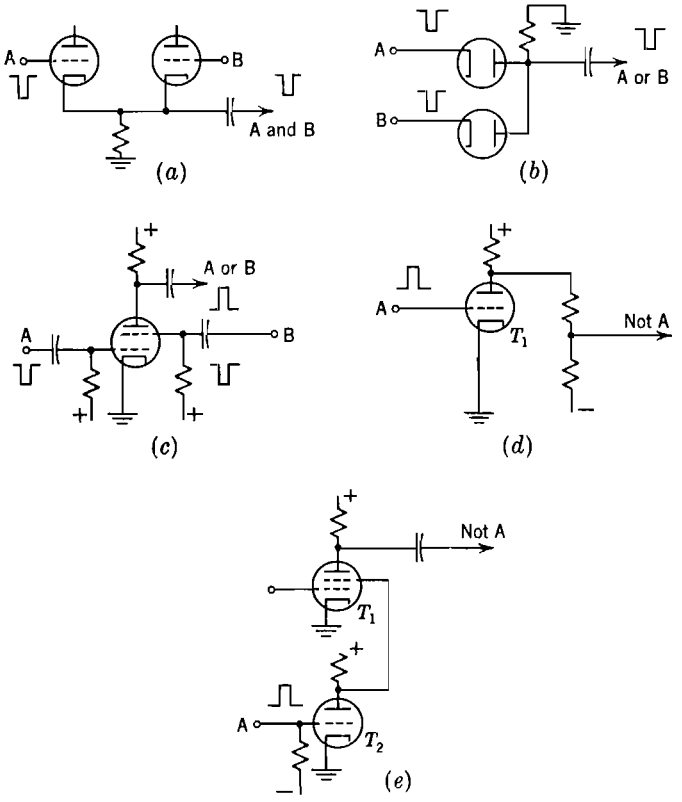



FIG. 9-65. (*a*) *And* circuit. (*b*, *c*) *Or* circuits. (*d*, *e*) *Not* circuits.

computer, it is necessary to provide certain *logical circuits* to obtain and use correctly the sum and carry digits or other results of an arithmetical operation. It has been found that these operations can be carried out if electronic circuits to perform the simple commands of *and*, *or*, *not* are available.

The statement *A and B* implies that, if a 1 appears at both *A* and *B*, an output will be given. Such an *and* circuit would have two inputs,

$A$ and $B$, and would give an output if a 1 signal pulse appeared at both inputs simultaneously. This is recognized as the function of any of the gating circuits of Fig. 9-61, and so these may also be called *logical and* circuits, one form of which is redrawn in Fig. 9-65*a*.

The statement $A$ *or* $B$ implies that, if a 1 appears at either $A$ or $B$, an output will be given. Thus an *or* circuit would have two inputs $A$ and $B$ and would give an output if an input signal appears at either. One form of circuit is shown in Fig. 9-65*b*, wherein a negative input pulse on either input causes a negative output pulse. The dual-grid tube of Fig. 9-65*c* is operated conducting by positive bias on both grids. A negative signal on either grid cuts off the tube, and a positive output pulse is obtained.

A *not* circuit has a signal on its output if there is no input signal and no signal on the output if there is a signal on the input. The circuit of Fig. 9-65*d* inverts the signal and for zero volts input gives a negative output, whereas for a negative input a zero output is given. The circuit of Fig. 9-65*e* operates with $T_2$ normally cut off and positive screen voltage on $T_1$, with $T_1$ conducting. Presence of a positive signal on $T_2$ causes that tube to conduct, dropping the screen potential on $T_1$ and cutting off that tube. Thus a positive signal on $A$ or $T_2$ inhibits an output from $T_1$, or gives a *not* $A$ output.

These circuits and the flip-flop are basic computer building blocks.

**9-34. Addition Logic.** While detailed computer circuits are highly specialized and beyond the scope of this book, it still seems desirable to indicate broadly the manner of operation and the application of some of the circuits discussed. An addition operation seems desirable as an example.

The flip-flop is the basic circuit usually employed in the computer registers, holding numbers supplied from the memory for the purpose of an immediate operation, say, of addition. To perform the addition between numbers held in two registers requires a study of the logical operations involved, and this leads to an application of the *and, or, not* circuits just described.

Consider first the accompanying table resulting from an addition oper-

| Addend $A$ | Augend $B$ | Sum $S_1$ | Carry $C_1$ |
|---|---|---|---|
| 0 | 0 | 0 | 0 |
| 0 | 1 | 1 | 0 |
| 1 | 0 | 1 | 0 |
| 1 | 1 | 0 | 1 |

ation in binary digits. Study of each line shows that the following logic holds:

$$S_1 = (A \text{ or } B) \text{ and } not\ (A \text{ and } B)$$
$$C_1 = A \text{ and } B$$

As in Sec. 9-33, *A or B* indicates a 1 at either $A$ or $B$ but not both, whereas *A and B* indicates that a 1 appears at both $A$ and $B$.

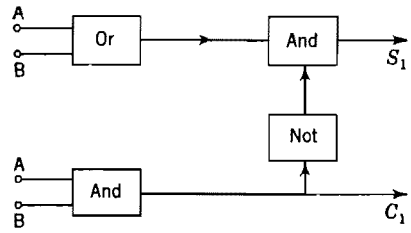



FIG. 9-66. Logical diagram for the first half of an addition operation.

Figure 9-66 illustrates a block diagram for *and*, *or*, *not* circuits which will operate in accordance with the above. These circuits would add pulses from two flip-flops $A$ and $B$, giving a pulse $S_1$ for a third-sum register flip-flop, and would provide a carry pulse $C_1$ to another circuit. Since the full addition including carries is not completed, such a circuit is called a *half-adder*.

Tracing along the upper line of the diagram, it can be seen that an output *A or B* is combined with a *not A and B* signal to give a sum $S_1$. Thus, if $A = 1$, $B = 0$, the *or* circuit will give an output which appears at $S_1$. The *not* circuit does not introduce an inhibitory action because both $A$ and $B$ are not 1.

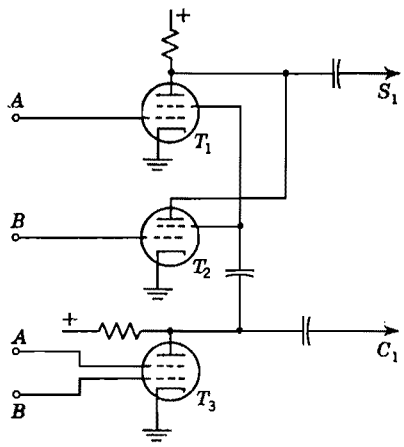



FIG. 9-67. Electronic half-adder.

If $A = 1$, $B = 1$, the lower *and* circuit gives an output pulse transmitted directly to $C_1$. In addition the *not* circuit is operated to inhibit a signal on the upper bus to $S_1$. Thus a zero is correctly transmitted to $S_1$.

An electronic form of the half-adder made up of circuits from Fig. 9-65 appears in Fig. 9-67. The first grids of $T_1$ and $T_2$ serve as the *or* circuit, the second grids acting as the *not* circuits. Tubes $T_1$ and $T_2$ will pass a pulse to $S_1$ if either $A$ or $B$ receives a positive 1 digit. Tube $T_3$ is an *and* tube, normally cut off and requiring positive pulses on both $A$ and $B$ for output. If a 1 pulse appears at both $A$ and $B$, $T_3$ conducts, lowering the second-grid potentials of $T_1$ and $T_2$ and inhibiting an output pulse to $S_1$. At the same time $T_3$ provides an output negative pulse to the carry circuit.

Actually the inputs $A$ and $B$ to $T_1$ and $T_2$ would have to be delayed slightly, so that, in the case of $A = 1$, $B = 1$, the inhibiting pulse could reach the second grids of $T_1$ and $T_2$ before the input pulses reach the first grids.

The development of the second half of the adder can be deduced from the accompanying table. Two other cases, $C_o = 0$, $S_1 = 1$, $C_1 = 1$ and

| Previous carry $C_o$ | First sum $S_1$ | First carry $C_1$ | New sum $S_2$ | New carry $C_2$ |
|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 1 | 0 | 1 |
| 0 | 1 | 0 | 1 | 0 |
| 1 | 0 | 0 | 1 | 0 |
| 1 | 0 | 1 | 1 | 1 |
| 1 | 1 | 0 | 0 | 1 |

$C_o = S_1 = C_1 = 1$, are omitted as impossible.

In the first half-adder a sum $S_1$ and a carry $C_1$ were formed. In the full adder provision has been made for accepting a carry $C_o$ from a preceding digit and for forming a new sum $S_2$ from this carry $C_o$ and the first sum $S_1$. A carry $C_2$ may also result from this summing operation.

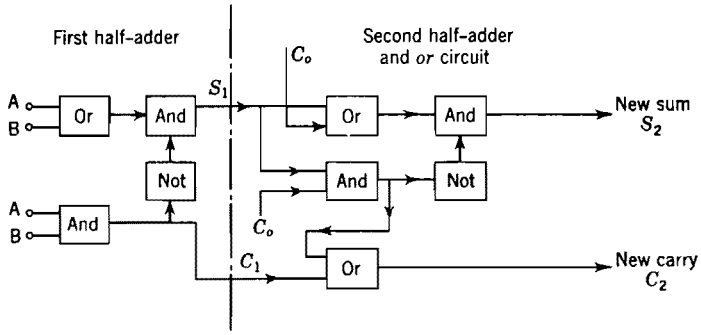



FIG. 9-68. Two half-adders and an *or* circuit for forming a full adder.

The rules of logic as indicated by the table for the second step, the formation of $S_2$ and $C_2$, can be stated as follows:

$$S_2 = (S_1 \textit{ or } C_o) \text{ and } \textit{not } (S_1 \textit{ and } C_o)$$
$$C_2 = (S_1 \textit{ and } C_o) \textit{ or } C_1$$

and it can be seen that these rules are identical in form with the first set except for the statement *or* $C_1$ in the second line. Thus a circuit identical to that needed for the first half-adder, plus the addition of an *or* circuit, will perform the second step in the addition problem.

A block diagram for both halves of the adder is shown in Fig. 9-68 and can be traced through for conformity with the above statements.

All such circuits require occasional pulse-reshaping circuits, so that the waveforms will be exact at critical points throughout the circuit. Most computers use pulses of 30 to 60 volts amplitude, to avoid troubles which might arise due to random noise and hum. To keep these voltage levels exact requires the use of carefully stabilized voltage supplies of considerable current capacity and low output impedance.

## PROBLEMS

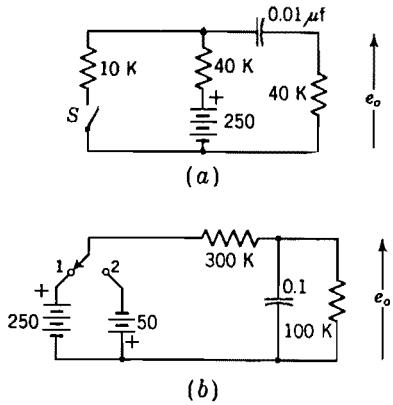
**9-1.** If the waveform of Fig. 9-69*a*, with $t_1 = 200$ μsec, $t_2 = 100$ μsec, $E_1 = -100$ volts, $E_2 = 100$ volts, is applied as $e_i$ to the circuit of Fig. 9-69*b*, plot $e_1$ and $e_2$ out to the end of the transient. Show all waves in proper relation and scale to the applied emf.

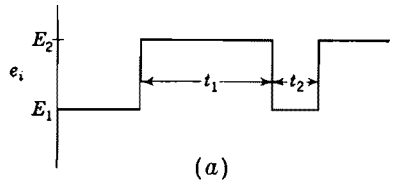
**9-2.** Repeat Prob. 9-1, with $t_1 = 500$, $t_2 = 500$ μsec, $E_1 = 0$, $E_2 = 200$ volts. What is the average value of $e_1$? Of $e_2$?

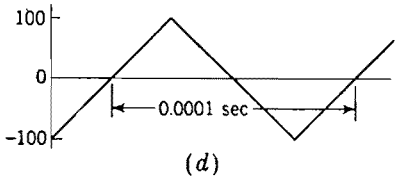
**9-3.** In Fig. 9-70*a* switch *s* has been open for a long time. It is closed at $t = 0$ for a period of 0.001 sec, and again opened. Plot $e_o$ against time. Compute the power loss in the 10,000-ohm resistor.

**9-4.** In Fig. 9-70*b* the switch oscillates between 1 and 2, the duration of contact on each point being 0.5 sec. Plot the waveform of $e_o$.

**9-5.** In Fig. 9-70*c* the duration of the switch contact on each of 1 and 2 is 500



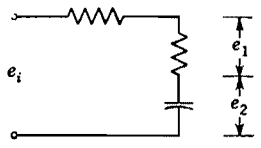




FIG. 9-69

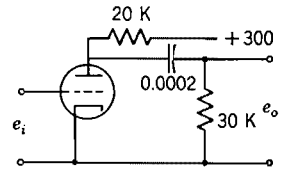



FIG. 9-70

$\mu$sec. Assume $r_c = 0$ when the grid is positive, infinite when negative, and plot grid-to-cathode voltage $e_c$ as a function of time.

**9-6.** An $RC$ integrating circuit of value $R = 100K$, $C = 0.01$ $\mu$f has applied to it the triangular wave of Fig. 9-70*d*. Plot the output waveform if the wave is symmetrical and has a period of 0.0001 sec.

**9-7.** The waveform of Fig. 9-70*d* as specified in Prob. 9-6 is applied to a differentiating circuit of value $RC = 100$ $\mu$sec. Plot the output waveform.

**9-8.** In the circuit of Fig. 9-19 $R_g = 1$ megohm, $R = 100{,}000$ ohms, $R_L = 30{,}000$ ohms, $E_{bb} = 300$ volts, and the tube is one section of a 12AU7. Assume $r_c = 1000$ ohms with positive grid. Sketch the plate voltage to scale, for an input of 100 volts rms, 1000 cycles.

**9-9.** A constant-current sweep uses a pentode to charge the capacitor. The discharge thyratron has an extinction potential of 15 volts and a firing potential of 110 volts. If the constant pentode current is 8 ma, find the value of charging capacitor needed for a sweep-repetition rate of 1800 per second, assuming discharge time negligible.

**9-10.** A trapezoidal voltage $e = K_1 + K_2 t$ is applied to a series $RL$ circuit. If $i_L = 0$ at $t = 0$, find an expression for the resulting current.

**9-11.** The circuit of Fig. 9-70*e* is operated with input a rectangular wave having zero volts for 200 $\mu$sec and $-100$ volts for 200 $\mu$sec. The tube is a 6C4 with $\mu = 17$. Plot $e_o$ to scale over one cycle of input.

**9-12.** The circuit of Fig. 9-71*a* is triggered with a single positive pulse on the grid of $T_1$. Find (*a*) the bias on $T_2$ before the trigger, (*b*) the bias on $T_1$ before the trigger, (*c*) the bias on $T_1$ during the period of the pulse, (*d*) the duration of the output pulses formed between plate and ground of $T_2$.

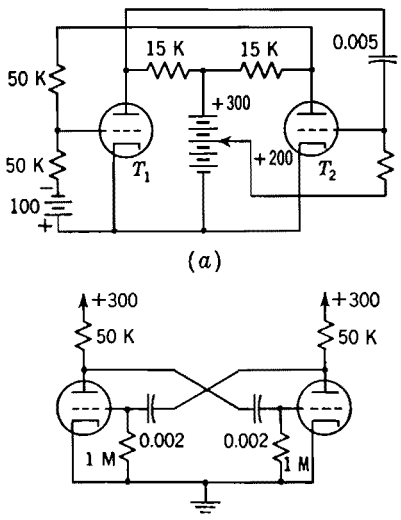



FIG. 9-71

**9-13.** Design a free-running multivibrator to operate at 1000 cycles. Assume $R_L = 10K$, $R_g = 500K$, 6C4 tubes, and $E_{bb} = 200$ volts.

**9-14.** Design a free-running unbalanced multivibrator to operate at 5000 cycles, with one pulse three times as long as the other. Assume the conditions of Prob. 9-13.

**9-15.** Using the multivibrator circuit of Fig. 9-71*b*, with $r_c = 500$ ohms for positive grid, determine $r_p$ for $e_c = 0$ from the characteristics for the 6C4 tubes.

(*a*) Find the discharge time constant for the capacitor circuits.

(*b*) Assuming the conducting-tube grid voltage to be zero, find the peak value of voltage on the other grid.

(*c*) Calculate the period of cutoff.

(*d*) Calculate the repetition frequency.

**9-16.** (*a*) Write the number 1274 in binary notation.

(*b*) Repeat in octal notation; in trinary notation.

**9-17.** (*a*) Write the numbers 1022 and 817 in binary notation, and add; check the result by decimal methods.

(*b*) Repeat with multiplication of the same numbers.

**9-18.** Divide 623 by 37 by binary methods. HINT: Carry out the process by methods identical to manual long division.

**9-19.** Draw a circuit diagram so as to extend the diode matrix of Fig. 9-64 to operate with signals up to binary 1010.

**9-20.** Following Fig. 9-67, devise an electronic circuit for a full adder, following the block diagram of Fig. 9-68.

**9-21.** You will operate two dpdt switches in such a way that the characteristics of the entire population passing a point on a downtown street are recorded as to male or female, over or under 5 ft 10 in., blond or brunet, and blue-eyed or not. Then devise a circuit which will ring a bell each time a blue-eyed blonde girl over 5 ft 10 in. passes.

## REFERENCES


1. Von Tersch, L. W., and A. W. Swago: "Recurrent Electrical Transients," Prentice-Hall, Inc., Englewood Cliffs, N.J., 1953.
2. Chance, B., V. Hughes, E. F. MacNichol, D. Sayre, and F. C. Williams: "Waveforms," McGraw-Hill Book Company, Inc., New York, 1949.
3. Mather, N. W.: Clipping and Clamping. *Electronics*, **20**: 111 (July, 1947).
4. Frank, Ernest: "Pulsed Linear Networks," McGraw-Hill Book Company, Inc., New York, 1945.
5. Batcher, R. R.: Circular and Polar Sweeps, *Electronic Inds.*, **3**: 80 (September, 1944).
6. Clarke, A. C.: Linearity Circuits, *Wireless Engr.*, **21**: 256 (June, 1944).
7. Chance, Britton: Some Precision Circuit Techniques Used in Waveform Generation and Time Measurement, *Rev. Sci. Instr.*, **17**: 396 (October, 1946).
8. Cocking, W. T.: Linear Sawtooth Oscillators, *Wireless World*, **16**: 425 (May, 1939).
9. Herrenden Harker, G. F.: The Mechanism of Synchronization in the Linear Time Base, *Phil. Mag.*, **26**: 193 (1938).
10. MacGregor-Morris, J. T., and J. A. Henley: "Cathode Ray Oscillography," Chapman and Hall, Ltd., London, 1936.
11. Potter, J. L.: Sweep Circuits, *Proc. IRE*, **26**: 713 (1938).
12. Puckle, O. S.: "Time Bases," John Wiley & Sons, Inc., New York, 1943.
13. Soller, J. T., M. A. Starr, and G. E. Valley, Jr.: "Cathode Ray Tube Displays," McGraw-Hill Book Company, Inc., New York, 1948.
14. Abraham, H., and E. Bloch: Notice sur les lampes-valves à 3 électrodes et leurs applications, Ministère guerre *Publ.* 27, April, 1918.
15. Mather, N. W.: Multivibrator Circuits, *Electronics*, **19**: 136 (October, 1946).
16. Burks, A. W.: Electronic Computing Circuits of the ENIAC, *Proc. IRE*, **35**: 756 (1947).
17. Page, C. H.: Digital Computer Switching Circuits, *Electronics*, **21**: 110 (September, 1948).
18. Eccles, W. H., and F. W. Jordan: Trigger Relay Utilizing Three Element Thermionic Vacuum Tubes, *Radio Rev.*, **1**: 143 (October, 1919).
19. Engineering Research Associates: "High-speed Computing Devices," McGraw-Hill Book Company, Inc., New York, 1950.
20. West, C. F., and J. F. DeTurk: A Digital Computer for Scientific Applications, *Proc. IRE*, **36**: 1452 (1948).
21. *Proceedings of the IRE*, Computer Issue, **41**: 1219–1568 (1953).
22. Richards, R. K.: "Arithmetic Operations in Digital Computers." D. Van Nostrand Company, Inc., Princeton, N.J., 1955.

CHAPTER 10

# POWER SUPPLIES AND FILTERS; VOLTAGE REGULATORS

Since almost all electric power is generated as alternating current, it is necessary to rectify that current when d-c supplies for amplifiers and electronic equipment are desired. This is the common practice for the supply of plate potentials for electron tubes, vacuum-diode rectifiers usually being employed.

**10-1. The Half-wave Diode Rectifier.** A vacuum diode will conduct only with positive anode, so that a circuit such as that of Fig. 10-1*a*, supplied by an alternating voltage, will be a *half-wave rectifier*. This circuit is the basic form from which more complex circuits are derived and is studied here as an example of rectifier analysis.

The diode will have a nonlinear volt-ampere relation as shown by the static curve of Fig. 10-1*b*. The diode resistance is measured as the reciprocal of the slope of this curve and is not a constant. When the diode is in series with a load $R$, the total circuit resistance is $r_p + R$ and a volt-ampere curve for the circuit, relating current to the voltage applied to the circuit, is much more linear than that of the tube alone. For ease of analysis an assumption of linearity for the static curve is usually made, with small error, and under this assumption $r_p = r_b$ for the diode.

The value of $r_b$ for a particular diode may be found by taking the reciprocal of the slope of a straight line passing through the origin and approximating the diode volt-ampere curve.

For the half-wave circuit of Fig. 10-1*a* the current for an input

$$e = E_m \sin \omega t$$

may be written under the linearity assumption as

$$i_b = \frac{E_m}{r_b + R} \sin \omega t \qquad 0 \leq \omega t \leq \pi \tag{10-1}$$

$$i_b = 0 \qquad \pi \leq \omega t \leq 2\pi \tag{10-2}$$

and the resultant half-sinusoid pulses are shown in Fig. 10-1*b*.

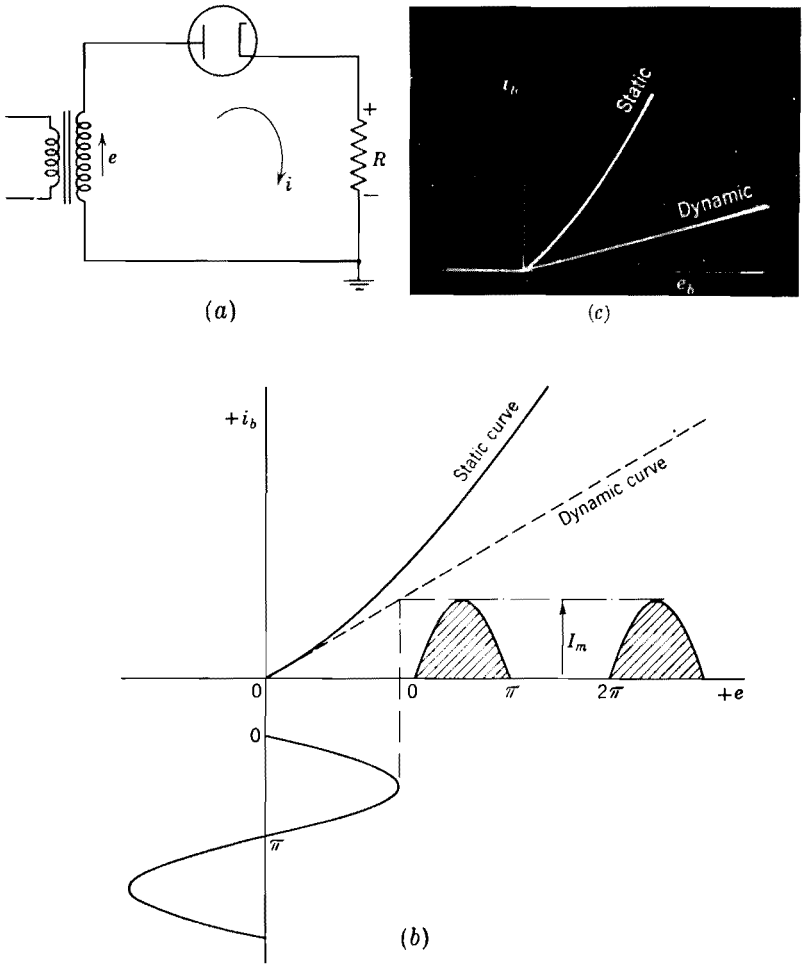



FIG. 10-1. (*a*) Half-wave circuit. (*b*) Graphical determination of waveform. (*c*) Experimental characteristic of diode, static and dynamic, with $R = 2000$ ohms.

The function of a rectifier is to produce direct current from an alternating source. Its ability to do so can be demonstrated by determining the d-c component, or average value, of the above current pulses, given by

$$I_{dc} = \frac{1}{2\pi}\int_o^{\pi} i_b \, d\omega t = \frac{1}{2\pi}\left(\int_o^{\pi} \frac{E_m}{r_b + R}\sin \omega t \, d\omega t + \int_{\pi}^{2\pi} 0 \, d\omega t\right)$$
$$= \frac{1}{\pi}\frac{E_m}{r_b + R} = \frac{I_m}{\pi} \qquad (10\text{-}3)$$

Equation (10-3) serves also to define the symbol $I_m$.

From Eq. (10-3)

$$I_{dc}(r_b + R) = \frac{E_m}{\pi}$$

from which the d-c voltage on the load is

$$E_{dc} = I_{dc}R = \frac{E_m}{\pi} - I_{dc}r_b \tag{10-4}$$

At no load, with $I_{dc} = 0$, the direct load voltage is the average of the positive half of the input sine wave, or $E_m/\pi$. As the load current increases, the d-c voltage will fall linearly, owing to the $I_{dc}r_b$ drop in the rectifier tube.

The transformer supplies an a-c power given by

$$P_{ac} = I_{rms}^2(r_b + R) \tag{10-5}$$

where $I_{rms}$ is the effective value of the load-current pulses, given by definition as

$$I_{rms} = \sqrt{\frac{1}{2\pi}\int_o^\pi i_b^2\, d\omega t} = \sqrt{\frac{1}{2\pi}\int_o^\pi I_m^2 \sin^2 \omega t\, d\omega t} = \frac{I_m}{2} \tag{10-6}$$

This is the value read by an a-c ammeter in series with the load.

A Fourier analysis of the half-sinusoid voltage pulses applied to the load gives

$$e = \frac{E_m}{\pi} + \frac{E_m}{2}\sin \omega t - \frac{2E_m}{\pi}\sum_{\text{even } k} \frac{\cos k\omega t}{(k+1)(k-1)} \tag{10-7}$$

where $k$ has values 2, 4, 6, 8, . . . . All these voltage terms produce corresponding current terms in the load. The first term on the right is recognized as the d-c or average value, effective in charging a battery or in an electrolytic process. The remaining terms are a-c harmonics, the currents producing losses in the resistive load but not being available as direct current. The lowest a-c frequency present is that of the supply.

It is desired that a rectifier produce direct current and minimize the losses due to flow of a-c harmonics in the load. The efficiency with which alternating current is converted to direct current is measured by a defined term, the *efficiency of rectification*, stated as

$$\eta_R = \frac{P_{dc}\text{ output}}{P_{ac}\text{ input}} \times 100\% \tag{10-8}$$

which for the half-wave diode becomes

$$\eta_R = \frac{[E_m/\pi(r_b + R)]^2 R \times 100\%}{[E_m/2(r_b + R)]^2(r_b + R)} = \frac{4R}{\pi^2(r_b + R)} \times 100\%$$
$$= \frac{1}{(1 + r_b/R)} \times 40.6\% \tag{10-9}$$

For the half-wave diode the presence of harmonics in the load limits the maximum efficiency to 40.6 per cent. The term $1 + r_b/R$ indicates the effect of diode resistance on efficiency, and the ratio $r_b/R$ should be small.

Circuits which separate the a-c and d-c components and bypass the a-c components around the load, or prevent their generation, are called *filters*. Filter circuits increase the efficiency of rectification toward the 100 per cent level and also provide smoother direct current, with smaller a-c components. Such well-filtered output is required for most electronic apparatus.

The *peak inverse voltage* is the maximum voltage appearing across the tube during the time in which the anode is negative. This value occurs at the negative sine-wave peak and for the half-wave circuit has the value of $PIV = E_m$. Tubes must be designed with sufficient insulation to withstand this voltage.

**10-2. The Full-wave Rectifier.** If the load could be supplied with current during both halves of the input a-c cycle, it would appear that greater d-c power output and increased rectification efficiency might be obtained. This result is achieved in the *full-wave circuit* of Fig. 10-2*a*, the most commonly used form of rectifier. The result is obtained by use of a second-diode half-wave circuit and a common load, the a-c input to the second diode being at 180° with respect to that of the first circuit. Since conduction occurs only during the half cycle in which an anode is positive, tubes $T_1$ and $T_2$ conduct on alternate half cycles, resulting in a current wave as in Fig. 10-2*b*.

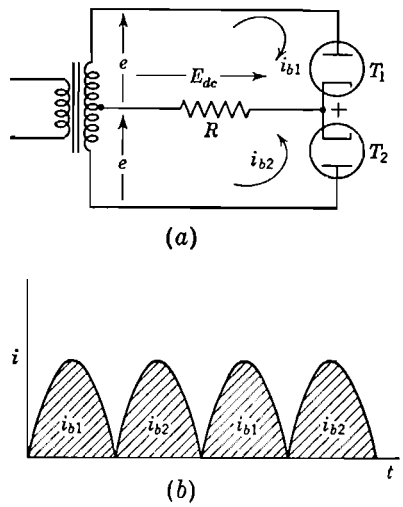



FIG. 10-2. (*a*) Full-wave rectifier. (*b*) Current pulses in load.

Assuming identical tubes, the respective tube currents can be written as

$$\left.\begin{aligned} i_{b1} &= \frac{E_m}{r_b + R} \sin \omega t \\ i_{b2} &= 0 \end{aligned}\right\} 0 \leq \omega t \leq \pi \qquad (10\text{-}10)$$

$$\left.\begin{aligned} i_{b1} &= 0 \\ i_{b2} &= -\frac{E_m}{r_b + R} \sin \omega t \end{aligned}\right\} \pi \leq \omega t \leq 2\pi \qquad (10\text{-}11)$$

The voltage $E_m$ is the peak of the sine wave for one-half of the transformer secondary, or $e = E_m \sin \omega t$ as before.

Each diode is then operating under identical conditions as in the half-wave circuit the load currents being combined in the common load. The d-c current can be found from the average of the half sinusoids of Fig. 10-2*b*, as

$$I_{dc} = \frac{2E_m}{\pi(r_b + R)} = \frac{2I_m}{\pi} \tag{10-12}$$

The output-voltage expression may be obtained as

$$E_{dc} = I_{dc}R = \frac{2E_m}{\pi} - I_{dc}r_b \tag{10-13}$$

which shows the no-load voltage as $2E_m/\pi$, or twice that of the half-wave circuit for the same value of $E_m$. The regulation is also improved over that of the half-wave circuit.

The value of $I_{rms}$ for each diode, and the a-c power input to each diode, is the same as for the half-wave circuit. The total a-c power input then is

$$P_{ac} = 2\left[\frac{E_m}{2(r_b + R)}\right]^2 (r_b + R) = \frac{{E_m}^2}{2(r_b + R)}$$

and the efficiency of rectification becomes

$$\eta_R = \frac{8}{\pi^2}\frac{1}{1 + r_b/R} \times 100\% = \frac{1}{1 + r_b/R} \times 81.2\% \tag{10-14}$$

which is twice that for the half-wave circuit, with a maximum of 81.2 per cent.

A Fourier analysis of the waveform of Fig. 10-2*b* leads to

$$e = \frac{2E_m}{\pi} - \frac{4E_m}{\pi}\sum_{\text{even } k}\frac{\cos k\omega t}{(k+1)(k-1)} \tag{10-15}$$

where $k$ has the even values 2, 4, 6, 8, . . . . The lowest frequency present is double the supply frequency, and this makes easier the duty of a filter in removal of the a-c harmonics. Where large currents are involved, the easier filtering and higher efficiency make the full-wave circuit more desirable than the half-wave form.

Writing an instantaneous potential equation around the outside loop of the circuit of Fig. 10-2*a*, with tube $T_1$ nonconducting, gives

$$e_{b1} = 2e + e_{b2} = -2E_m \sin \omega t + e_{b2}$$

The maximum value, or peak inverse voltage (PIV), for $T_1$ occurs when $\omega t = 3\pi/2$; if $E_m \gg e_{b2}$, then

$$\text{PIV} = 2E_m$$

and this is a criterion for selection of a tube with sufficient breakdown strength.

It should be noted that in most such circuits the negative terminal is grounded, placing the diode cathode at full d-c potential above ground. Filament transformers must be insulated for this potential, and if heater tubes are used, the cathode-heater insulation must be sufficient to prevent breakdown.

**10-3. Ripple Factor.** The pulsations, or *ripple*, caused by the a-c components in the rectifier output, appearing as audible hum in sound amplifiers, are very undesirable in the supply to most electronic circuits. The amount of ripple, compared with the direct component of current or voltage, is a measure of the purity of rectifier output and is called the *ripple factor*.

$$\text{Ripple factor} = \frac{\text{effective value of all a-c components}}{\text{average, or d-c, component}} \tag{10-16}$$

The load current of a rectifier is composed of the direct component $I_{dc}$ and an effective value for all the a-c components $I_{ac}$. By definition the effective value $I_{rms}$ of the total load current is then

$$I_{rms} = \sqrt{I_{dc}^2 + I_{ac}^2}$$

from which

$$I_{ac} = \sqrt{I_{rms}^2 - I_{dc}^2} \tag{10-17}$$

This supplies the numerator for the ripple definition in Eq. (10-16), so that in general

$$\text{Ripple factor} = \frac{\sqrt{I_{rms}^2 - I_{dc}^2}}{I_{dc}} = \sqrt{\left(\frac{I_{rms}}{I_{dc}}\right)^2 - 1} \tag{10-18}$$

The terms in Eq. (10-18) may be measured by a-c and d-c types of instruments and the ripple factor determined by experiment when the ripple is large. More refined measurements are required when the ripple is small and $I_{dc}$ nearly equals $I_{rms}$.

It can be seen that ripple factor is a function of the waveform of the current in the rectifier load. If $I_{dc} = I_{rms}$, the ripple would be zero and rectifier tubes would supply successive rectangular blocks of current to the load.

Use of values of $I_{dc}$ and $I_{rms}$ from the previous discussions for the half-wave and full-wave rectifiers allows ripple-factor computations for these circuits without filters, giving

Ripple factor:

Half-wave = 1.21
Full-wave = 0.48

The full-wave circuit has less ripple, and is more easily filtered owing to the higher ripple frequency.

**10-4. The Shunt-capacitor Filter.** A simple form of filter, in which the a-c harmonics are bypassed around the load by a capacitor, is drawn in Fig. 10-3 for a full-wave rectifier. Proper operation requires that $X_C \ll R$ for the lowest frequency present in the rectifier output.

The presence of the capacitor considerably affects the diode operating conditions, since the rectifier is no longer operating with a resistance load. It may be considered that when the rectifier input voltage is above that of the capacitor then the capacitor charges, reaching the peak voltage of the supply. When the input voltage falls, the diode disconnects the source since the diode will have a positive cathode, and the stored energy in the capacitor maintains the load voltage at a high value over a longer period than if no capacitor were used. The diode current flows in short pulses, rather than during the whole positive half cycle. The situation is illustrated in Fig. 10-4.

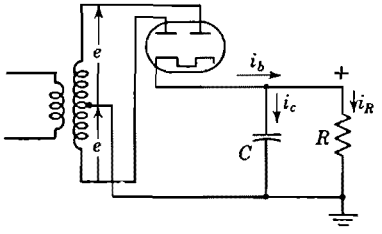



FIG. 10-3. Full-wave circuit with capacitor filter.

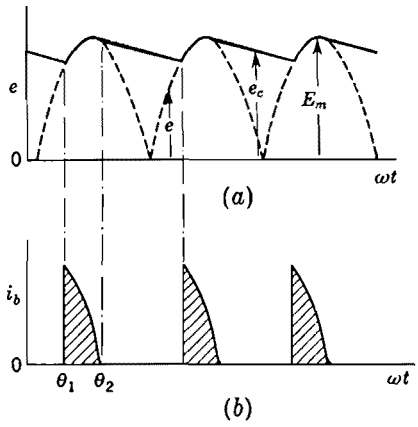



FIG. 10-4. (a) Voltage waveform. (b) Tube-current waveform.

For good efficiency $r_b \ll R$, and for simplicity of analysis it will be desirable to assume $r_b = 0$. The leakage reactance of the input transformer will also be neglected as small. Under these assumptions the diode serves merely as a switch and connects the source to the capacitor only during the charging interval from $\omega t = \theta_1$ to $\omega t = \theta_2$ in Fig. 10-4*b*.

It can be seen that during this time interval

$$i_b = i_C + i_R \tag{10-19}$$

and since the diode switch is closed,

$$e_C = E_m \sin \omega t \tag{10-20}$$

Then

$$i_b = C \frac{de_C}{dt} + \frac{e_C}{R}$$

so that during the charging interval

$$i_b = \omega C E_m \cos \omega t + \frac{E_m}{R} \sin \omega t \qquad \theta_1 \leq \omega t \leq \theta_2 \tag{10-21}$$

which may be reduced to the form

$$i_b = \frac{E_m}{R}\sqrt{\omega^2R^2C^2 + 1}\sin(\omega t + \phi) \qquad \theta_1 \leq \omega t \leq \theta_2 \qquad (10\text{-}22)$$

and

$$\phi = \tan^{-1}(\omega RC) \qquad (10\text{-}23)$$

as an expression for the current pulse through the diode during the conducting interval.

At $\omega t = \theta_2$ the voltage of the supply is falling faster than the capacitor can discharge through the load, the supply voltage becomes less than $e_C$, and the diode cathode becomes positive. The diode ceases conduction, disconnecting the source and load. At the instant of disconnection, at $\theta_2$, $i_b = 0$, and Eq. (10-21) leads to

$$-\omega CE_m \cos\theta_2 = \frac{E_m}{R}\sin\theta_2$$

from which

$$\theta_2 = \tan^{-1}(-\omega RC) \qquad (10\text{-}24)$$

Equation (10-22) introduces the dimensionless parameter $\omega RC$. If the time constant $RC$ is large, then the tube current peak will be high but the angle $\theta_2$ will decrease toward 90°, thus shortening the conduction angle. A small value of $\omega RC$ will reduce the tube current peak and will lengthen the angle of conduction. The large peak current for large $\omega RC$ may cause damage to the diode cathode, especially in the case of gas diodes.

Figure 10-5 shows the shortened conduction angle and high peak-current characteristic of a large $\omega RC$ value. Tube drop and transformer reactance present in an actual rectifier serve to reduce the expected abrupt rise on the leading edge of the current pulse and to extend the conduction angle.

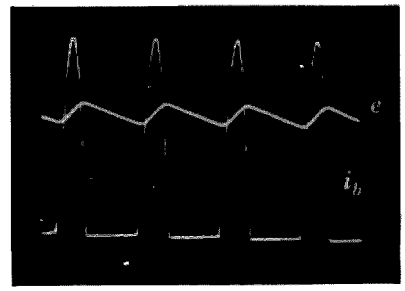



FIG. 10-5. Tube-current pulses and load voltage $e$, full-wave circuit, $E_{dc}$ = 130 volts, $\omega RC$ = 30.

In the interval $\theta_2 \leq \omega t \leq (\pi + \theta_1)$ it can be seen that

$$-i_C = i_R$$

or the capacitor supplies the load current. Then

$$-C\frac{de_C}{dt} = \frac{e_C}{R} \qquad (10\text{-}25)$$

Solution of this differential equation leads to

$$e_C = A\epsilon^{-t/RC}$$

It has already been stated that, at $\omega t = \theta_2$, $e_C = E_m \sin \theta_2$ so that $A$ may be evaluated, giving

$$e_C = E_m \sin \theta_2 \epsilon^{-(\omega t - \theta_2)/\omega RC} \qquad \theta_2 \leq \omega t \leq (\pi + \theta_1) \tag{10-26}$$

as the value of capacitor and load voltage during the discharge interval. This indicates that for no load, or $\omega RC = \infty$, the exponential term becomes unity and the output voltage is constant at the peak value $E_m$.

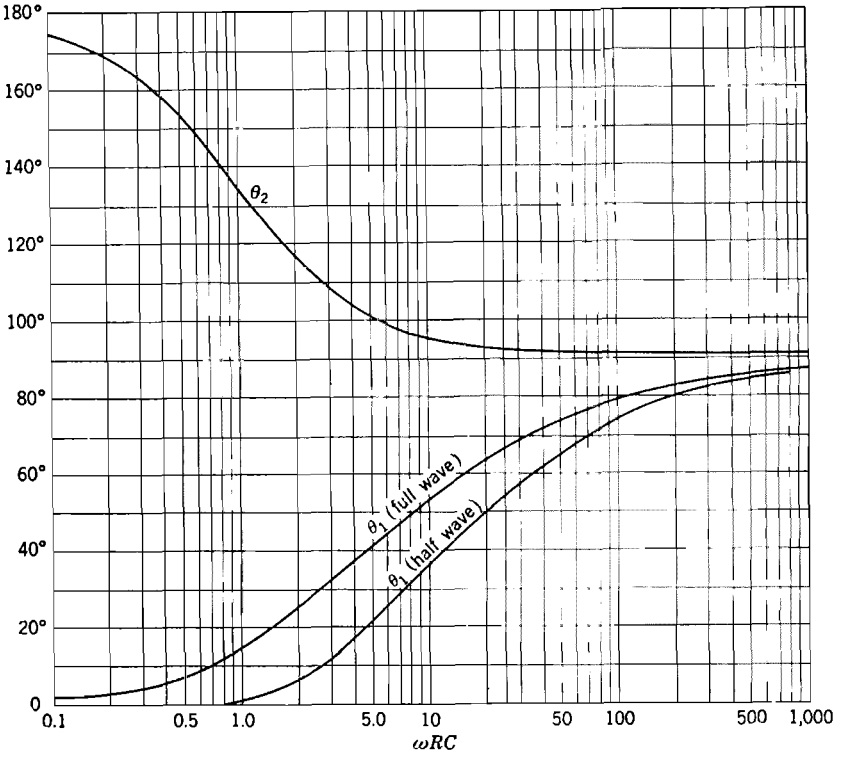



FIG. 10-6. Cutin point $\theta_1$ and cutoff point $\theta_2$ for capacitor filters.

For smaller $\omega RC$ values the exponential term causes the voltage to fall during discharge, introducing ripple into the output.

When the diode once more switches the transformer into the circuit at $\pi + \theta_1$, the capacitor voltage must equal the supply voltage, or

$$E_m \sin \theta_2 \epsilon^{-(\pi + \theta_1 - \theta_2)/\omega RC} = E_m \sin \theta_1$$
$$\sin \theta_1 = \sin \theta_2 \epsilon^{-(\pi + \theta_1 - \theta_2)/\omega RC} \tag{10-27}$$

Angle $\theta_2$ may be determined from Eq. (10-24). Equation (10-27) may then be solved graphically. Results for $\theta_1$ and $\theta_2$ for the full-wave rectifier, in terms of the parameter $\omega RC$, are plotted in Fig. 10-6.

It may be noted that Eqs. (10-23) and (10-24) show

$$\phi = \tan^{-1}(\omega RC) = \pi - \tan^{-1}(-\omega RC) = \pi - \theta_2$$

so that Eq. (10-22) for the waveform of current through the tube may be rewritten as

$$i_b = \frac{E_m}{R}\sqrt{\omega^2 R^2 C^2 + 1}\sin(\theta_2 - \omega t) \qquad \theta_1 \leq \omega t \leq \theta_2 \qquad (10\text{-}28)$$

Using data from Fig. 10-6, it is possible to determine tube-current waveforms, and several are plotted in Fig. 10-7 for typical $\omega RC$ values.

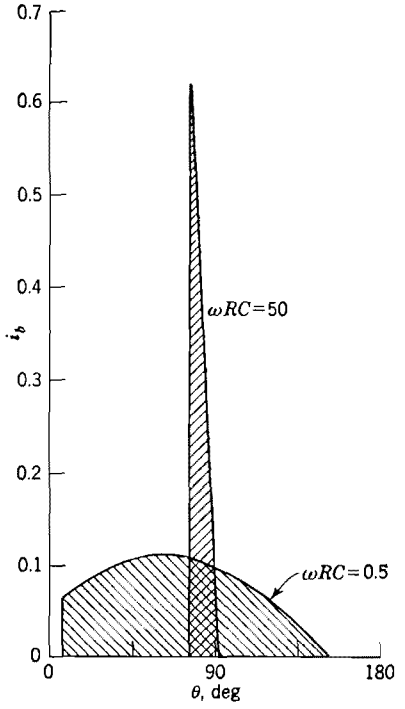



FIG. 10-7. Tube-current waveforms.

The d-c output voltage may be obtained by averaging Eqs. (10-20) and (10-26) over their respective portions of the rectifier cycle as

$$E_{dc} = \frac{1}{\pi}\int_{\theta_1}^{\theta_2} E_m \sin\omega t\, d\omega t + \frac{1}{\pi}\int_{\theta_2}^{\pi+\theta_1} E_m \sin\theta_2 \epsilon^{-(\omega t - \theta_2/\omega RC)}\, d\omega t$$

leading to

$$E_{dc} = E_m \frac{\sqrt{\omega^2 R^2 C^2 + 1}}{\pi}[1 - \cos(\theta_2 - \theta_1)] \qquad (10\text{-}29)$$

The ratio $E_{dc}/E_m$ for the full-wave rectifier with shunt capacitor filter is plotted in Fig. 10-8.

Ripple factor is difficult to calculate using the above methods, and an approximation to the rectifier output waveform is presented in Fig. 10-9*a* which leads to simplified and sufficiently accurate expressions for $E_{dc}$ and ripple factor. The approximation is supported by the oscillogram of Fig.

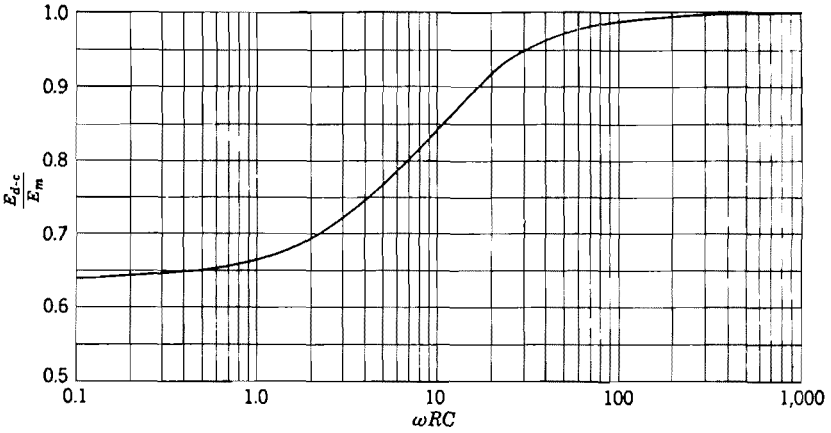
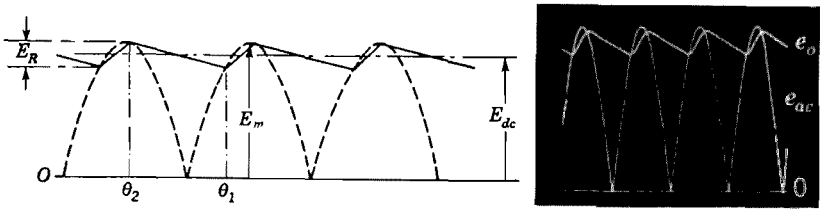



FIG. 10-8. Variation of $E_{dc}/E_m$ for shunt-capacitor filter on full-wave rectifier.




(*a*) (*b*)

FIG. 10-9. (*a*) Approximate load-voltage waveform on full-wave shunt-capacitor filter. (*b*) Actual output voltage $e_o$ and input a-c voltage $e_{ac}$ for full-wave rectifier; $\omega RC = 30$.

10-9*b*, which differs from theory largely because of the neglected tube resistance and transformer leakage reactance.

The approximate waveform consists of straight lines and has a peak value of $E_m$. The total change in capacitor voltage, or the ripple peak-to-peak value, is shown as $E_R$. The average, or d-c, value of such a triangular wave is halfway between peaks, or

$$E_{dc} = E_m - \frac{E_R}{2}$$

neglecting tube resistance and transformer reactance. During the discharge interval, $\theta_2$ to $\pi + \theta_1$, the capacitor is assumed to lose charge at a

constant rate and constant current $I_{dc}$. Since $e_C = q/C$, the uniform rate of loss of potential is

$$\frac{de_C}{dt} = \frac{E_R}{\pi + \theta_1 - \theta_2} = \frac{1}{\omega C}\frac{dq}{d\omega t}$$

and since $dq/d\omega t = I_{dc}$, then

$$E_R = \frac{(\pi + \theta_1 - \theta_2)I_{dc}}{\omega C} \tag{10-30}$$

The effective, or rms, value of the ripple component of the triangular wave is independent of the slopes or lengths of the straight lines, being

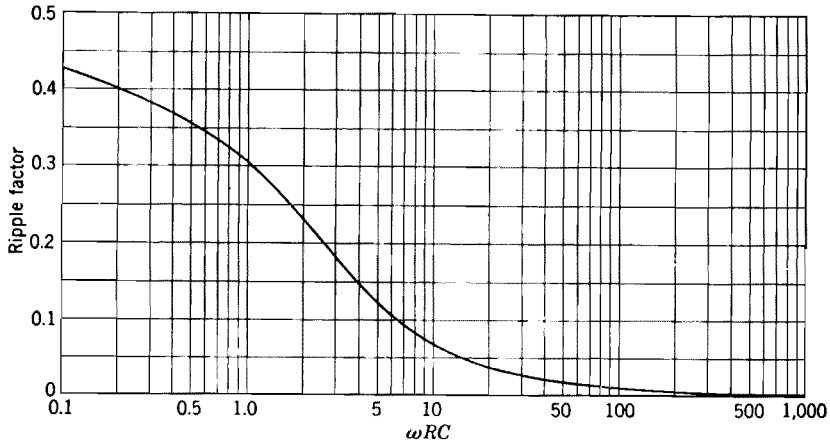



FIG. 10-10. Variation of ripple factor: full-wave rectifier with shunt-capacitor filter.

dependent only on the peak value. For a time axis along $E_{dc}$ it can be shown that the rms value of the a-c components is

$$E_{ac} = \frac{E_R}{2\sqrt{3}} = \frac{(\pi + \theta_1 - \theta_2)I_{dc}}{2\sqrt{3}\,\omega C} \tag{10-31}$$

Use of the definition of ripple factor gives

$$\text{Ripple factor} = \frac{E_{ac}}{E_{dc}} = \frac{(\pi + \theta_1 - \theta_2)I_{dc}}{2\sqrt{3}\,\omega C I_{dc} R} = \frac{\pi + \theta_1 - \theta_2}{2\sqrt{3}\,\omega C R} \tag{10-32}$$

for the full-wave rectifier with shunt-capacitor filter. A plot of this factor as a function of $\omega RC$ appears in Fig. 10-10.

The value of ripple increases with increase of load current (reduced $R$). For good voltage regulation Fig. 10-8 indicates that it is desirable to

operate with large $\omega RC$. This is the condition leading to large peak tube currents and may be undesirable for good tube life. Ordinarily the simple capacitor filter is employed only for small load-current applications, where peak-current ratings of the tubes are not endangered, and where the ripple will also be desirably low. For heavy current outputs the ripple will usually be too large to be satisfactory to supply electronic amplifier circuits.

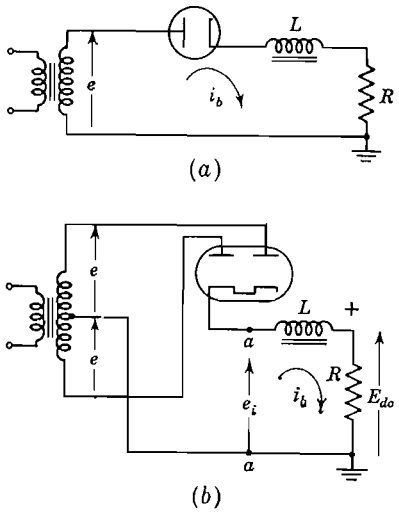



FIG. 10-11. (*a*) Half-wave rectifier with series-inductor filter. (*b*) Full-wave rectifier with series-inductor filter.

**10-5. The Series-inductor Filter.** An inductor in series with the load will serve also as a filter, representing a high series impedance to the alternating harmonic components in the rectifier output. The inductor may also be considered as storing magnetic energy when the current is above the average value and releasing that energy to the load when the current tends to fall below the average, thus reducing the ripple. Rectifiers with an inductor in series with the load are shown in Fig. 10-11.

For illustration and understanding of the individual diode current flow, consider first the action of the half-wave circuit. Writing the circuit differential equation for the conducting interval for $e = E_m \sin \omega t$,

$$L \frac{di}{dt} + Ri = E_m \sin \omega t \tag{10-33}$$

again neglecting the tube drop and transformer leakage reactance. Then

$$\frac{di}{dt} + \frac{Ri}{L} = \frac{E_m}{L} \sin \omega t \tag{10-34}$$

which has the transient solution

$$i_b = B\epsilon^{-Rt/L} \tag{10-35}$$

and a steady-state solution

$$i_b = \frac{E_m}{R\sqrt{(\omega L/R)^2 + 1}} \sin (\omega t - \phi) \tag{10-36}$$

where

$$\phi = \tan^{-1} \frac{\omega L}{R} \tag{10-37}$$

Therefore the actual tube current consists of repeated transients of the form

$$i_b = B\epsilon^{-Rt/L} + \frac{E_m}{R\sqrt{(\omega L/R)^2 + 1}} \sin(\omega t - \phi) \qquad 0 \leq \omega t \leq \theta_2$$

$$i_b = 0 \qquad \theta_2 \leq \omega t \leq 2\pi$$

Because of the negative portion of $E_m \sin \omega t$ the current must decrease to zero and tend to reverse before the beginning of the next cycle of the supply voltage. Since a reverse current is not possible, the current becomes zero at some cutoff angle $\theta_2$ and remains there until the supply

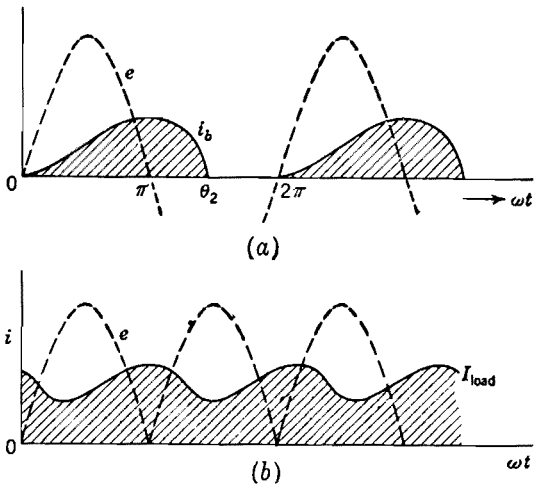



FIG. 10-12. (*a*) Load current: half-wave rectifier with inductor filter. (*b*) Load current: full-wave rectifier with inductor filter.

voltage again becomes positive. Therefore the current will always be zero at the instant the diode connects the source to the load at $\omega t = 0$, and the constant $B$ can be evaluated,

$$B = \frac{E_m}{R\sqrt{(\omega L/R)^2 + 1}} \sin \phi$$

so that the half-wave load-current pulses can finally be written as

$$i_b = \frac{E_m}{R\sqrt{(\omega L/R)^2 + 1}} \left[\sin(\omega t - \phi) + \epsilon^{-R\omega t/\omega L} \sin \phi\right] \qquad 0 \leq \omega t \leq \theta_2 \quad (10\text{-}38)$$

$$i_b = 0 \qquad \theta_2 \leq \omega t \leq 2\pi$$

As $\omega L/R$ is increased, the exponential term decays more slowly and the current flows for some time after the supply-voltage zero at $\omega t = \pi$, as shown in Fig. 10-12*a*. A positive emf is maintained on the diode

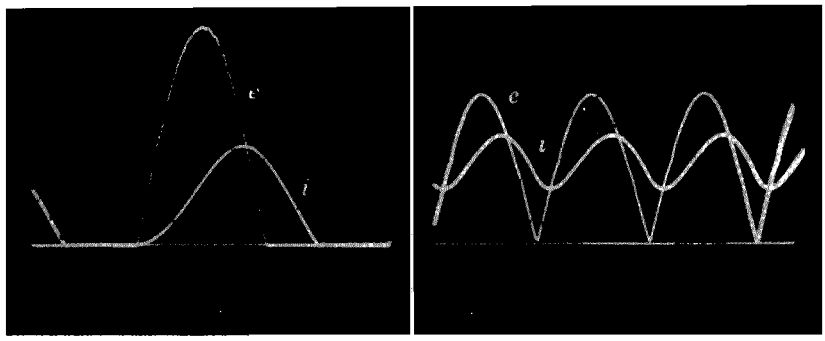
anode during the period $\pi$ to $\theta_2$ by reason of the $L\ di/dt$ voltage induced in $L$ by the decreasing current. For infinite $\omega L/R$ the cutoff angle would occur at $2\pi$, or 360° of the input cycle.

The cutoff angle $\theta_2$ may be found by considering that, when $\omega t = \theta_2$, $i_b = 0$; so

$$\sin(\theta_2 - \phi) + \epsilon^{-R\theta_2/\omega L} \sin\phi = 0$$

$$\frac{R}{\omega L} = \frac{\cos\theta_2 - \epsilon^{-R\theta_2/\omega L}}{\sin\theta_2} \tag{10-39}$$

and $\theta_2$ obtained by graphical means in terms of the parameter $R/\omega L$. The results for the half-wave rectifier are plotted in Fig. 10-14, and it may be seen that the conduction angle, from $\omega t = 0$ to $\omega t = \theta_2$, is




(*a*) (*b*)

FIG. 10-13. (*a*) Applied voltage and load current: half-wave rectifier with inductor filter. (*b*) Same for full-wave circuit. Note continuous current flow.

lengthened by increase of inductance or *decrease* of $R/\omega L$. Thus with increasing load current the conduction angle increases.

The d-c output current can be obtained by integration of Eq. (10-38) over the conduction interval 0 to $\theta_2$, and after some manipulation the result is

$$I_{dc} = \frac{E_m}{2\pi R}(1 - \cos\theta_2)$$

from which

$$E_{dc} = \frac{E_m}{2\pi}(1 - \cos\theta_2) \tag{10-40}$$

With large values of $L$ ($\theta_2 \cong 360°$) the d-c output voltage is small in terms of $E_m$. With small $L$ values ($\theta_2 \cong 180°$) the $E_{dc}$ value approaches the value of $E_m/\pi$, as for the half-wave rectifier without filter. The output voltage will thus be much lower than the capacitor filter, with the same supply voltage.

For the inductor filter with the full-wave circuit of Fig. 10-11*b* the current wave and conduction angle of one diode overlaps the current

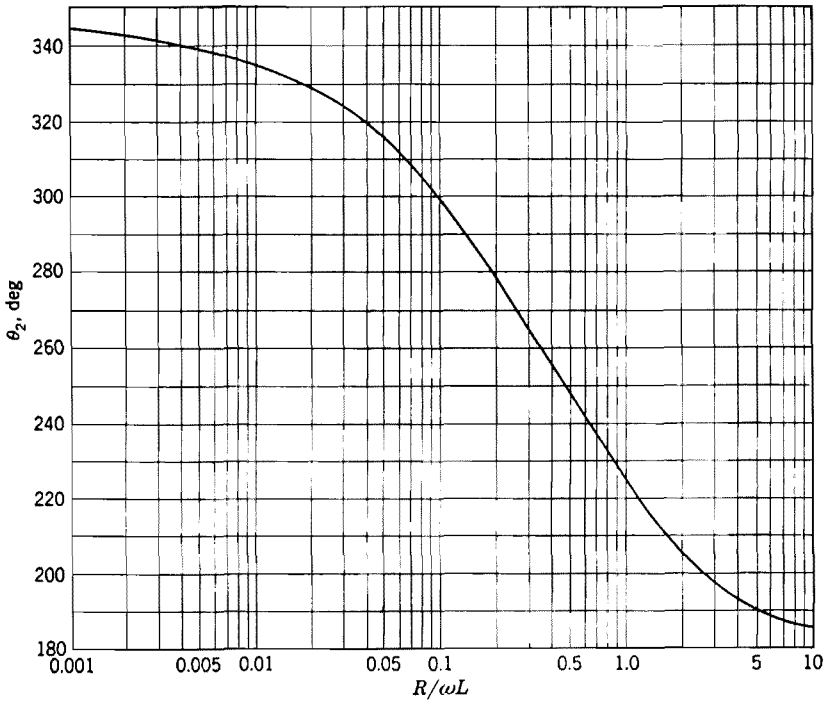



FIG. 10-14. Cutoff angle $\theta_2$ for the half-wave rectifier with inductor filter.

wave and conduction angle of the other diode, and with reasonable values of $L$ the conduction of current is continuous in the inductor, or *choke*, as indicated in Figs. 10-12*b* and 10-13*b*. The current commutates between diodes, the diode having a positive anode being the one conducting (neglecting transformer reactance). Thus a rectified sine voltage of the form indicated in Fig. 10-12*b* is applied to the filter input terminals at *a*, *a* of Fig. 10-11.

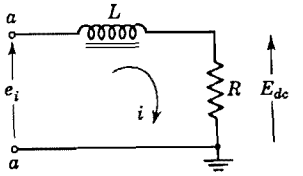



FIG. 10-15. Load circuit of inductor filter.

The filter input voltage $e_i$ is composed of the harmonic frequencies given in Eq. (10-15), a few of the low-frequency terms being

$$e_i = \frac{2E_m}{\pi} - \frac{4E_m}{3\pi}\cos 2\omega t - \frac{4E_m}{15\pi}\cos 4\omega t - \frac{4E_m}{35\pi}\cos 6\omega t - \cdots \quad (10\text{-}41)$$

the first term representing the d-c component of output.

If the above voltage is applied to the choke and load circuit of Fig. 10-15, with the choke having a resistance of $R_c$, then the direct load

current is that produced by the first emf term of Eq. (10-41),

$$I_{dc} = \frac{2E_m}{\pi} \frac{1}{R_c + R} \tag{10-42}$$

and the d-c voltage across the load is

$$E_{dc} = \frac{2E_m}{\pi} \frac{1}{1 + R_c/R} \tag{10-43}$$

The input impedance of the filter at $a$, $a$ for the lowest a-c frequency or the second harmonic is

$$\mathbf{Z}_2 = R_c + R + j2\omega L$$

The input second-harmonic current at $a$, $a$ is then

$$|I_2| = \frac{4E_m}{3\sqrt{2}\,\pi(R_c + R)\sqrt{[2\omega L/(R_c + R)]^2 + 1}} \tag{10-44}$$

and that of the fourth-harmonic component is

$$|I_4| = \frac{4E_m}{15\sqrt{2}\,\pi(R_c + R)\sqrt{[4\omega L/(R_c + R)]^2 + 1}}$$

The ratio $2\omega L/(R_c + R)$ should be large with respect to unity for good filtering, and under that assumption

$$\left|\frac{I_4}{I_2}\right| \cong \frac{1}{10}$$

and the sixth and higher orders are progressively smaller. Thus it is seen that the second-harmonic current is very largely responsible for the ripple in the output, and the ripple factor may be computed from Eqs. (10-42) and (10-44) as

$$\text{Ripple factor} \cong \frac{|I_2|}{I_{dc}} = \frac{\sqrt{2}}{3\sqrt{[2\omega L/(R_c + R)]^2 + 1}} \tag{10-45}$$

and if $2\omega L/(R_c + R) \gg 1$, then

$$\text{Ripple factor} \cong \frac{\sqrt{2}\,(R_c + R)}{6\omega L} = 0.236\,\frac{R_c + R}{\omega L} \tag{10-46}$$

Equation (10-46) shows that as $R$ decreases, or the load current increases, the ripple factor drops. This is not surprising since the energy stored in the inductor increases with average current. The effect is in contrast with the capacitor filter, in which the ripple becomes greater with increase of load current.

The inductor, or choke filter, produces no serious peak-current demands

on the rectifier tubes and has better voltage regulation than the capacitor-filter circuit but requires a high a-c input voltage from the transformer, possibly increasing the cost of that unit.

**10-6. The L-section, or Inductor-input, Filter.** Neither the shunt-capacitor nor the series-inductor forms of filters are economically capable of giving ripple factors sufficiently low for supplying direct current to much electronic equipment. A combination, as in the *L-section*, or *inductor-input*, filter of Fig. 10-16*a*, is frequently employed. The inductor serves to decrease the amplitude of the a-c components of current,

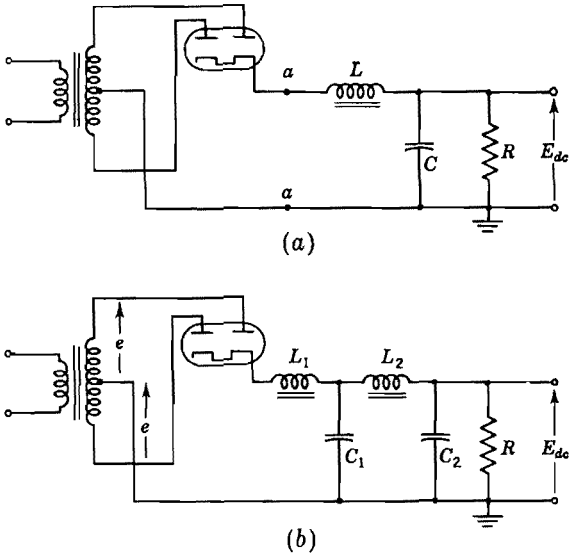



FIG. 10-16. Inductor input or L-section filter for full-wave rectifier.

and the capacitor bypasses them around the load, the result being a lower ripple than is possible with the same $L$ or $C$ alone.

Since the capacitor filter has low ripple on light loads and the inductor filter has a low ripple on heavy loads, it seems reasonable that the inductor-input filter might have suitable ripple characteristics at all loads. In addition, it is usually desirable that the impedance, as seen at the load looking back into the filter, be small for all signal frequencies in the amplifier supplied by the filter. This is desired to avoid feedback of signal components between amplifier stages. The shunt capacitor of the $L$ section may be chosen large enough to furnish this low filter output impedance. A large capacitor at this position also improves the dynamic output-voltage regulation, under conditions of highly variable d-c loads and surges.

For light loads, where $R \cong \infty$, the capacitor will charge to the peak

voltage, and the output will be $E_m$. As $R$ is decreased, the diodes switch the supply voltage onto the filter for an instant and $C$ charges to the peak, but the d-c voltage will be lowered owing to the average capacitor voltage being below the peak. The diode action is much as for the shunt-capacitor filter, since the current is so small that the energy storage and effect of the inductor are negligible. As the load current is further increased, the diode conduction angle lengthens owing to the more rapid

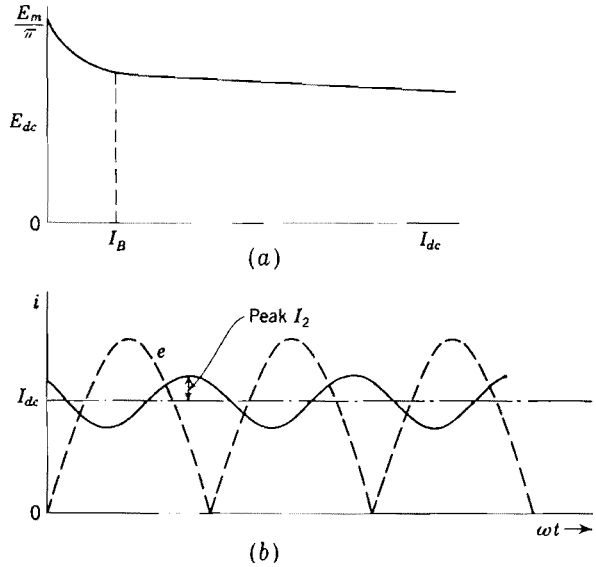



FIG. 10-17. (*a*) Regulation of L-section filter. (*b*) Output current for L-section filter.

discharge of the capacitor and to the effect of the inductor being appreciable with increased current. At some value of d-c current $I_B$ the conduction angle reaches 180° for each diode, the diodes commutate as for the inductor filter, and the load current is continuous.

In Fig. 10-17*a* current $I_B$ indicates the point on the regulation curve at which continuity of load current is obtained, load-voltage change beyond that current being small. The filter at that point begins to act more like an inductor filter and less like a shunt-capacitor filter. Because of the good regulation at current values above $I_B$ it is desirable that the rectifier be operated at the higher load currents, this being ensured by connection of a *bleeder resistor,* or minimum load, across the output. This resistor $R_B$ is designed to bleed, or draw, a current of about $I_B$ value, even with all external load removed. For purposes of analysis $R$ will designate the effective load formed by the actual load and $R_B$ in parallel.

Under conditions of load current continuity and 180° conduction per

diode the full-wave rectified voltage applied to the filter input $a$, $a$ of Fig. 10-16$a$ is again that of the Fourier series of Eq. (10-15) or (10-41). Values of $L$ and $C$ chosen for L-section filters may approximate $L = 20$ h, $C = 20$ μf. The second harmonic predominates, and for a 60-cycle source the second-harmonic reactances are $X_L = 2\omega L \cong 15{,}000$ ohms, $X_C = 1/2\omega C \cong 66$ ohms. It is reasonable to assume that $X_C$ will be negligibly small with respect to $R$, the load and bleeder in parallel, or that a-c components flow through $C$, the d-c components through $R$, as indicated in Fig. 10-18.

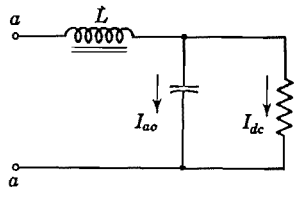


FIG. 10-18. Current paths assumed.

The second-harmonic rms current value is obtained by using the appropriate term from Eq. (10-41) as

$$I_2 = \frac{4E_m}{3\sqrt{2}\,\pi(2\omega L - 1/2\omega C)} = \frac{4E_m\omega C}{3\sqrt{2}\,\pi(4\omega^2 LC - 1)} \qquad (10\text{-}47)$$

This current is assumed to flow entirely through $C$ and in so doing produces a second-harmonic voltage drop across the load of

$$|E_2| = \frac{4E_m}{3\sqrt{2}\,\pi(4\omega^2 LC - 1)}$$

Once more assuming the ripple to be entirely due to the second harmonic, the ripple factor may be found as

$$\text{Ripple factor} = \frac{|E_2|}{E_{dc}} = \frac{|E_2|}{2E_m/\pi} = \frac{\sqrt{2}}{3(4\omega^2 LC - 1)} = \frac{0.47}{4\omega^2 LC - 1} \qquad (10\text{-}48)$$

Under the assumption of continuous current flow in the load the ripple factor is independent of the load current flowing. This is important, since it permits operation of the filter over all current values above $I_B$ without consideration of ripple. Ripple factor is plotted as a function of $\omega^2 LC$ in Fig. 10-19, and it can be seen that ripple with the L section is much smaller than that possible with $L$ or $C$ alone in the circuits previously described. The effect on ripple magnitude of changing the value of $C$ is indicated in Fig. 10-20.

The condition of $4\omega^2 LC = 1$ should be avoided in filter design, because of the large ripple which would be produced by this condition of series resonance at the second harmonic.

Additional sections of filter are frequently added, as in Fig. 10-16$b$. Following the methods used above, it is possible to determine the ripple for such cases as

$$\text{Ripple factor} = \frac{0.47}{(4\omega^2 L_1 C_1 - 1)(4\omega^2 L_2 C_2 - 1)} \cdots \qquad (10\text{-}49)$$

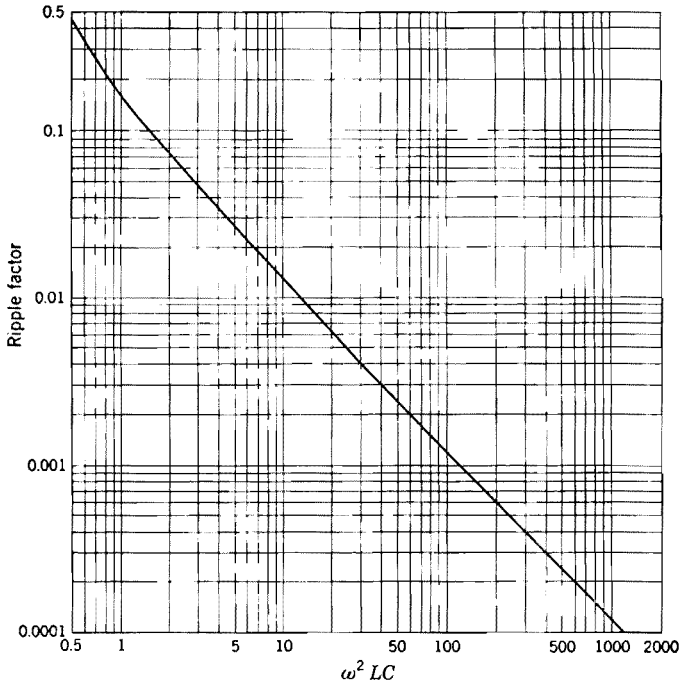



FIG. 10-19. Ripple factor vs. $\omega^2LC$ for choke-input single-section filter.

each section reducing the ripple by the factor $(4\omega^2LC - 1)$. If $n$ identical L sections are used and the permitted ripple $r$ is specified, then Eq. (10-49) may be inverted to give

$$LC = \frac{1}{4\omega^2}\left[1 + \left(\frac{0.47}{r}\right)^{1/n}\right] \quad (10\text{-}50)$$

for a full-wave rectifier, $L$ being in henrys, $C$ in farads.

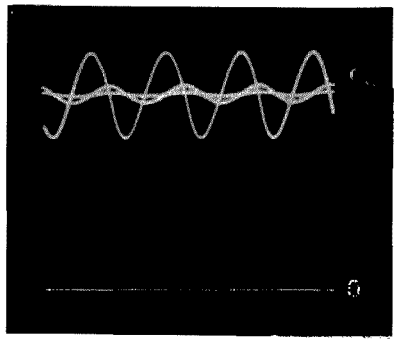


FIG. 10-20. Effect of increasing $C$ in the L-section filter. $C$ = 0, 2, 10 $\mu$f as ripple decreases; $L$ = 10 h.

Inductors in the range of 4 to 20 h are commonly used in these filters. Capacitors in the range of 4 to 80 $\mu$f are used; for voltages up to about 400 these are usually of the electrolytic type and are oil-impregnated paper dielectric for higher voltages.

**10-7. The Bleeder Resistor and Critical Input Inductance.** It is still necessary to determine the value of bleeder resistance to ensure operation beyond the steeply rising portion of the voltage-regulation curve in Fig.

10-17*a*. For the full-wave rectifier with inductor-input filter the load current should be just continuous at $I_{dc} = I_B$, the current taken by the bleeder. At this value the negative peak of the current wave is just tangent to the zero axis, and the wave appears as in Fig. 10-21. For this condition, with $R = R_B$ or no load other than the bleeder,

$$I_{dc} = I_{2m}$$

where $I_{2m}$ is the peak of the predominant second-harmonic current.

Using $I_{dc} = 2E_m/\pi R$, and the value of $I_{2m}$ from Eq. (10-47), the above equality gives

$$\frac{2E_m}{\pi R_B} = \frac{4E_m}{3\pi(2\omega L - 1/2\omega C)}$$

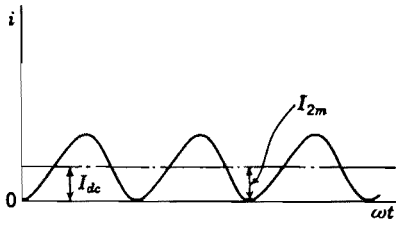



FIG. 10-21. Current wave through the load when $I_{dc} = I_B$.

It has previously been shown that $2\omega L \gg 1/2\omega C$ so that

$$R_B = 3\omega L \tag{10-51}$$

which, for 60-cycle supply, becomes

$$R_B = 1131L \tag{10-52}$$

and a value of $R_B = 1000L$ is frequently used, giving a small factor of assurance.

The value of the current taken by the bleeder is then

$$I_B = \frac{2E_m}{\pi R_B} = \frac{2E_m}{3\pi\omega L} = 0.000562\frac{E_m}{L} \quad \text{amp} \tag{10-53}$$

the latter expression applying for 60-cycle supply in the full-wave circuit.

It is desirable that $I_B$ be small, or $R_B$ large, to reduce the d-c power lost in the bleeder. The critical value for $I_B$ can be reduced by raising $R_B$ and $L$ according to Eq. (10-52), thus raising the second-harmonic input impedance of the filter, reducing $I_{2m}$. Increasing $L$ raises the cost of this item, and the change may dictate use of an inductor larger than needed to give the required filtering or ripple. The problem may be solved by using more turns and less iron, or reducing the air gap in the choke, allowing the iron to partially saturate at high values of $I_{dc}$, since at heavy loads the critical inductance given by Eq. (10-52) is small, and $L$ need be just large enough to limit the ripple.

As the $I_{dc}$ value is reduced, the inductance of the choke then rises or swings up to a value large enough at light loads to meet the desired criteria of Eq. (10-52), with a suitably large value of $R_B$. The inductor is then said to be a *swinging choke* and can be appreciably cheaper than

a choke with sufficient iron or air gap to avoid saturation at all load currents.

Such a choke can also improve the dynamic regulation of the filter on load-current swings. Since at heavy currents the inductor value is small, a sudden load-current increase can be immediately supplied from the rectifier tube and transformer, whereas if $L$ were constant at the high value dictated by Eq. (10-52), its time constant would prevent a rapid increase of current, and the load-current change could be only that immediately obtainable from the capacitor, and the capacitor voltage would fall. For this reason swinging chokes are almost universally found in filters supplying amplifiers having sudden load-current demands, as by Class B stages.

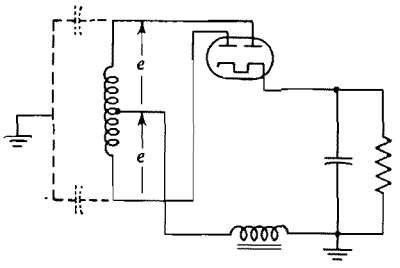



FIG. 10-22. Undesirable connection of filter choke in the negative lead.

Filters having more than one section employ the swinging choke on the input side of the filter; otherwise it could have no effect on the continuity of diode current.

The negative side of most electronic amplifier loads is grounded. This dictates that the filter choke not be placed in the negative lead of the rectifier. If so placed, a 60-cycle voltage equal to $e$ of the transformer is placed in series with the inductor and the secondary to primary capacitance of the transformer, as shown in Fig. 10-22. This causes a 60-cycle current to flow in the choke, which adds to the usual ripple.

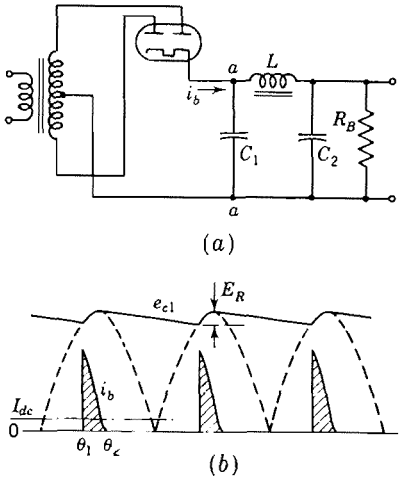



FIG. 10-23. (a) Capacitor-input filter. (b) Approximate tube-current waveform, and voltage on $C_1$.

**10-8. Capacitor-input Filters.** To provide a higher voltage, approximating $E_m$, from a given transformer, an input capacitor is sometimes added to the L-section filter, giving the circuit of Fig. 10-23*a*. Because of uncertainty as to tube cutin and cutout angles an exact analysis is difficult. However, if $C_1$ is of a size to aid in filtering, that is, 2 to 20 $\mu$f, then it charges to the peak voltage and the voltage $e_C$ behaves much as for the capacitor of the shunt-

capacitor filter. This is illustrated in Fig. 10-23*b*, which also shows the sharp and peaked current pulse. This peak current may be damaging to the tube cathode if heavy $I_{dc}$ loads are drawn.

Let it be assumed that the current out of the capacitor during the nonconducting interval of a diode is constant, as was done for the shunt-capacitor filter. This also means that the current in $L$ is constant, or that $L$ is large enough to keep the ripple component small. The average, or d-c, output voltage is then

$$E_{dc} = E_m - \frac{E_R}{2} - I_{dc}R_c$$

where $R_c$ is the choke resistance. The peak-to-peak magnitude $E_R$ may be found by assuming that $C_1$ is large and the conduction interval is near zero. The time of capacitor discharge may then be assumed equal essentially to one-half cycle of the second harmonic, or, if $T = 1/4f$,

$$E_R = \frac{q}{C_1} = \frac{I_{dc}T}{C_1} = \frac{\pi I_{dc}}{2\omega C_1}$$

The d-c output voltage then becomes

$$E_{dc} = E_m - I_{dc}\left(\frac{\pi}{2\omega C_1} + R_c\right) \tag{10-54}$$

showing that the regulation is a function of $C_1$, improving with large values. This is further illustrated in Fig. 10-24, which shows the effect of changing $C_1$ on both $E_{dc}$ and ripple. In fact, it is usually said that the major effect of $C_1$ is on regulation, that of $C_2$ on ripple.

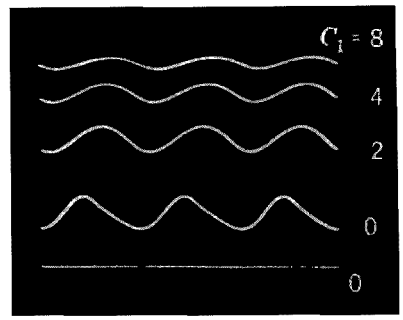



FIG. 10-24. Effect of increasing the input capacitor of the filter. All curves on same zero axis. $E_{dc}$ = 50, 120, 160, 190 volts, in ascending order.

If the ripple waveform of Fig. 10-23*b* is considered triangular, it has a Fourier series given by

$$e_R = \frac{E_R}{\pi}\sin 2\omega t - \frac{E_R}{2\pi}\sin 4\omega t + \frac{E_R}{3\pi}\sin 6\omega t - \cdots \tag{10-55}$$

This voltage is applied across terminals $a$, $a$ (Fig. 10-23*a*). Considering the second harmonic predominant in producing a current through the choke $L$, and if $1/2\omega C_2 \ll R$, where $R$ is the resistance of load and bleeder in parallel, then the portion of rms second-harmonic voltage appearing across $C_2$ or $R$ is

$$\mathbf{E}_2 = \frac{E_R}{\sqrt{2}\,\pi}\,\frac{-j/2\omega C_2}{j2\omega L - j/2\omega C_2} = \frac{E_R}{\sqrt{2}\,\pi}\,\frac{1}{4\omega^2 LC_2 - 1}$$

Use of the magnitude of $E_R$ determined for Eq. (10-54) gives

$$|E_2| = \frac{I_{dc}}{2\sqrt{2}\,\omega C_1} \frac{1}{4\omega^2 L C_2 - 1} \tag{10-56}$$

This is the ripple voltage across the load. The ripple factor follows directly as

$$\text{Ripple factor} = \frac{|E_2|}{E_{dc}} = \frac{1}{2\sqrt{2}\,\omega C_1 R(4\omega^2 L C_2 - 1)} \tag{10-57}$$

The ripple factor is a function of load on the rectifier for the capacitor-input filter, increasing with increasing current output. It may be remembered that for the L-section filter the ripple was independent of the load.

This circuit may burden the rectifier tubes with excessive current peaks and has poorer voltage regulation than does the inductor-input filter. It is used, however, to reduce cost of the transformer since a lower secondary voltage may be used.

**10-9. Voltage-multiplying Rectifiers.** For cost reasons it is sometimes desirable to eliminate the rectifier transformer; at the same time voltages higher than obtainable from the a-c line voltage may be needed. The circuit of Fig. 10-25 is a *voltage doubler* operating from the a-c line.

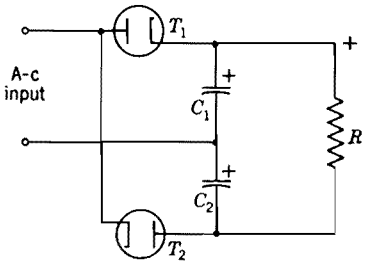


FIG. 10-25. Voltage doubler.

In this circuit with the upper a-c terminal positive diode $T_1$ conducts and charges capacitor $C_1$ to the peak of that half of the a-c wave. On the next half cycle $T_2$ conducts and charges $C_2$ to the peak voltage of the negative half cycle. Capacitors $C_1$ and $C_2$ supply the load current in series, giving a no-load output voltage roughly equal to the peak-to-peak value of the input wave.

Capacitors $C_1$ and $C_2$ must be large enough to store the energy required to supply the output circuit during the nonconducting intervals of the diodes. The voltage regulation of the output is directly a function of the amount of capacity used.

**10-10. Rectifier-circuit Design.** It should be noted that the relations in this chapter have been derived by neglecting the transformer leakage reactance, and the diode resistance in most cases. Thus transformer voltages should be increased over values obtained by analysis. Tube resistance may be estimated from the static characteristic curves.

Transformer secondaries for this service are usually rated in terms of the *a-c volts* each side of the center tap and of the *maximum d-c current* which can be drawn from the supply.

Cathode or filament heater transformers should be insulated to withstand the d-c voltage to ground, since most power supplies operate with the rectifier cathode above ground by the amount of the d-c voltage.

Tubes should be selected by consideration of the following factors:

*Peak Inverse Voltage.* A tube is insulated to withstand a limited voltage without breakdown. The PIV of the circuit is the highest voltage across the tube during a cycle, and this must be less than the rated maximum inverse voltage of the tube. For the half-wave circuit with capacitor filter the PIV is equal to $E_m + E_{dc}$, since the d-c output is so polarized as to add to the a-c input negative peak. For the full-wave circuit the PIV value is $2E_m$ in all cases.

*Peak Tube Current.* The loss of emitting material from the cathode usually ends the life of a tube. The rate of loss at current peaks is many times that at the average current value. Circuit designs should be checked for peak current values, to ensure the current peak being within the maximum current rating.

*Average Tube Current.* For inductor-input filters, where the tube current is fairly constant at the average value, the heat generated or the anode loss is a function of the average tube current. This heat loss must be limited to a safe value, so that the anode not be overheated and liberate gas, which might cause permanent tube damage.

For capacitor-input circuits the peak-current values will usually be exceeded before the anode-dissipation rating is reached.

**Example.** A full-wave rectifier, with choke-input filter, is to supply 250 volts at 100 ma d-c. Determine the needed ratings of tubes and transformer, the value of bleeder and ripple, if the tube resistance be 250 ohms per diode, $R_c$ of choke = 400 ohms, $L$ = 10 h, with $C$ = 20 $\mu$f.

The load (and bleeder) resistance will be given by

$$R = \frac{250}{0.100} = 2500 \text{ ohms}$$

Then, by use of Eq. (10-43) and the fact that $I_{rms} \cong I_{dc}$ in a choke-input filter, the supply voltage is obtainable as

$$E_m = \frac{\pi E_{dc}}{2}\left(1 + \frac{R_c}{R}\right) + I_{dc}r_b$$
$$= \frac{\pi \times 250}{2}\left(1 + \frac{400}{2500}\right) + 0.100 \times 250 = 480 \text{ volts}$$

The transformer should then supply 340 volts rms each side of the center tap. This still includes no allowance for transformer impedance, so that the transformer should be rated at about 360-0-360 volts, 100 ma d-c.

The tube rating should be 50 ma average per diode. In an inductor-input filter the peak rating need not be checked, since it will not greatly exceed the average value. For capacitor input the peak current should be computed since it may be the limiting factor.

The bleeder resistance can be found as

$$R_B = 1000L = 10{,}000 \text{ ohms}$$

and

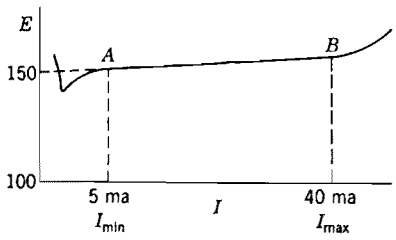
$$I_B = \frac{2E_m}{3\pi\omega L} = \frac{2 \times 480}{3\pi \times 377 \times 10} = 0.027 \text{ amp}$$

The ripple can be computed from Eq. (10-48) as

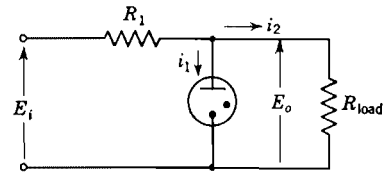
$$\text{Ripple factor} = \frac{0.47}{4\omega^2 LC - 1} = \frac{0.47}{4 \times 377^2 \times 10 \times 20 \times 10^{-6}} = 0.00413$$

which checks the value for $\omega^2 LC = 28.4$ in Fig. 10-19. This value of ripple means 4 mv of ripple per d-c volt output, or 0.1 volt rms ripple total. This is not sufficiently low for many electronic applications, and a second section of filter should be added.

**10-11. Voltage-regulator (VR) Tubes.** The internal impedance of the usual rectifier and filter is quite high, and the output-voltage regulation will be correspondingly poor for varying d-c load currents. The output voltage is also a function of the applied a-c line voltage. For proper operation of many electronic systems the supply voltages must be constant, and electronic devices and circuits have been developed for regulation.




(a)




(b)

FIG. 10-26. (a) Volt-ampere curve for a VR tube. (b) Glow-tube regulator.

A simple form of voltage regulator for loads up to 40 ma is a cold-cathode gaseous-glow tube. At ionization or breakdown potential the drop across the tube falls to a slightly lower level, and further current increases occur without change in tube voltage drop. The cathode is partially covered with a glow discharge at constant-current density, and as the current increases, the area covered increases at constant voltage. After the full cathode is covered with glow, additional current again causes a rise in voltage, as described in Sec. 2-13. Between the limiting currents of about 5 and 40 ma such tubes may have a variation in voltage drop of only 1 to 4 volts, as indicated in Fig. 10-26.

The operating voltage drop at which this constant-voltage range occurs is fixed by tube design and composition of the gas, and a variety of tubes are available to supply voltages such as 75, 90, 105, or 150 volts. Several tubes may be connected in series to give operating voltages equal to the

sum of individual tube ratings. The supply voltage must always be greater than the breakdown voltage to make ionization of the gas possible. Breakdown may run 20 to 50 volts above the rated operating drop.

At normal-load current the regulator-tube current $i_1$ should be about in the middle of the operating range. The sum $i_1 + i_2$ will cause a voltage drop in $R_1$ equal to the difference between the unregulated supply $E_i$ and the VR-tube voltage $E_o$. As the load current is increased, less current is taken by the tube; hence the total current and the $(i_1 + i_2)R_1$ drop remain constant. A change in $E_i$ will require a change in $i_1$ so as to adjust the drop in $R_1$ to equal the new difference between $E_i$ and the nearly constant $E_o$. The resistor $R_1$ may be chosen as

$$R_1 = \frac{E_i - E_o}{I_{max}} \tag{10-58}$$

The load may then draw any current from zero to $I_{max} - I_{min}$ without passing outside the operating range $A$, $B$ of the regulator tube, provided that $E_i$ is given its maximum expected value.

VR-tube regulators are normally employed to operate on amplifier plate-voltage supplies. With heater-type tubes a period will always exist during warm-up when the amplifier draws no plate current or $R_{load}$ is very large. To ensure breakdown of the VR tube it is then sufficient that the lowest expected value of $E_i$ should exceed $E_B$, the rated VR-tube breakdown voltage.

Such regulators are satisfactory where the limited current capacity and their fixed voltage ratings may be integrated into a design.

**Example.** A constant supply of 150 volts with load current of 30 ma is to be obtained from a source of 250 ± 30 volts. The 150-volt specification indicates choice of an OD3/VR 150 regulator rated at $E_o = 150$ volts, $E_B = 160$ volts, current range 5 to 40 ma. Regulation in this range will be within 4 volts.

Equation (10-58) gives

$$R_1 = \frac{280 - 150}{0.040} = 3250 \text{ ohms}$$

The load resistance is 150/0.030 = 5000 ohms.

**10-12. Vacuum-tube Regulators.** For greater flexibility in output voltage various electronic circuits are employed in preference to VR tubes.

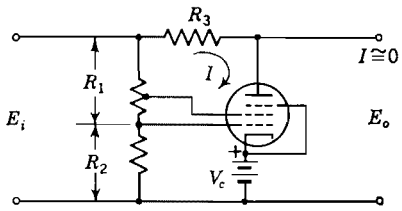



FIG. 10-27. Shunt-type regulator for small currents.

In regulation of anode voltages for cathode-ray tubes and similar loads where no appreciable current output is required, the circuit in Fig. 10-27 may be used. For a change $\Delta E_i$ in the input voltage the change in grid

voltage of the pentode will be

$$\Delta E_g = \frac{R_1}{R_1 + R_2} \Delta E_i \tag{10-59}$$

The resultant change in output is

$$\Delta E_o = \Delta E_i - \Delta I R_3 \tag{10-60}$$

and the plate current of the tube is given by

$$\Delta I = g_m \Delta E_g$$
$$= \frac{g_m R_2}{R_1 + R_2} \Delta E_i$$

so that

$$\Delta E_o = \Delta E_i - \frac{g_m R_2 R_3}{R_1 + R_2} \Delta E_i \tag{10-61}$$

The regulation sensitivity, or the change in output voltage per volt input change, is then

$$\frac{\Delta E_o}{\Delta E_i} = 1 - \frac{g_m R_2 R_3}{R_1 + R_2} \tag{10-62}$$

The change can be zero, or the regulation perfect, if

$$g_m = \frac{R_1 + R_2}{R_2 R_3} \tag{10-63}$$

If $R_1 = 20 \times 10^6$, $R_2 = 2 \times 10^6$, $R_3 = 2 \times 10^4$ ohms, a value of $g_m = 550$ $\mu$mhos would be required.

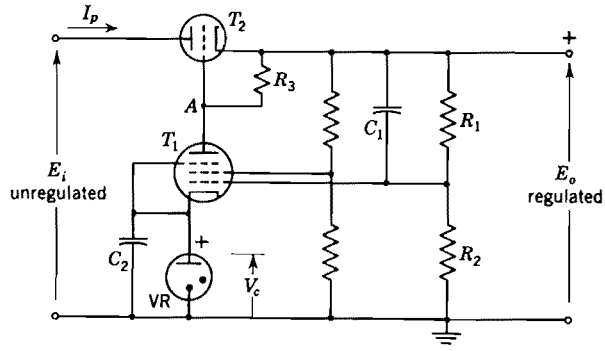



FIG. 10-28. Basic electronic regulator circuit.

For higher currents and improved regulation to variations in output load the basic circuit of Fig. 10-28 is used. Tube $T_2$, usually a power triode such as the 2A3 or 6AS7G, acts as a series variable resistor carrying the full-load current. A number of tubes may be used in parallel when current requirements dictate. Maximum current is fixed by the $T_2$ current rating and its plate dissipation.

The effective resistance of $T_2$ is controlled by the output of a d-c amplifier, here shown as pentode $T_1$. Tube $T_1$ compares a portion of the output voltage $E_o$ against a standard reference voltage such as that across the VR tube. Any difference between the output fraction and $V_c$ is amplified and applied to the control grid of $T_2$, altering the resistance of $T_2$. An increase of $E_o$ makes the grid of $T_1$ more positive; the increase in plate current of $T_1$ makes the grid of $T_2$ more negative, increasing the series resistance of $T_2$ and reducing the output voltage toward its original level.

The output voltage from such a regulator is not absolutely constant, since for an increase in input $\Delta E_i$ there must be a slight rise $\Delta E_g$ in the grid voltage of $T_1$ if its output across $R_3$ is to increase and change the resistance of $T_2$. However, if the gain of $T_1$ is large, then $\Delta E_g$ and the permitted change in $E_o$ can be very small.

If the regulator is to operate properly, all changes in voltage and current will be small and the equivalent circuit for $T_1$ used in analysis of performance. The input to the d-c amplifier $T_1$ is

$$E_g = \frac{R_2}{R_1 + R_2} E_o = KE_o \tag{10-64}$$

This voltage is so chosen with respect to $V_c$ that the $T_1$ grid is at $-3$ to $-5$ volts to cathode for the usual sharp-cutoff pentode. The gain of pentode $T_1$ is

$$A_1 = -g_{m1}R_3$$

Suppose that owing to a small change $\Delta E_i$ in the input a small change $\Delta E_o$ occurs in the output. The change in the voltage appearing at the grid of the d-c amplifier is

$$\Delta E_{g1} = K\,\Delta E_o$$

and at its output across $R_3$ appears

$$\Delta E_{g2} = -A_1K\,\Delta E_o \tag{10-65}$$

The amplification factor of $T_2$, for small changes, is

$$\mu_2 = -\frac{\Delta E_b}{\Delta E_{g2}} \qquad \text{or} \qquad \Delta E_b = -\mu_2\,\Delta E_{g2}$$

so that the anode-cathode voltage of $T_2$ is

$$\Delta E_b = \mu_2 A_1 K\,\Delta E_o = \Delta E_i - \Delta E_o$$

the tube $T_2$ having the difference between input and output as its anode voltage. Then

$$\Delta E_o = \frac{\Delta E_i}{1 + \mu_2 A_1 K} \tag{10-66}$$

The regulation sensitivity may again be defined as the change of output voltage per unit input change, so that

$$\frac{\Delta E_o}{\Delta E_i} = \frac{1}{1 + \mu_2 A_1 K} \cong \frac{1}{\mu_2 A_1 K} \tag{10-67}$$

This equation indicates the effectiveness with which a regulator compensates for input changes, the effectiveness being improved by increase of gain $A_1$ and $K$.

With a single-stage amplifier the sensitivity may be of the order of 0.005, or 5 mv change in output per 1 volt change in input. Additional d-c amplification, by the methods of Chap. 8, may reduce the sensitivity to figures of the order of 0.0001 to 0.00005. A block diagram of one such circuit, using a chopper d-c amplifier, appears in Fig. 10-29.

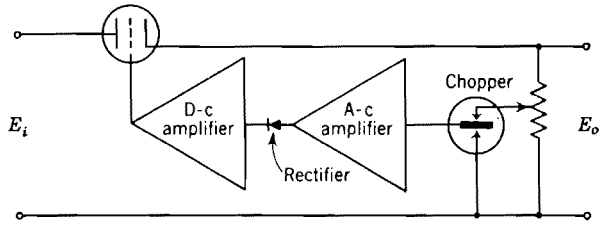



FIG. 10-29. High-precision voltage regulator.

Capacitor $C_1$ (Fig. 10-28) aids in reduction of ripple, since it makes $K = 1$ for a-c ripple voltages and sudden surges. With this capacitor a regulator becomes an important aid in reducing filter output ripple. Capacitor $C_2$ tends to prevent high-frequency gas oscillations that might occur in the VR tube, further stabilizing the supply.

The above analysis was made for varying input voltage. A change in load current may also cause a change in $E_o$, due to the internal resistance of the supply. By Eq. (10-65) the voltage at the grid of $T_2$ is

$$\Delta E_{g2} = A_1 K \, \Delta E_o$$

neglecting the minus sign. The transconductance of $T_2$ is

$$g_{m2} = \frac{\Delta I_p}{\Delta E_{g2}} \qquad \text{or} \qquad \Delta I_p = g_{m2} \, \Delta E_{g2}$$

so that

$$I_p = g_{m2} A_1 K \, \Delta E_o \tag{10-68}$$

The ratio of the change of output voltage to the change in output current is $R_s$, the effective internal resistance of the power supply. That is,

$$R_s = \frac{\Delta E_o}{\Delta I_p} = \frac{1}{g_{m2} A_1 K} \tag{10-69}$$

Equation (10-69) indicates that $T_2$ should have a high value of $g_m$. With high gain for $A_1$ the value of $R_s$ can be less than 0.2 ohm.

If the change in output load current also causes a change in the filter output value of $E_i$, then the internal impedance of transformer, rectifier tubes, and filter must be considered. If the internal resistance of these elements measured at the filter output is $R_f$, then by Eqs. (10-67) and (10-69)

$$R_s = \frac{1}{g_{m2}A_1K} + \frac{R_f}{\mu_2 A_1 K} \tag{10-70}$$

For a typical case it might be assumed that $R_f = 2000$ ohms, $\mu_2 A_1 K = 5000$, so that the added term represents 0.4 ohm due to the source impedance.

Regulation may also be affected by variation in the a-c supply voltage to the tube heaters, $T_1$ being especially sensitive to this. If long-time stability is important, resistors $R_1$ and $R_2$ should be wire-wound and mounted so as to be exposed to the same temperature. The VR tube should also be of a type especially manufactured for regulator service to reduce shifts in voltage level due to internal changes in gas content. With high-gain amplifiers stability better than 1 part in 10,000 can be held over long periods.

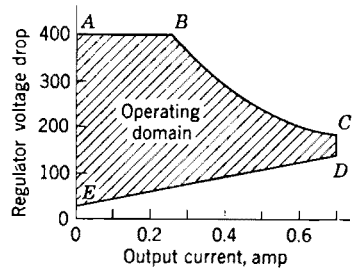



FIG. 10-30. Operating characteristics for electronic regulator.

When employed for manual variation of the output-voltage settings, restrictions on output appear as in Fig. 10-30. The *operating domain* is bounded by the line $A$-$B$, which indicates the greatest possible internal

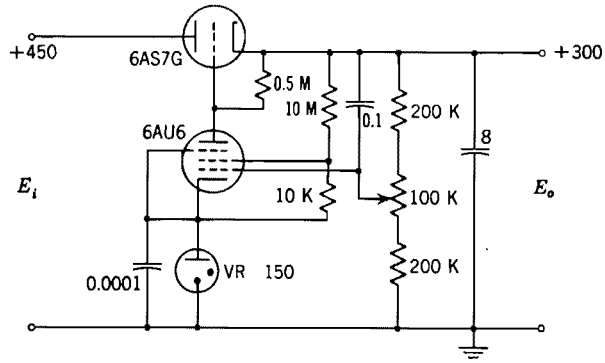



FIG. 10-31. A typical electronic regulator.

drop $E_i - E_o$, as fixed by the voltage required for the VR tube and other tube voltages. The segment $B$-$C$ is a portion of a hyperbola, determined by the plate-dissipation ratings of the series tubes. Line $C$-$D$ indicates

the maximum current value based on transformer or choke wire sizes and rectifier current ratings, and line $D$-$E$ is fixed by the minimum value of $E_b = E_i - E_o$ at which series tubes will pass the indicated current.

A complete regulator is diagramed in Fig. 10-31.

**10-13. A-C Regulators.** Regulation of tube-heater current may often be necessary in precision amplifiers and controls, particularly in d-c amplifiers. This may be done on the line side of the filament transformer by methods to be discussed, or it may be performed by current-regulator tubes. These are made with metallic filaments, usually iron, in a gas-filled bulb and change resistance rapidly as a function of filament temperature and current. A typical characteristic is shown in Fig. 10-32. These regulators are of particular value for tube heaters operated from storage batteries in portable equipment, where the tubes are subject to considerable variation in voltage during the charge-discharge cycle of the batteries. Current regulators are available for a variety of current and voltage ratings.

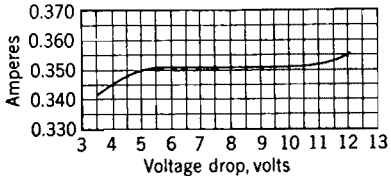



FIG. 10-32. Constant-current regulator characteristic.

Alternating-current line voltage regulators are of three classes, illustrated in Fig. 10-33. These differ in speed of response, waveform distortion, and the effect of frequency changes. They will usually hold voltage to $\pm 1$ per cent or better when operated with an input between 95 and 130 volts.

In Fig. 10-33*a* is shown a nonelectronic resonant stabilizer. Transformer $T_1$ is operated near saturation and partially resonated by $C$ at the supply frequency. As the applied voltage changes, the degree of saturation is altered, the inductance changes, and the resonant voltage across $T_1$ changes rapidly in phase. This causes a shift in magnitude and phase of the voltage on $P$ of $T_2$, such that the phasor sum of the voltage applied to $T_1$ and that on $S$ of $T_2$ tends to remain constant. Because of the saturation of $T_1$ considerable waveform distortion is present in the output. This type is sensitive to load-power-factor changes and to frequency shifts, the output-voltage level changing with either variable. Speed of response to transient changes in voltage or load is about 2 cycles. A major advantage is the lack of moving or electronic parts, the unit being entirely passive.

The second form, in Fig. 10-33*b*, is electromechanical. The motor $M$ drives the tap on transformer $T_1$ to adjust the secondary voltage on $S$, adding or subtracting the proper voltage to $E_i$ to maintain constant $E_o$. The small motor is controlled by electronic circuits to be described in Chap. 17. The exciting signal is obtained from the voltage-sensitive

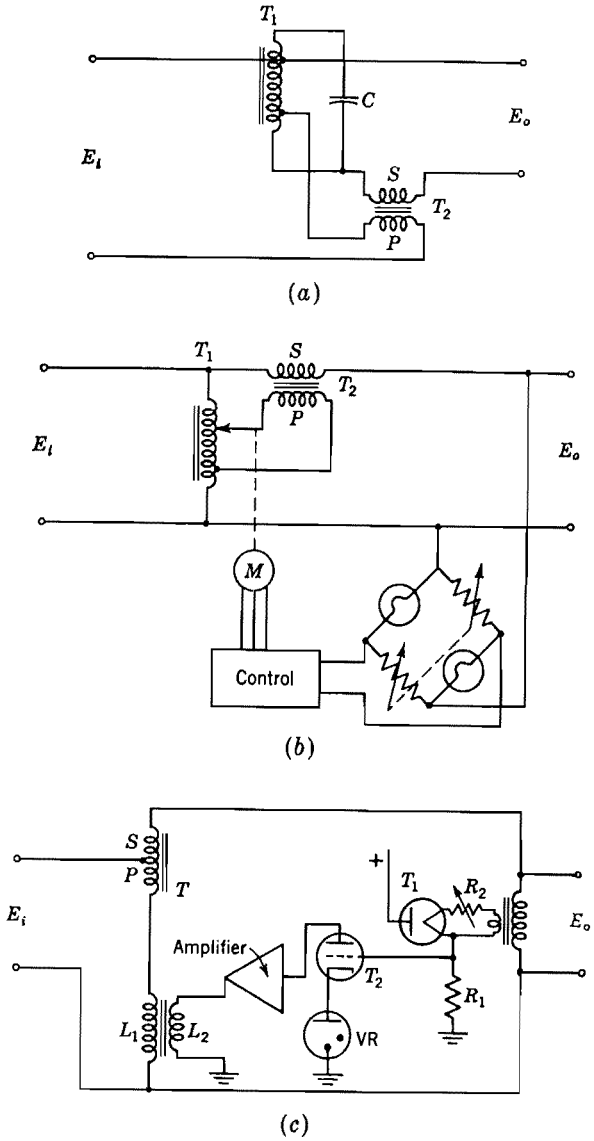



FIG. 10-33. Forms of a-c voltage regulators.

bridge circuit, having two opposite nonlinear resistance arms. These arms are frequently carbon- or tungsten-filament lamps, operated at a mid-point in their voltage range. The adjustable pair of bridge arms brings the bridge to balance for only one bridge-input voltage. A lower applied voltage reduces the lamp resistance (for tungsten), unbalancing

the bridge and giving an output of one a-c polarity. Increase of bridge voltage heats the lamp filaments, unbalancing the bridge in the opposite direction and giving an output of opposite a-c polarity.

The motor-control circuit recognizes these changes in polarity with some sort of phase detector and drives motor $M$ and the tap on $T_1$ in opposite directions as the bridge polarity reverses, but always back toward the desired standard voltage level, to restore the lamp bridge to balance. Speed of motor response will be determined by the amount of voltage unbalance.

Output waveform is not changed, nor is this type frequency-sensitive. The speed of response is dependent on motor speed, with full range of control being covered in a few seconds.

The third type, in Fig. 10-33*c*, is all-electronic in operation. The primary of $T$ is in series with a saturable reactor $L_1$, whose inductance can be varied by a small direct current in winding $L_2$. As the inductance of $L_1$ changes, the voltage applied to $P$ shifts and changes the voltage induced in the secondary series winding $S$, to maintain $E_o$ at the desired constant value.

The voltage-sensitive element is a temperature-saturated diode $T_1$, whose filament temperature is varied by $E_o$ and operated at saturation with all electrons emitted constituting the plate current. Changes in $E_o$ cause plate-current changes in $T_1$, with very considerable sensitivity because of the rapidly varying exponential in the Dushman law relating emitted current and temperature. Plate-current flow through $R_1$ produces a d-c voltage which is compared with the reference voltage across the VR tube. Any differential is amplified and applied to the saturable reactor winding $L_2$ so as to shift the voltage induced in $S$ to restore $E_o$.

As for the electromechanical form, no appreciable waveform distortion is introduced, nor is the circuit sensitive to frequency changes. Response to transients is completed within a few cycles. Variable resistor $R_2$ allows adjustment of diode-filament temperature to give any desired output $E_o$.

Factors to be considered in selection of a-c line-voltage regulators are the effect which a distorted waveform may have on connected apparatus, the stability of the frequency of the supply line, and the speed of response which will be needed.

**10-14. Voltage References.** In the described forms of regulators there have been sources of reference potential, or means of developing potentials representative of the deviation of the regulated voltage from a standard. A number of devices are capable of this action.

A *battery* is one of the simplest forms of reference potential and is satisfactory for comparisons with a d-c voltage. An a-c voltage can be rectified and compared with the battery potential. As the battery is usually employed, no current is drawn and the life is long. Some deviation

from a fixed voltage value may be expected as a function of temperature of the battery. For extreme precision of reference emf a standard cell may be used. It has a low potential, 1.01183 volts, and is usually employed as a comparison standard in connection with the chopper form of amplifier.

The *VR tube*, or gas-diode regulator, is most commonly used. Special types are available for precision applications, with high orders of stability with time and freedom from erratic drifts in potential.

The *temperature-saturated diode* uses a tungsten filament in a vacuum diode, the filament heated from the controlled source. Changes in source voltage cause changes in temperature and therefore in emitted current according to Dushman's law

$$J = A_o T^2 \epsilon^{-b_0/t} \qquad \text{amp/sq m}$$

as in Chap. 2. When combined with the Stefan-Boltzmann law determining temperature of a radiating source,

$$W = 5.77 \times 10^{-12} \sigma T^4 \qquad \text{watts/sq m}$$

it can be seen that $J$ will have a high rate of change with $T$ and input power $W$, and the emitted current is a very sensitive function of power input. Operation of the diode will be with high plate voltage so that saturation current is drawn to the anode, this current being independent of anode-voltage variations.

The *voltage-sensitive lamp and thyrite bridges* of Fig. 10-34 are also useful, particularly in control circuits. With tungsten lamps having a resistance variation as a high power of the applied voltage, there can be only one voltage at which the bridge will balance. The a-c output potential increases on either side of the balance voltage and reverses in phase as well. Thus information as to magnitude and direction of departure of the controlled voltage is available.

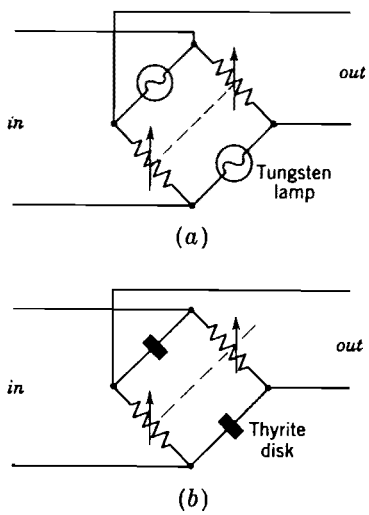



FIG. 10-34. Voltage-sensitive bridges.

The thyrite bridge is similar in operation, thyrite having a resistance given by $R = C/I^a$ with $a = 0.72$. There will again be only one value of voltage at which the bridge will balance.

The temperature-saturated diode and the voltage-sensitive lamp bridge depend on heat as their fundamental variable. They will give an output proportional to the rms value of the controlled voltage. Since the resist-

ance change of thyrite is instantaneous with voltage, it will distort the bridge-output waveform and this effect should be considered in circuit design. The bridge output will not be proportional to the effective value of applied voltage, and in fact will be somewhat dependent on bridge-input waveform.

## PROBLEMS

**10-1.** A half-wave diode, having a volt-ampere characteristic given by

$$i = 0.670V^{1.57} \quad \text{ma}$$

is in series with a 2000-ohm load.

(*a*) Plot the dynamic characteristic for the combination.

(*b*) Using the curve of (*a*), plot the waveform of voltage across the load, if a voltage $e = 50 \sin \omega t$ is applied to the combination.

**10-2.** A half-wave rectifier uses a diode with $r_b = 300$ ohms. If the input a-c voltage is 220 rms and the load is 1200 ohms, find $I_{dc}$, $I_{rms}$ in load, tube power loss, total a-c power input, and rectification efficiency.

**10-3.** A full-wave rectifier uses a 5Z3 rectifier with identical diodes of $r_b = 200$ ohms. A load of 2500 ohms is to be supplied at 300 volts d-c. Assume negligible transformer impedance, and find (*a*) $E_{rms}$ of transformer required, (*b*) peak diode current, (*c*) tube power loss, (*d*) a-c power rating of transformer, (*e*) rectification efficiency.

**10-4.** Find the minimum value of load $R$ permitted for maximum d-c power output from a full-wave rectifier using a 5Z3 tube rated PIV = 1550 volts, $I_m$ per diode = 675 ma, $I_{dc}$ per diode = 112 ma. Use $r_b = 150$ ohms. Also find the full-load to no-load change in voltage.

**10-5.** Derive Eq. (10-29).

**10-6.** Starting with Eq. (10-38), derive Eq. (10-40).

**10-7.** A full-wave rectifier, with 250 volts rms at 60 cycles on each side of the transformer center tap, supplies a shunt-capacitor filter having $C = 4$ $\mu$f, $R = 1000$ ohms. Find the cutin and cutout angles, and plot waveforms of $i_C$, $i_R$, $i_b$ over an input cycle. Neglect tube drop.

**10-8.** A full-wave rectifier operates into a shunt-capacitance filter having $C = 12$ $\mu$f, $R = 6000$ ohms, with supply of 300 rms volts on each half of the transformer at 60 cycles. Find the ripple factor, angles of cutin and cutout, and ratio of diode peak to average current.

**10-9.** Derive Eq. (10-49).

**10-10.** A full-wave rectifier has a series-inductor filter with $L = 10$ h. If the load is supplied with 300 volts, and 0.150 amp d-c, compute the ripple factor with primary supply at 60 cycles.

**10-11.** A filter of two cascaded L sections is to be used on the rectifier of Prob. 10-3, with 60-cycle supply. Capacitors $C_1$ and $C_2$ are to be equal. Design the filter, using the critical inductance value, and obtain a ripple factor of 0.001.

**10-12.** A full-wave rectifier to supply 250 ma, at 500 volts d-c, must have a ripple magnitude $E_R$ less than 10 volts. Specify the lowest cost elements for an L-section filter to provide this, if the supply frequency is 60 cycles, and if inductors of sufficient current-carrying capacity cost \$1.50 per henry, capacitors cost \$1 per microfarad.

**10-13.** A full-wave circuit, with tube resistance neglected, $I_{dc}$ per diode = 225 ma, supplies a load through an L filter with $L = 10$ h, $C = 8$ $\mu$f, from a 60-cycle line. The a-c voltage each side of the transformer center tap is 360 rms. If the choke has a

resistance of 200 ohms, find (*a*) full-load $E_{dc}$, (*b*) value of $R$ for a bleeder to prevent $E_{dc}$ from climbing at light load, (*c*) ripple voltage at full load.

**10-14.** A full-wave rectifier supplies a capacitor-input filter, consisting of two 8-$\mu$f capacitors and a 12-h choke. The output of the filter is 460 volts, and the output resistance of the filter is 600 ohms. The regulator of Fig. 10-31 is used on the filter output. If the load changes from 75 ma, at 300 volts, to 225 ma, find the change in output voltage at the regulator output.

**10-15.** For Prob. 10-14 compute the output ripple voltage at $E_{dc} = 300$ volts.

**10-16.** With $E_o = 300$ volts, $E_i$ changes from 425 to 450 volts for the regulator of Fig. 10-35. Compute the corresponding change in $E_o$.

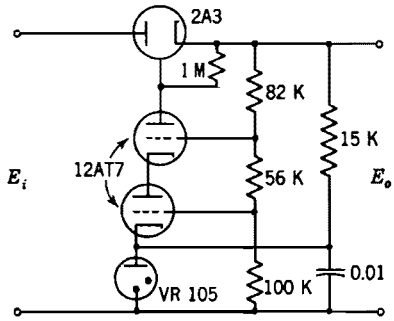



FIG. 10-35

**10-17.** If a thyrite disk has a characteristic $R = 580/I^{0.72}$, find the value of fixed resistors to use with two lamps in a bridge to balance with 22 volts applied.

## REFERENCES

1. Waidelich, D. L.: Diode Rectifying Circuits with Capacitance Filters, *Trans. AIEE*, **60**: 1161 (1941).
2. Dellenbaugh, F. S., Jr., and R. S. Quimby: The Important First Choke in High-voltage Rectifier Circuits, *QST*, **16**: 14, 27, 33 (February, March, April, 1932).
3. Waidelich, D. L.: Analysis of Full-wave Rectifier and Capacitive-input Filter, *Electronics*, **20**: 120 (September, 1947).
4. Terman, F. E., and S. B. Pickles: Note on a Cause of Residual Hum in Rectifier-filter Circuits, *Proc. IRE*, **22**: 1040 (1934).
5. Schade, O. H.: Analysis of Rectifier Operation, *Proc. IRE*, **31**: 341 (1943).
6. Hunt, V. F., and R. W. Hickman: On Electronic Voltage Stabilizers, *Rev. Sci. Instr.*, **10**: 6 (1939).
7. Langford-Smith, F.: "The Radiotron Designer's Handbook," 4th ed., The Wireless Press, Sydney, Australia, 1952.
8. Neher, H. V., and W. H. Pickering: Two Voltage Regulators, *Rev. Sci. Instr.*, **10**: 53 (1939).

CHAPTER 11

# OSCILLATORS AND CLASS C AMPLIFIERS; HIGH-FREQUENCY HEATING

The conversion of d-c power to high-frequency power, either in oscillators or in amplifiers, is an important duty of the vacuum tube. Since very large amounts of power are handled in these circuits when applied

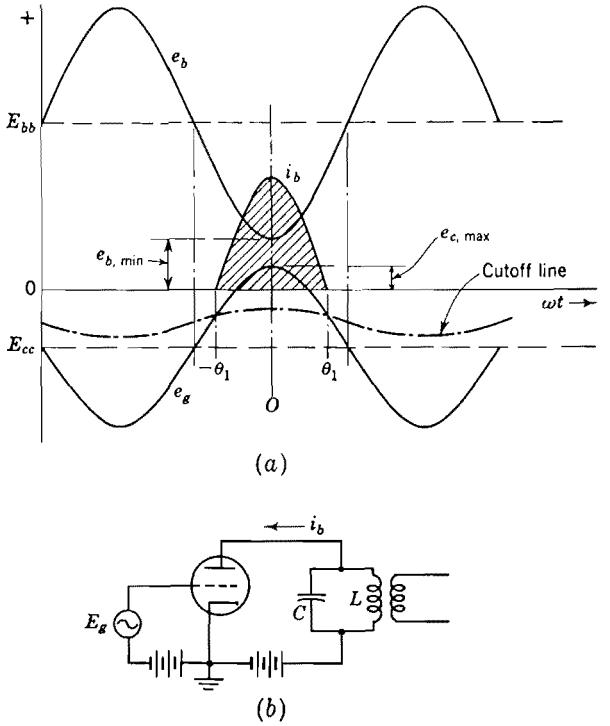



FIG. 11-1. (*a*) Waveforms in Class C amplifier. (*b*) Circuit for Class C amplifier.

to high-frequency heating of metals and dielectrics, power efficiency is important and the tubes are operated under Class C conditions. In this service the tubes act as synchronous switches, connecting d-c source and

load for small angles of conduction each cycle. The resonant loads then continue in free oscillation for the remainder of the cycle.

Because of the discontinuous nature of tube conduction in Class C amplifiers an equivalent-circuit approach is not possible, and graphical methods have been developed for analysis of Class C operation.

**11-1. Waveforms in Class C Amplifiers.** The Class C amplifier is operated with a bias of two or more times the cutoff grid voltage. A resonant-circuit load, or *tank circuit*, is used in the plate circuit, tuned to the frequency, or some multiple, of the grid input. With the grid well below cutoff only a small portion or angle of the positive grid input voltage will drive the tube into conduction, and the conduction angle will be less than 180°, occurring from $-\theta_1$ to $\theta_1$ as shown in Fig. 11-1*a*. The resonant load in the plate circuit, excited by these short current pulses, will go into oscillation in accordance with basic electric-circuit theory and develop a sinusoidal voltage across the tank.

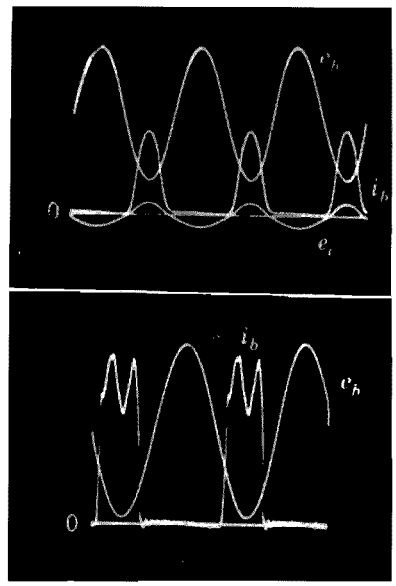



FIG. 11-2. (*a*) Plate-voltage, plate-current, and grid-voltage waves in Class C amplifier. (*b*) Distorted current wave in overdriven amplifier.

These waveforms and their phase relations are shown in Figs. 11-1 and 11-2. In the former $e_b$ illustrates the plate-voltage waveform as the resultant of $E_{bb}$ and the sinusoidal resonant-circuit voltage. The grid-voltage cutoff line is $e_b/\mu$ (for a triode) and also has a sinusoidal form for sinusoidal plate voltage. The sine applied voltage $e_g$ carries the grid above cutoff at $-\theta_1$, and the pulse of plate current appears while the grid is above cutoff. A pulse of grid current will also flow during the interval in which the grid is driven positive.

The vacuum tube, as a synchronous switch, connects $E_{bb}$ to the resonant load only for a short interval equal to $2\theta_1$. Sufficient energy must be given to the tank circuit during this interval to supply the power output and losses of the tank and to keep the circuit in oscillation until the time of the next pulse.

The major power loss of the circuit is in the vacuum tube. Its losses are given by $e_b i_b$ and can be reduced by causing the plate current to flow

in a very short high-amplitude pulse while the plate voltage is lowest. A high bias ensures a short conduction time at the peak of the input cycle while the plate voltage is minimum. This raises the plate power efficiency but requires a large input signal, which raises the input power to the grid.

An effective load resistance is inductively coupled into the tank circuit. This provides a means of taking power out of the tank but at the same time lowers the resonant impedance and $Q = \omega L/R$ of the tank. The efficiency may also be improved by increasing the resonant impedance or raising the $Q$ of the tank, thereby lowering the minimum plate voltage during the time of current flow. It is found, however, that, if $e_{b,\min}$ is allowed to fall below the positive peak of grid voltage $e_{c,\max}$, the grid current rises rapidly; and if current saturation is reached, the more positive grid will rob the anode of electrons and actually reduce the plate current, thereby lowering the power output. This is shown in Fig. 10-2*b*. It is found that maximum output usually occurs when $e_{b,\min} = e_{c,\max}$.

**11-2. Analysis of Class C Amplifier Performance.** Prediction of performance of Class C circuits with particular tubes is based on a knowledge of current and voltage waveforms present, determined from the tube characteristics by a graphical method. This approach is made necessary by the discontinuous nature of current flow. The method suffices to predict desirable or optimum values for load impedance, grid bias, a-c grid driving voltage, grid input power, and plate loss.

Operation of a Class C amplifier with tuned load is demonstrated in Fig. 11-3. Bias is usually fixed at two to five times cutoff and is shown as twice cutoff. The plate current flows for a conduction angle of $2\theta_1$, less than 180°. Grid current, if shown, would flow for a smaller angle during which the grid is positive. The instantaneous plate voltage applied through a resonant tank circuit of resistive impedance $R_L$ is

$$e_b = E_{bb} - \hat{I}_1 R_L \cos\theta \tag{11-1}$$

the *notation* $\hat{I}_1$ *indicating the peak value* of the fundamental component of a-c plate current. The instantaneous grid voltage is

$$e_c = E_{cc} + \hat{E}_g \cos\theta \tag{11-2}$$

For a theoretically linear dynamic curve, cutoff occurs at $\theta_1$ or where

$$e_c = -\frac{e_b}{\mu}$$

Cutoff bias is then calculated as

$$E_{cc} = -\frac{E_{bb}}{\mu} + \left(\hat{E}_g - \frac{\hat{I}_1 R_L}{\mu}\right)\cos\theta_1 \tag{11-3}$$

Equation (11-3) allows calculation of the bias required to give a particular value of conduction angle $\theta_1$.

Using $\hat{I}_1 R_L = \hat{E}_1$ and combining Eqs. (11-1) and (11-2) gives

$$e_b = E_{bb} + E_{cc}\frac{\hat{E}_1}{\hat{E}_g} - \frac{\hat{E}_1}{\hat{E}_g}e_c \tag{11-4}$$

Since $E_{cc}$ and $E_{bb}$ are circuit constants and $\hat{E}_1$ and $\hat{E}_g$ are peak values, the above appears as the equation of a straight line of form $x = A + B - ay$, indicating that the path of operation of the tube in Class C is a straight-line relation between $e_b$ and $e_c$.

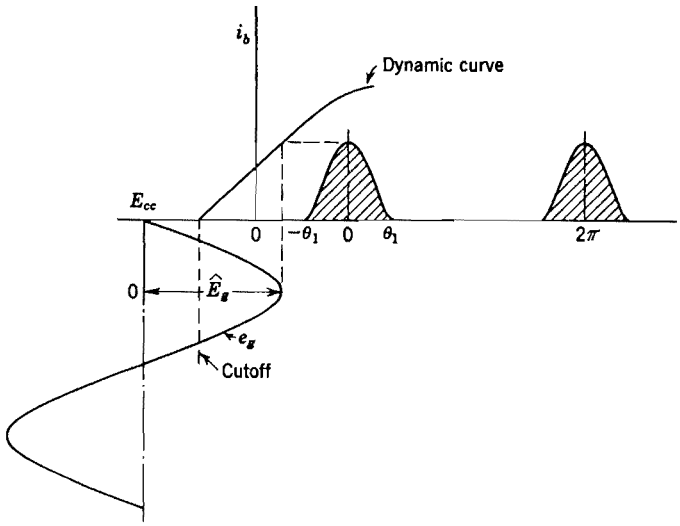



FIG. 11-3. Determination of waveforms.

Relations between $e_b$ and $e_c$ can be directly plotted on the constant-current family of tube characteristics as shown in Fig. 11-4. It is then possible to employ these curves to obtain the grid and plate voltages for various values of $\theta$. From these, waveforms may be plotted, or current values computed from which efficiency, plate loss, and other operating and design information can be obtained.

The operation line determined by Eq. (11-4) is drawn through any two points of the tube operation, or drawn with a slope equal to $-\hat{E}_1/\hat{E}_g$. A number of such lines may be drawn by trial selections of grid bias $E_{cc}$ and peak driving voltage $\hat{E}_g$, from which data covering a range of tube operating conditions can be obtained.

The first of the two points fixing the operation line may be the $Q$ point, or unexcited operating point, located by the value of $E_{bb}$ and $E_{cc}$ chosen. The value of $E_{cc} = -E_{bb}/\mu$ is a cut-and-try variable available to the

designer for use in finding the most desirable operating conditions. It can be seen that, as $E_{cc}$ is made more negative, the conduction angle decreases, with increased efficiency. Higher grid driving voltage and power are then required, and an optimum compromise is usually obtained where $e_{b,\min} = e_{c,\max}$. The second point fixing the operation line is usually for this condition, and in Fig. 11-4 a line for $e_{b,\min} = e_{c,\max}$ is drawn as a locus of the end of the operation line.

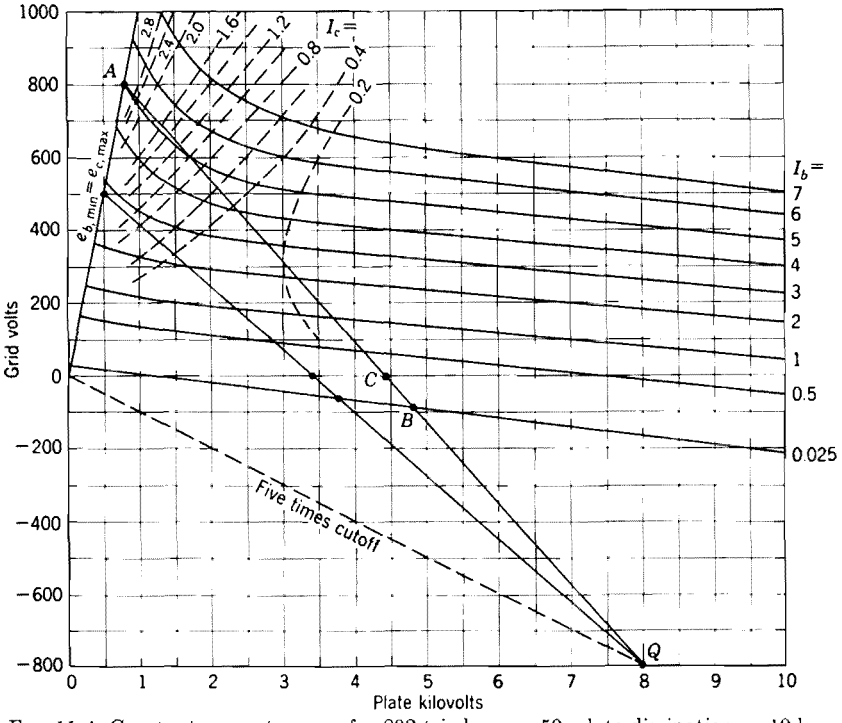



FIG. 11-4. Constant-current curves for 892 triode; $\mu = 50$, plate dissipation = 10 kw.

If the curves of Fig. 11-4 are used with a tube having $\mu = 50$, and $E_{bb} = 8000$ volts, a $Q$ point may be selected at $E_{cc} = -800$ volts, giving a bias five times cutoff. The second arbitrary choice of the designer is the selection of the voltage at which $e_{b,\min} = e_{c,\max}$. In the example of Fig. 11-4 this point $A$ is chosen for $e_c = 800$ volts, and a straight line for the path of operation is drawn between $Q$ and $A$. From this line values of $i_b$ and $i_c$ can be obtained as functions of $\omega t$.

Point $A$ determines that

$$\hat{E}_1 = 8000 - 800 = 7200 \text{ volts peak}$$
$$\hat{E}_g = 800 - (-800) = 1600 \text{ volts peak}$$

Point $B$ is placed at the $i_b = 0$ line, with $e_b = 4850$ volts, so that the load volts = 8000 − 4850 = 3150 volts. The angle of conduction $2\theta_1$ is then

$$2\theta_1 = 2\cos^{-1}\frac{3150}{7200} = 128°$$

At point $C$, where $e_b = 4400$ volts, the grid goes positive, and grid current flows. If the grid conduction angle is $2\theta_2$, then

$$2\theta_2 = 2\cos^{-1}\frac{8000 - 4400}{7200} = 120°$$

The grid conduction angle could also have been found from

$$2\theta_2 = 2\cos^{-1}\frac{E_{cc}}{\hat{E}_g}$$

Values of current and voltage for any part of the conduction angle can be read from the operation line. Results taken at 10° intervals are

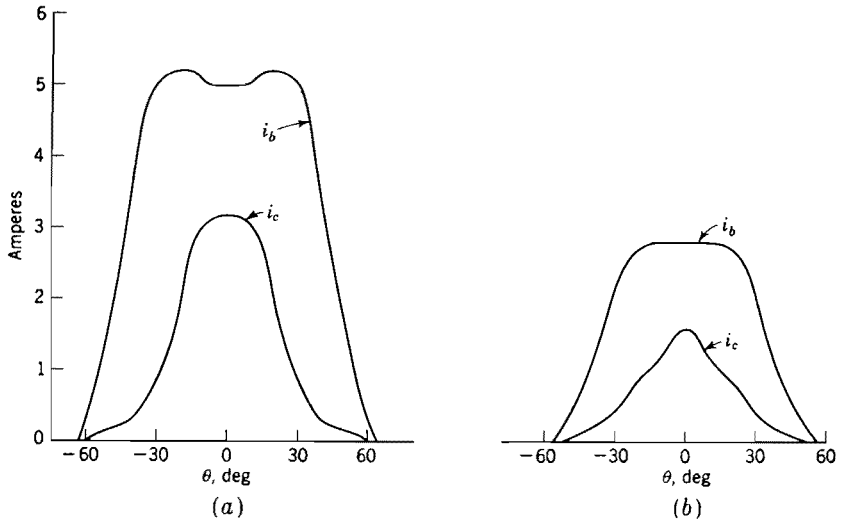



FIG. 11-5. Current waveforms. (*a*) $e_{b,\min} = e_{c,\max} = 800$ volts; (*b*) $e_{b,\min} = e_{c,\max} = 500$ volts.

tabulated in Table 19, and current waveforms are plotted in Fig. 11-5. Also presented for comparison are the waveforms obtained from a second operation line using the same $Q$ point at $E_{cc} = -800$ volts, $E_{bb} = 8000$ volts, but with reduced grid drive and increased load impedance so that $e_{b,\min} = e_{c,\max} = 500$ volts.

Graphical integration may be used to obtain the average values of $i_b$ and $i_c$ to give $I_b$ and $I_c$. If the base of the area under a curve $y = f(x)$

be divided, in an interval $a$ to $b$, into uniform intervals of length $\Delta x$, then

$$\int_a^b y\,dx = \Delta x\left(\frac{y_o}{2} + y_1 + y_2 + \cdots + y_{n-1} + \frac{y_n}{2}\right) \qquad (11\text{-}5)$$

Since

$$I_b = \frac{1}{2\pi}\int_{-\pi}^{\pi} i_b\,d\theta$$

then by the trapezoidal rule above, with $y_o$ taken for $\theta = 0°$, $y_n = 0$, and $\Delta x = 10° = \pi/18$,

$$I_b = \frac{2}{2\pi}\frac{\pi}{18}\left(\frac{y_o}{2} + y_1 + y_2 + \cdots + y_{n-1}\right) \qquad (11\text{-}6)$$

The coefficient 2 is present since $y_o, \ldots, y_n$ covers only one-half of the conduction angle under the assumed cosine symmetry.

TABLE 19. DATA FROM THE OPERATING LINE OF FIG. 11-4

$\mu = 50$ $E_{cc} = -800$ volts $e_{b,\min} = e_{c,\max} = 800$ volts
$E_{bb} = 8000$ volts $\hat{E}_g = 1600$ volts

| | $\theta$, deg | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 0 | 10 | 20 | 30 | 40 | 50 | 60 | 70 |
| $\cos\theta$ | 1.0 | 0.985 | 0.940 | 0.866 | 0.760 | 0.643 | 0.500 | 0.342 |
| $\hat{E}_1\cos\theta$ | 7200 | 7100 | 6770 | 6250 | 5470 | 4620 | 3600 | 2460 |
| $E_{bb} - \hat{E}_1\cos\theta$ | 800 | 900 | 1230 | 1750 | 2530 | 3380 | 4400 | 5540 |
| $i_b$, amp | 5.0 | 5.0 | 5.2 | 5.0 | 3.5 | 1.7 | 0.25 | |
| $i_c$, amp | 3.2 | 3.0 | 1.7 | 1.0 | 0.3 | 0.18 | 0.0 | |

Similar operations lead to the value of $I_c$, the grid current.

The peak value of the fundamental a-c component of plate current can be obtained from the Fourier expansion for the current wave,

$$i_b = I_b + \hat{I}_1\cos\theta + \hat{I}_2\cos 2\theta + \cdots + \hat{I}_n\cos n\theta$$

The coefficient $\hat{I}_1$ may be found by multiplying both sides by $\cos\theta\,d\theta$ and integrating between the limits of 0 and $2\pi$. All terms on the right except the second go to zero, leaving

$$\int_{-\pi}^{\pi} i_b\cos\theta\,d\theta = \int_{-\pi}^{\pi}\hat{I}_1\cos^2\theta\,d\theta = \pi\hat{I}_1 \qquad (11\text{-}7)$$

from which

$$\hat{I}_1 = \frac{1}{\pi}\int_{-\pi}^{\pi} i_b\cos\theta\,d\theta \qquad (11\text{-}8)$$

The trapezoidal rule may again be applied, taking successive products of $i_b$ and $\cos\theta$ at 10° intervals. If $x_o$ is the $i_b\cos\theta$ product at 0°, $x_1$ at

10°, and so on, then

$$\hat{I}_1 = \frac{2}{\pi}\frac{\pi}{18}\left(\frac{x_o}{2} + x_1 + x_2 + \cdots + x_{n-1}\right) \tag{11-9}$$

and a similar process leads to $\hat{I}_{g1}$ for the grid current.

These calculations have been carried out for the operation line of Fig. 11-4, and for the second operation line with $e_{b,\min} = e_{c,\max} = 500$ volts. The results appear at (*a*) and (*b*) in Table 20.

The method is equally applicable for a tetrode or beam tube. It should be noted, however, that cutoff bias for these types is determined as $-E_{c2}/\mu_{sg}$, where $\mu_{sg}$ is the grid-screen amplification factor. A tetrode acts much as a triode with the screen as anode, the actual anode serving largely as a current collector if the shielding is good.

TABLE 20. PERFORMANCE OF TRIODE FROM FIG. 11-4

| | (*a*) | (*b*) |
|---|---|---|
| $E_{bb}$, volts | 8000 | 8000 |
| $E_{cc}$, volts | −800 | −800 |
| $e_{b,\min} = e_{c,\max}$, volts | 800 | 500 |
| $2\theta_1$, deg | 128 | 56 |
| $\hat{E}_g$, volts | 1600 | 1300 |
| $\hat{E}_1$, volts | 7200 | 7500 |
| $I_b$, amp | 1.28 | 0.59 |
| $I_c$, amp | 0.43 | 0.21 |
| $\hat{I}_1$, amp | 2.28 | 1.08 |
| $\hat{I}_{g1}$, amp | 0.81 | 0.37 |

**11-3. Power Relations.** The power input to the plate circuit of the Class C amplifier is

$$P_{in} = E_{bb}I_b \tag{11-10}$$

and the fundamental a-c output to the load is

$$P_{out} = \frac{\hat{I}_1^2 R_L}{2} \tag{11-11}$$

where $R_L$ is the resistive impedance of the resonant load. The plate loss in the tube then is

$$P_d = P_{in} - P_{out}$$

and the plate-circuit efficiency is

$$\eta_p = \frac{P_{out}}{P_{in}} \times 100\% \tag{11-12}$$

Power computations from the two operation lines studied from Fig. 11-4

are tabulated in Table 21. Efficiencies of 80 and 85 per cent are indicated and may be taken as typical.

In the grid circuit the input a-c power is supplied from an external source for an amplifier and from the output circuit for an oscillator. Grid bias may be obtained from the d-c grid current and the grid resistor, of value

$$R_g = \frac{E_{cc}}{I_c}$$

or may be furnished by a rectified and filtered supply. The total power input to the grid circuit is

$$P_c = \frac{1}{2\pi} \int_{-\pi}^{\pi} i_c E_g \cos \theta \, d\theta \tag{11-13}$$

and these quantities are available for integration as in Eq. (11-6).

It may be noted that the larger grid-current values flow during the time near the peak of the grid-voltage wave, or when $\cos \theta$ approximates unity. If this assumption be made, then

$$P_c \cong \frac{\hat{E}_g}{2\pi} \int_{-\pi}^{\pi} i_c \, d\theta$$

However,

$$I_c = \frac{1}{2\pi} \int_{-\pi}^{\pi} i_c \, d\theta$$

so that the total grid input power is approximately

$$P_c \cong \hat{E}_g I_c \tag{11-14}$$

where $I_c$ is the d-c grid current. For the data of Table 21*a*, 688 watts would be obtained for the grid driving power, whereas Eq. (11-13) gives the more exact answer of 646 watts.

The grid current produces a loss in the grid leak or bias source of $P = E_{cc} I_c$ so that the power lost as heat on the grid structure owing to electron bombardment is

$$P_g = P_c - E_{cc} I_c \cong \hat{I}_c (\hat{E}_g - E_{cc}) \tag{11-15}$$

Since the power required to drive the grid circuit is expensive in terms of associated apparatus, it should be minimized, and a figure indicating the ability of an amplifier to operate on small relative input is the *power gain*, or the ratio of a-c plate power output to a-c grid input, or

$$\text{Power gain} = \frac{P_{out}}{P_c} \tag{11-16}$$

Tetrodes show very advantageous power gains when compared with triodes.

Electron tubes are rated in terms of allowable internal losses, rather than output as for most electrical machinery. The effect of operating conditions on efficiency is therefore very important. The calculated

TABLE 21. POWER RELATIONS FROM FIG. 11-4

| | (a) | (b) |
|---|---|---|
| $E_{bb}$, volts | 8,000 | 8,000 |
| $E_{cc}$, volts | −800 | −800 |
| $I_b$, amp | 1.28 | 0.59 |
| $\hat{E}_1$, volts | 7,200 | 7,500 |
| $I_c$, amp | 0.43 | 0.21 |
| $\hat{E}_g$, volts | 1,600 | 1,300 |
| $\hat{I}_1$, amp | 2.28 | 1.08 |
| $\hat{I}_{g1}$, amp | 0.81 | 0.37 |
| $2\theta_1$, deg | 128 | 112 |
| $R_L$, ohms | 3,160 | 6,950 |
| Power input, watts | 10,250 | 4,720 |
| Power output, watts | 8,200 | 4,050 |
| Plate loss, watts | 2,050 | 670 |
| $\eta_p$, % | 80 | 85.8 |
| Grid driving power, watts | 646 | 244 |
| Grid loss, watts | 302 | 84 |
| Power gain | 11.9 | 15.6 |

example shows the effect of changing load impedance on power input, efficiency, and power output. Relations giving optimum operating conditions may be found by assuming $e_{b,\min} = e_{c,\max}$ and trying several operating lines through a given $Q$ point. Tabulation of the results allows the designer to choose $R_L$, $\hat{E}_g$, $E_{cc}$ and to determine that plate loss is within tube ratings.

**11-4. The Resonant-load Circuit.** A parallel resonant circuit of impedance given by $R_L$ of the previous section is used as a load on amplifiers and oscillators. An external resistance is coupled in through the mutual inductance $M$ of Fig. 11-6, and $R$ represents the series resistance of both $L$ and the coupled-in resistance from the external load $R_x$.

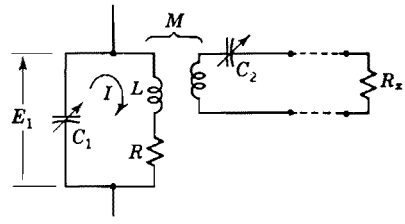



FIG. 11-6. Resonant-load circuit.

For resonant circuits a figure of merit, $Q$, is defined as

$$Q = \frac{\omega_r L}{R}$$

where $\omega_r$ is the resonant angular frequency of the circuit. For reasonable $Q$ values ($Q > 10$) $\omega LI$ approximates $E_1$, the load voltage, and

$$Q = \frac{(\omega_r LI)I}{RI^2} \cong \frac{E_1 I}{RI^2} = \frac{\text{volt-amp}}{\text{watts}} \tag{11-17}$$

and $Q$ approximates the ratio of the circulating volt-amperes to the power supplied to $R$. The input occurs during a short current pulse, but energy is taken from the circuit continuously. If the waveform is not to be distorted, which would indicate the presence of undesirable harmonic frequencies, the ratio of circulating energy to extracted energy must be high. Similarly the energy extracted from a pendulum during its swing must be small compared with the energy stored in the bob, or the rate of swing of the pendulum may be affected. This argument indicates a practical working minimum value of loaded $Q$ of approximately 10.

The resonant impedance of a parallel circuit is

$$R_L = \frac{L}{CR} \tag{11-18}$$

The power supplied to the circuit is

$$P_{out} = \frac{\hat{E}_1^2}{2R_L}$$

so that for design of a circuit of desired resonant impedance

$$\frac{L}{C} = \frac{R\hat{E}_1^2}{2P_{out}}$$

or

$$X_C = \frac{1}{\omega_r C} = \frac{\hat{E}_1^2}{2QP_{out}} = \frac{E_1^2}{QP_{out}} \tag{11-19}$$

where $E_1$ is the rms value. From resonant-circuit theory

$$X_L = \omega_r L = X_C\left(1 - \frac{1}{Q^2}\right) \tag{11-20}$$

which allows complete design of the output circuit.

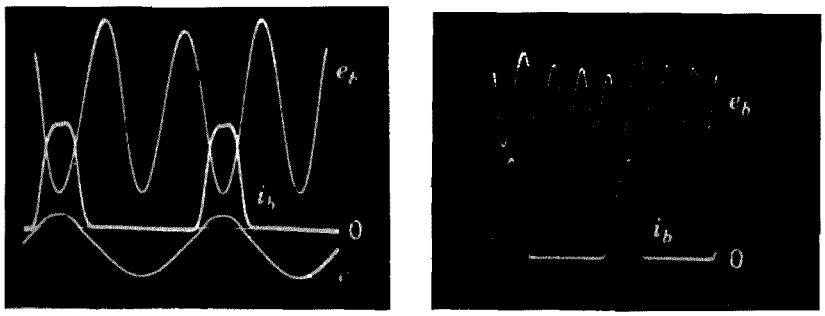
Computation of the effects of coupling a load $R_x$ into the tuned circuit is handled by usual coupled-circuit methods. Referring to Fig. 11-6, if $C_2$ is chosen to cancel the reactance of $L_2$ at $\omega_r$, then the impedance reflected into the resonant circuit is

$$R' = \frac{\omega_r^2 M^2}{R_2 + R_x} \tag{11-21}$$

where $R_2$ is the resistance of $L_2$. The load $R$ is then composed of $R_1$, the resistance of $L_1$, in series with $R'$.

Tuning of the load by varying $C_1$ produces resonance and maximum impedance and raises $\hat{E}_1$, thus reducing the average plate voltage and lowering the d-c plate current. Thus resonance and proper load tuning are indicated by a dip in the reading of the d-c plate-current meter of the amplifier. The circuit may also be tuned for maximum $\hat{E}_1$ or for maximum $I$ by use of suitable instruments.

**11-5. Frequency Multipliers.** Because of the large harmonic content in the plate current of a Class C amplifier it is possible to employ it as a *frequency multiplier* to double, triple, or quadruple the input frequency. The input signal is at the lower frequency, and the plate-circuit load is




(*a*) (*b*)

FIG. 11-7. (*a*) Double-frequency plate voltage, current, and grid input voltage of doubler. (*b*) Plate voltage and current in frequency quadrupler.

tuned to resonance at the desired harmonic frequency. A current pulse is supplied as controlled by the input frequency, and the resonant plate tank continues to oscillate between current pulses, as shown in Fig. 11-7.

For reasonable efficiency the conduction angle should not exceed the time of a half cycle of the plate voltage, and for doubling this fixes $2\theta_1$ at 90° or less of the input cycle. Larger angles lead to poor efficiency since current will flow at times of high plate voltage. This situation is indicated in the quadrupler of Fig. 11-7*b*, and low efficiency is to be expected.

To make the conduction angle short, a high bias voltage should be used, along with a large $\hat{E}_g$ value. A high-$\mu$ tube makes an excellent multiplier, and tetrodes or beam tubes are preferred, since the large grid drive can be easily obtained.

**11-6. Feedback in Oscillators.** Vacuum-tube oscillators are the usual generating sources for all frequencies above about 10 kc. While a wide variety of oscillator circuits exist, the fundamental requirement for oscillation is the presence of positive feedback from output to input.

In Fig. 11-8 is shown an amplifier of gain

$$\mathbf{A} = \frac{\mathbf{E}_2}{\mathbf{E}_1} \tag{11-22}$$

As in Chap. 7, a feedback network of characteristic $\beta$ provides a voltage

$$\mathbf{E}_f = \beta \mathbf{E}_2$$

As connected, the feedback network must so adjust the magnitude and phase of $\beta\mathbf{E}_2$ that $\mathbf{E}_f$ is identical with $\mathbf{E}_1$ in both phase and magnitude. If the $\beta$ network so operates that $\mathbf{E}_f = \mathbf{E}_1$, then

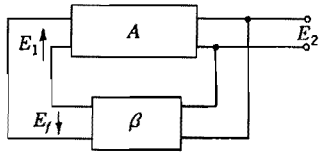


FIG. 11-8. Feedback oscillator.

$$\mathbf{A} = \frac{\mathbf{E}_2}{\beta \mathbf{E}_2} = \frac{1}{\beta}$$

$$\mathbf{A}\beta = 1 \tag{11-23}$$

becomes *a requirement for oscillation*, such that a circuit, closed upon itself, shall continue to provide an output $\mathbf{E}_2$.

Since $\beta$ can never be greater than unity in magnitude, then for oscillations to occur or for Eq. (11-23) to be satisfied the gain $\mathbf{A}$ can never be less than unity in magnitude.

The use of the general gain relation

$$\mathbf{A} = \frac{-\mu \mathbf{Z}_L}{r_p + \mathbf{Z}_L}$$

in Eq. (11-23) gives

$$\beta = -\left(\frac{1}{\mu} + \frac{1}{g_m \mathbf{Z}_L}\right) \tag{11-24}$$

which is known as *the Barkhausen criterion for sustained oscillation*. If $\mathbf{Z}_L$ is resistive, the minus sign represents the 180° phase reversal necessary in the $\beta$ network to provide the proper phase for $\mathbf{E}_1$. If $\mathbf{Z}_L$ is an impedance, an additional phase shift is necessary.

Equations (11-23) and (11-24) define the feedback specifications necessary for oscillation. The manner in which the feedback is obtained is the differentiating factor between the various circuit forms of Sec. 11-8.

Because of the infinite gain possible from the feedback relation

$$\mathbf{A}' = \frac{\mathbf{A}}{1 - \mathbf{A}\beta}$$

when $\mathbf{A}\beta = 1$, it is necessary that some circuit element operate to limit the amplitude and prevent it from growing without limit. This limiting action is usually provided by the nonlinear cutoff and saturation regions

of the tube characteristics, in which $\mu$ and $g_m$ reach values at which the Barkhausen criterion cannot longer be sustained. Thus distortion of waveform and harmonic generation are inherent in self-limitation of oscillator amplitude. If purer waveform is required, then external amplitude-limiting and other circuit elements must be employed.

**11-7. Circuit Criteria for Feedback Oscillators.** The circuit of Fig. 11-9*a* is known as a *tuned-plate oscillator*. A simple circuit analysis requires the assumption of zero grid current, constant tube coefficients, and sinusoidal waveforms, all requirements not entirely met in oscillator circuits, which ordinarily operate in Class C with discontinuous plate current and over regions of varying tube coefficients. However, the use here of the triode equivalent circuit is only an approximation helpful for an understanding of circuit requirements, and without quantitative intent.

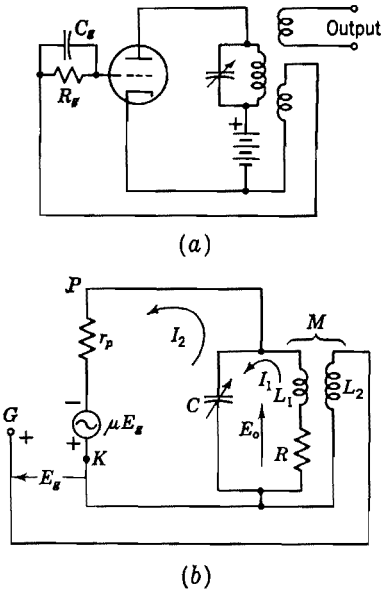



FIG. 11-9. (*a*) Tuned-plate oscillator. (*b*) Equivalent circuit for (*a*).

Under the assumptions the following equations may be written for the equivalent circuit of Fig. 11-9*b*:

$$\mathbf{E}_g - j\omega M\mathbf{I}_1 = 0 \tag{11-25}$$

$$\mu\mathbf{E}_g - (R + j\omega L_1)\mathbf{I}_1 - r_p\mathbf{I}_2 = 0 \tag{11-26}$$

$$\left(R + j\omega L_1 - \frac{j}{\omega C}\right)\mathbf{I}_1 + \frac{j\mathbf{I}_2}{\omega C} = 0 \tag{11-27}$$

The output voltage appears across the plate load as

$$\mathbf{E}_o = -(R + j\omega L_1)\mathbf{I}_1 \tag{11-28}$$

The voltage fed back to the input is $\mathbf{E}_g$, so that from Eqs. (11-25) and (11-28) it is possible to write an expression for the feedback $\beta$, as

$$\beta = \frac{\mathbf{E}_g}{\mathbf{E}_o} = \frac{j\omega M\mathbf{I}_1}{-(R + j\omega L_1)\mathbf{I}_1} = \frac{-j\omega M}{R + j\omega L_1} \tag{11-29}$$

Equations (11-26) and (11-27) may be solved for $\mathbf{I}_1$, allowing $\mathbf{E}_o$ to be written as

$$\mathbf{E}_o = \frac{-\mu\mathbf{E}_g(R + j\omega L_1)}{R + j\omega L_1 + r_p + j\omega C r_p(R + j\omega L_1)}$$

from which the gain may be obtained as

$$\mathbf{A} = \frac{\mathbf{E}_o}{\mathbf{E}_g} = \frac{-g_m}{1/r_p + 1/(R + j\omega L_1) + j\omega C} \tag{11-30}$$

For oscillation to occur, the criteria of Sec. 11-6 must be satisfied, or $\mathbf{A}\beta = 1$, and this will be required for the tuned-plate oscillator. Using Eqs. (11-29) and (11-30),

$$\mathbf{A}\beta = \frac{-j\omega M}{R + j\omega L_1} \frac{-g_m}{1/r_p + 1/(R + j\omega L_1) + j\omega C} = 1 \tag{11-31}$$

$$\frac{j\omega M}{R + j\omega L_1} = \frac{1}{\mu} + \frac{1}{g_m(R + j\omega L_1)} + \frac{j\omega C}{g_m}$$

Multiplying by $(R + j\omega L_1)/j\omega L_1$ leads to

$$\left(\frac{M}{L_1} - \frac{1}{\mu} - \frac{RC}{g_m L_1}\right) + j\left(\frac{R}{\mu\omega L_1} + \frac{1}{\omega L_1 g_m} - \frac{\omega C}{g_m}\right) = 0 \tag{11-32}$$

Equating the real terms to zero gives

$$\frac{M}{L_1} = \frac{1}{\mu} + \frac{RC}{g_m L_1}$$

and since by Eq. (11-18) $L_1/RC$ is the resonant impedance $R_L$ of the load circuit, then

$$-\frac{M}{L_1} = -\left(\frac{1}{\mu} + \frac{1}{g_m R_L}\right) \tag{11-33}$$

which is recognized as the Barkhausen criterion for oscillation. The circuit will satisfy this criterion provided that $M$ is negative. The ratio $M/L_1$ is equal to $\beta$, the feedback.

Equating the reactive terms of Eq. (11-32) to zero,

$$\omega^2 LC = 1 + \frac{R}{r_p}$$

from which

$$\omega = \sqrt{\frac{1}{LC}\left(1 + \frac{R}{r_p}\right)} \tag{11-34}$$

as a second condition required for oscillation, satisfied by the automatic adjustment of frequency to this value, making the reactive term zero. Since $R$ will ordinarily be small with respect to $r_p$, the oscillation frequency will closely approximate the resonant frequency of the tuned circuit alone. It should be noted, however, that $R$ includes the coupled-in load resistance, so that the oscillator frequency is a function of oscillator loading. For high stability of frequency it is desirable to isolate the oscillator tuned circuit from variable loading effects.

For oscillation the gain must be greater than unity, and the feedback must allow satisfaction of the Barkhausen criterion. The frequency will then adjust to satisfy Eq. (11-34). The Barkhausen criterion may be rearranged to give

$$g_m \geq \frac{1}{R_L(M/L_1 - 1/\mu)} \tag{11-35}$$

as a limiting value for $g_m$. For starting of oscillations with zero bias $g_m$ must be greater than the value given. During actual operation grid current flows through $R_g$, giving a value of bias sufficient for Class C conditions, and providing an average $\mu$ and $g_m$ to satisfy the Barkhausen condition.

The greater the amplitude of oscillation, the greater are the grid current and grid bias. This increased bias causes $g_m$ to fall, reducing oscillation amplitude until an equilibrium is reached. Thus the use of grid-leak bias permits a large $g_m$ value at zero bias for easy starting of oscillation and yet provides automatic limitation of oscillation amplitude. Low grid driving power and high frequency stability dictate large values of $R_g$, but too high values cause unstable and intermittent oscillation.

**11-8. Basic Feedback-oscillator Circuits.** The various types of feedback oscillators illustrated in Fig. 11-10 differ largely in the manner in which feedback from plate to grid is accomplished. The circuit of Fig. 11-9*a* obtains feedback through mutual-inductance coupling, with tuning in the plate circuit.

The circuit of Fig. 11-10*a* is the converse, with tuning in the grid circuit. The circuit is ordinarily used as a local signal source in receivers and is obviously a *tuned-grid oscillator*. The circuit of Fig. 11-10*b* employs both grid and plate tuning, with the respective coils shielded to prevent inductive coupling. Feedback takes place through the grid-plate capacity of the tube, and tetrodes may not oscillate readily unless parallel external capacity is added. The tuned circuit of higher $Q$ determines the frequency of oscillation. Since the plate circuit ordinarily supplies an external load, the grid circuit will usually have the higher $Q$ and will control the frequency. The plate circuit may then be tuned to adjust oscillator plate loading and Class C operating conditions. The circuit is known as a *tuned-plate–tuned-grid oscillator*.

The *Hartley circuit* of Fig. 11-10*c* and *d* is a popular form, with feedback obtained through mutual coupling between the two portions of the inductance. The position of the cathode tap is adjustable to control the magnitude of grid driving voltage and oscillator output and efficiency. The series-fed form of Fig. 11-10*c* is not ordinarily used because of the necessity for dividing the tank coil. The shunt form of Fig. 11-10*d* is usually employed, a high-frequency choke preventing the high-frequency

currents from flowing in the plate power source and shunting the tuned circuit.

The *Colpitts circuit* of Fig. 11-10*e* is the converse of the Hartley, the cathode tap being taken between the two tuning capacitors. Feedback occurs through the common tank coil. The tuning capacitors are ordinarily equal and gang-tuned, so that choice between Colpitts and Hartley

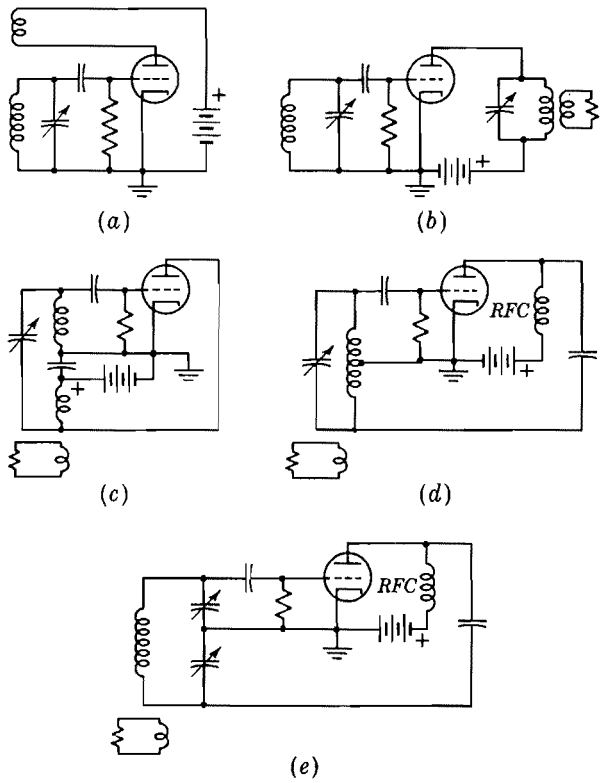



FIG. 11-10. (*a*) Tuned-grid oscillator. (*b*) Tuned-plate–tuned-grid oscillator. (*c*) Series-fed Hartley. (*d*) Shunt-fed Hartley. (*e*) Colpitts oscillator, shunt-fed.

circuits is largely a matter of desire for control of grid drive through the cathode tap in the Hartley or the two-terminal coil of the Colpitts.

**11-9. Electron Coupling; Crystal Control.** High stability of output frequency is usually a requirement for an oscillator. One method of improving stability was pointed out in Sec. 11-7, that of decoupling the load so that the $R/r_p$ term of Eq. (11-34) is made very small, reducing the effects of changes in both $R$ and $r_p$. This decoupling may be achieved by an intervening lightly loaded amplifier, or *buffer*, stage or by the *electron-coupled* circuit of Fig. 11-11.

In the latter circuit the cathode, control grid, and screen grid of a tetrode or pentode function as a triode Hartley oscillator, but with the screen grounded in place of the cathode. Electrons pass through the screen, reach the plate, and set up a voltage across the tuned circuit. Since in a well-screened tetrode the anode current is unaffected by anode voltage, then load-circuit changes have negligible effect on the frequency-determining grid circuit. The circuit is said to be electron-coupled because the coupling between grid and output circuits is only by the electron stream. In effect the circuit combines a triode oscillator and a buffer amplifier in one tube.

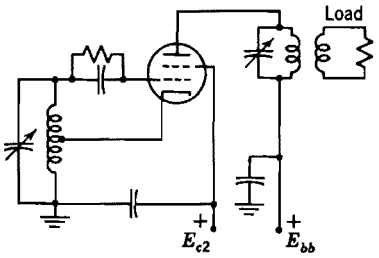



FIG. 11-11. Electron-coupled oscillator.

Even greater stability may be achieved by replacing the frequency-determining circuit of an oscillator with a piezoelectric crystal, usually of quartz. In such crystals electrical charges appear on one pair of faces if mechanical stresses are applied on another pair of faces. The converse effect, production of mechanical stress and displacement by electrical charge or potential, is also present. An alternating voltage applied to electrodes plated on opposite faces causes the crystal to vibrate mechanically, and at the mechanical resonant frequency these vibrations and charges become very large. The mechanical resonant frequency is dependent on crystal dimensions, the angle to the crystal axes at which the crystal was cut, and the allowed type of mechanical movement. Since quartz is very stable with time and temperature, the resonant frequency is highly constant.

Figure 11-12*a* shows the locations of the basic $X$ and $Y$ crystal cuts in the hexagonal quartz crystal. For the resultant crystal slabs the $X$ cut has a negative temperature coefficient of frequency and a frequency value given by

$$f = \frac{2.86 \times 10^6}{t}$$

where $t$ is in millimeters. The $Y$-cut crystal has a positive temperature coefficient and a frequency given by

$$f = \frac{1.96 \times 10^6}{t}$$

The above indicates the possibility of an intermediate crystal orientation having a zero temperature coefficient and these exist as $AT$, $BT$, and $GT$ cuts.

Electrically the crystal, with plated electrodes, has an equivalent circuit as in Fig. 11-12*b*, where $C_h$ is the capacity of the electrodes with quartz dielectric. The $Q$ of such a crystal will run very high, from

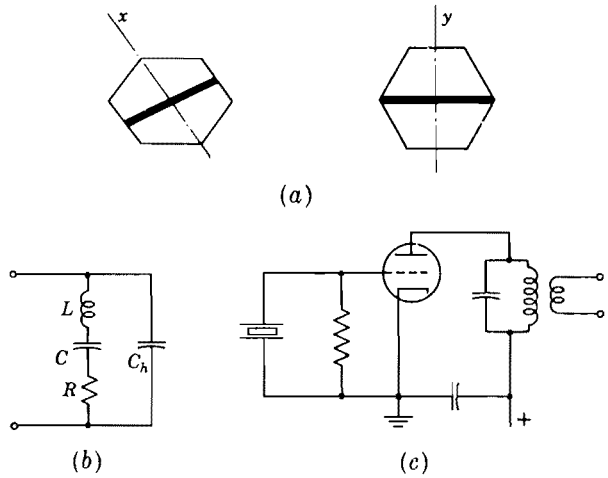



FIG. 11-12. (*a*) Basic crystal cuts. (*b*) Equivalent circuit for a quartz crystal. (*c*) Crystal oscillator.

10,000 to 500,000 for special mountings, and it is this high $Q$ which gives the crystal oscillator its frequency stability. The circuit of Fig. 11-12*c* is thus a tuned-plate–tuned-grid form, with the crystal forming the frequency-determining grid tank, and with feedback through $C_{gp}$. For proper phase of the feedback voltage the plate tank must be tuned on the inductive side of resonance.

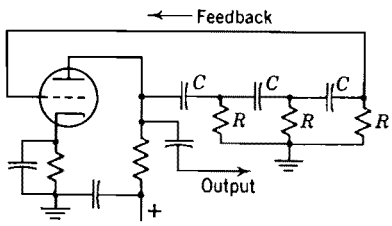



FIG. 11-13. Phase-shift oscillator.

Crystals can be made with fundamental frequencies ranging from about 4 kc to 10 Mc, and higher frequencies can be generated by overtone or harmonic vibration of the crystals. For zero-temperature-coefficient crystals operated in well-designed circuits, stability to an accuracy of 1 to 10 cycles per Mc may be obtained over long periods of time.

**11-10. Phase-shift Oscillators.** Oscillators need not employ inductive elements if the proper phase shift is provided by other means in the feedback network. The circuit of Fig. 11-13 oscillates at a frequency at which the total phase shift in the $RC$ network is 180°. It is suited to fixed-frequency operation.

A second circuit using $RC$ networks for positive feedback and phase shift is shown in Fig. 11-14. Neglecting the reactance of $C$ as small, and defining

$$\mathbf{Z}_1 = R_1 - \frac{j}{\omega C_1} \qquad \mathbf{Z}_2 = \frac{-jR_2/\omega C_2}{R_2 - j/\omega C_2}$$

then with amplifier gain $\mathbf{A}$ the requirement for oscillation is

$$\mathbf{A}\beta = 1 = \mathbf{A}\frac{\mathbf{Z}_2}{\mathbf{Z}_1 + \mathbf{Z}_2} \tag{11-36}$$

From this equation it is found that two circuit conditions must be fulfilled to provide proper feedback magnitude and phase for oscillation. These are

$$\omega = \sqrt{\frac{1}{R_1 R_2 C_1 C_2}} \qquad \text{and} \qquad \mathbf{A} = 1 + \frac{R_1}{R_2} + \frac{C_2}{C_1}$$

thus specifying frequency and amplifier gain.

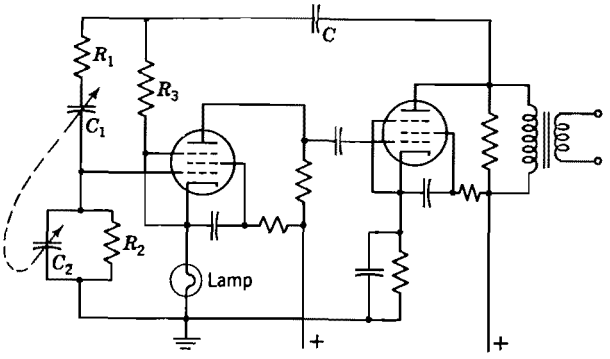



FIG. 11-14. Practical form of $RC$ oscillator.

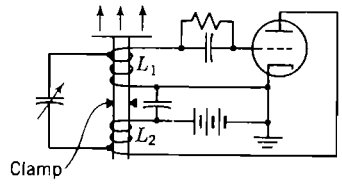
For a practical circuit $R_1 = R_2$, and $C_1 = C_2$, giving

$$\omega = \frac{1}{R_1 C_1} \qquad \text{and} \qquad \mathbf{A} = 3$$

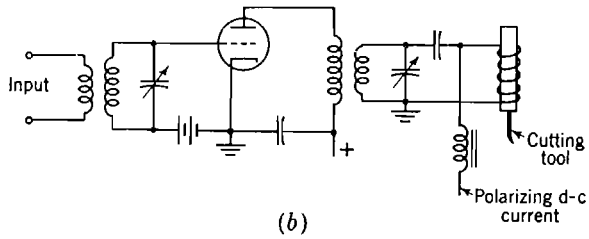
Negative feedback is added by $R_3$ and the tungsten lamp to limit the oscillation amplitude and improve the waveform. Increased output gives increased filament temperature and resistance in the lamp, and greater negative feedback, thus stabilizing output amplitude.

A major advantage of the circuit is the wide frequency range, frequency being inversely proportional to $C$ and not to $\sqrt{C}$ as in an $LC$ oscillator. Thus decade, or 10:1, frequency ranges are possible with the usual variable capacitors.

**11-11. Ultrasonic Oscillators.** It has been found that sound waves above audible frequencies, in the ultrasonic range, are valuable means of providing high mechanical accelerations and forces well suited to emulsification of liquids, speeding of chemical reactions and phenomena, metal cutting, and in medical processes. For development of intense sound fields in solutions and materials at frequencies from a few thousand cycles to many megacycles, two *transducers*, or energy converters, are ordinarily used with electronic oscillators. These are the magnetostrictive bar and the piezoelectric crystal, both useful for converting oscillating electric energy into mechanical sound energy.




(*a*)




(*b*)

FIG. 11-15. (*a*) Magnetostrictive oscillator. (*b*) Driven magnetostrictive transducer.

In the magnetostrictive oscillator of Fig. 11-15*a* a rod of nickel, invar, or other nickel alloy is employed as a coupling means between the two coils. When such a rod is placed in a magnetic field, magnetic stresses are set up which result in a change in length. To accentuate this change, the rod is made resonant or one-half wave in length, this length being determined as

$$\lambda = \frac{v}{f}$$

where $v$ = velocity of sound in material (about 5000 m/sec for nickel)
$\lambda$ = wavelength, m/sec
$f$ = frequency

A change in plate current changes the magnetic flux in the rod inside the plate coil. This causes an elongation of this end of the rod and sets up a longitudinal wave moving toward the other end of the rod. The

wave changes the length of the rod inside the grid coil, causing a change in flux and inducing a voltage in the grid coil. This grid voltage causes the plate current to increase further, and the circuit goes into oscillation at the resonant frequency of the rod, accentuated by the electrical tuning of the coils. While the circuit appears to be a Hartley, actually the coils are wound to oppose oscillation and are normally shielded from each other, coupling being only through the rod.

The rod must be normally polarized by a permanent magnet, since the change in length is of the same sign for both polarities of magnetization. Without polarization the rod would vibrate at twice the frequency of the applied magnetization. Necessity for the rods having a reasonable physical length limits the frequency to about 100 kc.

One end of the rod may be loaded to give a suitable impedance match and the resulting sound waves used to drive liquid or solid materials. At resonance the rod motion may approximate 0.001 in. The circuit of

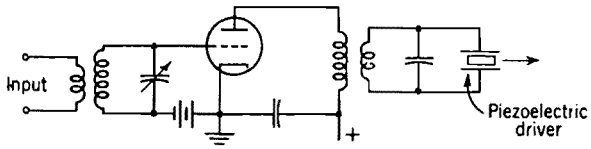



FIG. 11-16. Piezoelectric ultrasonic driver.

Fig. 11-15*b* illustrates the manner in which a magnetostrictive transducer may be driven by a Class C amplifier for greater power output and efficiency than is possible with the direct magnetostrictive-controlled oscillator. The amplifier is in turn driven by a low-powered oscillator.

Piezoelectric crystals provide transducers for the development of ultrasonic waves as in Fig. 11-16. As described in Sec. 11-9, these crystals develop mechanical motion upon proper application of electrical potentials. While quartz is common, other materials are crystals of Rochelle salts, and ammonium tartrate, these giving much higher output but being susceptible to damage by humidity. Barium titanate is more satisfactory, it being possible to mold it as a ceramic, the piezoelectric behavior being produced by polarization of the material under a high d-c electric-field intensity during manufacture. After removal of this field a residual polarization remains.

The barium titanate can be molded in the shape of a focusing radiator to concentrate the acoustic energy, the polarization being applied so that the mechanical motion is in the desired direction. Very sharp beams of energy are possible since the sharpness of the beam depends on the diameter of the radiator in wavelengths, and the radiator may be several inches in diameter with wavelengths of a few hundredths of an inch.

Operating temperatures must be limited to 90°C, and barium titanate transducers usually require oil or water cooling.

The *electromechanical coupling factor* is the ratio of the mechanical energy available to the total electrical-energy input. For quartz this represents 0.011, or 1.1 per cent, whereas for barium titanate the factor may be as high as 0.30. In all ultrasonic applications the matching of the acoustical output impedance to that of the load is important for good energy transfer. Such matching can be done by adjustment of the relative sizes of driving electrode and the load.

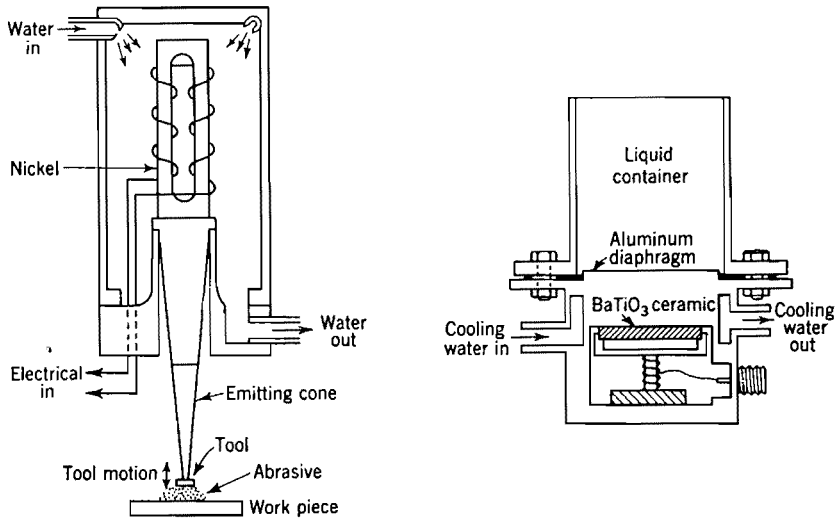



FIG. 11-17. (*a*) Magnetostrictive machine tool. (*b*) Ceramic driver for liquids.

The value of ultrasonic methods lies in the high particle velocities in the sound beam. The acceleration developed by molecules subjected to the action of moderate ultrasonic power (a few watts per square centimeter) at 1 Mc may approximate 1 million times that of gravity, a very violent agitation. Small particles tend to break apart, and cavitation is produced in liquids.

Figure 11-17*a* shows an application of a magnetostrictive tool to the drilling of very hard materials. Irregular holes having the exact cross section of the tool may be driven through Carballoy, glass, or other hard materials at rapid rates with tools of cold-rolled steel or brass. The up-and-down tool motion drives abrasive particles into the work, chipping it away. Figure 11-17*b* shows a mounting for producing agitation in a liquid over a barium titanate element.

TABLE 22. OPERATING FREQUENCIES FOR ULTRASONIC TRANSDUCERS

| *Type* | *Normal frequency range* |
|---|---|
| Ammonium tartrate | 0–250 kc |
| Rochelle salts | 0–300 kc |
| Barium titanate | 300 kc–1000 kc |
| Quartz | 10 kc–10 Mc |
| Magnetostriction | 10 kc–100 kc |

**11-12. Electromagnetic Fields in Metals.** Class C power amplifiers have an important application in the generation of large amounts of power for the processing of metals by induced currents or the heating of nonconductors by dielectric losses.

Metals may be more rapidly heated in this way than by heating in a furnace, and the heat may be produced in and confined to a localized area of the metal part. High-frequency current from an oscillator or amplifier flows through a coil surrounding the metal workpiece, and current is induced in the metal as a transformer secondary. Control of the heated area is possible because the primary coil can be located as desired and because the high-frequency currents tend to flow only in a shallow surface layer of the piece. This tendency for the current to crowd to the surface is known as *skin effect.* An example in which this is employed to advantage is the case-hardening of steel where it is desired to heat and harden only a thin outer skin of metal.

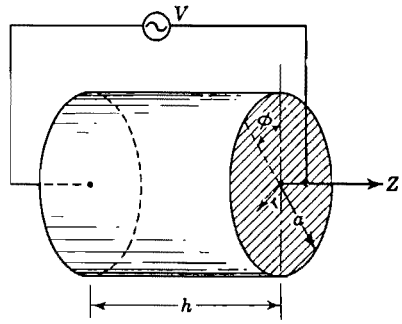



FIG. 11-18. Conducting cylinder, with perfectly conducting end plates. Note reference axes.

Skin effect may be qualitatively explained by remembering that, because of internal flux, a filament of current at the center of a conducting bar is linked by more flux and has a higher reactance than a filament of current at the surface of the bar. Current seeks the lowest impedance path and therefore flows at or near the surface. The phenomenon is evident at 60 cycles and becomes very pronounced at frequencies of a few thousand cycles in good conductors.

The effect may be quantitatively analyzed by consideration of the cylinder of Fig. 11-18, with length $h$ and radius $a$. The a-c generator $V$ is connected to perfectly conducting thin metal electrodes contacting each end of the cylinder. Neglecting any fringing field, the current and the fields present will then have axial symmetry and will vary only as

functions of time and radius $r$. The relations existing between the fields present are given by Maxwell's equations in cylindrical coordinates of $r$, $\phi$, and $z$, as

$$\begin{aligned}
\frac{1}{r}\frac{\partial \mathcal{E}_z}{\partial \phi} - \frac{\partial \mathcal{E}_\phi}{\partial z} &= -\mu \frac{\partial H_r}{\partial t} \\
\frac{\partial \mathcal{E}_r}{\partial z} - \frac{\partial \mathcal{E}_z}{\partial r} &= -\mu \frac{\partial H_\phi}{\partial t} \\
\frac{1}{r}\frac{\partial}{\partial r}(r\mathcal{E}_\phi) - \frac{1}{r}\frac{\partial \mathcal{E}_r}{\partial \phi} &= -\mu \frac{\partial H_z}{\partial t}
\end{aligned} \tag{11-37}$$

$$\begin{aligned}
\frac{1}{r}\frac{\partial H_z}{\partial \phi} - \frac{\partial H_\phi}{\partial z} &= \sigma \mathcal{E}_r + \epsilon \frac{\partial \mathcal{E}_r}{\partial t} \\
\frac{\partial H_r}{\partial z} - \frac{\partial H_z}{\partial r} &= \sigma \mathcal{E}_\phi + \epsilon \frac{\partial \mathcal{E}_\phi}{\partial t} \\
\frac{1}{r}\frac{\partial}{\partial r}(rH_\phi) - \frac{1}{r}\frac{\partial H_r}{\partial \phi} &= \sigma \mathcal{E}_z + \epsilon \frac{\partial \mathcal{E}_z}{\partial t}
\end{aligned} \tag{11-38}$$

For the symmetrical case under consideration all field components are zero except $H_\phi$ and $\mathcal{E}_z$, and the above equations reduce to

$$\frac{\partial \mathcal{E}_z}{\partial r} = \mu \frac{\partial H_\phi}{\partial t} \tag{11-39}$$

$$\frac{1}{r}\frac{\partial}{\partial r}(rH_\phi) = \sigma \mathcal{E}_z + \epsilon \frac{\partial \mathcal{E}_z}{\partial t} \tag{11-40}$$

where $\sigma$ is the conductivity of the material in mho-meters per square meter, $\epsilon$ and $\mu$ are the permittivity and permeability of the material, and

$$\epsilon = \epsilon_r \epsilon_v \qquad \mu = \mu_r \mu_v$$

The ordinary dielectric constant is $\epsilon_r$, $\mu_r$ is the relative permeability, and the space values $\epsilon_v$ and $\mu_v$ are

$$\epsilon_v = \frac{10^{-9}}{36\pi} \text{ farad/m} \qquad \mu_r = 4\pi \times 10^{-7} \text{ h/m}$$

If the applied potential $V$ be assumed to have a time variation

$$V = V_m e^{j\omega t}$$

then the fields will vary likewise and Eqs. (11-39) and (11-40) become

$$\frac{\partial \mathcal{E}}{\partial r} = j\omega \mu H \tag{11-41}$$

$$\frac{1}{r}\frac{\partial}{\partial r}(rH) = \frac{H}{r} + \frac{\partial H}{\partial r} = (\sigma + j\omega\epsilon)\mathcal{E} \tag{11-42}$$

effective values of the fields being employed and where $\mathcal{E} = \mathcal{E}_z$ and $H = H_\phi$ are implied. These equations describe the field relations in the bar of material of Fig. 11-18.

If Eq. (11-41) is solved for $H$ and the result substituted in Eq. (11-42), there results

$$\frac{\partial^2 \mathcal{E}}{\partial r^2} + \frac{1}{r}\frac{\partial \mathcal{E}}{\partial r} - j\omega\mu(\sigma + j\omega\epsilon)\mathcal{E} = 0 \tag{11-43}$$

This may be simplified by assigning the value

$$\gamma^2 = -j\omega\mu(\sigma + j\omega\epsilon) = \omega^2\mu\epsilon - j\omega\mu\sigma \tag{11-44}$$

so that

$$\frac{\partial^2 \mathcal{E}}{\partial r^2} + \frac{1}{r}\frac{\partial \mathcal{E}}{\partial r} + \gamma^2\mathcal{E} = 0 \tag{11-45}$$

This is a differential equation of Bessel form and has as its general solution

$$\mathcal{E} = AJ_o(\gamma r) + BK_o(\gamma r)$$

with $A$ and $B$ as arbitrary constants, $J_o(\gamma r)$ a Bessel function of the first kind and zero order, $K_o(\gamma r)$ a Bessel function of the second kind and zero order. Since $K_o(\gamma r)$ becomes infinite as $r$ approaches zero, the constant $B$ must be made zero to keep the field at the center of the bar finite as required by the physical situation. The solution then reduces to

$$\mathcal{E} = AJ_o(\gamma r)$$

The Bessel function $J_o(\gamma r)$ takes the form of an infinite series

$$J_o(\gamma r) = 1 - \frac{(\gamma r)^2}{2^2} + \frac{(\gamma r)^4}{2^2 \times 4^2} - \frac{(\gamma r)^6}{2^2 \times 4^2 \times 6^2} + \frac{(\gamma r)^8}{2^2 \times 4^2 \times 6^2 \times 8^2} - \cdots \tag{11-46}$$

and values are tabulated in Ref. 11.

To evaluate $A$, it may be noted that, at $r = 0$, $J_o(0) = 1$ and the electric-field intensity at $r = 0$ may be given the value $\mathcal{E}_o$ so that

$$\mathcal{E} = \mathcal{E}_o J_o(\gamma r) \tag{11-47}$$

presents the value of $\mathcal{E}$ at any radius $r$ in terms of the field on the axis, $\mathcal{E}_o$.

The magnetic intensity may be found from Eq. (11-41) as

$$H = \frac{1}{j\omega\mu}\frac{\partial \mathcal{E}}{\partial r}$$

and, taking the derivative of Eq. (11-46) with respect to $r$,

$$\frac{d}{dr}J_o(\gamma r) = -\gamma\left[\frac{\gamma r}{2} - \frac{(\gamma r)^3}{2^2 \times 4} + \frac{(\gamma r)^5}{2^2 \times 4^2 \times 6} - \frac{(\gamma r)^7}{2^2 \times 4^2 \times 6^2 \times 8} + \cdots\right] \tag{11-48}$$

The resulting series inside the brackets is known as the Bessel function of the first kind and first order, $J_1(\gamma r)$, so that

$$\frac{d}{dr} J_o(\gamma r) = -\gamma J_1(\gamma r) \tag{11-49}$$

With this result the magnetic-field intensity becomes

$$H = \frac{j\gamma \mathcal{E}_o J_1(\gamma r)}{\omega\mu} \tag{11-50}$$

at any radius in the cylinder. It is then possible to compute the fields and resultant currents at any point in the cylinder.

**11-13. Skin Effect in Metals.** If the cylinder of Fig. 11-18 be made of a good conductor, certain simplifying assumptions can be made. In particular, it can be shown that $\sigma \gg \omega\epsilon$ at even the highest attainable frequency, and thus Eq. (11-44) may be written for metals as

$$\gamma^2 = -j\omega\mu\sigma \tag{11-51}$$

Since the high-frequency currents tend to flow near the surface, a quantity $\delta$ may be defined as the *depth of penetration* and used as a yardstick for appraising the depth of current flow in high-frequency heating. Then

$$\delta = \frac{1}{\sqrt{\pi\mu\sigma f}} = \sqrt{\frac{2}{\omega\mu\sigma}} \quad \text{meters} \tag{11-52}$$

where $\delta$ is in meters. The parameter $\gamma$ becomes

$$\gamma = \sqrt{-j}\,\frac{\sqrt{2}}{\delta} \tag{11-53}$$

and from Eq. (11-47)

$$\mathcal{E} = \mathcal{E}_o J_o\left(\sqrt{-j}\,\frac{\sqrt{2}\,r}{\delta}\right) \tag{11-54}$$

in the conductive cylinder.

The above equation requires an interpretation of a Bessel function with a complex argument. Using Eq. (11-48) with the argument $\sqrt{-j}\,x$ gives

$$J_o(\sqrt{-j}\,x) = 1 + j\frac{x^2}{2^2} - \frac{x^4}{2^2 \times 4^2} - j\frac{x^6}{2^2 \times 4^2 \times 6^2} + \frac{x^8}{2^2 \times 4^2 \times 6^2 \times 8^2} \cdots$$

Two separate series are involved, the real part being named ber $(x)$, the imaginary part being bei $(x)$, so that

$$\text{ber}\,(x) = 1 - \frac{x^4}{2^2 \times 4^2} + \frac{x^8}{2^2 \times 4^2 \times 6^2 \times 8^2} - \frac{x^{12}}{2^2 \times 4^2 \times 6^2 \times 8^2 \times 10^2 \times 12^2} + \cdots \tag{11-55}$$

$$\text{bei}\ (x) = \frac{x^2}{2^2} - \frac{x^6}{2^2 \times 4^2 \times 6^2} + \frac{x^{10}}{2^2 \times 4^2 \times 6^2 \times 8^2 \times 10^2} - \cdots \quad (11\text{-}56)$$

so that $$J_o(\sqrt{-j}\,x) = \text{ber}\ (x) + j\ \text{bei}\ (x)$$

The derivatives of these functions may be taken in the series form and yield series designated as ber′ $(x)$ and bei′ $(x)$,

$$\text{ber}'\ (x) = -\frac{x^3}{2^2 \times 4} + \frac{x^7}{2^2 \times 4^2 \times 6^2 \times 8} - \frac{x^{11}}{2^2 \times 4^2 \times 6^2 \times 8^2 \times 10^2 \times 12} + \cdots \quad (11\text{-}57)$$

$$\text{bei}'\ (x) = \frac{x}{2} - \frac{x^5}{2^2 \times 4^2 \times 6} + \frac{x^9}{2^2 \times 4^2 \times 6^2 \times 8^2 \times 10} - \cdots \quad (11\text{-}58)$$

Values of these various functions are tabulated in Ref. 11, and selected values are given in Table 23.

In view of the above, Eq. (11-54) may be rewritten as

$$\mathcal{E} = \mathcal{E}_o\left[\text{ber}\left(\frac{\sqrt{2}\,r}{\delta}\right) + j\ \text{bei}\left(\frac{\sqrt{2}\,r}{\delta}\right)\right] \quad (11\text{-}59)$$

The current density at any radius is given by $\sigma\mathcal{E}$, so that

$$\text{amp/sq m} = \frac{I}{m^2} = \sigma\mathcal{E}_o\sqrt{\left[\text{ber}\left(\frac{\sqrt{2}\,r}{\delta}\right)\right]^2 + \left[\text{bei}\left(\frac{\sqrt{2}\,r}{\delta}\right)\right]^2} \quad (11\text{-}60)$$

in absolute value. This is plotted for a copper conductor for several frequencies in Fig. 11-19, showing that at the higher frequencies the current crowds to the surface.

In view of the definition

$$H = NI/2\pi r$$

it is possible to compute the total current flowing in the cylinder as

$$I_T = 2\pi a H_a = \frac{j2\pi a\gamma J_1(\gamma a)}{\omega\mu}$$

where $H_a$ is $H$ evaluated at $r = a$.

The term $\gamma J_1(\gamma a)$ may be evaluated by Eq. (11-49) as

$$\gamma J_1(\gamma a) = -\frac{d}{da}J_o(\gamma a) = -\frac{d}{da}\left[\text{ber}\left(\frac{\sqrt{2}\,a}{\delta}\right) + j\ \text{bei}\left(\frac{\sqrt{2}\,a}{\delta}\right)\right]$$
$$= -\frac{\sqrt{2}}{\delta}\left[\text{ber}'\left(\frac{\sqrt{2}\,a}{\delta}\right) + j\ \text{bei}'\left(\frac{\sqrt{2}\,a}{\delta}\right)\right]$$

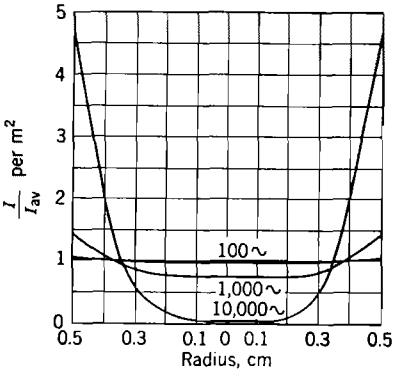



FIG. 11-19. Ratio of current density to average conductor current, showing skin effect.

TABLE 23. VALUES OF BER $(x)$, BEI $(x)$, BER′ $(x)$, BEI′ $(x)$

| $x$ | ber $(x)$ | bei $(x)$ | ber′ $(x)$ | bei′ $(x)$ |
|---|---|---|---|---|
| 0.0 | 1.000 | 0.00 | 0.00 | 0.00 |
| 0.5 | 0.999 | 0.0625 | −0.0078 | 0.250 |
| 1.0 | 0.984 | 0.250 | −0.0624 | 0.499 |
| 1.5 | 0.921 | 0.558 | −0.210 | 0.730 |
| 2.0 | 0.752 | 0.972 | −0.493 | 0.917 |
| 2.5 | 0.399 | 1.457 | −0.944 | 0.998 |
| 3.0 | −0.221 | 1.938 | −1.570 | 0.881 |
| 3.5 | −1.194 | 2.283 | −2.336 | 0.435 |
| 4.0 | −2.563 | 2.293 | −3.135 | −0.491 |
| 4.5 | −4.299 | 1.686 | −3.754 | −2.053 |
| 5.0 | −6.230 | 0.116 | −3.844 | −4.354 |
| 10.0 | 138.84 | 56.37 | 51.37 | 135.23 |
| 20.0 | 47,583.7 | 11,500.8 | 24,325.1 | 41,491.5 |
| ∞ | ∞ | ∞ | ∞ | ∞ |

and so the total current in the conducting cylinder is

$$I_T = \frac{-j2\sqrt{2}\,\pi a}{\omega\mu\delta}\left[\text{ber}'\left(\frac{\sqrt{2}\,a}{\delta}\right) + j\,\text{bei}'\left(\frac{\sqrt{2}\,a}{\delta}\right)\right] \tag{11-61}$$

The impedance per unit length of the cylinder is then

$$Z = \frac{E}{I} = \frac{\mathcal{E}_o\left[\text{ber}\left(\frac{\sqrt{2}\,a}{\delta}\right) + j\,\text{bei}\left(\frac{\sqrt{2}\,a}{\delta}\right)\right]}{\frac{-j2\sqrt{2}\,\pi a}{\omega\mu\delta}\left[\text{ber}'\left(\frac{\sqrt{2}\,a}{\delta}\right) + j\,\text{bei}'\left(\frac{\sqrt{2}\,a}{\delta}\right)\right]}$$

$$= \frac{j\omega\mu\delta}{2\sqrt{2}\,\pi a}\left[\frac{\text{ber}\left(\frac{\sqrt{2}\,a}{\delta}\right) + j\,\text{bei}\left(\frac{\sqrt{2}\,a}{\delta}\right)}{\text{ber}'\left(\frac{\sqrt{2}\,a}{\delta}\right) + j\,\text{bei}'\left(\frac{\sqrt{2}\,a}{\delta}\right)}\right] \tag{11-62}$$

This impedance per unit length has the form $Z = R_{ac} + jX_{ac}$ and may be normalized on $R_{dc}$ as

$$\frac{Z}{R_{dc}} = \frac{R_{ac}}{R_{dc}} + j\,\frac{X_{ac}}{R_{dc}}$$

where

$$R_{dc} = \frac{1}{\sigma\pi a^2}$$

The resulting coefficient

$$\left(\frac{\omega\mu\delta}{2\sqrt{2}\,\pi a}\right)\frac{1}{R_{dc}} = \frac{1}{2}\left(\frac{\sqrt{2}\,a}{\delta}\right) \tag{11-63}$$

so that

$$\frac{R_{ac}}{R_{dc}} = \frac{1}{2}\left(\frac{\sqrt{2}\,a}{\delta}\right)\left\{\frac{\operatorname{ber}\left(\frac{\sqrt{2}\,a}{\delta}\right)\operatorname{bei}'\left(\frac{\sqrt{2}\,a}{\delta}\right) - \operatorname{bei}\left(\frac{\sqrt{2}\,a}{\delta}\right)\operatorname{ber}'\left(\frac{\sqrt{2}\,a}{\delta}\right)}{\left[\operatorname{ber}'\left(\frac{\sqrt{2}\,a}{\delta}\right)\right]^2 + \left[\operatorname{bei}'\left(\frac{\sqrt{2}\,a}{\delta}\right)\right]^2}\right\} \tag{11-64}$$

$$\frac{X_{ac}}{R_{dc}} = \frac{1}{2}\left(\frac{\sqrt{2}\,a}{\delta}\right)\left\{\frac{\operatorname{ber}\left(\frac{\sqrt{2}\,a}{\delta}\right)\operatorname{ber}'\left(\frac{\sqrt{2}\,a}{\delta}\right) + \operatorname{bei}\left(\frac{\sqrt{2}\,a}{\delta}\right)\operatorname{bei}'\left(\frac{\sqrt{2}\,a}{\delta}\right)}{\left[\operatorname{ber}'\left(\frac{\sqrt{2}\,a}{\delta}\right)\right]^2 + \left[\operatorname{bei}'\left(\frac{\sqrt{2}\,a}{\delta}\right)\right]^2}\right\} \tag{11-65}$$

Equations (11-64) and (11-65) are plotted in Fig. 11-20 as functions of $a/\delta$, or the radius measured in terms of skin thickness.

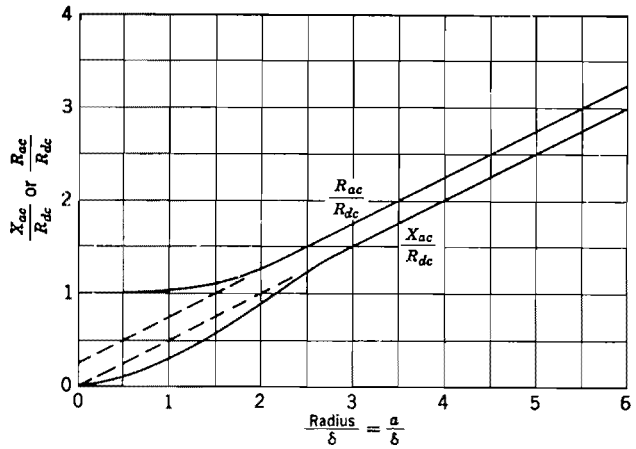



FIG. 11-20. Resistance and reactance at high frequency.

Figure 11-20 shows that above $a/\delta = 2$ the relations are approximately linear so that good approximations are given by the equations

For $\frac{a}{\delta} > 2$

$$\frac{R_{ac}}{R_{dc}} = \frac{a}{2\delta} + 0.25 \tag{11-66}$$

$$\frac{X_{ac}}{R_{dc}} = \frac{a}{2\delta} \tag{11-67}$$

Referring to Eq. (11-52), it is seen that increases in frequency $f$, $\sigma$, or $\mu$ all reduce the depth of penetration of current. The decrease in $\delta$, or reduction of effective conductor area, causes the increase in resistivity

at increasing frequency. This is illustrated by considering a direct current $I$ flowing uniformly over the solid conductor of Fig. 11-21*a*. The resistance per meter length would be

$$R_1 = \frac{E_1}{I} = \frac{1}{\sigma\pi a^2}$$

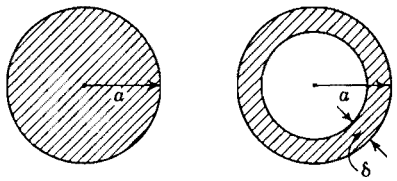



FIG. 11-21. Use of depth of penetration as wall thickness of equivalent tubular conductor.

Then consider the tubular conductor of Fig. 11-21*b*, with wall thickness $\delta$, with the same current uniformly flowing through its annular area. Then

$$R_2 = \frac{E_2}{I} = \frac{1}{\sigma[\pi a^2 - \pi(a - \delta)^2]}$$

The ratio $R_2/R_1$ is then

$$\frac{R_2}{R_1} = \frac{\sigma\pi a^2}{\sigma[\pi a^2 - \pi(a - \delta)^2]} = \frac{a}{2\delta}\left(\frac{1}{1 - \delta/2a}\right)$$

For $\delta/2a \ll 1$

$$\frac{1}{1 - \delta/2a} \cong 1 + \frac{\delta}{2a}$$

so that

$$\frac{R_2}{R_1} = \frac{a}{2\delta}\left(1 + \frac{\delta}{2a}\right) = \frac{a}{2\delta} + 0.25 \tag{11-68}$$

which is exactly Eq. (11-66), proving that the ratio $R_{ac}/R_{dc}$ is just that obtained if all the current is uniformly crowded into a skin $\delta$ m thick.

**11-14. Induction Heating of Cylindrical Bars.** Induction heating of metals frequently involves a configuration such as in Fig. 11-22, with a long cylindrical conductor surrounded by a current-carrying helix. Current induced in the bar produces $I^2R$ losses which raise the temperature of the bar.

To simplify, it will be assumed that the bar and coil are long, so that end effects may be neglected. It will also be assumed that $H$ is everywhere axially directed and varies only as a function of radius $r$ and time and that $\mathcal{E}$ is everywhere circumferential and is likewise a function only of radius and time.

By returning to Eqs. (11-37) and (11-38), and with $H_z$ and $\mathcal{E}_\phi$ the only field components present under the chosen symmetry, it is possible to arrive at

$$\frac{\partial \mathcal{E}_\phi}{\partial r} + \frac{\mathcal{E}_\phi}{r} = -j\omega\mu H_z \tag{11-69}$$

$$-\frac{\partial H_z}{\partial r} = (\sigma + j\omega\epsilon)\mathcal{E}_\phi \tag{11-70}$$

as describing the field situation. As in Sec. 11-12 these lead to

$$\frac{\partial^2 H}{\partial r^2} + \frac{1}{r}\frac{\partial H}{\partial r} - j\omega\mu(\sigma + j\omega\epsilon)H = 0 \tag{11-71}$$

where $H = H_z$ is implied. It may again be assumed that $\sigma \gg \omega\epsilon$ for a metal cylinder.

Equation (11-71) is identical in form to Eq. (11-43), $H$ being interchanged for $\mathcal{E}$, or the fields being oriented in the physical situation such that $H$ is axial and $\mathcal{E}$ is circumferential as induced by the coil. Thus the solution may be written from Eq. (11-59) as

$$H = C\left[\text{ber}\left(\frac{\sqrt{2}\,r}{\delta}\right) + j\,\text{bei}\left(\frac{\sqrt{2}\,r}{\delta}\right)\right] \tag{11-72}$$

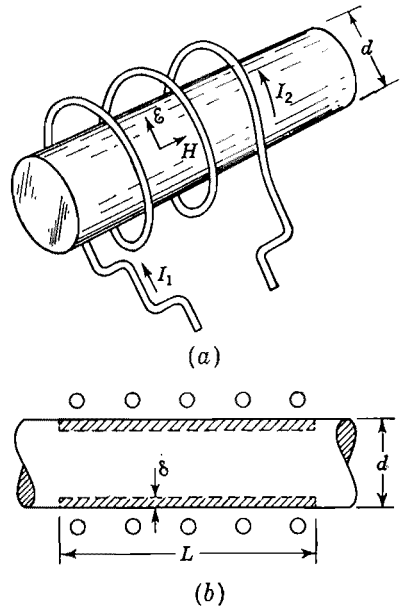



FIG. 11-22. (*a*) Work coil around cylindrical bar. (*b*) Illustration of depth of penetration.

The constant $C$ can be evaluated by considering that the field just outside the surface of the cylinder must be equal to the sum of the impressed intensity $H_o = NI/2\pi a$ of the surrounding coil and of the added intensity set up by currents flowing in the cylinder. For a perfect conductor these intensities would be equal and additive, and for ordinary conductors and wavelengths large with respect to coil-to-cylinder spacing this condition of equality is still very nearly true. Thus the total magnetic intensity at a radius $a$ at the cylinder surface may be taken as twice the impressed intensity, or $2H_o$. Thus $C$ can be evaluated at $r = a$, giving

$$H = 2H_o\left[\frac{\text{ber}\,(\sqrt{2}\,r/\delta) + j\,\text{bei}\,(\sqrt{2}\,r/\delta)}{\text{ber}\,(\sqrt{2}\,a/\delta) + j\,\text{bei}\,(\sqrt{2}\,a/\delta)}\right] \tag{11-73}$$

From Eq. (11-70)

$$\mathcal{E} = -\frac{1}{\sigma}\frac{\partial H}{\partial r}$$

$$\mathcal{E} = -\frac{2\sqrt{2}\,H_o}{\sigma\delta}\left[\frac{\text{ber}'\,(\sqrt{2}\,r/\delta) + j\,\text{bei}'\,(\sqrt{2}\,r/\delta)}{\text{ber}\,(\sqrt{2}\,a/\delta) + j\,\text{bei}\,(\sqrt{2}\,a/\delta)}\right]$$

with a magnitude given by

$$|\mathcal{E}| = \frac{2\sqrt{2}\,H_o}{\sigma\delta}\sqrt{\frac{[\text{ber}'\,(\sqrt{2}\,r/\delta)]^2 + [\text{bei}'\,(\sqrt{2}\,r/\delta)]^2}{[\text{ber}\,(\sqrt{2}\,a/\delta)]^2 + [\text{bei}\,(\sqrt{2}\,a/\delta)]^2}} \tag{11-74}$$

The power developed in the metallic cylinder per meter of length will be

$$P = \int_{r=0}^{r=a} 2\pi r\sigma\mathcal{E}^2\,dr \qquad \text{watts/m of length}$$

which is given by McLachlan[12] as

$$P = \frac{8\pi H_o^2}{\sigma}\left(\frac{\sqrt{2}\,a}{\delta}\right) \left\{\text{ber}\left(\frac{\sqrt{2}\,a}{\delta}\right)\text{ber}'\left(\frac{\sqrt{2}\,a}{\delta}\right) + \text{bei}\left(\frac{\sqrt{2}\,a}{\delta}\right)\text{bei}'\left(\frac{\sqrt{2}\,a}{\delta}\right) \Big/ \left[\text{ber}\left(\frac{\sqrt{2}\,a}{\delta}\right)\right]^2 + \left[\text{bei}\left(\frac{\sqrt{2}\,a}{\delta}\right)\right]^2\right\} \tag{11-75}$$

If a quantity $F$ is defined as

$$F = \sqrt{2}\left[\frac{\text{ber}\,(\sqrt{2}\,a/\delta)\ \text{ber}'\,(\sqrt{2}\,a/\delta) + \text{bei}\,(\sqrt{2}\,a/\delta)\ \text{bei}'\,(\sqrt{2}\,a/\delta)}{[\text{ber}\,(\sqrt{2}\,a/\delta)]^2 + [\text{bei}\,(\sqrt{2}\,a/\delta)]^2}\right]$$

then

$$P = \frac{8\pi H_o^2}{\sigma}\frac{a}{\delta}F \qquad \text{watts/m of length} \tag{11-76}$$

Values of $F$ may be found from Fig. 11-23 as functions of the ratio $a/\delta$.

For a cylinder of circumference $2\pi a$ and length 1 m the surface power density or power per unit area is

$$P_d = \frac{P}{2\pi a} = \frac{8\pi H_o^2}{2\pi a}\frac{a}{\sigma\delta} \qquad \text{watts/sq m}$$

and *for the limiting case of a flat sheet,* where $a$ may be taken as very large, this becomes

$$P_d = \frac{4H_o^2}{\sigma\delta} \qquad \text{watts/sq m} \tag{11-77}$$

For a given primary current in the exciting coil $H_o$ will be constant, and the power $P$ induced in the metal cylinder will rise with $aF/\delta$, which is a function of frequency. However, Brown[13] has called attention to the fact that $F$ does not increase rapidly with changes in $a/\delta$ when $a/\delta > 2.5$, and Fig. 11-23 shows this. The break in the curve indicates a critical frequency above which it is not economic to go. Actually it is found that, for frequencies greater than the critical value $f_c$, the power loss in the exciting coil rises as fast as the power induced in the work and no increase in efficiency is obtained. For frequencies less than $f_c$ the possible power

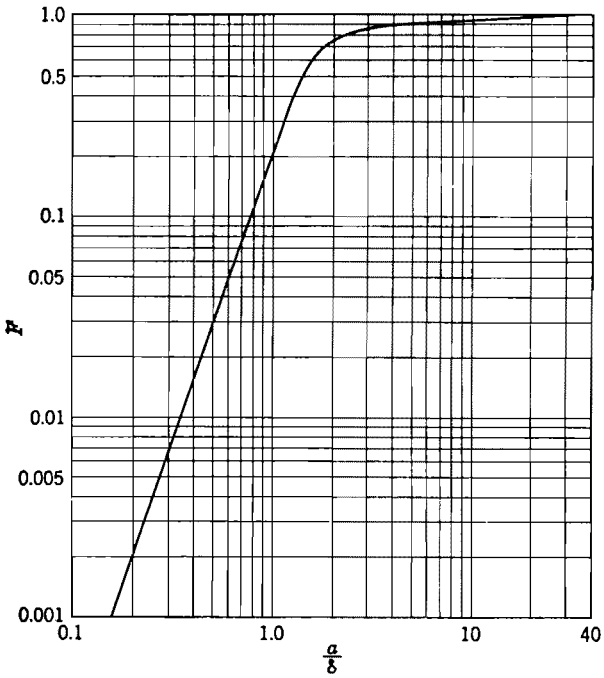



FIG. 11-23. $F$ as a function of $a/\delta$.

input drops rapidly. Thus the desirable frequency for a given workpiece should be above

$$\frac{a}{\delta} = \frac{a}{\sqrt{2/\omega\mu\sigma}} = 2.5$$

$$f_c = \frac{12.5}{2\pi\mu\sigma a^2} \cong \frac{2}{\mu\sigma a^2} \tag{11-78}$$

While all the above, except Eq. (11-77), has been derived on the basis of a cylindrical bar, many such problems are encountered and many others may be approximated by this theory.

**11-15. Heating of Dielectric Cylinders or Slabs.** The method of Sec. 11-12 will now be applied to a problem in the heating of dielectrics. Assume that the dielectric of Fig. 11-24 is perfect, or that $\sigma = 0$. For dielectrics $\mu = \mu_v$, and so

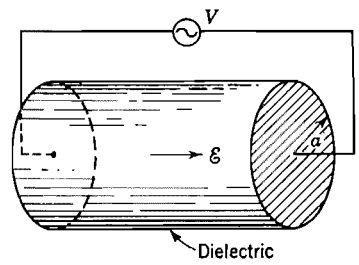



FIG. 11-24. Cylinder of dielectric between plane conducting electrodes.

$$\gamma = \sqrt{\omega^2\mu_v\epsilon_v\epsilon_r}$$

The velocity of electromagnetic propagation in free space is

$$\frac{1}{\sqrt{\mu_v \epsilon_v}} = 3 \times 10^8 \text{ m per sec} = c$$

and the free-space wavelength $\lambda_o = f/c$ m. Then

$$\gamma = \frac{\omega \sqrt{\epsilon_r}}{c} = \frac{2\pi \sqrt{\epsilon_r}}{\lambda_o} \tag{11-79}$$

where $\epsilon_r$ is the usual dielectric constant.

Equation (11-47) then expresses the electric-field intensity in the perfect dielectric as

$$\frac{\mathcal{E}}{\mathcal{E}_o} = J_o\left(\frac{2\pi r \sqrt{\epsilon_r}}{\lambda_o}\right) \tag{11-80}$$

This variation of field intensity in terms of the magnitude of $\mathcal{E}/\mathcal{E}_o$ is plotted in Fig. 11-25 as a solid line, showing that the intensity is not constant with depth in the cylinder at the higher frequencies.

A cylinder of perfect dielectric would not be heated by application of the electric field. In practice it is usually desired that dielectric materials be uniformly raised in temperature by application of high-frequency power to the end electrodes. Assuming a rather lossy material, or one in which $\sigma < \omega\epsilon$, but $\sigma \neq 0$, then from Eq. (11-44)

$$\gamma = \sqrt{\omega^2 \mu_v \epsilon_v \epsilon_r (1 - j\alpha)}$$

where $$\alpha = \frac{\sigma}{\omega \epsilon_v \epsilon_r} = \frac{\sigma}{\omega \epsilon}, \text{ a small quantity}$$

Then $$\gamma = \frac{2\pi \sqrt{\epsilon_r}}{\lambda_o} \sqrt{1 - j\alpha} \tag{11-81}$$

This value of $\gamma$ may be substituted in the series for $J_o(\gamma r)$ of Eq. (11-46). If the result is inserted in Eq. (11-47) for $\mathcal{E}$ and all terms involving the square or higher powers of $\alpha$ are dropped as negligible, there results

$$\frac{\mathcal{E}}{\mathcal{E}_o} = J_o\left(\frac{2\pi r \sqrt{\epsilon_r}}{\lambda_o}\right) + j\frac{\alpha}{2}\left(\frac{2\pi r \sqrt{\epsilon_r}}{\lambda_o}\right) J_1\left(\frac{2\pi r \sqrt{\epsilon_r}}{\lambda_o}\right) \tag{11-82}$$

The absolute value of this expression, with $\alpha = 0.1$, is also plotted in Fig. 11-25 for comparison with the perfect-dielectric case. The similarity is marked, except where the ideal case predicts zero field intensities. Since most dielectrics will have values of $\alpha$ between 0 and 0.1, most practical cases will be bounded by the curves of the figure.

Since the field intensity differs little between the lossy case and the perfect dielectric, and since Eq. (11-80) is considerably simpler than the more exact form of Eq. (11-82), the former equation will be used to deter-

mine the power dissipated in the dielectric, with the understanding that some error will arise near the points of field zero. Since the power in watts per cubic meter is $\sigma\mathcal{E}^2$, and since $\mathcal{E}$ is not uniform with $r$, it follows that the heat may not be uniformly dissipated throughout the dielectric.

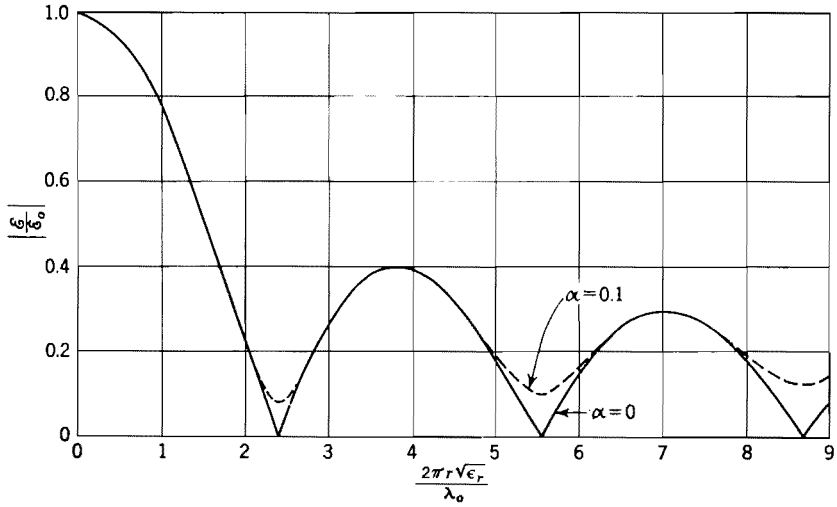



FIG. 11-25. $|\mathcal{E}/\mathcal{E}_0|$ in a dielectric cylinder as a function of $(2\pi r\sqrt{\epsilon_r})/\lambda_0$.

The power dissipated at any radius $r$ is

$$P_r = \sigma\mathcal{E}^2 = \sigma\mathcal{E}_0^2 J_0^2\left(\frac{2\pi r\sqrt{\epsilon_r}}{\lambda_0}\right) \quad \text{watts/cu m} \tag{11-83}$$

The total power per meter of length of the cylinder is

$$P = 2\pi\sigma\mathcal{E}_0^2 \int_{r=0}^{r=a} rJ_0^2\left(\frac{2\pi r\sqrt{\epsilon_r}}{\lambda_0}\right) dr$$

using $2\pi r\,dr \times 1$ m as the elemental volume, and after changing variable,

$$\begin{aligned} P &= \frac{2\pi\sigma\mathcal{E}_0^2}{(2\pi\sqrt{\epsilon_r}/\lambda_0)^2} \int_0^{2\pi a\sqrt{\epsilon_r}/\lambda_0} xJ_0^2(x)\,dx \\ &= \pi a^2\sigma\mathcal{E}_0^2\left[J_0^2\left(\frac{2\pi a\sqrt{\epsilon_r}}{\lambda_0}\right) + J_1^2\left(\frac{2\pi a\sqrt{\epsilon_r}}{\lambda_0}\right)\right] \quad \text{watts/m} \end{aligned} \tag{11-84}$$

The volume of the cylinder per meter of length is $\pi a^2 \times 1$ m, and the average power dissipated per cubic meter is

$$P_{av} = \sigma\mathcal{E}_0^2\left[J_0^2\left(\frac{2\pi a\sqrt{\epsilon_r}}{\lambda_0}\right) + J_1^2\left(\frac{2\pi a\sqrt{\epsilon_r}}{\lambda_0}\right)\right] \tag{11-85}$$

It is interesting to measure the uniformity of power dissipation or heating at any given radius as given by Eq. (11-83) by comparing it with the average value from Eq. (11-85). Thus

$$\frac{P_r}{P_{av}} = \frac{J_o^2(2\pi r \sqrt{\epsilon_r}/\lambda_o)}{J_o^2(2\pi a \sqrt{\epsilon_r}/\lambda_o) + J_1^2(2\pi a \sqrt{\epsilon_r}/\lambda_o)} \tag{11-86}$$

This relation is plotted in Fig. 11-26 for a dielectric slab, 50 cm in diameter and of arbitrary thickness, for frequencies of 60, 150, and 300 Mc. The two higher frequencies do not produce uniform heating.

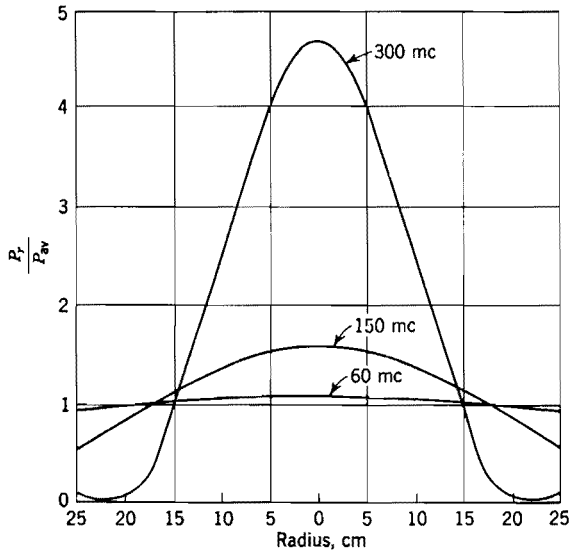



FIG. 11-26. Power distribution across a circular dielectric slab: diameter 50 cm, $\epsilon_r = 3.0$.

For many applications it is necessary to have uniform heating throughout the dielectric. A limiting frequency $f_c$ can be found, subject to a decision as to the permitted variation in power per unit volume. Assuming equal $\mathcal{E}_o$ values at all frequencies, then the permitted ratio at $r$ to that at the center is

$$\frac{P_r}{P_{center}} = \frac{J_o^2(2\pi r \sqrt{\epsilon_r}/\lambda_o)}{1} \tag{11-87}$$

since $J_o(0) = 1.0$. It follows then that if the edge is to receive 0.95 as much power as the center,

$$J_o\left(\frac{2\pi a \sqrt{\epsilon_r}}{\lambda_o}\right) = \sqrt{0.95} = 0.975$$

and from tables

$$\frac{2\pi a\sqrt{\epsilon_r}}{\lambda_o} = 0.3$$

$$f_c = \frac{0.3 \times 3 \times 10^8}{2\pi a\sqrt{\epsilon_r}} = \frac{14.31}{a\sqrt{\epsilon_r}} \quad \text{Mc}$$

as the upper frequency limit for which the edge power is not less then 95 per cent of the power at the center.

**11-16. Power Required for Rapid Heating.** A great advantage of induction or dielectric heating is the speed with which the heat may be produced. In fact, this speed is an important factor in the quality of the finished part, since it is usual to raise the temperature so rapidly that spreading of heat to adjacent portions of the workpiece by conduction, or loss by convection, is minimized. It is desirable to be able to predict theoretically the amount of power required to raise a workpiece to a given temperature in a given time under conditions of negligible conduction and convection loss. The heat input then is used only to raise the temperature, to supply the radiation loss, and to supply any latent heat required for melting or evaporation.

For a material of $M$ kg with specific heat $s$ joules per °K per kg the amount of heat needed to raise the temperature $\Delta\theta°$ is

$$\Delta h = sM\,\Delta\theta$$

and the rate of heat supply is

$$\frac{dh}{dt} = sM\frac{d\theta}{dt} \qquad \text{joules/sec} = \text{watts}$$

This is the heat needed to supply the transient stored heat during a temperature change.

Losses have been assumed due to radiation only. For a perfect radiator, or a *black body*, the power radiated is given by the Stefan-Boltzmann law as

$$P = CA(\theta^4 - \theta_o^4) \qquad \text{watts/sq m} \tag{11-88}$$

where $\theta_o$ is the ambient temperature of the body and its surroundings from which heating starts, in degrees Kelvin, and $A$ is the total radiating surface in square meters. For a black body the constant $C = 5.73 \times 10^{-8}$ watt per sq m per deg$^4$. All actual surfaces have an emissivity below that of a black body, or radiate at a lower rate. The power required to raise a black body to temperature $\theta$ will be greater than is needed by an actual body. Since conduction and convection losses have been neglected here, the power computed for the black body will somewhat compensate for the added losses in practice.

The temperature finally reached represents a balance between the power input and the power radiated. In final steady state the total power input is radiated and an ultimate temperature $\theta_u$ reached, the value of $\theta_u$ being given from Eq. (11-88) as

$$\theta_u = \sqrt[4]{\frac{P}{CA} + \theta_o^4} \qquad °K \tag{11-89}$$

The total power input to the workpiece at any time $t$ is the sum of that required to supply the stored energy and that required to maintain a steady state. Thus

$$P = sM\frac{d\theta}{dt} + CA(\theta^4 - \theta_o^4) \qquad \text{watts}$$

$$\frac{d\theta}{dt} = \frac{CA}{sM}\left(\frac{P}{CA} + \theta_o^4 - \theta^4\right)$$

Substituting Eq. (11-89) yields

$$\frac{d\theta}{dt} = \frac{CA}{sM}(\theta_u^4 - \theta^4) \tag{11-90}$$

as the differential equation of the system. Separation of variables gives

$$dt = \frac{sM}{CA}\left(\frac{d\theta}{\theta_u^4 - \theta^4}\right) \tag{11-91}$$

The right-side parentheses may be expanded in partial fractions as

$$\frac{1}{\theta_u^4 - \theta^4} = \frac{a}{\theta_u^2 - \theta^2} + \frac{b}{\theta_u^2 + \theta^2}$$

$$1 = (\theta_u^2 + \theta^2)a + (\theta_u^2 - \theta^2)b$$

This is true for all $\theta$'s, so let $\theta = \theta_u$ and

$$a = \frac{1}{2\theta_u^2}$$

Letting $\theta = -\theta_u$ gives

$$b = \frac{1}{2\theta_u^2}$$

and Eq. (11-91) may be written as

$$dt = \frac{sM}{CA}\frac{1}{2\theta_u^2}\left(\frac{d\theta}{\theta_u^2 - \theta^2} + \frac{d\theta}{\theta_u^2 + \theta^2}\right)$$

which may be integrated as

$$t = \frac{sM}{2CA\theta_u^3}\left[\tan^{-1}\frac{\theta}{\theta_u} - \tan^{-1}\frac{\theta_o}{\theta_u} + \frac{1}{2}\ln\frac{(\theta_u + \theta)(\theta_u - \theta_o)}{(\theta_u - \theta)(\theta_u + \theta_o)}\right] \qquad \text{sec} \tag{11-92}$$

This is the time in seconds for the black body to attain a temperature $\theta$ with a given input power $P$.

The use of the above equation is illustrated in Fig. 11-27, which shows the rate of heating curve for an iron bar of 2 cm diameter, 6 cm

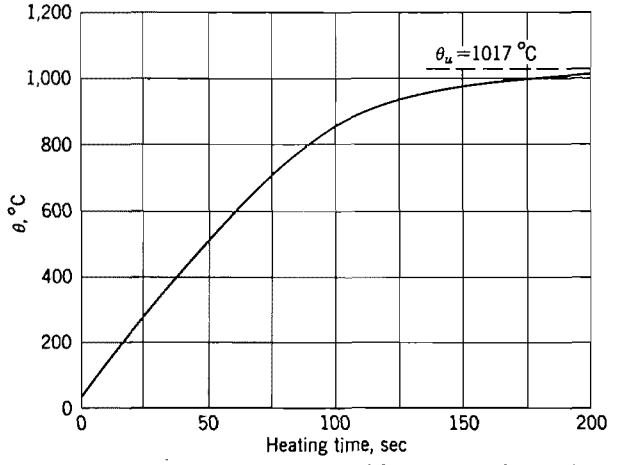



FIG. 11-27. Heating time of an iron bar with 700 watts input, 0.149 kg.

long, weighing 0.149 kg. The average specific heat of iron over the expected temperature range is taken as 461 joules per kg per °C, and $\theta_o = 27°C = 300°K$. Physical constants for some materials are given in Table 24.

TABLE 24. PHYSICAL CONSTANTS OF CONDUCTORS

| Metals | Av specific heat $s$, joules/°K/kg | Density, kg/cu m | Thermal conductivity, watts/sq m/°K | Electrical conductivity, mho-m/sq m | Melting point, °C |
|---|---|---|---|---|---|
| Aluminum | 962 | 2,560 | $2.03 \times 10^4$ | $3.54 \times 10^7$ | 660 |
| Brass (66-34) | 394 | 8,400 | $1.2 \times 10^4$ | $1.5 \times 10^7$ | 920 |
| Copper | 394 | 8,800 | $3.88 \times 10^4$ | $5.8 \times 10^7$ | 1083 |
| Iron | 461 | 7,860 | $0.67 \times 10^4$ | $1.04 \times 10^7$ | 1535 |
| Lead | 126 | 11,300 | $0.34 \times 10^4$ | $0.46 \times 10^7$ | 327 |
| Nickel | 460 | 8,900 | $0.6 \times 10^4$ | $1.15 \times 10^7$ | 1452 |
| Silver | 234 | 10,400 | $4.19 \times 10^4$ | $6.11 \times 10^7$ | 960 |
| Steel (18-8) | 500 | 7,700 | $0.16 \times 10^4$ | $0.011 \times 10^7$ | 1410 |
| Tin | 240 | 7,300 | $0.64 \times 10^4$ | $0.86 \times 10^7$ | 232 |

**11-17. Heating-power Calculations.** As an example, assume that 20-cm (6.6-in.) lengths of brass rod, of 2 cm diameter, are to be annealed

at 700°C, from an ambient temperature of 27°C. It is necessary to determine the required generator power rating at 300 kc.

For brass

$$\begin{aligned} \text{Density} &= 8.4 \times 10^3 \text{ kg/cu m} \\ \text{Average specific heat} &= 3.94 \times 10^2 \text{ joules/°K/kg} \\ \text{Average electrical conductivity} &= 1.5 \times 10^7 \text{ mho-m/sq m} \\ \text{Permeability} &= \mu_r = 1.0 \end{aligned}$$

To reach the desired ultimate temperature will require a power input given by

$$\begin{aligned} \text{Surface area} &= 2\pi al + 2\pi a^2 = 0.0126 + 0.0006 = 0.0132 \text{ sq m} \\ \text{Power radiated} &= 5.73 \times 10^{-8} \times 1.32 \times 10^{-2}(973^4 - 300^4) \\ &= 5.73 \times 10^{-8} \times 1.32 \times 10^{-2} \times 8.88 \times 10^{11} \\ &= 670 \text{ watts} \end{aligned}$$

To achieve 700°C in a practical time will require a greater power input, which might be assumed as 2000 watts. With this input $\theta_u$ can be found as

$$\begin{aligned} \theta_u &= \sqrt[4]{\frac{2 \times 10^3}{5.73 \times 10^{-8} \times 1.32 \times 10^{-2}} + 81 \times 10^8} \\ &= 1277°\text{K} = 1004°\text{C} \end{aligned}$$

assuming black-body conditions.

The volume of the bar $= \pi \times 0.01^2 \times 0.2 = 6.28 \times 10^{-5}$ cu m, and the weight is

$$M = 8.4 \times 10^3 \times 6.28 \times 10^{-5} = 0.527 \text{ kg}$$

Using Eq. (11-92),

$$\begin{aligned} \frac{sM}{2CA\theta_u^3} &= \frac{3.94 \times 10^2 \times 0.527}{2 \times 5.73 \times 10^{-8} \times 1.32 \times 10^{-2} \times 2.08 \times 10^9} \\ &= 66.0 \end{aligned}$$

$$\begin{aligned} \text{Then} \quad t &= 66\left(\tan^{-1}\frac{973}{1277} - \tan^{-1}\frac{300}{1277} + \frac{1}{2}\ln\frac{2250 \times 977}{304 \times 1577}\right) \\ &= 66(0.654 - 0.230 + 0.762) = 78.3 \text{ sec} \end{aligned}$$

as the heating time, assuming black-body conditions and neglecting conduction and convection losses.

The watts developed per meter of length of rod are

$$\frac{P}{m} = \frac{2000}{0.2} = 10{,}000 \text{ watts/m length}$$

$$\begin{aligned} \frac{a}{\delta} = a\sqrt{\frac{\omega\mu\sigma}{2}} &= 0.01\sqrt{\frac{3 \times 10^5 \times 2\pi \times 4\pi \times 10^{-7} \times 1.5 \times 10^7}{2}} \\ &= 0.01\sqrt{17.7 \times 10^6} = 42.0 \end{aligned}$$

From Fig. 11-23, $F = 1.0$ so that Eq. (11-76) gives

$$P = \frac{8\pi H_o^2}{\sigma}\frac{a}{\delta} \times 1.0 = 10{,}000 \text{ watts/m}$$

$$H_o = \sqrt{\frac{P\sigma}{8\pi a/\delta}} = \sqrt{\frac{10^4 \times 1.5 \times 10^7}{8\pi \times 42.0}}$$
$$= 11.9 \times 10^3 \text{ amp turns/m}$$

Assume the coil long and that the magnetomotive force at the center is equal to the ampere turns per meter of the coil. If 0.25-in. copper tubing is wound with turns spaced 0.25 in., the turns per meter will be

$$\frac{N}{l} = \frac{39.37}{0.5} = 78.7 \text{ turns/m}$$

The required helix current will then be

$$I = \frac{H_o}{N/l} = \frac{11.5 \times 10^3}{78.7} = 151 \text{ amp}$$

Thus the problem may be summarized as:

Generator power rating = 2 kw, plus coil losses
Time of heating to 700° = 78.3 sec
Coil wound 2 turns/in.
Coil current 151 amp

**11-18. Oscillators as High-voltage D-C Sources.** The tuned circuit of a high-frequency oscillator may be coupled to a secondary coil of many

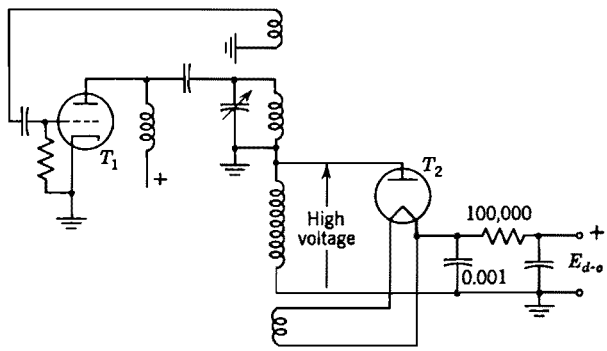



FIG. 11-28. High-frequency oscillator as high-voltage source for direct current.

turns and a high secondary a-c voltage developed. Rectification of this a-c voltage provides a low-cost method of obtaining a high voltage for low current requirements, as is frequently required for cathode-ray-beam acceleration. The chief advantage of the circuit is economic, since it eliminates the expensive and heavy transformers for stepping up the a-c voltage from the supply line. At the usual frequency of operation of the

oscillator (a few hundred kilocycles) a magnetic core if used will consist of only a small section of compressed powdered iron.

Tube $T_1$ in Fig. 11-28 is the oscillator in a conventional plate-tuned circuit. A high-voltage winding and a filament-heating winding are coupled to the plate tank, the alternating current then being rectified by high-voltage diode $T_2$. Filtering follows in a conventional $RC$ filter. Because of the high frequency employed the filtering is easily performed with small values of capacity.

Voltages as high as 80 to 100 kv are obtained in this way, with currents of a few hundred microamperes.

## PROBLEMS

**11-1.** The circuit of Fig. 11-29 is tuned to resonance at 100 kc with a capacitor of 0.02 $\mu$f. The $Q$ of the circuit is 50. For a current $i$ given by

$$i = 0.5 + \frac{4}{\pi} \sin 2\pi\, 10^5 t + \frac{1}{3} \sin 6\pi \times 10^6 t + \cdots$$

find the rms voltage across $C$ and the rms current read by ammeter $A$.

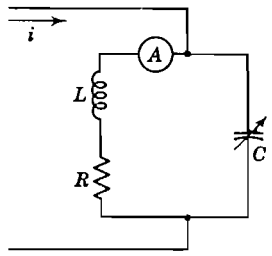


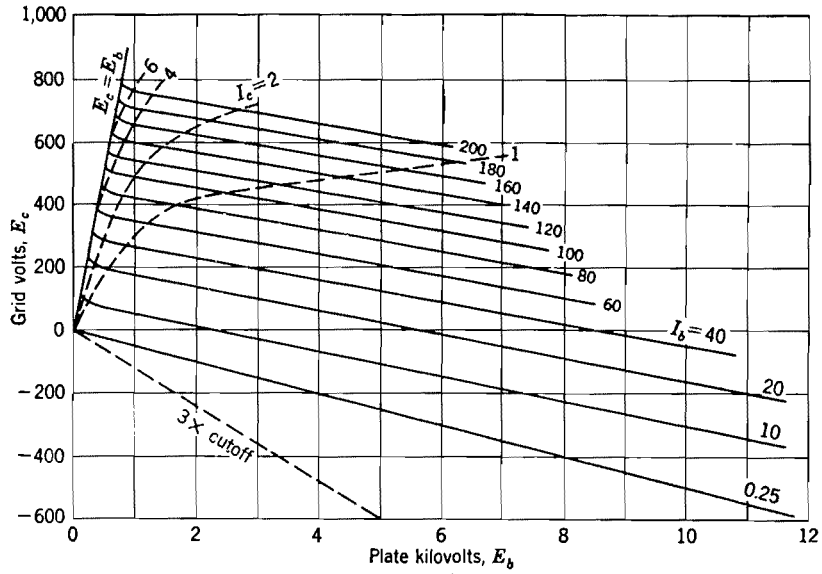

FIG. 11-29

**11-2.** For the tuned-grid oscillator of Fig. 11-10*a*, under the assumptions of Sec. 11-7 and the equivalent plate circuit, derive expressions for the frequency of oscillation and the $g_m$ criterion for oscillation.

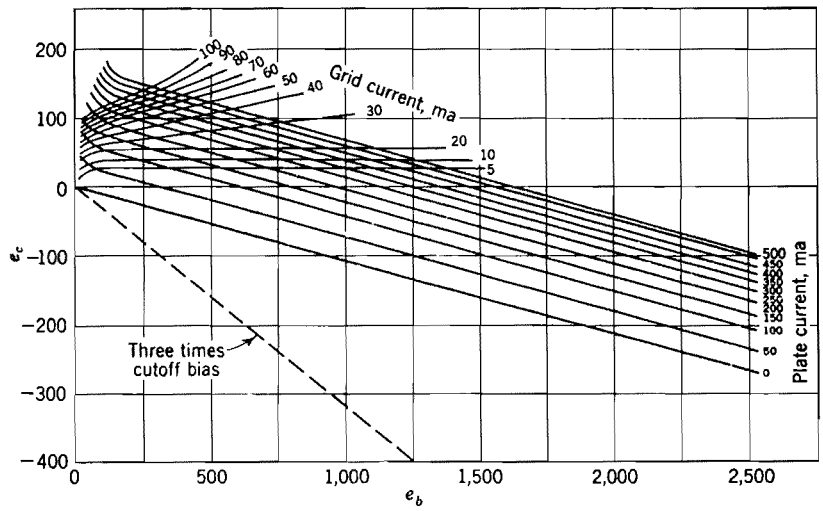
**11-3.** The triode of Fig. 11-30*a* operates as a Class C power amplifier at $E_{bb}$ = 5000 volts, $E_{cc}$ = three times cutoff, $\hat{E}_g$ = 1300 volts. If $e_{b,\min} = e_{c,\max}$ and $f$ = 1.5 Mc, compute (*a*) the d-c plate current, (*b*) rms value of $I_1$, (*c*) power output, (*d*) power input to the plate circuit, (*e*) plate-circuit efficiency, (*f*) the resonant load impedance, (*g*) plate loss, (*h*) grid driving power.

**11-4.** A tuned-plate oscillator uses a triode of the characteristics of Fig. 11-30*b* and operates at 1.0 Mc. $E_b$ = 1500 volts, $E_{cc}$ = twice cutoff, $\hat{E}_g$ = 400 volts, $e_{b,\min} = e_{c,\max}$, loaded tank $Q$ = 12, tuning capacity $C$ = 200 $\mu\mu$f. Find the power output and the plate-circuit efficiency, and design the tank circuit.

**11-5.** The 75T tube of Fig. 11-30*b* is used as a Class C amplifier with 1500 plate volts, $E_{cc}$ = three times cutoff, a peak plate current of 400 ma, and $e_{b,\min} = e_{c,\max}$. Find (*a*) a-c power output, (*b*) power input from power supply, (*c*) plate-circuit efficiency, (*d*) grid driving power, (*e*) plate loss in watts. (*f*) Design the parallel-tuned load if loaded $Q$ = 12 and frequency is 0.5 Mc.




(*a*)




(*b*)

FIG. 11-30. (*a*) $\mu = 25$, maximum $E_{bb} = 10{,}000$ volts. (*b*) 75T tube, $\mu = 10.6$, maximum $E_{bb} = 3000$ volts.

**11-6.** The tube of Fig. 11-30*b* has ratings of maximum $E_{bb}$ = 3000 volts, maximum $I_b$ = 175 ma, plate dissipation = 75 watts, maximum grid current $I_c$ = 30 ma. It is operated in Class C at 1.59 Mc, with $E_{bb}$ = 2000 volts, grid bias of twice cutoff, $e_{b,\min} = e_{c,\max}$, and a peak plate current of 450 ma. Find (*a*) a-c power output, (*b*) input power and plate-circuit efficiency, (*c*) plate loss, (*d*) power gain, (*e*) required impedance of plate-circuit load at resonance. (*f*) Which, if any, of the tube ratings are being exceeded?

**11-7.** Using the constant-current curves of Fig. 11-4, plot the waveforms of grid current and plate current if $E_{bb}$ = 5000 volts, $E_{cc} = -500$ volts, $e_{b,\min} = 1.5e_{c,\max} = +400$ volts. Find the angles of conduction for grid and plate current.

**11-8.** Compute and plot the value of electric-field intensity across the diameter of a copper conductor 0.6 cm in diameter for a frequency of 7960 cycles.

**11-9.** Find the a-c resistance per meter of the conductor of Prob. 11-8. If the center of the conductor is removed to yield a tube of 1 mm wall thickness, also find the a-c resistance, and compare.

**11-10.** Show that the depth of penetration given by Eq. (11-52) has dimensions of length. Also show that $1/\sqrt{\mu\epsilon}$ has dimensions of velocity.

**11-11.** A round nickel tube, 1 in. diameter, with ⅛-in. wall and 3-in. length, is to be induction-heated to a temperature of 700°C. Assuming that heat is lost only by radiation, determine heat input required with an ambient temperature of 27°C. If twice this value of input is supplied, determine the time required for heating. Determine the lowest practical frequency for good energy transfer.

**11-12.** A plastic has $\epsilon_r$ = 3.7, specific heat $s$ = 600, density 2000 kg/cum, and $\alpha$ = 0.05. The dielectric is in the form of a slab 10 in. diameter and 1 in. thick. Find the maximum frequency to be used if the heat developed at the edge is to be 0.9 of that at the center.

**11-13.** A steel bar 2 in. long by ¼ in. diameter is to be heated by a long coil of ¼-in. copper tubing, having three turns per inch, and of 1½-in. diameter. Assuming radiation losses only and relative permeability of unity, plot a heating-time curve for the steel if it is to be heated to 1000°C from 37°C. For a frequency of 100 kc determine the coil current required and find the time of heating if 4.0 kw is used.

**11-14.** A copper tube of ½ in. OD, ⅛ in. ID is 24 ft long. Find its resistance at 100,000 cycles and at 10 Mc. Repeat if the tubing is made of nickel, and explain why the ratio of increase over $R_{dc}$ is less.

## REFERENCES

1. Cady, W. G.: The Piezo-electric Resonator, *Proc. IRE*, **10**: 83 (1922).
2. Dow, J. B.: A Recent Development in Vacuum Tube Oscillator Circuits, *Proc. IRE*, **19**: 2095 (1931).
3. Llewellyn, F. B.: Constant-frequency Oscillators, *Proc. IRE*, **19**: 2063 (1931).
4. Pierce, G. W.: Magnetostriction Oscillators, *Proc. IRE*, **17**: 42 (1929).
5. Mason, W. P.: Low Temperature Coefficient Quartz Crystals, *Bell System Tech. J.*, **19**: 74 (1940).
6. Terman, F. E., et al.: Some Applications of Negative Feedback with Particular Reference to Laboratory Equipment, *Proc. IRE*, **27**: 649 (1939).
7. Ryder, J. D.: "Electronic Fundamentals and Applications," Prentice-Hall, Inc., Englewood Cliffs, N.J., 1950.
8. Ryder, J. D.: "Networks, Lines, and Fields," 2d ed., Prentice-Hall, Inc., Englewood Cliffs, N.J., 1955.

9. Sohon, Harry: "Engineering Mathematics," D. Van Nostrand Company, Inc., Princeton, N.J., 1944.
10. Reed, M. B., and Georgia Reed: "Mathematical Methods in Electrical Engineering," Harper & Brothers, New York, 1951.
11. Jahnke, E., and F. Emde: "Tables of Functions with Formulae and Curves," Dover Publications, New York, 1943.
12. McLachlan, N. W.: "Bessel Functions for Engineers," 2d ed., Oxford University Press, New York, 1955.
13. Brown, G. H., C. N. Hoyler, and R. A. Bierwirth: "Theory and Application of Radio-frequency Heating," D. Van Nostrand Company, Inc., Princeton, N.J., 1947.
14. Carlin, B.: "Ultrasonics," McGraw-Hill Book Company, Inc., New York, 1949.
15. Bierwirth, R. A., and C. N. Hoyler: Radio Frequency Applied to Wood Gluing, *Proc. IRE*, **31**: 529 (1943).
16. Stansel, N. R.: Induction Heating—Selection of Frequency, *Trans. AIEE*, **63**: 755 (1944).

CHAPTER 12

# SEMICONDUCTORS; TRANSISTORS

The properties of semiconducting crystal rectifiers have long been employed in radio reception. These semiconductors, with properties between those of good conductors and those of insulators, have recently been more thoroughly investigated, and this research has led to further applications and particularly to the *transistor*, a semiconductor control device having amplifying as well as rectifying properties.

The transistor is able to perform the functions usually assigned to the triode vacuum tube, although it is at present limited to a few megacycles as an upper frequency limit for amplification and has a higher internal noise level. The vacuum tube is a voltage-controlled device, while in contrast the transistor is current-controlled and is in fact the circuit dual of the vacuum tube. This duality aids in understanding transistor performance by analogy, but the concept should not be carried to extremes.

The transistor boasts a major advantage over the vacuum tube in that it requires no cathode heating power. With lower internal impedance levels it is operable at lower applied voltages than the vacuum tube, giving greater over-all power efficiencies. Since there is no emitting cathode, very long life is anticipated for transistors.

Analysis and design of transistor circuits follow the pattern already discussed for vacuum-tube circuits, provided due consideration is given to the requirements for operating impedance levels and the fact that the transistor has a finite input impedance even at low frequencies. The parameters of the transistor will be discussed and similarities to the vacuum tube pointed out, without repetition of the applicable circuit analysis and design methods of preceding chapters.

**12-1. Semiconductors.** A good electrical conductor such as copper and an insulator such as steatite differ greatly in electrical conductivity, this difference being of the order of $10^{17}$ to $10^{19}$. Intermediate between these extremes are the semiconductor materials, such as germanium, silicon, selenium, and various metallic sulfides and oxides. Of these the most important at present seem to be germanium and silicon.

In a good conductor there are available a large number of relatively free electrons to serve as charge carriers in an electric current. In an

insulator all the charges are tightly bound to the atoms and are not available for conduction at normal temperatures. A pure semiconductor at absolute zero has all its charges tightly bound and would behave as an insulator, but as the temperature is raised, a limited number of charges become available for conduction. Because the number of charge carriers in semiconductors is limited and dependent on temperature, these materials may have conduction properties quite different from those of good conductors. For instance, semiconductors may have a negative temperature coefficient of resistance over some temperature range and may exhibit polarity effects or not be strictly bilateral at a junction with a contacting conductor.

Metals and semiconductors occur in polycrystalline form, and the crystal boundaries introduce inconsistencies in conduction from crystal to crystal. For lowered resistance and more consistent conductive properties it has been found desirable to grow large single crystals of a semiconductor from which small wafers are cut. These large single crystals are grown by dipping a small seed crystal into a bath of molten semiconductor and withdrawing the seed at a constant rate as the crystal grows and cools.

The resultant crystal structure has a regular pattern of atom arrangement called a *lattice*. Silicon and germanium are elements from the fourth group of the periodic table and crystallize in the cubic-lattice form as does carbon in the diamond. Each atom of silicon, of atomic number 14, has its electrons held in groups of energy levels, or *shells*, containing 2, 8, 4 electrons respectively. The inner shells of 2 and 8 electrons have no excess orbits or energy levels and are stable and complete; the outer 4 electrons are known as *valence electrons* and contribute to the chemical properties of the atom. Germanium, of atomic number 32, has its electrons held in shells containing 2, 8, 18, 4 electrons and has chemical and electrical properties similar to silicon because of the four valence electrons in the outer, or incomplete, shell.

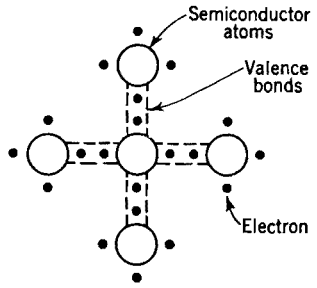


FIG. 12-1. Symbolic structure of a pure germanium crystal.

Each atom of germanium or silicon in its lattice is joined to its four nearest neighbors by four valence bonds each having two valence electrons, one from each atom as in Fig. 12-1. As a result, each atom is surrounded by a stable shell of eight electrons, and the atoms and electrons are held rigidly in place. Since no charges are free to move, then germanium and silicon in this stable condition will behave as insulators. As the temperature is raised, kinetic energy is given to the electrons and

a few of the valence bonds will be broken and these electrons will become free. At the same time an electron vacancy, or *hole*, is left in the valence bond, so that actually an *electron-hole pair* is formed. The hole acts as a positive charge, or is attractive to an electron, and both electron and hole may contribute to current flow under action of an applied electric field. At room temperature germanium may have an *intrinsic conductivity* of about 1.67 mho per m,* even though only about one electron is freed per $10^{10}$ germanium atoms.

In usual temperature ranges, conduction due to impurity atoms present in the lattice is of greater importance than that due to intrinsic or

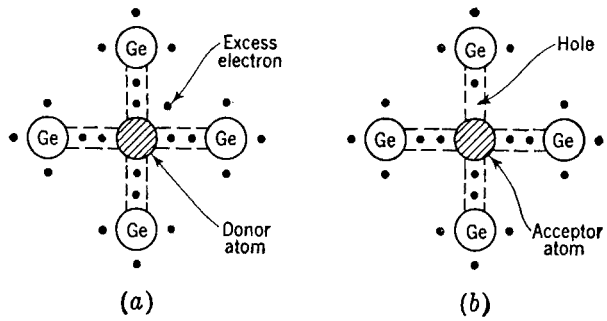



FIG. 12-2. (*a*) Symbolic *N*-type lattice. (*b*) Symbolic *P*-type lattice.

thermally released electrons. A few atoms of an element having five valence electrons, such as phosphorus, antimony, or arsenic from the fifth group of the periodic table, may be introduced into the high-purity germanium. The impurity atom will not be able to form a complete bond; four of its five valence electrons will be found in the germanium bonds, but the fifth will be left relatively free to wander through the crystal and contribute to current conduction. Such conduction is then by negative charges or electrons and is called *N-type*, the impurity atoms being called *donors*. The situation is illustrated in Fig. 12-2*a*.

Impurity atoms with three valence electrons from the third group of the periodic table, such as boron, aluminum, gallium, or indium, will incompletely fill the germanium valence bonds and leave a vacancy, or hole, in one bond. Under electric-field attraction an electron may come from a nearby bond and fill this hole. This electron in turn leaves a new hole, or vacancy, behind. Effectively, a hole appears to wander

* Conductivity between opposite faces of a meter cube. In much semiconductor literature the conductivity between faces of a centimeter cube or as mhos per centimeter is used. Conversion follows from

$$\text{mho/cm} = \tfrac{1}{100}\ \text{mho/m}$$

Likewise, for resistance,

$$\text{ohm-cm} = 100\ \text{ohm-m}$$

about the crystal and will contribute a positive charge to the conduction, which is then referred to as *P-type*. Such an impurity atom is called an *acceptor*, and the situation is symbolically shown in Fig. 12-2*b*.

More accurately, it should be stated that the electron takes on additional energy when it is freed from the valence bond. When the free electron enters a hole in a valence bond, it releases this excess energy as a quantum of heat or light. The quantum may be reabsorbed by another electron to break its valence bond and create a new electron-hole pair. Conduction by holes differs from that by electrons because the freedom of movement, or *mobility*, of the hole is only about half that of the electron. Mobility $\mu$ is defined as velocity per unit applied electric field.

It should be emphasized that throughout the above processes the material remains electrically neutral, and for every electron or hole there is always an opposite charge present in bound or unbound form.

Germanium is a hard, dense material melting at 937°C. It is obtained by reduction of the oxide, which is found in a flue residue from zinc smelting. After extreme purification the germanium melt may be appropriately doped under careful control to yield either $N$ or $P$ germanium, or impurities may be added at proper times as the crystal is being pulled from the melt to give separate layers of $N$ or $P$ conduction. Other methods of controlled diffusion of impurities into the initially purified semiconductor are also employed.

**12-2. Energy Levels and Conductivity in Semiconductors.** The electron-energy-level discussion of Sec. 2-4 will be useful in further explanation of the operation of semiconductors as rectifiers and transistors. The diagram of Fig. 12-3*a* shows the distribution of electron energies in a good conductor at 0°K and at a higher temperature $T$. It should be noted that there may be additional bands of electron energies to the left of the vertical axis, these representing the inner shells of bound electrons, not contributing to conduction or chemical properties.

The bands represent an enormous but finite number of discrete energy levels, created by the interactions of the closely packed nuclei. Since the energy differences between levels are so small, many electrons can acquire small amounts of energy from thermal or electric-field sources and move into higher energy levels, as shown by the curve for temperature $T$.

In a semiconductor it is found that the valence electrons are normally bound in an energy band in which all available states or valence-bond positions are filled or occupied. This filled band is separated from a normally unoccupied higher energy band by an unallowed or *forbidden energy gap*, of energy magnitude $\Delta E$. Quantum-mechanics theory requires that, when an energy band is completely filled, the electrons therein can-

not contribute to electrical conductivity. There is no open energy level to which they can move after absorbing energy from an electric field; therefore they do not absorb energy and cannot become conduction electrons. The conductivity of a pure semiconductor in Fig. 12-3b is zero at absolute zero temperature since the lower energy or valence band is filled and there are no electrons in the upper band.

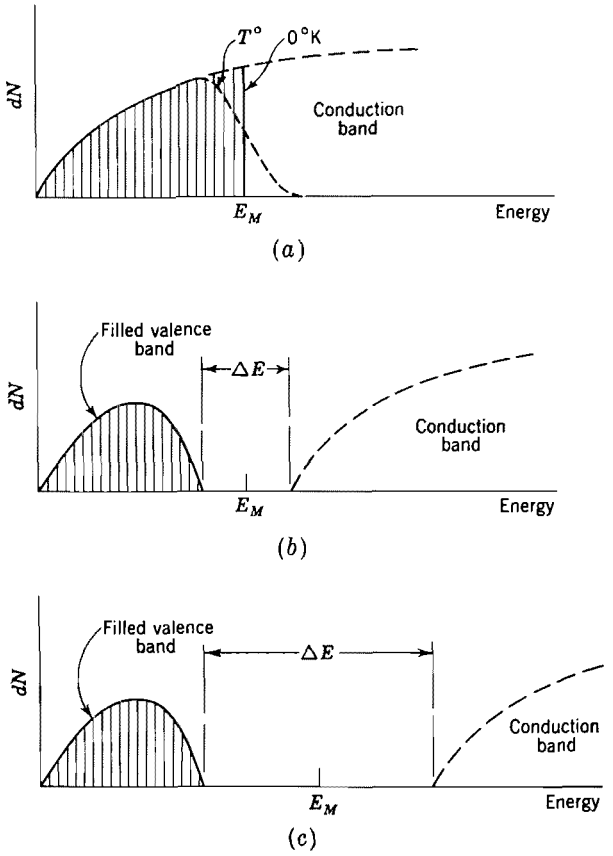



FIG. 12-3. Density of electron energy states, at 0°K, in (*a*) a conductor, (*b*) a pure semiconductor, (*c*) an insulator.

As the temperature is raised, it is ultimately possible for some electrons to receive energy equal to or greater than the gap energy $\Delta E$, and these will transfer to the upper energy band. Since these electrons are now in an incompletely filled band, they may contribute to conduction. In so transferring to the upper conduction band they left holes in the valence band, or formed electron-hole pairs. It is these electron-hole

pairs formed by thermal breaking of the valence bonds which give a semiconductor its intrinsic conductivity.

An insulator as in Fig. 12-3*c* differs from a semiconductor mainly in the magnitude of $\Delta E$, this gap energy running up to 7 to 12 ev. Thus much higher temperatures or thermal energies are required to make usual insulators conduct. The intrinsic resistivity of a semiconductor is related to the gap energy $\Delta E$ as

$$\rho = A\epsilon^{\Delta E/kT} \times 10^{-2} \text{ ohm-m} \tag{12-1}$$

where $A$ = a coefficient varying slightly with temperature
$k$ = Boltzmann's constant = $8.69 \times 10^{-5}$ ev/deg
$T$ = temperature, °K

The value of $\Delta E$ is obtainable from the slope of a plot of $\ln \rho$ against $1/T$, and a few typical values are given in Table 25.

TABLE 25. FORBIDDEN GAP ENERGY, $\Delta E$

| *Material* | *Gap energy, ev* |
|---|---|
| Diamond | 7 |
| Silicon | 1.12 |
| Germanium | 0.72 |
| Lead | 0 |

Pure germanium has an intrinsic resistivity of about 0.65 ohm-m (ohm per cubic meter) at 25°C, and this value is often used as a standard of purity. Germanium and silicon have $4.5 \times 10^{22}$ and $5.2 \times 10^{22}$ atoms per cubic centimeter, and the sensitivity of the resistivity to traces of impurity is shown by Fig. 12-4, where the resistivity is plotted as a function of the number of semiconductor atoms per impurity atom.

Rectifier or transistor action is obtained by reason of majority employment of only one type of charge carrier, either electron or hole, in a given material. Intrinsic conduction by electron-hole pairs is not desired. Very high-purity germanium or silicon must be employed to keep down the intrinsic conductivity at normal operating temperatures, the material then being doped with the selected impurity to provide the desired type of charge carriers for either $N$ or $P$ conduction. If operated at too high a temperature, intrinsic conduction increases and may mask the currents due to the impurity carriers. It can be seen from Fig. 12-5 that a germanium material having a resistivity of about 0.04 ohm-m (ohm per cubic meter) remains an impurity conductor up to possibly 80°C. Intrinsic conductivity places a definite upper limit on the operating temperature.

The gap energy of silicon is greater than that of germanium, and silicon requires higher thermal energy for the breaking of valence bonds and forming electron-hole pairs for intrinsic conduction. This means

that silicon devices may be operated at higher temperatures than germanium before intrinsic conduction destroys operating characteristics.

The energy-band picture explains why all elements with four valence electrons, namely, carbon, germanium, silicon, and lead, do not behave alike. Carbon in crystal form (diamond) is an insulator because $\Delta E$ is so large. Germanium and silicon have smaller gap energies and are

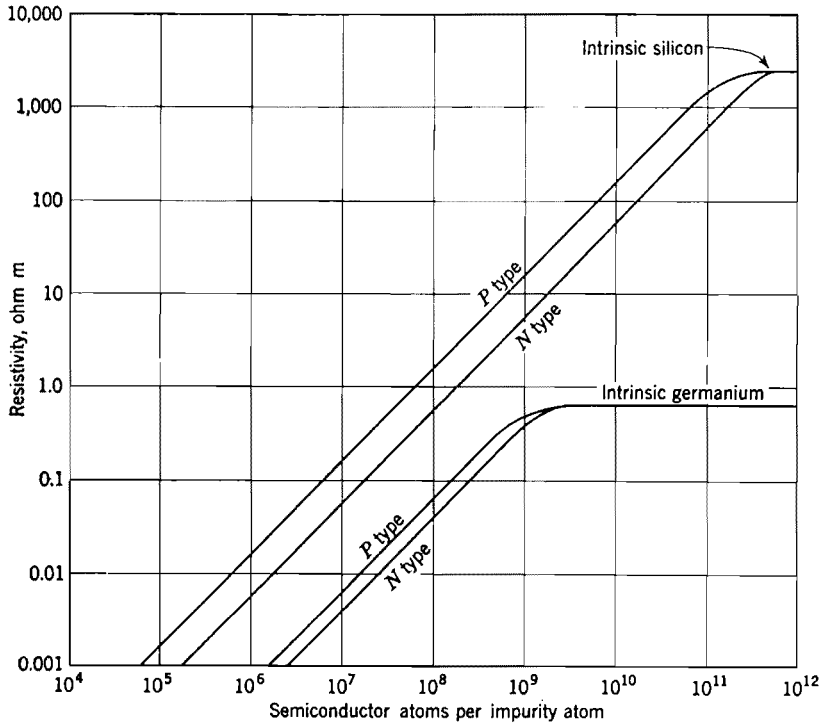



FIG. 12-4. Resistivity, at 25°C, as a function of impurity content.

semiconductors, while in lead the gap energy is zero and the metal is a conductor.

Conduction by holes or electrons contributed by acceptor or donor impurities may also be discussed in terms of the energy bands in the crystal. As shown in Fig. 12-6*a*, the absence of some electrons in the valence band, as produced in *P* material, leaves unfilled levels in the lower band. Since the band is no longer filled, it is possible for the holes therein to contribute to conduction, by transfer of electrons from hole to hole, that is, for the hole to move and ultimately absorb an electron from the external circuit. Likewise, in the *N* material represented in Fig. 12-6*b*, it is found that the relatively free excess electrons from the donor atoms

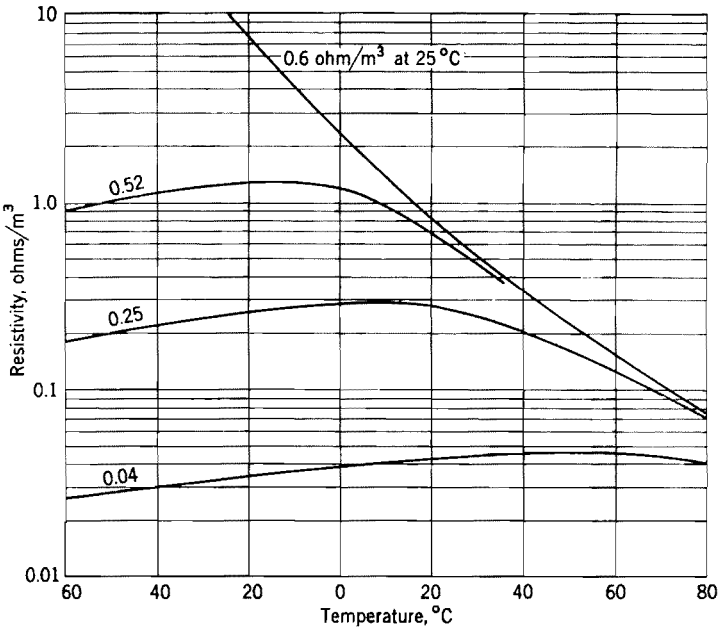



FIG. 12-5. Resistivity vs. temperature for $N$-type germanium.

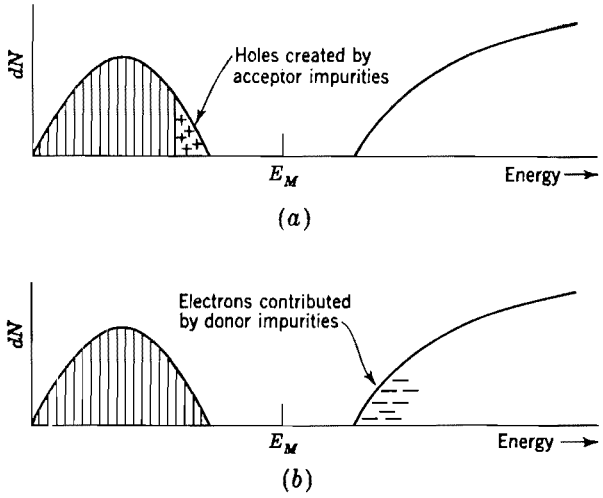



FIG. 12-6. (*a*) $P$-type material. (*b*) $N$-type material.

have energies sufficiently high to place them in the upper conduction band, allowing conduction by electrons.

**12-3. The *P-N* Junction.** The control of the conductivity of germanium or silicon by the addition of traces of other elements has been discussed. *N* or *P* conduction can be obtained as desired by introducing first one kind of impurity, then another, into the growing crystal, giving regions of alternating *N* and *P* conductivity, as in Fig. 12-7*a*. Such *junctions* may also be formed by fusing the impurity material onto thin germanium wafers, as in Fig. 12-7*b*, or by vapor diffusion or electroplating the impurity material onto the germanium or silicon. By careful manufacturing control the transition between the two regions can be made quite abrupt.

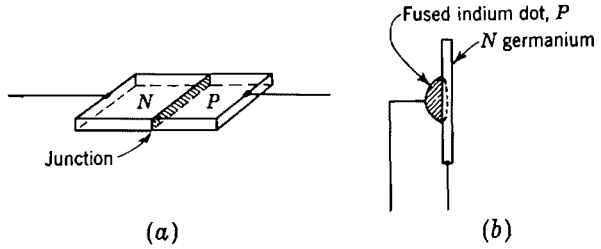



FIG. 12-7. (*a*) Grown-junction diode. (*b*) Fused-junction diode.

In the *N* material when formed there are many relatively free high-energy electrons. In the *P* material there are holes, with a few intrinsic electrons contributed by thermal energy. Some of the high-energy electrons from the *N* material are able to pass across the *N-P* boundary of the junction and fill some of the holes on the *P* side. Such a flow of electrons cannot continue since it places an excess of electrons, or a negative charge, on the *P* material and a deficiency of electrons, or a positive charge, on the *N* material. The result is a potential difference across the depleted charge region, or *barrier layer*, at the junction of *N* and *P*, opposing any further movement of electrons across the boundary, as illustrated in Fig. 12-8*c*.

In magnitude the potential difference, or barrier potential, across the *N-P* junction may be found from

$$E = \frac{kT}{e} \ln \frac{n_e n_h}{n_i^2} \quad \text{volts} \tag{12-2}$$

where $n_i$ is the density of electron-hole pairs and $n_e$ and $n_h$ the density of electrons and holes, respectively, in *N* and *P* regions. The electric field is directed from the *P* to the *N* region, or the *N* region is positive to the *P* region because the normal number of electrons in the *N* region is depleted. The normal *P-N* junction exists with this potential across the

barrier, or transition, layer, which may be about $10^{-4}$ cm thick. Of course, there has been no over-all change in total charge, and the junction as a whole remains electrically neutral.

The *P-N* junction may be shown to have rectification or diode properties. If an external emf is applied to a junction in the so-called *reverse direction*, as in Fig. 12-9*a*, with *P* region negative, the majority carriers on each side will be forced farther away from the junction, or the barrier

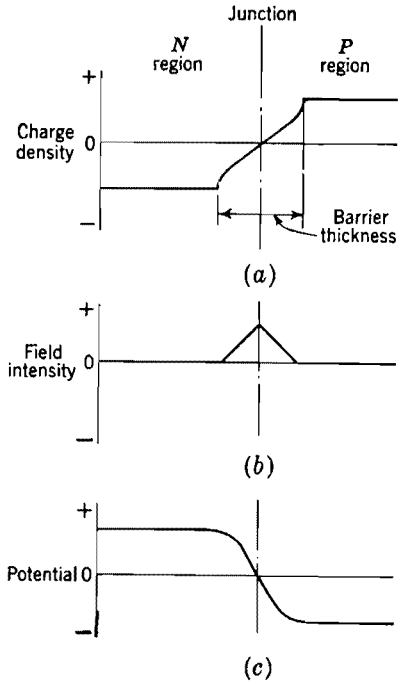



FIG. 12-8. (*a*) Showing depleted charge region due to migration across boundary. (*b*) The resulting field intensity. (*c*) The potential change across the boundary.

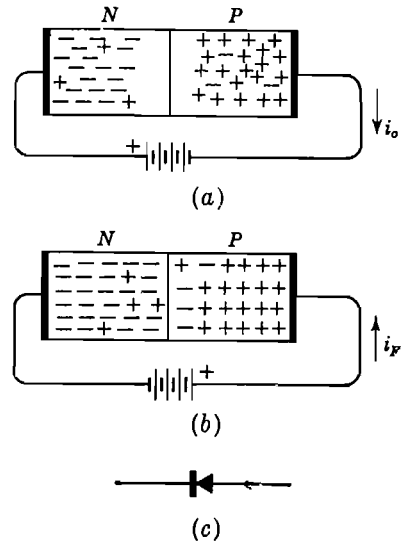



FIG. 12-9. (*a*) *P-N* diode reverse-biased. (*b*) *P-N* diode forward-biased. (*c*) Circuit symbol for rectifier diode; arrow in forward direction.

layer will be thickened, leaving only a few thermally generated minority carriers in the barrier to produce a very small current $i_o$. In this case the diode presents a very high resistance with an additional capacitive component, due to the dielectric of the barrier layer between *N* and *P* regions.

If an external emf is applied in the *forward direction*, as in Fig. 12-9*b*, with *P* region positive, the majority carriers from each side may move across the junction, or the potential barrier is wiped out. This is the easy flow direction, and a substantial current may flow. The current

will depend on the density of the $N$ or $P$ carriers, the area of the junction, and the external potential. The actual type of majority carrier is largely determined by the resistivity of the region or the impurity content and the mobilities, the type of carrier from the region of lower resistivity usually predominating.

**12-4. Junction and Point-contact Diode Rectifiers.** The $P$-$N$ junction is employed as a rectifying diode, for which a typical volt-ampere curve

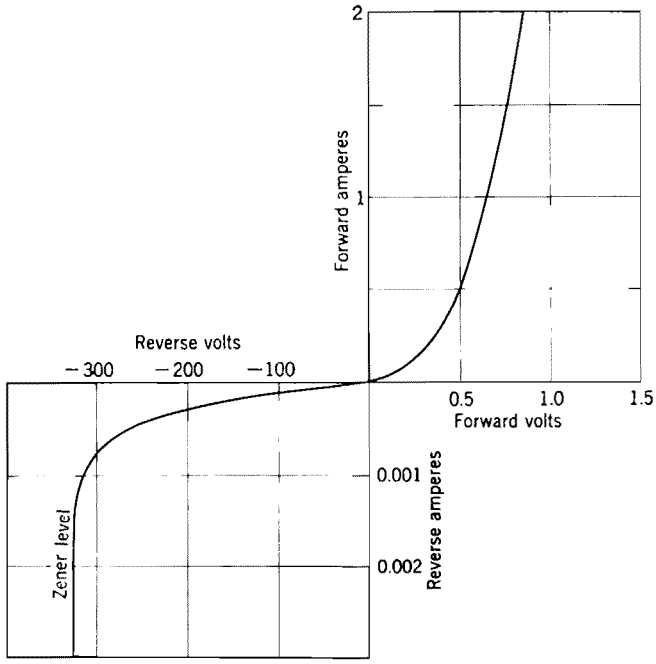



FIG. 12-10. Volt-ampere curve for $P$-$N$ junction diode. (Note scale changes.)

is shown in Fig. 12-10. The change in scale factors between forward and reverse quantities should be noted. A great advantage of the junction rectifier is its low forward voltage drop, which permits high power efficiencies.

If the reverse potential is increased beyond a certain value, the current increases rapidly, as can be seen in the figure at about −320 volts. This current is due to electrons being pulled out of the valence bonds by the high electric field present across the junction in the reverse direction. Beyond this breakdown point, usually called the *Zener level*, the voltage across the junction remains practically constant over a wide current range. An obvious application is found as a voltage regulator.

The extraction of electrons from the valence bonds is a reversible process and is not destructive to the crystal if excessive temperatures can be prevented. The voltage at which the Zener effect occurs is quite stable, and methods are available for controlling the voltage level at which the effect takes place.

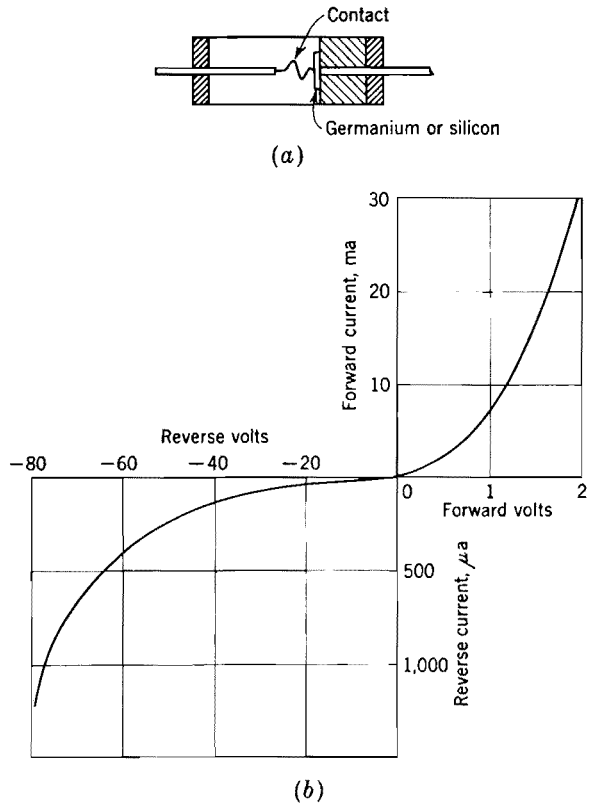



FIG. 12-11. (*a*) Point-contact rectifier. (*b*) Volt-ampere curve for 1N34 diode.

Because of the high internal capacity of the barrier layer the junction rectifier is largely useful only at low frequencies. However, the cat-whisker, or point-contact, type of rectifier of the early radio days has been revived and improved, with germanium or silicon as the semiconductor, and is now frequently employed as a low-shunt-capacity form of detector of microwaves. Construction of such a point-contact unit is shown in Fig. 12-11. Small size is another valuable property, the germanium wafer being of the order of 3 mm square and 0.6 mm thick. The small size and thin contacting wire lead to internal capacities of only a

few micromicrofarads, so that operation is possible to frequencies of thousands of megacycles.

The semiconductor is usually manufactured as $N$-type. During manufacture a high momentary current is passed between contact and semiconductor, and this is believed to form a $P$ region around the contact. Thus the theory of operation follows that already given for the junction unit.

The volt-ampere curves for a point-contact unit appear in Fig. 12-11*b*, differing from those of the junction unit only in current magnitudes. The forward resistance of the point-contact rectifier is considerably lower than that of the vacuum diode, giving improved performance when used in clipping and other wave-shaping circuits. A variety of germanium and silicon rectifiers are available, with various design features such as low forward resistance, high back resistance, low capacity, and high back voltage.

**12-5. The Point-contact Transistor.** In 1948 Bardeen and Brattain reported on work with a second point contact on a germanium diode, which showed that when the second contact was biased in the reverse direction its current could be controlled by the current through the forward-biased first rectifying contact. They called the resulting solid-state three-electrode device the *transistor*.

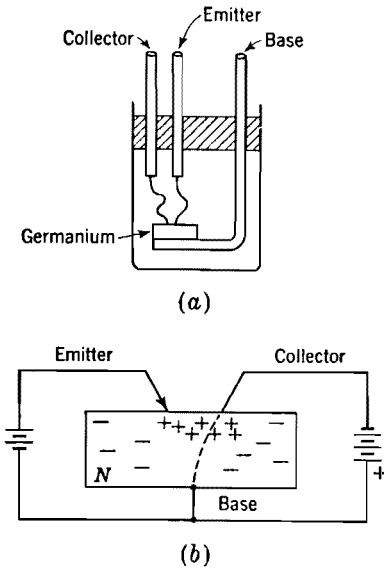



FIG. 12-12. (*a*) Point-contact transistor. (*b*) Action of point-contact transistor.

The forward-biased electrode is called the *emitter*, the reverse-biased electrode being the *collector*, and the germanium or silicon element is the *base*, as shown in Fig. 12-12. The unit uses an $N$-type semiconductor wafer about 0.5 mm thick, and the two wire contacts are spaced about 0.05 mm (0.002 in.).

The positive bias on the emitter immediately removes electrons from the valence bonds near the surface, creating holes. In fact the emitter is said to *inject* holes into the semiconductor, and the emitter current is composed of electrons. The holes immediately diffuse toward the negative collector. In the process many holes may meet and recombine with the free electrons in the $N$-type base material. To reduce this loss of holes, the emitter and collector must be close-spaced. Holes that reach

the collector area combine with excess electrons injected into the region by the collector, and this permits the reverse-biased collector to inject more electrons, increasing the collector current to the base.

The holes, in traveling from emitter to collector by random paths, set up a positive space charge. This space charge attracts more electrons from the collector and serves to further increase the electron flow from the collector. While some of these electrons serve to neutralize holes, the majority flow toward and reach the base terminal, forming a base-collector current.

The mobility of electrons is about twice that of holes, so that a hole is on the average able to neutralize the charge due to about two electrons, and the collector current may be expected to have a ratio of electrons to holes of about 2, with a total collector current about three times the emitter current. It may also be seen that this collector current depends on the number of holes injected at the emitter and can be controlled by the emitter current, since otherwise the current to the reverse-biased collector would be very small.

*The ratio of a change in collector current $\Delta i_c$ to a given change in emitter current $\Delta i_e$, with collector voltage constant, is called the short-circuit current amplification factor $\alpha_{ce}$.* That is, in the limit

$$\alpha_{ce} = \frac{\partial i_c}{\partial i_e}\bigg]_{v_c = \text{constant}} \tag{12-3}$$

From the above discussion it would be expected that $\alpha_{ce}$ for a point-contact transistor would be about 3. This ratio may not be reached, however, because of recombination in transit, and usual values are in the range of 2 to 3. There are also examples in which $\alpha_{ce}$ may exceed 3. Understanding of this phenomenon is not yet complete, but it is believed due to properties developed near the collector by the forming operation.

A point-contact transistor can also be made from $P$ material. The emitter is forward-biased but negative, the collector back-biased and positive.

The input resistance of the emitter-base circuit is low, of the order of 100 to 500 ohms, because of the forward bias. The output resistance between collector and base is relatively high, because of the back bias, and of the order of 10,000 to 50,000 ohms for point-contact units. For a unit with $\alpha_{ce} = 2.5$, input resistance of 300 ohms, and a load of 20,000 ohms the voltage gain may approximate 160 and the power gain 400 or 26 db. Thus, even with a low current gain the voltage and power gain from a transistor may be large because of the difference between input- and output-impedance levels.

**12-6. The Junction Transistor.** A germanium or silicon crystal may be grown with two $N$ regions separated by a very thin (0.002-cm) $P$

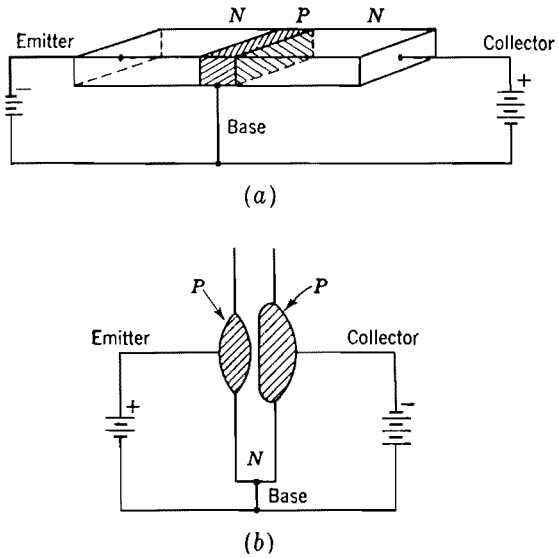



FIG. 12-13. (*a*) *N-P-N* grown-junction transistor. (*b*) Fused-alloy *P-N-P* junction transistor.

region, or with welded or fused *P* regions on a very thin *N* section, thus forming *N-P-N* or *P-N-P* junction transistors, as in Fig. 12-13.

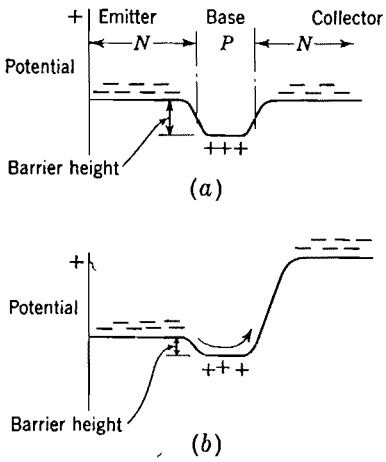



FIG. 12-14. (*a*) Potential levels present in unbiased *N-P-N* transistor. (*b*) Potential levels present in *N-P-N* unit biased for operation.

For the junction construction the first diode, or *emitter, junction* is biased in the forward direction. For the *N-P-N* unit this means the emitter is negative, and thus the emitter-junction potential barrier is lowered, as shown in Fig. 12-14*b*, in contrast with the unbiased level in Fig. 12-14*a*. Electrons enter the *N* region from the emitter, and those of higher energy are able to pass over the small barrier into the *P* region. The barrier height and the number of these high-energy electrons reaching the *P*, or base, region are a function of the emitter-base potential. In this operation the emitter barrier acts as an electron velocity or energy filter much as does the grid in a vacuum tube.

Some of the electrons (possibly a few per cent) combine with holes in the hole-rich base region, but because the base is very thin, most elec-

trons pass on into the $N$-type collector region under attraction of the back-biased but positive collector, and these electrons constitute the collector current. Though the emitter-collector current-amplification factor $\alpha_{ce}$ may have a value approximating 0.95, amplification occurs because electrons injected through the emitter into a low-impedance forward-biased circuit are extracted at the collector at a very much higher impedance level.

The collector potential is made high enough so that a small reverse current $i_{co}$ flows with zero emitter current. Ordinarily $i_{co}$ is small with respect to $i_c$, the total collector current. It is found that $i_{co}$ does not vary appreciably with $v_c$, the collector-base voltage. The current $i_{co}$ is highly temperature-sensitive, since it is due to intrinsic conduction, and may become of appreciable magnitude at high operating temperatures.

Another current component present consists of holes flowing from base, or $P$, layer to the emitter as in Fig. 12-15*a*, and this current is also a function of emitter-base potential. The ratio of the electron and hole currents at the emitter is expressible as

$$\frac{i_e}{i_h} = \frac{L\sigma_n}{W\sigma_p}$$

where $L$ = diffusion distance for a hole in $N$ region
$W$ = thickness of $P$ layer
$\sigma_n, \sigma_p$ = conductivities of $N$ and $P$ regions
all units being consistent. By choosing the ratio $\sigma_n/\sigma_p$ and making $W$ small, the ratio $i_e/i_h$ may be controlled. When an input signal is applied to the base, the emitter-junction barrier height is modulated, causing the number of electrons injected by the emitter and the number of holes from the base both to vary. As a result the collector current will vary. Thus a *transistor is a device in which an input current controls an output current.* Therefore transistor performance is usually specified with currents as independent variables, rather than voltages as for the vacuum tube.

Space-charge effects are not present in junction units as they are in point-contact transistors. The emitter-collector short-circuit current gain will be less than unity, usually in the range 0.90 to 0.99, owing to recombination in the base region. A second current-amplification factor, the ratio between a change in collector current $\Delta i_c$ and a given change in base current $\Delta i_b$, with collector voltage constant, is called the *collector-to-base short-circuit current amplification factor* $\alpha_{cb}$. That is, in the limit

$$\alpha_{cb} = \left.\frac{\partial i_c}{\partial i_b}\right]_{v_c = \text{constant}} \tag{12-4}$$

This parameter may reach values of 40 or more. Since by definition

$$\Delta i_c = \alpha_{ce}\, \Delta i_e \tag{12-5}$$

and the base current $i_b$ is

$$\Delta i_b = \Delta i_e - \Delta i_c = (1 - \alpha_{ce})\,\Delta i_e \tag{12-6}$$

then it follows that

$$\alpha_{cb} = \frac{\alpha_{ce}}{1 - \alpha_{ce}} \tag{12-7}$$

Typical junction-transistor input resistances lie in the range 300 to 1000 ohms, with output resistances as high as 1 megohm. With $\alpha_{ce} = 0.95$, voltage gains of 2000 to 3000 are possible with power gains in the range

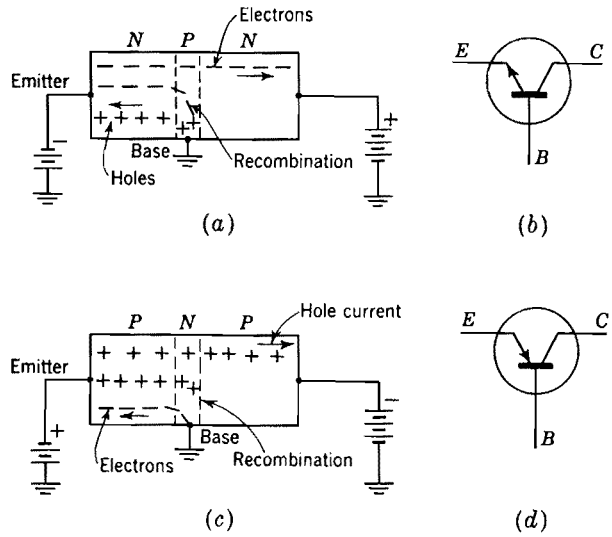



FIG. 12-15. (*a*) *N-P-N* action. (*b*) Proposed *N-P-N* transistor symbol. (*c*) *P-N-P* action. (*d*) Proposed *P-N-P* symbol.

of 30 to 40 db. These values compare favorably with those for a pentode vacuum tube.

The *P-N-P* transistor operates as in Fig. 12-15*b*, with reversed potentials and with conduction largely by holes. Otherwise no major performance differences are found.

**12-7. Transistors and Vacuum Tubes.** The transistor surpasses common forms of vacuum tubes in needing no heater power and in small size. However, the transistor is at present limited in upper frequency limit to a few megacycles per second as an amplifier and to a few hundred megacycles per second as an oscillator by reason of the drift time of the charges and the high internal capacities. Power output of the transistor is limited to a few watts, even with a high internal efficiency, because of the approximate 80°C operating-temperature limit, but higher-power units are under development. This temperature limit also prevents

transistor application in some equipment where ambient temperatures may be high, as in some industrial and aircraft installations. As yet transistors develop more random noise than do vacuum tubes, and this may be of some importance at low frequencies.

The junction transistor, because of its more nearly ideal and linear characteristics, permitting operation at very low voltages, is now replacing the point-contact type in most applications except those of switching

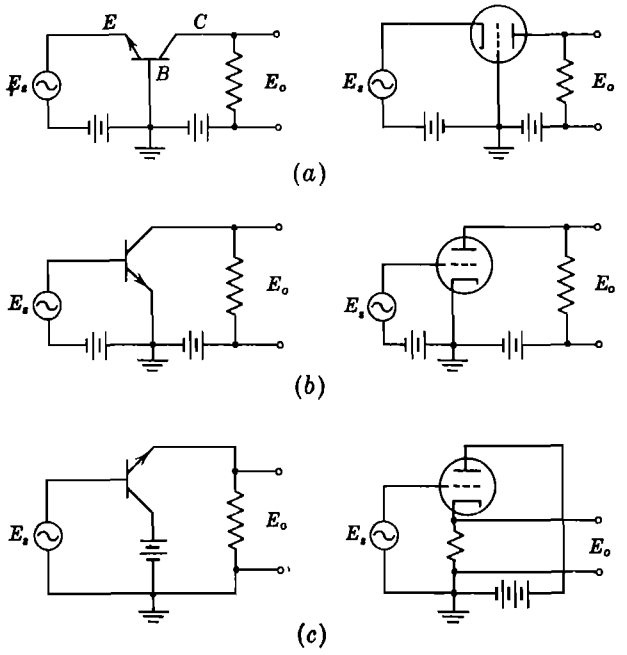



FIG. 12-16. (*a*) Common-base transistor amplifier; grounded-grid analogy. (*b*) Common-emitter amplifier; grounded-cathode analogy. (*c*) Common-collector amplifier; grounded-anode analogy.

or high-frequency amplification. It will be the junction unit which is studied here for amplification, unless otherwise stated.

Vacuum-tube linear Class A operation has been analyzed in terms of an equivalent circuit, starting from the functional relations

$$i_b = f(e_c,e_b)$$
$$i_c = 0$$

the second equation being implied in Class A operation. These equations indicate the choice of $e_c$ and $e_b$ as independent variables.

The transistor employs four analogous variables $v_e$, $v_c$, $i_e$, and $i_c$, or $v_b$ and $i_b$ may be introduced in place of one of the other pairs. Because

the transistor is a current-operated device and its performance is usually dependent on choice of operating currents, it is convenient to consider $i_e$ and $i_c$ as independent variables and to specify transistor performance by the functional relations

$$v_e = f_1(i_e,i_c)$$
$$v_c = f_2(i_e,i_c)$$

These choices are arbitrary, and it may occasionally be more convenient to choose other pairs of the variables as independent. Conventional network theory may then be applied to the transistor to determine its equivalent circuit and to analyze its performance.

The operation of the emitter of the transistor as the source of current carriers is analogous to that of the cathode of the vacuum tube. Likewise the base and collector serve functions which parallel those of the grid and anode of the vacuum tube. There are three basic circuits for the tube, and likewise there are three basic forms of transistor connection, the grounded, or common, base, the common emitter, and the common collector, analogous to the grounded grid, grounded cathode, and grounded anode, or cathode follower, with vacuum tubes. These are shown in Fig. 12-16.

Analysis of transistor circuits is more difficult, and the resultant expressions more complicated, than for the vacuum tube. Because of inherent transistor properties there are signal feed-through and signal feedback at all times, or the transistor has the properties of a four-terminal network, with all impedances or admittances finite.

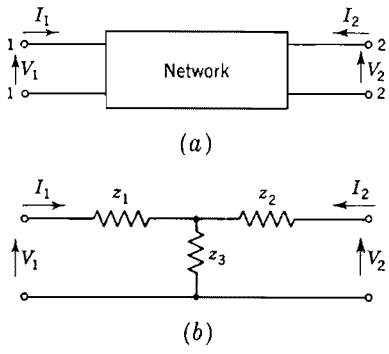



FIG. 12-17. (*a*) Arbitrary four-terminal network with current and voltage conventions. (*b*) T equivalent for (*a*).

**12-8. Four-terminal Networks.** A general four-terminal network is illustrated in Fig. 12-17*a*, with assumed currents and voltages, the latter being taken as potential rises in the direction of the arrows. Determination of eight quantities, the magnitudes and phase angles of $\mathbf{V}_1$, $\mathbf{V}_2$, $\mathbf{I}_1$, and $\mathbf{I}_2$, will completely specify the performance of any network in the box. It is possible to make eight measurements at the box terminals, and thus to determine the network without knowledge of its internal elements. These measurements are arbitrary but may be conveniently taken as the magnitudes and phase angles of the following impedances:

$$\begin{aligned} z_{11} &= \frac{V_1}{I_1} = \text{input impedance with output open or } I_2 = 0 \\ z_{22} &= \frac{V_2}{I_2} = \text{output impedance with input open or } I_1 = 0 \\ z_{12} &= \frac{V_1}{I_2} = \text{reverse transfer impedance with } I_1 = 0 \\ z_{21} &= \frac{V_2}{I_1} = \text{forward transfer impedance with } I_2 = 0 \end{aligned} \tag{12-8}$$

It is possible to replace any bilateral passive network, no matter how complicated, with a T network as in Fig. 12-17*b*. By inspection it is possible to write

$$V_1 = (z_1 + z_3)I_1 + z_3I_2 \tag{12-9}$$
$$V_2 = z_3I_1 + (z_2 + z_3)I_2 \tag{12-10}$$

It may also be observed that Eq. (12-8) gives for the circuit of Fig. 12-17*b*

$$z_{11} = z_1 + z_3 \qquad z_{21} = z_3$$
$$z_{12} = z_3 \qquad z_{22} = z_2 + z_3$$

The relation that $z_{12} = z_{21}$ holds in general only for a passive network and will not be expected to hold for the transistor as an active device.

The above values may be inserted in Eqs. (12-9) and (12-10), leading to

$$V_1 = z_{11}I_1 + z_{12}I_2 \tag{12-11}$$
$$V_2 = z_{21}I_1 + z_{22}I_2 \tag{12-12}$$

which are fundamental to any four-terminal network. For future reference it may be noted that the determinant of the network is

$$\Delta = z_{11}z_{22} - z_{12}z_{21} \tag{12-13}$$

The network behavior is specified by $V_1$, $V_2$, $I_1$, and $I_2$. In Eqs. (12-11) and (12-12) the currents $I_1$ and $I_2$ were selected as independent variables, thus allowing solution of these two equations for the other two variables. It may at times be convenient to choose $V_1$ and $V_2$, or $V_2$ and $V_1$, as independent variables, in which case Eqs. (12-11) and (12-12) become

$$I_1 = y_{11}V_1 + y_{12}V_2 \tag{12-14}$$
$$I_2 = y_{21}V_1 + y_{22}V_2 \tag{12-15}$$

and

$$V_1 = h_{11}I_1 + h_{12}V_2 \tag{12-16}$$
$$I_2 = h_{21}I_1 + h_{22}V_2 \tag{12-17}$$

Definitions for the y parameters can be found from the definitions for the z parameters. The h parameters can be found by simultaneous

solution of the z and y sets of equations as

$$\mathbf{h}_{11} = \frac{\mathbf{V}_1}{\mathbf{I}_1} = \frac{1}{\mathbf{y}_{11}} = \text{input impedance with } \mathbf{E}_2 = 0$$

$$\mathbf{h}_{12} = \frac{\mathbf{V}_1}{\mathbf{V}_2} = \mu_{12} = \text{reverse open-circuit voltage amplification}$$

$$\mathbf{h}_{21} = \frac{\mathbf{I}_2}{\mathbf{I}_1} = \alpha_{21} = \text{forward short-circuit current amplification}$$

$$\mathbf{h}_{22} = \frac{\mathbf{I}_2}{\mathbf{V}_2} = \frac{1}{\mathbf{z}_{22}} = \text{output admittance with } \mathbf{I}_1 = 0$$

The z's are known as open-circuit impedance parameters, the y's as short-circuit admittance parameters, and the h's as hybrid parameters, all being occasionally used in transistor analysis.

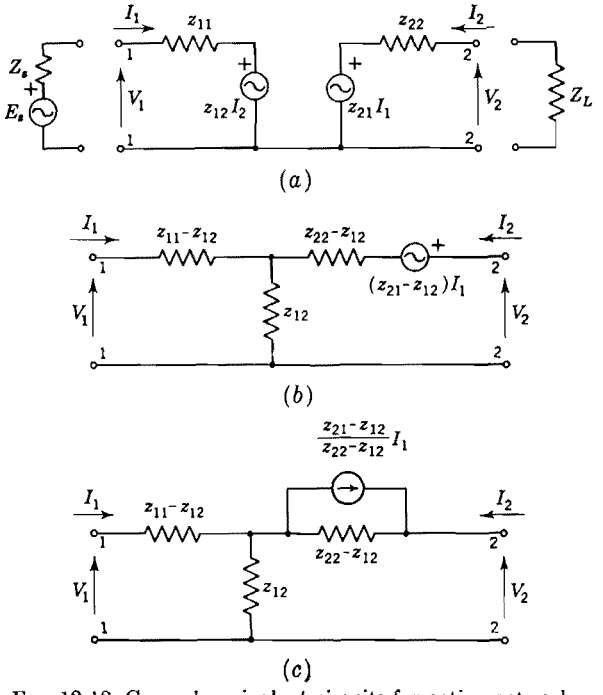



FIG. 12-18. General equivalent circuits for active networks.

Equation pairs, such as those of Eqs. (12-11) and (12-12), (12-14) and (12-15), or (12-16) and (12-17), are general and actually apply to any network whether active or passive. A T or $\pi$ section is an equivalent for any bilateral passive network, and a number of equivalent circuits for active four-terminal networks may also be found. As an example, Eqs. (12-11) and (12-12) may be taken.

$$\mathbf{V}_1 = \mathbf{z}_{11}\mathbf{I}_1 + \mathbf{z}_{12}\mathbf{I}_2$$
$$\mathbf{V}_2 = \mathbf{z}_{21}\mathbf{I}_1 + \mathbf{z}_{22}\mathbf{I}_2$$

These equations may be interpreted term by term as applying to the two-generator circuit of Fig. 12-18*a*, and therefore the circuit becomes equivalent for any network described by the above equations.

In more frequent usage is a one-generator form of equivalent circuit found by algebraic manipulation of Eqs. (12-11) and (12-12) to transfer the first mesh generator to an equivalent in the output circuit. Altering Eq. (12-11) by adding and subtracting $\mathbf{z}_{12}\mathbf{I}_1$ leads to

$$\mathbf{V}_1 = (\mathbf{z}_{11} - \mathbf{z}_{12})\mathbf{I}_1 + \mathbf{z}_{12}(\mathbf{I}_1 + \mathbf{I}_2) \tag{12-18}$$

Treatment of Eq. (12-12) by adding and subtracting $\mathbf{z}_{12}\mathbf{I}_1$ and $\mathbf{z}_{12}\mathbf{I}_2$ gives

$$\mathbf{V}_2 = (\mathbf{z}_{21} - \mathbf{z}_{12})\mathbf{I}_1 + (\mathbf{z}_{22} - \mathbf{z}_{12})\mathbf{I}_2 + \mathbf{z}_{12}(\mathbf{I}_1 + \mathbf{I}_2) \tag{12-19}$$

and these two equations are translated into circuit form in Fig. 12-18*b*. This is a one-generator form of equivalent circuit for an active network.

Another equivalent circuit appears in Fig. 12-18*c*. Here a current generator is used as an equivalent for the voltage generator $(\mathbf{z}_{21} - \mathbf{z}_{12})\mathbf{I}_1$. The current source bridged across $\mathbf{z}_{22} - \mathbf{z}_{12}$ must supply a current of value

$$\mathbf{I} = \frac{\mathbf{z}_{21} - \mathbf{z}_{12}}{\mathbf{z}_{22} - \mathbf{z}_{12}}\mathbf{I}_1 \tag{12-20}$$

Input and output impedances $\mathbf{Z}_{11}$ and $\mathbf{Z}_{22}$ are frequently required as design information for active networks. From Eq. (12-11)

$$\frac{\mathbf{V}_1}{\mathbf{I}_1} = \mathbf{Z}_{11} = \mathbf{z}_{11} + \mathbf{z}_{12}\frac{\mathbf{I}_2}{\mathbf{I}_1}$$

From Eq. (12-12), with a load designated as $\mathbf{Z}_L$,

$$\frac{\mathbf{V}_2}{\mathbf{I}_2} = -\mathbf{Z}_L = \mathbf{z}_{21}\frac{\mathbf{I}_1}{\mathbf{I}_2} + \mathbf{z}_{22}$$

from which

$$\frac{\mathbf{I}_2}{\mathbf{I}_1} = \frac{-\mathbf{z}_{21}}{\mathbf{z}_{22} + \mathbf{Z}_L} \tag{12-21}$$

so that the input impedance of a four-terminal general network is

$$\mathbf{Z}_{11} = \mathbf{z}_{11} - \frac{\mathbf{z}_{12}\mathbf{z}_{21}}{\mathbf{z}_{22} + \mathbf{Z}_L} = \frac{\Delta + \mathbf{z}_{11}\mathbf{Z}_L}{\mathbf{z}_{22} + \mathbf{Z}_L} \tag{12-22}$$

The output impedance may be found looking back into the 2, 2 terminals as

$$\frac{\mathbf{V}_2}{\mathbf{I}_2} = \mathbf{Z}_{22} = \mathbf{z}_{21}\frac{\mathbf{I}_1}{\mathbf{I}_2} + \mathbf{z}_{22} \tag{12-23}$$

From Eq. (12-11), with $\mathbf{Z}_s$ as the source impedance,

$$\frac{\mathbf{V}_1}{\mathbf{I}_1} = -\mathbf{Z}_s = \mathbf{z}_{11} + \mathbf{z}_{12}\frac{\mathbf{I}_2}{\mathbf{I}_1}$$

from which

$$\frac{\mathbf{I}_1}{\mathbf{I}_2} = \frac{-\mathbf{z}_{12}}{\mathbf{z}_{11} + \mathbf{Z}_s}$$

Therefore the output impedance of the network is

$$\mathbf{Z}_{22} = \mathbf{z}_{22} - \frac{\mathbf{z}_{12}\mathbf{z}_{21}}{\mathbf{z}_{11} + \mathbf{Z}_s} = \frac{\Delta + \mathbf{z}_{22}\mathbf{Z}_s}{\mathbf{z}_{11} + \mathbf{Z}_s} \tag{12-24}$$

**12-9. Equivalent Circuit for the Common-base Amplifier.** The common-base circuit of Fig. 12-19 was frequently used in early transistor work. Much circuit analysis has been performed in terms of the common-base transistor parameters developed from that circuit, and most transistors are rated in terms of emitter and collector currents, $i_e$ and $i_c$, as the independent variables. It is convenient to continue this practice, which allows writing

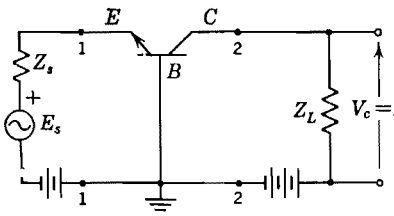



FIG. 12-19. Common-base circuit for *N-P-N* junction transistor.

$$v_e = f(i_e, i_c)$$
$$v_c = f(i_e, i_c)$$

for the transistor characteristics. For small changes of current $\Delta i_e$ and $\Delta i_c$ from chosen quiescent values, the resultant voltage changes may be written as

$$\Delta v_e = \frac{\partial v_e}{\partial i_e}\Delta i_e + \frac{\partial v_e}{\partial i_c}\Delta i_c \tag{12-25}$$

$$\Delta v_c = \frac{\partial v_c}{\partial i_e}\Delta i_e + \frac{\partial v_c}{\partial i_c}\Delta i_c \tag{12-26}$$

If the $\Delta$ changes are small, the transistor characteristics may be assumed linear over the operating range, in which case the partial derivatives, as slopes of the curves, will be constants.

If the $\Delta$ changes are made sinusoidal, then the partial-derivative terms take the dimensions of impedance and Eqs. (12-25) and (12-26) may be written for the transistor as

$$\mathbf{V}_e = \mathbf{z}_{11}\mathbf{I}_e + \mathbf{z}_{12}\mathbf{I}_c \tag{12-27}$$
$$\mathbf{V}_c = \mathbf{z}_{21}\mathbf{I}_e + \mathbf{z}_{22}\mathbf{I}_c \tag{12-28}$$

where the $z$ parameters are defined for the transistor as

$$\mathbf{z}_{11} = \left.\frac{\partial v_e}{\partial i_e}\right]_{i_c = \text{constant}} = \text{input impedance}$$

$$\mathbf{z}_{12} = \left.\frac{\partial v_e}{\partial i_c}\right]_{i_e = \text{constant}} = \text{reverse transfer impedance}$$

$$\mathbf{z}_{21} = \left.\frac{\partial v_c}{\partial i_e}\right]_{i_c = \text{constant}} = \text{forward transfer impedance}$$

$$\mathbf{z}_{22} = \left.\frac{\partial v_c}{\partial i_c}\right]_{i_e = \text{constant}} = \text{output impedance}$$

these impedances being recognizable as the similarly designated impedances of Eq. (12-8). Because of the term-by-term analogy between Eqs. (12-11) and (12-12) for a general network, and Eqs. (12-27) and (12-28) for the transistor, it can be seen that the previously developed four-terminal-network theory may be applied to the transistor. Therefore Eqs. (12-27) and (12-28) define *small-signal linear transistor performance in the common-base circuit.*

At low frequencies the parameters of the transistor become resistive, or $\mathbf{z}_{11} = r_{11}$, $\mathbf{z}_{12} = r_{12}$, $\mathbf{z}_{21} = r_{21}$, $\mathbf{z}_{22} = r_{22}$. These quantities may be measured directly by small signals applied to the transistor in circuits designed for measurement under the conditions specified in the definitions of Eq. (12-8). The parameters may also be interpreted as the slopes of characteristic curve families of which Fig. 12-20 gives two examples,

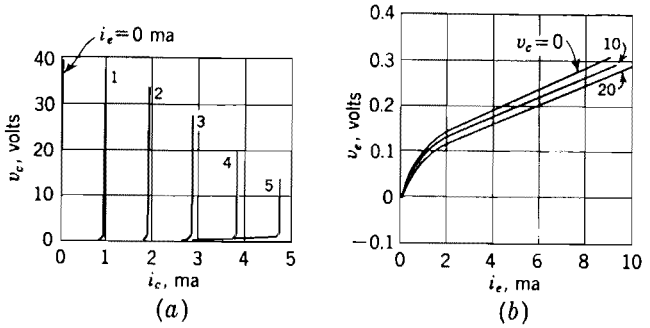



FIG. 12-20. Common-base characteristics. (*a*) Collector family, slope = $r_{22}$; (*b*) emitter family, slope = $r_{11}$.

yielding $r_{11}$ and $r_{22}$. It can be seen that $r_{22}$ is very high for a junction transistor and $r_{11}$ is low and relatively independent of $v_c$. Values for $r_{12}$ and $r_{21}$ may be found from similar curve families for $v_e$ vs. $i_c$ and $v_c$ vs. $i_e$ if desired.

Since general network theory applies, the one-generator equivalent circuits of Fig. 12-18 become available to represent the transistor. Figure 12-18*b*, redrawn as Fig. 12-21*a*, may be more readily adapted to the common-base circuit if a new set of transistor parameters are introduced. Definition of the new parameters $r_e$, $r_b$, $r_c$, and $r_m$ in terms of the fundamental network parameters $r_{11}$, $r_{12}$, $r_{21}$, and $r_{22}$ can be accomplished through writing the general relations of Eqs. (12-27) and (12-28) as

$$\mathbf{V}_e = r_{11}\mathbf{I}_e + r_{12}\mathbf{I}_c \tag{12-29}$$

$$\mathbf{V}_c = r_{21}\mathbf{I}_e + r_{22}\mathbf{I}_c \tag{12-30}$$

Network equations for the circuit of Fig. 12-21*b* can be written directly as

$$\mathbf{V}_e = (r_e + r_b)\mathbf{I}_e + r_b\mathbf{I}_c \tag{12-31}$$

$$\mathbf{V}_c = (r_b + r_m)\mathbf{I}_e + (r_b + r_c)\mathbf{I}_c \tag{12-32}$$

from which a term-by-term comparison leads to

$$r_{11} = r_e + r_b \qquad r_{21} = r_m + r_b$$
$$r_{12} = r_b \qquad r_{22} = r_c + r_b$$

which may also be checked by application of the fundamental definitions of Eq. (12-8). The general network relations then follow as

$$\begin{aligned} r_e &= r_{11} - r_{12}, \text{ the emitter resistance} \\ r_b &= r_{12}, \text{ the base resistance} \\ r_c &= r_{22} - r_{12}, \text{ the collector resistance} \\ r_m &= r_{21} - r_{12}, \text{ the mutual resistance} \end{aligned} \tag{12-33}$$

The equivalent circuits of Fig. 12-21 thus represent the transistor in the common-base circuit. The defined parameters $r_e$, $r_b$, and $r_c$ have

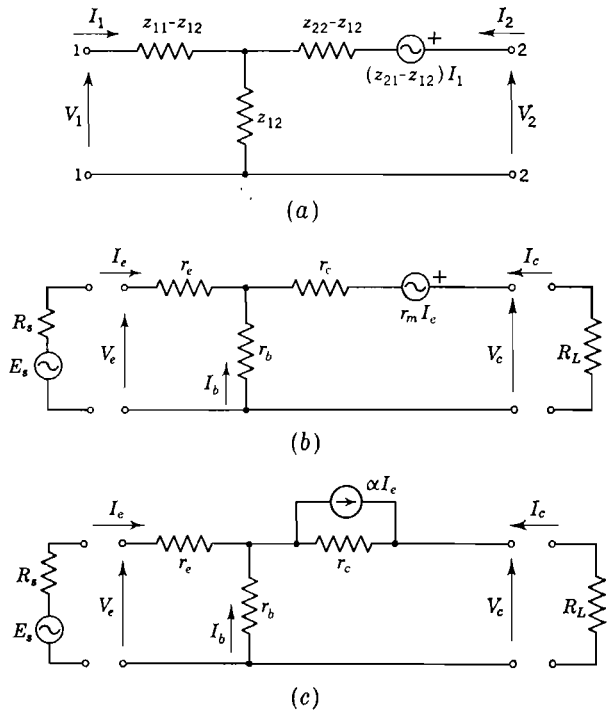



FIG. 12-21. Common-base equivalent circuits derived from Fig. 12-18.

some physical significance in the transistor, but the function of $r_m$ may not be so obvious. It may be noted, however, that the potential transferred into the collector circuit is $(r_b + r_m)\mathbf{I}_e$, and from Eq. (12-32) under

short-circuit conditions, or with $\mathbf{V}_c = 0$,

$$-(r_m + r_b)\mathbf{I}_e = (r_c + r_b)\mathbf{I}_c$$

$$\left|\frac{\mathbf{I}_c}{\mathbf{I}_e}\right| = \alpha_{ce} = \frac{r_m + r_b}{r_c + r_b} = \frac{r_{21}}{r_{22}} \tag{12-34}$$

For the junction transistor $r_b$ is usually negligible compared with either $r_m$ or $r_c$, so that

$$\alpha_{ce} \cong \frac{r_m}{r_c} \tag{12-35}$$

and so $r_m$ may be defined as

$$r_m \cong \alpha_{ce} r_c \tag{12-36}$$

Equation (12-34) allows computation of the collector-emitter short-circuit current-amplification factor $\alpha_{ce}$, a parameter always less than unity for junction transistors.

From Eq. (12-22) the input impedance of a four-terminal network was

$$\mathbf{Z}_{11} = \mathbf{z}_{11} - \frac{\mathbf{z}_{12}\mathbf{z}_{21}}{\mathbf{z}_{22} + \mathbf{Z}_L}$$

Using the defined parameters for the common-base circuit, the *input resistance* for low frequencies is

$$\begin{aligned} R_{11} &= r_e + r_b - \frac{r_b(r_m + r_b)}{r_c + r_b + R_L} \\ &= r_e + \frac{r_b(r_c - r_m + R_L)}{r_c + r_b + R_L} \end{aligned} \tag{12-37}$$

Because of the effects of shunt capacities and the matching of the output impedance $R_{22}$, the load $R_L$ will usually be chosen small with respect to $r_c$ and to $r_c - r_m$. The input resistance of the common-base circuit then reduces to

$$R_{11} \cong r_e + r_b(1 - \alpha_{ce}) \tag{12-38}$$

The *output resistance* for the circuit can be similarly developed from Eq. (12-24),

$$\mathbf{Z}_{22} = \mathbf{z}_{22} - \frac{\mathbf{z}_{12}\mathbf{z}_{21}}{\mathbf{z}_{11} + \mathbf{Z}_s}$$

and use of the parameter definitions leads to

$$R_{22} = r_c - r_b \frac{r_m - r_e - R_s}{r_e + r_b + R_s} \tag{12-39}$$

the output resistance being less than the collector resistance because of the active elements of the circuit, although the second term on the right

may be small with respect to $r_c$, in which case

$$R_{22} \cong r_c \tag{12-40}$$

Proper selection of $R_s$ and $R_L$ will produce impedance matches at the input and output terminals when desired.

The *voltage gain* $A_v$ of the transistor in the common-base circuit is

$$\mathbf{A}_v = \frac{\mathbf{V}_c}{\mathbf{V}_e} = \frac{-R_L\mathbf{I}_c}{R_{11}\mathbf{I}_e}$$

The current ratio is obtainable from Eq. (12-21) as

$$\frac{\mathbf{I}_2}{\mathbf{I}_1} = \frac{\mathbf{I}_c}{\mathbf{I}_e} = \frac{-\mathbf{z}_{21}}{\mathbf{z}_{22} + \mathbf{Z}_L} = \frac{-r_{21}}{r_{22} + R_L} \tag{12-41}$$

After use of Eq. (12-37) and $r_m = \alpha_{ce}r_c$,

$$\mathbf{A}_v = \frac{(r_m + r_b)R_L}{r_e r_b + r_c[r_e + r_b(1 - \alpha_{ce})] + (r_e + r_b)R_L} \tag{12-42}$$

It will usually be found that the first and last terms in the denominator are small. Also $r_b \ll r_m$; so the above reduces approximately to

$$\mathbf{A}_v \cong \frac{\alpha_{ce}R_L}{r_e + r_b(1 - \alpha_{ce})} \tag{12-43}$$

The *current gain* $\mathbf{A}_i$ is obtainable directly from Eq. (12-41) as

$$\mathbf{A}_i = \frac{\mathbf{I}_c}{\mathbf{I}_e} = \frac{-r_{21}}{r_{22} + R_L} = \frac{-(r_b + r_m)}{r_c + r_b + R_L} \tag{12-44}$$

Again under the assumptions of $r_b \ll r_m$, $r_b + R_L \ll r_c$, and by use of $r_m = \alpha_{ce}r_c$,

$$\mathbf{A}_i \cong -\alpha_{ce} \tag{12-45}$$

and is less than unity for the common-base circuit with the junction transistor.

The *power gain* of the circuit is then simply $|A_vA_i|$, and in the approximate form

$$\text{Power gain} \cong \frac{\alpha_{ce}^2R_L}{r_e + r_b(1 - \alpha_{ce})} \tag{12-46}$$

The common-base circuit, being analogous to the grounded-grid vacuum-tube circuit, is a form of feedback amplifier and thus is capable of considerable stability. It operates without phase reversal and with medium power gain, being largely used in impedance transformation to couple a low-impedance source to a high-impedance load.

Typical values for an *N-P-N* junction transistor in the common-base circuit are

$$\alpha_{ce} = 0.97 \qquad r_c = 2 \times 10^6 \text{ ohms}$$
$$r_e = 35 \text{ ohms} \qquad r_m = 1.94 \times 10^6 \text{ ohms}$$
$$r_b = 100 \text{ ohms}$$

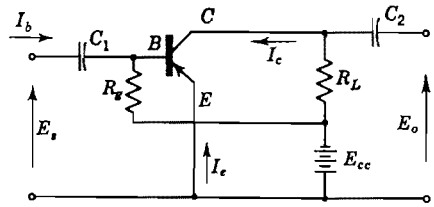
**Example.** Using the above transistor constants, calculate $R_{11}$, $R_{22}$, $A_i$, $A_v$ and the power gain for a common-base circuit with $R_g = 500$ ohms and $R_L = 10{,}000$ ohms.

Using the indicated equations the results may be tabulated as:

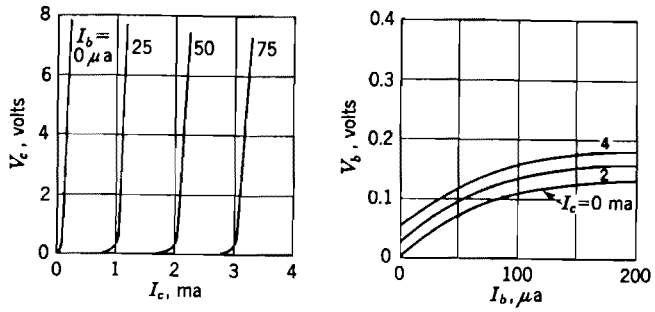
| | |
|---|---|
| $R_{11} = 38.5$ ohms | from Eq. (12-37) |
| $\cong 38$ ohms | approximate Eq. (12-38) |
| $R_{22} = 1.7 \times 10^6$ ohms | from Eq. (12-39) |
| $\cong 2.0 \times 10^6$ ohms | approximate Eq. (12-40) |
| $A_i = -0.965$ | from Eq. (12-44) |
| $\cong -0.97$ | approximate Eq. (12-45) |
| $A_v = 25.1$ | from Eq. (12-42) |
| $\cong 24.6$ | approximate Eq. (12-43) |

Power gain $= 0.965 \times 25.1 = 24.2$
$= 13.8$ db

**12-10. The Common-emitter Circuit.** The common-emitter transistor amplifier is analogous to the common-cathode vacuum-tube amplifier and is a widely used form because of its high gain. For discussion of




(*a*)




(*b*)

FIG. 12-22. (*a*) Typical low-frequency common-emitter amplifier; $R_g = 150K$, $R_L = 10K$, $C_1 = C_2 = 1$ μf. (*b*) Common-emitter characteristics.

the low-frequency form of the circuit as in Fig. 12-22*a*, it is convenient to choose $\mathbf{I}_b$ and $\mathbf{I}_c$ as the independent variables, with characteristic curve families as in Fig. 12-22*b*. The general network relations then may be written as

$$\mathbf{V}_b = r_{11}\mathbf{I}_b + r_{12}\mathbf{I}_c \tag{12-47}$$

$$\mathbf{V}_c = r_{21}\mathbf{I}_b + r_{22}\mathbf{I}_c \tag{12-48}$$

with the parameters redefined for the common-emitter circuit.

The elements of the equivalent circuit for the transistor as derived in Fig. 12-21*b* for the common-base circuit may be rearranged into a form

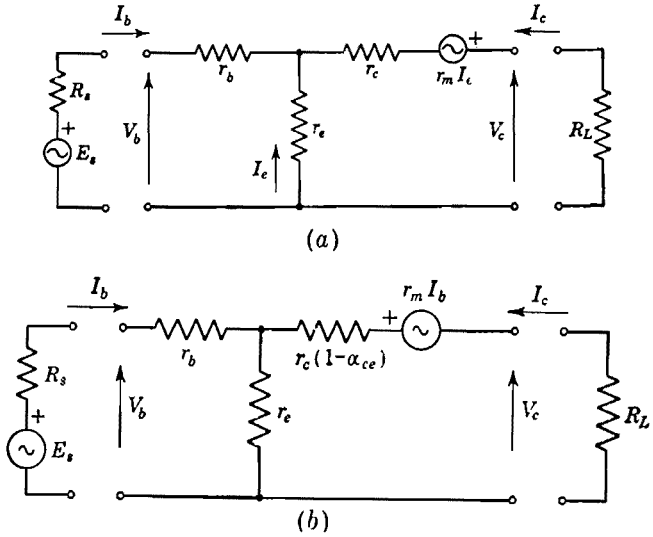



FIG. 12-23. (*a*) Equivalent circuit for the common-emitter amplifier. (*b*) More usual form for the common-emitter equivalent.

suited to the common-emitter amplifier as shown in Fig. 12-23*a*. This follows the familiar method used with the vacuum tube, in which the elements of one equivalent circuit are rearranged for the other two forms of amplifiers. It may be noted that, in so far as the transistor itself is concerned, the only difference between the common-base and common-emitter circuits is the reversal of the polarity of the signal with respect to the emitter and base terminals. This change accounts for the phase reversal inherent in the circuit.

Network equations for the circuit of Fig. 12-23*a* can be written directly, but it must be noted that the equivalent generator is a function of $\mathbf{I}_e$.

Then

$$\mathbf{V}_b = r_b\mathbf{I}_b - r_e\mathbf{I}_e$$

$$\mathbf{V}_c = (r_m - r_e)\mathbf{I}_e + r_c\mathbf{I}_c$$

and

$$\mathbf{I}_e = -(\mathbf{I}_b + \mathbf{I}_c)$$

from which

$$\mathbf{V}_b = (r_b + r_e)\mathbf{I}_b + r_e\mathbf{I}_c \tag{12-49}$$

$$\mathbf{V}_c = (r_e - r_m)\mathbf{I}_b + (r_c + r_e - r_m)\mathbf{I}_c$$
$$= (r_e - r_m)\mathbf{I}_b + [r_e + r_c(1 - \alpha_{ce})]\mathbf{I}_c \tag{12-50}$$

Equation (12-50) shows the mutual-generator term $r_m\mathbf{I}_b$ negative and a function of input current $\mathbf{I}_b$, which is more convenient for this circuit. Equations (12-49) and (12-50) lead to a circuit of the form of Fig. 12-23*b*, and this is the usual form for the common-emitter circuit. It should be noted that it is not a new circuit but was derived by starting with the elements of the common-base amplifier equivalent circuit.

Term-by-term comparison with the general relations of Eqs. (12-47) and (12-48) allows the network parameters to be defined for the general network as

$$r_{11} = r_b + r_e \qquad r_{21} = r_e - r_m$$
$$r_{12} = r_e \qquad r_{22} = r_e + r_c(1 - \alpha_{ce})$$

While these general network definitions differ from those for the common-base circuit, the transistor parameters $r_e$, $r_b$, $r_c$, and $r_m$ are identical with those for the common-base circuit.

The *input resistance* $R_{11}$ for the common-emitter circuit can be found from the general relation [Eq. (12-22)],

$$R_{11} = r_b + \frac{r_e(r_c + R_L)}{r_c + r_e - r_m + R_L} \tag{12-51}$$

In practical circuits it can usually be assumed that $r_e \ll R_L \ll r_c(1 - \alpha_{ce})$; so

$$R_{11} \cong r_b + \frac{r_e}{1 - \alpha_{ce}} \tag{12-52}$$

Similarly, with Eq. (12-24) the *output resistance* is

$$R_{22} = r_c(1 - \alpha_{ce}) + \frac{r_e(r_b + r_m + R_s)}{r_b + r_e + R_s} \tag{12-53}$$

$$\cong r_c(1 - \alpha_{ce}) + \frac{r_e r_m}{r_b + r_e + R_s} \tag{12-54}$$

The *current gain* $\mathbf{A}_i$ of the circuit is obtainable from Eq. (12-21) as

$$\mathbf{A}_i = \frac{\mathbf{I}_2}{\mathbf{I}_1} = \frac{\mathbf{I}_c}{\mathbf{I}_b} = \frac{-r_{21}}{r_{22} + R_L}$$

from which

$$\mathbf{A}_i = \frac{-(r_e - r_m)}{r_e + r_c - r_m + R_L} = \frac{1}{1 + (r_c + R_L)/(r_e - r_m)} \tag{12-55}$$

For $r_e \ll R_L \ll r_c$

$$\mathbf{A}_i \cong \frac{\alpha_{ce}}{1 - \alpha_{ce}} \tag{12-56}$$

With values of $\alpha_{ce}$ in the range of 0.95 to 0.99 the current gain of the circuit can be quite high. Actually this is due to the small magnitude of the base current in comparison with the current flowing between emitter and collector.

The *voltage gain* of the common-emitter circuit can be written as

$$\mathbf{A}_v = \frac{\mathbf{V}_c}{\mathbf{V}_b} = \frac{-R_L I_c}{R_{11} I_b} = \frac{R_L}{r_b + [(\overline{r_b + r_e})(r_c + R_L)/(\overline{r_e - r_m})]} \tag{12-57}$$

Assuming the usual situation, $r_e \ll r_m$, $R_L \ll r_c$, then

$$\mathbf{A}_v \cong \frac{-\alpha_{ce} R_L}{r_e + r_b(1 - \alpha_{ce})} \tag{12-58}$$

the negative sign indicating the phase reversal. High gain is possible with large $R_L$ values, but frequency range suffers.

The *power gain* then is $|A_v A_i|$, and, using the approximate forms,

$$\text{Power gain} = \frac{\alpha_{cc}^2 R_L}{(1 - \alpha_{cc})r_e + r_c(1 - \alpha_{ce})^2} \tag{12-59}$$

**Example.** Using the values given for the *N-P-N* junction transistor of Sec. 12-9, compute the values of $R_{11}$, $R_{22}$, $A_i$, $A_v$ and the power gain in the common-emitter circuit, with $R_s = 500$ ohms and $R_L = 10{,}000$ ohms.

The results may be tabulated as

| | |
|---|---|
| $R_{11} = 1105$ ohms | from Eq. (12-51) |
| $\cong 1266$ ohms | approximate Eq. (12-52) |
| $R_{22} = 167{,}000$ ohms | from Eq. (12-53) |
| $\cong 167{,}000$ ohms | approximate Eq. (12-54) |
| $A_i = 28.6$ | from Eq. (12-55) |
| $\cong 32.3$ | approximate Eq. (12-56) |
| $A_v = -250$ | from Eq. (12-57) |
| $\cong -255$ | approximate Eq. (12-58) |
| Power gain $= 28.6 \times 250 = 7150$ | from Eq. (12-59) |
| $= 38.5$ db | |

**12-11. The Common-collector Amplifier.** The common-collector amplifier has a high input impedance, low output impedance, less than unity voltage gain, and no phase reversal, as does its analog, the vacuum-tube cathode-follower circuit.

For the basic circuit of Fig. 12-24*a* it is possible to rearrange the elements of the transistor equivalent to form the equivalent circuit of Fig. 12-24*b*, keeping all current directions and elements consistent with usage in the common-base equivalent. The independent variables are $\mathbf{I}_b$ and $\mathbf{I}_e$, and the basic four-terminal network relations at low frequency are

$$\mathbf{V}_b = r_{11}\mathbf{I}_b + r_{12}\mathbf{I}_e \tag{12-60}$$

$$\mathbf{V}_e = r_{21}\mathbf{I}_b + r_{22}\mathbf{I}_e \tag{12-61}$$

Network equations may be written by inspection of the equivalent circuit in Fig. 12-24*b*, using the branch currents as

$$\mathbf{V}_b = r_b\mathbf{I}_b - r_c\mathbf{I}_c - r_m\mathbf{I}_e$$
$$\mathbf{V}_c = r_e\mathbf{I}_e - r_c\mathbf{I}_c - r_m\mathbf{I}_e$$

A more convenient form of equivalent circuit may be obtained if $r_m\mathbf{I}_b$ is added and subtracted from each of the above equations. After it is noted

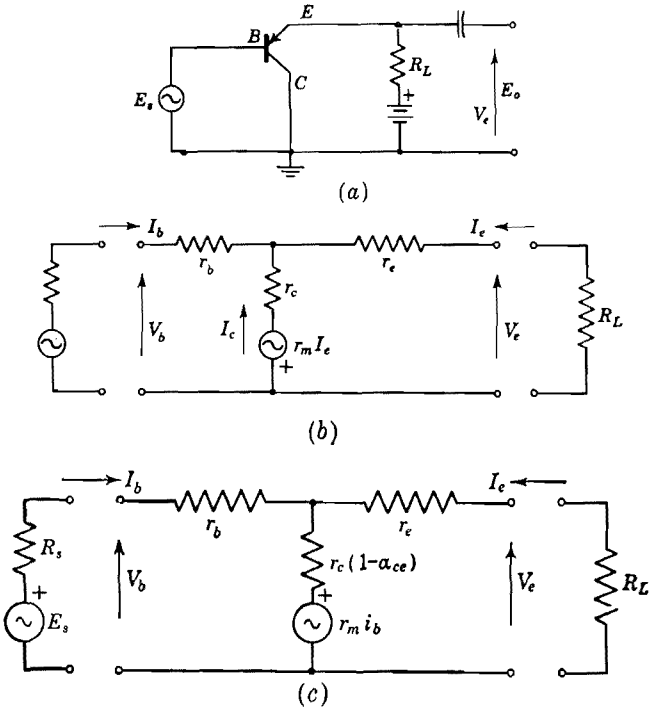



FIG. 12-24. (*a*) Common-collector amplifier. (*b*) Equivalent circuit for (*a*). (*c*) More usual form for (*b*).

that $\mathbf{I}_c = -(\mathbf{I}_b + \mathbf{I}_e)$, the equations reduce to

$$\begin{aligned}\mathbf{V}_b &= (r_b + r_c - r_m)\mathbf{I}_b + r_m\mathbf{I}_b + (r_c - r_m)\mathbf{I}_e \\ &= [r_b + r_c(1 - \alpha_{cc})]\mathbf{I}_b + r_m\mathbf{I}_b + r_c(1 - \alpha_{cc})\mathbf{I}_c \qquad (12\text{-}62)\end{aligned}$$
$$\begin{aligned}\mathbf{V}_e &= (r_c - r_m)\mathbf{I}_b + r_m\mathbf{I}_b + (r_e + r_c - r_m)\mathbf{I}_e \\ &= r_c(1 - \alpha_{ce})\mathbf{I}_b + r_m\mathbf{I}_b + [r_e + r_c(1 - \alpha_{ce})]\mathbf{I}_e \qquad (12\text{-}63)\end{aligned}$$

Equations (12-62) and (12-63) lead to the equivalent circuit of Fig. 12-24*c*, which is the usual form for the common-collector circuit. Note that again only the common-base circuit was used as a starting point, rearranged to suit the actual circuit configuration, and that the new

equivalent was then derived through algebraic manipulation of the circuit equations. Thus the procedure is analogous to that employed with the vacuum triode, wherein a single equivalent circuit is rearranged for use in the three basic circuit forms. The result is more complicated with the transistor and leads to three different equivalent circuits for convenience, but all employ the same transistor parameters. The additional complication is a result of the fact that the transistor is a more elaborate device with its finite input impedance.

Again, term-by-term comparison with the general network relations of Eq. (12-60) and (12-61) reveals that the general network parameters may be defined for the common-collector circuit as

$$r_{11} = r_b + r_c \qquad r_{21} = r_c$$
$$r_{12} = r_c(1 - \alpha_{ce}) \qquad r_{22} = r_e + r_c(1 - \alpha_{ce})$$

These same parameters might be obtained through use of the definitions of Eq. (12-9) and Eqs. (12-62) and (12-63). As before, the transistor parameters are as employed in the common-base and common-emitter circuits.

Performance of the transistor in the common-collector circuit may be obtained from the general network relations [Eqs. (12-21) to (12-24)], giving the following results for the common-collector circuit:

Input resistance:

$$R_{11} = r_b + \frac{r_c(r_e + R_L)}{r_e + r_c(1 - \alpha_{ce}) + R_L} \tag{12-64}$$

If $r_b \ll r_c$ and $r_e \ll R_L \ll r_c(1 - \alpha_{ce})$, the above reduces to

$$R_{11} \cong \frac{R_L}{1 - \alpha_{ce}} \tag{12-65}$$

For the junction transistor with $\alpha_{ce}$ less than but near unity the input resistance can be quite high. However, for $\alpha_{ce}$ near unity the approximation that $R_L \ll r_c(1 - \alpha_{ce})$ may not be fulfilled. In that case Eq. (12-64) must be used.

Output resistance:

$$R_{22} = r_e + \frac{r_c(1 - \alpha_{ce})(r_b + R_s)}{r_b + r_c + R_s} \tag{12-66}$$

For $r_c$ large

$$R_{22} \cong r_e + (1 - \alpha_{ce})(r_b + R_s) \tag{12-67}$$

This can be seen as low in value, not greatly exceeding $r_e$.

Current gain $\mathbf{A}_i$:

$$\mathbf{A}_i = \frac{-r_{21}}{r_{22} + R_L} = \frac{-r_c}{r_e + r_c(1 - \alpha_{ce}) + R_L} = \frac{-1}{1 + r_e - \dfrac{r_m + R_L}{r_c}} \tag{12-68}$$

For $r_e$ and $R_L$ small with respect to $r_m$, then

$$\mathbf{A}_i \cong \frac{-1}{1 - \alpha_{ce}} \tag{12-69}$$

Voltage gain $A_v$:

$$\mathbf{A}_v = \frac{\mathbf{V}_e}{\mathbf{V}_b} = \frac{-R_L I_e}{R_{11} I_b} = \frac{r_c R_L}{r_b(r_c - r_m + r_e + R_L) + r_c(r_e + R_L)} \tag{12-70}$$

Making the usual assumptions regarding $r_b$ and $r_e$ with respect to $r_c$ and $R_L$, the above reduces to

$$\mathbf{A}_v \cong \frac{1}{1 + [r_b(1 - \alpha_{ce})/R_L]} \tag{12-71}$$

which reduces to unity in the limit. The positive sign indicates no phase reversal.

Power gain:

$$A_v A_i = \frac{1}{(1 - \alpha_{ce})\left[1 + \dfrac{r_b(1 - \alpha_{ce})}{R_L}\right]} \cong \frac{1}{1 - \alpha_{ce}} \tag{12-72}$$

by use of the approximate relations.

**Example.** For the *N-P-N* junction-transistor values of Sec. 12-9 compute $R_{11}$, $R_{22}$, $A_i$, $A_v$ and the power gain when used in the common-collector circuit, with $R_s$ = 500 ohms and $R_L$ = 10,000 ohms. The results may be tabulated as

| | |
|---|---|
| $R_{11}$ = 328,000 ohms | from Eq. (12-64) |
| $\cong$ 333,000 ohms | approximate Eq. (12-65) |
| $R_{22}$ = 53 ohms | from Eq. (12-66) |
| $\cong$ 53 ohms | approximate Eq. (12-67) |
| $A_i$ = −29.4 | from Eq. (12-68) |
| $\cong$ −33.3 | approximate Eq. (12-69) |
| $A_v$ = 0.9996 | from Eq. (12-70) |
| $\cong$ 0.9997 | approximate Eq. (12-71) |
| Power gain = 29.39 | from Eq. (12-72) |
| = 14.7 db | |

**12-12. Summary of Transistor Amplifier Characteristics.** The common-base amplifier is characterized by low input and high output impedance and no phase inversion. It provides voltage and power gain and is suited to d-c-coupled circuits and for impedance matching.

The common-emitter circuit is the most flexible and generally used of the group. Its input resistance is higher than that of the common-base circuit, and it has a lower output resistance. It provides maximum voltage and power gain for a given transistor. For low values of $R_L$, less than $r_c(1 - \alpha_{ce})$, the input resistance of both common-base and common-emitter circuits is low and independent of load value. For intermediate values of $R_L$, approaching $r_c$, the input resistance approximates $r_b + r_e$,

as shown in Fig. 12-25. These variations are features which make the transistor more flexible than the vacuum tube but which complicate circuit design for the transistor.

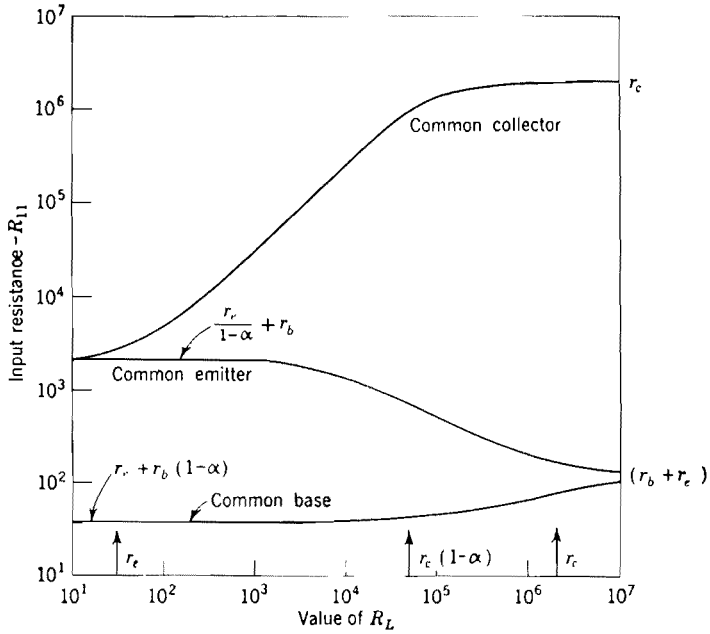



FIG. 12-25. Variation of $R_{11}$ with $R_L$ for the transistor of the several examples.

The common-collector circuit has a voltage gain approaching but less than unity. It has a low output resistance, a high input resistance which

TABLE 26. SUMMARY OF APPROXIMATE RELATIONSHIPS*

| Parameter | Common-base | Common-emitter | Common-collector |
|---|---|---|---|
| $R_{11}$. . . . | $r_e + r_b(1 - \alpha_{ce})$ | $\frac{r_e + r_b(1 - \alpha_{ce})}{1 - \alpha_{ce}}$ | $\frac{R_L}{1 - \alpha_{ce}}$ |
| $R_{22}$. . . . | $r_c$ | $r_c(1 - \alpha_{ce}) + \frac{r_e r_m}{r_b + r_e + R_s}$ | $r_e + (1 - \alpha_{ce})(r_b + R_s)$ |
| $A_i$. . . . | $-\alpha_{ce}$ | $\frac{\alpha_{ce}}{1 - \alpha_{ce}}$ | $\frac{-1}{1 - \alpha_{ce}}$ |
| $A_v$. . . . | $\frac{\alpha_{ce} R_L}{r_e + r_b(1 - \alpha_{ce})}$ | $\frac{-\alpha_{ce} R_L}{r_e + r_b(1 - \alpha_{ce})}$ | 1 |
| Power gain | $\frac{\alpha_{ce}^2 R_L}{r_e + r_b(1 - \alpha_{ce})}$ | $\frac{\alpha_{ce}^2 R_L}{(1 - \alpha_{ce})[r_e + r_b(1 - \alpha_{ce})]}$ | $\frac{1}{1 - \alpha_{ce}}$ |

* It is assumed that $r_e + r_b \ll R_L \ll r_c(1 - \alpha_{ce})$.

is a function of $R_L$, and no phase reversal. It is used largely as an impedance-matching or buffer amplifier.

Power-handling capabilities of transistors are limited by allowable temperature rise and high ambient temperatures, since intrinsic conduction of the semiconductor enters. Maximum collector voltage is a function of the Zener voltage of the material. While power output is limited by temperature rise, yet reasonable power output can be obtained because of the high internal efficiency of the transistor, reaching 49 per cent in Class A service.

Minimum voltage and current limits are much lower than for the vacuum tube, as can be discovered by noting the linearity of the collector family of curves at extremely low voltages.

TABLE 27. USUAL PERFORMANCE RANGES FOR JUNCTION TRANSISTORS

| Circuit | $R_{11}$, ohms | $R_{22}$, ohms | Power gain, db | Phase reversal |
|---|---|---|---|---|
| Common-base | 30–1000 | 100,000–2,000,000 | 15–30 | No |
| Common-emitter | 300–1500 | 5,000–100,000 | 30–40 | Yes |
| Common-collector | 100,000–500,000 | 100–20,000 | 12–16 | No |

**12-13. Graphical Analysis of Large-signal Transistor Performance.** Graphical analysis of transistor performance with large signals follows directly from vacuum-tube methods and so will be only briefly treated here. Transistors are usually rated in terms of maximum collector current and voltage, maximum emitter or base current, and allowable power dissipation at a specified ambient temperature. A maximum or limiting dissipation line may be drawn on a set of collector characteristics as in Fig. 12-26, by computing points from the hyperbolic relation

$$E_c I_c = \text{rated dissipation, watts}$$

The point of operation should not rise above this limiting curve at any instant in the cycle. This is especially important with reactive loads, where the load line becomes an ellipse. In this case, even though the $Q$ point is below the maximum dissipation hyperbola, a portion of the ellipse may rise above that critical curve. The dissipation will then momentarily rise above the rated value and may lead to transistor damage.

For the figure it may be assumed that the rated dissipation is 50 mw, the collector-supply voltage $E_{cc} = 25$ volts, and maximum base current $I_b = 250$ $\mu$amp. Using the rated $E_{cc}$ value, a load line may be drawn tangent to the maximum dissipation line, giving a slope corresponding to a load $R_L = 3200$ ohms. The swing in $I_b$ for low distortion may be 0 to

200 $\mu$amp, with the $Q$ point then located at $I_b = 100$ $\mu$amp. For usual junction transistors the curves are linear to low collector voltages, and operation is possible down to about 1 to 2 volts. Transistors with power-output ratings of 65 to 100 watts are being developed, these ratings assuming excellent heat dissipation or forced cooling.

Performance calculations follow directly from the load line as for the vacuum tube. Transformer or choke coupling would be handled by drawing the d-c load line before placing the $Q$ point and the a-c load line.

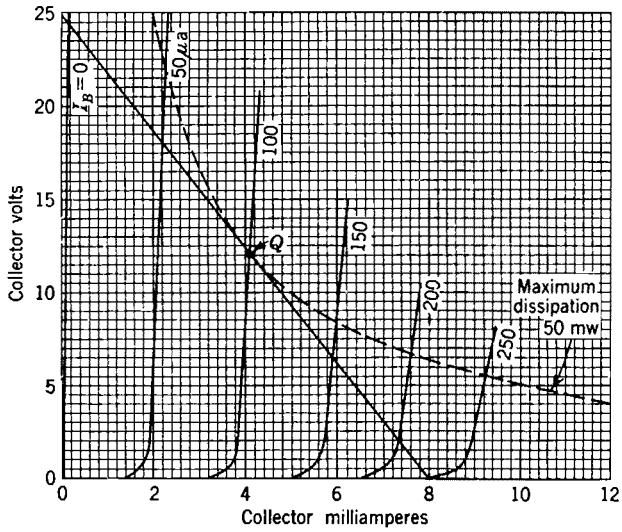



FIG. 12-26. Graphical solution of transistor operating conditions.

**12-14. Bias and Stabilization.** The $Q$ point for a transistor is fixed by several currents, and for proper maintenance of the $Q$ point these currents should be constant. However, most supply sources are of constant-voltage form, and thus maintenance of constant-current supply is not inherent, although use of a large series resistor with a high-voltage supply provides an approximation. In addition transistors differ one from another, and use of a replacement unit may cause different currents to be drawn from constant-voltage supplies and shift of the $Q$ point into an undesirable operating region.

Another of the major transistor-operating problems is the variation of transistor parameters with temperature. The saturation current $I_{co}$ which flows with zero emitter current is particularly sensitive to temperature, increasing exponentially with rising semiconductor temperature. An increase of $I_{co}$ further raises the dissipation and the temperature, again raising $I_{co}$, transistor burnout being the possible ultimate result.

Fixed current-bias operation is possible in the circuit of Fig. 12-27*a*, in which $I_B$ may be set by making

$$R_1 = \frac{E_{CC}}{I_B} \tag{12-73}$$

this assuming that the emitter to base resistance $r_e + r_b$ is small with respect to $R_1$, which is usually the case if $E_{cc}$ is large. If $I_B = 400\ \mu$amp at the desired $Q$ point, $E_{CC} = 20$ volts, then $R_1$ becomes 50,000 ohms.

The circuit of Fig. 12-27*a*, while providing all voltages from one source, is very sensitive to change in $I_{co}$ due to temperature. To stabilize the

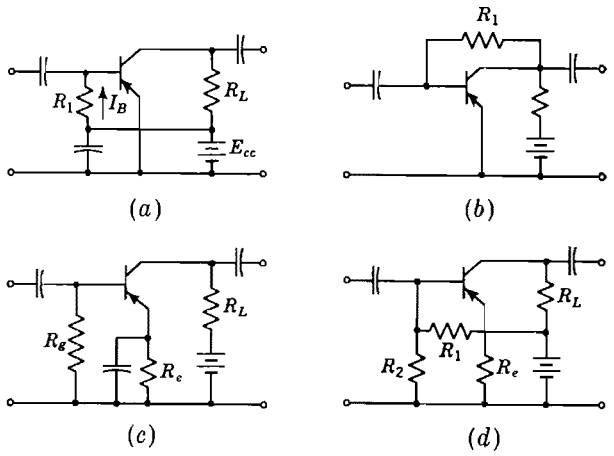



FIG. 12-27. Bias circuits.

operating point, the circuit needs degeneration at zero frequency, such as is provided by a cathode resistor in a vacuum-tube circuit.

The circuit of Fig. 12-27*b* provides negative feedback from collector to base for both bias current and signal. The feedback resistor $R_1$ should have a value

$$R_1 = \frac{V_C - V_B}{I_B}$$

where $V_C$, $V_B$, and $I_B$ are values at the desired $Q$ point. If the collector voltage tends to be low, the base bias current will be reduced, tending to restore the desired collector voltage. A high collector voltage will cause high base current, causing the $Q$ point to move down the load line toward the desired location.

The circuit of Fig. 12-27*c* grounds the emitter only for alternating current, giving negative feedback and stabilization of the $Q$ point for d-c current values. Use of 100 to 500 ohms will usually be sufficient for $R_e$, and this will not detract from a-c gain if suitably bypassed for the operating frequency range.

A combination of fixed current bias and an unbypassed emitter resistor $R_e$ appears in Fig. 12-27*d*. The value of $R_1$ may be chosen to give the desired base current. The effect of $R_e$ is to introduce negative feedback in both d-c and a-c operation and to hold $I_C$ constant with variations of temperature or $\alpha_{ce}$. In circuit analysis the $R_e$ value may be added to $r_e$ of the transistor in the performance equations. From Table 26 it can

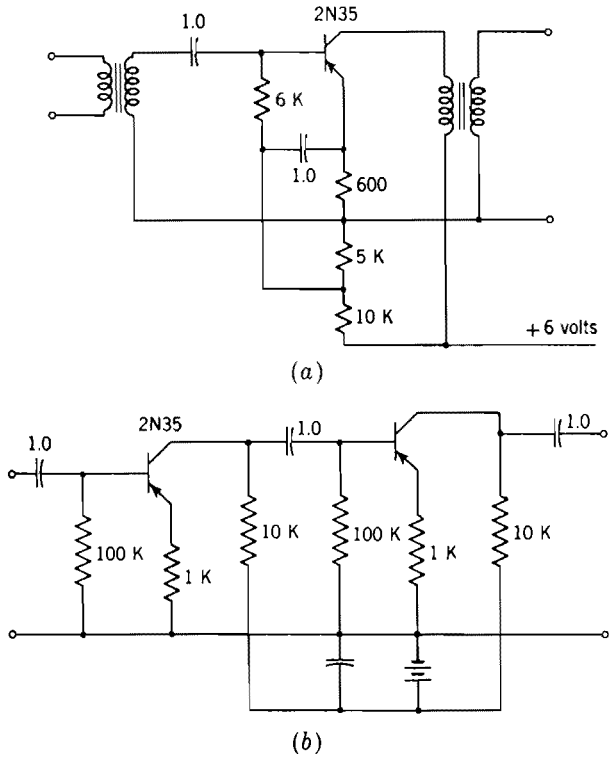



FIG. 12-28. (*a*) Transformer-coupled amplifier. (*b*) *RC*-coupled for audio frequencies.

be seen that increase of emitter resistance will reduce $A_v$ and the power gain. Another important effect is to increase the input resistance in both common-base and common-emitter circuits. In the latter circuit, for the transistor used in the example of Sec. 12-9, a resistor $R_e$ of 300 ohms increases the input resistance from 1105 to 11,300 ohms. The voltage gain is reduced from $-255$ to $-29$; yet the improved load condition for the preceding stage may make the addition of $R_e$ worthwhile, in addition to the stabilization of the $Q$ point.

**12-15. Cascaded Amplifiers.** The common-emitter circuit is frequently employed in cascaded amplifiers as in Fig. 12-28. In Fig. 12-28*a*

transformer coupling is used to permit impedance matching between output and input of succeeding stages, and low values of supply voltage may be sufficient for all except power stages. Emitter-resistance stabilization of the $Q$ point is employed, bypassed for alternating current.

Figure 12-28*b* illustrates a cascaded $RC$-coupled common-emitter amplifier in which impedance mismatch is tolerated because of the small physical size possible without transformers. Degeneration supplied by the 1000-ohm emitter resistors raises the input resistance for a better match to the output impedance and increases stability markedly. Power supply may be 20 to 25 volts, and the constants shown will permit satisfactory coverage of the audio-frequency range.

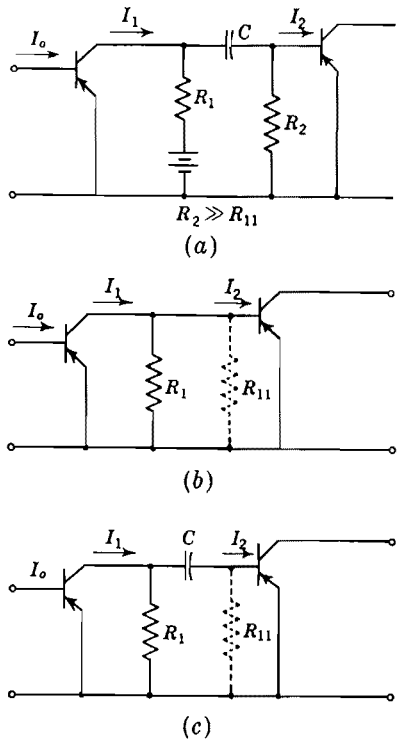



FIG. 12-29. (*a*) $RC$ coupling. (*b*) Medium-frequency equivalent. (*c*) Low-frequency equivalent.

**12-16. Frequency Response of the $RC$ Amplifier.** For analysis of the response of the $RC$-coupled amplifier in the middle- and low-frequency ranges, one stage is redrawn in Fig. 12-29*a*. With $r_e \ll r_m$, Eq. (12-55) for the current gain of the common-emitter amplifier may be modified to give the output current $\mathbf{I}_1$ as

$$\mathbf{I}_1 \cong \mathbf{I}_o \frac{\alpha_{ce} r_c}{r_c(1 - \alpha_{ce}) + R_L} \tag{12-74}$$

where $R_L$ is the complete a-c load of the transistor. If $R_2 \gg R_{11}$, where $R_{11}$ is the input resistance of the second stage as shown in Fig. 12-29*b*, then

$$R_L = \frac{R_1 R_{11}}{R_1 + R_{11}}$$

in the mid-frequency range, where the reactance of $C$ may be neglected with respect to $R_{11}$.

The over-all current gain $I_2/I_o$ in the mid-frequency range then can be computed as

Mid-frequency: $$\mathbf{A}_i = \frac{\mathbf{I}_2}{\mathbf{I}_o} \cong \frac{\alpha_{ce} r_c}{r_c(1 - \alpha_{ce}) + R_L} \frac{R_1}{R_1 + R_{11}} \tag{12-75}$$

The term $R_1/(R_1 + R_{11})$ is present as a current-division factor, the output current $\mathbf{I}_1$ dividing between $R_1$ and $R_{11}$ of the second transistor.

The result may be further reduced as

$$\mathbf{A}_i \cong \frac{\alpha_{ce}}{1 - \alpha_{ce}} \frac{1}{1 + \dfrac{R_L}{r_c(1 - \alpha_{ce})}} \frac{R_1}{R_1 + R_{11}} \tag{12-76}$$

which brings out the short-circuit current-gain term $\alpha_{ce}/(1 - \alpha_{ce}) = \alpha_{cb}$ for the common-emitter circuit.

At low frequencies the reactance of $C$ becomes large with respect to $R_{11}$, and the over-all current gain then may be written as

Low-frequency:

$$\mathbf{A}_i \cong \frac{\alpha_{ce} r_c}{r_c(1 - \alpha_{ce}) + \dfrac{R_1}{1 + j\omega C R_1}} \frac{R_1}{R_1 + \dfrac{1}{j\omega C}} \tag{12-77}$$

from the low-frequency circuit of Fig. 12-20*c*. This equation may be reduced to

$$\mathbf{A}_i = \frac{\alpha_{ce}}{1 - \alpha_{ce}} \frac{1}{1 - \dfrac{j}{\omega C}\left[\dfrac{1}{r_c(1 - \alpha_{ce})} + \dfrac{1}{R_1}\right]} \tag{12-78}$$

which shows the effect of the reactance of $C$.

It is apparent that at a corner frequency

$$f_1 = \frac{1}{2\pi C}\left[\frac{1}{r_c(1 - \alpha_{ce})} + \frac{1}{R_1}\right] \tag{12-79}$$

the current gain will be reduced 3 db and the phase angle will be +45°, with the gain falling at the rate of 6 db per octave at lower frequencies.

The current gain at low frequencies may then be written as

$$\mathbf{A}_i = \frac{\alpha_{ce}}{1 - \alpha_{ce}} \frac{1}{1 - jf_1/f} \tag{12-80}$$

and the phase angle is

$$\phi = \tan^{-1} \frac{f_1}{f} \tag{12-81}$$

These expressions are analogous to low-frequency equations for the $RC$-coupled vacuum-tube amplifier.

The response of the transistor at high frequencies is complicated by internal reactances and diffusion times, and an exact equivalent circuit

is not easily drawn. First approximations usually assume capacitive reactances in shunt with $r_e$ and $r_c$ of the low-frequency equivalent circuit. Since the emitter capacitance $C_e$ is in shunt with the low value $r_e$, $C_e$ can usually be neglected even though it may reach values of 100 to 150 $\mu\mu$f. Collector capacitance $C_c$, being in shunt with the high value of $r_c$, can be very effective in reducing high-frequency response and output impedance and in some circuits may be effective down to 1 kc.

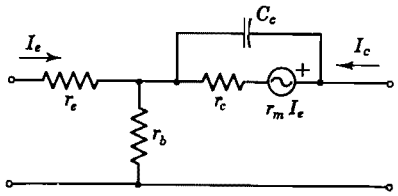



FIG. 12-30. High-frequency equivalent circuit of simple form for the common-base circuit.

An approximate high-frequency equivalent circuit for the common-base form appears in Fig. 12-30, it being possible to derive equivalents for the other amplifier forms therefrom. It should be also noted that both $r_c$ and $C_c$ may be functions of frequency, falling with increasing frequency.

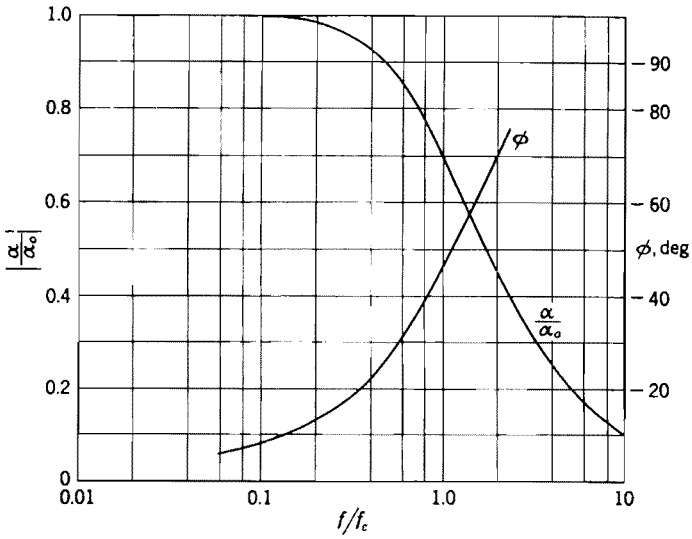



FIG. 12-31. Variation of $\alpha$ with frequency.

It has also been found that, owing to hole and electron diffusion in the crystal, the magnitude of $\alpha_{ce}$ is a function of frequency and at high frequencies there may be a phase angle in $\alpha_{ce}$ as well.

Because of the complex nature of the internal actions in a transistor at high frequencies it is customary to define a corner frequency for the $\alpha_{ce}$ vs. frequency curve in terms of a cutoff frequency $f_c$ at which the magnitude of $\alpha_{ce}$ is 70.7 per cent, or 3 db below its low-frequency value $\alpha_o$.

Under this definition it is found that over a reasonable frequency range $\alpha_{ce}$ follows the relation

$$\frac{\alpha_{ce}}{\alpha_o} = \frac{1}{1 + jf/f_c} \tag{12-82}$$

fairly well.

The variation of $\alpha_{re}$ as predicted by Eq. (12-82) is plotted in Fig. 12-31. The specification of $f_c$ for a transistor allows a reasonable estimate of its possible high-frequency performance, and junction units with $f_c$ values up to several megacycles are available. The performance of the associated circuit as a function of frequency may be calculated and introduced in decibel response by direct addition to or subtraction from the decibel response curve plotted, for the transistor alone, by use of $f_c$ as a corner frequency.

For greater accuracy in prediction of characteristics at high frequencies it is necessary to take account of the distributed nature of the charge carriers and their paths in the transistor and to utilize equations based on distributed resistance and capacity within the transistor, as well as considering the inherent mobilities of the carriers.

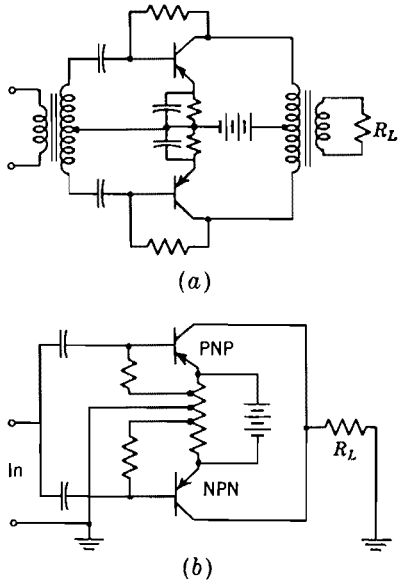



FIG. 12-32. (*a*) Push-pull Class A or B circuit, common-emitter form. (*b*) Complementary symmetry Class B amplifier.

**12-17. Push-Pull and Complementary Symmetry Circuits.** For power-output stages it is possible to employ transistors in circuits approximating conventional push-pull forms under either Class A or Class B conditions. In such operation Class A efficiencies of 49 per cent and Class B efficiencies of 75 per cent are obtainable.

The limit on power output is fixed by the dissipation and the rated $V_{c,max}$ on the collector. The maximum collector voltage is limited by the Zener voltage of the material, or the voltage at which electric-field breakdown occurs in the reverse or high-resistivity direction. High-voltage units must use high-resistivity or high-purity material, and the high-frequency response may be poor. Also at high voltages it is necessary to prevent overheating of the material; otherwise thermal instability and burnout may follow. Operational analysis follows the graphical methods developed for the vacuum tube.

Push-pull circuits using transistors of opposite conductivity types (*N-P-N*, *P-N-P*) are particularly useful for Class B operation, since a signal of one polarity will simultaneously drive one base into the operating region and at the same time drive the base of the opposite unit beyond cutoff. Thus the input signals can be in phase, and no phase inverter is required. An example of the circuit is shown in Fig. 12-32*b*. The output circuit receives the signal from the two transistors in parallel. Advantages of complementary symmetry include elimination of input transformers and cancellation of collector-current variations. The output transformer may also be eliminated if load matching is otherwise achieved.

It is necessary that changes in collector current be identical for both units, and there is sometimes difficulty in finding suitable matched pairs.

**12-18. Instability in Point-contact Transistor Circuits.** A set of typical point-contact transistor characteristics appears in Fig. 12-33. They

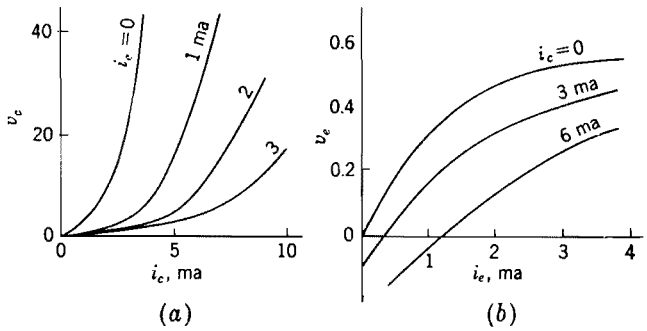



FIG. 12-33. Point-contact transistor characteristics.

deviate from linearity much more than do curves for junction units, and it is found that $r_{22}$ as measured by the slope of the curves in Fig. 12-33*a* is much lower than for junction transistors. Another very important difference which has been discussed is that $\alpha_{ce}$ is greater than unity. It can immediately be seen that the term

$$r_c - r_m = r_c(1 - \alpha_{ce})$$

which appeared frequently in previously developed transistor relations, will be negative, and this will lead to some negative parameters. The equivalent circuit, however, will not differ from that for the junction transistor and is drawn in Fig. 12-34 for the common-base circuit.

From Eq. (12-38), the input resistance of the common-base circuit,

$$R_{11} = r_e + r_b(1 - \alpha_{ce})$$

it may be seen that for sufficiently large values of $r_b$, and with $\alpha_{ce} > 1$, the input resistance will be negative. For instance, with $r_e = 35$ ohms,

$r_b$ = 100 ohms, and $\alpha_{ce}$ = 2.5, the input resistance will be −115 ohms. The negative value is a result of positive feedback provided by coupling through the base resistance and, if the feedback is sufficient, can lead to current instability. In this case the emitter current increases spontaneously, allowing the collector current to increase until limited by circuit resistances or by transistor damage. Care must be taken to limit the external resistance in the base circuit to avoid this dangerous characteristic of short-circuit instability in point-contact transistors.

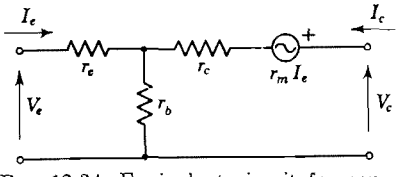



FIG. 12-34. Equivalent circuit for common-base amplifier.

Because the common-base circuit best lends itself to keeping base resistance low, it is generally employed when point-contact transistors are used for amplification. Analysis for performance follows methods previously developed. Other circuits for oscillators or switches capitalize on the current instability and may employ common-emitter or common-collector configurations with intentionally increased base resistance.

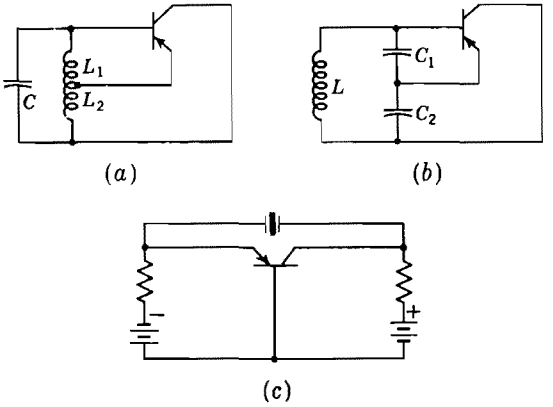



FIG. 12-35. Basic transistor feedback oscillators. (*a*) Hartley; (*b*) Colpitts; (*c*) crystal.

**12-19. Transistor Feedback Oscillators.** As with any amplifying device the transistor can be made to oscillate by meeting necessary frequency and feedback conditions. In appearance the usual transistor feedback-oscillator circuits of Fig. 12-35 look much like their vacuum-tube counterparts. Circuit loading is more severe because of the low input and output resistances of transistors. Higher $C/L$ ratios in the tuned circuits and methods of isolating the transistor from the frequency-determining circuit are desirable for good frequency stability. In these feedback circuits either junction or point-contact units may be used.

The effect of the transistor parameters on frequency of oscillation and the critical feedback level may be found by analysis of the Colpitts oscillator, as an example. The circuit is drawn with the transistor equivalent in Fig. 12-36. Using $C_s = C_1C_2/(C_1 + C_2)$, the circuit equations may be written as

$$j\left(\omega L - \frac{1}{\omega C_s}\right)\mathbf{I}_1 - \frac{j}{\omega C_1}\mathbf{I}_b + \frac{j}{\omega C_2}\mathbf{I}_c = 0 \quad (12\text{-}83)$$

$$-\frac{j}{\omega C_1}\mathbf{I}_1 + \left(r_b + r_e - \frac{j}{\omega C_1}\right)\mathbf{I}_b + r_e\mathbf{I}_c = 0 \quad (12\text{-}84)$$

$$\frac{j}{\omega C_2}\mathbf{I}_1 + (r_e - r_m)\mathbf{I}_b + \left(r_c + r_e - r_m - \frac{j}{\omega C_2}\right)\mathbf{I}_c = 0 \quad (12\text{-}85)$$

It is known that, for oscillations to occur, the determinant of the above equations must be zero, and equating the real terms and the reactive terms of the determinant separately to zero gives two equations from which the stability or feedback condition and the operating frequency may be found.

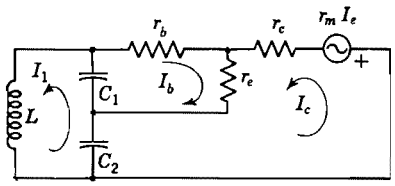



FIG. 12-36. Equivalent circuit for Colpitts oscillator.

The real terms of the determinant are found to be

$$\frac{r_c + r_e - r_m}{\omega^2 C_1^2} + \frac{2r_e - r_m}{\omega^2 C_1 C_2} + \frac{r_b + r_e}{\omega^2 C_2^2} + \left(\omega L - \frac{1}{\omega C_s}\right)\left(\frac{r_c + r_e - r_m}{\omega C_1} + \frac{r_b + r_e}{\omega C_2}\right) = 0$$

The frequency at which oscillation starts will be very close to the resonant frequency of the resonant circuit, at which $\omega L - 1/\omega C_s = 0$. Also, it may be assumed that $r_e \ll r_m$. The equation may then be solved, leading to

$$\frac{C_1}{C_2} = \frac{r_m \pm \sqrt{r_m^2 - 4(r_c - r_m)(r_b + r_e)}}{2(r_b + r_e)} \quad (12\text{-}86)$$

For the usual transistor the second term under the radical will be small with respect to the first, and so for oscillations to start and be maintained in the circuit

$$\frac{C_1}{C_2} \geq \frac{r_m}{r_b + r_e} \quad (12\text{-}87)$$

The reactive terms of the impedance determinant may be equated to zero as

$$\left(\omega L - \frac{1}{\omega C_s}\right)\left(r_b r_c + r_b r_e + r_e r_c - r_b r_m - \frac{1}{\omega^2 C_1 C_2}\right) - \frac{1}{\omega^3}\left(\frac{1}{C_1^2 C_2} + \frac{1}{C_2^2 C_1}\right) = 0$$

Writing $A = r_b r_c + r_b r_e + r_e r_c - r_b r_m$, the above reduces to

$$f = \frac{1}{2\pi}\sqrt{\frac{1}{LC_s} + \frac{1}{AC_1C_2}} \tag{12-88}$$

as the frequency of oscillation. The transistor is effective in modifying the resonant frequency of $L$ and $C_s$ through the term $A$, involving all the transistor parameters, and frequency stability is affected by changes in those parameters.

The Hartley oscillator is found to require

$$\frac{L_2 + M}{L_1 + M} \geq \frac{r_m}{r_b + r_e} \tag{12-89}$$

as the feedback condition for oscillation. The frequency of oscillation is

$$f = \frac{1}{2\pi}\sqrt{\frac{1}{C(L_1 + L_2 + 2M) - (L_1L_2 - M^2)/A}} \tag{12-90}$$

$M$ being the mutual inductance between $L_1$ and $L_2$.

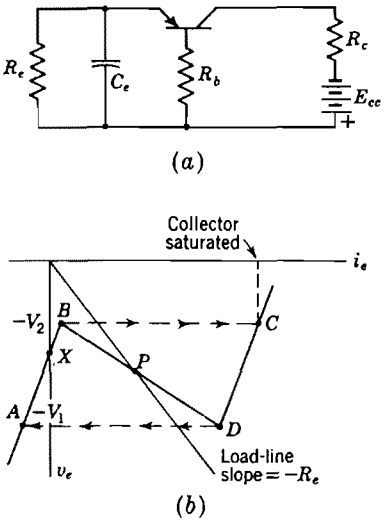



FIG. 12-37. (*a*) Relaxation oscillator using *N*-type point-contact unit. (*b*) Emitter negative-resistance characteristic.

**12-20. Negative-resistance Oscillators and Triggers.** The instability of the point-contact transistor with $\alpha_{ce} > 1$ and with a high base resistance is applicable in a number of circuits employing negative resistance. A simple relaxation oscillator is shown in Fig. 12-37*a*, where $R_e$, $R_b$, and $R_c$ are external resistors, $R_e$ to properly locate the operating region, $R_b$ to increase the base resistance and accentuate instability, and $R_c$ to limit collector current to a safe value. Explanation of operation follows from the idealized emitter characteristic of Fig. 12-37*b*.

Region $A$-$B$ is one of very small emitter current or cutoff, $B$-$D$ one of negative resistance and instability, and $C$-$D$ a region of high $I_c$ and positive-input resistance. Assume that at $t = t_o$ the capacitor $C_e$ has a negative potential to emitter of $-V_1$

which places the emitter beyond cutoff with $i_c = 0$, and operation is at point $A$ on the characteristic. Since the emitter current is so small, the capacitor $C_e$ discharges through $R_e$ with a time constant $T_1 = C_eR_e$. At time $t_1$ the potential of the emitter has risen to $-V_2$ at point $B$. The emitter rises above cutoff, collector current begins to flow, and because of the positive feedback when $\alpha_{ce} > 1$, the emitter current rises spontaneously to point $C$, the emitter voltage remaining at $-V_2$ because a capacitor cannot rapidly change its potential. At $C$ the collector current has approximately its full forward saturation value $E_{cc}/R_c$, and oscillation has reached one of its limits. The potential on $C_e$ then starts to increase negatively toward its original value of $-V_1$, charging through the parallel resistance of $R_b$ and $R_c$ with a time constant

$$T_2 = C_eR_bR_c/(R_b + R_c)$$

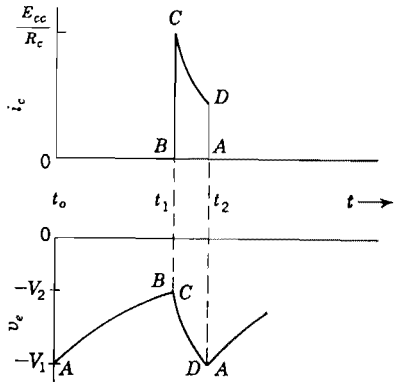



FIG. 12-38. Waveforms of current and voltage in relaxation oscillator.

When the potential reaches $-V_1$ at $t_2$, the operating point is at $D$ and the emitter and collector currents drop spontaneously to cutoff values, taking operation from $D$ to $A$, where the cycle is repeated. Waveforms of current and voltage are presented in Fig. 12-38.

Since the time of operation is determined by the two time constants, the frequency of operation may be stated as

$$f = \frac{1}{C_e[R_e + R_bR_c/(R_b + R_c)]} \tag{12-91}$$

If $T_2$ is made small, the current pulse becomes very short, and the emitter voltage wave approaches a saw-tooth form. Variation of $C_e$ and $R_c$ can control the width of the current pulse, for pulse generation.

By replacing $C_e$ with a series $LC$ circuit more nearly sinusoidal output may be obtained at approximately the $LC$ resonant frequency. Other forms of negative-resistance oscillators are shown in Fig. 12-39.

The transistor is unstable whenever the operating point is in region $B$-$D$. To initially place it there and start oscillation, resistor $R_e$ must be chosen so that the intersection of the $-R_e$ load line and emitter characteristic occurs in the region, as at $P$ (Fig. 12-37*b*). By placing the operating point at $X$ in a stable region, as by the removal of $R_e$ from the circuit as in Fig. 12-40*a*, the circuit becomes a *monostable trigger*. By

biasing the emitter three operating points are created as at $X_1$, $X_2$, and $X_3$, and the result is a *bistable trigger circuit* as in Fig. 12-41.

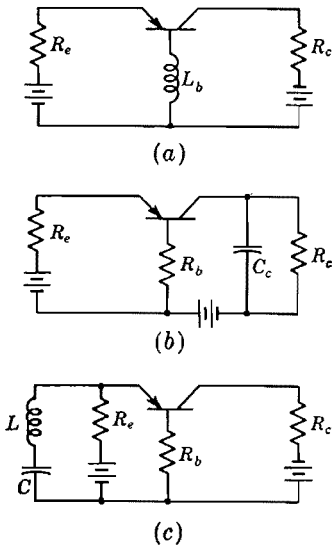



FIG. 12-39. Forms of negative-resistance oscillators. (*a*) Base-controlled; (*b*) collector-controlled; (*c*) sinusoidal.

In Fig. 12-40*a* the removal of $R_e$ makes the emitter current bias equal zero and places the operating point at $X$ on the $-v_e$ axis. This is in the stable region $A$-$B$, the circuit being biased in the cutoff region. Now, if a current pulse of sufficient magnitude ($>i_p$) is applied to the emitter, the operating point is forced past $B$ into the negative-resistance region. Because of the cumulative action the current jumps to $D$ with saturating $i_c$ flowing. The negative potential on $C_e$ starts to build up, and the operating point moves down to $C$, into the unstable $B$-$C$ region, and the operating point jumps to $A$. The charge on $C_e$ leaks off through the emitter-base circuit of $r_e + r_b + R_b$ until the operating point returns to $X$, ready for another trigger pulse. The $C_e(r_e + r_b + R_b)$ time constant is large due to $r_c$ (emitter in cutoff), and this factor limits the possible repetition rate.

In the bistable form of Fig. 12-41 the emitter must be so voltage-biased that the emitter load line intersects the characteristic at three points, $X_1$ and $X_3$ in a stable region, $X_2$ between in the unstable region. Operation may start with the "off" state at $X_1$, the current being small.

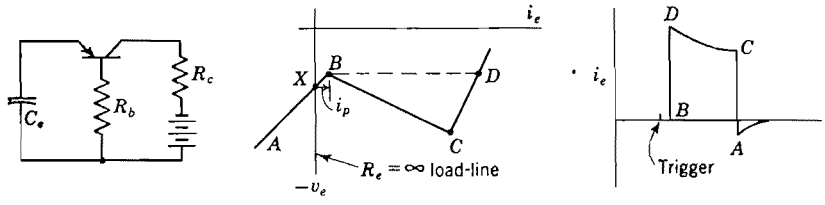



FIG. 12-40. (*a*) Monostable circuit. (*b*) Emitter characteristic. (*c*) Emitter current waveform.

Upon application of a positive current pulse of magnitude sufficient to drive the operating point past $B$, the circuit enters the unstable region, and operation jumps past $X_2$ to $X_3$, where the circuit is once more stable, high $i_c$ flows, and the circuit is "on." To move back to the "off" state

requires a negative emitter current pulse $i_p$ at least sufficient to drive the operation past $C$ into the unstable region, where it then swings back to the stable "off"-state point at $X_1$.

The sensitivity of operation is determined by the closeness of $X_1$ to $B$, or by the magnitude of $V_{ee}$. A large value of $V_{ee}$ decreases sensitivity but reduces chances of unwanted triggering. Point $X_3$ can be brought as close to $C$ as desired by varying $R_e$, the choice again being dictated by the reliability desired.

Actual pulse output waveforms are not as perfect as might be desired for short pulses, owing to inherent transistor properties. A time delay may be present due to the finite time of travel of the input holes ($N$-type material) from the emitter to the collector. Also, since the paths taken by the holes may differ, all the holes do not arrive at the collector at the same instant, giving a spread in rise time. A similar event delays the

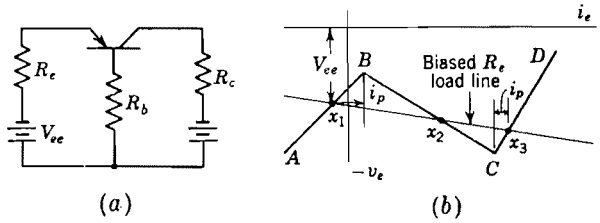



FIG. 12-41. (*a*) Bistable circuit. (*b*) Emitter characteristic.

fall of the pulse and causes a trailing edge to develop. A further addition to the trailing edge is due to storage of holes in the remote portions of the crystal while the circuit is at saturation level, with low applied voltage. The amount of such storage is a function of the magnitude of input current and the length of time the transistor is in the saturation state.

**12-21. The Tetrode Transistor.** The tetrode junction transistor is a form developed for improved high-frequency operation. In Fig. 12-42 a fourth lead is added to the base opposite the normal base lead, and a negative bias is added across the base layer. The bias $V_{bb}$ is made more negative than the emitter bias $V_{ee}$ for an $N$-$P$-$N$ unit.

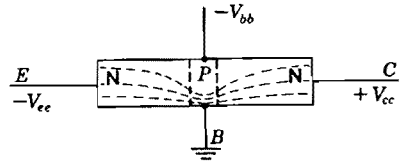



FIG. 12-42. The tetrode junction transistor.

Because of the transverse voltage drop between upper and lower base leads, only at the lower part of the base will the emitter be more negative than the base, a condition required for proper $N$-$P$-$N$ transistor operation. Therefore the emitter and collector currents are confined to a small region near the lower base lead.

The result is a reduced internal capacity and improved high-frequency response, although $\alpha$ is also reduced.

**12-22. Large-area Rectifiers.** Prior to an understanding of the semiconductor properties of germanium and silicon there were developed a number of rectifiers employing properties of selenium, copper sulfide, and cuprous oxide as semiconductors. These consist of a layer of the semiconductor bonded to a metal base. They are usually of $P$ type, the forward, or low-resistance, direction occurring with the semiconductor positive.

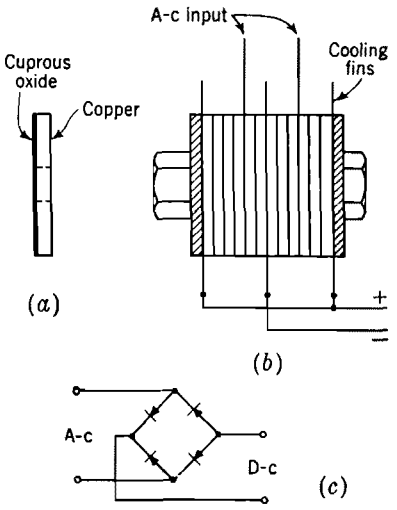



FIG. 12-43. (*a*) Single copper–cuprous oxide disk. (*b*) Bridge-connected rectifier having three disks in series in each arm. (*c*) Bridge circuit for instrument rectifiers.

In the copper–cuprous oxide cell developed by Grondahl in 1926 the cuprous oxide is formed on a copper washer by controlled oxidation at high temperatures. These washers are assembled on an insulated bolt and separated by lead washers to provide firm, large-area contact between washers. Each disk is limited to 8 or 10 volts in the reverse direction, so that many units are stacked in series for higher-voltage use as in Fig. 12-43. The reverse voltage falls with rising temperature, top limit for copper–cuprous oxide being about 45°C. Current-carrying capacity is proportional to area and to cooling.

The selenium rectifier, used in many electronic devices, is similar, the selenium being deposited on iron or aluminum. These disks will withstand up to 26 volts reverse emf per disk and may be operated up to 85°C. A volt-ampere characteristic for selenium is shown in Fig. 12-44. Rectifiers of smaller size and weight are possible with selenium than with

TABLE 28. CHARACTERISTICS OF RECTIFIERS

| | Silicon junction | Germanium junction | Selenium stack | Vacuum diode |
|---|---|---|---|---|
| PIV | 600 | 300 | 350 | 1500 |
| D-c current, ma | 300 | 75 | 150 | 270 |
| Volts drop (full load) | 1.0 | 0.5 | 6 | 45 |
| Maximum operating temperature, °C | 150 | 85 | 85 | 250 |

copper oxide. Efficiency vs. load current for selenium is plotted in Fig. 12-45.

Characteristics of several types of rectifiers are given in Table 28, for complete units of somewhat comparable ratings.

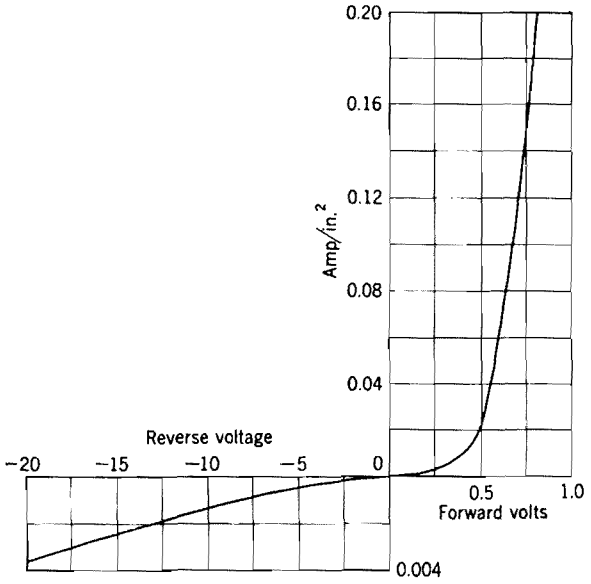



FIG. 12-44. Volt-ampere curve for a single selenium disk.

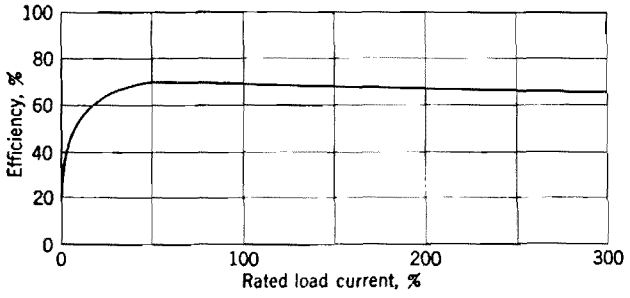



FIG. 12-45. Efficiency of a selenium bridge rectifier.

**12-23. Thermistors.** The rapid variation of resistance with temperature for certain semiconductors is employed in the thermally sensitive resistor, or *thermistor*, to measure temperature. Mixtures of various metallic oxides are used to give units with useful resistance variations in desired temperature ranges. In one type the resistance varies about 4 per cent per °C, at room temperature, and the resistance may change by a factor of $10^3$ in a 300°C range.

Thermistors are made as small beads or rods with wire leads, and bead diameters down to 0.015 cm (0.006 in.) are made. It has been found that many such units obey an exponential law of the form

$$\rho = \rho_o \epsilon^{(B/T - B/T_0)} \tag{12-92}$$

where $T$ = temperature, °C

$\rho_o$ = $\rho$, ohm-cm, when the temperature is at $T_o$

$B$ = constant having dimensions of temperature

The thermistor may be used in standard bridge circuits for the measurement of resistance and will usually show a great increase in sensitivity over that possible with metallic resistance thermometers. The sensitivity and their small size make them adaptable to much research work, particularly in medicine. High-resistance thermistor resistance bridges operated from alternating current may have their output amplified by vacuum-tube amplifiers, thus eliminating galvanometers or other delicate instruments.

## PROBLEMS

**12-1.** Solve for the values of the $h$'s in Eqs. (12-16) and (12-17) in terms of the $z$'s of Eqs. (12-11) and (12-12).

**12-2.** Find a constant-current generator form of equivalent circuit for the common-emitter amplifier. Repeat for the common-collector amplifier.

**12-3.** Devise and draw circuits for the measurement of $\alpha_{ce}$, $r_{11}$, $r_{22}$, $r_{12}$, and $r_{21}$.

**12-4.** Compute, using exact relations, the input resistance, output resistance, current amplification, and power gain for a common-emitter transistor amplifier having $r_e = 25$ ohms, $r_b = 500$ ohms, $r_c = 500{,}000$ ohms, $\alpha_{ce} = 0.96$, $R_L = 25{,}000$ ohms, $R_s = 100$ ohms.

Compute the same quantities by the approximate relations, and discuss the differences in results.

**12-5.** With $R_s = 1000$ ohms, $R_L = 5000$ ohms find the voltage gain for the transistor of Prob. 12-4 in common-base, common-emitter, and common-collector circuits.

**12-6.** A junction transistor having $r_e = 25$ ohms, $r_b = 350$ ohms, $r_c = 2$ megohms, $\alpha_{ce} = 0.98$ is to deliver an output of 4 volts with a load of 12,000 ohms. Find the required input current.

**12-7.** For $r_e = 30$ ohms, $r_b = 500$ ohms, $r_c = 2$ megohms, $\alpha_{ce} = 0.98$ plot current gain, voltage gain, $R_{11}$, and power gain against $R_L$ in the range of zero to $10^6$ ohms. Determine value of $R_L$ for maximum power gain in the common-emitter circuit.

**12-8.** Repeat Prob. 12-7 for the common-base circuit.

**12-9.** Repeat Prob. 12-7 for the common-collector circuit.

**12-10.** Derive an expression for the load $R_L$ giving maximum power gain in the common-emitter circuit. What load will be used for maximum power gain with a transistor having $r_e = 30$ ohms, $r_b = 250$ ohms, $r_c = 750{,}000$ ohms, $R_s = 1000$ ohms, $\alpha_{ce} = 0.97$?

**12-11.** For the transistor of Fig. 12-26 compute power-output efficiency, second-harmonic distortion, and average power dissipation and peak power dissipation for a load of 2000 ohms, $E_{cc} = 20$ volts, $I_{b,max} = 200\mu$a.

**12-12.** The constants of a junction transistor in a common-base circuit at 1 kc are $\alpha_{ce} = 0.96$, $r_c = 870{,}000$ ohms, $r_b = 400$ ohms, $r_e = 35$ ohms, $f_c = 220$ kc, $C_c = 20$ $\mu\mu$f. Find the complex value of $\alpha_{ce}$ at 400 kc.

**12-13.** An *RC*-coupled amplifier consists of two identical junction transistors in common-emitter circuits. A 12-volt battery is the common supply, and $R_L$ is 12,000 ohms for each stage. With $\alpha_{CB} = 50$, $r_c = 2 \times 10^6$ ohms, $r_b = 500$ ohms, $r_e = 35$ ohms find (*a*) the over-all current gain, (*b*) the input impedance of the first stage, (*c*) the decibel power gain with $R_s$ matched to $R_{11}$ of the first stage and $R_L$ for both stages as indicated.

**12-14.** Compute the value of coupling capacitor between stages in Prob. 12-13, for a response down 1 db at 30 cycles.

**12-15.** Compute the 1000-cycle current gain of Fig. 12-28*b*, with the **2N35** characteristics as $r_e = 40$ ohms, $r_b = 400$ ohms, $r_c = 1.5 \times 10^6$ ohms, $\alpha_{CE} = 0.98$.

**12-16.** A certain transistor with $\alpha_o = 0.96$ has $f_c = 0.6$ Mc. Find the value of $\alpha_{CE}$ at 0.1 Mc.

**12-17.** From the common-base equivalent circuit of Fig. 12-46 show the condition by which the point-contact transistor will be stable, or the impedance determinant positive, under short-circuit conditions.

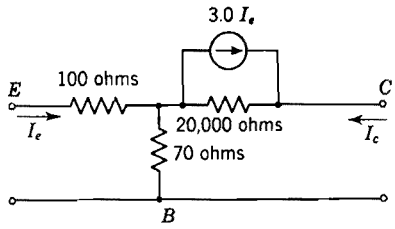



FIG. 12-46

**12-18.** For the circuit of Fig. 12-46 plot variation of $R_{11}$ as a function of $R_L$ for the range zero to 500,000 ohms.

**12-19.** Derive the starting conditions for a point-contact negative-resistance oscillator with a series-resonant circuit across $R_E$ in the grounded-base circuit.

**12-20.** Show that the condition for positive input resistance of any network configuration is that the impedance determinant be greater than zero.

**12-21.** Find an active one-generator equivalent circuit using the $h$ parameters of Eqs. (12-16) and (12-17).

**12-22.** Find an active one-generator equivalent circuit for the transistor using the $y$ parameters of Eqs. (12-14) and (12-15).

## REFERENCES

1. Seitz, F.: "The Modern Theory of Solids," McGraw-Hill Book Company, Inc., New York, 1940.
2. Cornelius, E. C.: Germanium Crystal Diodes, *Electronics*, **19**: 118 (February, 1946).
3. Torrey, H. C., and C. A. Whitmer: "Crystal Rectifiers," McGraw-Hill Book Company, Inc., New York, 1948.
4. Bardeen, J., and W. H. Brattain: The Transistor, a Semiconductor Triode, *Phys. Rev.*, **74**: 230 (1948).
5. Ryder, R. M., and R. J. Kirchner: Some Circuit Aspects of the Transistor, *Bell System Tech. J.*, **28**: 367 (1949).
6. Bardeen, J., and W. H. Brattain: Conductivity of Germanium, *Phys. Rev.*, **75**: 1216 (1949).

7. Shockley, W.: "Electrons and Holes in Semiconductors," D. Van Nostrand Company, Inc., Princeton, N.J., 1950.
8. Shockley, W.: The Theory of *P-N* Junctions in Semiconductors and *P-N* Junction Transistors, *Bell System Tech. J.*, **29**: 435 (1949).
9. *Proceedings of the IRE*, Special Issue on Transistors, **40**: 1283–1632 (1952).
10. Sziklai, G. C.: Symmetrical Properties of Transistors and Their Applications, *Proc. IRE*, **41**: 71 (1953).
11. Shea, R. F.: "Principles of Transistor Circuits," John Wiley & Sons, Inc., New York, 1953.
12. Coblenz, A., and H. L. Owens: "Transistors: Theory and Applications," McGraw-Hill Book Company, Inc., New York, 1955.
13. Ramsey, G.: The Selenium Rectifier, *Elec. Eng.*, **63**: 425 (1944).
14. *Proceedings of the IRE*, Special Issue on Solid-state Materials, **43**: 1701–2000 (1955).

CHAPTER 13

# PHOTOELECTRIC DEVICES AND APPLICATIONS

Photoelectric release of electrons or conduction by impact of light, as discussed in Chap. 2, is employed in four basic types of cells. These are:

*Photoemissive cells,* or phototubes, employing the phenomena of emission of electrons when light strikes a cathode of low-work-function material. The sensitive surface is enclosed in an evacuated or gas-filled envelope of glass or quartz, the electrons being captured by a positive anode.

*Photovoltaic cells,* in which the radiant energy striking the sensitive surface causes generation of an emf. With proper design the short-circuit current can become an almost linear function of the incident light flux.

*Photoconductive cells,* which vary their electrical resistance from a high value in the dark to a low value when exposed to light. They consist of a thin coating of selenium on a glass plate. An early form, they are now rarely used and will not be discussed here.

*Phototransistors,* in which the semiconductive properties of germanium or silicon are varied in accordance with incident light.

**13-1. Light.** *Light* is the name given to radiant electromagnetic energy having wavelengths roughly in the range from 0.01 to 100 microns, where the *micron* = $10^{-6}$ m and is abbreviated by the symbol $\mu$. Wavelength may also be stated in terms of the *angstrom unit* = A = $10^{-10}$ m, or the *millimicron* = m$\mu$ = $10^{-9}$ m.

The frequency of the light may be calculated from the wavelength through use of the velocity of light = $c = 2.99776 \times 10^8$ m per sec $\cong 3 \times 10^8$ m per sec, as

$$f = \frac{c}{\lambda}$$

where $\lambda$ is in meters.

The extremes of the wavelength range stated above lie in the far ultraviolet and the far infrared. Intermediate between these limits is the *visible range* extending only from about 0.4 $\mu$ in the violet to 0.7 $\mu$ in the red end of the spectrum. *Luminous light flux* is that portion of the radiant energy which produces a visual sensation in the human eye. Since the sensitivity of the human eye is not uniform to all colors in the

visible range, being most sensitive in the blue-green near 0.53 $\mu$, the eye cannot be used as a standard of intensity measurement.

The output of a photoelectric cell may be stated in terms of current output per unit of *total radiant power* striking the cell surface. This is frequently given as microamperes per microwatt of incident power over a specified color, or wavelength, range or from a given source. The color, or wavelength, range or the source must be stated since the conversion

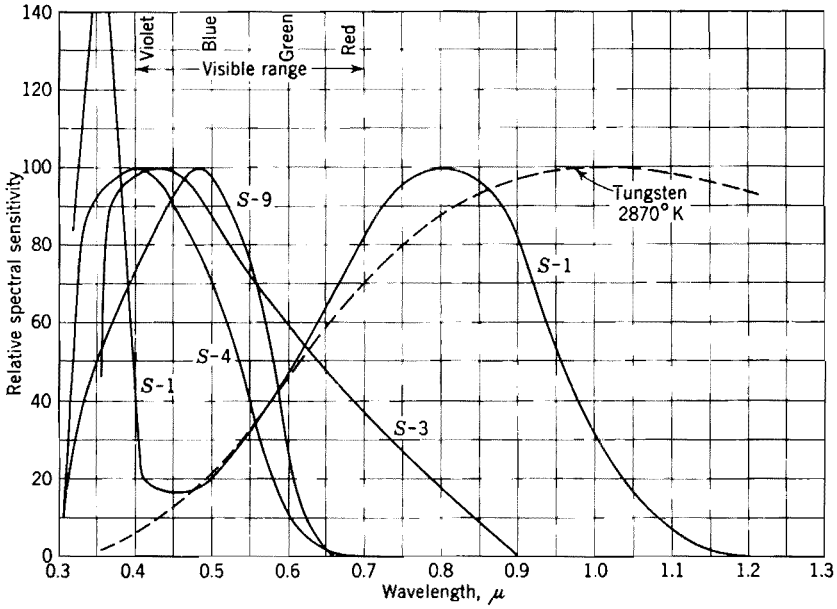



FIG. 13-1. Relative spectral sensitivity of various surfaces. Actual sensitivities are not equal at 100.

efficiency of various photoelectric cathodes varies radically with incident wavelength. A standard source for comparison purposes is frequently a tungsten lamp operating at a temperature such that its spectrum of radiant energy is that of a black body operating at 2870°K. The relative distribution of energy at different wavelengths for a tungsten filament operating at this *color temperature* is shown in Fig. 13-1.

Sensitivity may also be stated in terms of visible light only and is then known as *luminous sensitivity*, or microamperes per lumen of visible light. It is still necessary to state the source used. In this definition the *lumen* is used as a unit of luminous, or visible-light, flux.

The *candle* is the unit of luminous flux produced by a standard candle burning under specified conditions. Such a candle emits light in all directions and gives a total light flux of $4\pi$ lumens. Since the total solid

angle surrounding a point is $4\pi$ steradians, the *lumen* is the light flux emitted per unit solid angle from a standard candle.

The *foot-candle* is the illumination produced on the spherical surface by a standard candle at the center of a sphere of 1 ft radius. Since the total area of the spherical surface is $4\pi$ sq ft and the candle emits $4\pi$ lumens, it follows that a foot-candle equals a lumen per square foot.

While the lumens per unit solid angle from a point source remain constant, the flux incident per unit area decreases inversely with the square of the distance. To convert the candlepower rating of a point source into lumens, the following relation may be used,

$$L = \frac{CA}{d^2} \tag{13-1}$$

where $L$ is the illumination in lumens on $A$ sq ft at a distance $d$ ft from a point source of $C$ candlepower.

**13-2. Cathode Materials.** It was pointed out in Chap. 2 that photoemission of electrons was governed by the Einstein equation

$$hf = eE_w + \frac{mv^2}{2} \tag{13-2}$$

where $h$ = Planck's constant = $6.57 \times 10^{-34}$ joule-sec. Since the energy of an individual light photon is given by $hf$, the equation states that the energy of a photon is absorbed to overcome the work-function energy $eE_w$ and to supply any kinetic energy of emission $mv^2/2$.

A photon must have sufficient energy to at least supply the work-function energy, so that a threshold frequency $f_o$, below which emission is not possible, is determined as

$$f_o = \frac{eE_w}{h} \tag{13-3}$$

Only a few materials have low work functions which permit emission by photons of visible light. These materials are in the alkali-metal group of lithium, sodium, potassium, rubidium, and cesium. The latter, prepared as a monomolecular layer on a layer of cesium oxide on a silver base, provides a surface with considerable sensitivity in the infrared, shown as the curve for surface $S$-1 of Fig. 13-1. This extended long-wave sensitivity is of value when used with incandescent-light sources, as indicated by the tungsten radiation curve of the figure. The cutoff at about 0.3 $\mu$ is due to failure of the glass to transmit light farther into the ultraviolet. When the range must be extended to shorter wavelengths, a quartz bulb must be used.

In view of the differing color response of the various available surfaces, care must be taken to consider the color and spectral distribution of the light source to be used with any given phototube.

**13-3. The Vacuum Photoemissive Cell (Phototube).** The vacuum phototube consists of a cathode surface treated with cesium or other low-work-function material, with a small wire serving as anode, made small so as to cast no shadow on the cathode. Plate voltages are made

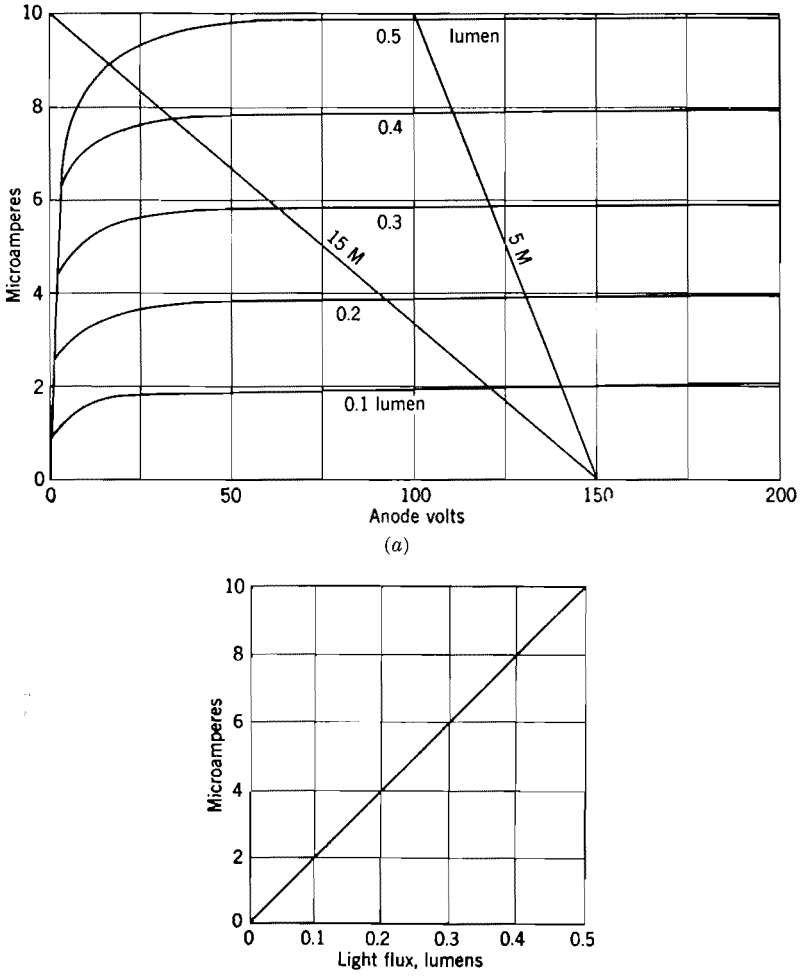



FIG. 13-2. (*a*) Vacuum-photocell characteristic (type 925). (*b*) Linearity of type 925 photocell indicated.

large enough so that all electrons emitted are attracted to the anode, or current saturation exists; for most tubes this occurs with $e_b$ greater than about 40 volts. A typical volt-ampere family appears in Fig. 13-2*a*, showing the independence of current and applied voltage.

The similarity between these curves and those for a pentode suggests the use of a current-source equivalent circuit for the vacuum phototube. The slopes of the characteristics in Fig. 13-2*a* indicate $r_p$ to be many megohms for $e_b > 40$ volts, so that the equivalent circuit simplifies to that of Fig. 13-3*a*, the tube becoming a generator of current of value $mF$, where $F$ is the incident-light flux. Actually the slight upward slope of the characteristics is due to the Schottky effect, in aiding emission of the electrons by the increased field at the cathode surface due to the rising anode voltage.

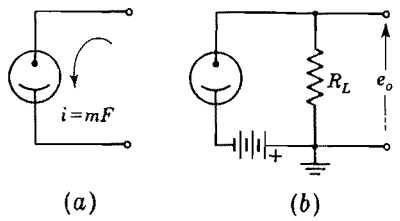



FIG. 13-3. (*a*) Equivalent circuit for vacuum photocell. (*b*) Load circuit added.

The constant of proportionality $m$, in microamperes per lumen, is usually stated in the ratings of a given photocell or may be found from the slope of Fig. 13-2*b*.

A practical output voltage may be obtained by connecting a load to the cell output as in Fig. 13-3*b*, where $e_o$ is a rise in voltage from ground, or

$$e_o = -iR_L = -mFR_L \tag{13-4}$$

For changes in light flux $\Delta F$ this becomes

$$e_o = -m\,\Delta F\,R_L$$

and $\Delta F$ may be considered the peak-to-peak variation if sinusoidal light-source modulation is employed.

Load-line analysis may also be employed as illustrated by Fig. 13-2*a*. It can be seen that for high linear output voltages it is desirable that $R_L$ and $E_{bb}$ both be high. In fact, since $e_o$ is proportional to $R_L$, it is desirable that $R_L$ be increased to a value of many megohms. A practical limit exists, however, in the inability to maintain tube-socket and other insulation values sufficiently high in practice. Since these insulation paths are in parallel with the load resistor, the inevitable leakages limit practical $R_L$ values to about 5 to 20 megohms unless extreme precautions, amounting to laboratory cleanliness, are undertaken.

**Example.** A vacuum phototube as in Fig. 13-2*a* has a sensitivity of 20 $\mu$amp per lumen and is used with an incandescent-lamp source. If $\Delta F = 0.2$ lumen, find the load resistance needed for a 50-volt output change with $E_{bb} = 200$ volts.

Since

$$\Delta e_o = m\,\Delta F\,R_L$$

$$R_L = \frac{\Delta e_o}{m\,\Delta F} = \frac{50}{20 \times 10^{-6} \times 0.2} = 12.5 \text{ megohms}$$

This result may also be found graphically from Fig. 13-2*a*. The needed load line will start at $E_{bb} = 200$ volts and proceed upward to intersection of the 150-volt 0.2-lumen line at 3.95 $\mu$amp. The reciprocal of its slope then gives

$$R_L = \frac{50}{3.95 \times 10^{-6}} \cong 12.5 \text{ megohms}$$

A vacuum phototube is linear with light intensity and has quite stable output with time. Its operation is independent of applied voltage above 30 to 40 volts and is essentially instantaneous, since emission occurs within $10^{-9}$ sec of light incidence. Most such cells must not operate at temperatures above 125°F because of possible damage to the emitting surface.

**13-4. The Gas-filled Phototube.** The current handled in vacuum phototubes is minute, being only a few microamperes. This magnitude

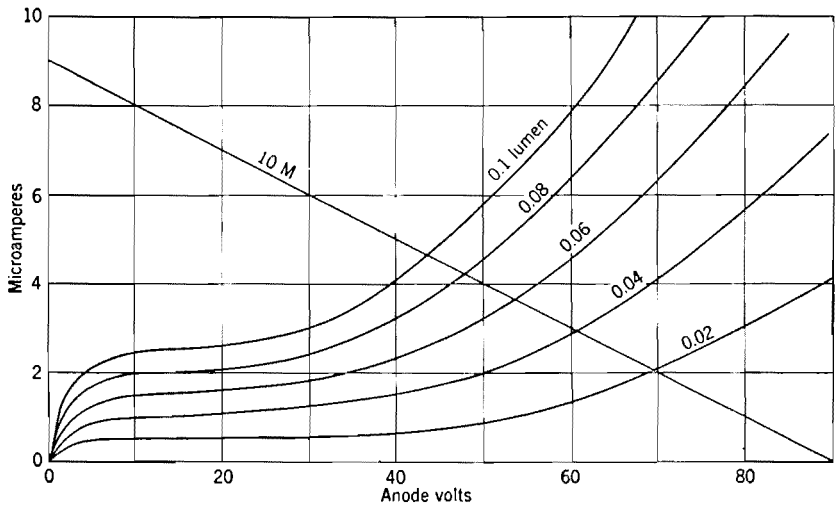



FIG. 13-4. Gas-filled photoemissive-cell characteristics (type 930).

may be increased by introduction of a low-pressure gas, usually argon, into the tube, giving characteristics as in Fig. 13-4. At a given light flux but with very low anode voltage, the current is limited by space charge and increases with anode voltage. After the saturation voltage is reached at about 10 volts, the current becomes independent of anode voltage. At such low anode voltages the electrons in passing between cathode and anode are not given sufficient energy to ionize any of the gas atoms, but some ionization does take place as the anode voltage is further increased. This ionization causes additional electrons and ions to reach the electrodes and thus adds to the current. Further increases in voltage give more electrons sufficient ionizing energy, and the total current increases further, as indicated by the curves.

The process must be limited, however, at a voltage insufficient to cause secondary emission from the cathode due to bombardment of the positive ions, since such bombardment would soon destroy the monomolecular surface. It would also cause the discharge to run away or be self-maintaining and not controlled by the incident light, and this is not desired.

The safe limit is usually set at about 90 volts, although this may have to be reduced for very bright light. The current increase due to the presence of gas is then of the order of seven to ten times that of the vacuum phototube. This ratio of current increase in the gas phototube is called the *gas-amplification factor*.

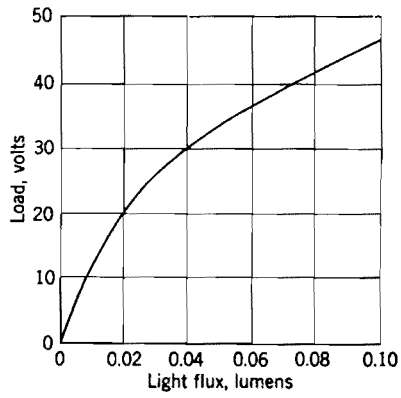



FIG. 13-5. Output vs. light flux for 930 gas photocell; $R_L$ = 10 megohms.

Gas phototubes do not have a linear light flux vs. output-current relation and are not well suited to use with modulated light above 2000 cycles, since the time required for steady state in the gas-ionization process is appreciable, and waveform distortion may occur. The nonlinear characteristic is shown in Fig. 13-5.

Analysis of gas-photocell circuits must be carried out by use of the load line, as illustrated in Fig. 13-4. This method yields accurate results for the usual gas-tube applications in the field of on-off control under dim-light conditions.

**Example.** A gas phototube is operated at $E_{bb}$ = 90 volts, with $R_L$ = 10 megohms. Plot a curve of output voltage vs. light input for the tube of Fig. 13-4.

The load-line intercepts will be at $E_{bb}$ = 90 volts and at $I = 90/(10 \times 10^6)$ = 9 $\mu$amp. Values for the curve of Fig. 13-5 may then be selected.

**13-5. Phototube Relays.** Since the phototube has a high output impedance of many megohms, it works well into the grid circuit of a vacuum tube, the tube providing a power gain and a step-down in impedance to a more useful level. Typical relay-operating applications are shown in Fig. 13-6. The voltage divider $P$ determines the value of light required to cause the relay to pick up or drop out. The sensitivity of the circuit, or range of light variation between pickup and drop-out, is a function of $R_g$ and amplifier gain. To make the gain reasonable, high-resistance relays must be used, and values up to 10,000 ohms are common.

The polarity of the phototube connection and its applied voltage determines whether the relay picks up or drops out on incidence of light.

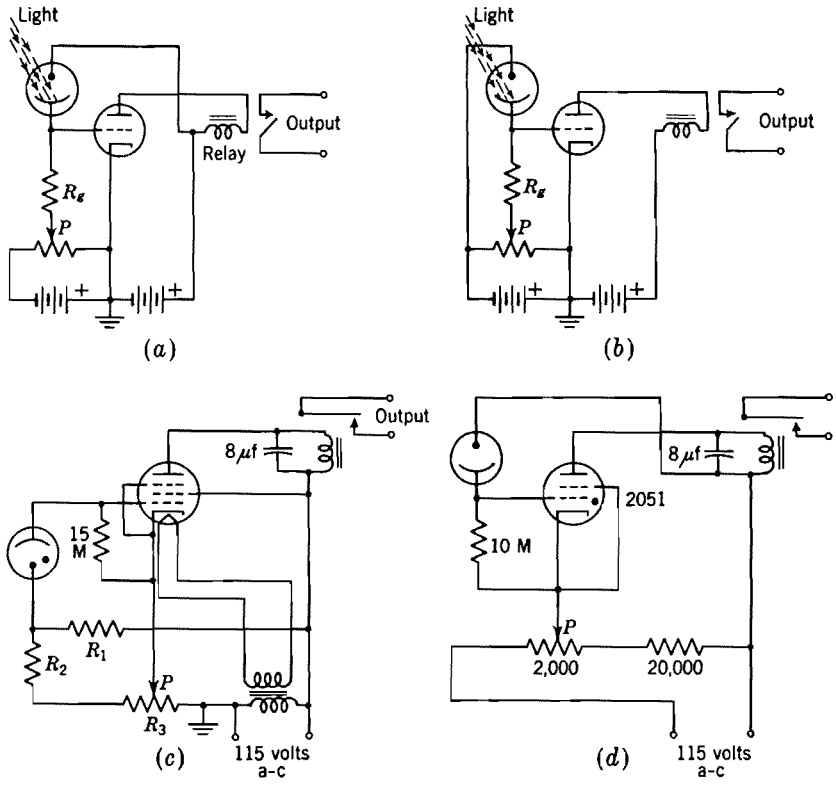



FIG. 13-6. (*a*) Relay closed on light incidence. (*b*) Relay open on light incidence. (*c*) Alternating-current-operated relay with gas phototube. (*d*) Alternating-current-operated with thyratron tube.

**Example.** A triode of $\mu = 20$, $r_p = 8000$ ohms has in its plate circuit a relay of 10,000 ohms resistance. This relay will pick up its armature at 4.0 ma and drop out at 2.4 ma. Assuming a vacuum-photocell sensitivity of 20 $\mu$amp per lumen and linear operation of the triode, determine the value of photocell load resistor to operate the relay armature on a change of light of 0.05 lumen.

For linear operation the triode equivalent circuit gives

$$\Delta I_p = \frac{\mu \, \Delta E_g}{r_p + R}$$

where $R$ is the relay resistance = 10,000 ohms. The change $\Delta I_p = 4.0 - 2.4 = 1.6$ ma. Then

$$\Delta E_g = \frac{\Delta I_p(r_p + R)}{\mu}$$

$$= \frac{1.6 \times 10^{-3} \times 18{,}000}{20} = 14.4 \text{ volts}$$

Then for the photocell, since $\Delta F = 0.05$ lumen,

$$\Delta e_o = 14.4 = m\,\Delta F\,R_L = 20 \times 10^{-6} \times 0.05 \times R_L$$

$$R_L = \frac{14.4}{10^{-6}} = 14.4 \text{ megohms}$$

Figure 13-6*c* illustrates a sensitive a-c-operated relay using a gas-filled phototube and a pentode amplifier. Resistor network $R_1$ and $R_2$ serves to limit the voltage across the phototube so as not to exceed the rated 90 volts. Resistor $R_3$ allows adjustment of the light level for operation. The a-c supply voltage is rectified in the plate circuit, with operation only on the positive half cycles. The capacitor across the relay serves to filter the rectified current pulses and prevent relay chatter due to the intermittent current flow.

The circuit in Fig. 13-6*d* employs a gas-filled triode or thyratron (see Chap. 15) as a relay, giving a larger output current than is obtainable from most vacuum tubes. The variable resistor allows adjustment of pickup light intensity. The capacitor again filters the relay coil current because of the a-c supply.

Many more elaborate forms of these circuits are to be found described in the literature.

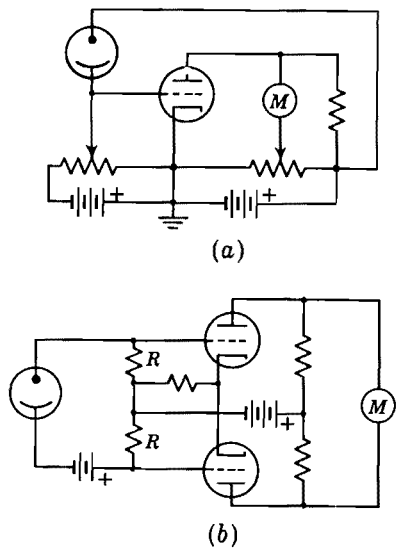



FIG. 13-7. (*a*) Circuit for balancing out the zero current. (*b*) Differential amplifier with photocell input.

**13-6. Phototube Circuits for Measurement Purposes.** If an indication of light flux reaching the phototube is required, the relays of Fig. 13-6*a* or *b* may be replaced with a microammeter, but because of nonlinearity of calibration and dependence on constancy of applied voltages the circuits are not very useful. For most light-measuring instruments, as in spectrophotometers, measurement of light transmission through turbid liquids or photographic images, and the like, a combination of the phototube with the d-c amplifier techniques of Chap. 8 is usually made.

Figure 13-7*a* shows the combination of a phototube and means for balancing out the zero light current of the triode. This circuit may be operated in the linear region of the triode but must use regulated supply voltages. The differential amplifier of Fig. 13-7*b* eliminates the latter difficulty to a great extent with matched tubes. In both cases the sensi-

tivity and range of light covered depend on the magnitude of $R_g$ which can be maintained as a cell load.

A basic circuit of very high light sensitivity is that of Fig. 13-8*a*, in which a second phototube $T_2$ is substituted for $R_g$. Since the plate resistance of $T_2$ may be very high, the output voltage of $T_1$ is increased, and since the glass-enclosed cell is well insulated, this high value can be reasonably well maintained in practical situations, particularly for laboratory measurement work. The circuit takes advantage of the fact that the volt-ampere curves of the vacuum phototubes continue to rise very slightly at high plate voltages owing to the Schottky effect. Since $T_2$ acts as a load on $T_1$, the volt-ampere curve for $T_2$ may be plotted as a load line for a particular characteristic of $T_1$, these curves appearing as solid lines in Fig. 13-8*b*.

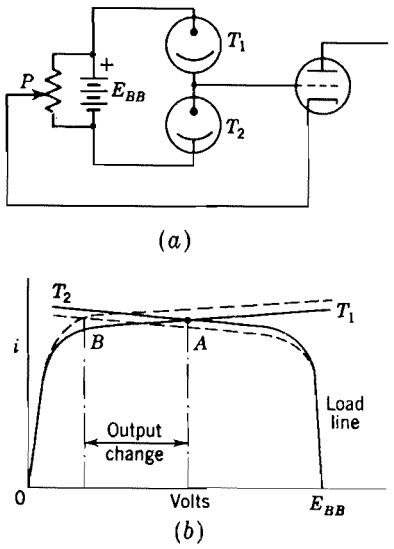



FIG. 13-8. (*a*) High-sensitivity circuit. (*b*) Load-line calculation for (*a*).

A small differential shift in illumination, increasing light on $T_1$ and reducing it on $T_2$, results in the dashed characteristics for the two cells and a shift in the output-voltage point from $A$ to $B$. The magnitude of this voltage change can be very large for a small illumination shift, giving the circuit a high sensitivity.

The output of the circuit is relatively independent of absolute illumination value since the phototubes will divide the applied voltage equally for any value of illumination which is balanced between the cells. Initial balance adjustment is made by voltage divider $P$.

**13-7. Use of Modulated Light.** To avoid the use of a d-c amplifier, many modulated-light systems are employed with a-c amplifiers. Modulated light is also used in systems where it is desired to avoid accidental operation by stray constant light, as in burglar-alarm systems; modulated light is the exciting source for the phototube in sound motion pictures.

In Fig. 13-9*a* is shown a rotating light chopper to give a photocell output which may be amplified by a conventional $RC$-coupled amplifier. The a-c component is dependent on the amplitude of light received when a tooth of the wheel is in the beam and the light received when the beam passes through a slot of the wheel. A $Q$ point for photocell operation is

established at the average value of illumination during shutter rotation. Frequency of output and amplifier design depend on the number of teeth passing through the beam per second. Other mechanical chopper forms as in Chap. 8 are also employed.

A second modulated-light system, suitable for comparison of light transmission of sample $A$ and a standard $B$, is shown in Fig. 13-9*b*. The oscillating mirror throws the light beam alternately through $A$ and $B$. A steady, or d-c, photocell output is obtained only if the light transmission of both $A$ and $B$ is identical, and this will give zero output from

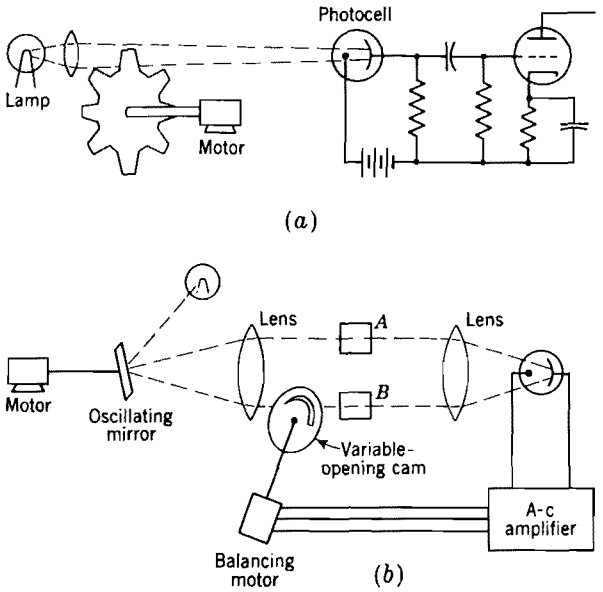



FIG. 13-9. (*a*) Light chopper for modulating light beam. (*b*) Oscillating mirror system for comparison of samples $A$ and $B$.

the a-c amplifier. If one path passes more light than the other, the phototube output will have a superimposed a-c component, of magnitude dependent on the differential transmission characteristics and with frequency determined by the mirror revolutions per second. The relative phase of this a-c component will depend on whether sample or standard gives the greater transmission.

The a-c amplifier uses this information from magnitude and phase of signal to operate the balancing motor in proper speed and direction so as to restore the light balance and give zero photocell output again. The balancing motor does this by rotating a light cam or plate with a variable-width slit in the light beam reaching the standard $B$, reducing or increasing the transmission of the $B$ path until its light output matches that

of the sample in $A$. This is a form of null balance system, and therefore the maintenance of constant characteristics by the photocell and amplifier is not important, and a gas cell may be used without regard for its nonlinearity since it serves merely as an unbalance detector.

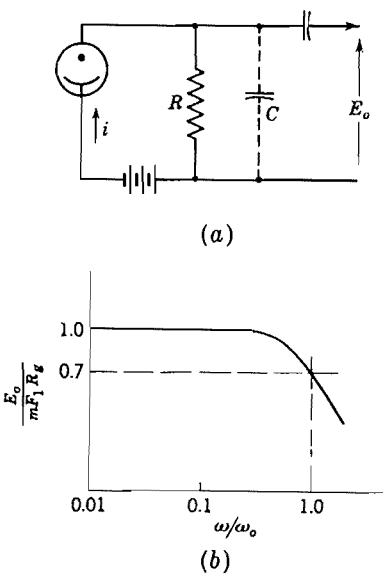



FIG. 13-10. ($a$) Alternating-current output circuit of phototube. ($b$) Response as a function of $\omega/\omega_o$.

In any modulated-light system, and particularly in those for pickup of audio-frequency signals from film as in sound motion pictures, the circuit of the photocell output and amplifier input becomes as in Fig. 13-10$a$, with $C$ representing the internal phototube and amplifier input capacities. This input circuit may then be followed by a conventional $RC$-coupled amplifier.

The phototube $Q$ point is determined by an average value of illumination $F_o$, with the total light flux being

$$F = F_o + F_1 \sin \omega t$$

The output current is

$$i = m(F_o + F_1 \sin \omega t)$$

The a-c component only will be effective in the following amplifier, and the a-c rms voltage across $R$ and $C$ in parallel is

$$E_o = \frac{mF_1}{\sqrt{2}} \left( \frac{-jR/\omega C}{R - j/\omega C} \right)$$

which may be placed in the form

$$E_o = \frac{mF_1R}{\sqrt{2}\sqrt{1 + (\omega/\omega_o)^2}} \underline{/\tan^{-1}(-\omega CR)}$$

where $\omega_o = 2\pi/CR$. Relative output $E_o/mF_1R$ is plotted in Fig. 13-10$b$ as a function of $\omega/\omega_o$. As long as $\omega \ll \omega_o$, the output is constant and $E_o = mF_1R/\sqrt{2}$. Since $C$ may reach a value of 100 $\mu\mu$f, then $\omega_o$ may be that due to only a few thousand cycles unless $R$ is lowered below usual phototube loads and frequency distortion avoided.

Gas phototubes cannot ordinarily be employed in such circuits above about 2000 cycles without considerable additional droop in high-frequency response due to the time delay in ionization. This may be compensated

by $RC$ circuit design and gas tubes used for pickup in sound motion pictures on a small-signal, and therefore approximately linear, basis.

**13-8. Electron Multipliers.** Emitted currents from photoelectric surfaces are very small, especially with low light levels. These currents may be directly amplified in a device known as an *electron multiplier*, utilizing the principles of secondary emission.

Electrons emitted from a photoelectric cathode $K$ by light are attracted to a positive electrode $D_1$, as in Fig. 13-11. They bombard the treated

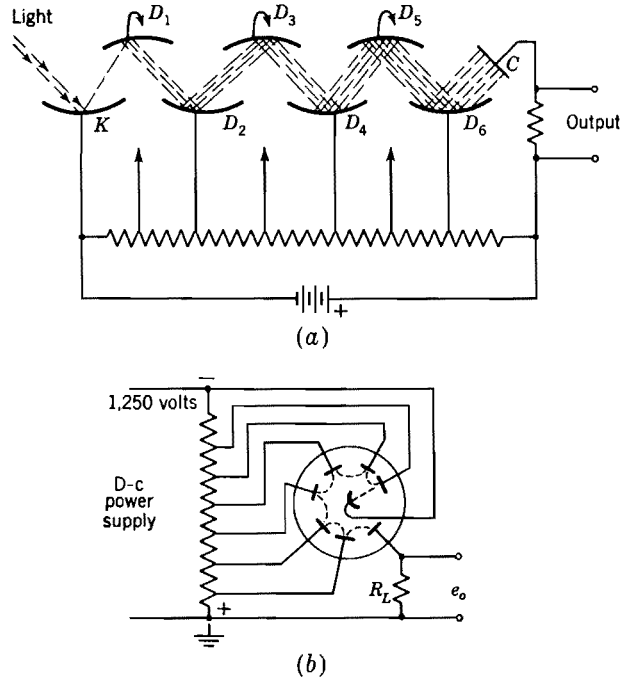



FIG. 13-11. (*a*) Curved electrode multiplier. (*b*) Multiplier circuit.

surface of $D_1$, having a secondary-emission coefficient $\delta$ above unity. Then $\delta$ electrons leave $D_1$ for every electron striking it, on the average. These electrons may again be attracted toward another positive electrode $D_2$, at a potential above that of $D_1$. This surface may also be treated to yield a similar secondary-emission coefficient $\delta$, so that $\delta^2$ electrons leave $D_2$ for every electron originally leaving the photoelectric cathode $K$.

This process may be repeated $n$ times on $n$ electrodes, or *dynodes* as they are called, giving an over-all current gain of $\delta^n$. In practice $\delta$ may be in the range of 5 to 10, and as many as nine cascaded dynodes are used for current gains as high as 2,000,000. The dynodes of Fig. 13-11*a*

are shaped to form curved electric fields which will focus the emitted electrons accurately onto each succeeding dynode, since if any electrons stray and miss a dynode, the value of gain will be reduced. The focusing may also be achieved by a combination of electric and magnetic fields, but electric-field focusing is usually employed for simplicity.

The potential gain of a multiplier is limited by the noise or random variation in the output current. With very low light intensities and very small currents the quantized nature of the emission introduces a small irregularity or noise into the output of the photoelectric cathode.

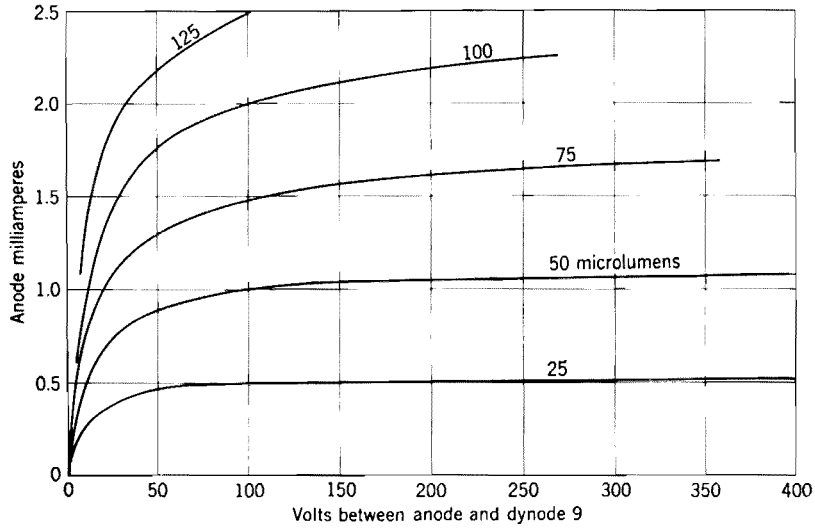



FIG. 13-12. Characteristics of type 931-A photomultiplier. Volts per dynode = 100.

Assuming some variation in $\delta$ from electron to electron, the noise may be further increased by the additional dynode secondary emission. It is found, however, that if normal statistical laws apply then the presence of the multiplying stages increases the noise of the original photoelectric emission only by the factor

$$\sqrt{\frac{\bar{\delta} - \bar{1}}{\delta}}$$

which is small if $\delta$ is of the order of 9 or 10. Thus the random output noise is not appreciably different from that of the original photoelectric current.

There are two other factors also serving as limits on the usable gain of a multiplier. These are the effects of positive gas ions due to imperfect vacuum and thermionic emission from the dynodes. Because of the low

work function the thermionic emission may be appreciable from the dynodes even at room temperature.

The characteristics of a photomultiplier depend on the voltage per dynode and on the potential between the last dynode and the collector anode. The voltage per dynode usually is chosen in the range of 75 to 100 volts since satisfactory sensitivity is then obtained. Performance in terms of last-dynode collector voltage is shown in Fig. 13-12. Comparison with Fig. 13-2*a* shows that the multiplier gives output in milliamperes for illumination in microlumens, whereas the usual vacuum photocell gives output in microamperes for illumination of the order of lumens.

**13-9. The Photovoltaic Cell.** The *photovoltaic cell* generates an emf upon incidence of light. Such cells consist of a layer of a semiconductor on a metal base plate and are of two types, either copper oxide on copper or iron selenide on iron. The semiconductor may be covered by a thin film of a transparent conducting lacquer. Circuit connections are then made to the base and to a spring contact made on the periphery of the lacquered area, as indicated in Fig. 13-13.

Upon incidence of light energy, electrons are given enough energy to pass across the forbidden energy gap, and if the semiconductor is placed on a metal base, these high-energy electrons can flow over into the metal. As long as the illumination continues, there is a supply of these electrons in the conduction band and current continues.

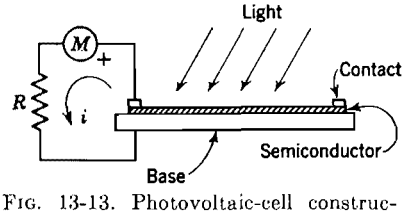



FIG. 13-13. Photovoltaic-cell construction.

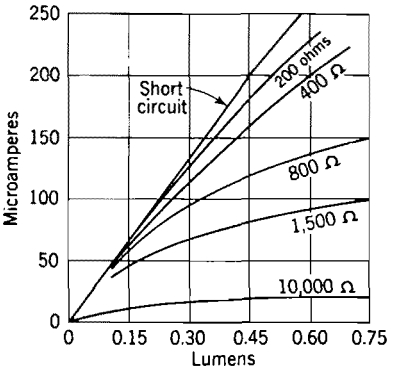



FIG. 13-14. Output vs. illumination for a photovoltaic cell with various loads. Area = 1.1 sq in.

The voltage generated is of the order of a few hundred millivolts in bright light, and this voltage is not a linear function of light flux. The cell resistance is small, and in sunlight currents of a few milliamperes may flow. It is found that the short-circuit current is a linear function of light flux, as indicated in Fig. 13-14, and is proportional to the total light flux or number of photons received, as might be expected from the conduction theory.

If a small resistance such as that of a microammeter is used as a load, the response is approximately linear with light and the instrument may

be calibrated in terms of incident illumination. This is the application in the photographic exposure meter. The current may also be used directly to operate sensitive relays for illumination control, such as in school classrooms. Since spectral response may closely follow that of the human eye, this is a desirable control application.

Cells of this nature are sensitive to high temperature, and the selenide cell may be damaged if operated above 55°C. Fortunately, it is found that both emf and internal resistance are temperature functions, so that the short-circuit current is nearly independent of ambient-temperature changes.

Since the photovoltaic cell develops such a small d-c voltage, amplification directly by vacuum tubes is not suitable, except by a chopper amplifier. Use with modulated light must also be limited to very low frequencies because of the shunt capacity between semiconductor and base. This layer is so thin and the capacity so high that the output may be reduced one-half even with a 60-cycle variation.

Photovoltaic cells may be connected in parallel to increase the total area and the total received light flux and the output current. They may also be connected in series for working into high-resistance loads.

**13-10. The Phototransistor.** While early photoconductive cells employed selenium as the semiconductor whose resistance decreased with light, more modern units employ germanium or silicon in forms similar to the point-contact or junction diodes or transistors.

Radiant energy is absorbed in raising an electron from the top of the filled band across the gap into the conductivity band. A hole is then left in the filled band, and this hole may migrate, as described in Chap. 12. Both the electron and hole are able to contribute to the photocurrent.

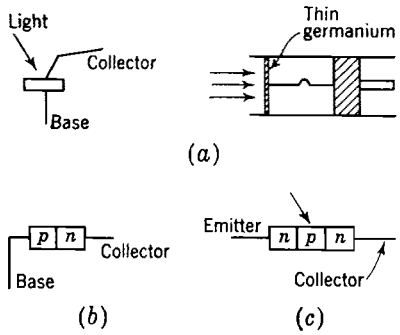



FIG. 13-15. Phototransistor forms. (*a*) Point contact; (*b*) *P-N* junction; (*c*) *N-P-N* junction.

The point-contact and junction types of Fig. 13-15*a* and *b* operate with reverse bias, that is, with positive voltage applied to the collector, a negative voltage to the *P* side. Light incident near the point contact or junction liberates carriers as mentioned above, and these diffuse to the electrodes, constituting a current. It appears that the holes form a positive space charge near the collector, making it relatively easy for the freed electrons to pass through the barrier; thus efficiencies somewhat in excess of the theoretical 1 electron-hole pair per quantum of incident light energy may be ob-

tained in point-contact or *P-N* units. This may be increased to as much as 100 electron-hole pairs per quantum in the case of the *N-P-N* unit of Fig. 13-15*c* because of the inherent current gain of this configuration as a transistor.

Owing to impurity atoms there will be some electrons in the conduction band at room temperature. These will produce a *dark current* of a

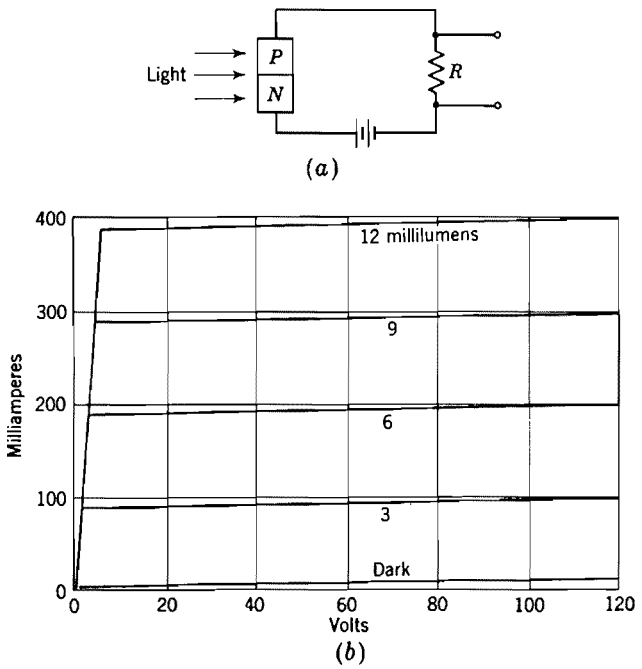



FIG. 13-16. (*a*) *P-N* junction circuit. (*b*) Approximate characteristics.

few microamperes. As long as this current is small, it will not interfere with the photocurrent.

Germanium and other semiconductors give a response further into the infrared than is obtainable from the photoemissive surfaces. The work function of the usual cesium photoemissive surface is about 1 ev, whereas the width of the unallowed energy gap in germanium is of the order of 0.7 ev. The light quantum required to produce photoconductivity is smaller than that required for photoemission in the approximate ratio 0.7:1.0. The gap energy for germanium being taken as 0.7 ev, then the threshold, or longest exciting wavelength, for germanium will be

$$\lambda_0 = \frac{hc}{eE} = \frac{6.55 \times 10^{-34} \times 3 \times 10^8}{1.60 \times 10^{-19} \times 0.7} = 1.75\ \mu$$

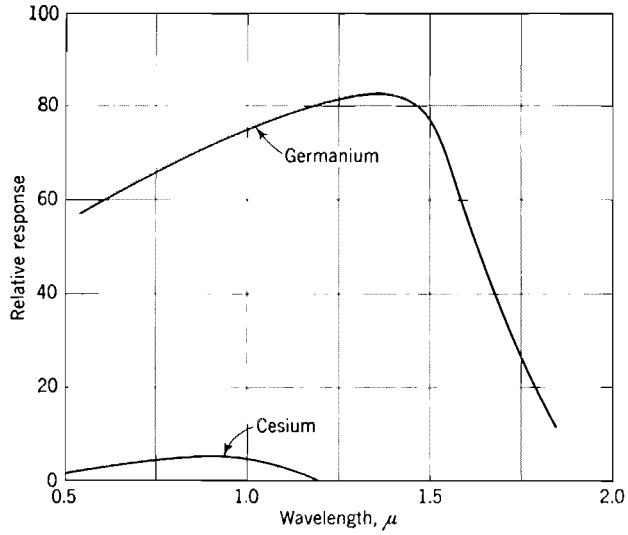



FIG. 13-17. Comparison of photoresponse of germanium and cesium.

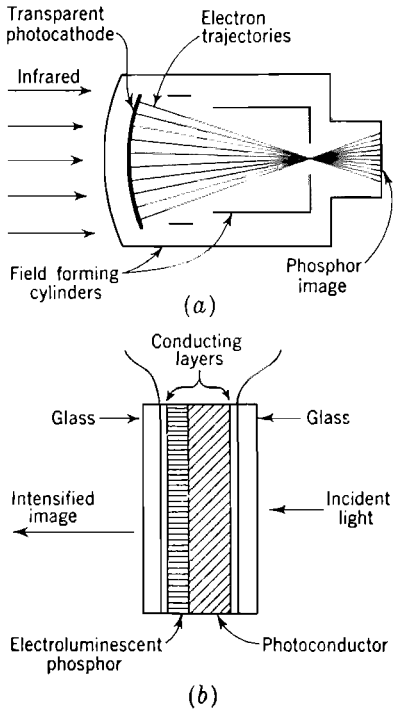



FIG. 13-18. (*a*) Image intensifier. (*b*) Solid-state light amplifier.

This wavelength is in the infrared, and a response curve for germanium is compared with one for cesium in Fig. 13-17.

**13-11. Light Amplifiers.** A further application of the light sensitivity of semiconductors is the *light amplfiier*, in which an image of a scene at low intensity is amplified to produce a similar image at a higher intensity. This is similar to but differs from the *image converter*, or "snooperscope," in which amplification of input was coupled with a frequency conversion from light in the invisible infrared to light in the visible range. A similar device is also employed to translate X-ray images into the visible range with increased intensity.

The image converter of Fig. 13-18*a* utilizes a transparent photocathode, receiving nonvisible illumination on one side and emitting electrons from the other. These

electrons are focused by electric fields to a phosphor on which the amplified and visible image appears.

The direct light amplifier of Fig. 13-18*b* utilizes semiconductor materials in a sandwich between two glass plates and two transparent conducting coatings. The electroluminescent phosphor layer is excited into luminescence by the application of a potential difference. Electrically in series with the phosphor layer is a layer of photoconductor of high dark resistance. A potential is applied between the conducting coatings, and in the dark this voltage will appear largely across the photoconductor.

When an image is focused on the photoconductor, its volume resistance decreases from point to point on its surface and the voltage across the electroluminescent layer correspondingly increases. The electroluminescent layer will emit light in direct proportion to the applied electric field, and thus in proportion to the incident illumination from point-to-point on the photoconductor. A gain in image brightness of about 20 has been achieved, and greater gains are anticipated with further research.

## PROBLEMS

**13-1.** Plot current vs. illumination curves for a 930 gas phototube with loads of 1, 7.5, and 15 megohms, $E_{bb} = 90$ volts.

**13-2.** A vacuum photocell has a sensitivity of 16 $\mu$amp per lumen. A change of 7 volts on the triode grid is required to operate the relay in the plate circuit of Fig. 13-6*a*. Specify $\Delta F$ if $R_g$ is 4 megohms.

**13-3.** Two vacuum phototubes have identical areas and physical characteristics except that one has an *S*-1 and the other an *S*-3 surface. Determine the ratio of currents obtained from the two cells when illuminated by light from a tungsten filament at 2870°K color temperature.

**13-4.** A 32-candlepower auto head lamp (assumed a point source) is located 6 ft from a phototube of 1.5 sq in. sensitive area and sensitivity of 20 $\mu$amp per lumen. Calculate the focal length of a lens 1.5 in. in diameter which will be required between source and phototube to give a phototube output of 40 volts across a load of 12 megohms.

**13-5.** A 925 vacuum phototube receives a change in illumination from 0.12 to 0.27 lumen. With $E_{bb} = 150$ volts determine the output-voltage change across a 2-megohm load.

**13-6.** The circuit of Fig. 13-6*b* is used to control the illumination in a schoolroom with a 925 vacuum phototube and a 6J5 triode having $r_p = 8000$ ohms, $\mu = 20$. With illumination on, $P$ is adjusted for grid bias of $-2$ volts and plate current of 5.5 ma. Assuming the triode operates in the linear region, find the values of room illumination at which the lamps will be turned on and off, with $R_g = 10$ megohms and a relay of 7500 ohms resistance with pickup and drop-out currents of 4.0 and 2.3 ma, respectively.

**13-7.** A 32-candlepower point-source lamp is used as a source for a 930 gas-filled photocell with 90 volts applied through an 8-megohm load. The area of the cell surface is 0.4 sq in. How far away can the lamp be from the cell and still produce a current of 5 $\mu$amp in the load?

**13-8.** In Fig. 13-6*c* designate values for $R_1$, $R_2$ to prevent the voltage across the gas photocell exceeding 90 volts at any instant. Resistor $R_3$ is 2000 ohms.

**13-9.** Two 925 vacuum photocells are employed in the circuit of Fig. 13-8*a* as $T_1$ and $T_2$. The steady illumination is 0.2 lumen, and $E_{bb}$ = 200 volts. If the illumination shifts +0.01 lumen on $T_1$, −0.01 lumen on $T_2$, determine the change in grid voltage of the triode.

**13-10.** (*a*) A secondary-emission surface in an electron multiplier has $\delta = 7$. Determine the number of stages needed to provide an output current of 1 ma from an emitted photoelectric current of 0.015 $\mu$amp.

(*b*) The above tube is followed by a pentode of $g_m$ = 5000 $\mu$mhos. Compute the over-all transconductance of the combination.

**13-11.** A vacuum cell of 925 type is used with 2000-cycle modulated light, with load $R$ = 4 megohms and combined photocell and amplifier input capacity of 78 $\mu\mu$f. If the triode amplifier has a load of 50,000 ohms, $\mu = 30$, $r_p$ = 45,000, determine the rms light-flux change required to give an amplifier output of 50 volts rms.

**13-12.** The 931-A photomultiplier of Fig. 13-12 is used with a load of 50,000 ohms and anode–last dynode supply of 300 volts. When the input light swings sinusoidally,

$$L = 75 + 25 \sin \omega t$$

with $L$ in $\mu$lumens, determine the output rms voltage across the load.

**13-13.** A certain photocell with an *S*-3 surface gives a current of 5.7 $\mu$amp when irradiated by a tungsten source at 2870°K color temperature. Assuming equal peak sensitivities, calculate the relative current when the same source is used with a phototube with an *S*-4 surface.

**13-14.** A slotted metal plate is mounted 3 in. from the filament of a 32-candlepower point-source bulb. The slot is $\frac{1}{2}$ in. long by $\frac{1}{32}$ in. wide. A 925 phototube receives all the light coming through the slot. The phototube load is 10 megohms, $E_{bb}$ = 150 volts. Compute the change in voltage across the load when the slot is open and closed.

**13-15.** The characteristics of a type 930 gas-filled phototube are given in Fig. 13-4. Such a tube has an anode area of 1.1 sq in. and is operated with $E_{bb}$ = 90 volts. A 50-candlepower lamp is placed 12 in. from the cell and is moved in steps of 2 in. to a distance of 36 in. Plot the output voltage across a 5-megohm load.

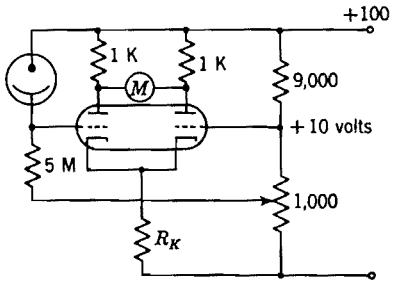



FIG. 13-19

**13-16.** The circuit of Fig. 13-19 is used with a 925 cell and a 12AT7 double triode. (*a*) Specify the cathode resistor if the plate current at balance, or zero meter current, is to be 2.5 ma. (*b*) If the meter is rated 200 $\mu$amp full-scale, 500 ohms resistance, compute the light-flux change to give full-scale meter reading.

## REFERENCES

1. Metcalf, G. F.: Operating Characteristics of Photoelectric Tubes, *Proc. IRE*, **17**: 2064 (1929).
2. Richter, W.: "Fundamentals of Industrial Electronic Circuits," McGraw-Hill Book Company, Inc., New York, 1947.
3. Shepard, F. H.: Some Unconventional Vacuum Tube Applications, *RCA Rev.*, **2**: 149 (1937).
4. Zworykin, V. K., G. A. Morton, and L. Malter: The Secondary-emission Multiplier, *Proc. IRE*, **24**: 351 (1936).
5. Pierce, J. R.: Electron-multiplier Design, *Bell Labs. Record*, **16**: 305 (1938).
6. Henney, Keith, and J. D. Fahnestock: "Electron Tubes in Industry," 3d ed., McGraw-Hill Book Company, Inc., New York, 1952.
7. Sweet, M. H.: Logarithmic Photometer, *Electronics*, vol. 105, November, 1946.
8. Ryder, J. D.: "Electronic Engineering Principles," 2d ed., Prentice-Hall, Inc., Englewood Cliffs, N.J., 1952.
9. Hughes, A. L., and L. A. DuBridge: "Photoelectric Phenomena," McGraw-Hill Book Company, Inc., New York, 1932.

CHAPTER 14

# POWER RECTIFICATION

To provide large amounts of direct current in industry, it is customary to employ the mercury-vapor or gas-filled diode as a rectifier. The gas diode is well suited to industrial application because of its large current capacity and high power efficiency.

Because of special characteristics due to the gas present, and because the electric power is usually supplied at three or more phases, the operating characteristics and circuits differ from those of Chap. 10, where single-phase low-current rectifiers for supply of electronic amplifiers were studied. In industrial applications a filter circuit is rarely needed, owing to tolerance for considerable ripple, and if employed can be simpler than those previously studied, because of the higher ripple frequencies occasioned by polyphase operation.

**14-1. Arc Discharges.** The gaseous arc discharge takes place between an anode and a cathode in an envelope containing an inert gas at pressures ranging from a few microns to a few millimeters of mercury. The anode may be a metal plate or a graphite block, the cathode being of the heated oxide-coated form or a pool of liquid mercury. Gases commonly employed are mercury vapor for large power applications, and argon, xenon, or hydrogen, for special applications where characteristics are not to be affected by temperature.

Only minute negative-space-charge-limited currents flow with anode-cathode potentials much less than the ionization potential of the gas. As was described in Sec. 2-13, at potentials somewhat above the ionization potential of the particular gas a few electrons will obtain sufficient energy to ionize upon collision with a gas atom, the result being formation of positive ions and more free electrons. The positive ions serve to neutralize part of the negative space charge owing to their low mobility and the length of time they are in the space. As a result more electrons acquire ionizing energy; these electrons produce more ions, which further reduce the negative space charge in a cumulative effect.

The process results in an *arc discharge* in the gas, with complete neutralization of the negative space charge by the slow-moving positive ions, with a large current flowing and a small voltage drop. The time from

the first electron-atom collision to complete formation of the arc discharge may be 10 $\mu$sec or less.

The positive ions are attracted to the negative cathode and build up a positive-ion space charge, or *sheath*, around the cathode. Electric-field lines from the negative cathode largely terminate on these positive charges close by, only a few lines penetrating to the anode. This results in an intense field, accelerating to electrons, in a narrow region round the cathode. Since in passing through this sheath an electron, on the average, must be able to acquire ionizing energy to maintain the discharge, the potential drop through the cathode sheath will approximate the ionizing potential of the gas.

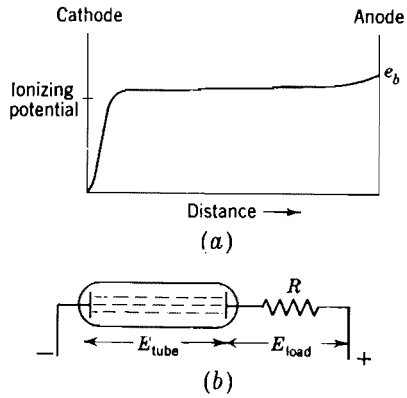



FIG. 14-1. (*a*) Potential distribution across tube in circuit of (*b*).

No electric field can exist between positive charges, and so there will be relatively little field, and consequently only a small potential drop, between the ion sheath and the anode. This region, called the *plasma*, contains about equal numbers of positive ions and negative electrons, all moving randomly with only a slow drift component toward the electrodes, due to the low value of electric field intensity. Only a small additional potential appears across the plasma, so that the total voltage drop across the tube is essentially that of the ionizing potential. The distribution of potential is shown in Fig. 14-1*a* for a typical gas tube.

If the tube voltage tends to fall, ion production and current fall, reducing $E_{load}$ and raising $E_{tube}$ in Fig. 14-1*b*. This increase in tube voltage raises ion production, reduces negative space charge, and raises the current and drives the tube voltage back to normal. This is a sensitive phenomenon, and changes in voltage with changes in current are not large, the tube potential being essentially independent of tube current in the normal operating ranges.

Since no positive ions are present initially, the voltage rises to an initial *breakdown potential*, possibly twice the ionization value, and then falls abruptly to the lower value once ionization is established. In mercury vapor, with an ionizing potential of 10.39 volts, the breakdown value may approximate 15 to 25 volts, with usual operating tube drops in the range of 10 to 20 volts, depending on the type of cathode employed and the directness of the electron and ion paths. In a few cases where ionization is by repeated impacts and steps between successive levels

or metastable states in the atom are small, slightly lower potentials may be encountered.

**14-2. Thermionic Cathodes in Gas Discharges.** The arc in a gas-discharge device can be maintained only by a continuing source of electrons. These electrons may be obtained by thermionic, secondary, or high-field emission.

Most thermionic cathodes used in gas tubes are of the oxide-coated type because of their high relative emission efficiency. When used in a gas discharge, the cathode is under bombardment by the heavy positive gas ions. If the tube drop is sufficiently great, the ionic energies may become large enough to mechanically damage the oxide coating and in extreme cases this coating may be completely blasted from the metal base in a few seconds, destroying the emission abilities of the cathode.

It has been found that the rate of destruction can be reduced and a reasonable oxide-cathode life obtained if the energies of the positive ions, as determined by the tube voltage, can be kept below a certain limiting value. This limiting value for mercury ions has been found to be about 24 volts. Since the mercury ionizing potential is 10.39 volts and normal tube voltages approximate this value under usual gas temperatures and pressures, then mercury-vapor tubes with oxide-coated cathodes will have a reasonable factor of safety and a life of some thousands of hours. Other gases of lower atomic weight can be operated at higher voltages before destructive damage results.

If the tube voltages are below the destructive limit, then ion energies are also below the value at which appreciable secondary emission can occur at the cathode due to ion bombardment, and secondary emission cannot be depended on as an additional source of emitted electrons. In fact, if secondary electrons are obtained, then cathode damage will probably result. Cathodes must be designed with sufficient emitting area to furnish the maximum rated current of the tube at the cathode operating temperature.

If, because of accidental reduction of anode load resistance, the current drawn from the cathode becomes greater than the emitted value, the tube voltage will rise until sufficient additional emission is obtained by secondary emission under positive-ion bombardment. At the same time, however, this bombardment may remove some of the cathode coating.

If the anode potential is applied simultaneously with the cathode-heater potential, the current required by the external anode circuit will be greater than the emitted value during part of the cathode warm-up period. This may result in very serious cathode bombardment and damage, even to visible removal of cathode surface as apparent sparks.

To avoid this damage, hot-cathode gas tubes should be given sufficient time for cathode warm-up before plate potential is applied.

Low ambient temperature will cause low mercury-vapor pressure, and a reduced number of ions in the space owing to the smaller number of gas atoms available for collision. If insufficient ions are present to neutralize the space charge, the tube voltage will rise and the electrons will acquire further energy, making ionization more probable. The positive ions then receive sufficient energy to damage the cathode, so that operation of mercury-vapor hot-cathode tubes in low ambient temperatures must be avoided, even to the extent of adding electrical heaters in some applications. Tubes containing other gases are not radically affected, since their internal gas pressure is determined by Charles's law rather than by the vapor pressure of a liquid.

Relative emission efficiency has been defined in Chap. 2 as milliamperes emitted per watt of heating power. If heat loss by radiation from the cathode can be reduced, heat input to provide a given temperature may also be reduced and the relative emission efficiency increased. Heat shielding or radiation shielding around a cathode will improve the efficiency in this way.

This shielding is not possible in vacuum tubes because any confined spaces tend to build up high negative space-charge values and reduce the emission. However, in a gas tube the positive ions are present to neutralize negative space charge, and no limiting of the current is present. Heat shields are usually made of polished nickel and may take the form of Fig. 14-2, surrounding an indirectly heated cathode. The shields may also be used around directly heated filamentary cathodes and may reduce cathode-heating power by a factor of 10 to 20.

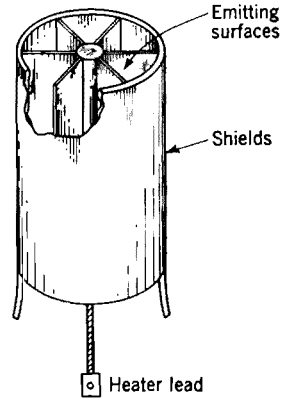



FIG. 14-2. Heat-shielded cathode partly cut away to show internal construction.

**14-3. Mercury-pool Cathodes.** An early form of mercury-vapor rectifier employed a pool of mercury as a cathode, with a side-arm glass bulb as in Fig. 14-3. Electron emission from the mercury pool becomes possible when a *cathode spot* is formed on the cathode-pool surface. The cathode spot shows itself as a bright point of light which moves over the mercury surface, the spot being initiated by a small spark on the mercury. The spark produces the first electrons and positive ions; if one of the anodes is positive, the electrons are attracted and produce additional ions which neutralize any negative space charge and an arc is established.

The exact theory of cathode-spot formation and emission is not known, but Langmuir has suggested that it is a form of high-field emission. The positive-ion cathode sheath is very thin, so that the electric-field intensity therein can reach the magnitude necessary for high-field emission. Consideration of rates of evaporation of mercury from the spot shows that temperatures there are only of the order of 250 to 350°C, too low for emission of thermionic nature. Current densities are very high, and it cannot be conclusively proved that present theories of emission actually apply.

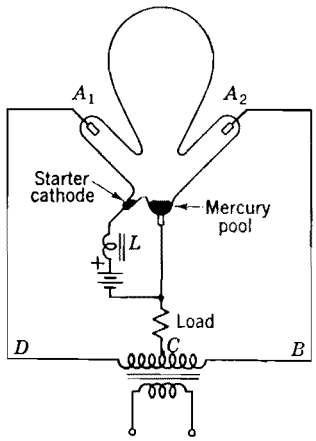



FIG. 14-3. Early mercury-pool rectifier.

The spark for starting the discharge can be produced in several ways. In the old design of Fig. 14-3 it was accomplished by mechanically tipping the tube until the mercury in the main pool and that in the small pool flowed together. Upon righting the tube the inductive circuit was opened when the pools separated, and a spark jumped between the pools, forming the cathode spot. The small mercury pool has now been replaced by a steel plunger operated by an external solenoid magnet to dip into the main mercury pool, the inductive circuit being broken upon release and the first spark being obtained.

Such a mechanical device is slow, and if the spark is not initiated at each half cycle of the input voltage, the arc will not restart and conduction will cease, since the positive ions will diffuse to the anode and walls during the nonconducting period and recombine to form neutral atoms and the space will become *deionized.*

In modern designs a spot is formed by magnetically depressing the starter electrode into the mercury pool, after which an auxiliary anode, continuously supplied with direct current, picks up an arc of a few amperes and maintains a cathode spot and ionization in the space. The main arc then picks up from this small cathode spot, current flowing to whichever anode is most positive. Such a tube is known as an *excitron,* shown in Fig. 14-4. Another form, using an electrical starter, will be discussed in Chap. 15.

The tube drop is usually higher for a tube with a mercury-pool cathode than for a thermionic cathode, because the work-function energy must be supplied by the arc drop rather than from the external cathode heating circuit. Bombardment of a pool cathode by positive ions can do no damage, since more mercury atoms are always available, the evaporated

mercury condensing at the coolest spot on the tube wall. Peak currents are therefore not limited by cathode emission, but by instantaneous heating and magnetic stresses imposed on the leads due to high currents.

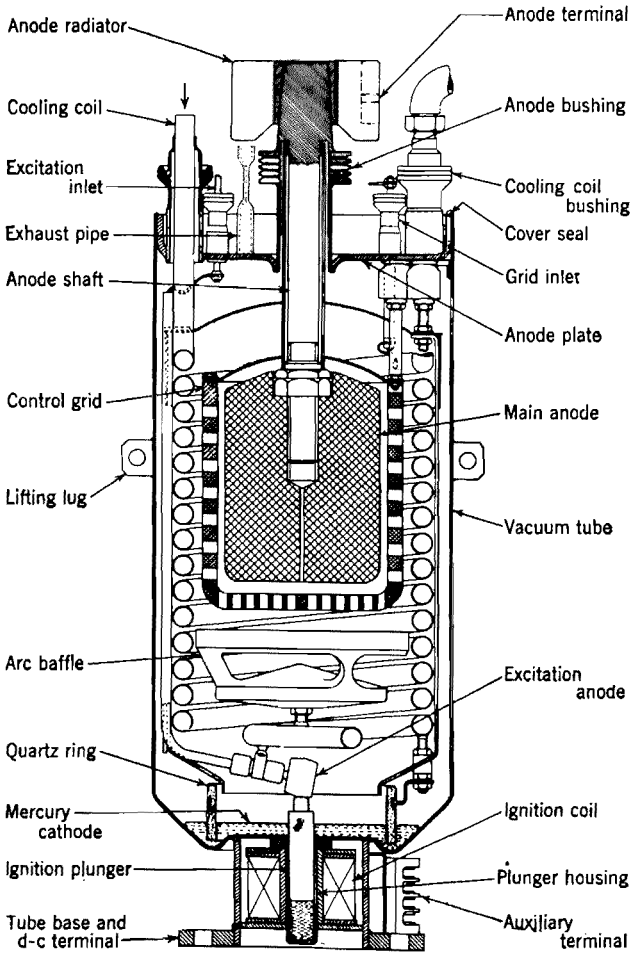



FIG. 14-4. Cross section of excitron rectifier.

High peak currents can be carried, and therefore a tube with a mercury-pool cathode can be operated in more severe service than can the hot-cathode type.

The glass-bulb multiple side-arm tube of Fig. 14-3 was replaced with a form of evacuated steel tank with 6, 12, or 18 anodes for polyphase operation. Modern forms as in Fig. 14-4 employ an individual container for each anode and pool cathode, because of the resulting decreased losses,

easier manufacturing, and more flexible operation and maintenance. These are then employed in multiple banks, as in Fig. 14-5, to provide

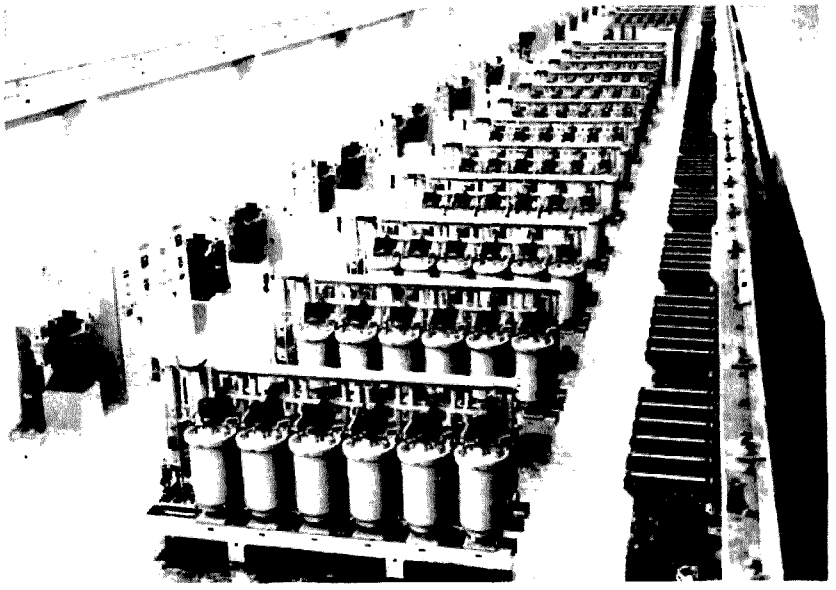


FIG. 14-5. Rectifier room of aluminum-reduction plant containing twenty-four 6000-amp 750-volt d-c rectifier assemblies and control. (*Courtesy of Allis-Chalmers Mfg. Company.*)

polyphase operation and large d-c outputs. The steel-tank form is still frequently used in Europe.

**14-4. Mercury-vapor-tube Characteristics and Ratings.** Any of the inert gases may be used, but because of its ease of handling and its low ionizing potential the most frequently employed gas is mercury vapor, provided by distilling mercury into the evacuated tube during the exhaust process. Unfortunately the vapor pressure will then be a function of temperature, as shown in Fig. 14-6, causing certain of the tube characteristics to vary with the temperature at the point of mercury condensation in the tube.

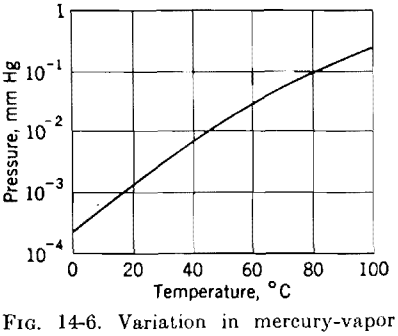



FIG. 14-6. Variation in mercury-vapor pressure as a function of temperature.

The usual operating temperature of the mercury will be in the range of 20 to 60°C, in ambient range, or 30 to 80°C, condensing mercury temperature, resulting in vapor pressures of 0.001 to 0.1 mm Hg and

tube-voltage drops of 10 to 15 volts. The lower temperature limit prevents operation with too few mercury atoms in the space and a high tube drop, possibly above the cathode-disintegration value. The high temperature limit is fixed by frequency of arc-back, a phenomenon discussed in Sec. 14-5. The variation in arc drop as a function of temperature is shown in Fig. 14-7, cathode-disintegration voltages being reached below about 10°C.

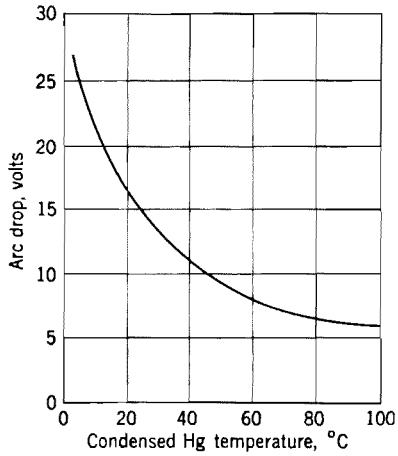



FIG. 14-7. Variation of arc drop with temperature in mercury-vapor diode.

An equation for the circuit of Fig. 14-8, during the positive-anode, or conducting, interval, is

$$e = e_b + i_b R$$

Since $e_b$ is relatively constant, current $i_b$ must increase if $R$ is reduced. If $R$ were zero, the current would increase to a very high value at which the tube would be destroyed. Thus *arc-discharge tubes must always be operated in series with sufficient resistance load to limit the current to a safe value.*

Tube internal heat loss will be shown to be proportional to average rather than effective current, and the *average current rating* is that value which can be carried without excessive heating. Since irregular waveforms and operating cycles may be employed, it is customary for the manufacturer to specify a time over which the current average is to be taken. This time depends on the size and heat-storage capacity of the tube, and may run from 15 to 60 sec or more.

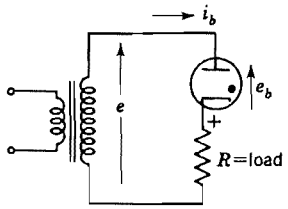


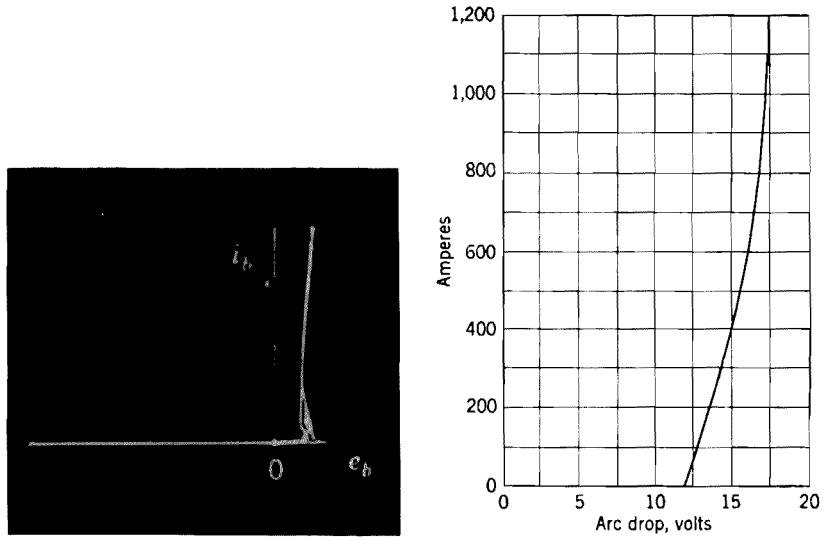

FIG. 14-8. Basic circuit for hot-cathode gas diode; gas filling is indicated by the dot in the symbol.

The *instantaneous peak current* is also an important operating factor. While 100 amp for 1 sec or 10 amp for 10 sec have the same average value, the 100-amp peak may be far above the emission capability of a thermionic cathode, or current capacity of the leads in a pool-cathode tube, and damage may result. This peak current is the maximum permitted instantaneous value.

*Tube drop*, or *arc drop*, is the tube voltage while the tube is conducting. Circuit analysis is usually simplified by assuming the tube drop constant or independent of load current, as supported by Fig. 14-9.

The maximum voltage across the tube during the time the anode is negative to cathode is called the *peak inverse voltage*. The anode has no source of electrons so that the tube cannot conduct unless a gas breakdown occurs. The peak inverse voltage must not exceed the gas sparking potential at the highest expected gas pressure.

Temperature limits on mercury-vapor tubes, usually in the range of 20 to 60°C, ambient temperature, have already been discussed. At low temperatures the tube drop rises, and cathode damage may occur, and




(*a*) (*b*)

FIG. 14-9. (*a*) Arc-drop variation with current for type 866 hot-cathode rectifier. Peak current = 1.2 amp. (*b*) Tube drop vs. current for type 5555 mercury-pool rectifier. Peak current rating = 1200 amp.

at high temperatures arc-back and possible tube damage may become frequent.

**14-5. Arc-back.** In Fig. 14-3 the full transformer secondary voltage is applied between anodes $A_1$ and $A_2$. Because the anodes are of poor emitting material and operate at low temperature, there is normally no source of electrons on either anode. However, if a source of emission should occur on the anode that is negative or the gas between electrodes should break down, then a short circuit between anodes $A_1$ and $A_2$ is placed across the transformer. The only circuit impedance is that of the transformer reactance and resistance since the tube drop remains small with respect to the transformer voltage, and currents of several thousand per cent of normal may flow.

An accidental reverse discharge of this nature is called an *arc-back*. These are random in occurrence, and not an indication of a faulty

rectifier unless they continue at frequent intervals. Frequency of arc-back may be increased by high operating temperature, since this reduces the gas sparking potential, as shown for a particular tube in Fig. 14-10, or by splashing of mercury on the anode surface, causing a momentary cathode spot thereon. Gas may be momentarily released by positive-ion bombardment while the anode is negative, and this appears to increase the probability of arc-back. In hot-cathode tubes cathode material may be sputtered onto an anode, leading to an anode emission source.

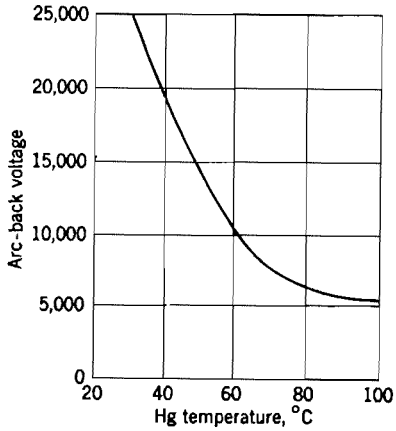



FIG. 14-10. Variation of arc-back voltage with mercury temperature.

If the arc-back is not quickly interrupted, it may overheat and destroy the rectifier. During an arc-back the current through the defective unit is in reverse direction, and this reverse current may be used to trip a single-pole high-speed breaker in that anode circuit, limiting the duration of the fault current. Opening of the main breaker is then not necessary, and only a slight voltage dip occurs while the faulty anode is being cleared and restored to service.

**14-6. Basic Half-wave Gas-diode Rectifier.** As shown in Fig. 14-9*a* and *b*, the arc drop varies only slightly over the rated operational current

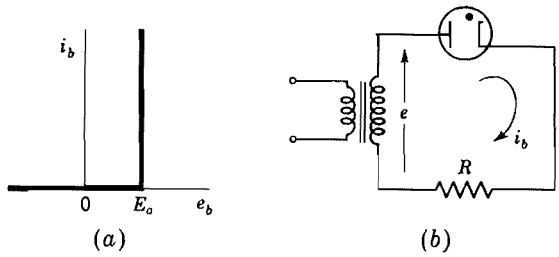



FIG. 14-11. (*a*) Idealized volt-ampere curve for a gas diode. (*b*) Half-wave circuit.

range of a gas diode. It is customary to consider the arc drop constant at $E_o$ volts to simplify circuit analysis. Also, since the breakdown or starting voltage is usually only a few volts higher than the tube drop, no great error will be introduced by assuming that breakdown and tube voltage are equal, if the applied a-c voltage is large with respect to the tube drop. These assumptions cause the actual volt-ampere curve of Fig. 14-9*a* to take the idealized form of Fig. 14-11*a*.

Writing an equation for the circuit of Fig. 14-11*b*, *during the conduction interval*, gives

$$e = i_b R + E_o \tag{14-1}$$

If the input voltage is $e = E_m \sin \omega t$, then the current becomes

$$i_b = \frac{E_m \sin \omega t - E_o}{R} \tag{14-2}$$

A reverse, or negative, current is not possible; therefore $i_b$ must be zero for all negative values of the numerator. Conduction will start and end when the numerator becomes zero, so that the starting angle of conduction is

$$\alpha_1 = \sin^{-1} \frac{E_o}{E_m} \tag{14-3}$$

and conduction will cease at $\alpha_2$, where

$$\alpha_2 = \pi - \sin^{-1} \frac{E_o}{E_m} \tag{14-4}$$

The total conduction angle is $\pi - 2\alpha_1$ under these circumstances. These angles are demonstrated in Fig. 14-12. From observation of Eq. (14-2)

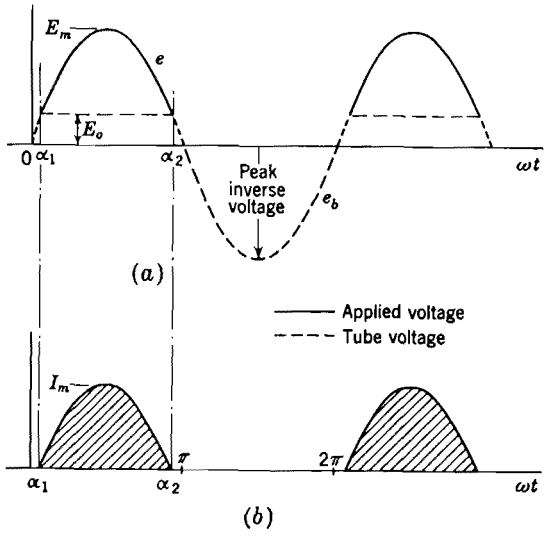



FIG. 14-12. (*a*) Voltage waveform. (*b*) Current waveform.

or the waveform of Fig. 14-12*b* the current waveform is seen as the top of a sinusoid, $E_o/R$ units being subtracted from the base. This is a typical gas-diode waveform, as would be expected from a device with the idealized volt-ampere curve of Fig. 14-11*a*. That practice conforms

well is shown in Fig. 14-13, taken with $E_m = 30$ volts. The constancy of $E_o$ and the slight effect of ignoring the breakdown-voltage pip can be seen.

The dashed curve of Fig. 14-12*a* is the voltage appearing across the tube during the cycle. The voltage rises until the ionization starts; then the voltage falls slightly to the nearly constant $E_o$ value, which approximates 11 volts. At $\omega t = \alpha_2$ the applied voltage becomes less then $E_o$, or the numerator of Eq. (14-2) goes negative, and conduction ceases. The tube then represents an open circuit, and the entire applied voltage appears across the tube, until the voltage again rises positively and exceeds the starting value in the next cycle.

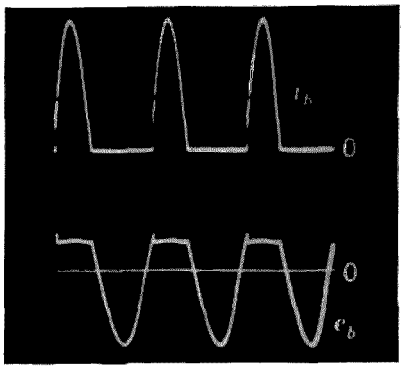



FIG. 14-13. Current and tube voltage waveforms for 872 hot-cathode gas diode. Note small breakdown pip; $E_m = 30$ volts.

A gas diode is ordinarily employed to rectify alternating current into a direct output. The output component of interest is then $I_{dc}$, or the average value of $i_b$ over a cycle. Expressing the current as

$$i_b = \frac{E_m \sin \omega t - E_o}{R} \qquad \alpha_1 < \omega t < \alpha_2 \tag{14-5}$$

$$i_b = 0 \qquad \alpha_2 < \omega t < 2\pi + \alpha_1 \tag{14-6}$$

the average current is obtained as

$$I_{dc} = \frac{1}{2\pi} \int_o^{2\pi} i_b \, d\omega t = \frac{1}{2\pi} \int_{\alpha_1}^{\alpha_2} \frac{E_m \sin \omega t - E_o}{R} \, d\omega t$$

Use of the definitions for $\alpha_1$ and $\alpha_2$ leads to

$$I_{dc} = \frac{E_m}{\pi R} \left[ \sqrt{1 - \left(\frac{E_o}{E_m}\right)^2} - \frac{E_o}{E_m} \cos^{-1} \frac{E_o}{E_m} \right] \tag{14-7}$$

The output d-c voltage across the load is $E_{dc} = I_{dc}R$.

It can be seen that as the applied voltage is increased $E_m$ becomes considerably larger than the $E_o$ value of 10 to 15 volts; consequently $E_o/E_m$ becomes small, and the conduction angle approaches 180°. Both $I_{dc}$ and $E_{dc}$ rise and approach limiting values of

$$I_{dc} = \frac{E_m}{\pi R} \qquad E_{dc} = \frac{E_m}{\pi} \tag{14-8}$$

The ratio $E_{dc}/E_m$ has an upper limit of $1/\pi = 0.318$ for $E_o/E_m$ very small. For many applications with high-voltage rectifiers this condition and

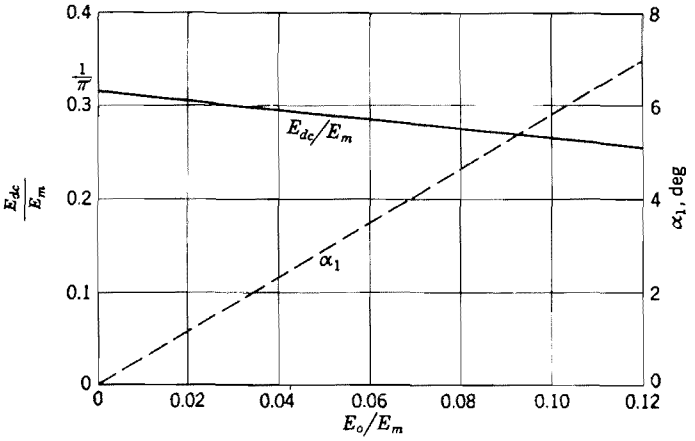



FIG. 14-14. $E_{dc}/E_m$ and $\alpha_1$ as a function of $E_o/E_m$.

$\alpha_1 = 0$ may be assumed to simplify the analysis. The effect of the assumption may be estimated from Fig. 14-14.

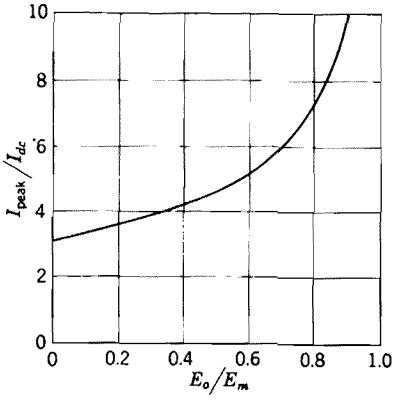



FIG. 14-15. Ratio of peak to average current as a function of $E_o/E_m$ for the half-wave gas diode.

The peak current through the rectifier is a fundamental operational limit, and for small values of $E_m$ the conduction angle may become short with the peak current forced up to supply a given d-c output. For the simple resistive load the tube current is a maximum at $\omega t = \pi/2$. The ratio of the peak current to the average is a function of $E_o/E_m$ and is plotted in Fig. 14-15. Very high ratios of peak to average currents may be reached for short conduction angles.

The power lost in the rectifier tube must be radiated or removed by the cooling water. To calculate the amount of such tube loss,

$$P_T = \frac{1}{2\pi}\int_o^{2\pi} e_b i_b \, d\omega t = \frac{E_o}{2\pi}\int_{\alpha_1}^{\alpha_2} \frac{E_m \sin \omega t - E_o}{R} \, d\omega t \qquad (14\text{-}9)$$

The factor $e_b = E_o$ has been assumed constant and taken outside the integral, which is then recognizable as that which led to the average

current, so that

$$P_p = E_o I_{dc} \tag{14-10}$$

a result independent of current waveform or of angle of conduction, and of considerable practical utility. For instance, the usual voltmeters cannot be used to read $E_o$ in an operating rectifier. However, by noting that

$$E_o = \frac{P_p}{I_{dc}} \tag{14-11}$$

it is possible to use a wattmeter connected to read tube loss, from which the arc drop may be determined in operating rectifiers. Another method involves averaging measurements taken from oscillograms and lacks the speed of the above method.

The average power input to the rectifier circuit may be calculated from the basic power relation as

$$P_{ac} = \frac{1}{2\pi}\int_o^{2\pi} e i_b \, d\omega t = \frac{1}{2\pi}\int_{\alpha_1}^{\alpha_2} E_m \sin \omega t \frac{E_m \sin \omega t - E_o}{R} \, d\omega t$$

After integration and the use of certain trigonometric identities the a-c power input becomes

$$P_{ac} = \frac{E_m^2}{2\pi R}\left[\cos^{-1}\frac{E_o}{E_m} - \frac{E_o}{E_m}\sqrt{1 - \left(\frac{E_o}{E_m}\right)^2}\right] \tag{14-12}$$

For high-voltage rectifiers where $E_o/E_m = 0$ the power input is

$$P_{ac} \cong \frac{E_m^2}{4R}$$

The efficiency with which the rectifier converts a-c power to d-c power has already been called the efficiency of rectification, and, using Eqs. (14-7) and (14-12),

$$\eta_R = \frac{P_{dc}}{P_{ac}} = \frac{2[\sqrt{1 - (E_o/E_m)^2} - (E_o/E_m)\cos^{-1}(E_o/E_m)]}{\pi[\cos^{-1}(E_o/E_m) - (E_o/E_m)\sqrt{1 - (E_o/E_m)^2}]} \times 100\% \tag{14-13}$$

The rectification efficiency of a half-wave gas-tube rectifier with resistive load is seen to be independent of the load current, being only a function of the $E_o/E_m$ ratio, and approaching the maximum value of 40.6 per cent when $E_o/E_m$ is negligible. Use of inductance or capacitance as filters will alter the current waveform and the efficiency.

Rectifiers of this nature are frequently used for the charging of batteries or in electrochemical processes where the load includes a counter emf. If the fundamental processes of writing a mesh equation and determining

$\alpha_1$ and $\alpha_2$ as the angles at which the current expression goes to zero are followed, then equations for such operation can be obtained. In such cases the resistance of the battery is made part of the load $R$, and the counter emf adds to $E_o$ of the tube.

The d-c load current flows through the transformer secondary and may lead to d-c saturation of the iron. Because of the expense of

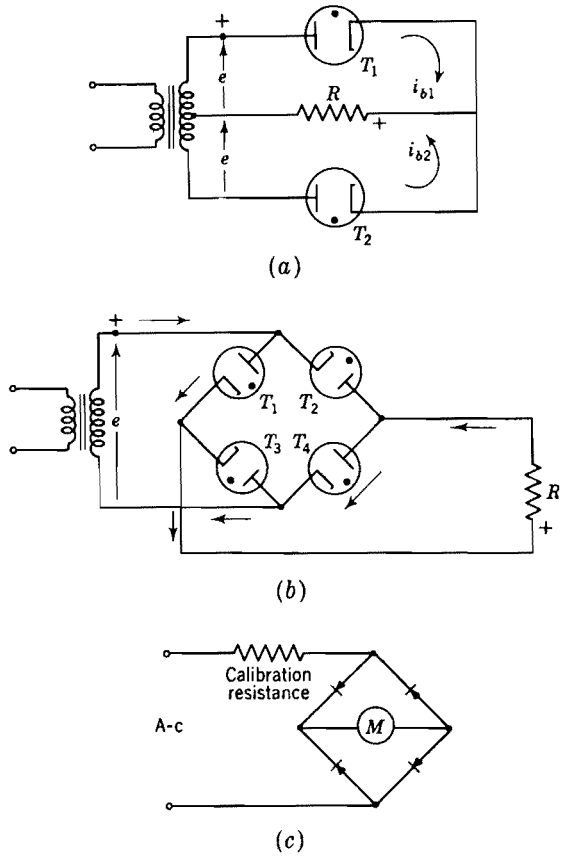



FIG. 14-16. (*a*) Full-wave diametric circuit. (*b*) Bridge rectifier. (*c*) Bridge rectifier as adapted for instrument use.

special transformers with air gaps, the half-wave circuit is not ordinarily employed for power purposes but has been presented here as a basic circuit form, to demonstrate methods which will be applied in the analysis of more complex circuits.

**14-7. Full-wave and Bridge Rectifiers.** The methods of the previous section may be followed in analysis of the *full-wave diametric rectifier* of Fig. 14-16*a*. Here each tube conducts for an angle of $\pi - 2\alpha_1$, approxi-

mating 180°, and the load current comprises a series of pulses as in Fig. 14-12*b* but with a second pulse occurring in the $\pi$ to $2\pi$ interval. The analysis will give expressions for $I_{dc}$ and $E_{dc}$ exactly twice those for the half-wave circuit. In addition it may be found that the peak inverse voltage equals the peak of the full transformer-secondary voltage, or $2E_m$, where $e = E_m \sin \omega t$.

Since the d-c tube currents flow in opposite directions in the transformer secondary, no saturation of the transformer core will result. The two halves of the secondary do not carry sinusoidal currents, and the transformer will not be so efficiently used as if it were supplying a-c continuously; for a given power output this makes a larger transformer necessary.

The *bridge rectifier* allows full-wave rectification from a transformer without a center tap. During the half cycle in which the upper transformer terminal is positive, diodes $T_1$ and $T_4$ have positive anodes and conduct in series as shown by the arrows. Tubes $T_2$ and $T_3$ have negative anodes and are nonconducting. On the next half cycle, tubes $T_2$ and $T_3$ conduct, causing another load-current pulse in the same direction as for the first half cycle. The circuit is a full-wave rectifier.

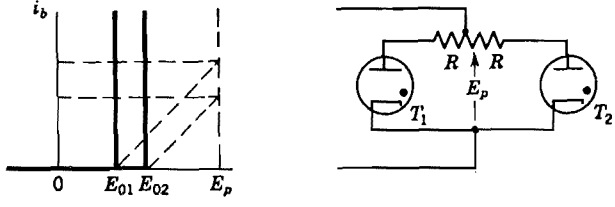
Since the bridge circuit has a tube drop equal to that of two tubes, it is most suited to high voltages, where the tube drop can be neglected. It also permits use of a standard untapped transformer, giving the required d-c voltage with only half the secondary turns needed in the full-wave diametric circuit. Currents flow continuously in the secondary, and a transformer need be only two-thirds as large as one for the full-wave circuit of Fig. 14-16*a*. In addition the peak inverse rating of the tubes need be only half that in the circuit in Fig. 14-16*a*, for equal d-c output voltage.

Rectifier-type a-c instruments usually employ the bridge circuit with copper-oxide rectifier elements, as in Fig. 14-16*c*. The resultant direct current is indicated by a sensitive galvanometer or microammeter of the D'Arsonval type. These instruments are usually satisfactory up to a few thousand cycles. Since they have scales marked in terms of the rms value of a sine wave but actually measure the average value, their indications may be in error if the applied waveform differs from the sine form.

**14-8. Parallel Operation of Gas Diodes.** Figure 14-17*a* shows the ideal volt-ampere curves for two gas diodes, with tube drops $E_{o1}$ and $E_{o2}$. If these tubes are connected in parallel, the voltage at the beginning of a cycle will rise until it reaches the starting value $E_{o1}$, at which time the tube conducts. The voltage across the parallel combination can never rise above $E_{o1}$, and tube $T_2$ never will conduct, the tube with the lowest tube drop (actually the lowest starting voltage) carrying all the load.

If a small resistor of value $R$ is connected in series with each anode as in Fig. 14-17*b*, the individual volt-ampere characteristics change to the sloping dashed lines of Fig. 14-17*a*. At some voltage $E_p$, greater than the starting voltage of either tube, tube $T_2$ will start and conduct, sharing the load at the common supply voltage $E_p$ with currents $i_1$ and $i_2$, respectively.

Tube drops may differ as much as 8 to 10 volts between new tubes and those of lengthy service, and if resistor $R$ can produce this drop at the desired tube current, then operation will be satisfactory. A small center-tapped reactor may also be used in place of the resistor. It is usually more economical to purchase one large tube than two smaller ones, so that parallel operation is not considered desirable.




(*a*) (*b*)

FIG. 14-17. (*a*) Ideal volt-ampere curves for two gas diodes in parallel. (*b*) Circuit for parallel operation of gas diodes.

**14-9. Three-Phase Half-wave Rectifier Relations.** Most power rectifiers employ some modification of a star-connected transformer secondary, with a diode connected to each point of the star. Usually the circuits have a common cathode connection and are therefore suited to the individual diode unit employing thermionic or mercury-pool cathodes or to the older form of steel-tank rectifier which used a number of anodes and one common mercury-pool cathode.

In Fig. 14-18*a* is shown a simple three-phase half-wave circuit as an example of the type under discussion. The load $R$ is connected between common cathode and the transformer neutral; it is customary to use the transformer neutral for voltage reference. The effects of transformer reactance will be neglected here and discussed in a later section; sinusoidal voltages to neutral as in Fig. 14-18*b* will be assumed. With negligible transformer resistance and reactance a circuit equation around phase 1 of the rectifier during its conducting interval is

$$e_{o1} - E_o - e_k = 0$$

where $e_k$ is the instantaneous cathode potential above transformer neutral. Then

$$e_k = e_{o1} - E_o \tag{14-14}$$

so that the cathode of $A_1$ is only 10 to 15 volts lower in potential than the anode of $A_1$, and this is indicated by the dashed line of Fig. 14-18*b*. Anode $A_2$ is negative with respect to cathode and unable to conduct

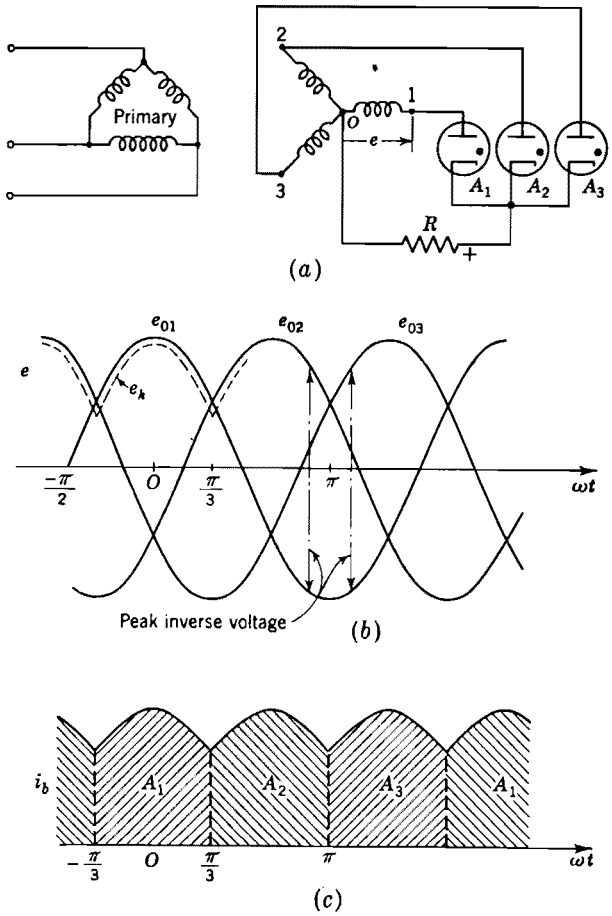



FIG. 14-18. (*a*) Three-phase half-wave rectifier. (*b*) Voltages to neutral. (*c*) Load current.

until its voltage exceeds $e_k$ by the amount of its tube drop. Anode $A_2$ may conduct when

$$e_{o2} - E_o \geq e_k = e_{o1} - E_o$$

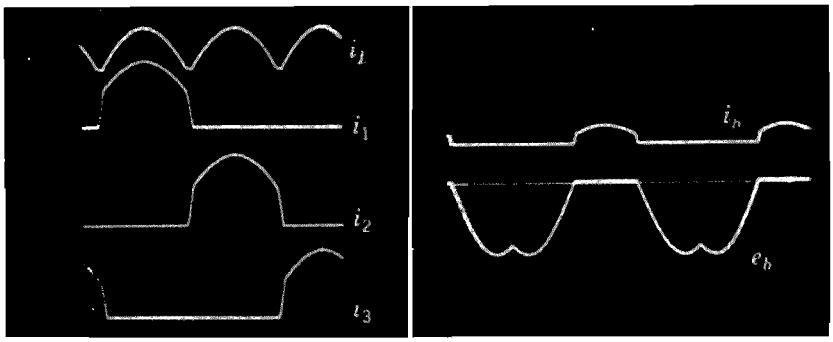
With equal tube drops this expression allows anode $A_2$ to conduct when

$$e_{o2} \geq e_{o1} \tag{14-15}$$

which occurs at $\pi/3$ or 60° with respect to the representation of $e_{o1}$ as a

cosine wave. This choice of the cosine is made because of symmetry which will later come into integral limits.

Since at $\pi/3$ the voltage $e_{o1}$ is dropping in value, anode $A_1$ ceases conduction and the arc is said to have transferred, or *commutated*, to $A_2$. At $\omega t = \pi$ the arc transfers from $A_2$ to $A_3$, each anode conducting for 120° of the cycle, current flowing to the most positive anode at any instant. The current waveform per anode is shown diagrammatically in Fig. 14-18c. In practice these current pulses take the form of Fig. 14-19, the slope of the first and last portions of each current pulse being due to the neglected transformer reactance.




(*a*) (*b*)

FIG. 14-19. Three-phase half-wave rectifier. (*a*) Load and individual tube currents; (*b*) Tube-current and voltage waves.

The three-phase half-wave circuit may be analyzed for performance data. The equation around mesh 1 is

$$E_m \cos \omega t - E_o - i_b R = 0 \tag{14-16}$$

To simplify analysis, the tube drop will be neglected, as well as the transformer reactance and resistance, since these effects are small. The anode current is then

$$i_b = \frac{E_m \cos \omega t}{R} \qquad -\frac{\pi}{3} < \omega t < \frac{\pi}{3} \tag{14-17}$$

The direct load current will be three times the average current of one anode, or

$$I_{dc} = \frac{3}{2\pi} \int_{-\pi/3}^{\pi/3} \frac{E_m \cos \omega t}{R} \, d\omega t = \frac{0.827 E_m}{R} \tag{14-18}$$

Writing

$$E_{dc} = I_{do} R = 0.827 E_m = 1.17 E_{rms}$$

provides a relation between d-c output voltage and the a-c transformer voltage required to produce it, on the ideal basis of zero tube drop and zero transformer reactance.

The ripple will be due to three pulses per cycle, or will have a fundamental frequency three times the supply. The ripple magnitude may be found from Eq. (10-18) by computing $I_{rms}$ as

$$I_{rms} = \sqrt{\frac{3}{2\pi}\int_{-\pi/3}^{\pi/3}\frac{E_m^2\cos^2\omega t}{R^2}\,d\omega t} = \frac{0.838E_m}{R} \tag{14-19}$$

The ripple then is

$$\text{Ripple} = 100\%\sqrt{\left(\frac{I_{rms}}{I_{dc}}\right)^2 - 1} = 100\%\sqrt{\left(\frac{0.838}{0.827}\right)^2 - 1}$$
$$= 17\%$$

The peak inverse voltage will occur when the maximum voltage appears across a tube while nonconducting. When $A_1$ is nonconducting, the voltage across it is $e_{o1} - e_{o2}$ or $e_{o1} - e_{o3}$, considering each as a rise from neutral. Then

$$E_m\cos\omega t - E_m\cos\left(\omega t + \frac{\pi}{3}\right) = e_b \tag{14-20}$$

Taking the derivative and setting it equal to zero,

$$1.5\sin\omega t_1 = 0.866\cos\omega t_1$$

from which

$$\frac{\sin\omega t_1}{\cos\omega t_1} = \tan\omega t_1 = 0.577$$

The maximum inverse voltage occurs at

$$\omega t_1 = 60^\circ + \frac{n\pi}{2} \qquad n = \text{odd} \tag{14-21}$$

Therefore the peak inverse voltage occurs at 150°, or $5\pi/6$ of the $e_{o1}$ cycle, as shown in Figs. 14-18*b* and 14-19*b*. The value of the peak inverse voltage for the three-phase half-wave circuit is then

$$\text{PIV} = \sqrt{3}\,E_m = 2.09E_{dc}$$

The circuit as discussed suffers from possible d-c saturation of the transformer core, since the d-c component flows through each transformer secondary. Modifications of the circuit, as discussed in Sec. 14-16, are more usually employed. The above analysis has demonstrated certain fundamental principles of polyphase-rectifier study and can be readily adapted to more complex circuits studied later, by proper interpretation of phase and winding voltages and currents.

**14-10. *m*-Phase Rectifier Relations.** In practice it is frequently found that optimum rectification conditions or costs are obtained with 6, 12, or more phases. For a general study of polyphase rectifiers it is con-

venient to use a star-connected transformer of $m$ phases, connected to $m$ anodes, each anode conducting for an interval of $360/m°$, or $2\pi/m$ radians per cycle. The load-current waveform for resistance load will appear as in Fig. 14-20, utilizing a cosine wave because of the symmetry obtainable. It should be noted that the assumption of the angles of conduction rules out the single-phase half-wave rectifier, since therein load-current flow is not continuous.

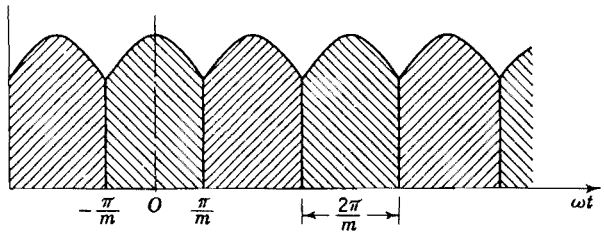



FIG. 14-20. Current pulses in the resistance load of the $m$-phase rectifier.

Using the assumptions of negligible tube drop and zero transformer reactance and resistance, a general expression can be written for $I_{dc}$ in the resistance load as

$$I_{dc} = \frac{m}{2\pi} \int_{-\pi/m}^{\pi/m} \frac{E_m \cos \omega t}{R} d\omega t = \frac{E_m}{R} \frac{m}{\pi} \sin \frac{\pi}{m} \tag{14-22}$$

where $E_m$ is the peak anode to neutral voltage of the transformer. Since $E_{dc} = I_{dc}R$, then

$$\frac{E_{dc}}{E_m} = \frac{m}{\pi} \sin \frac{\pi}{m} \tag{14-23}$$

and values of this dimensionless ratio can be computed in terms of $m$ as in the accompanying tabulation, showing that $E_{dc}$ increases with number of phases, although the increase is not great above $m = 6$.

| $m$ | 2 | 3 | 4 | 6 | 12 | $\infty$ |
|---|---|---|---|---|---|---|
| $\frac{E_{dc}}{E_m}$ | 0.636 | 0.826 | 0.896 | 0.955 | 0.999 | 1.00 |

The rms current per diode is

$$I_{rms} = \sqrt{\frac{1}{2\pi} \int_{-\pi/m}^{\pi/m} \frac{E_m^2 \cos^2 \omega t}{R^2} d\omega t}$$

$$= \frac{E_m}{R} \sqrt{\frac{1}{2\pi} \left( \frac{\pi}{m} + \sin \frac{\pi}{m} \cos \frac{\pi}{m} \right)} \tag{14-24}$$

and the rms value of current in the load having $m$ pulses per cycle is $\sqrt{m}$ times the above.

The ripple in the load current may be calculated from

$$\text{Ripple} = \sqrt{\left(\frac{I_{rms}}{I_{dc}}\right)^2 - 1} \times 100\%$$

$$= 100\% \sqrt{\frac{(m/2\pi)[(\pi/m) + \sin(\pi/m)\cos(\pi/m)]}{[(m/\pi)\sin(\pi/m)]^2} - 1} \quad (14\text{-}25)$$

Values of ripple computed from the above are as in the accompanying

| $m$ | 2 | 3 | 4 | 6 | 12 | $\infty$ |
|---|---|---|---|---|---|---|
| Ripple | 0.47 | 0.17 | 0.089 | 0.04 | 0.014 | 0 |
| Ripple frequency for 60-cycle supply | 120 | 180 | 240 | 360 | 720 | $\infty$ |

tabulation. The ripple decreases rapidly with the number of phases, and this with the higher ripple frequency makes the output of rectifiers of six or more phases easy to filter. Distortion of the primary-current waveform is also less, and this is important, since harmonics introduced into the primary may cause interference in adjacent telephone lines.

Because use of large numbers of phases reduces ripple, raises d-c output voltage, and increases the load current, rectifiers of 12, 18, 24, or more phases are employed, particularly to supply electrolytic processes, as in aluminum reduction.

**14-11. Transformer Utility Factor.** Transformers which supply rectifiers usually do not carry sinusoidal currents in their windings or may carry current in various coils only over portions of a cycle. The harmonic components of these distorted waves contribute to transformer heating but produce no useful d-c output. To provide for these harmonic currents, a transformer rating in watts greater than the contemplated d-c power output is required; that is, the harmonic currents increase the volt-amperes without adding to the output d-c watts.

The *utility factor* of a rectifier transformer is given by the ratio of the d-c power output of the winding to the volt-amperes carried; the factor rates the efficiency with which a given winding is being used. The utility factor is a function of waveform and the ratio of rms to average voltage. For instance, the utility factor of a secondary winding used in a single-phase diametric rectifier is low because of the half-sinusoid secondary current. It has the value 0.574 for that circuit. In the single-phase bridge circuit with sinusoidal secondary current the secondary utility factor has the value 0.813 with resistance rectifier load.

It is possible to find $m$-phase general expressions for secondary and primary utility factors. For a rectifier such as the delta-star six-phase circuit of Fig. 14-21 the volt-amperes per secondary phase are obtainable

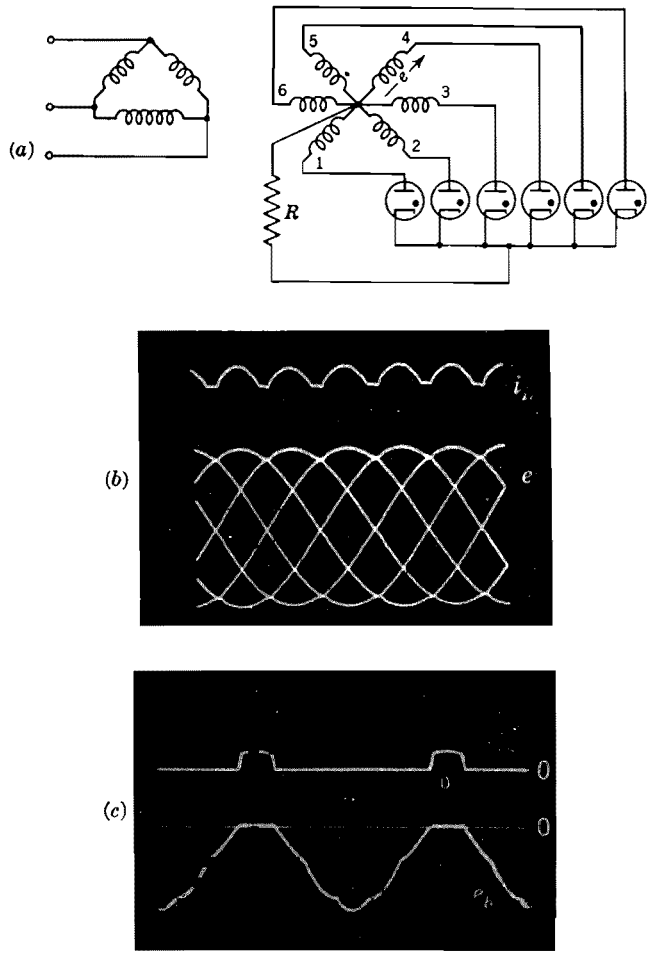



FIG. 14-21. (*a*) Six-phase rectifier. (*b*) Current in load and voltages to neutral. (*c*) Individual tube current and tube voltage.

by use of Eq. (14-24) as

$$\text{Total SVA} = m \frac{E_m}{\sqrt{2}} \frac{E_m}{R} \sqrt{\frac{1}{2\pi}\left(\frac{\pi}{m} + \sin\frac{\pi}{m}\cos\frac{\pi}{m}\right)} \qquad (14\text{-}26)$$

By use of $I_{dc}^2 R$ the d-c power output is obtained from Eq. (14-22) as

$$P_{dc} = \frac{E_m^2}{R}\left(\frac{m}{\pi}\right)^2 \sin^2\frac{\pi}{m} \qquad (14\text{-}27)$$

so that the secondary utility factor is

$$\text{SUF} = \frac{2[(m/\pi)\sin^2(\pi/m)]}{\sqrt{\pi[(\pi/m) + \sin(\pi/m)\cos(\pi/m)]}} \tag{14-28}$$

Calculation of secondary utility factor as a function of $\pi/m$, or the half conduction angle, gives for the resistance-loaded rectifier the values in the accompanying tabulation. The maximum appears near $2\pi/m = 120°$,

| $m$..... | 2 | 3 | 4 | 6 | 12 | 24 |
|---|---|---|---|---|---|---|
| $2\pi/m$, deg.. | 180 | 120 | 90 | 60 | 30 | 15 |
| SUF.... .. | 0.57 | 0.675 | 0.636 | 0.551 | 0.399 | 0.286 |

or $m = 3$. Because of this, circuits employing adaptations of the three-phase half-wave circuit to give 120° conduction angles are frequently used, but with modifications to remove the possibility of core saturation of the transformer. These are the three-phase zigzag, the three-phase double-Y, and other branched circuits to be discussed later.

For star-connected circuits where the d-c saturation is balanced out, as in the circuit of Fig. 14-21*a*, two secondary phases are supplied by a single primary phase. The effective current in a primary winding is the effective value of two current pulses so that

$$\text{Primary } I_{\text{rms}} = \sqrt{2}\,\frac{E_m}{R}\sqrt{\frac{1}{2\pi}\left(\frac{\pi}{m} + \sin\frac{\pi}{m}\cos\frac{\pi}{m}\right)} \tag{14-29}$$

assuming 1:1 voltage transformation. If the circuit employs $p$ primary phases, and $E = E_m/\sqrt{2}$, then the primary volt-amperes will be

$$\text{PVA} = \frac{pE_m^2}{R}\sqrt{\frac{1}{2\pi}\left(\frac{\pi}{m} + \sin\frac{\pi}{m}\cos\frac{\pi}{m}\right)} \tag{14-30}$$

The primary utility factor is then

$$\text{PUF} = \frac{(m/\pi)^2\sin^2(\pi/m)}{p\sqrt{(1/2\pi)[(\pi/m) + \sin(\pi/m)\cos(\pi/m)]}}$$

$$= \frac{\sqrt{2}\,m}{p}\left\{\frac{(m/\pi)\sin^2(\pi/m)}{\sqrt{\pi[\pi/m + \sin(\pi/m)\cos(\pi/m)]}}\right\} \tag{14-31}$$

$$= \frac{m}{\sqrt{2}\,p}\,\text{SUF} \tag{14-32}$$

The factor $1/\sqrt{2}$ is present because of the use of one primary phase to supply two secondary phases at an angle of 180°. For circuits of the form assumed Table 29 represents a summary of performance data, including primary and secondary utility factors.

TABLE 29. RESISTANCE-LOADED-RECTIFIER SUMMARY

| | $1\phi$ Full-wave | $3\phi$ Half-wave | $3\phi$ Bridge | $6\phi$ Double-Y | $6\phi$ Half-wave |
|---|---|---|---|---|---|
| $m$ | 2 | 3 | 6 | 6 | 6 |
| $I_{dc}$ | $0.636E_m/R$ | $0.827E_m/R$ | $1.65E_m/R$ | $1.65E_m/R$ | $0.955E_m/R$ |
| Peak current | $1.57I_{dc}$ | $1.21I_{dc}$ | $1.21I_{dc}$ | $0.60I_{dc}$ | $1.05I_{dc}$ |
| $E_{dc}$ | $0.636E_m$ | $0.827E_m$ | $1.65E_m$ | $1.65E_m$ | $0.955E_m$ |
| Ripple % | 47 | 17 | 4 | 4 | 4 |
| PVA / $P_{dc}$ | 1.23 | 1.23 | 1.05 | 1.05 | 1.28 |
| PUF | 0.81 | 0.83 | 0.95 | 0.95 | 0.78 |
| SVA / $P_{dc}$ | 1.75 | 1.50 | 1.05 | 1.48 | 1.82 |
| SUF | 0.57 | 0.67 | 0.95 | 0.67 | 0.55 |
| Theoretical maximum efficiency, % | 81.2 | 96.5 | 99.8 | 99.8 | 99.8 |
| PIV | $3.14E_{dc}$ | $2.09E_{dc}$ | $1.05E_{dc}$ | $2.09E_{dc}$ | $2.09E_{dc}$ |

**14-12. $m$-Phase Rectifiers with Inductance Filters.** To reduce ripple, it is common practice to connect an inductor in series with the d-c load as a filter element. Because of the high ripple frequencies and the small ripple values of polyphase rectifiers, nominal values of inductance give

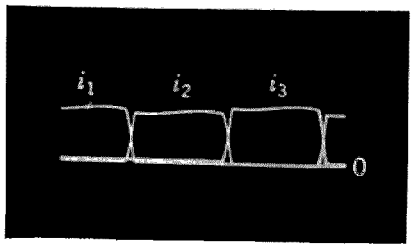



FIG. 14-22. Current pulses in load of three-phase half-wave rectifier with large series inductor.

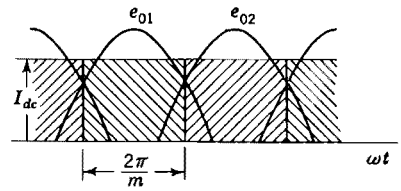



FIG. 14-23. Current pulses in an $m$-phase rectifier with inductive load.

good filtering. The flattening of the top of the current pulses with a practical inductor is apparent in Fig. 14-22, where all load pulses are shown on the same zero axis. The overlap of the waves at the end of tube conduction will be explained in Sec. 14-13.

It can be seen that the tube currents with inductance present may be assumed constant and of value $I_{dc}$ during the conduction interval, as shown in Fig. 14-23, neglecting transformer reactance. The conduction angle per tube will again be $2\pi/m$ radians per cycle.

In the inductively loaded $m$-phase rectifier it can be seen that the average voltage will be

$$E_{dc} = \frac{m}{2\pi} \int_{-\pi/m}^{\pi/m} E_m \cos \omega t \, d\omega t = E_m \frac{m}{\pi} \sin \frac{\pi}{m} \tag{14-33}$$

The instantaneous value of current in any diode is $i_b = I_{dc}$, so that the a-c power input from the $m$ transformer secondaries is

$$P_{ac} = \frac{m}{2\pi} \int_{-\pi/m}^{\pi/m} I_{dc} E_m \cos \omega t \, d\omega t = I_{dc} E_m \frac{m}{\pi} \sin \frac{\pi}{m} \tag{14-34}$$

The theoretical rectification efficiency is then obtainable as

$$\eta_R = \frac{P_{dc}}{P_{ac}} = \frac{I_{dc} E_m (m/\pi) \sin (\pi/m)}{I_{dc} E_m (m/\pi) \sin (\pi/m)} \times 100\% = 100\% \tag{14-35}$$

The presence of the inductance has raised the possible rectification efficiency to 100 per cent. This is not surprising since a constant load current implies zero harmonic losses in the resistive load.

The effective current per diode is

$$I_{rms} = \sqrt{\frac{1}{2\pi} \int_{-\pi/m}^{\pi/m} I_{dc}^2 \, d\omega t} = \frac{I_{dc}}{\sqrt{m}} \tag{14-36}$$

The rms value of the load current can be obtained by multiplying the individual diode current by $\sqrt{m}$. This shows that $I_{rms} = I_{dc}$, which is correct for the steady load current.

The a-c volt-ampere output from the secondary is

$$\text{SVA} = \left( \frac{E_m}{\sqrt{2}} \frac{I_{dc}}{\sqrt{m}} \right) m = \sqrt{\frac{m}{2}} E_m I_{dc} \tag{14-37}$$

from which the secondary utility factor may be computed as

$$\text{SUF} = \frac{P_{dc}}{\text{VA}} = \frac{I_{dc} E_m (m/\pi) \sin (\pi/m)}{\sqrt{\frac{m}{2}} E_m I_{dc}}$$

$$= \frac{\sqrt{2m}}{\pi} \sin \frac{\pi}{m} \tag{14-38}$$

This expression may be readily maximized with respect to $m$, showing that the maximum secondary utility factor occurs at $m = 2.7$, between the biphase case and the three-phase condition. Therefore three-phase operation, or *a circuit which provides a current conduction angle of 120°*, will be near optimum in utilization of the transformer-secondary windings.

For circuits in which the d-c components cancel in a transformer secondary, so that d-c saturation is avoided, and in which the primary current per phase consists of two pulses, the primary phase current is

$$\text{Primary } I_{rms} = \frac{\sqrt{2}\, I_{dc}}{\sqrt{m}} \tag{14-39}$$

assuming 1:1 voltage transformation. The total primary volt-amperes are then obtained by multiplying the primary-phase volt-amperes by $p$, the number of primary phases. Then

$$\text{PVA} = \frac{pE_m}{\sqrt{2}} \frac{\sqrt{2}\, I_{dc}}{\sqrt{m}} = \frac{p}{\sqrt{m}} E_m I_{dc} \tag{14-40}$$

and the primary utility factor is

$$\text{PUF} = \frac{(m/\pi) \sin (\pi/m)}{p/\sqrt{m}} = \frac{m\sqrt{m}}{\pi p} \sin \frac{\pi}{m} \tag{14-41}$$

which shows that the primary utility factor is related to the secondary utility factor as

$$\text{PUF} = \frac{m}{\sqrt{2}\, p} \text{SUF}$$

as it was for the resistance-loaded rectifier.

TABLE 30. COMPARISON OF UTILITY FACTORS

| | 1$\phi$ Full-wave | 3$\phi$ Half-wave | 3$\phi$ Bridge | 6$\phi$ Double-Y | 6$\phi$ Half-wave |
|---|---|---|---|---|---|
| | $m$ | | | | |
| | 2 | 3 | 3 | 6 | 6 |
| PUF: | | | | | |
| $R$ load | 0.81 | 0.83 | 0.95 | 0.95 | 0.78 |
| $L$ load* | 0.90 | 0.83 | 0.95 | 0.95 | 0.78 |
| SUF: | | | | | |
| $R$ load | 0.57 | 0.67 | 0.95 | 0.67 | 0.55 |
| $L$ load* | 0.64 | 0.67 | 0.95 | 0.67 | 0.55 |

* Theoretically infinite inductance.

It may be seen from both Tables 29 and 30 that, for values of $m = 3$ or more, there is negligible difference in performance and transformer ratings between rectifiers with resistance and inductance loads. When

$m = 3$ or more, the load-current waveforms are close to the rectangular ideal without any added filter inductance.

Because of the rectangular form of current it is possible to calculate the primary utility factor for the three-phase half-wave circuit with inductive load. The secondary current $I_{dc}$ of Fig. 14-24*a* flows for 120°, and the corresponding primary phase also carries load for 120°. The primary-current magnitude during load is

$$I_1 = \tfrac{2}{3} I_{dc}$$

assuming a 1:1 voltage ratio. During the remaining two-thirds of the cycle the current in the primary has the value

$$I_1 = -\tfrac{1}{3} I_{dc}$$

because the average current in the primary must be zero.

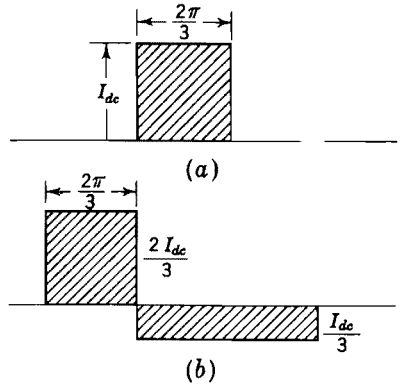



FIG. 14-24. Secondary (*a*) and primary (*b*) phase currents in three-phase half-wave rectifier with inductive load.

The rms primary current then is

$$I_{rms} = \sqrt{\frac{1}{2\pi}\left[\int_0^{2\pi/3}\left(\frac{2I_{dc}}{3}\right)^2 d\omega t + \int_{2\pi/3}^{2\pi}\left(\frac{I_{dc}}{3}\right)^2 d\omega t\right]} = \frac{\sqrt{2}\, I_{dc}}{3} \qquad (14\text{-}42)$$

The primary volt-amperes are

$$\text{PVA} = \frac{3E_m}{\sqrt{2}} \frac{\sqrt{2}\, I_{dc}}{3} = E_m I_{dc} \qquad (14\text{-}43)$$

From Eq. (14-33) $E_m = 1.21 E_{dc}$ for $m = 3$ so that

$$\text{PVA} = 1.21 E_{dc} I_{dc}$$

The primary utility factor then is $1/1.21 = 0.827$ for the three-phase half-wave circuit with inductive load. This analysis illustrates the manner in which primary waveforms may be built up from the various secondary-phase contributions and the primary rms value found.

**14-13. Effect of Transformer Reactance.** The leakage reactance of rectifier transformers has been neglected to this point, and it has been assumed that the current commutates instantly, or that current to one anode ceases as current to another begins. Owing to the presence of the leakage inductances $L_s$ in Fig. 14-25 current starts to a second anode before the first-anode current ceases, because of the inductive emf gener-

ated in $L_s$. The total current during the overlap interval will tend to remain constant, so that one tube picks up load as the other tube drops it, and this overlap is well illustrated by the oscillogram of Fig. 14-26.

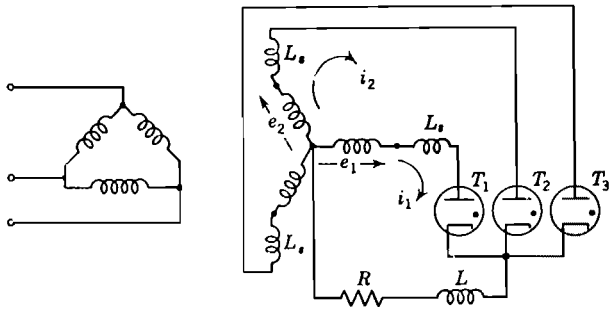



FIG. 14-25. Circuit showing equivalent transformer leakage inductances $L_s$.

While the current commutates between the two anodes during the overlap interval, the voltage across the leakage reactance subtracts from the transformer voltage and a lower average load voltage is obtained, or the rectifier regulation is increased. Figure 14-27 shows the current overlap and its effect on the magnitude of the supply voltage.



FIG. 14-26. Load-current pulses showing overlap angle in six-phase rectifier, inductive load.

Under the assumption that the anode of $T_2$ has just become positive to its cathode and has started conducting before $T_1$ has ceased, the paths from neutral to cathode through the two transformer phases are in parallel and thus

$$e_1 - L_s \frac{di_1}{dt} - e_{b1} = e_2 - L_s \frac{di_2}{dt} - e_{b2} = e_{\text{load}} \tag{14-44}$$

Neglecting tube drop,

$$e_1 - L_s \frac{di_1}{dt} = e_2 - L_s \frac{di_2}{dt} \tag{14-45}$$

As has been stated, the inductive load maintains the total current approximately constant, or

$$i_1 + i_2 \cong I_{\text{dc}}$$

during the commutation period. It follows from this statement that

$$\frac{di_1}{dt} + \frac{di_2}{dt} \cong 0 \tag{14-46}$$

and from Eq. (14-44)

$$e_1 + e_2 - L_s \left( \frac{di_1}{dt} + \frac{di_2}{dt} \right) = 2e_{\text{load}}$$

Using Eq. (14-46), then

$$e_{\text{load}} = \frac{e_1 + e_2}{2} \tag{14-47}$$

so that the load voltage during the commutation interval is the average of the phase supply voltages during that interval. The voltage curve for the overlap angle $\theta$ in Fig. 14-27 is thus determined.

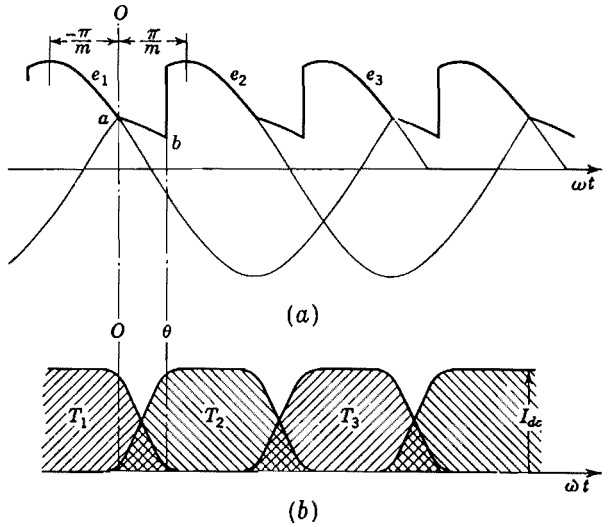



FIG. 14-27. (*a*) *m*-Phase voltage wave, showing effect of leakage reactance. (*b*) Current wave, showing overlap.

Use of Eqs. (14-45) and (14-46) gives

$$\frac{di_1}{dt} = \frac{e_1 - e_2}{2L_s} \tag{14-48}$$

During the overlap angle $\theta$ the phase voltages, based on the indicated zero time location at the beginning of the overlap period, are

$$e_1 = E_m \cos\left(\omega t + \frac{\pi}{m}\right)$$

$$e_2 = E_m \cos\left(\omega t - \frac{\pi}{m}\right)$$

Expanding the angle and subtracting gives

$$e_1 - e_2 = -2E_m \sin\frac{\pi}{m} \sin \omega t \tag{14-49}$$

so that Eq. (14-48) becomes

$$\frac{di_1}{dt} = -\frac{E_m}{L_s} \sin\frac{\pi}{m} \sin \omega t \tag{14-50}$$

After integration the above becomes

$$i_1 = \frac{E_m}{\omega L_s} \sin \frac{\pi}{m} \cos \omega t + A$$

where $\omega L_s$ is the leakage reactance of the transformer per phase at the supply frequency. Since $i_1 = I_{dc}$ at $t = 0$,

$$A = I_{dc} - \frac{E_m}{\omega L_s} \sin \frac{\pi}{m} \tag{14-51}$$

and the current in $T_1$ during commutation takes the form

$$i_1 = I_{dc} - \frac{E_m}{\omega L_s} \sin \frac{\pi}{m} (1 - \cos \omega t) \tag{14-52}$$

Since $i_2 \cong I_{dc} - i_1$, the current in $T_2$ may also be found during commutation.

The average voltage over the conduction period of $T_2$ (or any other recurrent portion of the cycle) is reduced from the value

$$E_{dc} = E_m \left(\frac{m}{\pi}\right) \sin \left(\frac{\pi}{m}\right)$$

for an ideal $m$-phase rectifier without reactance drop to

$$E_{dc} = E_m \frac{m}{\pi} \sin \frac{\pi}{m} - \frac{m}{2\pi} \int_0^{\theta} (e_2 - e_{load})\, d\omega t \tag{14-53}$$

Use of Eqs. (14-47) and (14-49) allows the integrand to be expanded so that

$$\begin{aligned} E_{dc} &= E_m \frac{m}{\pi} \sin \frac{\pi}{m} + \frac{m}{2\pi} \int_0^{\theta} E_m \sin \frac{\pi}{m} \sin \omega t\, d\omega t \\ &= E_m \frac{m}{\pi} \sin \frac{\pi}{m} \left(1 - \frac{1 - \cos \theta}{2}\right) \end{aligned} \tag{14-54}$$

The commutation angle $\theta$ may be found by noting that, at $t = 0$, $i_1 = I_{dc}$ in Eq. (14-52), giving

$$1 - \cos \theta = \frac{\omega L_s I_{dc}}{E_m \sin (\pi/m)} \tag{14-55}$$

so that

$$\theta = \cos^{-1} \left[1 - \frac{\omega L_s I_{dc}}{E_m \sin (\pi/m)}\right] \tag{14-56}$$

Using Eq. (14-55) in Eq. (14-54) leads to

$$\begin{aligned} E_{dc} &= E_m \frac{m}{\pi} \sin \frac{\pi}{m} \left[1 - \frac{\omega L_s I_{dc}}{2E_m \sin (\pi/m)}\right] \\ &= E_m \frac{m}{\pi} \sin \frac{\pi}{m} - \frac{m}{\pi} \frac{\omega L_s I_{dc}}{2} \end{aligned} \tag{14-57}$$

Thus the d-c voltage for the ideal rectifier, as given by the first term on the right, is reduced linearly with current by the second term on the right owing to the transformer reactance. Equation (14-57) is an expression for rectifier regulation due to the transformer reactance.

The angle of overlap is seen to range from 10° to above 30° in Fig. 14-28. For 12-phase rectifiers where the angle of conduction is 30°, this curve would indicate that a third anode may start conduction before the

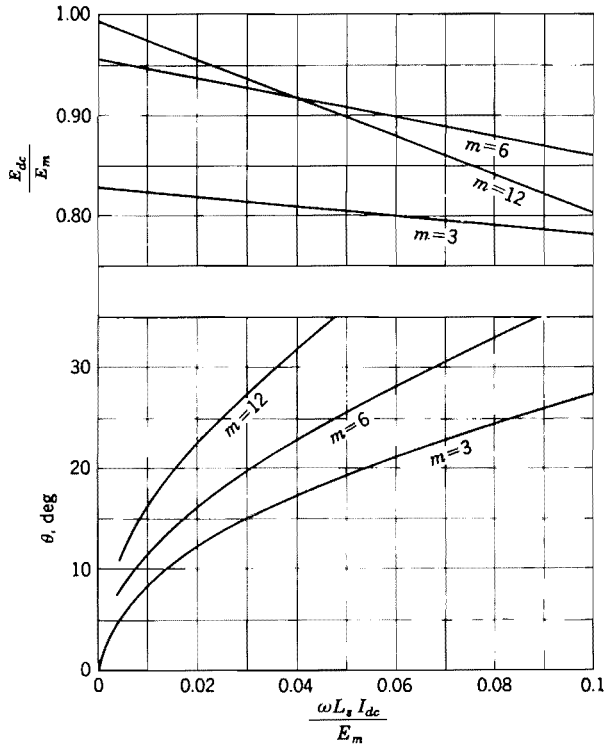



FIG. 14-28. Regulation curve and angle of overlap as a function of load current.

first stops. The results of the analysis of this section are not applicable with more than two anodes conducting at once.

**14-14. The Three-phase Bridge Circuit.** The circuit of Fig. 14-29 is a three-phase adaptation of the bridge circuit, sometimes also called a three-phase full-wave circuit. Each transformer secondary carries current on both half cycles, and this gives the circuit a high transformer utility factor. From the current pulses of Fig. 14-29*b* it can be seen that each tube conducts for two 60° pulses per cycle, in series with one other tube. That is, while phase *a* is positive, $T_1$ conducts for 120°, during

the first 60° of which it is in series with $T_5$, whose cathode is most negative on phase $b$, and during the second 60° in series with $T_6$, whose cathode is then most negative on phase $c$.

The shape of the current pulses in the load is that of a six-phase rectifier, and the general rectifier relations may be applied with $m = 6$. How-

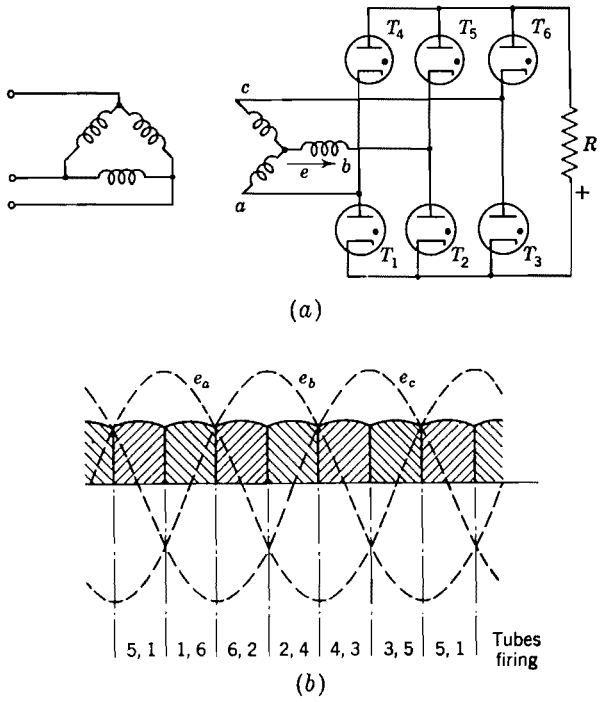



FIG. 14-29. (*a*) Three-phase full-wave or bridge circuit. (*b*) Voltages to neutral and load current.

ever, $E_{dc}$ is $\sqrt{3}$ times greater if $E_m$ is that of one phase to neutral, since two 120° voltages are always acting. Thus

$$E_{dc} = 1.65E_m$$

The peak inverse voltage may be computed from the voltage waves as $1.732E_m$.

The circuit is usually employed for high-voltage rectifiers, where the losses due to two tubes in series are not appreciable. Since the transformer secondaries carry no d-c component, there will be no difficulty due to d-c saturation of the core.

**14-15. Double-Y Rectifiers.** The circuit of Fig. 14-30*a* employs two three-phase circuits, displaced by 60°, and connected in parallel at the

neutrals by an interphase transformer. At any instant there are two anodes carrying current, one in each group as shown in Fig. 14-30*b*, with the addition of the two tube currents in the load resulting in a six-phase load ripple as in Fig. 14-30*c*.

A small excitation current flows through the interphase transformer from its mid-point to the Y group having the highest instantaneous potential, each Y group being at a higher potential than the other group

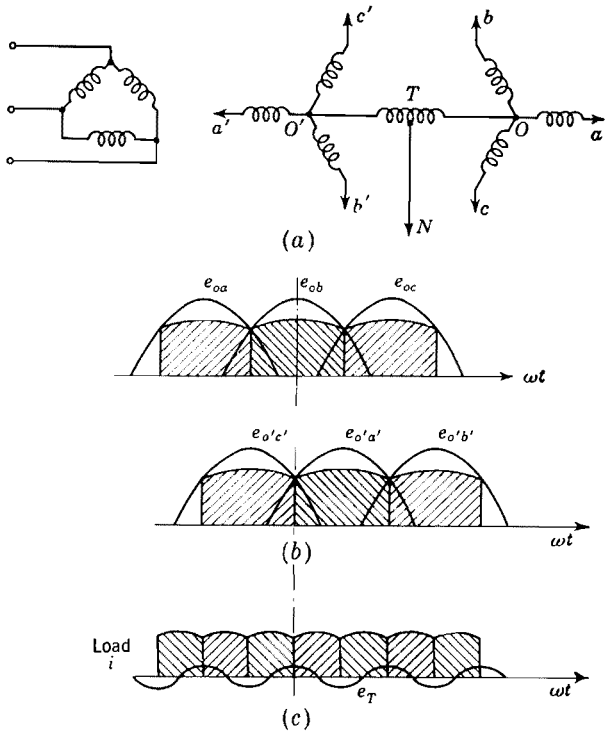



FIG. 14-30. (*a*) Six-phase double-Y circuit. (*b*) Tube-current waveforms. (*c*) Load-current and interphase voltage $e_T$.

three times per cycle. The excitation current for the transformer thus has a frequency three times that of the supply, as shown by $e_T$ in Fig. 14-30*c*, and induces a voltage which opposes the potential of the high-potential group and adds to that of the low-potential Y. The interphase transformer potentials behave as a teeter-totter board might if balanced on the neutral point, the ends going up and down at a frequency three times that of the supply, as the instantaneous Y neutral potentials move up and down. If the interphase transformer were not present, the circuit would become a six-phase star-connected one with 60° tube conduc-

tion. If the d-c load current falls below the excitation current of the interphase transformer, the system changes over to six-phase operation with a rise in voltage of 15 per cent.

With an inductor in the load each half of the interphase reactor carries half the load current, so that the effective reactor current is

$$I_{rms} = 0.5I_{dc}$$

The voltage across the interphase transformer will have a value equal to the difference between the two neutrals, or

$$E_{peak} = 0.5E_m$$

where $E_m$ is that of one phase. Use of Eq. (14-23) with $m = 6$ gives

$$E_{dc} = 0.826E_m$$

The volt-ampere rating of the interphase transformer can then be obtained as

$$\text{VA} = \frac{E_{peak}I_{rms}}{\sqrt{2}} = \frac{0.5E_m}{\sqrt{2}}\frac{I_{dc}}{2} = 0.214E_{dc}I_{dc}$$

The d-c components cancel, and only a-c harmonic components appear in the voltage across the transformer.

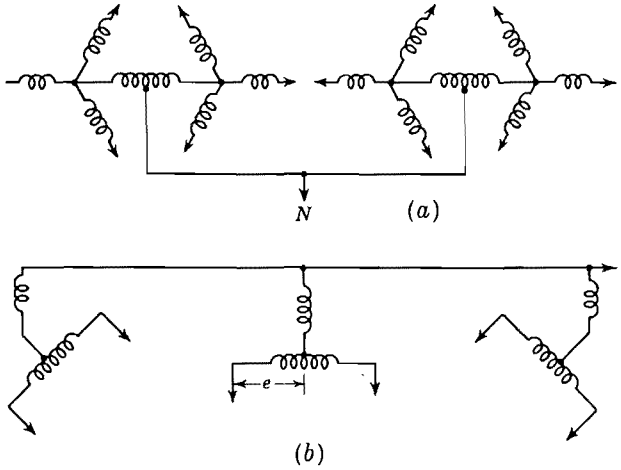



FIG. 14-31. (*a*) Six-phase parallel double Y. (*b*) Six-phase triple diametric connection for 180° conduction.

A six-phase parallel double-Y connection is illustrated in Fig. 14-31*a*. Two interphase transformers are employed with their neutrals paralleled. Also shown in Fig. 14-31*b* is a triple diametric connection employing a three-winding interphase transformer which forces the tube conduction angle to 180°. This circuit requires fewer current commutations per

cycle, and thus the voltage drop due to transformer reactance is only four-ninths that for the six-phase double Y, or more reactance may be tolerated in the transformer for the same drop. Increased reactance reduces the fault current flowing on arc-back and is an advantage of this circuit, as well as the lower peak to average rating required of the tubes.

Most large rectifier installations employ some variant of the double-Y system of connections, because it maintains tube conduction angles near the optimum value for high utility and reduces the number of commutations per cycle.

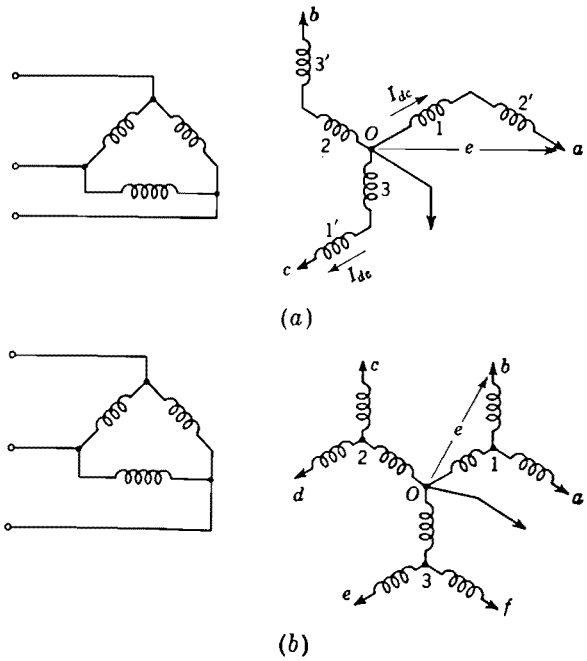



FIG. 14-32. (*a*) Three-phase zigzag transformer connection. (*b*) Six-phase branched circuit.

**14-16. Polyphase Branched Connections.** It was previously mentioned that the three-phase half-wave circuit suffered from d-c saturation of the transformer core and was rarely used. A modification which avoids saturation is shown in Fig. 14-32*a* as the three-phase zigzag connection. Windings on the same core are drawn parallel to each other and indicated by the numbers. On each core there are two windings which carry d-c components $I_{dc}$ in opposite directions as shown, thereby neutralizing the direct magnetomotive force. Since the anode voltage $e$ is derived from two windings, where the voltage of each winding is $0.575e$, a larger and special transformer is required and utility factor is

decreased. Avoidance of the saturation difficulty makes the circuit worthwhile, however.

Branched circuits derived from that in Fig. 14-32*b* are also common. In this case a six-phase voltage is arrived at, in a manner similar to the zigzag circuit. Three secondary windings derive power from one primary phase; secondaries 0-1, 3-*e*, and 2-*d* have windings on the same core. It should be noted that secondary 0-1 carries two 60° pulses of current for each one carried by 3-*e* and 2-*d*; thus the secondaries 0-1, 0-2, and 0-3 must have volt-ampere ratings 1.41 times larger than those of the branched secondaries. This system provides improved utility and allows use of a smaller transformer. Direct-current magnetization of the cores is also avoided.

Further branching connections can be made at *a*, *b*, *c*, *d*, *e*, and *f* to provide larger numbers of phases as required.

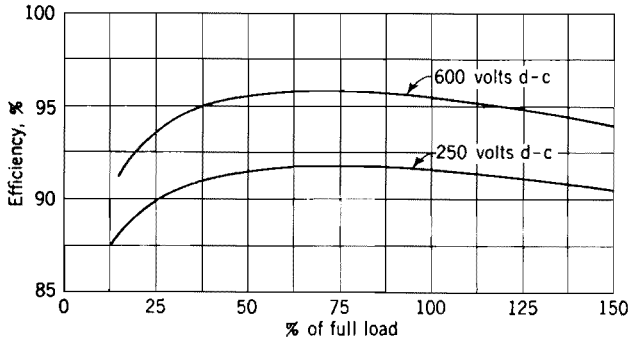



FIG. 14-33. Typical rectifier over-all efficiency curves.

**14-17. Performance.** The theoretical efficiency of rectification has been discussed in connection with various circuits as an inherent circuit property dependent on load-current waveform. When the rectifier has a large number of phases, or for sufficiently inductive load, the theoretical rectification efficiency reaches 100 per cent. Practical operating losses then reduce this figure to the operating value, usually in the range above 90 per cent.

Included in the operating losses are transformer iron and copper losses, rectifier arc-drop loss, and losses due to auxiliary equipment needed for normal operation.

The efficiency curve of a rectifier, of which those in Fig. 14-33 are typical, is fairly flat from one-quarter to above full load. At light load the tube losses are small since they vary in proportion to d-c load current. The copper loss varies as the square of the d-c load current. Transformer iron losses are small and almost independent of load.

The auxiliary losses, most of which are independent of load, include the loss in the keep-alive arc, if one is used, and the power required for operation of vacuum pumps and cooling-water pumps. Losses in the cooling water are not included since they appear in the calculation as arc-drop loss. Since the fixed iron loss and fixed auxiliary losses are small, the rectifier efficiency is good at light load. The variable losses, including the transformer copper and arc-drop losses, cause a droop in the efficiency curve at high loads.

The arc-drop loss of a low-voltage unit is a higher percentage of the kilowatt rating than for a high-voltage unit, since arc-drop loss is proportional only to current. Thus the 250-volt unit of Fig. 14-33 has a full-load efficiency of 91.5 per cent, whereas the 600-volt unit has a full-load efficiency of 96 per cent. A unit operating at 3000 volts might approach 98 per cent efficiency over all.

The power factor of a rectifier is lagging, caused by transformer excitation, harmonic load components, and transformer reactance, with the latter the predominating element at normal load.

Voltage regulation can be calculated, using factors already obtained for the various circuits. Voltage drop is due to transformer copper drop, transformer reactance drop due to commutation, and arc drop. For instance, the output voltage of a star-connected rectifier can be obtained as

$$E_{\text{load}} = E_m \frac{m}{\pi} \sin \frac{\pi}{m} - \frac{m}{\pi} \frac{X I_{\text{dc}}}{2} - \frac{P_c}{I_{\text{dc}}} - E_o \qquad (14\text{-}58)$$

where $X$ = commutating reactance of transformer, ohms
$P_c$ = transformer copper loss, watts
$E_o$ = arc drop of rectifier

Usual regulation figures for rectifiers are in the range of 5 to 6 per cent from full load to light load.

In the case of the six-phase double Y this may be increased at the point at which the load is insufficient to excite the interphase transformer. At this light load the unit changes from a six-phase double-Y circuit, where $E_{\text{dc}} = E_m(m/\pi) \sin (\pi/m) = 0.827E_m$, to a six-phase star rectifier, for which $E_{\text{dc}} = 0.955E_m$. This action may be most easily controlled by a fixed minimum load on the rectifier.

## PROBLEMS

**14-1.** A single-phase full-wave rectifier has a load-current waveform as in Fig. 14-34, due to unequal tube drops. (*a*) Compute the d-c load current. (*b*) Find the total power lost in the resistive load of 25 ohms.

**14-2.** A gas diode is used in a single-phase half-wave circuit to supply 50 volts at 12 amp d-c to a resistive load. If the tube drop is 12 volts, find (*a*) the rms trans-

former voltage required, (*b*) the peak anode current rating required, (*c*) the power input from the transformer, (*d*) the efficiency of rectification. (*e*) Account for all power input. (*f*) Plot to scale waveforms of anode current and tube voltage.

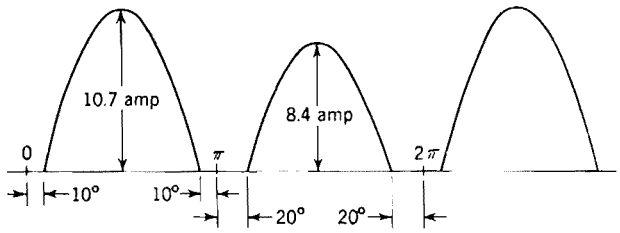



FIG. 14-34

**14-3.** A single-phase full-wave circuit uses a double gas diode rated at 0.675 amp peak and 0.225 amp average per diode. With tube drop of 15 volts, and with a load requiring 300 ma at 600 volts, find (*a*) rms transformer voltage, (*b*) a-c power input from transformer, (*c*) tube-loss total, (*d*) required PIV rating for each diode.

**14-4.** A gas diode rated at 15 amp peak plate current, 5.0 amp average plate current, is used in a half-wave circuit to charge a 12-volt storage battery of 0.02 ohm resistance from a 24-volt-rms source. (*a*) Determine the maximum charging rate possible without exceeding any tube rating. (*b*) Assuming sufficient resistance is added to prevent exceeding the peak-current rating, how long will it take to add a charge of 100 amp-hr?

**14-5.** By accident, the battery in Prob. 14-4*b* is reversed in polarity. Determine the peak and average currents and the conduction angle.

**14-6.** A transformer is rated 1000 volts with a secondary center tap. Find the ratio of d-c power output in the full-wave diametric circuit to that obtainable using the same transformer in a bridge circuit. Neglect tube drop.

**14-7.** Compute the theoretical rectification efficiency for the three-phase half-wave circuit, neglecting thbe-drop and transformer losses.

**14-8.** A six-phase mercury-tube rectifier is operated from a transformer with 230-volt delta-connected primary and star-connected secondary to supply 500 amp at 600 volts d-c. Neglect transformer reactance, but assume the arc drop is 20 volts. Find (*a*) current and voltage ratings for all transformer windings, (*b*) the total a-c input power to the transformer, if its efficiency is 96.5 per cent.

**14-9.** A three-phase half-wave rectifier uses a mercury-vapor diode rated at 1800 amp peak, 200 amp average, 900 volts PIV. (*a*) Neglecting tube drop and transformer reactance, find the maximum d-c current and voltage which may be obtained from such a rectifier with resistance load. (*b*) State transformer primary and secondary volt-ampere ratings required.

**14-10.** Compute the peak inverse voltage for the tubes in the circuit of Fig. 14-32*b* in terms of $E_m$ as the peak of one-winding potential.

**14-11.** Compute the ratio of peak inverse voltage to $E_{dc}$ for the six-phase star circuit of Fig. 14-21*a*.

**14-12.** Maximize the expression of Eq. (14-38), and show the value of $m$ at which the utility factor is maximum.

**14-13.** Using the principles of Fig. 14-32*b*, develop a 12-phase circuit. Calculate its peak inverse voltage.

**14-14.** Design a 200-kw 600-volt d-c rectifier using the three-phase zigzag circuit. Neglecting tube drop and transformer reactance, specify (*a*) voltage rating for all

transformer windings, (*b*) utility factor of secondary windings, (*c*) the kva rating of primary and secondary windings.

**14-15.** Calculate the utility factor for primary and secondary for the circuit of Fig. 14-32*b*, with inductive load.

**14-16.** Mercury-vapor tubes having a rating of 2500 volts peak inverse, 100 amp peak, 35 amp average are available. Find the maximum d-c power available from such tubes when used in the single-phase full-wave diametric, three-phase half-wave zigzag, and three-phase bridge circuits, neglecting tube drop and transformer reactance.

**14-17.** Calculate the utility factor of the secondary of the transformer of Fig. 14-32*a*, and compare with the value for the three-phase half-wave circuit of Fig. 14-18, with same $E_{dc}$ in both cases.

**14-18.** The six-phase half-wave circuit supplies 600 kw at 600 volts d-c. The load inductor is large, and the leakage reactance in series with each anode makes $\omega L I_{dc}$ equal to 5 per cent of $E_m$. The arc drop is 15 volts. (*a*) Find the rms voltage rating of the secondary windings. (*b*) Compute the voltage regulation of the circuit in terms of no-load and full-load voltage. (*c*) Calculate the volt-ampere rating of the transformer.

**14-19.** If the transformer copper loss is 10 kw at full load, the iron loss 7.5 kw, and 2.5 kw is used for pumps, plot a curve of over-all efficiency against load up to 150 per cent load for the rectifier of Prob. 14-18.

**14-20.** The circuit of Fig. 14-32*b* is to supply a load of 200 volts, 1000 amp d-c, with high load inductance. If the tube drop is 15 volts, specify the voltage and volt-ampere ratings of all transformer windings, neglecting transformer reactance.

## REFERENCES

1. Hull, A. W.: Gas-filled Thermionic Tubes, *Trans. AIEE*, **47**: 753 (1928).
2. Marti, O. K., and H. Winograd: "Mercury Arc Power Rectifiers—Theory and Practice," McGraw-Hill Book Company, Inc., New York, 1930.
3. Cobine, J. D.: "Gaseous Conductors," McGraw-Hill Book Company, Inc., New York, 1941.
4. Winograd, H.: Development of the Excitron Rectifier, *Trans. AIEE*, **63**: 969 (1944).
5. Jolley, L. B. W.: "Alternating-current Rectification and Allied Problems," 3d ed., John Wiley & Sons, Inc., New York, 1928.
6. Chin, P. T.: Gaseous Rectifier Circuits, *Electronics*, **18**: 138 (April, 1945), 132 (May, 1945).

CHAPTER 15

# POWER CONTROL AND INVERSION

It has been found possible to control the starting time of the arc discharge by means of a grid in the hot-cathode gas diode and thereby add a form of electrostatic control to the advantages of high current capacity and high efficiency already held by that device. The result of the addition of the grid by Langmuir and Hull was the *thyratron* tube (meaning "door tube"). Later Slepian developed an electrical starter for the mercury-pool cathode tube, eliminating the need for the slow-acting mechanical starters in circuits where control of starting of the tube was desired. This device was called the *ignitron.*

Emission is always present in the thyratron, whereas the igniter element in the ignitron provides the initial spark and electrons. In the thyratron the grid *prevents* the arc until it is desired, while in the ignitron the arc is *started* when desired.

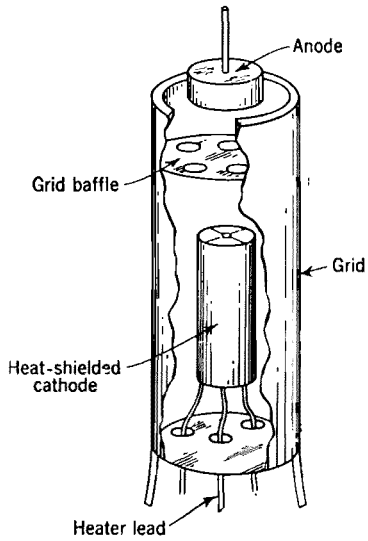
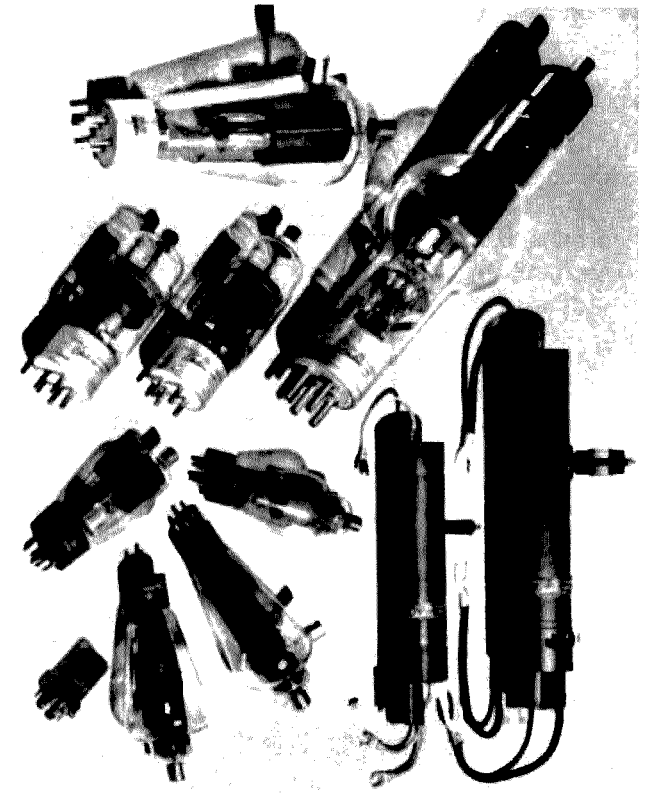


FIG. 15-1. Cutaway view of internal structure of a thyratron.

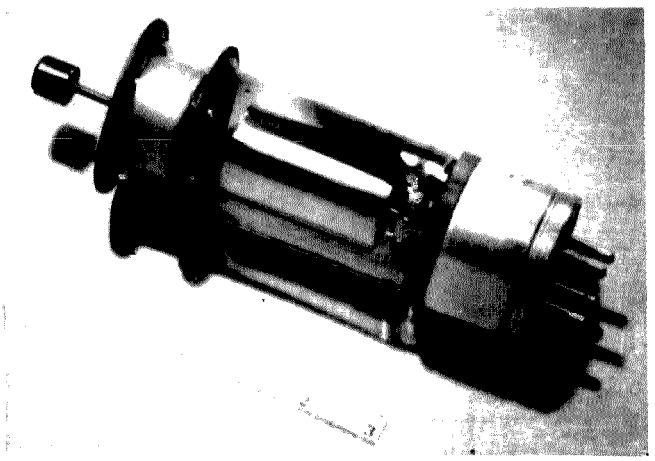
These inventions have opened a new field of control, equivalent in many ways to that by the vacuum tube, but at a much higher level of power and efficiency.

**15-1. The Thyratron.** In order to assure complete control of the anode current, the grid structure in a thyratron forms a very complete electrostatic shield around the cathode. Because of the industrial service for which the tube is intended this structure is also made mechanically rugged, as seen in Figs. 15-1 and 15-2. An oxide-coated heat-shielded cathode is employed, with the electrons forced to pass through one or more holes in a grid baffle between cathode and anode.

With sufficiently negative grid voltage the effect of the anode potential is nullified at the cathode, and no electrons are accelerated to ioniz-



(*a*)



(*b*)

FIG. 15-2. (*a*) Group of thyratrons. (*b*) Modern type GL-6087 thyratron, 6.4 amp average rating. (*Courtesy of General Electric Company.*)

ing energy. At some less negative grid voltage the anode field is able to attract a few electrons and accelerate them to ionizing velocity. The resulting positive ions neutralize some of the negative-grid field near the baffle holes, allowing more electrons to pass through the holes and be accelerated, the process being cumulative until an arc discharge forms between cathode and anode. A particular grid-voltage value is critical in starting the arc.

After ionization begins, the positive ions are attracted to the grid structure and form a sheath around it. Electric-flux lines, starting on the negative grid, reach only to some of the surrounding positive ions, and the grid becomes completely insulated from the discharge by its sheath of positive ions. Changes in grid potential change the number of ions or the sheath thickness, but these sheath changes have no effect on the current to the anode passing beyond the grid sheath. Thus, after the start of the discharge, *the grid has no control over the magnitude of current flowing, nor can it stop the arc.* The anode current after firing is solely a function of resistance or impedance in the anode circuit, and not a function of grid potential. The only way the arc can be stopped is by removal of the positive anode potential.

*In the vacuum tube the anode current is under control of the grid voltage at every instant, but in the gas triode or thyratron the grid can only initiate the arc.*

With the grid negative before the arc starts there will be a very small negative electron flow to the grid due to chance bombardment of high-energy electrons. After ionization starts, the grid current reverses, since it is then due to flow of positive ions to the negative grid. This positive-ion current may be quite large and is limited usually only by grid-circuit resistance or impedance.

If small holes are placed in the grid baffle, the cathode may be so shielded from the anode as to require a small positive grid voltage to initiate the arc. Such a tube is called a *positive-grid tube.* If large holes are placed in the grid baffle, a negative grid voltage will be required to prevent the discharge and such a tube is called a *negative-grid* tube. Figure 15-1 shows an arrangement which would probably provide a positive-grid characteristic.

A curve relating the anode potential and the corresponding critical grid potential which will just prevent the arc is called the *critical grid-control characteristic.* Curves for a negative-grid tube and a positive-grid tube are presented in Fig. 15-3*a* and *b*. For mercury-vapor tubes the control characteristic is a function of condensed-mercury temperature, whereas it becomes a single curve for a tube with inert-gas filling. Circuits for mercury-vapor types must be designed to tolerate this possible shift in characteristics and still give reliable operation.

It should be noted that, for the negative-grid type, the grid voltage

must be applied to *prevent* the discharge. For the positive-grid tube the grid voltage must be applied to *initiate* the discharge. The latter type is useful if no operation is to occur if the grid circuit is accidentally opened.

A considerable variety of thyratron characteristics are available in tubes having a number of anode current and voltage ratings.

**15-2. Ionization and Deionization Times.** It takes a short but finite time for the cumulative process of ionization in a thyratron. Of importance as a time delay in some applications, it is usually in the range of 0.1 to 10 μsec, depending on ion paths, type of gas, and gas pressures.

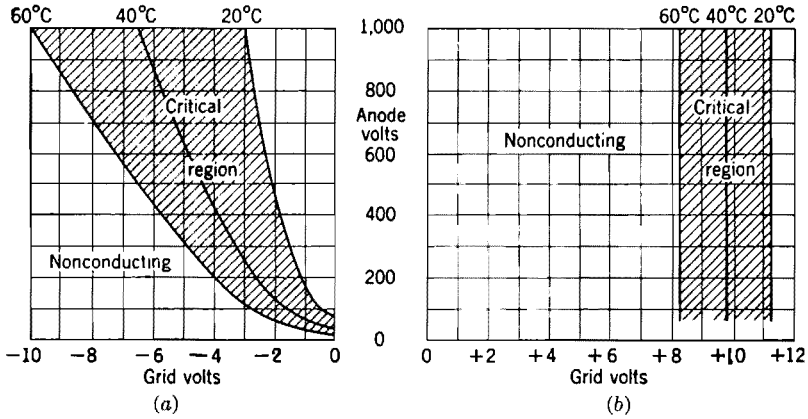



FIG. 15-3. Critical grid characteristic. (*a*) Type 5559/FG-57 negative-grid thyratron; (*b*) type 5720/FG-33 positive-grid thyratron.

Of more importance in most thyratron applications is the *deionization time*, or the time for all ions to recombine and return the space to normal un-ionized gas. It is desired that the grid regain control of arc starting time during each half cycle or period of negative anode voltage, and this means that sufficient time must be allowed for the grid sheath to dissipate.

The deionization time varies with tube structure and amount of electrode surface, since recombination takes place largely on negative surfaces. It is also dependent on temperature, since a higher gas pressure increases deionization time. Large negative anode voltages are helpful in reducing the deionization time since they drive the positive ions to the anode for recombination. The deionization time is usually in the range from 10 to 1000 μsec for mercury-vapor tubes.

Hydrogen-filled thyratrons have been developed for giving very short high-current pulses. Their ionization time has been reduced to less than 0.1 μsec, with a deionization time of the order of 5 μsec, both being results of the lightweight hydrogen ion.

Since the anode must be negative for a time longer than the deionization time if the grid is to regain control of firing time in the succeeding

cycle of a-c anode voltage, then the deionization time must be shorter than the time of a half cycle of the anode voltage. This requirement sets an upper frequency limit for thyratron operation.

**15-3. Grid Current in Thyratrons.** As was mentioned, before firing a very small grid current flows, made up of electrons. At the instant of firing this current increases very rapidly and after ionization is established may reach values of many amperes unless limited by series resistance. The action of the grid current prior to and at the firing instant is shown by the curves of Fig. 15-4 and the oscillogram of Fig. 15-5. In particular the latter shows the increase

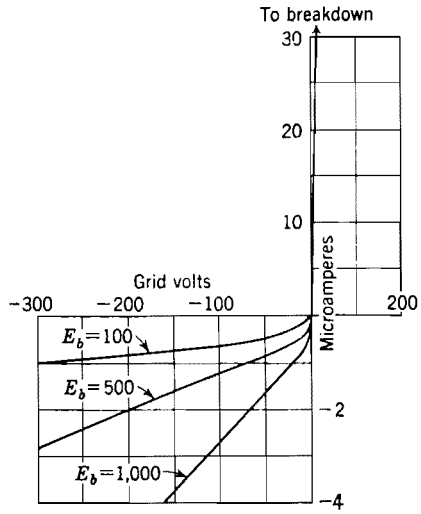



FIG. 15-4. Grid current prior to conduction.

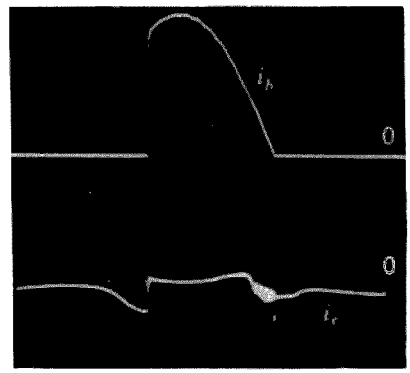



FIG. 15-5. Anode current and grid current in a thyratron with delayed firing.

in negative current just prior to firing, as the positive sinusoidal anode voltage rises. This increase in current with increased anode voltage is predicted by the curves of Fig. 15-4, in the negative-grid region. At the instant of firing the grid current reverses, being then due to positive-ion flow. At the end of the positive-anode half cycle the grid current becomes somewhat erratic or oscillatory in going back to its original negative value.

If there is a large resistance in the thyratron grid circuit, the small electron flow may bias the grid so far negative that the discharge cannot start, even at the anode-voltage peak. This places an upper limit on allowable resistance in the thyratron grid circuit. For use with devices of very high resistance, such as emissive photocells, it is necessary to use thyratron types having extremely small predischarge grid currents, and these tubes are usually of the shield-grid type, to be described.

After conduction starts, it is usually necessary to limit the grid current with resistance to prevent burnout of grid leads inside the tube or other circuit damage.

**15-4. The Gas Tetrode or Shield-grid Thyratron.** It has been found possible to reduce the current flowing in the control-grid circuit prior to firing, by improved shielding of the control-grid structure and reduction of its area. Figure 15-6 shows a diagrammatic view of such a tube. The surrounding shield is isolated from the control grid and may be connected to cathode or biased positively or negatively as desired. The control grid is only a ring or sleeve interposed in the electron path to the anode. Electrons leaving the cathode under nonconducting conditions strike the shield grid instead of the small area of control grid, reducing the control-grid current. Grid-circuit resistances up to 100 megohms may then be tolerated without excessive bias build-up.

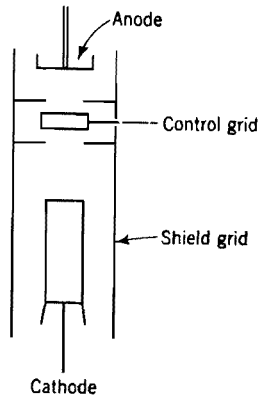


FIG. 15-6. Cross section of a shield-grid thyratron.

If the shield grid is made positive, the tube develops a negative control characteristic, whereas if the shield grid is made negative, a positive control characteristic is obtained. The input capacitance of the gas tetrode may be only a fraction of a micromicrofarad, and this raises the grid-circuit impedance and improves the firing reliability for very short input pulses.

**15-5. Bias Control of Average Current.** If an alternating voltage is applied to the anode of a thyratron, the arc is extinguished on each negative half cycle and the grid regains control of the firing time in each positive half cycle. Thus, while the grid cannot control the instantaneous anode current, as in a vacuum tube, it can control the d-c, or average, value of the anode-current pulses by delaying the starting time in any positive anode cycle.

If the anode voltage is plotted as $e$ in Fig. 15-7*a*, the critical grid voltage corresponding to each anode value may be read from a tube characteristic and plotted as a *critical control locus* for that particular anode-voltage waveform. Then, if the d-c grid voltage is set at a value $-E_1$, the grid voltage will become more positive than the critical locus at angle $\theta_1$ and the tube will fire.

Since the grid then loses control, conduction will continue until the circuit voltage falls below the tube drop at $\theta_2$. If the grid voltage were made more negative than the most negative value of the critical locus, then the tube would not fire at all. The effect of delaying the firing angle to 70° of the cycle is demonstrated in Fig. 15-8.

The anode current can be written as

$$i_b = \frac{E_m \sin \omega t - E_o}{R} \tag{15-1}$$

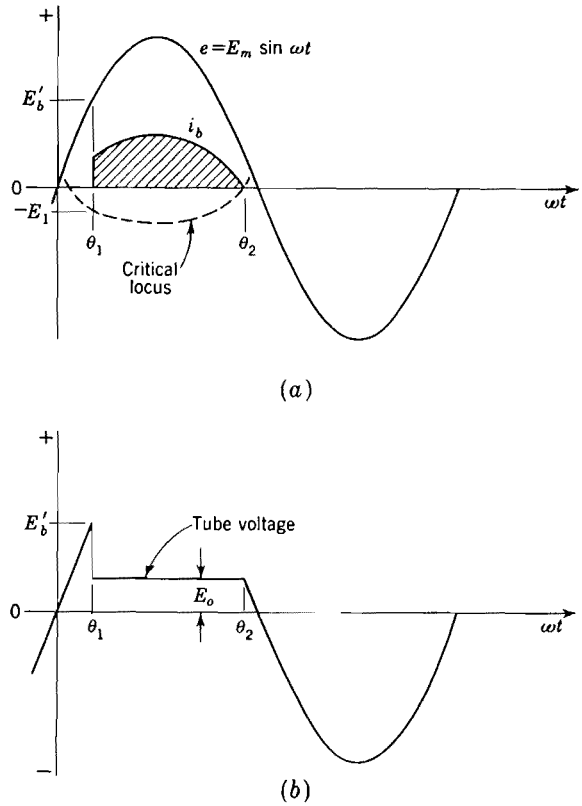



FIG. 15-7. (*a*) Critical control locus for sine voltage applied, with firing angle = $\theta_1$. (*b*) Tube voltage.

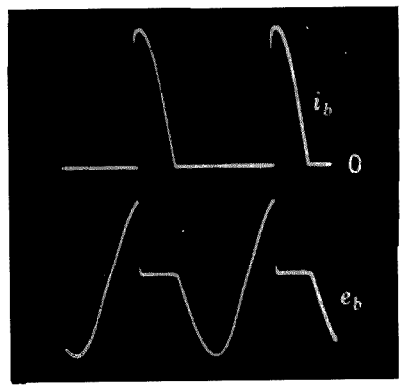



FIG. 15-8. Current and tube voltage with firing angle delayed to 70°.

The firing angle $\theta_1$ may be found by determining the anode voltage $E_b'$ from the control characteristic which corresponds to grid voltage $-E_1$, as from Fig. 15-3. Then

$$\theta_1 = \sin^{-1} \frac{E_b'}{E_m} \tag{15-2}$$

At $\theta_1$ the tube voltage drops abruptly to $E_o$, the arc-drop value, and without inductance in the load the excess voltage causes a current step $(E_b' - E_o)/R$ in value.

Since the grid loses control at $\theta_1$, the tube becomes a diode and conduction continues to the angle $\theta_2$, at which the numerator of Eq. (15-1) goes to zero, or

$$\theta_2 = \pi - \sin^{-1} \frac{E_o}{E_m} \tag{15-3}$$

for a resistance load. The case of inductive load will be discussed in Sec. 15-13.

The direct current produced by such a controlled firing time may be computed by the methods of Chap. 14 as

$$\begin{aligned} I_{dc} &= \frac{1}{2\pi} \int_0^{2\pi} i_b \, d\omega t = \frac{1}{2\pi} \int_{\theta_1}^{\theta_2} \frac{E_m \sin \omega t - E_o}{R} \, d\omega t \\ &= \frac{E_m}{2\pi R} \left[ \sqrt{1 - \left(\frac{E_o}{E_m}\right)^2} + \cos \theta_1 - \frac{E_o}{E_m} (\theta_2 - \theta_1) \right] \end{aligned} \tag{15-4}$$

For circuits in which the tube drop $E_o$ is small with respect to $E_m$, then

$$I_{dc} \cong \frac{E_m}{\pi R} \frac{1 + \cos \theta_1}{2} \tag{15-5}$$

Even though the grid loses control after ignition, it is still possible to control the *average* value of anode current by delaying the time of firing of the thyratron. Thyratron control circuits must then be designed to provide variation of the first point in the cycle at which the actual grid voltage instantaneously exceeds the value of the critical-control-locus voltage. It is possible to design low-power circuits to achieve this variation and to provide smooth, stepless, and practically lossless control of the magnitude of an average current.

A simple but not too satisfactory form of *bias control* is shown in Fig. 15-9*a*. It can be seen from Fig. 15-7 that, if $e_c$ becomes more negative than any point on the critical control locus for the applied anode voltage, then the tube will not fire during the cycle. If the grid is made less negative, the first intersection of the bias-voltage line and the critical locus must occur to the left of the maximum locus voltage, or for $\theta_1 < \pi/2$.

Thus bias control can give a range of current variation only from the maximum of $E_m/\pi R$ to half this value. Figure 15-9*b* shows this variation for firing angles $\theta_1$ produced by change of d-c grid bias.

It is also apparent that, if the critical control locus moves up or down by reason of temperature variation, a given grid voltage may or may not fire the tube and considerable instability exists, particularly near $\theta_1 = \pi/2$. Thus the circuit lacks precision in firing angle as well as being hampered by the limited range of current control and is usually employed only with large bias changes for on-off control.

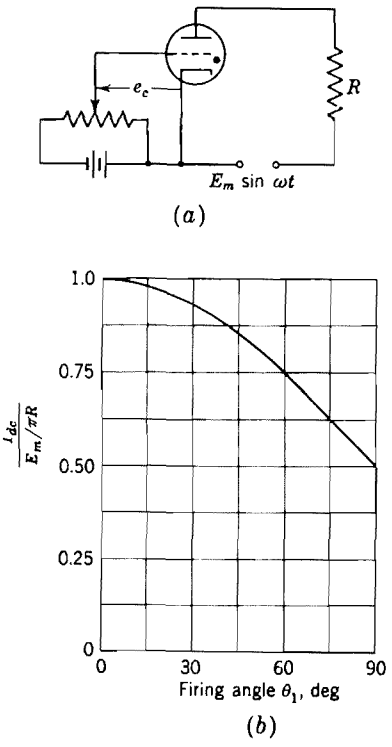



FIG. 15-9. (*a*) Bias control of average current. (*b*) Variation of current produced by (*a*), tube drop neglected.

Several other more satisfactory circuits for variation of ignition angle will be considered in succeeding sections. Analysis of these circuits for tube-performance values such as $I_{rms}$, tube loss, and a-c power input follows the general methods of Chap. 14, with due consideration of the definitions of $\theta_1$ and $\theta_2$.

**15-6. Phase-shift Control.** If the d-c bias of the previous section be replaced with an a-c bias of variable phase, it is possible to cause the positive-going bias wave to intersect the critical-grid-locus curve at any point in the positive-anode half cycle, or to vary the average value of anode current from maximum to zero as would be predicted by Eq. (15-5), and as plotted in Fig. 15-10 for resistive load. It is apparent that a leading grid voltage will always start the arc at $\omega t = 0°$ of the anode wave, and so leading grid voltage provides no control.

If the amplitude of the a-c grid-voltage wave is large with respect to any voltage value on the critical locus, then the grid-voltage wave crosses the critical locus at a steep slope and the firing angle $\theta_1$ will be very nearly equal to $\theta$, the angle of lag of the grid voltage behind the a-c plate voltage, as can be seen in Fig. 15-11. This is equivalent to saying that the critical-control-locus voltage is zero for all anode voltages, and this is a reasonable assumption for the usual anode voltages.

Because of the steepness with which a large-amplitude a-c bias wave crosses the critical locus, any variations which tend to move the locus up or down, such as mercury temperature, can have only a very slight effect on the ignition angle. Thus the a-c phase-control circuit also provides precision of firing angle, as well as the full range of current control not available with simple bias control.

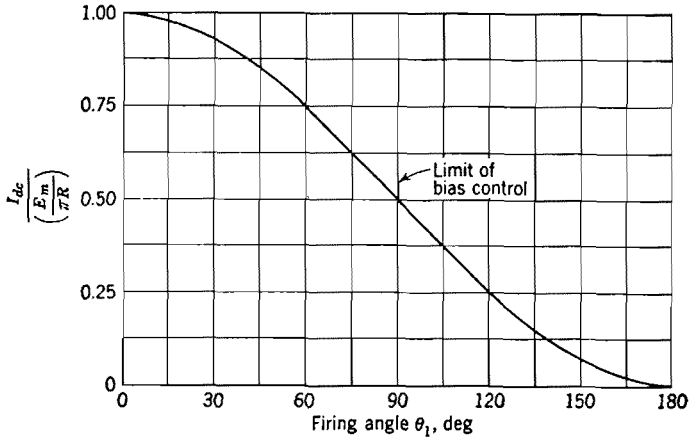



FIG. 15-10. Variation of average anode current for resistive load, as given by Eq. (15-5). Tube drop neglected.

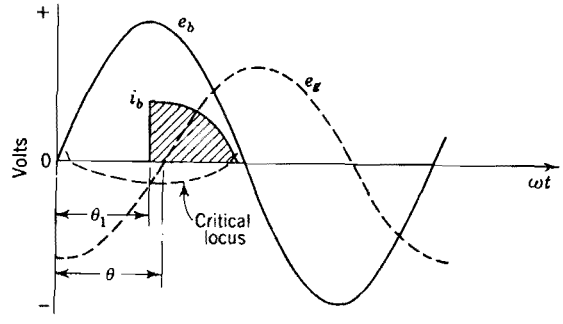



FIG. 15-11. Phase-shift control with a-c grid voltage lagging by angle $\theta \cong 100°$.

This form of control was suggested in 1928 by the French engineer Toulon. Circuits for achieving the required phase shift will be discussed later.

**15-7. Combined Phase and Bias Control.** In some situations it is desirable to control a thyratron by variation of a d-c bias yet necessary to have a wider range of current control than is possible with simple bias control. This may be accomplished by *bias-phase control*, a variable d-c bias being combined with a shifted- but fixed-phase a-c voltage. The

a-c component is shifted by about 90°, as shown in Fig. 15-12*a*, by placing a resistor $R_T$ in series with the grid-transformer primary $T$, which takes a lagging current.

Figure 15-12*b* shows that, if the bias $-E_{cc}$ is increased or decreased, the zero axis for the a-c grid component is raised or lowered and the intersection of $e_g$ and the critical control locus, or the firing angle, can be shifted from 180° back to an early time in the cycle, considerably less

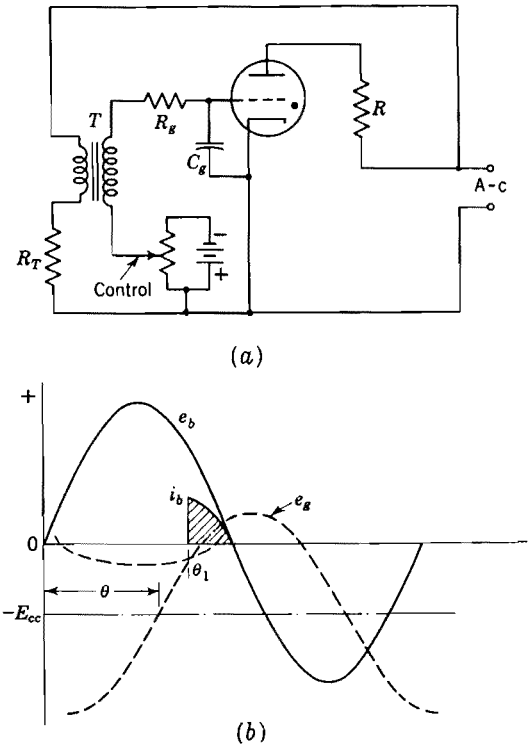



FIG. 15-12. (*a*) Bias-phase control. (*b*) Effect of bias-phase control.

than 90°. Since the a-c wave is usually quite steep at its point of intersection with the locus, then the firing angle is reasonably independent of temperature and other tube factors, as well as grid-voltage magnitude. It is again convenient in analysis to take the control-locus voltage as zero, or to assume that the tube fires when $e_g$ crosses the zero-voltage axis.

The capacitor $C_g$ between grid and cathode is advisable in most thyratron circuits to prevent accidental ignition due to surges in the anode voltage. Because of the capacity between anode and grid a sudden anode-voltage surge may raise the grid potential instantaneously above

firing level. To prevent this the grid-cathode capacity is increased by $C_g$. This increases the grid-cathode capacity with respect to the anode-grid capacity and stabilizes the grid potential for sudden surges. The value of $C_g$ will usually approximate 100 $\mu\mu$f.

This circuit is well suited to control of thyratrons by vacuum tubes, since the variation of the d-c bias is a convenient form of output for the vacuum tube.

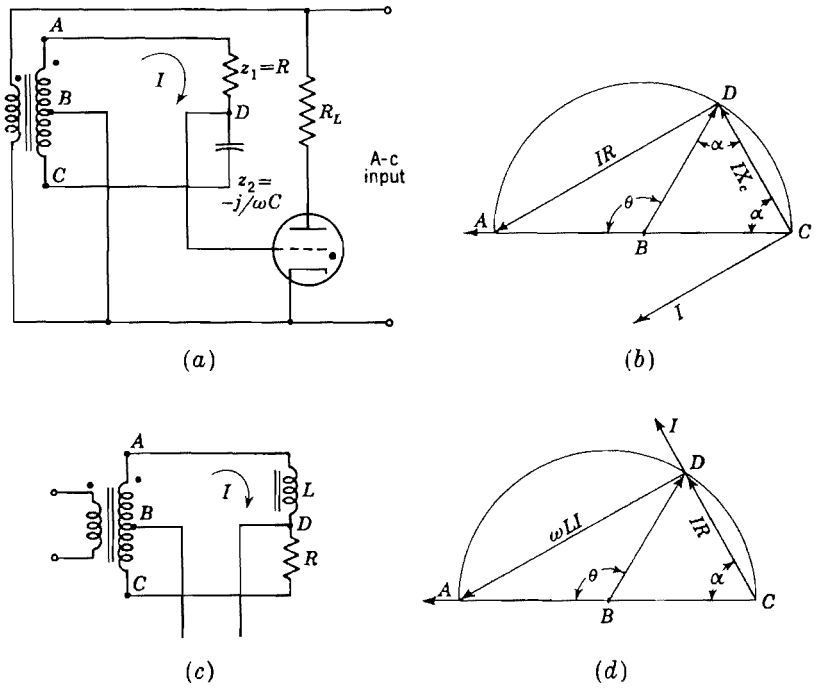



FIG. 15-13. Phase-shift bridges.

**15-8. The Phase-shifting Bridge.** An economical form of circuit for shifting the phase of an a-c voltage is the *phase-shift bridge* of Fig. 15-13*a*. If the transformer is polarized as indicated by the dots, the voltage of $A$ is in phase with the anode supply; voltage $CA$ is drawn as a phasor in Fig. 15-13*b*. With $\mathbf{Z}_1 = R$ and $\mathbf{Z}_2 = -j/\omega C$ it is known that the drops $IR$ and $IX_C$ must be at right angles and must add to the constant voltage $CA$. Thus the locus of point $D$ is a circle of diameter $CA$, as shown in Fig. 15-13*b*. Phasor $BD$ is then the cathode-grid voltage, which will be of constant magnitude, and angle $\theta$ is the lag of grid voltage behind the anode voltage. Variation of either $R$ or $X_C$ will change the grid-voltage phase angle, with $\theta = 0°$ at $R = 0$ and $\theta = 180°$ for $R = \infty$.

From the geometry of the phasor diagram it is apparent that triangle $BCD$ is isosceles. Then

$$\theta + (180° - 2\alpha) = 180°$$

or

$$\theta = 2\alpha = 2 \tan^{-1} \frac{\mathbf{Z}_1}{\mathbf{Z}_2} = 2 \tan^{-1} \omega CR \tag{15-6}$$

for the circuit using $C$ and $R$ elements.

If $\mathbf{Z}_1 = j\omega L$ and $\mathbf{Z}_2 = R$, a phase-shifting circuit using variable $R$ or $L$ is also possible. A circle diagram as in Fig. 15-13*d* is obtained, with $AD = \omega LI$ and $DC = IR$. It may be found that

$$\theta = 2 \tan^{-1} \frac{\mathbf{Z}_1}{\mathbf{Z}_2} = 2 \tan^{-1} \frac{\omega L}{R} \tag{15-7}$$

This circuit modification is shown in Fig. 15-13*c* and is frequently employed because variable inductances are of practical size at supply

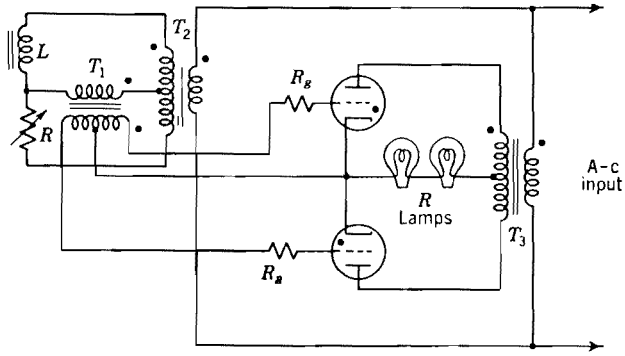



FIG. 15-14. Full-wave controlled thyratron circuit with $RL$ phase shifter. Transformer polarity indicated by dots.

frequencies. It should be pointed out that the full 180° range of shift is not usually available in these circuits, since no variable $L$, $C$, or $R$ ever smoothly covers the range from zero to $\infty$.

Both of the equations for $\theta$ were obtained on the assumption that the grid circuit of the thyratron takes negligible current, and this is usually true prior to starting the arc. After ignition the phase angle and waveform may change because of grid current, but the result is not important at that time. A complete full-wave control circuit employing a variable resistor for phase shift is drawn in Fig. 15-14. Transformer polarities are indicated by dots. Controlled d-c output in the load is then possible by variation of the resistance $R$, and this circuit has been much used for control of lighting.

If the control is to be by variable $L$, then the inductor may take the form of a movable iron core in a solenoid coil, where mechanical motion

is available as the primary control function. In cases where an electrical current or voltage is the primary quantity it is more usual to avoid any mechanical moving parts and employ a *saturable reactor* as the variable-inductance element.

**15-9. Saturable-reactor Control.** A saturable reactor is an inductance whose magnitude is controllable by a direct current. Referring to Fig.

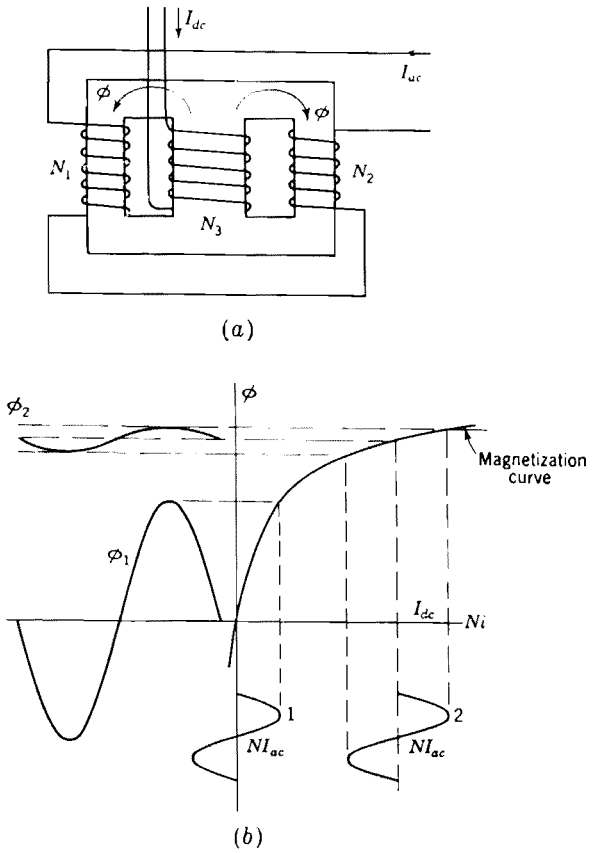



FIG. 15-15. The saturable reactor.

15-15, the direct current flows in the winding $N_3$ around the center leg of a three-legged core. Alternating current flows through $N_1$ and $N_2$ on the outer legs, so connected that their fluxes add in the outside path of the core. While shown in series, the a-c coils may also be connected in parallel, the performance differing only slightly and being due to flow of harmonic currents in the parallel connection.

The a-c fluxes oppose in the center leg, and no a-c voltage will be induced in the d-c coil $N_3$, provided a-c windings and cores are identical.

Since $N_3$ may have many turns so as to operate with small d-c current values, it would be undesirable to have a large a-c voltage present in the d-c circuit.

With no d-c current and flux present the operation is shown as condition 1 in Fig. 15-15*b*. The current $I_{ac}$ produces a magnetomotive force of $(N_1 + N_2)I_{ac} = NI_{ac}$, and from the magnetization curve of the figure this results in a flux $\phi_1$, as shown plotted to the left of the flux axis. Such a large flux variation causes a large counter emf, which results in a large inductive reactance.

When a current $I_{dc}$ is applied to $N_3$, the d-c flux moves the zero axis for $NI_{ac}$ out to some point as for condition 2 of the figure. Because of the flatness of the magnetization curve in the saturation region only a small flux variation is produced. This sets up only a small counter emf, and the reactor appears to have a small reactance.

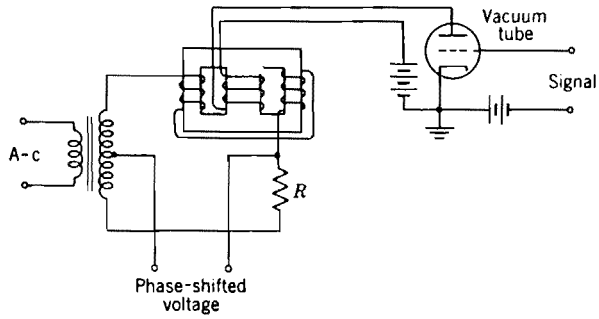



FIG. 15-16. Use of saturable reactor as control element in a phase-shift bridge.

Reactance is stated by $2\pi fN\, d\phi/di$. The derivative represents the slope of the magnetization curve, and this slope under the two chosen conditions confirms the above reasoning.

The reactance of the a-c coils can be varied smoothly over a wide range, possibly as high as 20:1. The phase of the bridge output voltage of Fig. 15-16 is then shifted in accordance with the theory of the preceding section. Since $N_3$ can be wound with many turns of fine wire, the phase shift can be produced by very small currents if desired. For such applications magnetic steels having very high permeabilities and small hysteresis loops of nearly rectangular shape are employed for high-sensitivity to small d-c currents.

While operating near the knee of the magnetization curve, there will be considerable distortion of the a-c flux. However, the harmonics produced in one a-c leg will partially oppose those produced in the other leg and reduce the distortion somewhat.

Saturable reactors of small size are used as indicated, to control thyratron operation, or to vary an alternating current in accordance with the

magnitude of a d-c current. In larger sizes, with the d-c supplied by controlled thyratrons, they furnish the alternating current for operation of electric furnaces, theater lights, and other loads not affected by a distorted current waveform. In theater-light dimming they provide stepless variation of illumination, high efficiency because the thyratron losses are small, and reduction of the theater switchboard to a small control desk.

**15-10. Pulse Control of Thyratrons.** In many thyratron applications it is desirable that the firing angle be closely representative of some

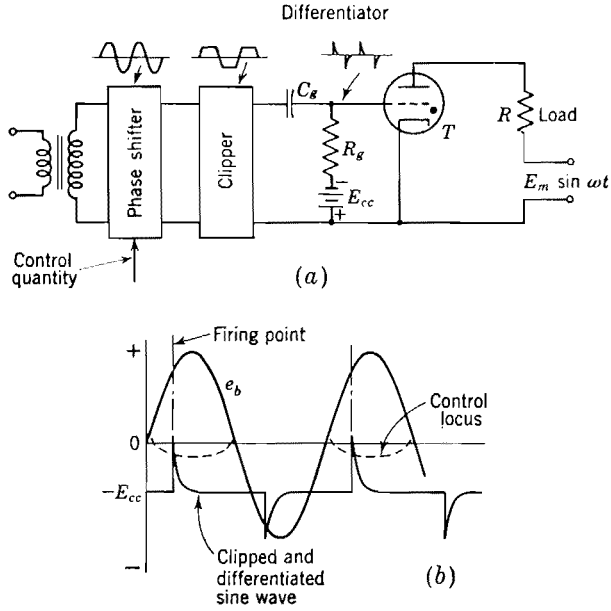



FIG. 15-17. Pulse firing a thyratron.

physical condition and independent of tube characteristics, condensed mercury temperature, bias voltages, and other tube variables. This can be accomplished by firing the thyratron by a sharp, positive-going pulse of grid voltage at the desired instant. The exact level of the critical control locus as a function of mercury temperature then does not enter into determining the time of firing.

One means of providing sharp pulses for precision firing of thyratrons is illustrated in Fig. 15-17, where the sine-wave supply is shifted by the usual phase shifter in accordance with the controlling quantity, then clipped to approximate rectangular form and followed by an *RC* differentiator. The positive-going pips of the resultant wave can be used to fire an overbiased thyratron at the exact instant desired, as shown in

Fig. 15-17*b*. The actions of the clipper and differentiator circuits have been discussed in Chap. 9.

Another device which produces a reasonably sharp pulse without other electronic devices is the *peaking transformer*, which may be used between the phase-shift bridge and the thyratron grids. Such a device may consist of a small transformer in series with a resistor to produce a large voltage drop at the peak of the magnetizing current, giving a flattened flux wave, or a transformer of special design may be used to achieve the same effect, essentially replacing the series resistor with a high internal transformer reactance. A peaking transformer of this design is shown in Fig. 15-18.

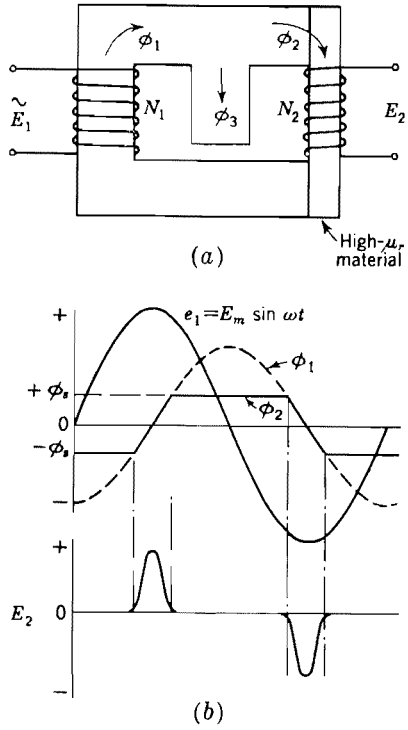



FIG. 15-18. (*a*) Peaking transformer with magnetic shunt. (*b*) Action of peaking transformer.

The primary leg is magnetically shunted by a core piece having an air gap, and the secondary winding $N_2$ is placed on a leg of small cross section, made from a magnetic material of high initial permeability which saturates at a low magnetomotive force. As the applied emf wave rises from zero, $\phi_1 = \phi_2$, since the reluctance of the air-gap shunt is much larger than that of the secondary path. However, the secondary saturates at some low flux level $\phi_s$, and above this level the primary flux is forced to pass through the shunt path containing the air gap. Change of flux through the secondary then occurs only during the interval when the flux passes through zero, going from $-\phi_s$ to $+\phi_s$. As a result the only secondary emf will be a voltage peak during this same interval, giving a peaked wave for $E_2$ as in Fig. 15-18*b*.

Under the assumption that the magnetomotive-force drop through the iron is negligible with respect to that required to force the flux across the air gap, and that the coupling to the secondary can be neglected because it affects the primary current only over the very short period in which the flux goes from $-\phi_s$ to $+\phi_s$, then from magnetic-circuit theory

$$\phi_3 = \frac{N_1 i_1}{l_g/\mu A_g} \quad \text{webers} \tag{15-8}$$

where $l_g$ = length of air gap, m
$A_g$ = area of air gap, sq m
$\mu = 4\pi \times 10^{-7}$ for air

The counter emf of the primary is related to the primary flux $\phi_1$ as

$$e = E_m \sin \omega t = N_1 \frac{d\phi_1}{dt}$$

from which, under the assumptions,

$$\phi_3 = \phi_1 = \frac{-E_m}{\omega N_1} \cos \omega t \tag{15-9}$$

the flux lagging the applied voltage by 90° as shown in Fig. 15-18. From Eqs. (15-8) and (15-9) the instantaneous current in the primary is

$$i_1 = \frac{-E_m l_g}{4\pi \times 10^{-7} A_g \omega N_1^2} \cos \omega t \tag{15-10}$$

giving for the rms primary current

$$\mathbf{I}_1 = \frac{-\mathbf{E}_1 l_g}{4\pi \times 10^{-7} A_g \omega N_1^2} \quad \text{amp} \tag{15-11}$$

where $\mathbf{E}_1$ is the applied rms voltage. It can then be seen from the maximum relations in Eq. (15-10) that the input impedance of the primary is

$$\mathbf{Z}_1 = \frac{j4\pi \times 10^{-7} \omega N_1^2 A_g}{l_g} \quad \text{ohms} \tag{15-12}$$

The transformer may then be designed so that this impedance will not severely load the phase-shift bridge with which it is to work.

Use of core material for the secondary leg which saturates at low density will give a short secondary pulse, the only requirement being that $\phi_s$ should be small with respect to the peak of $\phi_1$. Prediction of secondary-voltage waveform is dependent on the magnetization curve of the material and on secondary-circuit loading and capacities and will not be attempted here.

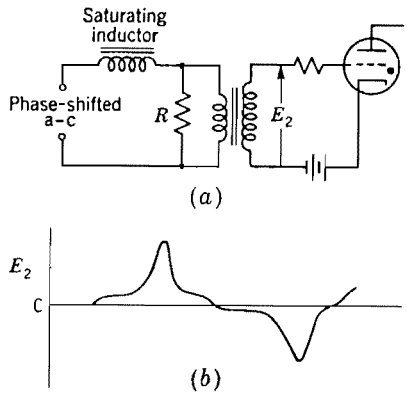



FIG. 15-19. Saturating-inductor peaking circuit.

Another peaking circuit uses an inductor which readily saturates near the peak of the applied a-c wave. The high peak current flowing when the inductor saturates produces a peaked voltage across $R$ (Fig. 15-19), and this gives a peaked secondary voltage $E_2$ to more precisely fire the thyratron.

**15-11. The Ignitron.** The *ignitron* was developed by Slepian in 1933 to provide a fast electrical starter for the mercury-pool type of tube. It utilizes a graphite anode, pool cathode, and an igniter electrode, all contained in a steel envelope. For small sizes these are pumped and sealed off, but large units are usually operated continuously connected to vacuum pumps to remove any gases which may be released during operation. Most sizes are water-jacketed for removal of heat developed

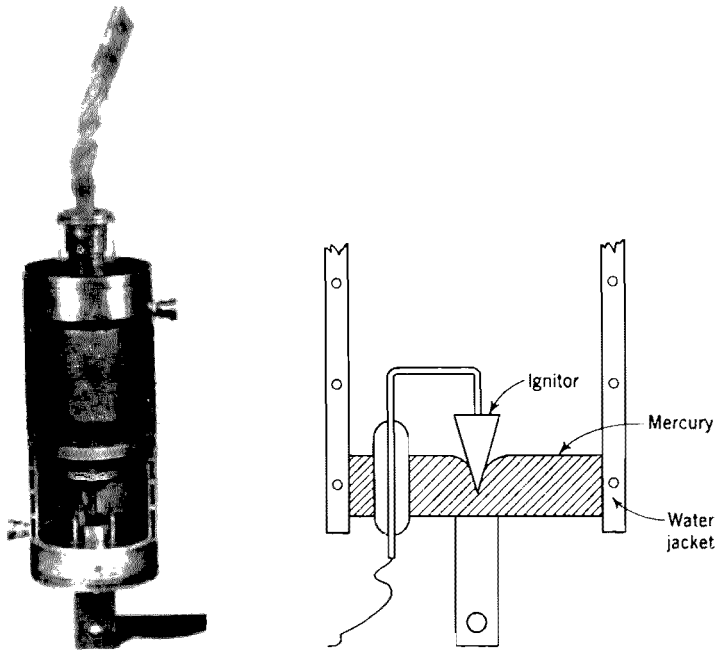



FIG. 15-20. (*a*) Cutaway view of type 5555 ignitron. (*Courtesy of General Electric Company.*) (*b*) Cross section showing igniter.

in the tube during operation. Figure 15-20 shows a cutaway view of one type of ignitron.

The igniter is made of a refractory semiconductor, usually silicon carbide or boron carbide, which is nonwetting in mercury. Rod-shaped and pointed, the igniter dips into the mercury pool, and the tube is fired by passing a current from igniter to pool. The theory of operation is not well understood, but it is believed that the first spark develops when the voltage gradient in the thin gas layer surrounding the unwetted rod exceeds gas breakdown. This first spark grows into a full cathode spot in a few microseconds. The igniter *initiates* the arc, in contrast to the

action of the grid in the thyratron, which *prevents* the arc until it is desired.

The igniter permits starting the arc at any desired point in the positive-anode half cycle, and the average anode current may be varied in the same manner as with a thyratron. At present the ignitron is suited to low-voltage operation, whereas the thyratron is operable as high as 15,000 volts. Since the ignitron is a mercury-pool tube, there are no severe limitations on available cathode emission as for the hot-cathode thyratron, and high peak- or surge-current ratings are possible. Because of the pool cathode the tube drop may lie in the range of 15 to 25 volts.

In the thyratron's favor is its high-impedance input or grid circuit, since the ignitron requires currents of 10 to 30 amp at 50 to 400 volts for firing. This represents a low-impedance circuit, but the average power demand is low because the firing operation usually lasts only 50 to 150 $\mu$sec. To avoid this low-impedance burden on the controlling circuits, it is common practice to employ thyratrons to trigger the igniter circuit.

The cathode current in mercury tubes actually flows from a number of cathode spots acting in parallel, each spot carrying between 3 and 15 amp. If the total tube current becomes less than about 5 amp, the one remaining spot may become unstable, so that it is sometimes necessary to provide a dummy load, or an auxiliary anode and small load to operate continuously, to give smooth operation.

Internal heating is the usual limiting factor on current ratings, and cooling water must be supplied to the jackets to maintain the tube temperature in the desired 10 to 40°C range. Interlocks must be used so that the tubes cannot carry current without cooling-water flow. Thermostats on outlet-water temperature may be used to regulate the water flow rate. With the steel tube jackets at cathode potential there will usually be necessity for insulation from the grounded water line. This is accomplished by introducing 10 to 15 ft of rubber hose in the water line, fitted with suitable internal copper electrodes to prevent destruction of pipe and fittings by electrolysis from the leakage current.

The temperature rise of the water is fixed by the tube ratings, usually to the range of 4 to 9°C rise. It may be computed as

$$\text{Temperature rise, °C} = \frac{0.014 \text{ kw loss}}{\text{kg water/min}} = \frac{\text{kw loss}}{0.263 \times \text{gpm}} \tag{15-13}$$

Ignitrons are frequently assembled as multiple-unit rectifiers to operate in the circuits of Chap. 14, the igniter control being employed for output-voltage regulation or for circuit-breaker use. They are also used in pairs as synchronous contactors for a-c loads such as resistance welders,

where their ability to turn on and off for integral numbers of cycles is utilized.

**15-12. Ignitron Firing Circuits.** Maximum and minimum igniter voltage and current are shown in Fig. 15-21, and a pulse having instantaneous values falling within the shaded area should be applied for several hundred microseconds to ensure firing of the ignitron. Longer pulses of igniter current should not be needed to fire the tube and will lead to igniter overheating and possible damage. Negative igniter current should not be permitted, because a reverse current may cause a cathode spot to form on the igniter, resulting in ultimate overheating and igniter burnout. Reverse current can be prevented by insertion of semiconductor diodes or other rectifier elements in the igniter circuits.

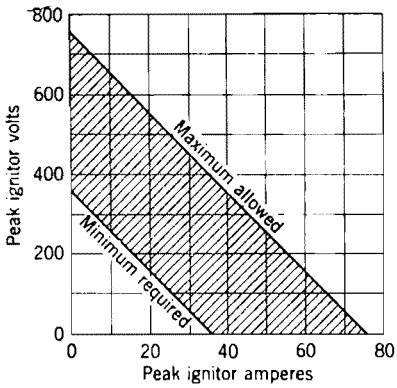



FIG. 15-21. Maximum and minimum voltage and current for firing igniter of type 5555 ignitron.

Ignitron firing circuits are of two forms, *self-excited* from the anode power source, and *separately excited* from an independent source, usually a timed d-c pulse. Examples are illustrated in Fig. 15-22, in which $T_1$ is the ignitron. In Fig. 15-22*a* the rectifier $D_1$ serves to prevent reverse igniter current, and the inductance $L$ limits the peak value of the igniter pulse. When $S$ is closed, the full value of $E_{ss}$ is placed between igniter and cathode pool. The igniter fires the tube, and the voltage between igniter and pool falls to that of the tube drop, which is insufficient to maintain igniter current. Thus a firing impulse is given until the tube fires; then the tube itself cuts off the igniter current. The fuse is present to prevent excessive current should the igniter short to the mercury cathode.

The self-excited circuit allows no control of firing time, and it should also be noted that the igniter current flows through the load, which may not always be desirable.

In Fig. 15-22*b* the igniter element is not connected to its anode voltage until the thyratron $T_2$ is fired by the controlling signal. This circuit has a controlled time of firing, with the self-excitation feature, and is the basis of a synchronous switch which can be used to replace mechanical forms of switches and contactors, giving half-cycle operation.

The circuit in Fig. 15-22*c* is an example of the separately excited form, in which a separate rectifier provides a d-c anode potential for the trig-

gering thyratron $T_2$. Triggering of $T_2$ discharges capacitor $C$ through the thyratron and igniter of $T_1$, giving a short high-current pulse. The capacitor $C$ then recharges during the negative half cycle of ignitron anode voltage. High energy values can be stored in $C$, giving ample firing currents for the ignitron, along with control of timing of the pulse. The thyratron duty is severe in terms of current peak, and a rather large

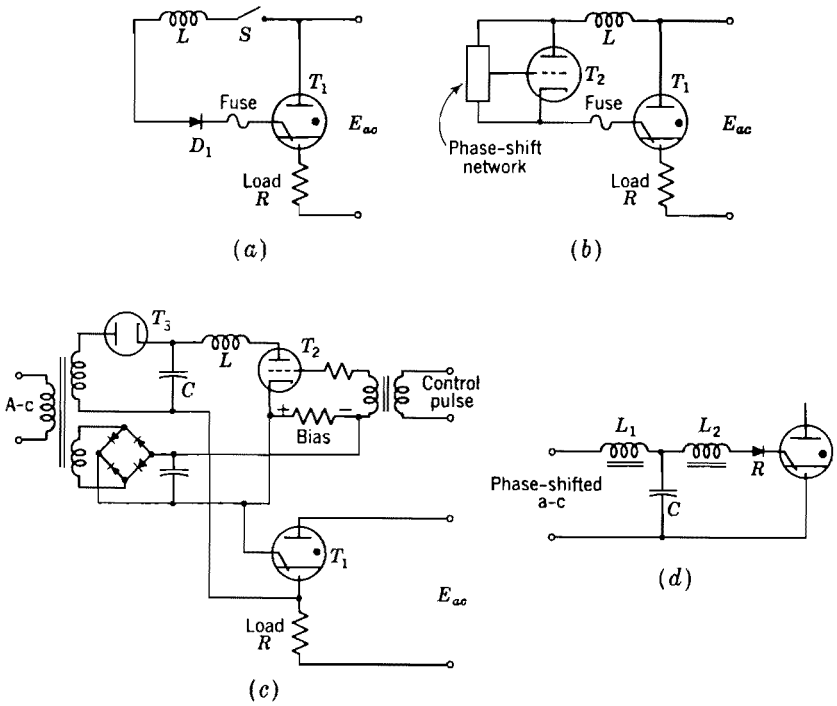



FIG. 15-22. (*a*, *b*) Self-excited firing circuits. (*c*) Separately excited firing circuit. (*d*) Reactor firing.

and expensive tube may be needed, but very positive control of firing is obtained.

In Fig. 15-22*d* is shown another firing circuit for giving a peaked firing current. Capacitor $C$ is charged during the cycle through current-limiting inductor $L_1$. Inductor $L_2$ is designed to readily saturate with the normal peak current. The magnetizing current of this reactor is large and rises sharply. As the voltage across $C$ nears its peak, the reactor saturates on the igniter current and the igniter current rises abruptly. The capacitor then discharges through the reactor, the rectifier, and the igniter, with a high peak value. The rectifier prevents reverse current on the negative cycle. The method is known as *reactor firing*.

**15-13. Controlled Thyratron or Ignitron Rectifiers.** Thyratron or ignitron tubes employed in the conventional rectifier circuits, but with control of firing time, will provide a source of variable d-c voltage. Performance with resistance load will be similar to that of diodes, except that portions of the current waveform may not be present, and overlap conditions may differ somewhat.

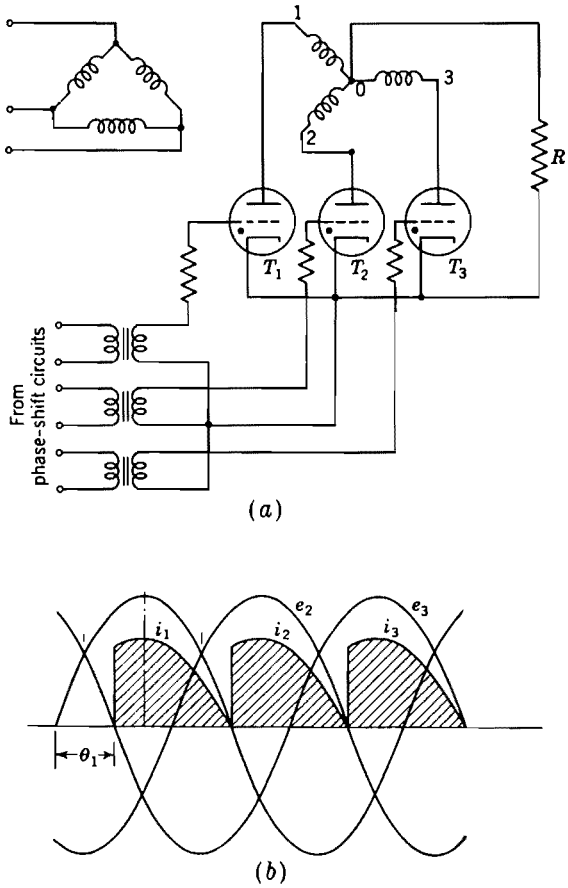



FIG. 15-23. (*a*) Three-phase controlled rectifier. (*b*) Current pulses in load of (*a*).

For the three-phase half-wave circuit of Fig. 15-23 as an example, the firing angle may be delayed by an angle $\theta_1$, although $\theta_1$ may never be less than $\pi/6 = 30°$ since for smaller values the anode of the particular tube is not positive to its cathode. Neglecting transformer reactance, and with $\pi/6 < \theta_1 < \pi/3$, the current conduction will be continuous, and the current of $T_1$ will continue to flow past the angle $\pi/6 = 30°$ until $T_2$

is allowed to fire. With $\theta_1 > 60°$ current flow will be discontinuous for the resistive-load case, and tube $T_1$ will carry current out to $\pi/3$ as is shown in Fig. 15-23*b*. Because of variations in angle it is best to calculate the performance for each operating condition directly by the methods of Chap. 14, using the particular conduction angles which apply. The results of such calculation appear in the curve of Fig. 15-24 for resistive load on the three-phase half-wave rectifier.

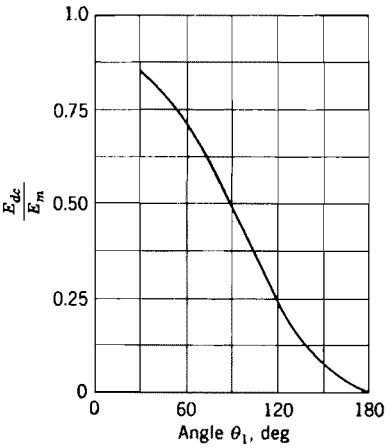
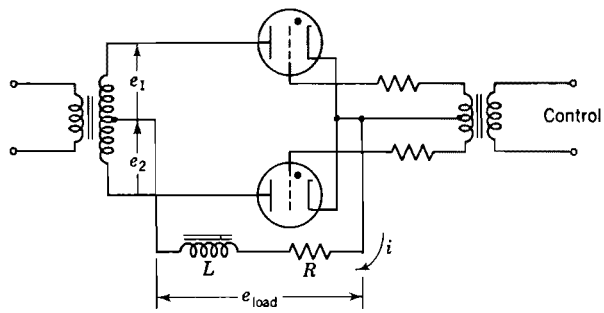
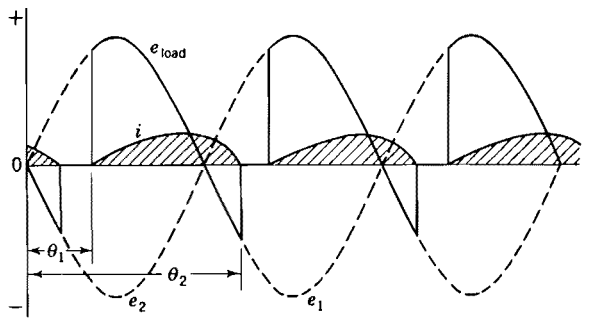



FIG. 15-24. Variation of $E_{dc}$ for three-phase half-wave circuit with $\theta_1 > 30°$.

Very commonly the controlled rectifier is used to supply variable d-c




(*a*)




(*b*)

FIG. 15-25. (*a*) Controlled rectifier with *RL* load. (*b*) Waveforms of load current and voltage with discontinuous conduction.

voltage to loads containing both resistance and inductance as in Fig. 15-25*a*. With inductance present conduction will continue after the cycle zero. If firing of the next tube is sufficiently delayed, then there may be a period of zero conduction and zero load voltage, as demonstrated in Fig. 15-25*b* for the full-wave circuits or Fig. 15-26 for the half-wave circuit. For earlier firing, or for large inductance, the current flow will be continuous, or the firing angle $\theta_1$ is equal to and determines the cutoff angle $\theta_2$ for the preceding tube.

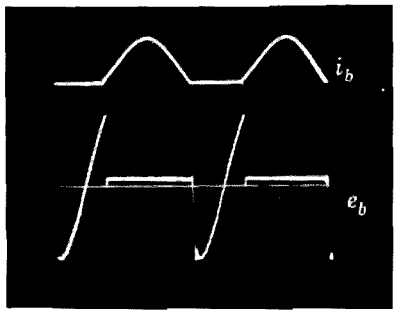



FIG. 15-26. Half-wave controlled rectifier current and tube voltage with delayed firing.

If tube drop is considered negligible, then for the circuit of Fig. 15-25*a*

$$\frac{di}{dt} + \frac{Ri}{L} = \frac{E_m}{L} \sin \omega t \quad (15\text{-}14)$$

similar to the expression in Sec. 10-5. By conventional methods the complete solution for this differential equation is

$$i_b = A\epsilon^{-R\omega t/\omega L} + \frac{E_m}{\sqrt{R^2 + \omega^2 L^2}} \sin (\omega t - \phi) \quad (15\text{-}15)$$

where

$$\phi = \tan^{-1} \frac{\omega L}{R} \quad (15\text{-}16)$$

For the condition of discontinuous current flow $\theta_1 > \theta_2 - \pi$, and the above equation is useful only in the range $\theta_1 < \omega t < \theta_2$, as shown in Fig. 15-25*b*. For these conditions $i_b = 0$ at $\omega t = \theta_1$, so that

$$A = \frac{-E_m}{\sqrt{R^2 + \omega^2 L^2}} \epsilon^{R\theta_1/\omega L} \sin (\theta_1 - \phi)$$

and then

$$i_b = \frac{E_m}{\sqrt{R^2 + \omega^2 L^2}} [\sin (\omega t - \phi) - \epsilon^{-(R/\omega L)(\omega t - \theta_1)} \sin (\theta_1 - \phi)] \quad (15\text{-}17)$$

describes the current waveform in Fig. 15-25*b*.

The current is again zero at $\omega t = \theta_2$, and setting Eq. (15-17) equal to zero leads to

$$\epsilon^{R\theta_2/\omega L} \sin (\theta_2 - \phi) = \epsilon^{R\theta_1/\omega L} \sin (\theta_1 - \phi) \quad (15\text{-}18)$$

This is a transcendental relation which can be solved graphically in terms of the circuit parameter $\phi$ and the independent variable $\theta_1$. The result is plotted in Fig. 15-27 and shows that, for greater delay in firing angle $\theta_1$, the cutoff angle $\theta_2$ lies earlier in the cycle. It also indicates that a recti-

fier with highly inductive load cannot be fired later than 90° in the anode cycle when carrying discontinuous current pulses.

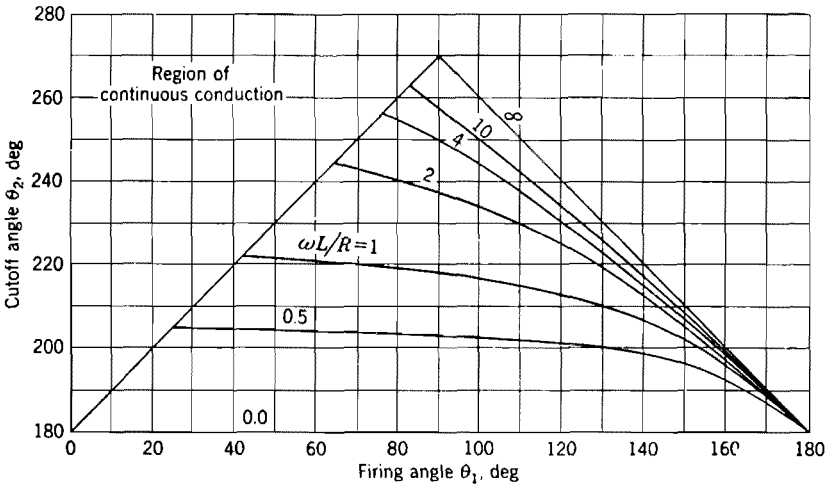



FIG. 15-27. Variation of cutoff angle $\theta_2$ with time of firing $\theta_1$, for various $\omega L/R$ ratios, for discontinuous conduction.

In the range of firing angles giving discontinuous conduction, or within the triangular area of Fig. 15-27, the direct current in the $RL$ load of a single-phase full-wave controlled rectifier can be found as

$$I_{dc} = \frac{E_m}{\sqrt{R^2 + \omega^2 L^2}} \int_{\theta_1}^{\theta_2} [\sin(\omega t - \phi) - \epsilon^{-(R/\omega L)(\omega t - \theta_1)} \sin(\theta_1 - \phi)]\, d\omega t \tag{15-19}$$

After use of Eq. (15-18) and certain trigonometric relations derived from Eq. (15-16)

$$I_{dc} = \frac{E_m}{\pi R} (\cos\theta_1 - \cos\theta_2) \tag{15-20}$$

There can be no average voltage across an inductance so that the voltage across $R$ is

$$E_{dc} = \frac{E_m}{\pi} (\cos\theta_1 - \cos\theta_2) \tag{15-21}$$

in the region of discontinuous conduction.

With firing angles giving continuous conduction the current at $\theta_1$ equals the current at $\theta_1 + \pi$ for the full-wave circuit, and the average current may be found as

$$I_{dc} = \frac{2E_m}{\pi R} \cos\theta_1$$

and

$$\frac{E_{dc}}{E_m} = \frac{2}{\pi} \cos\theta_1 \tag{15-22}$$

At the point at which conduction just becomes continuous $\theta_2 = \pi + \theta_1$. For this case Eq. (15-21) reduces to Eq. (15-22), as it should.

The results of Eqs. (15-21) and (15-22) are consolidated in Fig. 15-28, in terms of $\omega L/R$ of the load and the ignition angle $\theta_1$. The action of the gas-tube rectifier (either thyratron or ignitron) is very much dependent on the type of load. For controlled rectifiers supplying d-c motor-field circuits, and the like, these results are important.

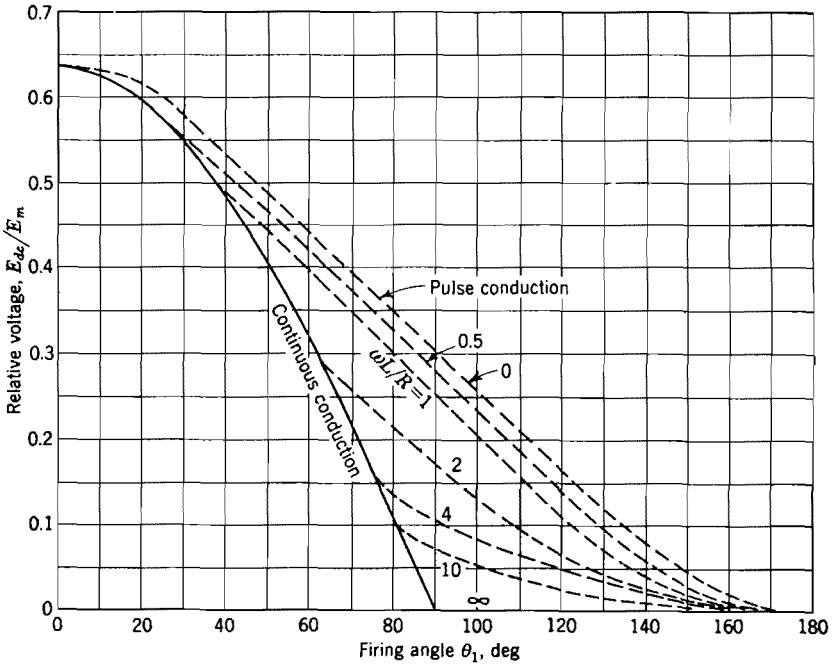



FIG. 15-28. Output voltage for firing angle $\theta_1$ and various $\omega L/R$ ratios, full-wave rectifier.

Another circuit suitable for supplying highly inductive loads is shown in Fig. 15-29*a*. If thyratron $T_1$ fires at $\theta_1$ as shown in Fig. 15-29*b*, it would normally conduct to some angle greater than 180°. However, as the current in $L$ starts to decrease, the voltage across the load reverses because of $L\,di/dt$ of the inductance, and this voltage makes the anode of $T_2$ positive to its cathode and the diode conducts. The energy stored in the inductance is dissipated, causing current to flow in the inductor and load over most of the negative half cycle, and possibly until $T_1$ fires again. The earlier $T_1$ fires in each cycle, the greater is the stored energy. This causes a greater current during the period when $T_1$ is off.

The circuit requires only one expensive thyratron and yet achieves a continuity of current flow comparable with the full-wave circuit. It also

makes a center-tapped transformer unnecessary, and there is no tendency to go from continuous to intermittent conduction. However, the circuit does require use of a larger thyratron, since the conduction time of $T_1$ is shortened and it must still carry the same average current in a shorter period.

The circuit is frequently employed with the thyratron operating into the highly inductive d-c coils of saturable reactors or of motor fields.

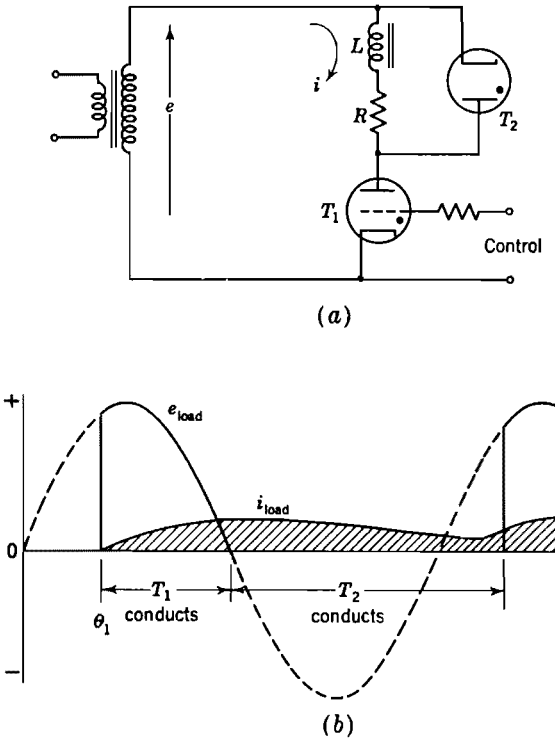



FIG. 15-29. Follow-up circuit for highly inductive load.

**15-14. Control of Alternating Current.** By use of two gas tubes in a back-to-back or inverse parallel connection, load circuits requiring alternating current can be supplied and controlled. The tubes are usually employed as a high-speed nonmechanical switch, which will operate synchronously with the supply voltage or current. The firing angle is controllable, although this causes transients in the load which are frequently undesirable, and it is more common to use the grid or igniter simply to turn the tubes on or off, to give full half-cycle conduction in each tube of a pair.

Figure 15-30 illustrates such a back-to-back connection for a pair of ignitron tubes with self-excitation, used as a synchronous contactor to turn a resistance-welding machine on and off. Operation of the control may be discussed by considering an instant when the anode of $T_2$ has

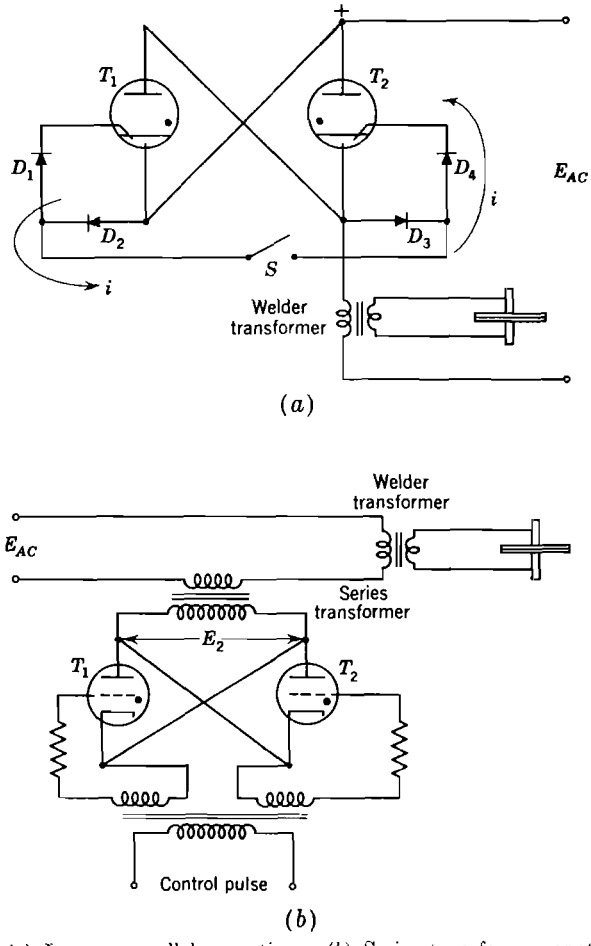



FIG. 15-30. (*a*) Inverse parallel operation. (*b*) Series transformer control of alternating current.

just gone positive. Current $i$ flows as indicated through diodes $D_2$ and $D_4$ to the igniter of $T_2$. When conduction starts in $T_2$, the voltage applied to the igniter drops to less than the tube drop and igniter current ceases. During this process diode $D_3$ has prevented a short circuit across tube $T_2$. On the next half cycle current flows through $D_3$ and $D_1$ to the igniter of $T_1$, a short circuit is prevented by diode $D_2$, and $T_1$ fires.

Therefore conduction will start upon closing $S$ and cease at the end of the first half cycle after $S$ is opened. If $S$ is operated by an electronic circuit to count cycles, then conduction periods of one or more integral cycles may be applied to the welder transformer primary and accurate control of the time and heat developed in the weld can be obtained. This circuit will be further discussed in Sec. 15-15.

In the back-to-back circuit the tubes are required to handle the full line current of the load. The circuit of Fig. 15-30*b* uses a form of impedance matching, so that line currents considerably above tube-current ratings can be handled. This is possible because thyratrons frequently have current ratings below the values taken by loads, but thyratron-voltage ratings are usually considerably higher than load voltages. The series transformer thus matches the relatively high-impedance thyratrons to the somewhat lower-impedance load.

The series transformer has a primary rated for line voltage and a step-up ratio such that, with normal load current flowing, full rated thyratron current will flow in the secondary. When the thyratrons are nonconducting, the series transformer has a high impedance and only the magnetizing current flows in the primary, producing a small voltage across the welder transformer. When the thyratrons are fired, they effectively short circuit the transformer and place nearly full line voltage across the load. The ratio of load current to tube current is equal to the $N_2/N_1$ ratio of the transformer.

**Example.** A pair of thyratrons rated 15 amp average each, 15,000 PIV, are to control a 450-volt 150-kva welder. Calculate the rating of the series transformer and the tube loads.

The welder line current is

$$I_{rms} = \frac{150{,}000}{450} = 334 \text{ amp}$$

Since the d-c tube current $I_{dc}$ and the tube peak current $I_m$ are related as

$$I_{dc} = \frac{I_m}{\pi}$$

then to operate the tubes at full current ratings of $I_{dc} = 15$ amp per tube will require that the peak current per tube be

$$I_m = 15\pi = 47.2 \text{ amp}$$

assuming full half-cycle current conduction per tube. The secondary rms current is then

$$I_2 = \frac{I_m}{\sqrt{2}} = \frac{47.2}{\sqrt{2}} = 33.4 \text{ amp}$$

The transformer turns ratio is

$$\frac{1}{a} = \frac{N_2}{N_1} = \frac{I_1}{I_2} = \frac{334}{33.4} = 10$$

The secondary a-c voltage is then $450 \times 10 = 4500$ volts. The tube peak inverse voltage when not firing is $\sqrt{2} \times 4500 = 6200$ peak volts, well within the rated value. Therefore the transformer should be rated 450/4500 volts, 150 kva. Each tube will carry 15 amp average.

**15-15. Ignitron Ratings.** Ignitron tubes are largely employed in two types of industrial service, either as controlled rectifiers or in intermittent operation as welding-machine contactors in the inverse parallel circuit.

Ratings of ignitrons for controlled rectifier service are essentially of the same nature as those for gas diodes, with the addition of ratings for the igniter. The igniter requires a large current at low voltage or a small current at high voltage to ensure firing, as shown in Fig. 15-21. There are also ratings on *maximum instantaneous igniter potential*, both positive and negative, *maximum instantaneous igniter current*, and *maximum average igniter current*, all as top limits on operation.

For a-c welder control service the tubes behave as a high-speed switch, as mentioned in Sec. 15-14. In resistance or spot welding the welding transformer electrodes pass heavy currents through the metal sheets to be joined, rapidly generating heat at the junction and causing a weld. The machine is turned on and off to form successive welded spots in the metal sheets, overlapping as in seam welding or widely spaced as determined by the operation. Recurrent heavy-duty operation of this kind leads to mechanical trouble in the usual contactors, and accurate timing of the "on" time and heat generated in the weld is not possible for short weld times of a few cycles. Ignitron tubes, without mechanical moving parts and with their ability to turn on and off in integral half cycles, are admirably adapted to replace the mechanical contactor.

For such intermittent tube service the tube ratings must be modified to ensure economic tube application. Allowance may be made for the cooling possible during "off" periods, and heavier currents carried for the short "on" periods.

It is assumed that the tubes conduct for full half cycles, and the output of the two tubes in inverse parallel is then normal alternating current. The *rms demand current* is this effective a-c value carried by the pair of tubes.

The product of supply voltage and rms demand current is the *kva demand*, and this has a maximum value stated by the rating of the load or welder controlled.

The greatest d-c current allowable per tube for *continuous operation* is called the *maximum average current*, $I_{dc,max}$, and is stated in the tube ratings. The actual direct current in on-off contactor operation may be increased over $I_{dc,max}$ if allowance is made for adequate cooling periods. The rated value for $I_{dc,max}$ is determined by the ability of the tube to dissipate its losses in the cooling water and not by possible cathode emission, since with a pool cathode the emission may be very large.

Gas-tube losses are a linear function of average current, and if these losses are averaged over a period known as the *maximum averaging time*, the current during conduction may greatly exceed $I_{dc,max}$ for continuous conduction. That is, the maximum average current $I_{dc,max}$ may be exceeded as a d-c current value but may not be exceeded as an average current over the stated period. If a given tube has an averaging time of 6 sec or 360 cycles and $I_{dc,max} = 50$ amp, the tube may carry 50 amp continuously, 100 amp for 180 cycles out of each 360, or 200 amp for 90 cycles out of each 360, and so on.

This system of rating recognizes that with a pool cathode the ratio of peak to average current may be very high and that a tube has a certain thermal storage capacity. Rigorously the amount of heat storage allowable should be stated in watt-seconds, but since the tube drop is assumed constant, the heat loss is proportional to average current, and thus the thermal capacity of a tube is expressed in ampere-seconds. For a given tube the thermal capacity can be stated by $I_{dc,max}T_a$, where $T_a$ is the rated averaging time in seconds. While primarily a heat-storage factor, averaging time as usually employed is also a function of operating voltage, since the severity of operation is affected by voltage.

The *per cent duty* is the ratio of "on" time to maximum averaging time in a particular application. It may be calculated from the demand kva starting with

$$I_{rms} = \frac{\text{volt-ampere demand}}{E_{ac}} \tag{15-23}$$

For half-wave operation per tube in the back-to-back circuit

$$I_m = \sqrt{2}\, I_{rms}$$

If $I_{dc}$ is the average current per tube flowing during a given conduction interval of length $T_c$, then

$$I_{dc} = \frac{I_m}{\pi} = \frac{\sqrt{2}\, I_{rms}}{\pi} \tag{15-24}$$

If $T_c$ is the actual conduction time during a given averaging interval, the actual ampere-second losses may equal but not exceed the rated ampere-second capacity, or

$$I_{dc,max}T_a = I_{dc}T_c = \frac{\sqrt{2}\, I_{rms}T_c}{\pi} \tag{15-25}$$

or

$$T_c = \frac{\pi I_{dc,max}T_a}{\sqrt{2}\, I_{rms}} \tag{15-26}$$

and the allowable per cent duty of the tubes is then

$$\text{Per cent duty} = \frac{T_c}{T_a} \times 100\% = \frac{\pi I_{dc,max}}{\sqrt{2}\, I_{rms}} \times 100\% \tag{15-27}$$

This fixes the allowable percentage of "on" time, during any given averaging time, to avoid tube overheating. The per cent duty is frequently measured in terms of "on" cycles during an averaging time of a certain number of cycles. In case the averaging time is shorter than a given welding operation, the per cent duty is always figured on the worst possible loading situation.

To ensure coverage of the range of welder sizes by a coordinated line of tubes, the industry has set up formal rating charts, to which standardized tube sizes have been designed. These charts divide the welder kva demands into four ranges which are served by four tube designs, known as sizes A, B, C, and D. These tubes have been designed for nominal 250- and 500-volt service and have the ratings given in Table 31.

TABLE 31. IGNITRON RATINGS FOR WELDER OPERATION

| Size | $I_{dc,max}$, amp | Nominal $I_{rms}$/pair | Rated averaging time, sec | Maximum $I_{rms}$/pair |
|---|---|---|---|---|
| | | | At 250 Volts | |
| A | 22.4 | 50 | 22 | 1200 |
| B | 56 | 125 | 18 | 2400 |
| C | 130 | 300 | 14 | 4800 |
| D | 355 | 800 | 11 | 9600 |
| | | | At 500 Volts | |
| A | 22.4 | 50 | 11 | 600 |
| B | 56 | 125 | 9 | 1200 |
| C | 130 | 300 | 7 | 2400 |
| D | 355 | 800 | 5.5 | 4800 |

It may be noted that the average current ratings increase by a factor of 2.5 per size, whereas the maximum $I_{rms}$ ratings increase by a factor of 2.

Certain other features of tube rating and application have also been standardized through industry action, among these being the per cent duty curve. Equation (15-27) represents a hyperbolic relation between per cent duty and $I_{rms}$ demand in the a-c line, with $I_{dc,max}$ and $I_{rms,max}$ as rated constants. This relation may be plotted as in Fig. 15-31*a*, for a size A tube. However, it becomes a more convenient tube-performance characteristic if placed on log-log paper, when the curve becomes a straight line as in Fig. 15-31*b*.

It has been found by experience that the per cent duty at the higher rms currents should be reduced, to give somewhat greater cooling time and longer tube life, and this has been done by arbitrary definition. The reduction begins at a value of 33 per cent of maximum rms amperes

at $X$ in Fig. 15-31*b* and takes a new slope up to $Z$, at which point the per cent duty is reduced to 54 per cent of the value it would have had at $Y$. This reduction is illustrated by the dashed curve of Fig. 15-31*b*.

In Fig. 15-32 the family of such curves is drawn for all four ignitron sizes for both 250 and 500 volts. Percentage of duty determined from these curves multiplied by the averaging time in cycles gives the largest number of cycles conduction allowed under the ratings at the given

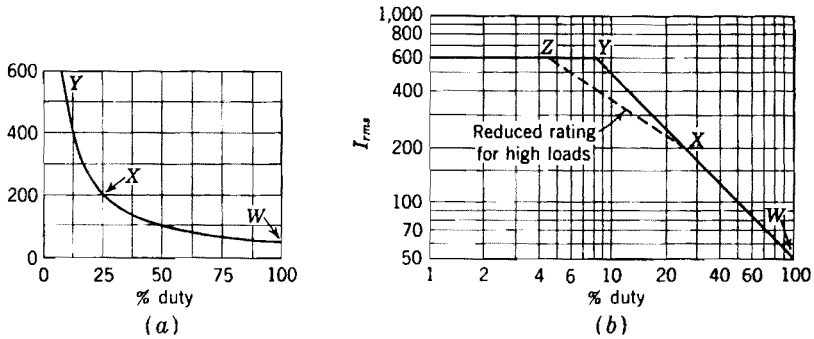



FIG. 15-31. Methods of plotting per cent duty rating curves for a size A ignitron.

demand current. The same length current conduction can be repeated at an interval equal to the averaging time.

**Example.** A particular spot-welding job requires six spots, each 6 cycles long, repeated at intervals of 30 cycles. The current demand is 950 amp rms, supplied at 500 volts. Find the ignitron size required if a period of 7 sec is required to adjust the next workpiece.

The actual welding time is $6 \times 6 = 36$ cycles. The "off" time during the welding period is $5 \times 30 = 150$ cycles; therefore the total time for the operation is 186 cycles, to which may be added the down time of $7 \times 60 = 420$ cycles, for a total time per operation of 606 cycles = 10.1 sec.

The per cent duty, if size B tubes with an averaging time of 9 sec are used, is

$$\text{Per cent duty} = {}^{36}\!/_{540} \times 100\% = 6.7\%$$

From Fig. 15-32*b* size B tubes may carry 1060 amp at 6.7 per cent duty and are suitable for the operation. Size C tubes, with an averaging time of 7 sec, might also be considered, since 7 sec exceeds the actual operating time of 186 cycles and is less than the total operating time. Then for size C tubes

$$\text{Per cent duty} = {}^{36}\!/_{420} \times 100\% = 8.6\%$$

Size C tubes can carry 2000 amp rms demand at this duty and might be considered large for the job.

**15-16. Switching Gas Triodes with D-C Anode Potential.** Thyratrons can be operated with d-c anode potential, and used as d-c switches, if circuits are devised that will develop transient voltages capable of driving

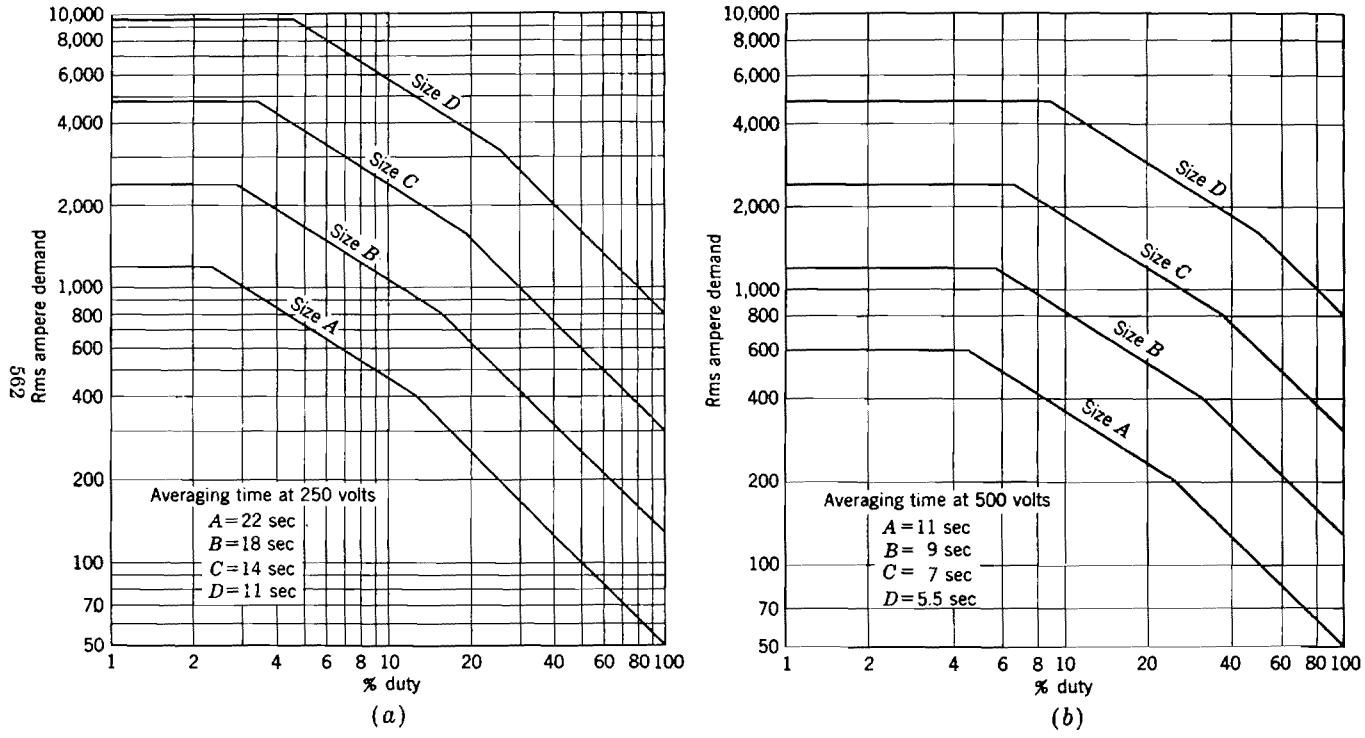



FIG. 15-32. (*a*) Per cent duty for two tubes in inverse parallel at 250 volts rms. (*b*) Same for two tubes at 500 volts rms.

the anode negative to the cathode for a time long enough for deionization to take place. The circuit in Fig. 15-33*a* is an example.

Assume that the thyratron has been fired by a positive impulse on its grid, reducing the potential at the anode, or point $A$, to the tube drop, or about 15 volts. No current flows in $R_2$, and the potential of $B$ is that of the supply, or +200 volts, so that $B$ is 185 volts above $A$, and this is the capacitor potential.

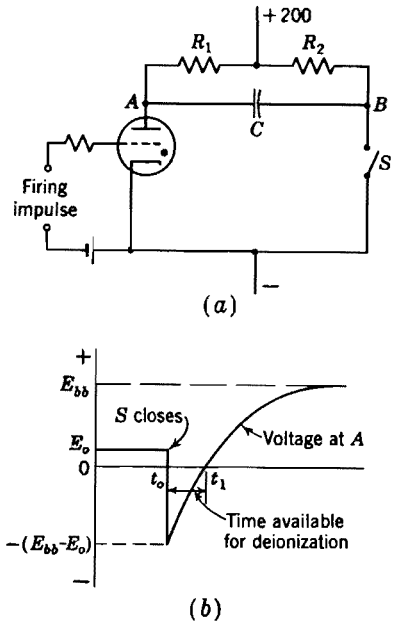



FIG. 15-33. (*a*) Circuit for switching a thyratron with d-c anode voltage. (*b*) potential on anode at $A$.

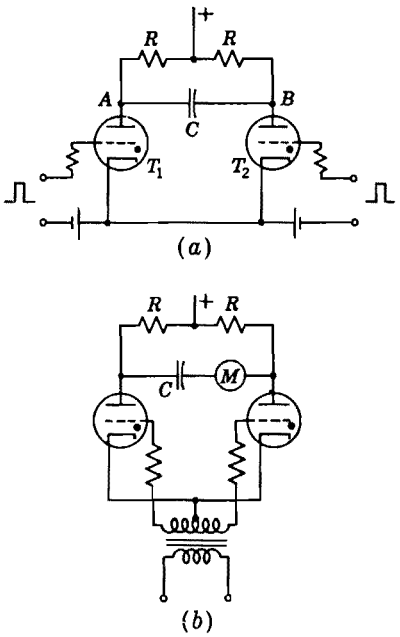



FIG. 15-34. (*a*) Tube switching with d-c anode potential. (*b*) Frequency meter.

If switch $S$ is closed, point $B$ is connected to the cathode. Since $C$ cannot discharge instantaneously, point $A$ will continue to be at −185 volts to $B$, but this means that point $A$ is now −185 volts with respect to cathode as indicated in Fig. 15-33*b*. Capacitor $C$ now starts to discharge, and its potential goes toward zero. During the interval from the closure of $S$ at $t_o$ to time $t_1$ when $A$ is at zero, the gas in the thyratron will deionize, and the grid will regain control. The anode voltage may rise toward the $E_{bb}$ value, but the arc will not restrike until another grid impulse is received. The time interval $t_1 - t_o$ is dependent on the $(R_1 + R_2)C$ time constant and must be greater than the tube-deionization time.

The mechanical switch $S$ may be inconvenient and can be replaced by another thyratron as in Fig. 15-34*a*. The only difference in per-

formance between Figs. 15-33 and 15-34 is that, with tube $T_1$ on, when $T_2$ fires the potential of $B$ goes only to about 15 volts above cathode instead of to zero. This is unimportant in the circuit operation.

If the grids are excited from an alternating voltage as in Fig. 15-34*b*, then each half cycle will cause the tubes to switch, the capacitor $C$ will receive a reversal of charge each time, and every cycle will cause a definite amount of charge to pass through the capacitor. This charge passes only during the transient portions of the anode-voltage wave of Fig. 15-33*b*, and the transient time may be made shorter than the shortest applied half cycle, so that the transient is completed each time. If this is the case, then a full-wave rectifier milliammeter may be placed in series with $C$ and its reading will be proportional to the frequency applied to the tube grids and independent of the applied grid voltage. This makes a simple frequency meter.

**15-17. Gas-triode Counting Circuits.** In Fig. 15-34, after $T_2$ fires it may in turn be extinguished by the firing of $T_1$, because of the symmetry of the circuit. Since a positive grid pulse is required for firing either tube, the grids may be paralleled as in Fig. 15-35. Only the tube which is off at a particular instant will be fired by a positive pulse, the other tube being already conducting and a grid pulse having no effect. Thus a succession of positive grid pulses will cause conduction to pass from $T_1$ to $T_2$ and back to $T_1$ again. Tube $T_2$ will fire for every second grid pulse, and the circuit is a *scale-of-2 counter*.

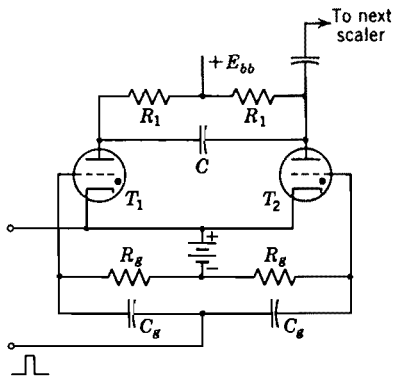


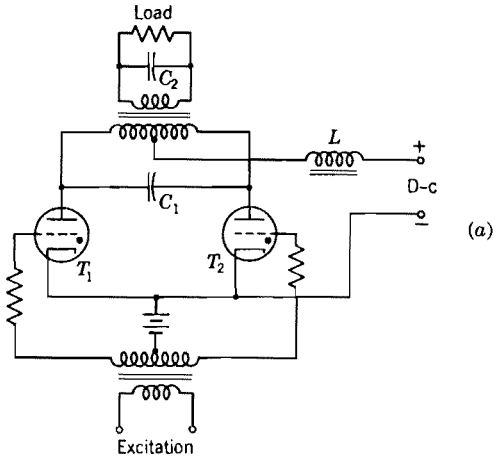

FIG. 15-35. Scale-of-2 circuit.

A positive pulse can be obtained from the anode of $T_2$ every time it is extinguished, and this pulse may be applied to a second circuit, giving a scale of 4, and so on. A basic limit in such operation is the time of deionization of the thyratrons, and such circuits are usually limited to operation at audio-frequency rates. Vacuum tubes are suited to higher rates.

**15-18. Inverters.** A circuit that will change direct to alternating current is called an *inverter*. Gas tubes are well suited to such operation at the lower frequencies, owing to their high power efficiency. The circuits of Fig. 15-34 operate on direct current and provide alternating voltages and may be modified by insertion of an output transformer as in Fig. 15-36.

If $T_1$ is turned on by a positive grid wave, capacitor $C_1$ charges and

current flows through $L$ and half the transformer to $T_1$. On the next half cycle $T_2$ ignites, the charge on $C_1$ instantaneously drives the anode of $T_1$ negative, and $T_1$ deionizes, allowing its grid to regain control. Current flows through the other half of the transformer to $T_2$. This cycle of operation is repeated every grid-voltage cycle, each tube kicking the other off. The transient currents in the transformer induce an



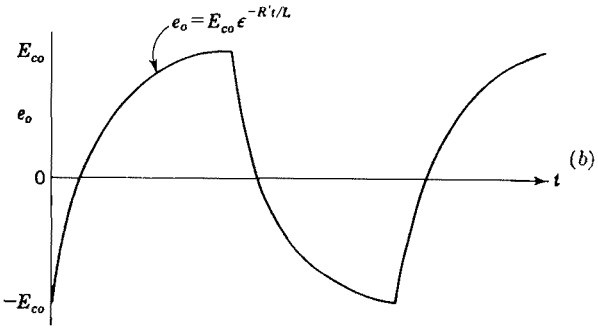




FIG. 15-36. Separately excited parallel inverter.

alternating secondary voltage, and capacitor $C_2$ aids in shaping this to sinusoidal form.

If the transformer is assumed ideal, then $C_1$ charges through the reflected load resistance $R'$ and the inductance $L$, between the limits of $E_{co}$ and $-E_{co}$, as the extremes of capacitor voltage. For the resistance $R'$ small (heavy load) the effect of $C_1$ will be shunted out, as seen in

Fig. 15-37, and the $RL$ time constant will be the determining factor, with $e_o$ of the form

$$e_o = E_{co}\epsilon^{-R't/L} \tag{15-28}$$

Upon repeated firing of the thyratrons the output voltage will appear as a succession of repeated transients, as in Fig. 15-36*b*.

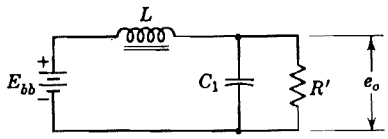



FIG. 15-37. Equivalent circuit for inverter.

For the resistance $R'$ large (light load), then, $R'$ may be neglected with respect to $C_1$, and the circuit becomes a series-resonant one with $\omega_r = 1/\sqrt{LC_1}$. For light load the output will then be nearly sinusoidal. However, the excitation frequency and $\omega_r$ must be carefully adjusted. If the a-c load is reactive, its reflected elements may be considered as parts of $L$ and $C_1$. The load waveform is therefore near sinusoidal for light loads and departs therefrom for heavy loads.

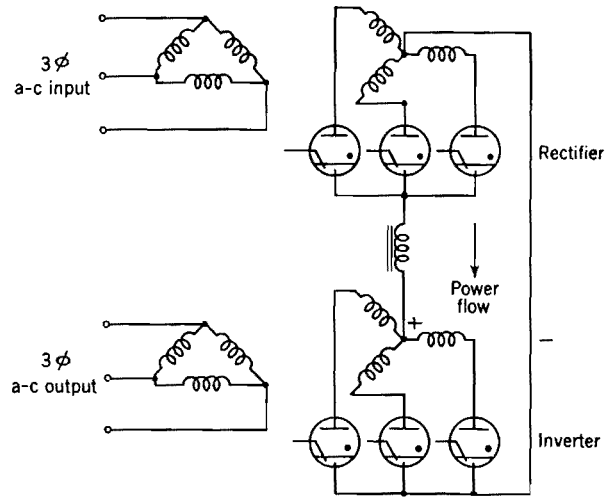



FIG. 15-38. Three-phase rectifier and inverter.

It can be seen that, if one tube fails to start or one fails to turn off, the circuit becomes a short circuit on the d-c input. More complex circuits are employed to avoid this.

By combining a rectifier and an inverter it is possible to have an electronic frequency changer; these have been employed in some steel mills where 25-cycle power is required, and a simplified circuit is shown in Fig. 15-38. Grid excitation for the inverter is obtained from the receiv-

ing system, which is assumed to contain other generating sources. If the d-c link is lengthened, the circuit may be employed for d-c transmission.

**15-19. The Transistor Inverter.** There is frequently a need for conversion of low-voltage direct current, as from storage batteries, into direct current suitable for supplying the high-voltage filter systems of electronic apparatus. This has been done in the past by electromagnetic vibrators, chopping the d-c input and rectifying the a-c output, but contact difficulties may lead to short life. An inverter, or feedback oscillator, employing power transistors with adequate cooling has been found useful, and one form is illustrated in Fig. 15-39.

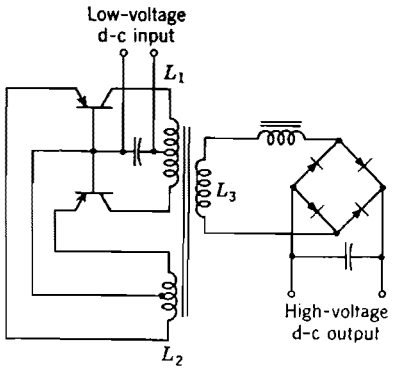



FIG. 15-39. Transistor inverter.

Windings $L_1$ and $L_2$ and the transistors are connected as a common-base push-pull feedback oscillator, winding $L_2$ providing the feedback. The resultant a-c output is coupled out by $L_3$, stepped up to the desired voltage level, and then rectified by the germanium diode bridge. The frequency of operation is determined by the inductance and saturation effects in the transformer and is usually in the 60- to 100-cycle range. The transistors operate in a 50 per cent duty cycle, being cut off half the time, and with good cooling conditions are reported to have given outputs of several hundred watts at over-all efficiencies of 80 per cent.

## PROBLEMS

**15-1.** The critical grid characteristic for a mercury-vapor thyratron is given by

| $e_b$, volts | 10 | 20 | 40 | 80 | 140 | 240 |
|---|---|---|---|---|---|---|
| $e_c$, volts | 0 | −1 | −2 | −3 | −4 | −5 |

With an alternating anode voltage of 120 volts rms, plot the grid control-locus curve. At what angle in the anode cycle will the tube fire for each of the following grid voltage conditions: (*a*) −2 volts, (*b*) −4.5 volts, (*c*) an a-c voltage lagging the anode voltage by 45° with 24-volt amplitude, (*d*) same, lagging 150°?

**15-2.** A thyratron in a half-wave controlled rectifier has a current rating of 2.5 amp peak and 0.6 amp average. With an applied a-c voltage of 240 volts rms determine the lowest value of load resistance that may be used, without exceeding ratings, for a variation of 0 to 90° in firing angle.

**15-3.** The thyratron of Prob. 15-2 is to be controlled by an *RC* phase-shift bridge with $C = 5$ μf. What range of variable $R$ will be needed? If the circuit is made

into an $RL$ type, what value of $L$ will be needed for the same $R$? Supply frequency is 60 cycles.

**15-4.** The type FG-105 thyratron is rated 40 amp peak, 6.4 amp average. What is the greatest delay in the firing angle which may be used if rated average current flows in a resistive load?

**15-5.** The circuit of Fig. 15-40 is to shift the phase of $E_s$ by 90°. The ratio $N_2/N_1 = 3$, and the transformer $T_1$ is ideal. The characteristics of the 12AT7 tube

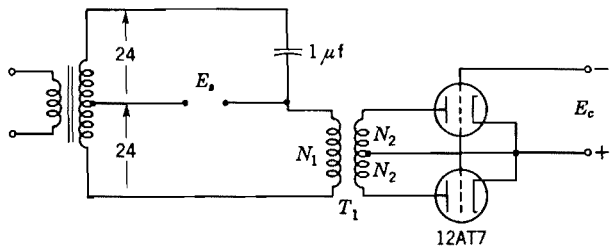



FIG. 15-40

may be used to determine its resistances. Plot a curve of $E_c$ vs. phase shift of $E_s$, with 60-cycle supply, and explain your method of obtaining the vacuum-triode resistances.

**15-6.** The critical grid voltage of a thyratron is assumed zero for all anode voltages. The grid voltage is made up of an a-c voltage of 24 volts peak, lagging 90°, and a d-c bias variable from −24 to +24 volts. Calculate firing angles as a function of d-c grid bias, and plot a curve of per cent maximum average anode current vs. grid bias.

**15-7.** An ignitron with anode supply of 120 volts rms, 60 cycles, applied through a load of 100 ohms, is fired at 50° of the anode cycle. Assuming tube drop of 20 volts, find (*a*) average anode current, (*b*) power input, (*c*) tube loss in watts, (*d*) reading of an a-c ammeter in series with the load.

**15-8.** A full-wave single-phase ignitron rectifier with tube drop of 15 volts charges a 240-volt storage battery of 1.0 ohm resistance. If the transformer supplies 300 volts rms each side of center tap, plot a curve of charging rate as a function of the firing angle.

**15-9.** A 50-kw 600-volt six-phase half-wave rectifier is to operate at 80 per cent load for 8 hr per day and at an average 10 per cent load for 16 hr per day. Six suitable thyratrons have heaters taking 5 volts, 20 amp each, with 12-volt tube drops. Six ignitrons have tube drops of 22 volts, with negligible igniter losses. If power costs 2 cents per kilowatthour, which type of tube will be the cheapest to operate? How much would the annual cost difference be?

**15-10.** A half-wave thyratron circuit operates a solenoid coil of 10 ohms and 0.15 h from a 60-cycle source. The source voltage is 115 volts rms. Find the average and peak currents for a firing angle of 50°.

**15-11.** Starting with Eq. (15-19), derive Eq. (15-22) for the d-c voltage of a full-wave circuit.

**15-12.** A welding transformer which takes 50 kva is to be controlled from a 240-volt 60-cycle circuit by a series transformer in the circuit of Fig. 15-30*b*. If the series-transformer primary voltage is 220 on open secondary and if the tubes available are rated at 15,000 peak inverse volts, find (*a*) average current per tube, (*b*) transformer voltage and kva ratings.

**15-13.** The GL-5553 ignitron is a size C tube. Determine the per cent duty when supplying a 750-kva welder in the back-to-back circuit, and the tube average current during the weld time, at 500 volts.

**15-14.** A spot welder is rated at 250 kva, 500 volts. It is to produce welds with four 10-cycle "on" periods, separated by 20-cycle "off" periods, with a 140-cycle rest period, for a total duty cycle of 4 sec. Determine the size of tubes to be used, and specify their operating average current, their average current over the averaging time, and the demand rms current.

**15-15.** Two ignitrons are used in the back-to-back circuit to supply a welder of 500 kva demand at 250 volts, 60 cycles. The tubes are size A.

(*a*) Find the total number of conduction cycles permitted in any averaging time.

(*b*) If the tube drop is 20 volts and cooling water is supplied at 15°C, what will be the outlet temperature of the water if the tubes are operated repetitively in the cycle of (*a*)?

**15-16.** In the circuit of Fig. 15-33*a* $R_1 = 100$ ohms, $R_2 = 5000$ ohms. The tube is conducting and is to be extinguished by closing $S$. The arc drop is 20 volts, and the deionization time is 500 $\mu$sec. How large must $C$ be to ensure that the arc will not restrike after closing $S$, assuming that the arc may restrike when the anode is positive by the amount of the tube drop, if ionization is present.

## REFERENCES


1. Hull, A. W.: Hot Cathode Thyratrons, *Gen. Elec. Rev.*, **32**: 213, 390 (1929).
2. Slepian, J., and L. R. Ludwig: A New Method of Initiating the Cathode of an Arc, *Trans. AIEE*, **52**: 693 (1933).
3. Nottingham, W. B.: Characteristics of Small Grid-controlled Hot-cathode Mercury Arcs or Thyratrons, *J. Franklin Inst.*, **211**: 271 (1931).
4. Steiner, H. C., A. C. Gable, and H. T. Maser: Engineering Features of Gas-filled Tubes, *Elec. Eng.*, **51**: 313 (1932).
5. Cobine, J. D.: "Gaseous Conductors," McGraw-Hill Book Company, Inc., New York, 1941.
6. Herskind, C. C., and E. J. Remscheid: Excitation, Control, and Cooling of Ignitrons, *Trans. AIEE*, **65**: 632 (1946).
7. Myers, H. C., and J. H. Cox: Excitation Circuits for Ignitron Rectifiers, *Trans. AIEE*, **60**: 943 (1941).
8. Holubow, H.: DC Saturable Reactors for Control Purposes, *Electronic Ind.*, March, 1945.
9. Palmer, H. L.: Electronic Welding Control, *Electronics*, **10**: 20 (June, 1937).
10. Bivens, M. E.: Seam and Pulsation Welding Controls, *Electronics*, **15**: 55 (September, 1942).
11. Tompkins, F. N.: The Parallel Type of Inverter, *Elec. Eng.*, **52**: 253 (1933).
12. Willis, C. H., B. D. Bedford, and F. R. Elder: Constant-current D-C Transmission, *Elec. Eng.*, **54**: 102 (1935).
13. Willis, C. H., R. W. Kuenning, E. F. Christensen, and B. D. Bedford: Design of Electronic Frequency Changers, *Trans. AIEE*, **63**: 1070 (1944).
14. Wagner, C. F.: Parallel Inverter with Inductive Load, *Elec. Eng.*, **55**: 970 (1936).

CHAPTER 16

# RELAYS, TIMERS, AND RESISTANCE-WELDING CONTROLS

Relays are used where a small amount of power must control a larger power, and this is a requirement well met by the electron tube. Important applications of vacuum and gas tubes are found in the operation of electrical contacts with small current and in the timing of various operations.

For time intervals of a few seconds to many minutes it is common to employ a synchronous-motor-driven switch. For less accurate control a bimetal strip and heating element or a thermistor may serve over the same range of time. With time intervals of a minute down to a fraction of a second the vacuum tube controlled by transient voltages in an $RC$ circuit is usual, and for fractions of a second down to millisecond or microsecond intervals various forms of counter circuits, oscillators, and multivibrators are employed.

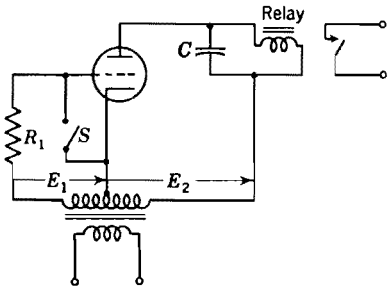


FIG. 16-1. Contact relay.

**16-1. The Contact Relay.** Very delicate contacts, such as those sometimes employed with electrical instruments, are incapable of carrying appreciable current. These delicate contacts may be employed at $S$ in the grid circuit of Fig. 16-1, in series with resistor $R_1$ of many megohms, to operate the plate-circuit relay.

Voltage $E_1$ is greater than $E_2/\mu$ so that with $S$ open the grid is always sufficiently negative to cut off the tube. When $S$ closes, the grid will be connected to cathode and plate current will flow. The capacitor across the relay serves to smooth the half-wave rectified-plate-current pulses.

By placing $S$ in series with $R_1$ the circuit can be made to open the relay when contact $S$ closes.

**16-2. D-C-operated Time-delay Relays.** The properties of $RC$ circuits may be used to obtain time-delay characteristics in the operation of relays, as illustrated in Fig. 16-2*a*. With $S$ closed $E_{cc}$ is large enough

to cut off the tube. Capacitor $C$ will charge to $E_{cc}$ value, and when $S$ is opened, $C$ will begin to discharge through $R$, with its potential determined by

$$e_C = E_{cc}\epsilon^{-t/RC}$$

where both $E_{cc}$ and $e_C$ will be negative quantities. As $e_C$ drops, the bias on the grid becomes less negative, greater plate current flows, and after a time $t_1$ the plate current reaches the pickup value for the relay. If the grid bias at this plate current is $E_1$, as determined by a load line for the relay resistance, the time delay after operation of $S$ is given by

$$T = RC \ln \frac{E_{cc}}{E_1} \quad \text{sec} \quad (16\text{-}1)$$

under the assumptions that $E_{bb}$ is constant and relay inductance and capacitor leakage are negligible.

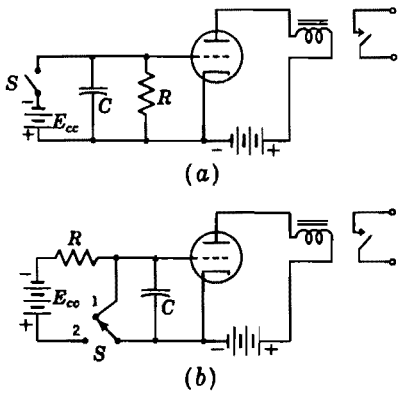



FIG. 16-2. $RC$-operated timing circuits.

In Fig. 16-2*b* there appears a circuit using a series $RC$ circuit in which the capacitor is shorted and relay closed with $S$ in position 1. When $S$ is switched to 2, the capacitor voltage $e_C$ starts to go negative in accordance with

$$e_C = E_{cc}(1 - \epsilon^{-t/RC})$$

After a delay depending on the time constant the plate current reaches the drop-out current of the relay and it opens. If $E_2$ is the grid voltage giving the drop-out plate current as determined from a load line, then

$$T = RC \ln \frac{E_{cc}}{E_{cc} - E_2} \quad \text{sec} \quad (16\text{-}2)$$

where $E_{cc}$ and $E_2$ are usually negative.

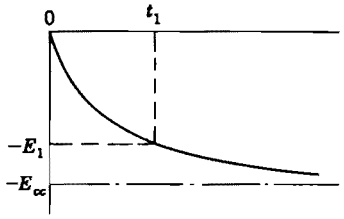



FIG. 16-3. Exponential variation of grid potentials for Fig. 16-2*a*.

Circuits such as these are not precise when used for delay times above a few seconds because they will involve operation far to the right on the curve of Fig. 16-3, where the slope is very flat. Variation of $E_{bb}$ and $E_{cc}$ then affects the timing.

If the $RC$ value is increased to restore operation to a steeper slope, precision still may not be achieved. If $RC$ is increased by raising $R$, then the uncertain grid current of the tube may be appreciable with

respect to the charging current. If $C$ is increased, its leakage current may increase and adversely affect accuracy.

**16-3. A-C-operated Time-delay Relays.** For industrial use it is much more convenient and economical to have relays and timers operate directly from the a-c line, without rectifier-power-supply complexities. The fundamentals of two such circuits are illustrated in Fig. 16-4.

In Fig. 16-4*a*, with switch $S$ closed, the voltage at $A$ when $C$ is positive is sufficient to keep the tube cut off. However, with $S$ closed, grid rectification occurs on the other half cycles when point $A$ is positive, and the $R_1C_1$ circuit clamps the grid-voltage peaks at zero or develops a negative voltage to the grid as shown across $C_1$. The $R_1C_1$ time constant is made long with respect to the time of a cycle, so that with $S$ closed the grid

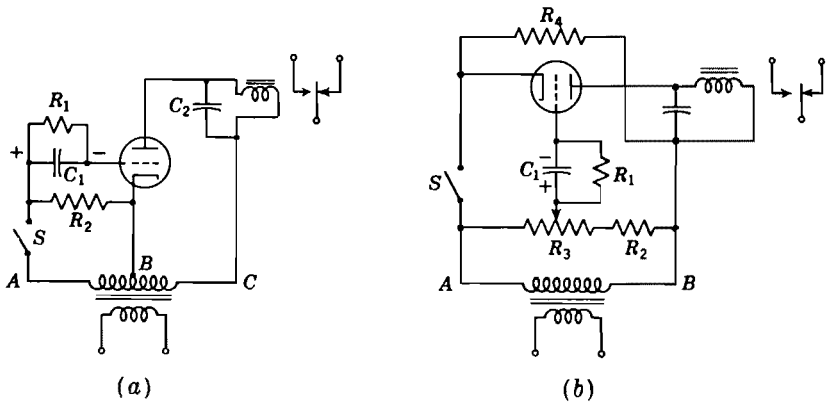



FIG. 16-4. Alternating-current-operated time-delay relays.

acquires a negative d-c bias approximating the peak of voltage $e_{AB}$, sufficient to cut off the tube and keep the relay nonoperated.

When $S$ is opened, the grid circuit is cut loose from the a-c input and connected to cathode through $R_2$, of possibly 50,000 ohms. Capacitor $C_1$ discharges through $R_1$, and the grid voltage decays exponentially. The grid potential rises above cutoff, the plate current flowing in pulses, and soon these have a sufficiently large average value to pull in the relay. Time delay is fixed by choice of $R_1$ and $C_1$.

The circuit in Fig. 16-4*b* is similar but adds an a-c voltage in the grid circuit which improves the accuracy for short time intervals by tending to synchronize operation with integral cycles of the input voltage. This circuit gives a time delay after $S$ closes rather than after $S$ opens, as in the previous circuit. With $S$ open, grid rectification occurs on positive-anode half cycles from point $A$ through the $R_1C_1$ circuit to grid and cathode, with return through $R_4$. Capacitor $C_1$ charges quickly to the

peak of the a-c voltage applied across $R_2$ and the portion of $R_3$ and biases the tube to cutoff.

When $S$ is closed, the negative bias across $C_1$ is applied directly between grid and cathode with an a-c component due to the drop across the included portion of $R_3$, as illustrated at time $t = 0$ in Fig. 16-5. It can be seen that as $C_1$ discharges through $R_1$ the average grid voltage rises, and eventually the positive peaks go above the tube cutoff value, plate-current pulses appear, and the relay operates.

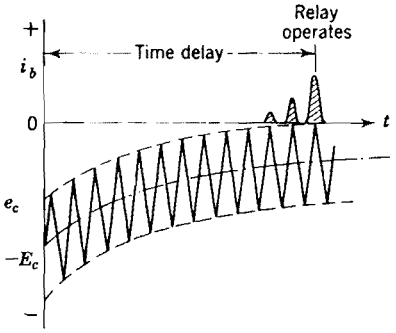



FIG. 16-5. Grid voltage in circuit of Fig. 16-4*b* after $S$ is closed.

By moving the slider on $R_3$ to the left the value of $E_c$ is increased, and this increases the time delay. At the same time the a-c voltage added in the grid circuit is decreased, also tending to increase the delay interval, and giving a rather wide useful range from possibly 0.5 to 12 sec. The values of $R_1$ and $C_1$ are also controlling factors in the delay obtained.

Pentodes may be used as well as triodes, and small thyratrons of the inert-gas-filled type also operate well in such circuits, allowing use of more rugged relays.

Alternating-current-operated relays are limited in time of response by the periodic nature of their supply. For fast action the use of d-c supply, with stable components, is dictated.

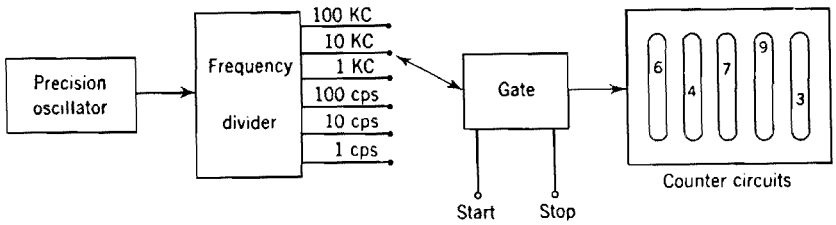



FIG. 16-6. Precision time-interval measurement.

**16-4. Time-interval Measurements.** Highly precise measurement of time intervals between successive events is possible by application of the decimal counter circuits of Chap. 9. As in Fig. 16-6, a fast-operating gate circuit has two terminals, one for the "start," or opening, operation, the other for the "stop," or closing, operation. When the gate is open, a standard frequency from an internal precision 100-kc oscillator and multivibrator frequency-dividing circuit is applied to the decimal counter circuits. The counters record the number of cycles of the standard fre-

quency that have passed the gate, or the time between the "start" and "stop" signals. The counters are so arranged as to operate one neon light for each number of counts 0 to 9 that each decade has received. By placing these neon lights to illuminate the appropriate digit in a translucent scale the count may be read off directly in decimal fashion.

A five-place display system as shown in the figure permits up to 100,000 counts to be indicated, and for the reading shown the duration of the event was 64,793 × 100 μsec, or 6.4793 sec. For the assembly shown the extreme range of interval measurement would be from 10 μsec to 100,000 sec. It is possible to add further counting decades to increase the number of significant figures indicated in the register.

The device may be used in inverse fashion to measure unknown frequency by counting cycles in unit time or to measure *events per unit time.*

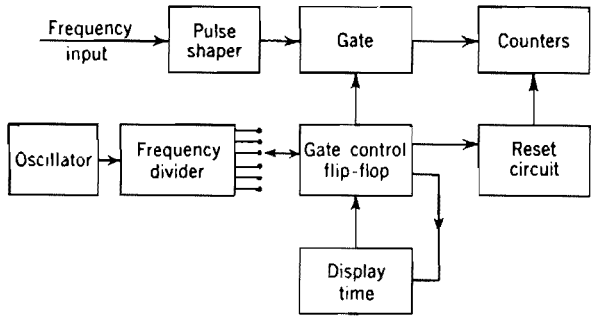



FIG. 16-7. Events per unit time measurement.

For this purpose the internal timing circuits from the 100-kc crystal oscillator are used to provide "start" and "stop" signals, and the unknown frequency, or event, signals are counted after pulse shaping. This usage is indicated in Fig. 16-7. When the gate is open, the input pulses are passed to the counters, and when the gate is closed, the counters display the number of pulses counted. The gate is opened for precisely determined time intervals from 0.001 to 10 sec, depending on the frequency, or rate of events, being counted. The frequency-dividing circuits control the gate through a flip-flop, which also causes the counters to be reset to zero shortly before each opening of the gate. For repeated measurement of the input the display time generator causes the circuits to go repeatedly through the cycle of measurement and display, thus following varying input rates by means of frequent repeated counts.

The precision crystal oscillator, operating at 100 kc, is usually temperature-controlled, and accuracy in measurement of either frequency or time interval is of the order of 0.001 per cent. Input signals of several volts peak are required.

**16-5. The Resistance-welding Process.** Resistance welding is a process for joining metals, usually in sheet or plate form. The welding heat is produced by passage of an electric current through the materials at the point to be welded. The resulting weld is of high quality, having been subjected to a forging or pressure action from the electrodes while the metal was at welding temperature. Spot welding implies a succession of spaced welded spots joining two metal sheets, whereas seam welding is produced by a succession of overlapping and timed spots, using water-cooled roller electrodes. Pulsation welding applies a number of successive current pulses to the same weld, used mainly to prevent electrode overheating in the welding of thick plate.

Electronic control of the duration and magnitude of the welding current has made precise and reproducible welding possible, and materials such as aluminum and stainless steel can be satisfactorily welded in production only because of electronic controls, giving precise welding times of a few cycles.

A complete control will perform a number of timed automatic functions, all initiated by the operator's closure of a foot switch. The control first applies air to a cylinder, which forces the electrodes against the work with the desired pressure. This operation is called the *squeeze time*. The ignitron contactor in the circuit of Fig. 16-8*a* is then operated by the control, the ignitrons being fired for the *weld time* of a few cycles to several seconds. The welding transformer steps up the line current to a value dependent on the material being welded, usually to many thousands of amperes. Passage of the heavy welding current produces heat where the $I^2R$ losses are greatest, at the junction of the materials as in Fig. 16-8*b*.

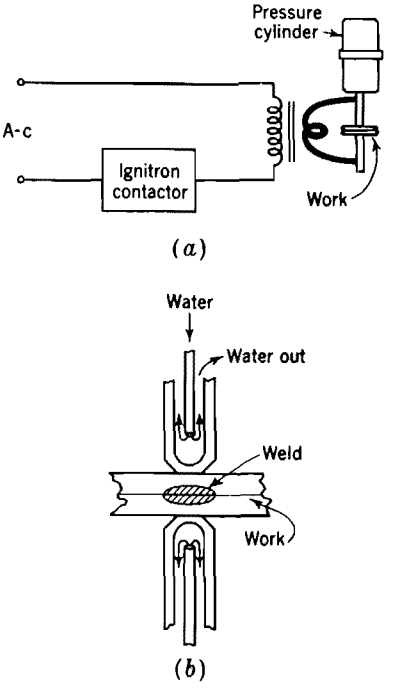



FIG. 16-8. (*a*) Resistance-welding circuit. (*b*) Water-cooled electrodes.

The water-cooled electrodes carry away some surface heat, but the time of the weld is usually made so short that the material brought to welding temperature is concentrated at the point of weld of the two workpieces, without appreciable softening of the surface of the work.

Thus short weld times are needed to reduce the time available for heat flow. The electrode pressure is maintained during a *hold time* after the welding interval, to allow the metal to solidify. The control then removes the air pressure, and the electrodes open. In many automatic machines a large number of electrodes are pressed onto the work and fired in succession, thus producing all welds on a piece without movement of the work.

The welding current and time required are a function of the work thickness and shape and of the electrical resistivity and thermal conductivity. High-resistivity materials such as the ferrous metals are easily welded, while aluminum and copper require more careful handling. Materials of low thermal conductivity weld more easily because the

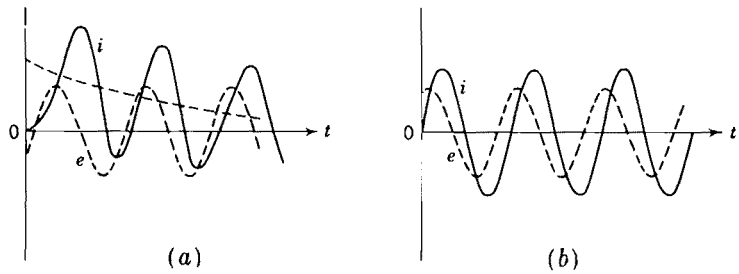



FIG. 16-9. Transients in a-c circuits. (*a*) Offset due to improper time of circuit closure. (*b*) No offset due to closing circuit at time of current zero.

heat cannot readily flow away from the weld volume. Electrode pressure is also important, too little producing an improperly worked weld, and too much leading to dimpling of the surface.

Precise control of the weld heat requires that the current initiating point be at the true current zero, or natural power-factor angle of the circuit, or that the control be *synchronous*, to prevent transient current offset as illustrated in Fig. 16-9. In *nonsynchronous* welding this offset leads to differing amounts of heat generated in the first several cycles and to inconsistent weld quality in very short welding times. Nonsynchronous operation is satisfactory, however, for weld times of a considerable fraction of a second or longer.

**16-6. Automatic Sequence Timer.** A simple nonsynchronous welder control may be assembled from four time-delay circuits of the type of Fig. 16-4*b*. A functional diagram is shown in Fig. 16-10, where $S$ is the operator's initiating foot switch. After locating the work, he presses $S$, turning on the air solenoid valve to the cylinder forcing the electrodes together, and starting the squeeze timer $T_1$ through line $A$. When $T_1$ operates at the end of the squeeze-time interval, it initiates the welding current by firing the ignitrons and starts the weld timer $T_2$ through line $B$.

When $T_2$ operates, it extinguishes the ignitron switch and starts the timing action of $T_3$, which is set for the "hold" period. At the end of the timing cycle of $T_3$ the air solenoid is released, the electrodes separate, and through lead $D$ all timers are reset for another operation. Timer $T_4$ also starts to operate, and after its period, which is set to permit the operator to shift the work to a new spot, $T_4$ will initiate a new complete timing sequence if the operator continues to press $S$.

The operator, by keeping $S$ closed, causes the welder to operate through as many complete cycles as he wishes, with his hands free to move the work in the "off" time between welds. Each time interval in the cycle may be independently adjusted by $R_3$ of Fig. 16-4*b* or a similar control on other timer circuits.

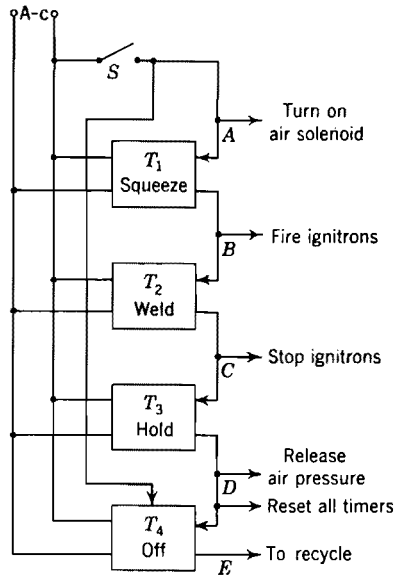



FIG. 16-10. Automatic sequence timer.

**16-7. Synchronous Weld Control.** Firing of the ignitrons at the natural current zero, or the power-factor angle, has been mentioned as desirable, to avoid the transient offset at current initiation and to prevent erratic welds. A synchronous control, of which Fig. 16-11 is one example,

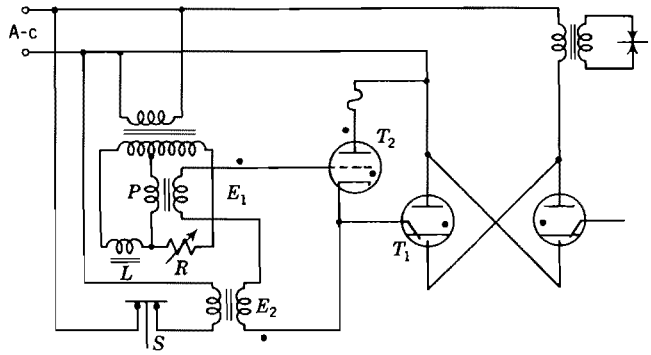



FIG. 16-11. Synchronous-control circuit.

permits adjustment of the circuit for such firing in synchronism with the current zero.

Essentially this circuit is that of the phase-shift-controlled thyratron, with a peaking transformer $P$ supplying a peaked grid voltage $E_1$ for the

thyratron $T_2$. Voltage $E_2$ is negative to the grid in the half cycle of positive anode voltage of $T_2$ and is of sufficient magnitude to override $E_1$ and hold thyratron $T_2$ off. Thus $E_2$ is simply a negative a-c bias which may be removed by opening switch $S$ (as might be done by the squeeze timer), leaving the peaks of $E_1$ to act alone in the circuit and fire the thyratron at an angle determined by the setting of $R$ in the phase-shifting bridge. Resistor $R$ may be set to fire the ignitron at the point on the a-c voltage wave required for elimination of the starting transient offset. This provides maximum consistent weld heat.

In other forms of the circuit voltage $E_2$ is replaced with a d-c bias, which is removed by action of the starting circuit.

**16-8. Follow-up, or Trailing-tube, Circuit.** The transient offset in the welding current due to improper firing angle can be eliminated by the

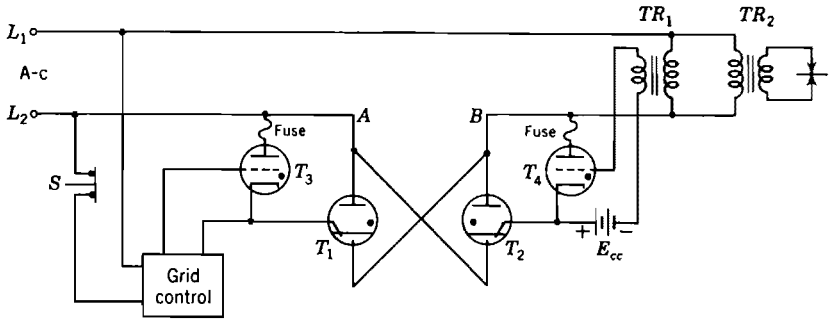



FIG. 16-12. Follow-up circuit with $T_1$ leading, $T_2$ trailing.

synchronous control described above. A somewhat similar erratic heating current is produced if the effect of residual magnetism in the welding transformer is not considered. The effect of such residual magnetism, due to the last half cycle of welder current, may be eliminated by forcing welder current to flow only in multiples of complete cycles and requiring the first half cycle of conduction always to flow through the transformer in the same direction.

In the circuit of Fig. 16-12 ignitron $T_1$ is fired by operation of $S$ and a synchronous grid control, so that the first half cycle of conduction always occurs when line $L_2$ is positive. During this half cycle the secondary voltage of $TR_1$ is connected to make the grid of thyratron $T_4$ negative. On the next half cycle this grid voltage becomes positive and overrides the small fixed bias $E_{cc}$. However, $T_4$ does not immediately fire because $T_1$ is still conducting owing to the lagging current taken by the welder, and the arc drop across points $A$-$B$ is too low to ionize $T_4$. At the instant of welder current zero the tube voltage rises abruptly, as shown in Fig. 16-13, and thyratron $T_4$ fires, followed by conduction in ignitron $T_2$.

In every cycle conduction will be initiated by ignitron $T_1$, and there will be a *follow-up*, or *trailing*, operation of ignitron $T_2$ to complete the weld operation with a half cycle of opposite polarity. Besides providing the necessary conditions for zero offset due to residual magnetism, the circuit also simplifies the control circuits, since control need be provided for thyratron $T_3$ only, $T_4$ always trailing along in operation.

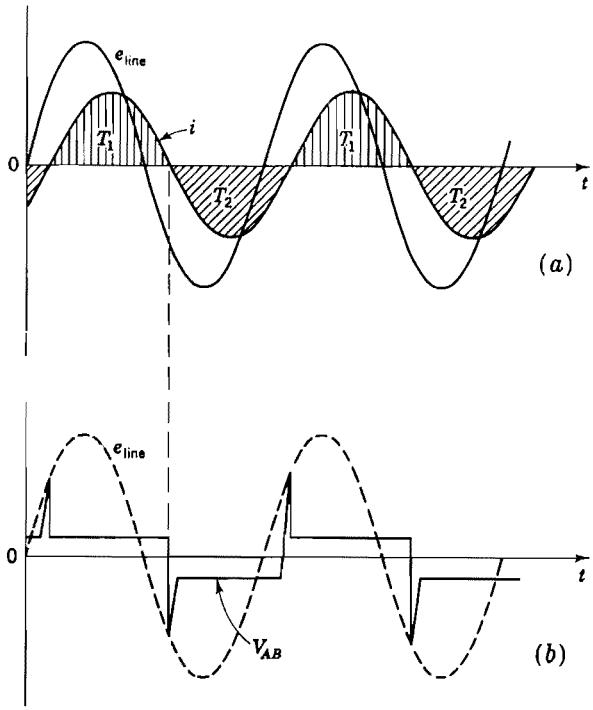



FIG. 16-13. (*a*) Line voltage and current. (*b*) Voltage across *A*-*B*, showing rise of voltage just before firing of ignitrons.

**16-9. Timing the Weld.** The circuits discussed above provide "on" and "off" switching of the welder current, but without the timing function. This may be readily incorporated to control firing tube $T_3$ by the addition of tubes $T_5$ and $T_6$ as in Fig. 16-14. This is a simplified version of an industrial control, which includes several other interesting circuit features.

Tube $T_5$ is known as the "start" tube, and $T_6$ as the "stop" tube, both being thyratrons. Starting is controlled by the phase-shift bridge of $L$ and $R_3$, the latter being adjusted to give synchronous firing with the current zero. The timing function is started when $T_5$ fires, and the

interval is measured by the increasing potential on $C_2$, charging through $R_2$; the potential ultimately fires the stop tube $T_6$ and ends the operation.

Contacts $S_1$ and $S_2$ start the operation, $S_1$ connecting the anodes of $T_5$ and $T_6$ to the d-c supply (actually a filtered rectifier), while $S_2$ removes the discharge path from timing capacitor $C_2$. Following $S_1$ closure, start tube $T_5$ is fired on the first positive peak from the peaking transformer $P$, the relative firing angle being set for synchronous operation.

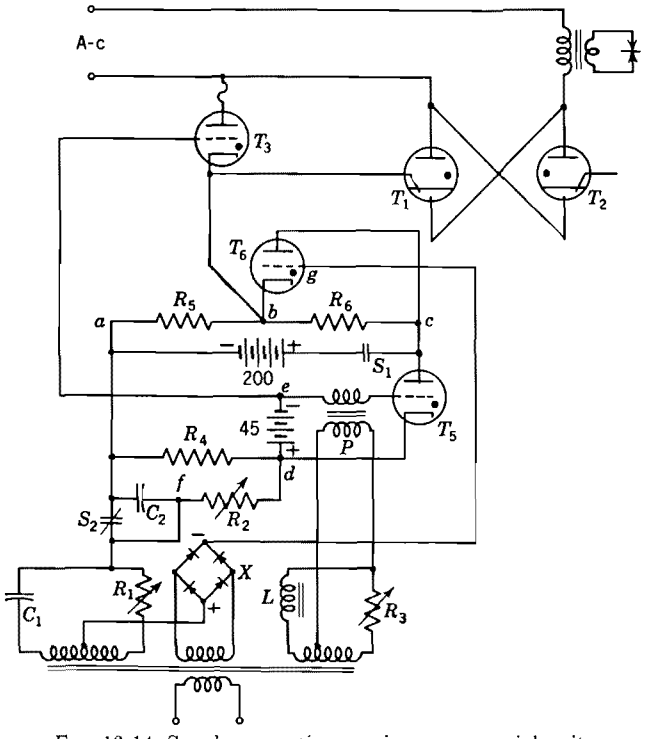



FIG. 16-14. Synchronous timer as in a commercial unit.

However, prior to $T_5$ firing and with resistors $R_5$ and $R_6$ equal, the potential of point $b$ and the cathode of $T_3$ is $+100$ volts with respect to point $a$. There is also no drop across $R_4$, and the potential of $d$ equals that of $a$. Therefore with $T_5$ not yet firing the grid of $T_3$ is at $-45$ volts to $a$, or the total bias on the grid of $T_3$ is $-145$ volts. When $T_5$ fires, it effectively connects point $d$ to $c$ (neglecting tube drop), and the potential between $b$ and $e$, or between cathode and grid of $T_3$, is $+100 - 45 = +55$ volts, and $T_3$ fires also. This initiates the firing action for the ignitron $T_1$, and if the trailing-tube operation is used for $T_2$, then both ignitrons will continue to fire as long as $T_3$ conducts.

With $T_5$ firing and point $d$ essentially connected to $c$ the full supply voltage appears across $R_2$ and $C_2$, but with capacitor $C_2$ initially uncharged point $f$ is at the potential of $a$, or $-100$ volts to the cathode of $T_3$ at $b$. Owing to the $T_5$ current drop across $R_4$ the capacitor $C_2$ begins to charge through $R_2$, and the potential of $f$ rises above the $-100$-volt level. The potential of $f$ supplies a rising d-c component to the grid of the stop tube $T_6$. An unusual combination of other voltages appears in series with this d-c component, and this circuit merits further discussion.

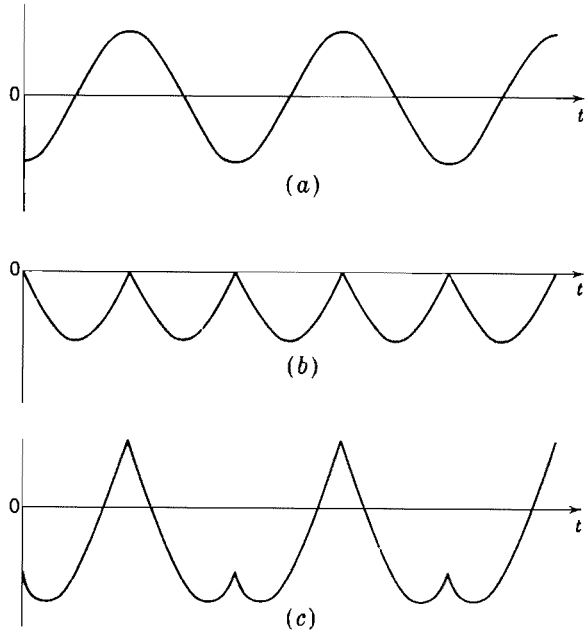



FIG. 16-15. Peaked voltage obtained by adding sine wave and rectified wave.

In addition to the varying d-c component across $R_2$ and $C_2$ there is a phase-shifted a-c component derived from the bridge composed of $C_1$ and $R_1$, and an unfiltered full-wave rectified component contributed by rectifier $X$. The result of addition of the a-c and rectified components is shown in Fig. 16-15$c$ as a wave with sharp positive peaks occurring at the positive maxima of the a-c wave, or once each cycle. This resultant wave, when added to the exponentially increasing voltage across $C_2$, gives a wave of the form of Fig. 16-16 for $e_{gb}$, the grid-to-cathode potential of $T_6$. At the completion of the desired number of welding cycles a voltage peak intersects the critical grid-potential locus of $T_6$, and stop tube $T_6$ fires. Flow of its current through $R_5$ makes the cathode of $T_3$ more positive and prevents $T_3$ from firing again, stopping the whole operation.

Reset of all tubes occurs when contact $S_1$ is opened at the end of the "hold" cycle by the hold timer. Additional timers for "squeeze," "hold," and "off" time must be added when the complete welding sequence is to be controlled. These may include the simpler timing circuits previously described.

The peculiar peaked waveform of Fig. 16-16 greatly improves the precision of timing in terms of cycles of conduction. In this respect it behaves in a manner similar to the synchronized multivibrator. Change of timing is achieved by variation of $R_2$. All electronic control of this nature has contributed greatly to the precision of welders on short welding intervals of a few cycles.

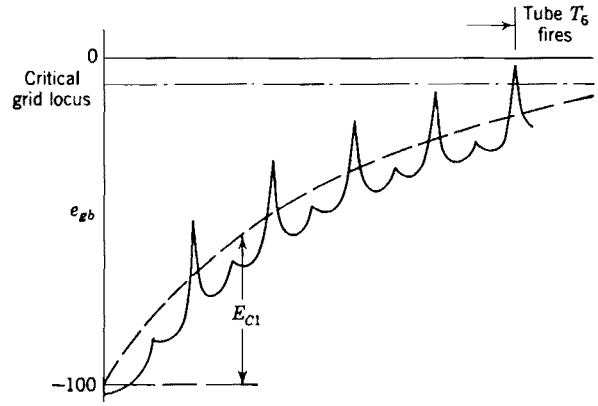



FIG. 16-16. Accurate weld timing by superimposed peaked wave.

**16-10. Heat Control.** A good weld depends not only on the duration but also on the magnitude of the welding current. While it is possible to control the weld current by transformer taps, this gives step control and cannot be changed while the welder is operating. A more desirable form of current or *heat control* is obtained by means of taps combined with phase-shift control on the ignitrons. This delays the firing of the ignitrons to angles later than the natural power-factor angle, and the current flow is then in alternating pulses, instead of sinusoidal alternating current.

Such stepless control may be achieved by conventional phase-shift methods, but in this case the leading-trailing circuit cannot be employed because it provides no means of delay of firing time for the trailing tube. By use of the stop-tube circuit of Fig. 16-16, combined with a phase-shift circuit controlling both ignitrons, it is possible to retain the advantages of even half-cycle conduction given by the trailing-tube circuit. Since the sharp positive-waveform peak occurs just once each cycle, the transformer polarities may be chosen so that this peak comes in the half

cycle in which ignitron $T_2$ is firing. The stop tube then fires and prevents $T_3$ from firing ignitron $T_1$ on the next half cycle. The circuit may be designed to always start operation with $T_1$ firing, and by use of stop-tube excitation the circuit will always stop with a half cycle of $T_2$ conduction, giving complete cycle conduction as with the trailing tube.

**16-11. Energy-storage Welding Systems.** Single-phase welding systems operating for such short periods of time place heavy demands on the power system, leading to excessive voltage drop and accompanying flicker of lights. Energy-storage systems, in which the average rate of

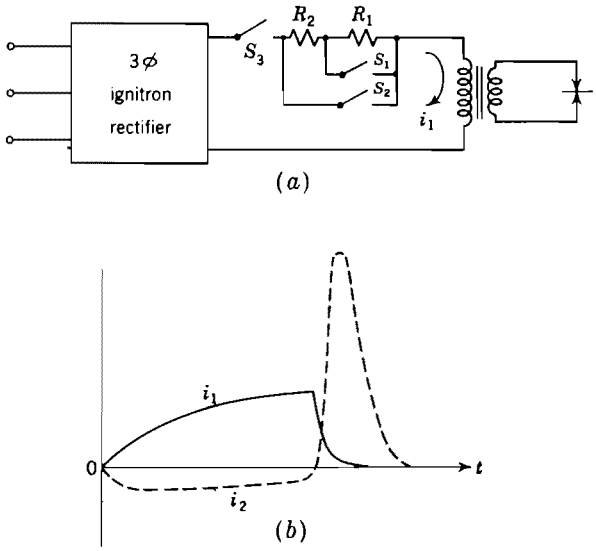



FIG. 16-17. Magnetic energy-storage welder.

energy input is low, have been devised, and these also serve to spread the demand equally over all three phases of the supply line.

The magnetic energy-storage system is illustrated in principle in Fig. 16-17. A three-phase ignitron rectifier supplies d-c current $i_1$ to the primary of the welding transformer upon closure of contactors $S_1$, $S_2$, and $S_3$. The primary current builds up exponentially because of the large primary inductance, and a small secondary current $i_2$ is induced which serves to preheat the weld. When $i_1$ reaches a value representative of sufficient energy storage in the core, as given by

$$W = \frac{Li^2}{2}$$

the contactors are opened in very rapid sequence $S_2$, $S_1$, and $S_3$. The rapid decay of magnetic flux induces a high secondary current through

the electrodes, as shown in Fig. 16-17*b*, and this surge performs the welding operation.

The amount of energy delivered to the weld can be controlled by the value of current $i_1$ at which the contactors open. The very brief high-current pulse is particularly valuable for welding aluminum, which requires a high current for a very short time, owing to its low resistivity and high thermal conductivity.

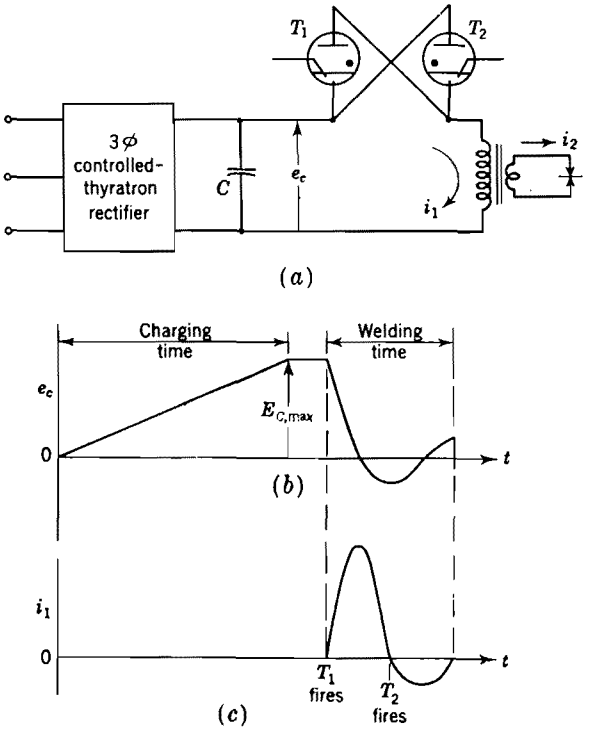



FIG. 16-18. Capacitor storage welder.

A comparable capacitor energy-storage system is shown in Fig. 16-18. The capacitor bank $C$ of 200 to 3000 $\mu$f is charged to a potential of several thousand volts by a controlled thyratron rectifier. At the instant the weld is desired, the energy stored in the capacitor is transferred to the welding transformer primary by firing ignitron $T_1$. The pulse of primary current produces a high secondary current of very short duration. The primary circuit may be oscillatory with a current of the form

$$i_1 = \frac{E_C}{\omega_o L} \epsilon^{-Rt/L} \sin \omega_o t$$

and ignitron $T_2$ may be fired, by trailing-tube action, from the reversed half cycle of current. The equivalent circuit of the welder during discharge is that of Fig. 16-19, where $R'$ is the primary resistance plus the reflected secondary-load value. The natural angular frequency of such a circuit is given by

$$\omega_o = \sqrt{\frac{1}{LC} - \frac{R^2}{4L^2}}$$

under oscillatory conditions.

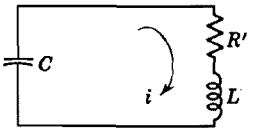



FIG. 16-19. Equivalent circuit of capacitor storage welder.

Accurate control of $E_{C,\max}$ by the controlled rectifier allows the amount of energy supplied to the weld to be adjusted in accordance with

$$W = \frac{CE^2}{2}$$

Additional controls are also required to reverse the transformer primary, or to bring its residual flux to zero between welds, so as to avoid transient offset and erratic welds due to the residual core flux. Equipment such as this is also of value in the welding of aluminum or other materials having low resistivity and high thermal conductivity.

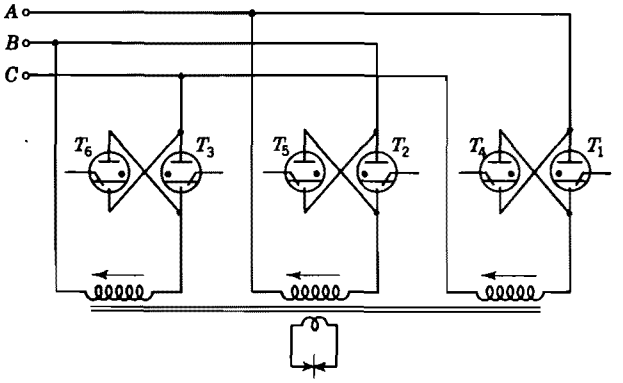



FIG. 16-20. Three-phase welder.

**16-12. Polyphase Welding Systems.** Another system that distributes the load to all three phases of a power system, although it still imposes short-time heavy demand, is the polyphase circuit of Fig. 16-20. Each phase winding on the transformer is controlled by an inverse parallel pair of ignitrons, but these are not controlled in the conventional manner. With phase rotation $A$-$B$-$C$ three tubes will be fired in order $T_1$, $T_2$, $T_3$, $T_1$, $T_2$, $T_3$ during the time of 2 cycles, giving two pulses of current in each winding in the direction shown by the arrows. During the next 2 cycles the firing sequence will be $T_4$, $T_5$, $T_6$, $T_4$, $T_5$, $T_6$, giving two

pulses of current in each winding in a direction opposite to the arrows. During the first tube sequence a secondary current will build up in one direction, and during the second tube sequence the secondary current will reverse. This cycle may be repeated as many times as is required to complete a weld.

As explained, the secondary frequency will be one-fourth that of the supply line and in actual practice may be reduced to about 5 cycles by alteration of the above firing schedule. The circuit gives equal phase loading, and the low secondary frequency is helpful in overcoming the reactance of the secondary circuit and workpiece.

## PROBLEMS

**16-1.** A relay of 8000 ohms resistance has a pickup current of 8.0 ma and drop-out current of 2.5 ma. Design a contact relay using a tube of $\mu = 8$, $r_p = 10{,}000$ ohms, $E_f = 6.3$ volts, $I_f = 0.3$ amp to operate completely from a 120-volt d-c source.

**16-2.** In the circuit of Fig. 16-21, with a 6J5 tube, the relay picks up at 6 ma and drops out at 3.5 ma. How long after $S$ is opened will the relay pick up?

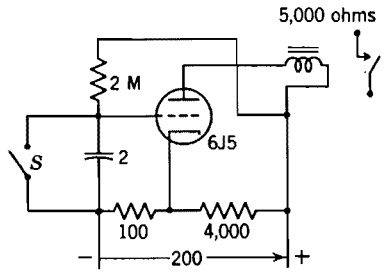



FIG. 16-21

**16-3.** In the circuit of Fig. 16-2*b* the tube is a 6J5 with $E_{bb} = 250$ volts and relay of 8000 ohms resistance and negligible inductance. If $E_{cc} = 50$ volts and the relay picks up at 5.0 ma and drops out at 2.7 ma, discuss the relay operation and timing on switching $S$ from 1 to 2 and back to 1 very slowly.

**16-4.** A welder takes 200 kva at 500 volts with 0.75 power factor for a particular job. At what angle should the thyratrons fire the ignitrons relative to line-voltage zero if maximum kva is to be consistently delivered to the weld?

**16-5.** A welding transformer has a primary resistance of 0.15 ohm and time constant of 40 sec. At what current value should the contactors open in a magnetic-storage welder, if supplied by 500 volt d-c, to make a weld requiring 7500 watt-sec of energy?

**16-6.** It has been found that a particular welding job requires 200 kva for 0.05 sec to produce a weld. The input power factor is 0.75. A capacitor welder with 2500-$\mu$f capacitor is to be used on the same job. To what voltage should the capacitor be charged?

**16-7.** For a capacitor storage welder, what value of reflected resistance $R'$, in terms of $L$ and $C$, will be needed if only a unidirectional current is to flow in a weld?

## REFERENCES

1. Wald, S.: Precision Interval Timer, *Electronics*, **21**: 88 (December, 1948).
2. Frase, S. H.: Electronic Timer, *Electronics*, **21**: 136 (April, 1948).
3. Chance, B.: Some Precision Circuit Techniques Used in Wave-form Generation and Time Measurement, *Rev. Sci. Instr.*, **17**: 396 (1946).
4. Chute, G. M.: "Electronics in Industry," McGraw-Hill Book Company, Inc., New York, 1946.
5. Chute, G. M.: "Electronic Control of Resistance Welding," McGraw-Hill Book Company, Inc., New York, 1943.
6. Bendz, W. I.: "Electronics for Industry," John Wiley & Sons, Inc., New York, 1947.
7. Bichsel, H. J., and E. T. Hughes: An Improved Electronic Control for Capacitor-discharge Resistance Welding, *Trans. AIEE*, **63**: 1150 (1944).
8. Engineers of the Westinghouse Electric Corp: "Industrial Electronics Reference Book," John Wiley & Sons, Inc., New York, 1948.

CHAPTER 17

# ELECTRONIC MOTOR CONTROL

Much industrial control equipment requires movement, either rotary or translatory, as its ultimate output. The electric motor, either a-c or d-c, is the usual means of providing such a mechanical output.

The a-c motor avoids the complication of separate field supply and for this reason is frequently employed in the smaller sizes, particularly where controlled speed is not a necessity. In the larger sizes the unstable speed characteristics and low efficiency of the a-c motor have caused the circuit designer to turn to the time-proven shunt-field d-c motor, in spite of the field-excitation requirements. The separately excited d-c motor has excellent variable-speed characteristics and is readily controllable at good efficiency with electronic circuits. Although a combination of tubes and motor is expensive, it is possible to incorporate other features, such as control of speed or torque, current limiting for starting or motor protection, and accurate and simple speed variation, which justify the added cost. Also the electronic control makes possible the operation of d-c motors from the a-c lines, avoiding the need for separate d-c power generation.

This chapter will consider some of the basic electronic principles utilized in electronic control of motors and will briefly discuss their synthesis in commercial controls.

**17-1. D-C Motor Relations.** The expression for the voltages present in the separately excited d-c motor may be written from the equivalent circuit of Fig. 17-1*a* as

$$E_{dc} - I_{dc}R_a - E_a = 0 \tag{17-1}$$

$$E_a = K_1\phi_f N \tag{17-2}$$

where $E_a$ = counter emf of motor
$\phi_f$ = effective field flux
$N$ = speed, rpm
$R_a$ = armature resistance
$K_1$ = a proportionality constant

The inductance $L$ of the armature has no effect under steady-current conditions.

Rewriting Eqs. (17-1) and (17-2) gives

$$N = \frac{E_{dc} - I_{dc}R_a}{K_1\phi_f} \tag{17-3}$$

With increasing $I_{dc}$ and constant field $\phi_f$ the speed drops off with load, owing to the increasing $I_{dc}R_a$ drop of the armature, and this is illustrated by curve (1) of Fig. 17-1*b*. If means were introduced to cancel the effect of the $I_{dc}R_a$ drop, and if the effect of armature reaction on field flux were made negligible, then the ideal speed-current curve of (2) might be obtained. These are duties which are assigned to the electronic circuits to be discussed.

The speed $N$ at which the motor operates with rated applied voltage, rated armature current, and with full field current may be called the *rated*, or *base speed*. Under these conditions the motor torque $T$ and horsepower output, related as

$$\text{hp} = \frac{2\pi NT}{33{,}000} \tag{17-4}$$

would also be at rated values.

Since the field flux cannot be appreciably increased owing to saturation of the iron field structure, then it is apparent from Eq. (17-3) that the only means available for reducing the motor speed below the rated value is by reduction of the applied voltage $E_{dc}$. Equation (17-3) indicates a linear relation between $E_{dc}$ and speed, and this is shown in Fig. 17-1*c*, below rated speed. Thus a variable-speed control circuit must provide a means of reducing $E_{dc}$, preferably in a stepless and lossless manner.

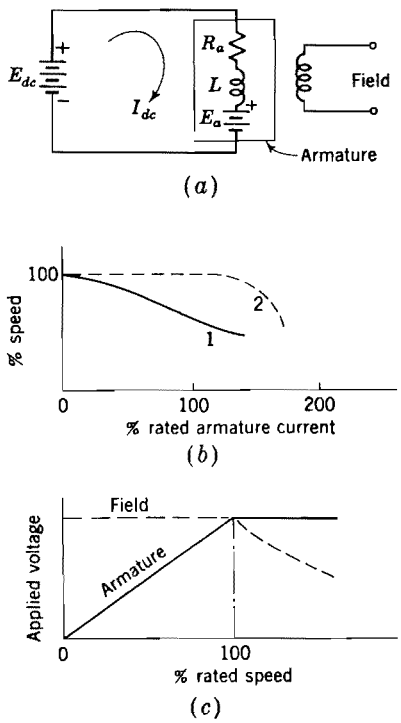



FIG. 17-1. (*a*) Equivalent circuit for separately excited d-c motor. (*b*) Speed-current relations. (*c*) Speed-voltage-field relation.

Since armature heating is a controlling factor in motor operation, then the rated armature-current value is a top limit. To raise the speed above rated value by increasing the applied voltage $E_{dc}$ with constant field flux will call for larger than rated currents and for continuous operation; this cannot be permitted. However, Eq. (17-3) indicates that the result may be achieved within the voltage limit by reducing the field

flux, and this is another function which may be assigned to the electronic control circuits.

It may be noted that

$$T = K_2\phi_f I_{dc} \tag{17-5}$$

is a basic relation between torque $T$, field flux, and armature current. If the flux is reduced, the torque load must likewise be reduced if the armature current is not to exceed the limiting rated value.

It appears that motor operation may be divided into two regions, that below the rated speed and that above. In the region below rated speed, with normal field flux and with a maximum value set on $I_{dc}$, then, Eq. (17-5) indicates a constant maximum torque value. Above rated speed, where $\phi_f$ is reduced but rated $I_{dc}$ remains constant, Eq. (17-5) shows that

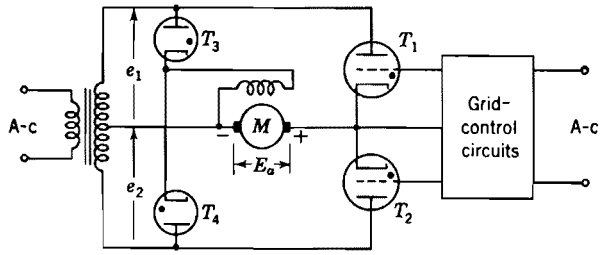



FIG. 17-2. Thyratron-controlled rectifier for d-c motor supply.

the permissible torque falls as speed increases, and thus Eq. (17-4) indicates that this is a region of constant permissible horsepower output. The terms *constant torque* and *constant horsepower* are occasionally used as descriptive of these two regions, not because operation is at these constant limits, but because maximum operation is fixed by them.

**17-2. Thyratron Supply for the D-C Motor.** While polyphase rectifiers and ignitrons are used for very large motors, the greatest number of rectifier-supplied motors are from fractions up to several horsepower in size, taking power from single-phase sources, as in Fig. 17-2. Thyratrons $T_1$ and $T_2$ furnish controlled direct current to the motor armature, and gas diodes $T_3$ and $T_4$ supply constant direct current to the separate field circuit.

Figures 17-3 and 17-4 show the waveforms of voltage and current supplied to the armature with a firing angle $\theta_1$, delayed so that conduction is discontinuous, even though the armature inductance is present. Just before the instant of firing at $\theta_1$ the current is zero, and the armature voltage is equal to the counter emf $E_a$. At $\theta_1$ thyratron $T_1$ fires, and the armature voltage jumps to $e_1$, neglecting tube drop. The armature current then builds up slowly, owing to the armature inductance, and carries over to angle $\theta_2$, when conduction ceases and the voltage jumps to $E_a$.

From the circuit of Fig. 17-2 it is apparent that the net voltage available to build up current in the circuit is only $e_1 - E_a$ or $E_m \sin \omega t - E_a$. Thus only during the interval from $\theta_1$ to $\theta_x$ at which $E_m \sin \omega t = E_a$ can the applied potential produce a positive current in the circuit. The remainder of the conduction period from $\theta_x$ to $\theta_2$ is due to the $L\,di/dt$ potential arising from the decreasing current in the armature inductance.

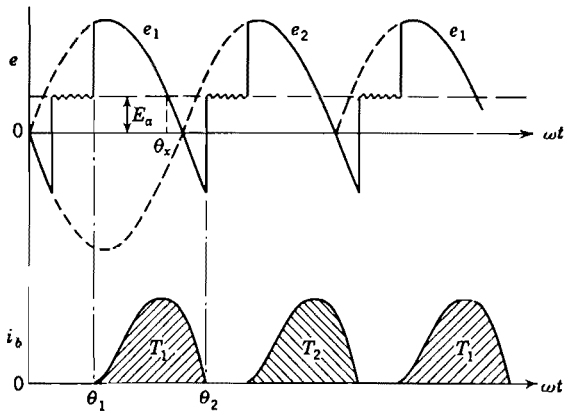



FIG. 17-3. Waveforms in a thyratron-controlled motor.

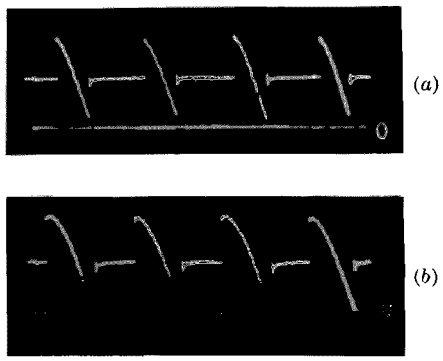



FIG. 17-4. Armature voltage from a controlled rectifier. (*a*) $\theta_1 = 100°$ and light load. (*b*) $\theta_1 = 85°$ with heavier load.

If $\theta_1$ is advanced, the average voltage will increase and the speed will rise. If the torque load on the motor rises, the armature current must increase by Eq. (17-5), and this can happen only as a result of a reduced $E_a$ value since the firing angle $\theta_1$ is fixed. Since current can flow for only a short interval, the emf $E_a$ must drop further than it would if the motor were supplied continuously from a d-c line. Since $E_a$ is proportional to speed, the result of the increase in load torque is an adverse

effect on speed regulation, making the speed-current relation of (1) (Fig. 17-1*b*) droop much more than if the motor were supplied from a d-c line. This accentuates the need for speed control as an effective part of the over-all thyratron motor-control system.

During the operation of a thyratron-controlled motor the current in the load may change from discontinuous to continuous flow, depending on the speed, firing angle, and armature inductance.* This situation is parallel to the one discussed in Sec. 15-13 for inductive loads on controlled rectifiers. Since the analysis differs only by the inclusion of the effect of $E_a$, the analysis for the tube and armature current will be sketched only briefly here, first for the case of discontinuous conduction.

For the motor of Fig. 17-1*a* and the current waveform of Fig. 17-3 the differential equation during conduction is

$$L\frac{di_b}{dt} + R_a i_b = E_m \sin \omega t - E_a \tag{17-6}$$

which has a solution

$$i_b = A\epsilon^{-R_a t/L} + \frac{E_m}{\sqrt{R_a^2 + \omega^2 L^2}} \sin (\omega t - \phi) - \frac{E_a}{R_a} \tag{17-7}$$

where

$$\phi = \tan^{-1} \frac{\omega L}{R_a} \tag{17-8}$$

A new coefficient $a = E_a/E_m$ may be defined. Since $E_a = K_1\phi_f N$, then $a$ is directly proportional to speed $N$ and may be called the *speed coefficient*. The constant $A$ may then be evaluated by noting that, at $\omega t = 0$, $i_b = 0$, so that the result for $i_b$ is

$$i_b = \frac{E_m}{R_a}\left\{\frac{R_a}{\sqrt{R_a^2 + \omega^2 L^2}} \sin (\omega t - \phi) - a + \left[a - \frac{R_a}{\sqrt{R_a^2 + \omega^2 L^2}} \sin (\theta_1 - \phi)\right]\epsilon^{-(R_a/\omega L)(\omega t - \theta_1)}\right\} \tag{17-9}$$

In view of the definition of Eq. (17-8)

$$\cos \phi = \frac{R_a}{\sqrt{R_a^2 + \omega^2 L^2}} \tag{17-10}$$

so that

$$i_b = \frac{E_m}{R_a}\{\cos \phi \sin (\omega t - \phi) - a + [a - \cos \phi \sin (\theta_1 - \phi)]\epsilon^{-(R_a/\omega L)(\omega t - \theta_1)}\} \tag{17-11}$$

which holds over the time range $\theta_1 \leq \omega t \leq \theta_2$ provided that $\theta_1 > \theta_2 - \pi$, or that conduction is discontinuous.

* As an example, for a particular 230-volt 1-hp d-c motor the armature resistance is 6.6 ohms, armature inductance 0.05 h.

It may be seen from the waveform that $i_b = 0$ at $\omega t = \theta_2$ also, for discontinuous conduction, so that

$$\epsilon^{R_a\theta_2/\omega L}\frac{\cos\phi\sin(\theta_2-\phi)-a}{\cos\phi\sin(\theta_1-\phi)-a}=\epsilon^{R_a\theta_1/\omega L} \tag{17-12}$$

This is a transcendental relation which may be solved graphically in terms of $\theta_1$, $a$, and $\omega L/R_a$ as parameters, the result appearing in Fig.

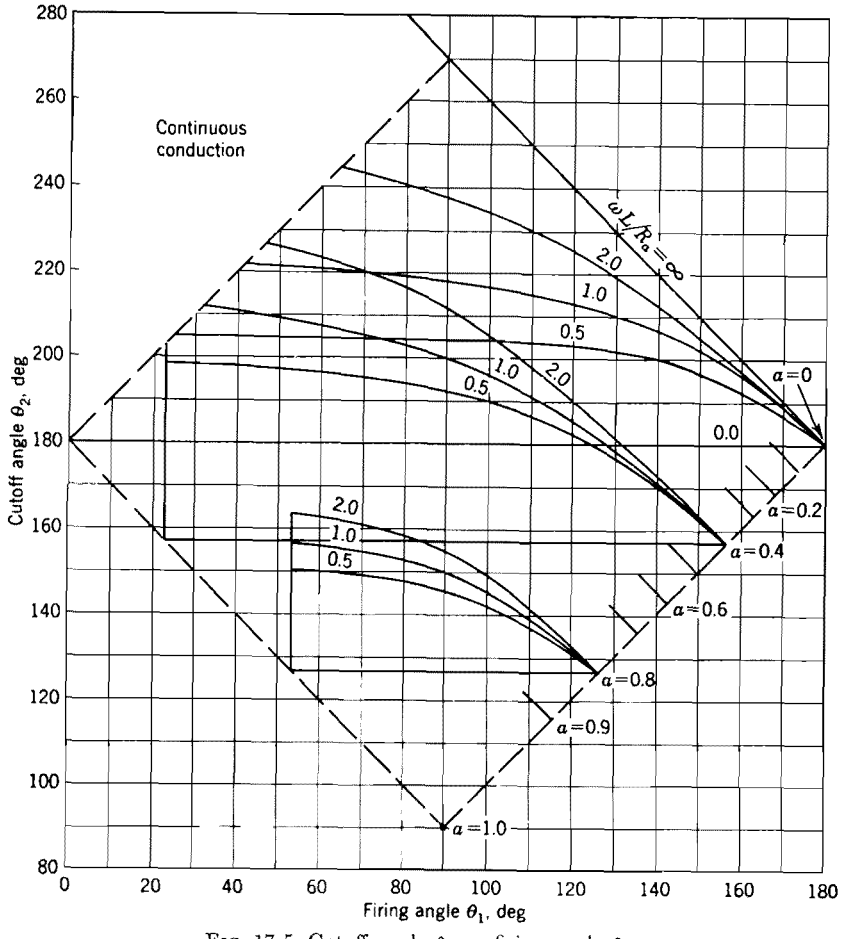



FIG. 17-5. Cutoff angle $\theta_2$ vs. firing angle $\theta_1$.

17-5, which is analogous to Fig. 15-27. The increased complexity of Fig. 17-5 is due to the various curve families for parameter $a$, that for $a = 0$ being taken directly from Fig. 15-27.

It can be reasoned that the tube cannot be fired at angles later or earlier than those at which $E_m \sin \omega t = E_a$, since this would make the

anode voltage negative. Thus the earliest firing angle is that at which

$$\theta_{1,\min} = \sin^{-1} a$$

This angle sets a limit for all the $\omega L/R_a$ curves of the family for a given $a$, as shown by the vertical limits on the $a = 0.8$ and $a = 0.4$ families.

The latest time for firing will occur for

$$\theta_{1,\max} = \pi - \sin^{-1} a$$

and this determines the locus shown as the dashed line from $a = 1.0$ at $\theta_1 = 90°$, sloping upward to the right with $a = 0$ at $\theta_1 = 180°$.

It can also be reasoned that for a conduction angle greater than 180°, or $\theta_2 - \theta_1 \geq \pi$, the conduction would be continuous, and the above analysis based on distinct current pulses does not hold. A limit line for

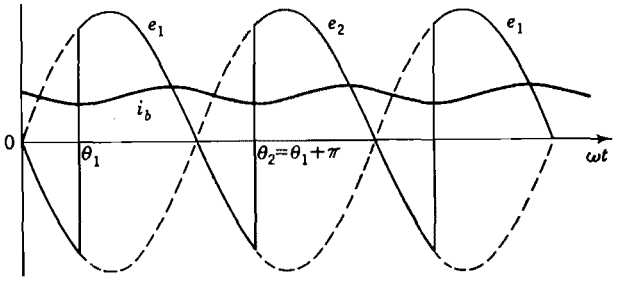



FIG. 17-6. Current and voltage under continuous conduction.

$\theta_2 - \theta_1 = \pi$ is shown dashed on the figure, starting at $\theta_1 = 0$, $\theta_2 = 180°$ and sloping up to the right, thus setting off the area of continuous conduction as shown.

The current under the continuous-conduction situation may also be studied, starting with Eq. (17-6) as before, but noting that the conditions for evaluation of the constant $A$ are different. For continuous conduction, as in Fig. 17-6, then $\theta_2 = \theta_1 + \pi$. Writing Eq. (17-7) for $\omega t = \theta_1$ and for $\omega t = \theta_2 = \theta_1 + \pi$, and equating, allows $A$ to be evaluated as

$$A = \frac{(2E_m/\sqrt{R_a^2 + \omega^2 L^2}) \sin (\theta_1 - \phi)}{(\epsilon^{-R_a\pi/\omega L} - 1)\epsilon^{R_a\theta_1/\omega L}} \tag{17-13}$$

and then $i_b$ for continuous conduction may ultimately be written as

$$i_b = \frac{E_m}{R_a}\left[\cos\phi \sin(\omega t - \phi) - a - \frac{2\cos\phi \sin(\theta_1 - \phi)}{1 - \epsilon^{-R_a\pi/\omega L}} \epsilon^{-(R_a/\omega L)(\omega t - \theta_1)}\right] \tag{17-14}$$

This will predict the current waveform of Fig. 17-6. Continuous conduction will occur whenever the theoretical value of $\theta_2$ is greater than $\theta_1 + \pi$, as determined from Fig. 17-5. The actual value of $\theta_2$ for continuous conduction may be taken as equal to $\theta_1 + \pi$.

It may be noted that all these possible current waveforms, as shown in Figs. 17-3 and 17-6, have appreciable harmonic content, and therefore $I_{rms} > I_{dc}$. This means that for a given average torque, which will be proportional to $I_{dc}$, a motor will run hotter when supplied by a rectifier than when operated from a d-c line. Motors in controlled rectifier service are usually derated therefore, a 1-hp frame being given a ¾-hp rating, as an example.

**17-3. Speed-Torque Relations.** Equation (17-1) gives the circuit relations for a separately excited motor supplied by a d-c line. It may be normalized with respect to applied voltage $E_{dc}$ as

$$1 - \frac{I_{dc}R_a}{E_{dc}} - \frac{E_a}{E_{dc}} = 0 \tag{17-15}$$

The second term may be rearranged as

$$\frac{I_{dc}}{E_{dc}/R_a}$$

The denominator $E_{dc}/R_a$ is the current which would flow in the armature at standstill. By Eq. (17-5) the torque is proportional to current, assuming no armature reaction and constant field $\phi_f$. It is convenient to define a *torque coefficient*

$$T_{dc} = \frac{T}{T_o} = \frac{I_{dc}}{E_{dc}/R_a} \tag{17-16}$$

that expresses the ratio between the actual torque and the standstill torque of the motor, when connected to a d-c line.

The counter emf $E_a$ is proportional to speed in accordance with Eq. (17-2), again assuming no armature reaction. For a d-c supply a *speed coefficient* may be defined as

$$a_{dc} = \frac{E_a}{E_{dc}} \tag{17-17}$$

as the ratio of the actual speed to a theoretical maximum speed at which $E_{dc}$ is the counter emf. This term is obviously related to the previously employed $a = E_a/E_m$ for the rectifier-supplied motor.

From Eq. (17-15)

$$\frac{I_{dc}}{E_{dc}/R_a} + \frac{E_a}{E_{dc}} = 1 \tag{17-18}$$

$$T_{dc} + a_{dc} = 1 \tag{17-19}$$

A plot of $a_{dc}$ vs. $T_{dc}$ appears in Fig. 17-7, as a theoretical speed-torque curve for a motor supplied from a d-c line.

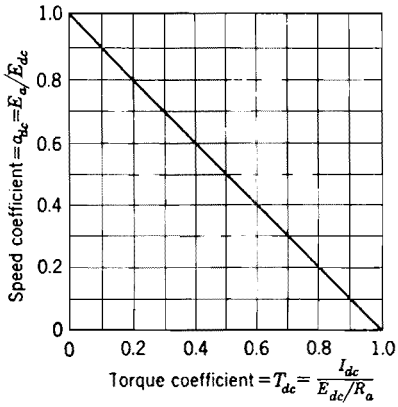



FIG. 17-7. Theoretical speed-torque characteristic for a motor supplied by a d-c line. (Armature reaction neglected.)

It is desirable to obtain a similar characteristic for a motor driven from rectifiers, in order to compare the system with normal d-c line operation. Such a curve may be obtained by finding $I_{dc}$, even though in discontinuous operation the motor torque will occur in pulses. The motor and load inertia will smooth these out, and an average torque proportional to $I_{dc}$ will be obtained. Referring to the equivalent motor of Fig. 17-1*a*, a d-c current cannot produce a voltage across an inductance; so

$$I_{dc} = \frac{1}{\pi} \int_{\theta_1}^{\theta_2} \frac{E_m \sin \omega t - E_a}{R_a} \, d\omega t \tag{17-20}$$

for the rectifier-supplied motor of Fig. 17-2, allowing for the full-wave rectification and neglecting tube drop. The appropriate waveform is that of Fig. 17-3 with discontinuous conduction.

After use of $a = E_a/E_m$ and integration

$$I_{dc} = \frac{E_m}{\pi R_a} [\cos \theta_1 - \cos \theta_2 - a(\theta_2 - \theta_1)] \tag{17-21}$$

This expression may be rearranged as

$$\frac{I_{dc}}{2E_m/\pi R_a} = \frac{1}{2} [\cos \theta_1 - \cos \theta_2 - a(\theta_2 - \theta_1)] \tag{17-22}$$

The ratio $2E_m/\pi R_a$ is the current from the full-wave rectifier into the armature resistance at standstill. Thus Eq. (17-22) may be interpreted as the ratio of the motor torque $T$ under operating conditions to the torque at standstill $T_o$. Hence for rectifier supply the torque coefficient of Eq. (17-16) may be redefined as

$$T_r = \frac{T}{T_o} = \frac{I_{dc}}{2E_m/\pi R_a} \tag{17-23}$$

Then

$$T_r = \frac{I_{dc}}{2E_m/\pi R_a} = \frac{1}{2} [\cos \theta_1 - \cos \theta_2 - a(\theta_2 - \theta_1)] \tag{17-24}$$

Since $a$ is the speed coefficient for the rectifier-supplied motor, the above expression is a dimensionless relation between torque and speed for discontinuous conduction of the rectifiers.

For continuous conduction the average current will be given by Eq. (17-20) with $\theta_2 = \theta_1 + \pi$. This results in

$$I_{dc} = \frac{E_m}{\pi R_a} (2 \cos \theta_1 - a\pi) \tag{17-25}$$

which is just Eq. (17-21) with $\theta_2 = \theta_1 + \pi$. Then the torque coefficient for continuous conduction is

$$T_r = \frac{I_{dc}}{2E_m/\pi R_a} = \cos \theta_1 - \frac{a\pi}{2} \tag{17-26}$$

With data taken from Fig. 17-5 and its source equations it is possible to plot speed-torque curves for a thyratron-controlled motor from Eqs.

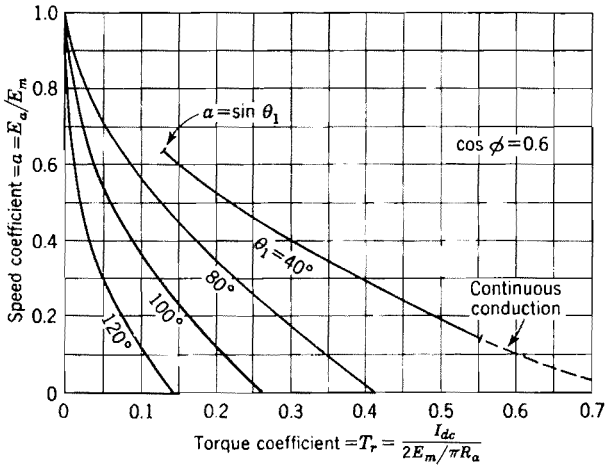



FIG. 17-8. Theoretical speed-torque curves for thyratron-supplied motor; cos $\phi$ = 0.6.

(17-24) and (17-26). These are presented in Fig. 17-8; in Fig. 17-9 is an experimentally determined curve to confirm the form of the theoretical curves of Fig. 17-8.

Comparison of the curves of Fig. 17-8 with that of Fig. 17-7 for a motor operated on a d-c line shows that the addition of thyratron supply causes the speed to fall off rather rapidly with torque. This indicates the need for a control circuit to maintain a desired operating speed.

It should be remembered that the rectifiers cannot start conducting unless $E_m \sin \omega t - E_a$ is positive in Eq. (17-20). This can be true only for $a \leq \sin \theta_1$, and this sets a limit on the curves for $\theta_1 = 80°$ and $40°$.

For the latter curve, since sin 40° = 0.64, the tubes cannot fire for values of $a > 0.64$, and the curve for $\theta_1 = 40°$ ends at that value.

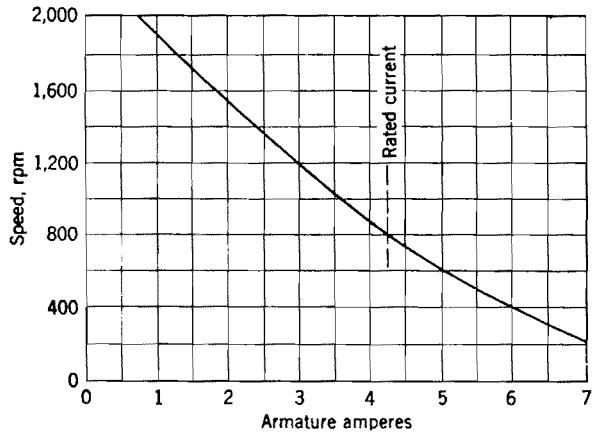



FIG. 17-9. Actual speed-torque curve for $\theta_1$ = constant, with thyratron-controlled 1-hp 1750-rpm 230-volt motor.

The average voltage required across the armature is

$$E_{dc} = E_a + I_{dc}R_a$$

and, dividing by $E_m$,

$$\frac{E_{dc}}{E_m} = \frac{E_a}{E_m} + \frac{I_{dc}R_a}{E_m} = a + \frac{2}{\pi} T_r \tag{17-27}$$

The armature voltage required to drive a thyratron-controlled motor is then expressed in terms of a dimensionless voltage coefficient

$$v_a = \frac{E_{dc}}{E_m} = a + \frac{2}{\pi} T_r \tag{17-28}$$

and a curve of $v_a$ against torque coefficient $T_r$ is plotted in Fig. 17-10. This allows estimation of the voltage required to achieve a desired output torque. Figure 17-11 is an experimentally determined curve for a thyratron-controlled motor, to confirm Fig. 17-10.

**17-4. Speed Regulation by Armature-voltage Control.** In the two most commonly employed forms of thyratron motor control the actual grid control of the main rectifier thyratrons is performed in the one case by the bias-phase method of control and in the other case by a saturable reactor and phase-shift bridge. Since these methods have already been discussed, emphasis will here be given to measuring and control circuits which are used to provide the voltages or currents needed by the above control methods.

If the counter emf be regulated, then the motor speed will be con-

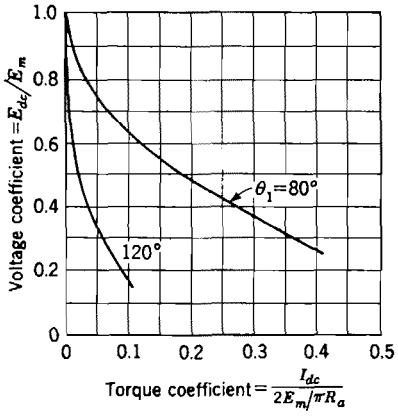



FIG. 17-10. Theoretical voltage-torque characteristic for thyratron-controlled motor; cos $\phi$ = 0.6.

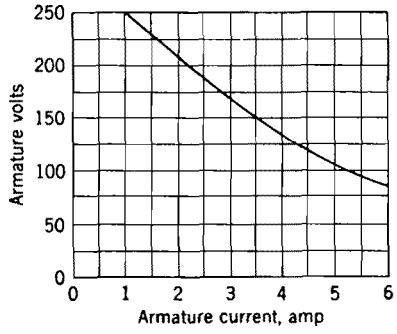



FIG. 17-11. Experimental voltage-torque curve for $\theta_1$ = constant, with 1-hp 1750-rpm 230-volt motor.

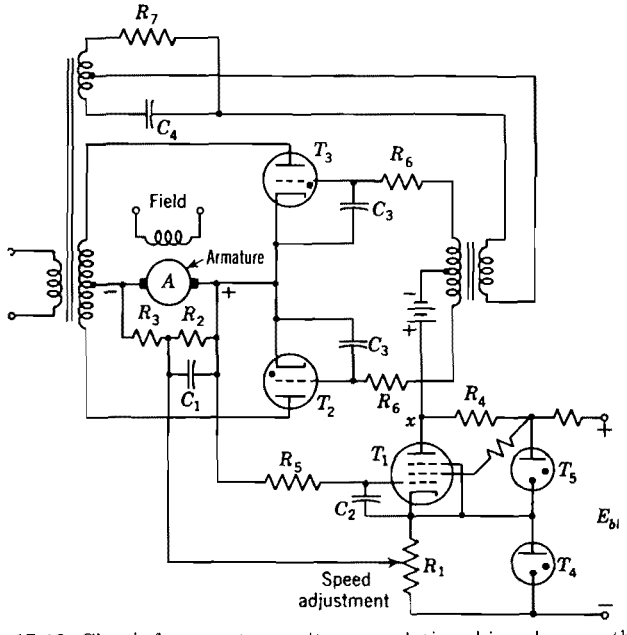



FIG. 17-12. Circuit for armature-voltage regulation; bias-phase method.

trolled, except for the effect of armature reaction. The armature voltage differs from the counter emf only by the amount of the armature resistance drop, which is usually small, so that controlling the armature voltage will give a reasonable speed control. This is the basic principle of the circuit of Fig. 17-12.

A fraction of the armature voltage determined by the ratio of

$$\frac{R_2}{R_3 + R_2}$$

is used to actuate the control. The $C_1$, $R_5$, $C_2$ filter removes much of the ripple in the armature voltage. The fraction of the armature voltage, representative of speed, is opposed by a d-c reference voltage developed across the upper part of $R_1$, supplied by the regulator $T_4$. The desired speed is represented by the voltage picked off from $R_1$, and any difference in this desired voltage and the actual armature-voltage

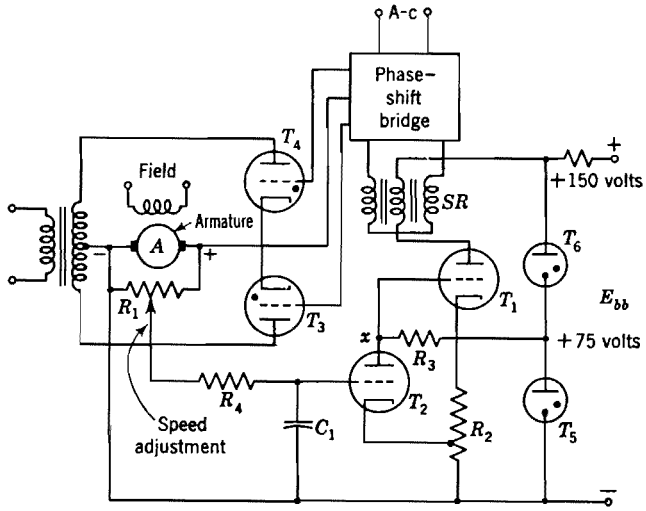



FIG. 17-13. Armature-voltage regulation for speed control; saturable-reactor method.

fraction is applied to the grid of $T_1$. If the actual speed is low, then the drop across $R_1$ exceeds that across $R_2$, makes the grid of $T_1$ negative and point $X$ more positive. This voltage added to the phase-shifted a-c voltage from the $R_7$, $C_4$ bridge advances the firing angle of the thyratrons, supplying a larger armature current and increasing the motor speed. Equilibrium occurs when the voltage across $R_2$ equals the selected voltage across $R_1$, and thus $R_1$ becomes a speed adjustment.

In the circuit of Fig. 17-13 the speed voltage is selected across $R_1$ and balanced against a fixed voltage from $R_2$. Any difference voltage is amplified by the d-c amplifier consisting of $T_1$ and $T_2$ to control the saturable reactor $SR$. If the speed is below standard, the difference voltage makes the grid of $T_2$ negative. This makes point $X$ and the grid of $T_1$ positive, increasing the plate current of $T_1$ and saturating the reactor $SR$. The firing angle of the thyratrons is advanced, increasing

armature current and raising the motor speed, until the speed-indicating voltage across $R_1$ equals the set voltage across $R_2$. High motor speeds will result from moving the tap on $R_1$ to the left, low speeds from moving the tap to the right.

**17-5. $IR$-drop Compensation.** Because the armature voltage is the sum of $E_a$ and $I_{dc}R_a$, the use of armature voltage to represent speed is not entirely accurate. A subtraction of $I_{dc}R_a$ from the armature voltage can be carried out electronically, by inserting into the armature-voltage-regulator circuit a bucking signal proportional to armature current. This will make control of speed proportional to the value of $E_a$ only.

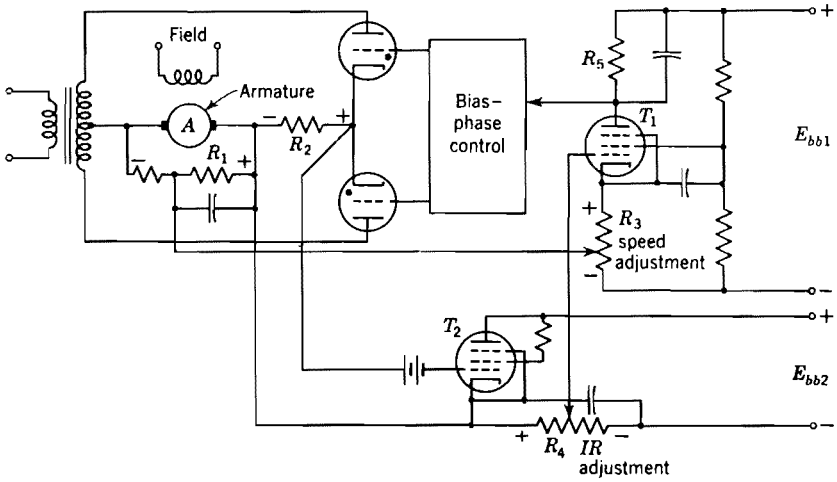



FIG. 17-14. Simple $IR$-drop compensation.

A simple means of obtaining a voltage proportional to $I_{dc}R_a$ is to add a very small resistor $R_2$ in series with the armature as in Fig. 17-14. This method requires that the voltage across $R_2$ be subtracted from that across $R_1$, in spite of the fact that their circuit relation is additive. Because of the common terminal, and because of the smallness of $R_2$ if its drop is not to be objectionable, tube $T_2$ is added as an amplifier and to invert the $IR$-drop voltage, before adding it into the armature-voltage-measuring circuit. The input circuit to $T_1$, the control amplifier, includes a rise in voltage across $R_4$ proportional to $I_{dc}R_a$, a drop across $R_1$ proportional to $E_a + I_{dc}R_a$, and a rise across the reference potentiometer $R_3$ to cathode. This results in a subtraction of $I_{dc}R_a$ from the armature voltage, and the net voltage is compared with the speed reference on $R_3$.

A more flexible and useful form of circuit is that of Fig. 17-15, in which a current transformer $CT$ is used to provide a secondary voltage proportional to motor current. The use of two primaries cancels the d-c satu-

rating flux and also sums the currents of the two thyratrons. Rectification of the secondary voltage gives an output on $R_5$ proportional to the motor current, with $C_3$ of aid in filtering this output. A portion of the output is then inserted in the measuring loop of either Fig. 17-12 or 17-13, to alter the speed-indicating voltage.

As applied in Fig. 17-15, the input to $T_2$ starts with a fall of potential through $R_1$, proportional to $E_a + I_{dc}R_a$, goes through a fixed rise of potential in $R_2$, and another rise in $R_5$ proportional to $I_{dc}R_a$, and thence to cathode. The net signal voltage is proportional to $E_a$, and this is

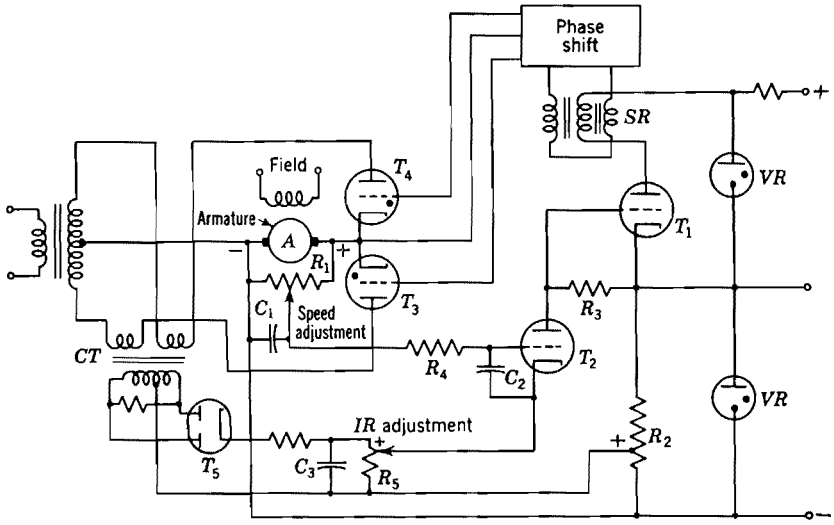



FIG. 17-15. *IR*-drop compensation signal from a current transformer.

balanced against the fixed reference voltage on $R_2$. The operation of $T_1$ is as previously discussed.

An interesting circuit feature is the use of a fixed reference voltage on $R_2$. The adjustment on $R_1$ for counter emf results in a constant $E_a + I_{dc}R_a$ voltage input to the control for all speeds, and since the current range and magnitude of $I_{dc}R_a$ will be the same for any speed, then the reference, $E_a$, and the $IR$-drop compensation are all maintained of appropriate magnitude, regardless of actual motor speed. If an adjustable reference were compared directly with $E_a + I_{dc}R_a$, then the controls for $IR$ drop and for speed would interlock.

It should be remembered that the counter emf of a separately excited motor is representative of speed only under the assumption of zero armature reaction, and $IR$-drop compensation is accurate only if the armature resistance does not change with temperature. It is apparent that the circuits discussed will not give speed control of high accuracy, although

they will maintain it within a few percent. The results for several different speed settings are shown in Fig. 17-16. It is probable that the droop of the curves at high currents is due not to failure of speed regulation but to the action of the current-limiting circuit to be discussed in a later section.

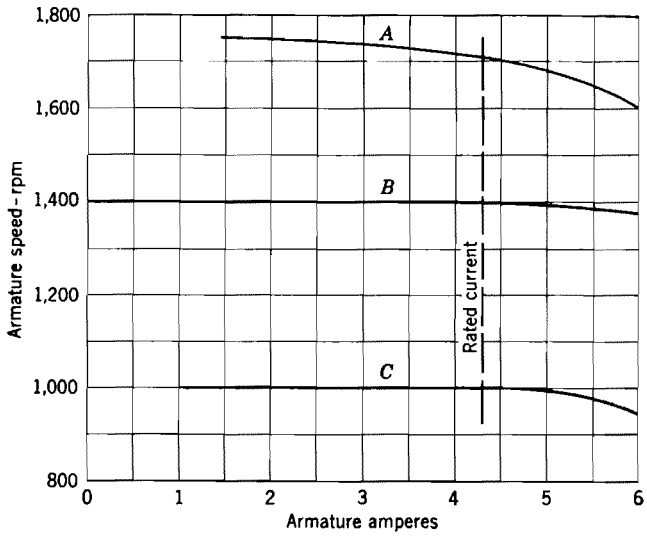



FIG. 17-16. Performance of constant-speed regulator.

**17-6. Field Weakening.** To increase the speed above the rated value with rated armature current requires that the field current be reduced. This may be easily accomplished by modifying the potentiometer $R_1$ used for speed adjustment in the preceding circuits. For half its rotation the potentiometer may be designed to perform its function as discussed, then its contact transferred to a segment giving a fixed voltage representative of the highest motor speed with full field.

A second potentiometer is coupled on the same shaft, and at its midpoint its contact transfers to a resistive sector used as a phase-shift control for the field rectifiers. Thus after reaching maximum speed with full field a grid-control circuit is applied to the field rectifiers to reduce the field current, thereby allowing the motor speed to be further increased. The counter emf remains constant in this range and the armature-voltage speed control remains effective to regulate the armature rectifiers.

**17-7. Armature-current Limiting.** It is comparatively easy to add armature-current limit control to the preceding circuits, to protect the motor and thyratrons from current surges due to starting or operation with excessive load. Such protection is of particular importance because of the limited peak-current ratings of thyratrons.

Figure 17-17 shows the addition of current limit to the previous circuits incorporating speed control and $IR$-drop compensation. The current-limit circuit consists of tube $T_6$, control resistor $R_6$, and a bias set by $R_2$. While $T_2$, the speed-control tube, and $T_6$, the current-limit tube, both operate from the current-transformer rectifier $T_5$, the inputs to the amplifiers are oppositely phased. An increasing armature current tends to make the grid of $T_6$ more positive, while it makes the grid of $T_2$ more negative because of $IR$ drop.

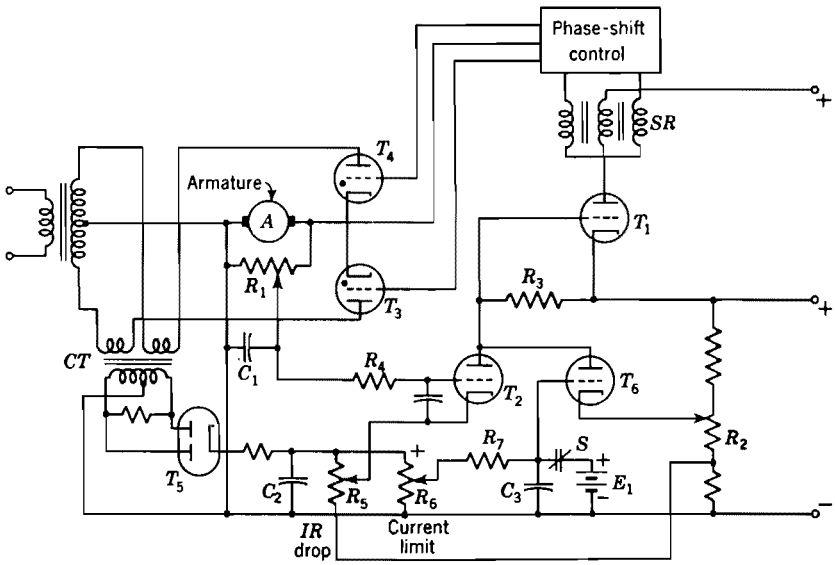



FIG. 17-17. Current-limit circuit addition.

Within the normal operating range of armature currents, from no load to approximately 150 per cent load, tube $T_6$ is biased beyond cutoff by $R_2$ and is ineffective in the circuit. Within that range tube $T_2$ is effective for speed control as previously described. When the armature current exceeds the predetermined maximum, then tube $T_6$, which may be a sharp-cutoff pentode, starts to conduct owing to the increased positive voltage across $R_6$. Because of its higher gain $T_6$ swamps the effect of $T_2$ in altering the grid voltage of $T_1$. As a result $T_6$ causes the grid of $T_1$ to become more negative. This decreases the plate current of $T_1$ through the saturable reactor $SR$ and causes the ignition angle of the thyratrons to be retarded. This prevents the armature current from exceeding a safe predetermined limit. The action of $T_6$ is illustrated by Fig. 17-18, which shows that the motor may be stalled without the current exceeding 200 per cent of the rated value.

At the instant of starting there is no armature current and no voltage across $R_6$, tube $T_6$ is cut off, and the ignition angle of the thyratrons is well advanced. The current-limit control has no effect on the magnitude of the first half cycle of current carried by the thyratrons. This first half cycle of current, limited only by armature resistance and inductance, may have a sufficiently high peak as to damage a thyratron. This will be particularly true if the setting of $R_1$ calls for high speed, since the firing angle set by $T_2$ will be far advanced.

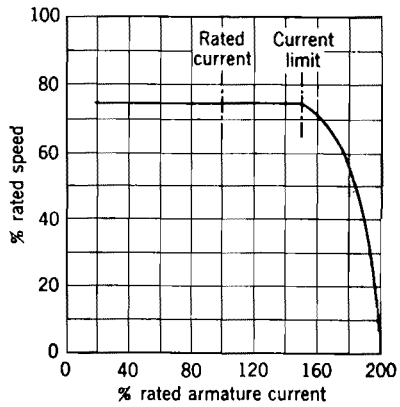



FIG. 17-18. Effect of current-limiting circuits.

A special protective circuit, consisting of $R_7$, $C_3$, switch $S$, and an auxiliary potential, is added to provide control of starting current during the initial stage of acceleration. Switch $S$ is normally closed and charges $C_3$ to $E_1$ volts, positive to the grid of $T_6$. At the instant the main-line contactor closes, the switch $S$ opens, leaving a positive potential across $C_3$ so that $T_6$ is made conducting, causing $T_1$ to be cut off and retarding the firing angle of the thyratrons so that they do not conduct.

Capacitor $C_3$ discharges through $R_7$ and $R_6$, and very quickly the plate current of $T_6$ begins to fall, allowing $T_1$ to conduct and begin to advance the thyratron firing angle. As $C_3$ discharges, the firing angle is further advanced, and as $C_3$ completely discharges, the current-limit voltage from $R_6$ is then available to take control of $T_6$ in a normal manner. The time constant of $C_3$ and its associated resistors should be equal to only a few cycles so as to allow the current-limit circuits quickly to take over the control of thyratron firing angle. The current-limit circuit then permits the acceleration of motor and load at constant current, as fixed by the setting of $R_6$.

The effect of thus shaping the first few half cycles of armature current is illustrated by Fig. 17-19. Following the first four half cycles, the current is maintained at its maximum permitted level until the acceleration period is completed; then tube $T_2$ and the speed-control circuits take over control of armature current.

**17-8. Auxiliary Controls.** Reversing of the thyratron-controlled motor can be handled by a magnetic contactor which reverses the armature connections. If the motor is running at the time of reversal, the time-delay starting circuit is brought into play so that the thyratrons will be

cut off and the contactor can reverse the terminals with no current flowing. The current which then flows after reconnection is aided by the generated emf and is in such a direction as to slow the motor down. The tube current flows during much of the time the a-c anode supply is negative, positive anodes being maintained by $E_a$. Current flow at that time represents power returned to the line, and the process is known as *regenerative braking*. After the motor is brought to zero speed, it starts in a normal manner in the reverse direction.

Dynamic braking is an added feature of most controls. When the "stop" button is pressed, the thyratrons are disconnected and the motor is connected across a low-value resistor. This absorbs the inertial energy and gives a quick stop. Any type of switching operation brings the time-delay starting circuit into play, so that a delay of a few cycles is present.

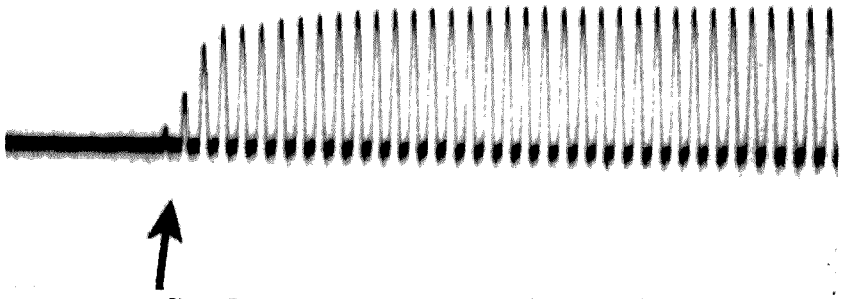


FIG. 17-19. Armature-current pulses during starting.

Time delay by synchronous motor timer is also incorporated to allow the thyratron cathodes to come to operating temperature before an initial operation of these motor controls.

**17-9. Speed Control by Tachometer.** Small motors have high resistances and poor inherent speed regulation, and for very accurate speed control the previously discussed circuits are insufficient. A means by which the motor speed is actually measured and compared with a standard reference is required, rather than obtaining the speed measurement by inference. One such means employs a small d-c tachometer generator, usually with permanent-magnet field, so that output emf without load is linearly proportional to speed. The tachometer generator is coupled to the output shaft of the controlled motor, and the tachometer output voltage is compared with a stable reference voltage. The value $E_r$ of this reference is a calibrated function of speed, and any difference between $E_r$ and the tachometer voltage $E_t$ is amplified and applied as a control signal to the main thyratrons to restore the speed to the desired standard value.

If the d-c amplifier has sufficient gain, an accuracy of a fraction of 1 per cent is possible at all except the very low speeds. Also line voltage and temperature variations have negligible effect, since the measurement of speed is a direct matter. Armature-voltage and $IR$-drop measurements are no longer needed.

Complete control systems incorporating armature-current limiting as well as tachometer speed measurement are employed. Since the system uses feedback, the question of stability naturally arises and will be further discussed in general in the next chapter.

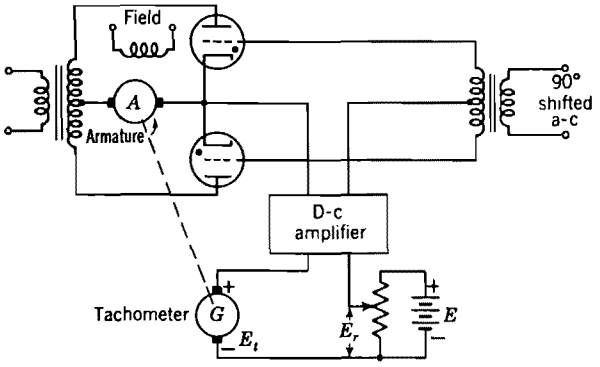



FIG. 17-20. Tachometer speed control.

**17-10. Control of Low-torque A-C Motors.** The control of small a-c motors, of single- or two-phase type, is an electronic application of importance in many recording instruments and low-power control systems. These motors are usually caused to rotate a bridge-balancing resistor and to move a recording pen, or indicator, as well as varying the position of a valve or other mechanical or electrical device. In this service the motor rotates intermittently in the direction required to bring the bridge or potentiometer into balance with the primary variable, with a speed dependent on the magnitude, or rate, of the departure of the actual variable from its indicated position.

The electronic circuits must give these small motors information on speed and direction of rotation in accordance with an input signal. Magnitude and phase of the input are usually the factors supplying this information needed by the control.

A simple form of control circuit, suited to a single-phase capacitor motor, is shown in Fig. 17-21. It employs saturable reactors $S_1$ and $S_2$, operated by vacuum tubes or small thyratrons, to control the motor. The tubes are operated with a-c plate voltage and must be given a differentially varying d-c input signal to reverse the motor.

When $T_1$ has a more positive grid voltage than $T_2$, then $T_1$ carries more plate current and reactor $S_1$ has a smaller a-c voltage drop than $S_2$. Field $F_1$ is then directly excited and field $F_2$ is partially excited from $S_1$ through the capacitor, thus giving a starting torque in one direction. When $T_2$ predominates in conduction, field $F_2$ is directly excited and $F_1$ is supplied through the capacitor, and the motor operates in the reverse direction.

Additional circuits for the control of small motors appear in Fig. 17-22. That in Fig. 17-22*a* is suited to a two-phase motor and employs a-c plate voltage, with oppositely polarized anodes, and a common a-c input signal. Each tube then compares the phase of the input voltage with its plate

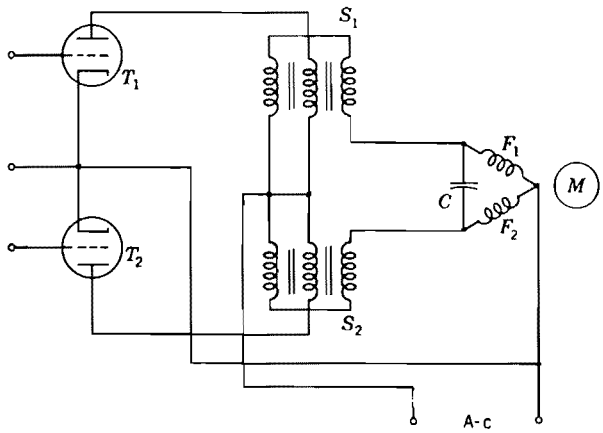



FIG. 21. Control of small a-c motor.

voltage, and that tube conducts most which has both voltages instantaneously positive. Reversal of motor rotation is obtained, depending on which tube is conducting as determined by the phase of the input signal.

The circuit of Fig. 17-22*b* employs a two-field reversing motor, controlled by series transformers $L_1$ and $L_2$, with secondaries polarized as indicated by the dots. Because of saturation by the d-c plate-current component, as well as loading of the transformers on the conducting half cycles, a ratio of primary impedance change of at least 10:1 may be obtained for $L_1$ and $L_2$ between nonconduction and full conduction. This allows most of the line voltage to appear across the motor field when a tube conducts, and drives the motor in the desired direction. The capacitors $C$ increase the voltage applied to the field by series resonance.

Reversal of phase of the input signal will operate one tube or the other. The magnitude of the a-c input determines the shunting effect on the series transformers and controls the speed of the motor. A circuit such

as this is well suited to operate from an a-c bridge, where 180° phase change denotes departure in one direction or the other from the balance position.

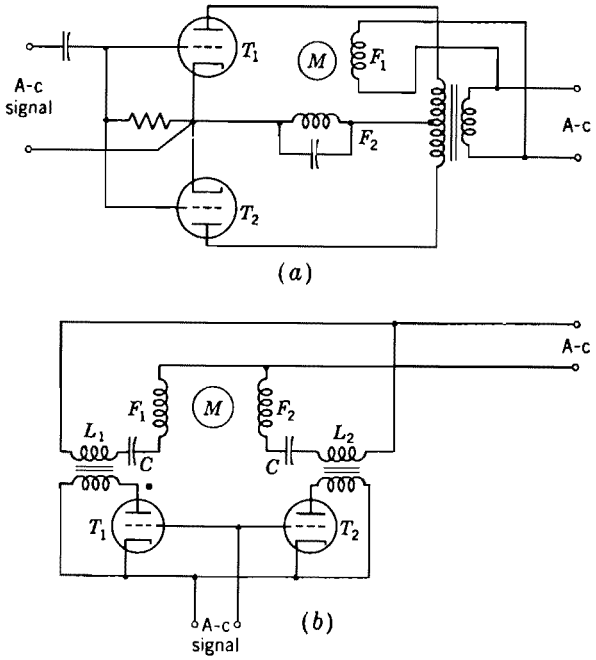



FIG. 17-22. Phase-sensitive small-motor control circuits.

**17-11. Phase-selective Circuits.** Phase discrimination, or selection, as a means of obtaining directional or polarity information, has been employed in the last several circuits, which are representative of a broad general class of phase-selective detectors. Such circuits are commonly used with various indicating systems or with two-phase motors, and a basic circuit form appears in Fig. 17-23.

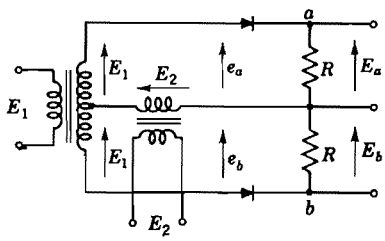



FIG. 17-23. Fundamental phase-selective detector.

Two a-c signals, of rms value $E_1$ and $E_2$, from a common source of frequency but of arbitrary phase, are introduced into the circuit. These voltages may be defined as

$$e_1 = E_{1m} \sin \omega t \tag{17-29}$$

$$e_2 = E_{2m} \sin (\omega t + \phi) \tag{17-30}$$

where $\phi$ is arbitrary but is usually zero or 180°. In the secondary circuit voltages $e_a$ and $e_b$ become

$$e_a = E_{1m} \sin \omega t + E_{2m} \sin (\omega t + \phi) \quad (17\text{-}31)$$

$$e_b = E_{1m} \sin \omega t - E_{2m} \sin (\omega t + \phi) \quad (17\text{-}32)$$

For the purpose of analysis the rectifiers may be assumed to have identical linear dynamic characteristics, and this requirement will be met if $r_p \ll R$. After rectification the pulsating d-c output voltage will be

$$E_a = \frac{1}{2\pi} \int_0^{\pi} E_{1m} \sin \omega t \, d\omega t + \frac{1}{2\pi} \int_0^{\pi} E_{2m} \sin (\omega t + \phi) \, d\omega t$$

and

$$E_b = \frac{1}{2\pi} \int_0^{\pi} E_{1m} \sin \omega t \, d\omega t - \frac{1}{2\pi} \int_0^{\pi} E_{2m} \sin (\omega t + \phi) \, d\omega t$$

After carrying out the integration and setting $b = E_{2m}/E_{1m}$, then

$$E_a = \frac{E_{1m}}{\pi} (1 + b \cos \phi) \quad (17\text{-}33)$$

$$E_b = \frac{E_{1m}}{\pi} (1 - b \cos \phi) \quad (17\text{-}34)$$

Thus the magnitude of the two d-c voltages to the common point is dependent on the angle $\phi$, and if $b = 1$, then the output will vary from zero to $2E_{1m}/\pi$ in either direction as $\phi$ goes from zero to 180°.

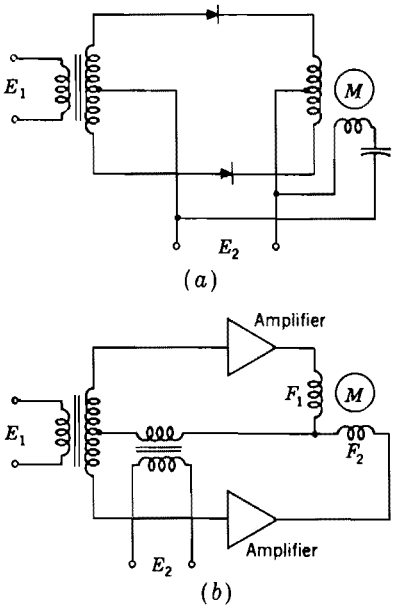



FIG. 17-24. Alternating-current motor control by phase-selective detectors.

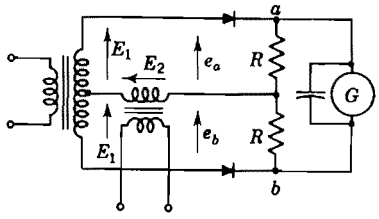



FIG. 17-25. Phase indicator.

Two-phase motor control by circuits of this nature is demonstrated in Fig. 17-24. In the circuit of Fig. 17-24*b* a value $\phi = \pm 90°$ is ordinarily employed, and the motor undertakes the detection operation without separate rectification. This is a common circuit for servomechanisms.

The basic phase detector becomes a convenient phase indicator by a modification, as in Fig. 17-25, in which the rectifiers should follow a square-law characteristic. This usually requires that $r_p \gg R$. Then it

may be assumed that

$$i_b = a_o + a_1 e + a_2 e^2 \tag{17-35}$$

and if $e_a$ and $e_b$ are as in Eqs. (17-31) and (17-32), then the currents in the two diode circuits will be

$$i_{b1} = a_o + a_1 E_{1m} \sin \omega t + a_1 E_{2m} \sin (\omega t + \phi) + a_2 E_{1m}^2 \sin^2 \omega t + 2a_2 E_{1m} E_{2m} \sin \omega t \sin (\omega t + \phi) + a_2 E_{2m}^2 \sin^2 (\omega t + \phi) \tag{17-36}$$

$$i_{b2} = a_o - a_1 E_{1m} \sin \omega t + a_1 E_{2m} \sin (\omega t + \phi) + a_2 E_{1m}^2 \sin^2 \omega t - 2a_2 E_{1m} E_{2m} \sin \omega t \sin (\omega t + \phi) + a_2 E_{2m}^2 \sin^2 (\omega t + \phi) \tag{17-37}$$

These currents will produce a net voltage across terminals $a$-$b$ equal to

$$e_{ab} = (i_{b1} - i_{b2})R = 2a_1 E_{1m} R \sin \omega t + 2a_2 E_{1m} E_{2m} R \sin \omega t \sin (\omega t + \phi)$$

Expanding the second term on the right in accordance with the identity

$$\sin a \sin b = \tfrac{1}{2}[\cos (a - b) - \cos (a + b)]$$

leads to

$$e_{ab} = 2a_1 E_{1m} R \sin \omega t - a_2 E_{1m} E_{2m} R \cos (2\omega t + \phi) + a_2 E_{1m} E_{2m} R \cos \phi \tag{17-38}$$

The fundamental and second-harmonic terms may be bypassed by a capacity, leaving

$$E_{ab} = a_2 E_{1m} E_{2m} R \cos \phi \tag{17-39}$$

as a d-c component. If $E_{1m}$ and $E_{2m}$ are made constant by limiting, then the voltage $E_{ab}$ is proportional to cos $\phi$ or the power-factor angle and an indicating instrument across $a$-$b$ will measure angle $\phi$. A circuit such as this, with output bypassed for alternating current, might well be used as input for the motor control of Fig. 17-21.

## PROBLEMS

**17-1.** A ¾-hp motor is operated in the circuit of Fig. 17-2. The armature resistance is 2.48 ohms, inductance is 0.021 h. When tested wth supply for the armature from a d-c line of 115 volts, the motor operates at 1800 rpm and armature current of 5.7 amp with ¾ shp. Neglecting tube drop, calculate and plot a curve of $T/T_o$ against the speed coefficient $a$ for $\theta_1 = 90°$ when this motor is supplied by thyratrons. The transformer input is 300 volts total on the secondary.

**17-2.** The motor of Prob. 17-1 is operated at a speed of 1000 rpm from the thyratrons. Calculate the average current.

**17-3.** Explain how an $IR$-drop compensating circuit may be employed for both over- and undercompounding of a motor.

**17-4.** Compute the ratio $I_{rms}/I_{dc}$ for the motor of Prob. 17-1 with $\theta_1 = 80°$ at a speed of 1700 rpm.

**17-5.** The motor of Prob. 17-1 is used in the circuit of Fig. 17-12. If $T_1$ is a 6AU6 with $g_m = 4000$ μmhos and very high $r_p$, $R_2 = 1000$ ohms, $R_3 = 9000$ ohms, $R_1 =$

$R_4$ = 100,000 ohms, compute the change in output voltage at $X$ for a speed change from 1700 to 1750 rpm.

## REFERENCES

1. Ryder, J. D.: Electronic Control for D-C Motors, *Electronics*, **12**: 20 (December, 1938).
2. Vedder, E. H., and K. P. Puchlowski: Theory of Rectifier–D-C Motor Drives, *Trans. AIEE*, **62**: 863 (1943).
3. Moyer, E. E.: Electronic Control of D-C Motors, *Electronics*, **16**: 98 (May, 1943), 119 (June, 1943), 118 (July, 1943), 133 (September, 1943), 128 (October, 1943).
4. Puchlowski, K. P.: Electronic Control of D-C Motors, *Trans. AIEE*, **62**: 870 (1943).
5. Chute, G. M.: "Electronics in Industry," McGraw-Hill Book Company, Inc., New York, 1946.
6. Schmidt, A., Jr.: Operating D-C Motors from Controlled Rectifiers, *Elec. Eng.*, **67**: 516 (1948).
7. Engineers of the Westinghouse Electric Corporation: "Industrial Electronics Reference Book," John Wiley & Sons, Inc., New York, 1948.
8. Chance, F. B., et al.: "Waveforms," McGraw-Hill Book Company, Inc., New York, 1949.
9. Hague, B.: "Alternating-current Bridge Methods," Pitman Publishing Corporation, New York, 1930.

CHAPTER 18

# SERVOMECHANISMS

Man, with his inherently slow responses, is a limiting factor on the speed and complexity of many modern devices, and the importance of automatic control to replace him is being increasingly realized. Because of the speed of operation of electronic devices the electronic art has been frequently called upon in the design of automatic control systems, and these systems now constitute a very large class of electronic applications.

Almost all such systems employ feedback, and the methods of analysis are closely allied to those of the feedback amplifier, as developed by Black, Nyquist, and Bode. Because of this unity of thought and method the analysis and design of servomechanisms and other control systems are an area of considerable interest to those working in the industrial electronic field.

**18-1. Servomechanisms and Control.** An *open-loop system* is a control in which no check of output magnitude against input, or no feedback, is employed. For example, a generator to supply power to a laboratory is remotely installed with its field rheostat. The scale of the rheostat is calibrated in line voltage, but no voltmeter is available. It is assumed by the designer that the voltage supplied to the laboratory is that indicated by the original calibration of the rheostat. Actually such a calibration would not be reliable, since the voltage could be affected by load, generator speed, field voltage, and other variables. There is no check on the controlled quantity, and the operation is that of an open-loop system.

If to the above were added a contact-making voltmeter in the laboratory to measure the load voltage, with a follow-up device such that every movement of the voltmeter needle were duplicated by a motor driving the field rheostat, so that the line voltage was maintained at a preset value, then the operation becomes that of a *closed-loop system.* In this equipment a continuous check would be kept on the load voltage, and any error, or difference between the desired value and the actual value, would immediately call for a correcting action.

The term servomechanism is ordinarily applied to a closed-loop controller of position, be it rotary or translatory. Thus the device which would follow the motion of the voltmeter needle in the above example,

and cause a motor to drive the field rheostat to a corresponding position, is a servomechanism.

There are two additional implications of this term, however. First, the usage implies that the energy required from the output must come from a local source and is not furnished by the input. In fact the output energy required is usually much greater than the possible input energy, and the system is usually a power amplifier. Second, the system is caused to operate by an error signal, or the difference (or error) between the input and a desired corresponding output. That is, if $\theta_i$ is the input angular position at any instant and $\theta_o$ is the output angular position at the same instant, then the error is

$$\text{Error} = \epsilon = \theta_i - \theta_o \tag{18-1}$$

and this is the actuating signal. The fundamentals of a servo system

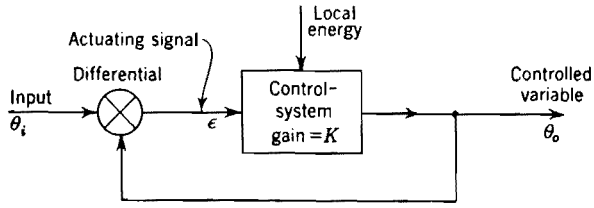



FIG. 18-1. Fundamental servomechanism.

are illustrated in Fig. 18-1. Control by error, or input-output difference, is the essence of closed-loop control, as well as of the feedback amplifier.

From the figure it can be seen that by reason of the feedback the input always knows what the output is doing, and if the output position does not correspond to the input position at a given instant, then the control reacts by driving $\theta_o$ toward the value of $\theta_i$.

In contrast, in an open-loop system, the input is always ignorant of the output position; it can only tell the output to change. Whether the output actually accomplishes the change is dependent on system parameters, whereas a closed-loop system is almost independent of system parameters. The price paid for this improved performance is the increased design complexity, to ensure a system which will be stable and nonoscillatory.

Controllers of physical quantities, such as pressure, temperature, or voltage, are known as *regulators*. Thus the over-all mechanism for control of the generator voltage discussed above, including the voltmeter and the field rheostat, might be considered a regulator, whereas the device for transmitting voltmeter pointer position to rheostat position would be a servomechanism. Another example is the gas-pressure regulator of Fig.

18-2. The distinction between servomechanisms and regulators is not always clear, and may not be necessary from a fundamental view. The definition of a servomechanism as related to position, however, does explain the usual system analysis in terms of an instantaneous angle $\theta$ as the independent variable.

A *continuous control* is one in which the output is under control of the input at all times. This is in contrast to a form in which output corrections are made only at fixed intervals.

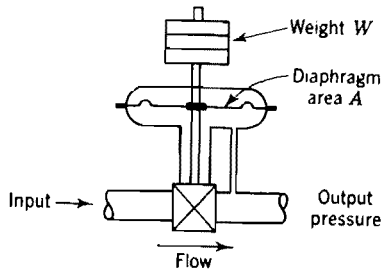



FIG. 18-2. Gas-pressure regulator: output pressure = $W/A$ psi.

A *proportional control* develops an output torque more or less proportional to the difference between the output and the desired value of the output. The correcting torque, or force, may be proportional to other factors as well, such as rate of change of error, or the integral of the error.

*On-off control* has only two operating conditions or states, and the correction applied is dependent only on the direction of the error, not on its magnitude. A home thermostat is usually of this type.

Analysis for performance and stability of a design depends on determination of the transient and steady-state solutions for various types of input. In particular the transient response to a step input and the steady-state response to a sinusoidal input give much useful information and will be studied in some detail. These analyses should be considered as applying to feedback control systems in general, irrespective of whether the amplifier gain is obtained by an electrical, mechanical, or hydraulic system.

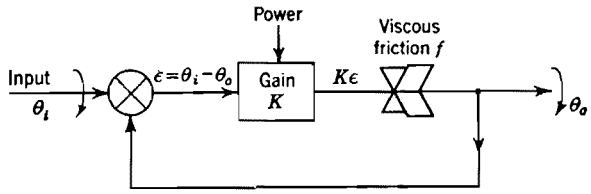



FIG. 18-3. Simple first-order servomechanism.

**18-2. First-order Servomechanism.** A simple idealized servomechanism system which leads to a differential equation of the first order is shown in Fig. 18-3. The system is intended to make the output shaft angle $\theta_o$ follow the input shaft angle $\theta_i$ at all times.

The output position $\theta_o$ is fed back and subtracted from the input $\theta_i$ to give the error $\epsilon = \theta_i - \theta_o$, which serves as input to the amplifier of gain

$K$ in torque developed per radian of error input. The output is then a torque of value $K\epsilon = K(\theta_i - \theta_o)$ exerted on the output shaft. Acting as a load on the output shaft is a device with negligible inertia, but with a viscous-damping constant, or coefficient of viscous friction, $F$. This assumes that the friction torque exerted as a load is proportional to the velocity of rotation, or the friction load torque is $F\,d\theta/dt = Fp\theta$, using the $p$ notation for the time derivative. Such a load device might be a small lightweight indicator of some kind.

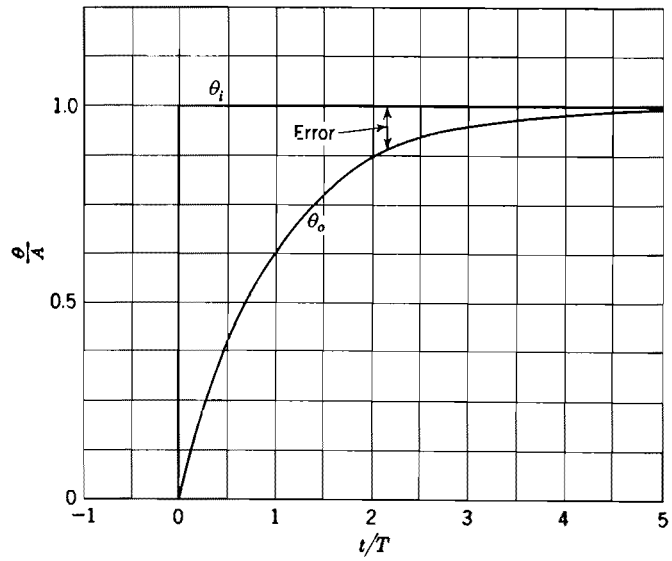



FIG. 18-4. Response $\theta_o$ of first-order servomechanism to step-function input $\theta_i$.

The differential equation* of the system may be obtained by setting the sum of the output shaft torques to zero, or

$$Fp\theta_o - K(\theta_i - \theta_o) = 0 \tag{18-2}$$

$$\theta_o = \frac{K\theta_i}{K + Fp} = \frac{1}{1 + Fp/K}\,\theta_i \tag{18-3}$$

The latter equation expresses the relation that must exist between $\theta_o$ and $\theta_i$.

A *step-function* input of the form of $\theta_i$ (Fig. 18-4) is a useful form for testing the response of a servomechanism, just as it was for an amplifier, although in the latter case it was repetitive and called a square wave.

* English units are most usually employed, and so the units of $\theta$ are radians, torque $T$ is in foot-pounds, viscous damping $F$ in foot-pounds per radian per second, and inertia $J$ in slug-feet$^2$.

This step-function input has the property that

$$\theta_i = 0 \qquad \text{for } t \leq 0$$
$$\theta_i = A \qquad \text{for } t > 0$$

and is sometimes called a *constant-amplitude input.* Then from Eq. (18-2), for $t > 0$,

$$\frac{F}{K} p\theta_o + \theta_o = \theta_i = A \tag{18-4}$$

The second term $\theta_o$ of the equation is an angle and therefore dimensionless. From the first term, since $p = d/dt$, it can be seen that the coefficient $F/K$ must have the dimensions of time in order that the first term may also be dimensionless. If then

$$T = \frac{F}{K} \tag{18-5}$$

the transient solution may be obtained by writing Eq. (18-4) as

$$(Tp + 1)\theta_o = 0 \tag{18-6}$$

which is recognized as having a solution

$$\theta_{ot} = C\epsilon^{-t/T} \tag{18-7}$$

where $C$ is a constant to be determined. From similar relations which apply in $RC$ and $RL$ circuits the parameter $T$ can be seen as the *time constant.*

The steady-state solution to Eq. (18-4) must next be obtained. For a feedback system it can be reasoned that the steady-state solution will have the same functional form as the driving function $\theta_i$. The steady-state value of $\theta_i = A$, and so in steady state $p\theta_i = 0$. From Eq. (18-4) the steady-state solution is then

$$\theta_{os} = A$$

and the complete solution is

$$\theta_o = A + C\epsilon^{-t/T} \tag{18-8}$$

If the output reference position be such that, at $t = 0$, $\theta_o = 0$, then

$$C = -A$$

and Eq. (18-8) becomes

$$\theta_o = A(1 - \epsilon^{-t/T}) \tag{18-9}$$

for $t > 0$. Figure 18-4 is a plot of the step response, in terms of the dimensionless terms $\theta_o/A$ and $t/T$. The transient is of exponential form, and after considerable time, or for $t$ large, the output $\theta_o$ will equal the

input $A$, or there will be no steady-state error for a first-order system with a step input. This result also corresponds to that for a first-order electrical circuit involving either $R$ and $C$ or $R$ and $L$.

Since $T = F/K$, the response may be speeded by increasing the gain $K$, or designing a larger and more powerful amplifier. In so doing, however, the larger components will introduce appreciable inertia effects, and the assumed first-order system becomes one of second order, with a much more complex response.

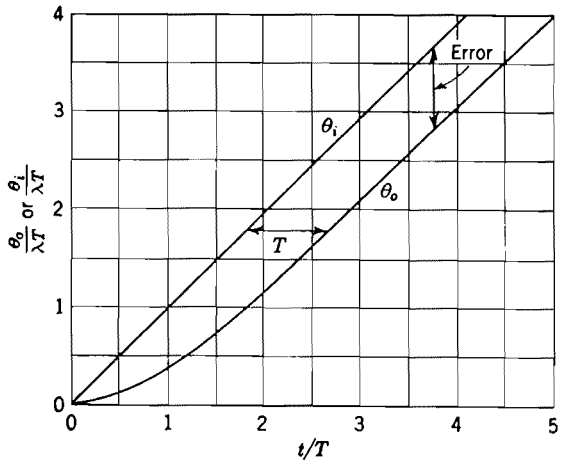



FIG. 18-5. Ramp response of a first-order servomechanism.

The response of the first-order system to a so-called *ramp input* $\theta_i$ of the form of Fig. 18-5 is also of interest. The input function is

$$\theta_i = 0 \qquad \text{for } t \leq 0$$
$$\theta_i = \lambda t \qquad \text{for } t > 0$$

which is also called a *constant-velocity input function.* The transient solution is, of course, independent of the driving function, so that the form of Eq. (18-7) is once more obtained. It is logical to assume a steady-state solution having the same form as $\theta_i$; thus

$$\theta_{os} = B_1 t + B_2 \tag{18-10}$$

Using Eq. (18-10) and with $\theta_i = \lambda t$ in Eq. (18-4) leads to

$$B_1 T + B_1 t + B_2 = \lambda t \tag{18-11}$$

Equating like powers of $t$ on both sides of the equation gives

$$B_1 = \lambda \qquad B_2 = -\lambda T$$

Combining these results with Eq. (18-10), and adding the transient solution, gives the value of $\theta_o$ as

$$\theta_o = C\epsilon^{-t/T} + \lambda(t - T) \tag{18-12}$$

Applying the initial condition that $\theta_o = 0$ at $t = 0$, then $C = \lambda T$, and the complete solution for the ramp input is

$$\theta_o = \lambda(t - T + T\epsilon^{-t/T}) \tag{18-13}$$

If the output were to follow the input exactly, then $\theta_o$ would equal $\lambda t$. Instead, Eq. (18-13) shows that there is an exponential transient error plus a steady-state error or lag in time of $T$ sec.

A dimensionless plot of $\theta_o/\lambda T$ vs. $t/T$ appears in Fig. 18-5. The steady-state error can be reduced by making $K$ large to reduce $T$, but once again some previously neglectable inertias would become important, and the system actually would become one of second order.

Steady-state response to a sinusoidal input is also of importance, either because knowledge of this response may be used to predict the response to any arbitrary periodic function through a Fourier series, or because the knowledge is helpful in predicting the stability of the system in closed-loop operation. The input may be assumed of the form

$$\theta_i = D \cos \omega t$$

which may also be written as

$$\theta_i = D \operatorname{Re} (\epsilon^{j\omega t}) \tag{18-14}$$

where Re implies the real part of $\epsilon^{j\omega t}$, which is cosine $\omega t$.

The steady-state output may be assumed of the form

$$\theta_{os} = H\epsilon^{j\omega t} \tag{18-15}$$

Since the original differential equation was

$$(Tp + 1)\theta_o = \theta_i$$

substitution of Eq. (18-15) leads to

$$(j\omega T + 1)H\epsilon^{j\omega t} = D\epsilon^{j\omega t} \tag{18-16}$$

where the real part is to be taken. Then

$$H = \frac{D}{j\omega T + 1} \tag{18-17}$$

Writing in polar form,

$$H = \frac{D\epsilon^{-j\phi}}{\sqrt{1 + \omega^2 T^2}}$$

where

$$\phi = \tan^{-1} \omega T \tag{18-18}$$

Substitution in Eq. (18-15) and taking the real part gives the steady-state solution as

$$\theta_{os} = \frac{D \cos (\omega t - \phi)}{\sqrt{1 + \omega^2 T^2}} \tag{18-19}$$

For $\omega T \ll 1$ the ratio of output amplitude to input amplitude is virtually unity, and the phase angle near zero. As $\omega T$ becomes appreciable with respect to unity, the amplitude starts to fall and a phase lag appears. For $\omega T \gg 1$ the amplitude falls as $1/\omega T$, and the phase angle approaches $-90°$. The time constant $T$ may be readily found from the sinusoidal response by noting that $\omega T = 1$, or $T = 1/\omega$, when the response $|\theta_o|/D$ is 0.707 of the low-frequency or zero-frequency value.

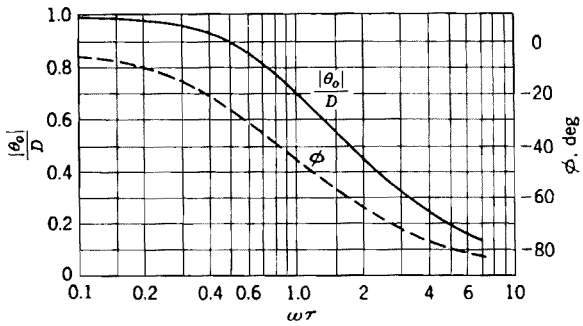



FIG. 18-6. Sinusoidal response of a first-order servomechanism.

Plots of the dimensionless ratio $|\theta_o|/D$ as well as $\phi$ appear in Fig. 18-6 and are similar in shape to amplifier response curves where first-order $RC$ systems are involved. In servomechanisms the frequency range of interest usually lies in the region from zero to a few tens of cycles.

The first-order system is not realized except in a few special cases, but it serves to introduce the methods of analysis that will be followed. A more realistic form is a second-order system, in which the normal equipment inertias are present, and which adds the problem of stability to those of transient and steady-state response already considered for first-order systems.

**18-3. Second-order Systems.** When inertia is included in the simple system, as in Fig. 18-7, the system becomes a more usual one, with a second-order equation as a description of performance. The total system inertia referred to the output shaft is $J$, and $F$ is the viscous friction, any Coulomb or nonlinear friction being neglected as small. Also neglected are any external torque loads. Amplifier gain $K$ is again in pound-feet torque per radian of input error angle.

The torque on the output shaft may be summed to obtain the system differential equation

$$J\frac{d^2\theta_o}{dt^2} + F\frac{d\theta_o}{dt} - K(\theta_i - \theta_o) = 0$$

which may be written as

$$\left(\frac{J}{K}p^2 + \frac{F}{K}p + 1\right)\theta_o = \theta_i \tag{18-20}$$

and from which

$$\theta_o = \frac{1}{(J/K)p^2 + (F/K)p + 1}\theta_i \tag{18-21}$$

By dividing out the fraction on the right-hand side the equation may be put into an infinite series in powers of the operator $p$. That is,

$$\theta_o = \left[1 - \frac{F}{K}p + \left(\frac{F^2}{K^2} - \frac{J}{K}\right)p^2 - \cdots\right]\theta_i \tag{18-22}$$

and this form is generally useful in determining the steady-state response to algebraic time-function inputs. For instance, when $\theta_i$ is a step function, then $p\theta_i = p^2\theta_i = p^n\theta_i = 0$ in the steady state and $\theta_o = \theta_i$, or the second-order servomechanism has zero steady-state error for a step or constant-amplitude input.

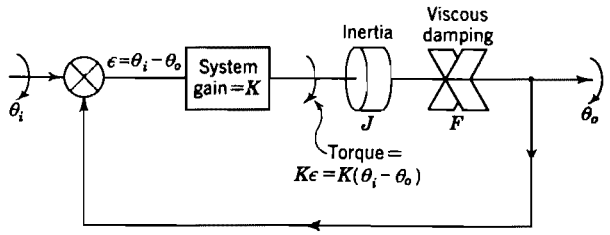



FIG. 18-7. Second-order position control.

For a ramp input or one with constant velocity, with $\theta_i = \lambda t$, then $p\theta_i = \lambda$, $p^2\theta_i = 0$, $p^n\theta_i = 0$ for $n > 1$. From Eq. (18-22) the steady-state response to a ramp input is

$$\theta_{os} = \left(1 - \frac{F}{K}p\right)\theta_i = \lambda\left(t - \frac{F}{K}\right) \tag{18-23}$$

Since the input is $\theta_i = \lambda t$, then it is seen that the output lags behind the input by $F/K$ sec. This shows that the controlled variable $\theta_o$ always has an error sufficiently large so that a torque can be developed to balance the steady-state viscous-damping torque $F$, and to provide the velocity of the output, since the ramp input may be considered as a constant velocity. Because of this lag the second-order system is seen to behave much as a first order on such an input.

The usage of Eq. (18-22) may be further demonstrated by the assumption of an input for which

$$\theta_i = \lambda t^2 \tag{18-24}$$

Then $p\theta_i = 2\lambda t$, $p^2\theta_i = 2\lambda$, $p^n\theta_i = 0$ for $n > 2$. The steady-state solution then follows as

$$\theta_{os} = \lambda t^2 - 2\lambda\left(\frac{F}{K}t - \frac{F^2}{K^2} + \frac{J}{K}\right) \tag{18-25}$$

The first term on the right is the input function. Thus the original acceleration is continued, but the error increases as a function of time. This may be analyzed as due to the necessity for the error signal to develop enough output torque to overcome the friction load and also to take care of the resisting torque due to the acceleration of the inertia load.

**18-4. Transient Response of a Second-order System to a Step Input.** Returning to the second-order-system differential equation from Eq. (18-20),

$$\left(\frac{J}{K}p^2 + \frac{F}{K}p + 1\right)\theta_o = \theta_i \tag{18-26}$$

for the system of Fig. 18-7, a step-function input may be applied, where

$$\theta_i = 0 \qquad \text{for } t \leq 0$$
$$\theta_i = A \qquad \text{for } t > 0$$

It has already been shown that the steady-state solution is $\theta_{os} = A$. The transient solution may be obtained by assuming its form as

$$\theta_{ot} = Ce^{\gamma t}$$

where $C$ is arbitrary. Substitution in Eq. (18-26) gives

$$\frac{J}{K}\gamma^2 + \frac{F}{K}\gamma + 1 = 0$$

from which the two roots are obtained as

$$\gamma_{1,2} = -\frac{F}{2J} \pm \sqrt{\frac{F^2}{4J^2} - \frac{K}{J}} \tag{18-27}$$

From Eq. (18-27) it is apparent that the two roots are of the form

$$\gamma_{1,2} = -a + jb \tag{18-28}$$

where

$$a = \frac{F}{2J} \tag{18-29}$$

$$b = \sqrt{\frac{K}{J} - \frac{F^2}{4J^2}} \tag{18-30}$$

and $a$ is the reciprocal time constant of the system. Equation (18-30) shows that $b$ may be real, zero, or imaginary, depending on the system parameters, particularly on the gain $K$ and the damping $F$.

The transient solution is given by

$$\theta_{ot} = \epsilon^{-at}(C_1\epsilon^{jbt} + C_2\epsilon^{-jbt}) \tag{18-31}$$

The transient dies out at a rate dependent on the value of the exponent $a = F/2J$, a direct function of the damping. The critical value of $F$, for which $b = 0$, is

$$F_c = 2\sqrt{KJ}$$

and it is convenient to define a *damping ratio* $\xi$,

$$\xi = \frac{F}{2\sqrt{KJ}} \tag{18-32}$$

as the ratio of the actual damping to the value for critical damping.

If the damping $F$ were zero, the solution would reduce to

$$\begin{aligned}\theta_{ot} &= C_1\epsilon^{j\sqrt{K/J}t} + C_2\epsilon^{-j\sqrt{K/J}t} \\ &= D_1 \cos\sqrt{\frac{K}{J}}\,t + D_2 \sin\sqrt{\frac{K}{J}}\,t\end{aligned} \tag{18-33}$$

and the result is oscillatory with an *undamped natural angular frequency* of

$$\omega_n = \sqrt{\frac{K}{J}} \tag{18-34}$$

It may then be noted that

$$a = \xi\omega_n \tag{18-35}$$

$$b = \omega_n\sqrt{1 - \xi^2} \tag{18-36}$$

and $\xi$ and $\omega_n$ may be considered convenient parameters for use in describing the system performance.

The transient solution of Eq. (18-31) then becomes

$$\theta_{ot} = \epsilon^{-\xi\omega_n t}(C_1\epsilon^{j\omega_n\sqrt{1-\xi^2}t} + C_2\epsilon^{-j\omega_n\sqrt{1-\xi^2}t}) \tag{18-37}$$

There are three possible cases, depending on the relation of $\xi^2$ to unity.

*Case* I. $\xi < 1$. It has been previously reasoned that for steady-state inputs $\theta_o = \theta_i = A$, so that for $\xi < 1$ the complete solution may be written from Eq. (18-37) as

$$\theta_o = A + \epsilon^{-\xi\omega_n t}(D_1 \cos\omega_n\sqrt{1 - \xi^2}\,t + D_2 \sin\omega_n\sqrt{1 - \xi^2}\,t) \tag{18-38}$$

In order to evaluate the constants, the initial conditions

$$\theta_o = 0,\ p\theta_o = 0 \qquad \text{for } t = 0$$

may be applied to Eq. (18-38) and its derivative, leading to

$$D_1 = -A \qquad D_2 = -\frac{A\xi}{\sqrt{1-\xi^2}}$$

The response of the second-order servomechanism, with $\xi < 1$, may then be written as

$$\frac{\theta_o}{A} = 1 - \epsilon^{-\xi\omega_n t}\left(\cos \omega_n \sqrt{1-\xi^2}\, t + \frac{\xi}{\sqrt{1-\xi^2}} \sin \omega_n \sqrt{1-\xi^2}\, t\right) \quad (18\text{-}39)$$

The result is oscillatory, as would be expected when the roots have imaginary parts or are complex. The system with $\xi < 1$ is said to be *underdamped.*

*Case* II. $\xi > 1$. The roots of Eq. (18-27) are both real, and the result is a heavily damped and slow system. The complete solution is

$$\theta_o = A + \epsilon^{-\xi\omega_n t}(D_1 \epsilon^{\omega_n \sqrt{\xi^2-1}\, t} + D_2 \epsilon^{-\omega_n \sqrt{\xi^2-1}\, t}) \quad (18\text{-}40)$$

Use of the initial conditions on Eq. (18-40) and its derivative leads to

$$\frac{\theta_o}{A} = 1 - \frac{\epsilon^{-\xi\omega_n t}}{2}\left[\left(1 + \frac{\xi}{\sqrt{\xi^2-1}}\right)\epsilon^{\omega_n \sqrt{\xi^2-1}\, t} + \left(1 - \frac{\xi}{\sqrt{\xi^2-1}}\right)\epsilon^{-\omega_n \sqrt{\xi^2-1}\, t}\right] \quad (18\text{-}41)$$

Thus, when $\xi > 1$, the response to a constant-amplitude input is exponential, and the system is said to be *overdamped.*

*Case* III. $\xi = 1$. For this case the two roots of Eq. (18-27) become one, or $b = 0$. The transient solution may be found as

$$\theta_{ot} = \epsilon^{-\omega_n t}(C_1 + C_2 t) \quad (18\text{-}42)$$

and $A$ may be added to obtain the complete solution. Use of the initial conditions gives

$$\frac{\theta_o}{A} = 1 - \epsilon^{-\omega_n t}(1 + \omega_n t) \quad (18\text{-}43)$$

for the case of $\xi = 1$. This is called the condition of critical damping, since it represents the least damping which may be used without an overshooting response. Physically the case of repeated roots cannot occur, since it is not possible practically to set or maintain a system with $\xi$ exactly equal to unity.

It is obvious that none of these solutions allows $\theta_o$ to instantaneously follow the step input $\theta_i$, because of inertia and damping effects. The ability of $\theta_o$ to follow a step input is illustrated in Fig. 18-8, where the response $\theta_o/A$ is plotted against time with damping ratio $\xi$ as a parameter. Values of $\xi$ between 0.3 and 1 give fast response, with $\xi = 0.7$ favored if only a small overshoot can be tolerated.

The response time can be reduced by increasing the gain $K$, which increases $\omega_n$, but the damping $F$ must then also be increased to maintain the same damping ratio $\xi$, and this increases the energy loss. Large $K$ creates a stiff system, but this is limited by the effects of gear backlash and other time lags.

The presence of complex conjugate roots has led to a possible oscillatory system. However, the oscillation was ultimately damped out *if the real part of the roots was negative.* Thus a test for stability of such

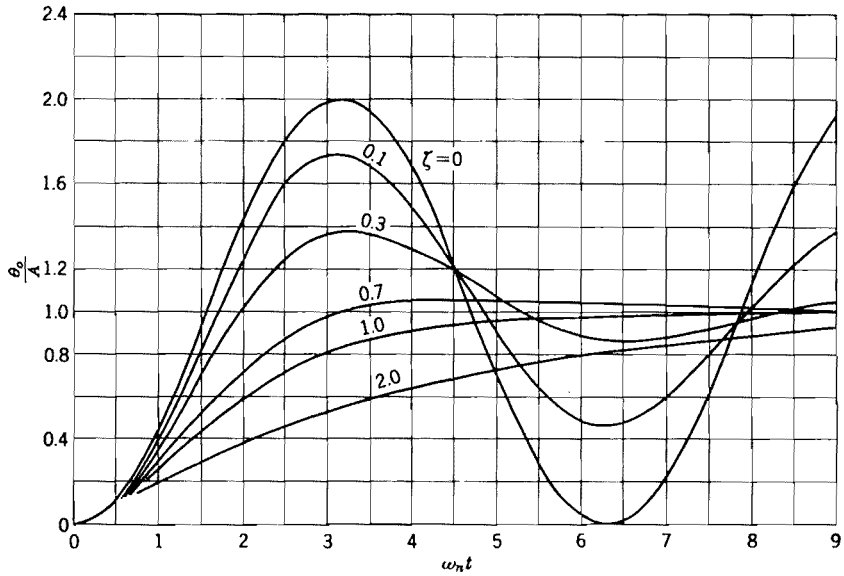



FIG. 18-8. Step-function response of a second-order system.

systems involves determining whether or not any roots exist with positive real parts.

It was shown in Eq. (18-23) that for a ramp input function the steady-state error was $F/K$ sec. If this is a form of input likely to be encountered, then $\xi$ should be reduced, and more overshoot tolerated on sudden input position changes, in order to reduce the steady-state error for sudden changes in velocity. Thus the choice of a damping factor is not clear-cut but is dependent upon the nature of the input.

**18-5. Derivative Control.** In the previous system the input to the torque amplifier was the error signal $\theta_i - \theta_o$, equal to the position error, and such a system is said to have position control. The effectiveness of this control has been shown to depend very largely on the amount of viscous damping employed, with the choice dependent on several factors.

The value of gain $K$ may also be increased to improve response, but this requires a corresponding increase in damping with a resultant energy loss.

It is also found in many actual servo systems that the viscous friction inherent in the mechanism, gears, and the like, is such that $\xi < 0.3$. In order to make the transient response satisfactory, it is necessary to add viscous damping, and this may dissipate a large share of output energy of the motor or other driving device. It has been found that a more satisfactory correction of the system characteristics can be achieved by adding a *derivative-feedback* loop.

Such a system is diagramed in Fig. 18-9, where not only is the output position $\theta_o$ fed back and subtracted from the input, but also the derivative $p\theta_o$ is fed back and subtracted. This is equivalent to subtracting a quantity proportional to the output velocity. Such a velocity indication

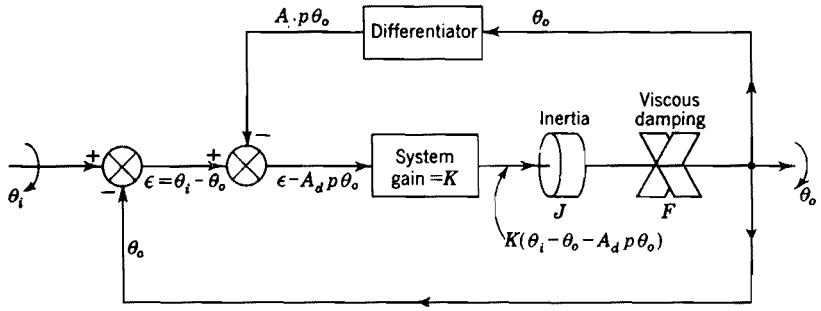



FIG. 18-9. Derivative-feedback control.

can be obtained by use of a tachometer generator on the output shaft, or by electrical differentiating circuits in the amplifier, if $\theta_o$ can be represented in terms of a voltage. The damping torque is proportional to velocity, and it can be seen that the torque developed due to the feedback velocity signal is also proportional to velocity. A means is provided in this way for increasing or decreasing the effective total damping torque of the system.

If $A_d p\theta_o$ is the derivative feedback, with $A_d$ a constant with dimensions of time, then the differential equation of the system of Fig. 18-9 may be written by summing torques on the output shaft, as

$$Jp^2\theta_o + Fp\theta_o - K(\theta_i - \theta_o - A_d p\theta_o) = 0$$

which may be rearranged to

$$[Jp^2 + (F + KA_d)p + K]\theta_o = K\theta_i \tag{18-44}$$

It can be seen that the coefficient of the $p\theta_o$ term is $F + KA_d$ and this is the *effective* damping constant as a result of derivative feedback. Even

though the damping is increased, no additional energy is dissipated by the $KA_d$ term.

From Eq. (18-32) the damping ratio $\xi$ then becomes

$$\xi = \frac{F + KA_d}{2\sqrt{KJ}} \tag{18-45}$$

By this additional derivative feedback it is possible to obtain any desired damping ratio in a linear second-order system.

The inherent damping $F$ may actually be quite small; yet the value of $\xi$ can be sufficient to yield small overshoot on sudden position changes. The steady-state error on ramp or constant-velocity inputs is $F/K$ sec, and $F$ may be reduced to cut this error, but $\xi$ maintained to prevent oscillatory response to a step input by use of the derivative-feedback term $KA_d$.

**18-6. Error-rate Control.** Another method for providing control of the effective system damping is the *error-rate control* of Fig. 18-10. In

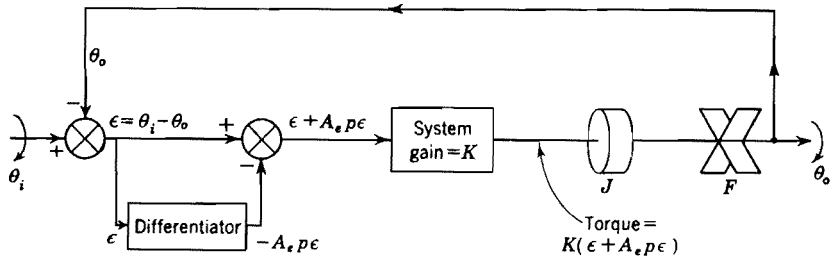



FIG. 18-10. Error-rate control system.

addition to the usual position feedback the error signal is differentiated and the resulting error rate is fed into the amplifier along with the error signal itself.

Summing the output shaft torques gives

$$Jp^2\theta_o + Fp\theta_o - K(\epsilon + A_e p\epsilon) = 0 \tag{18-46}$$

Since $\epsilon = \theta_i - \theta_o$, then

$$[Jp^2 + (F + KA_e)p + K]\theta_o = (1 + A_e p)K\theta_i \tag{18-47}$$

The effective viscous-damping ratio for the system is then

$$\xi = \frac{F + KA_e}{2\sqrt{KJ}} \tag{18-48}$$

and, by adjusting the error-rate coefficient $A_e$, any desired damping ratio can be obtained.

As was found for derivative control, such alteration of $\xi$ may be used to provide an opportunity to reduce $F$ to the vanishing point, thereby

eliminating the steady-state error for a ramp or constant-velocity input, still retaining the small overshoot of a well-damped system for a step-displacement input.

Due to the differences in the right side of Eqs. (18-44) and (18-47), there will be a difference in the response to step-input changes. The

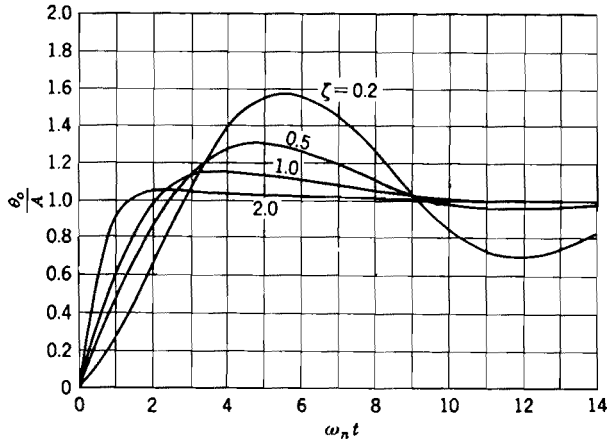



FIG. 18-11. Step response with error-rate damping.

term $A_e p\theta_i$ calls for an infinite impulse at the instant of step application. Physical systems cannot so respond, and there will always be less actual damping than the theoretical analysis indicates, owing to the finite output of realizable differentiators. This is demonstrated by the response curves of Fig. 18-11 for step displacement $\theta_i = A$, wherein some overshoot is always present, regardless of the effective value of $\xi$.

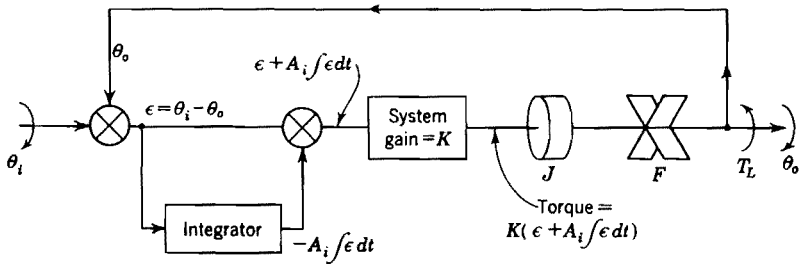



FIG. 18-12. Integral-control system.

**18-7. Integral Control.** Another interesting and useful control form is *integral control*, in which the time integral of the error is taken and introduced into the amplifier along with the error signal. This type of system is also called *reset control* in the process-control industry.

Figure 18-12 diagrams an integral-control system, to which has been added an opposing load torque $T_L$ on the output. Summation of the

torques at the output gives

$$Jp^2\theta_o + Fp\theta_o - K(\epsilon + A_i\int \epsilon\, dt) + T_L = 0$$

To illustrate another form of solution, the above may be rewritten in terms of error $\epsilon$ instead of output $\theta_o$. Then, by adding $-Jp^2\theta_o - Fp\theta_o$ to both sides, the equation may be written

$$\left(Jp^2 + Fp + K + \frac{KA_i}{p}\right)\epsilon = T_L + (Jp^2 + Fp)\theta_i \qquad (18\text{-}49)$$

The resulting equation is of third order, and the transient solution will have three roots, at least one or all three being real since complex roots always occur in pairs.

Of particular interest is the steady-state response to step amplitude and ramp (constant-velocity) inputs, and these may be studied by performing the division indicated by

$$\epsilon = \frac{T_L + (Jp^2 + Fp)\theta_i}{Jp^2 + Fp + K + KA_i/p} \qquad (18\text{-}50)$$

which leads to

$$\epsilon = \left[\frac{F}{KA_i}p^2 + \left(\frac{J}{KA_i} - \frac{F}{KA_i^2}\right)p^3 + \cdots\right]\theta_i + \left(\frac{1}{KA_i}p - \frac{1}{KA_i^2}p^2 + \cdots\right)T_L \qquad (18\text{-}51)$$

Since the input is $\theta_i = A$ for a step input and $\theta_i = \lambda t$ for a ramp, or constant-velocity, input, the above indicates that with integral control

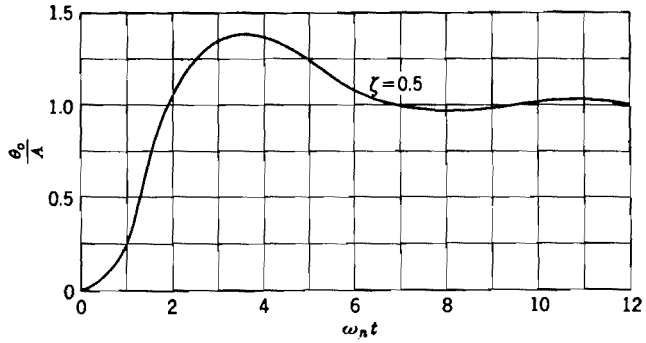



FIG. 18-13. Step response for integral control. $A_i/\omega_n = 0.2$.

the steady-state error will be zero, since the lowest derivative of $\theta_i$ appearing in the $\theta_i$ term in the error is the second.

It can also be seen from Eq. (18-51) that, if the torque load $T_L$ is constant, then no steady-state error will be introduced to carry this load.

The response of a typical integral control to a step-amplitude input is shown in Fig. 18-13. A combination of integral control, to provide zero steady-state error, and error rate or derivative feedback, to give damping sufficient for satisfactory transient response, is frequently used. When employed in this way, with several feedback loops, questions of stability must be investigated.

**18-8. Temperature Regulator with Integral Control.** The effect of integral control may be illustrated by its application to the temperature-control problem represented by Fig. 18-14, where a liquid is heated by

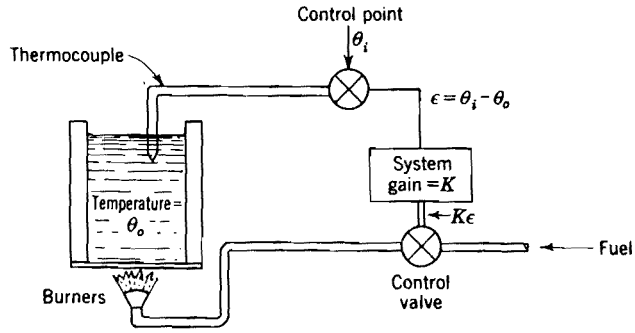



FIG. 18-14. Temperature regulator.

fuel admitted by the control valve. The heat input is made proportional to the error signal

$$H_i = K(\theta_i - \theta_o)$$

this system being known as *proportional control* in the process-control industry. The assumption of proportionality requires a constant fuel condition, and a linear relation between valve position and fuel flow for proper operation.

The output $\theta_o$ and the input $\theta_i$ must here be interpreted as actual and desired temperature. The fuel supplies the stored heat $S\, d\theta_o/dt$ whenever the temperature changes and makes up the conduction losses which are proportional to the difference between bath temperature and environmental temperature. The system equation then is

$$K(\theta_i - \theta_o) = Sp\theta_o + \lambda(\theta_o - \theta_e) \tag{18-52}$$

If $K/\lambda = B$, then

$$\frac{S/\lambda}{1+B}\, p\theta_o + \theta_o = \frac{B}{1+B}\, \theta_i + \frac{1}{1+B}\, \theta_e \tag{18-53}$$

This is in the form of Eq. (18-4) for a first-order system, and the transient response will be an exponential like Fig. 18-4, with a time constant $(S/\lambda)/(1 + B)$.

For steady state, with $\theta_i$ and $\theta_e$ constant, then $p\theta_o = 0$, and the temperature will be

$$\theta_{os} = \frac{B}{1 + B}\theta_i + \frac{1}{1 + B}\theta_e \tag{18-54}$$

Any difference between $\theta_o$ and $\theta_i$ may be removed in the calibration of the dial used to set $\theta_i$, if $B = K/\lambda$ remains fixed. However, any change in $K$, the heat input per unit temperature error, will cause a shift in the output temperature $\theta_{os}$. Likewise, while the effect of environmental-temperature change is reduced by the factor $1/(1 + B)$, yet it will still have some effect. If this is reduced by raising $K$, then other dynamic effects become appreciable and the system will become one of second order.

In order to reduce the steady-state error and the effects of changes in $K$, $\lambda$, or $\theta_e$, it is possible to add integral control, or *reset* as it is called by the process industry. The valve position and resultant heat input are made to change in accordance with

$$H_i = K(\theta_i - \theta_o) + K_1\int(\theta_i - \theta_o)\,dt$$

Introducing this for the left side of Eq. (18-52), and differentiating to remove the integral, leads to

$$\left(\frac{S}{\lambda B_1}p^2 + \frac{1 + B}{B_1}p + 1\right)\theta_o = \left(\frac{B}{B_1}p + 1\right)\theta_i + \frac{p}{B_1}\theta_e \tag{18-55}$$

where $B_1 = K_1/\lambda$.

This is of the form of Eq. (18-47), for a second-order system with error-rate control, and this system will perform in similar fashion with

$$\xi = \frac{\lambda + K}{2\sqrt{K K_1 S}} \tag{18-56}$$

The steady-state error will be zero and independent of parameter variations in $K$ and $\lambda$. Environmental changes will affect the transient but will not alter the steady-state operating temperature $\theta_{os}$. The addition of integral control has improved the performance considerably. In particular is this true because of changes in system load or environment, where without integral control the steady-state temperature will drift from the desired control point.

**18-9. Sinusoidal Response of a Second-order System.** The performance of the second-order system of Fig. 18-15, with position and derivative feedback, was derived in Eq. (18-44) as

$$[Jp^2 + (F + KA_d)p + K]\theta_o = K\theta_i \tag{18-57}$$

The undamped natural frequency of a second-order system was defined by Eq. (18-34) as

$$\omega_n = \sqrt{\frac{K}{J}}$$

while the damping ratio for this system is obtained from Eq. (18-45) as

$$\xi = \frac{F + KA_d}{2\sqrt{KJ}}$$

Using $\omega_n$ and $\xi$, Eq. (18-57) becomes

$$\left(\frac{1}{\omega_n{}^2}p^2 + \frac{2\xi}{\omega_n}p + 1\right)\theta_o = \theta_i \tag{18-58}$$

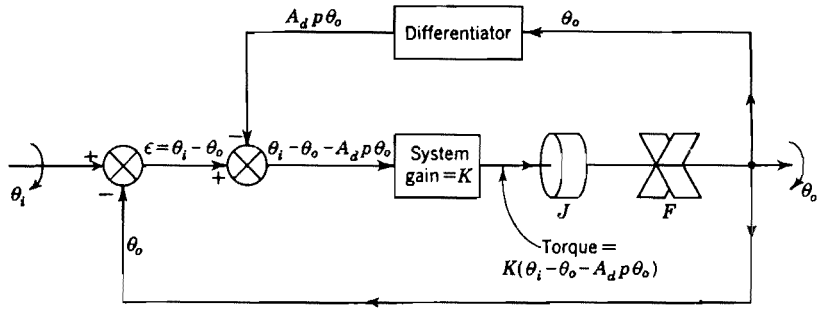



FIG. 18-15. Second-order derivative and position feedback control.

For a sinusoidal input given by $\theta_i = A_1 \cos \omega t = A_1 \operatorname{Re} \epsilon^{j\omega t}$, the steady-state response can be found by assuming $\theta_o = A_2 \operatorname{Re} \epsilon^{j(\omega t+\phi)}$ so that in the usual manner

$$\frac{\theta_o}{\theta_i} = \frac{1}{1 - (\omega/\omega_n)^2 + j2\xi\omega/\omega_n} = Y(j\omega) = M\epsilon^{j\phi} \tag{18-59}$$

where $M$ is the magnitude of $Y(j\omega)$, the closed-loop transfer function, and $\phi$ is its phase angle. From the above

$$M = \frac{1}{\sqrt{(1 - \omega^2/\omega_n{}^2)^2 + (2\xi\omega/\omega_n)^2}} \tag{18-60}$$

$$\phi = -\tan^{-1}\frac{2\xi\omega/\omega_n}{1 - \omega^2/\omega_n{}^2} \tag{18-61}$$

It is convenient to study $Y(j\omega)$ by plots of $M$ and $\phi$ as functions of $\omega/\omega_n$ and $\xi$, and these appear in Figs. 18-16 and 18-17.

Figure 18-16 indicates a resonant effect for $\omega/\omega_n \cong 1$. These curves are reminiscent of Fig. 5-18 for the compensated wideband vacuum-tube amplifier. The systems are indeed related since they are described by second-order equations of similar form.

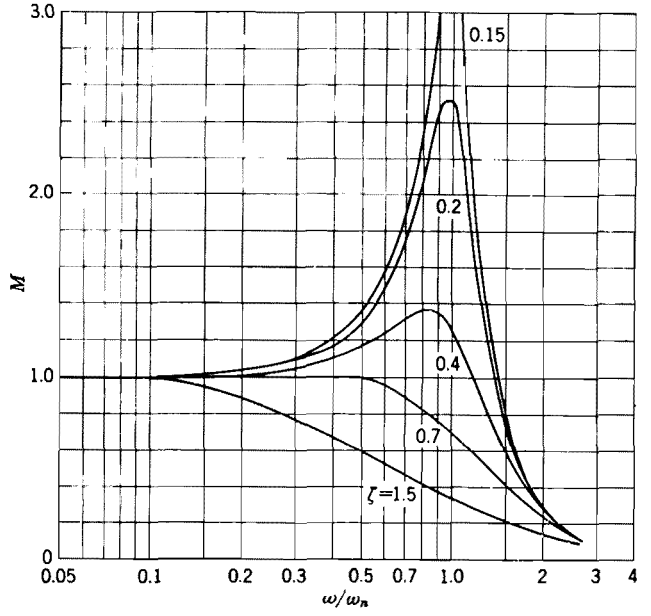



FIG. 18-16. Magnitude of $Y(j\omega)$.

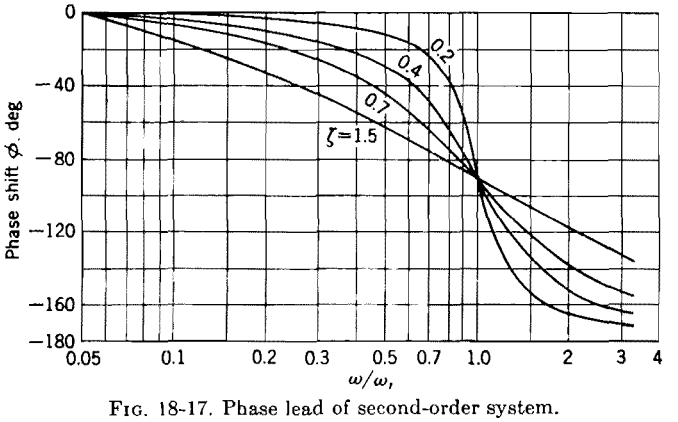



FIG. 18-17. Phase lead of second-order system.

The resonant peak occurs when the denominator of Eq. (18-60) goes through a minimum. Taking the derivative of the denominator with respect to $\omega$ and setting equal to zero gives

$$\omega_p = \omega_n \sqrt{1 - 2\xi^2} \tag{18-62}$$

as the frequency at which the peak response occurs. If $\xi > 0.707$, no maximum or overshoot appears in the response.

Using $\omega = \omega_p$ in Eq. (18-60) gives

$$M_p = \frac{1}{2\xi\sqrt{1 - \xi^2}} \tag{18-63}$$

as the peak value of the transfer-function magnitude.

From Fig. 18-17 it appears that the phase angle is $-90°$ for all values of $\xi$ at $\omega/\omega_n = 1$ and approaches $-180°$ for $\omega/\omega_n$ large.

**18-10. Transfer-function Analysis.** The application of each of the foregoing systems is general, in terms of types of apparatus, but idealized amplifiers of gain $K$ have been included and no means indicated for studying the effect of torque amplifiers of other than ideal characteristics. A method leading to more general results employs the operator known as the *transfer function,* to describe the effect of each subgroup of system elements on the over-all response. The transfer function is an expression for the ratio of output to input of a device or subsystem, in terms of the operator $p$, such that

$$\frac{\text{Output}}{\text{Input}} = \text{transfer function} = Y(p) \tag{18-64}$$

For an amplifier this is simply an expression for gain in terms of operator $p$.*

For instance, Fig. 18-10 shows the output torque of a control-system amplifier as

$$T_o = K(\epsilon + A_e p\epsilon) = K(1 + A_e p)\epsilon$$

The gain of this amplifier, in terms of torque developed per radian of error, is

$$\frac{T_o}{\epsilon} = K(1 + A_e p) = Y_1(p) \tag{18-65}$$

and this defines the amplifier transfer function $Y_1(p)$, including the differentiator as part of the amplifier.

The subsystem comprising the inertia $J$ and viscous damping $F$ may also be represented by a transfer function. The input to the subsystem is torque $T_o$, and the output is $\theta_o$. The equation of motion relating these quantities is obtained from Eq. (18-46) as

$$(Jp^2 + Fp)\theta_o = T_o$$

from which

$$\frac{\theta_o}{T_o} = \frac{1}{Jp^2 + Fp} = \frac{1/F}{p\left(\frac{J}{F}p + 1\right)} = Y_2(p) \tag{18-66}$$

* When the Laplace transform is employed, the transfer-function operator will be a function of $s$, as $Y(s)$, thus including the concept of complex frequency of Chap. 7.

Thus $Y_2$ is the transfer function of that portion of the system including the output inertia and viscous damping. Figure 18-10 can then be redrawn using transfer functions $Y_1$ and $Y_2$ as in Fig. 18-18*a*.

The method may be further illustrated by noting that the amplifier of the integral control of Fig. 18-12 has a transfer function given by

$$Y_1 = \frac{T_o}{\epsilon} = K(1 + A_i p^{-1}) \tag{18-67}$$

This amplifier is followed by an inertia and damping subsystem described by $Y_2$ of Eq. (18-66). Thus the block diagram of Fig. 18-18*a* will also serve for the integral control of Fig. 18-12 and by proper definition of $Y_1$ will in fact illustrate position, error rate, and integral control equally.

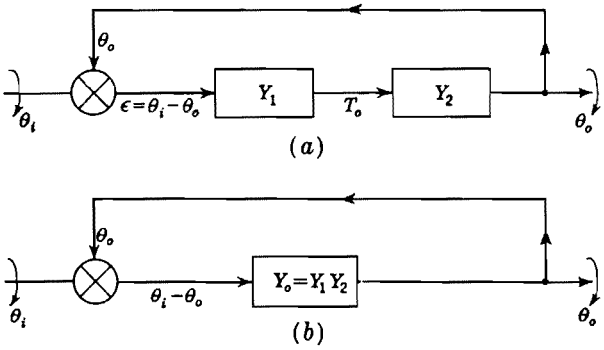



FIG. 18-18. Control system described by transfer functions.

If the position feedback loop of Fig. 18-18*a* is broken, there remains a system with an output $\theta_o$, and an input of some value $\epsilon$, although $\epsilon$ will not then represent $\theta_i - \theta_o$. From the transfer-function definition it is possible to write

$$\theta_o = Y_2 T_o = Y_1 Y_2 \epsilon = Y_o \epsilon \tag{18-68}$$

for Fig. 18-18. This serves to define

$$Y_o(p) = Y_1 Y_2 = \frac{\theta_o}{\epsilon} \tag{18-69}$$

as the *open-loop transfer function* of the system, or the transfer function which would be obtained if the feedback were removed.

In view of this definition of $Y_o$ a block diagram of any system with a single feedback loop can always be reduced to that of Fig. 18-18*b*. When the feedback loop is closed, then $\epsilon = \theta_i - \theta_o$, and so

$$\theta_o = Y_o \epsilon = Y_o(\theta_i - \theta_o)$$

$$\theta_o = \frac{Y_o}{1 + Y_o} \theta_i \tag{18-70}$$

The transfer function for a single-loop feedback system is then $Y_o/(1 + Y_o)$, in terms of the open-loop transfer function $Y_o$.

**18-11. Transfer Function for Multiple-loop Feedback.** Servomechanisms frequently include numerous elements with their individual transfer functions, and for stability reasons feedback may be used over portions of a system, as well as over-all. Such a possible system, expressed in terms of transfer functions, appears in Fig. 18-19.

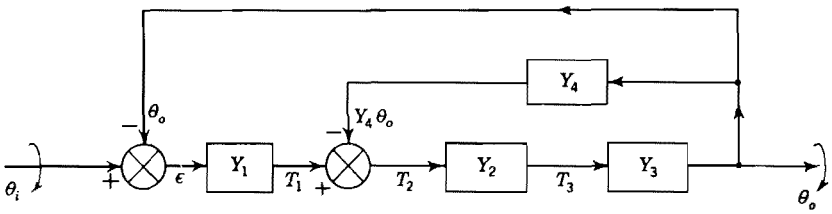



FIG. 18-19. Multiple-loop servomechanism.

After passing the signal through the first element of the system

$$T_1 = Y_1\epsilon = Y_1(\theta_i - \theta_o)$$

and

$$T_2 = Y_1(\theta_i - \theta_o) - Y_4\theta_o \tag{18-71}$$

as the input to the last two elements. Then

$$T_3 = Y_2[Y_1(\theta_i - \theta_o) - Y_4\theta_o]$$

and

$$\theta_o = Y_2Y_3[Y_1\theta_i - (Y_1 + Y_4)\theta_o] \tag{18-72}$$

The complete system then is described by a transfer function

$$\frac{\theta_o}{\theta_i} = \frac{Y_1Y_2Y_3}{1 + Y_1Y_2Y_3 + Y_2Y_3Y_4} \tag{18-73}$$

Other systems with multiple-loop feedback may be similarly analyzed by this step-by-step procedure.

**18-12. Transfer Functions of Several Circuits.** Because of the value of the transfer function in system analysis it is helpful to have available the transfer functions for certain commonly used circuits and system elements. Among these are the differentiating and integrating circuits of Fig. 18-20, previously discussed in Chap. 9.

From Fig. 18-20*b*

$$Ri + \frac{1}{C}\int_o^t i\,dt = e_i \tag{18-74}$$

Using $p = d/dt$, and noting that $e_o = Ri$, then

$$e_o = \frac{RCp}{1 + RCp}\,e_i$$

and the transfer function for the circuit of Fig. 18-20*b* is

$$Y(p) = \frac{e_o}{e_i} = \frac{RCp}{1 + RCp} \tag{18-75}$$

For $RCp \ll 1$ the system behaves as a differentiator of transfer function $RCp$. This circuit will later be referred to as a *lead network*, because of

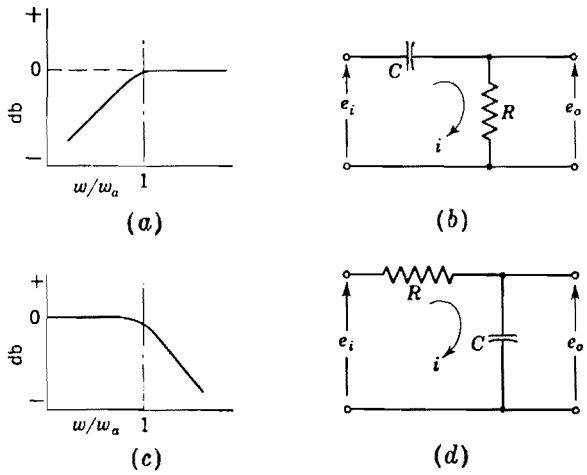



FIG. 18-20. (*a*, *b*) Differentiating, or lead, network. (*c*, *d*) Integrating, or lag, network.

its sinusoidal steady-state operation where $p = j\omega$ and $\omega_a = 1/RC$, so that

$$\frac{e_o}{e_i} = \frac{\omega/\omega_a}{\sqrt{1 + \omega^2/\omega_a^2}} \underline{/\phi_a} \tag{18-76}$$

where

$$\phi_a = \tan^{-1} \frac{1}{\omega/\omega_a} \tag{18-77}$$

and thus $e_o$ leads $e_i$ by the angle $\phi_a$.

For the circuit of Fig. 18-20*d*, again

$$Ri + \frac{1}{C} \int_o^t i \, dt = e_i$$

but $e_o = i/Cp$ so that

$$e_o = \frac{1}{1 + RCp} e_i \tag{18-78}$$

and the transfer function for Fig. 18-20*d* is

$$Y(p) = \frac{e_o}{e_i} = \frac{1}{1 + RCp} \tag{18-79}$$

For $RCp \gg 1$ the system operates as an integrator with transfer function $1/RCp$. It is also known as a *lag network* because of its sinusoidal operation, where $p = j\omega$, $\omega_a = 1/RC$, and

$$\frac{e_o}{e_i} = \frac{1}{\sqrt{1 + \omega^2/\omega_a{}^2}} \underline{/\phi_b} \tag{18-80}$$

$$\phi_b = -\tan^{-1}\frac{\omega}{\omega_a} \tag{18-81}$$

and thus $e_o$ lags $e_i$ by the angle $\phi_b$.

Two additional circuits serving as lead or lag networks are shown in Fig. 18-21. These were previously discussed in Sec. 5-4 as circuits suited

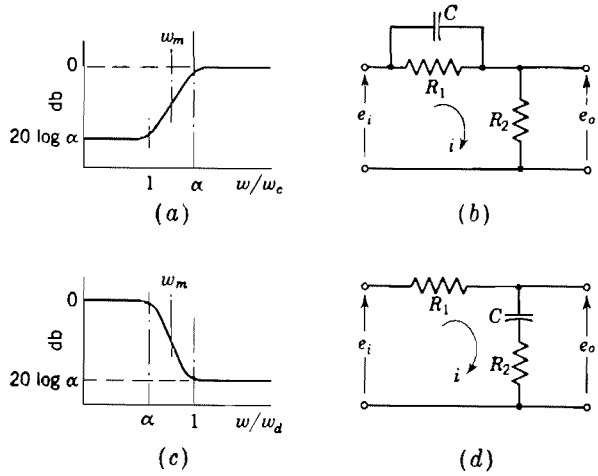



FIG. 18-21. (*a*, *b*) Lead network. (*c*, *d*) Lag network.

to the alteration of phase and magnitude in the frequency domain. From that section, and writing $j\omega = p$, the transfer function for Fig. 18-21*b* is

$$Y(p) = \frac{e_o}{e_i} = \alpha \frac{1 + R_1Cp}{1 + \alpha R_1Cp} \tag{18-82}$$

where

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

The circuit operates as a lead network with a sinusoidal response as indicated in Fig. 18-21*a*, where $\omega_c = 1/R_1C$, and the point of maximum phase shift is $\omega_m$, given by

$$\omega_m = \frac{\omega_c}{\sqrt{\alpha}}$$

The transfer function for Fig. 18-21*d* is

$$Y(p) = \frac{e_o}{e_i} = \frac{1 + R_2Cp}{1 + R_2Cp/\alpha} \tag{18-83}$$

where again $\alpha = R_2/(R_1 + R_2)$. The circuit operates as a lag network in sinusoidal response as indicated in Fig. 18-21*c*, where $\omega_d = 1/R_2C$, and the point of maximum phase shift $\omega_m$ is given by

$$\omega_m = \omega_d \sqrt{\alpha}$$

The lead network of Fig. 18-21*b* is usually preferred over that of Fig. 18-20*b*, because the latter has a transfer function which becomes zero at $\omega = 0$. This would lead to steady-state error. The circuits of Fig. 18-21 also provide some additional flexibility in phase angle since the angle of maximum lead or lag can be chosen by the designer.

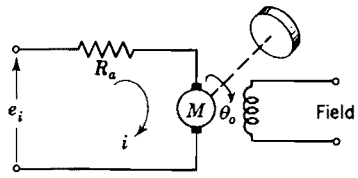



FIG. 18-22. Direct-current motor circuit.

**18-13. Transfer Functions of Several Physical Elements.** A d-c motor with separate field excitation is frequently used to provide the output torque of a servo system. Therefore the transfer characteristic of the motor must be included in the transfer characteristic of the torque amplifier. The inertia and damping of the motor should be included as part of the load elements.

The equation for the motor of Fig. 18-22 is

$$e_i = R_a i + C_n p\theta_o \tag{18-84}$$

where $R_a$ is the armature resistance and $C_n$ is the generated voltage constant in volts per radian per second.

The torque output of the motor is proportional to current and

$$\text{Torque} = C_T i = Jp^2\theta_o \tag{18-85}$$

with $C_T$ the torque per ampere of armature current. The moment of inertia is that of the armature and connected load, friction damping being neglected since it is small in such motors. Using Eq. (18-85) in Eq. (18-84), then

$$Y = \frac{\theta_o}{e_i} = \frac{1/C_n}{p\left(\dfrac{R_aJ}{C_TC_n}\,p + 1\right)} = \frac{1/C_n}{p(T_mp + 1)} \tag{18-86}$$

and this result is identical in form to that of Eq. (18-66) for a system with inertia and viscous damping. The viscous-damping effect is provided by the counter emf of the motor.

Most two-phase a-c motors have transfer functions of this same form.

In many cases the position feedback is handled by a synchrotransmitter, and the subtraction of output position from input position is

carried out in a synchrodifferential as in Fig. 18-23. The transmitters employ a single-phase rotor with a three-phase stator. The three-phase voltages induced in the stator are then representative of the angular position of the rotor. These voltages from the input synchro are connected to the stator windings of the synchrodifferential unit, the three-phase voltages from the output transmitter being connected to the three-phase rotor windings of the differential. The rotor of the differential

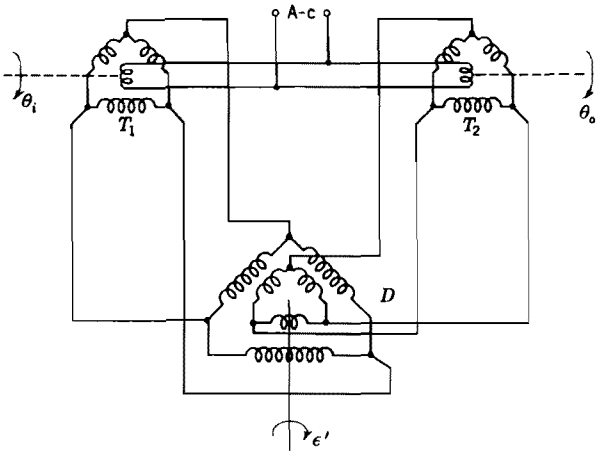



FIG. 18-23. Synchrodifferential system.

synchro provides a rotation equal to the error $\theta_i - \theta_o$ if a negligible torque load is applied.

The transfer function of the complete synchrodifferential system can be found as

$$\frac{\epsilon}{\epsilon'} = \frac{1}{\bar{p}^2/\omega_n + (2\xi/\omega_n)\bar{p} + 1} \tag{18-87}$$

where $\epsilon$ is the true error $\theta_i - \theta_o$ and $\epsilon'$ is the actual output indication of the synchro. It is usually necessary to add additional damping to prevent excessive overshoot of the synchro system.

The hydraulic torque amplifier of Fig. 18-24 has advantages over the electric motor, giving a much wider effective range of speed and much higher ratios of torque output to system inertia. Fluid is admitted to chamber $A$ or $B$ of the actuating piston by very small motion of the pilot valve as the balls are displaced from their "neutral" position by the input signal. As the piston moves, the pilot-valve casing is repositioned to restore the pilot valve to neutral. Thus every output motion imparted by the signal provides a measured and controlled motion of the driving piston, with speed of piston motion approximately proportional to pilot-valve displacement and actuating motion $\epsilon_s$. The transfer func-

tion of such a device is found to be

$$\frac{\epsilon_s}{d} = \frac{K}{p(T_a p + 1)} \tag{18-88}$$

where $d$ is the piston displacement, $K$ is the gain magnitude, and $T_a$ appears as the time constant of the torque amplifier. This is identical in form with that of the d-c motor in Eq. (18-86).

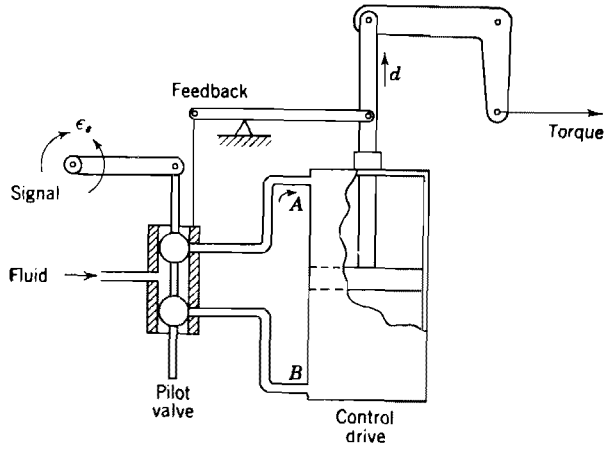



FIG. 18-24. Hydraulic, or air-operated, valve.

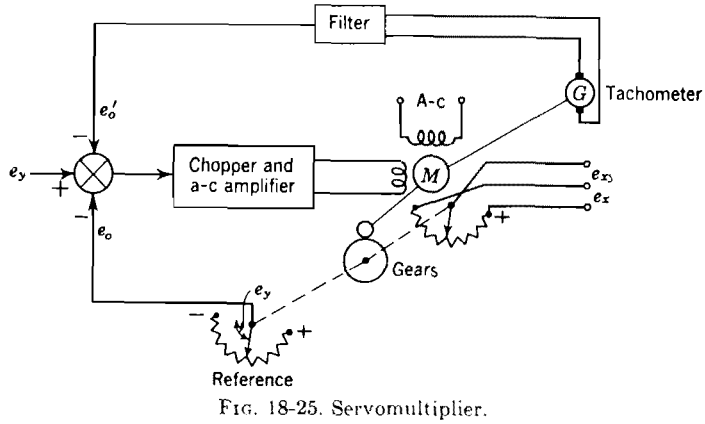



FIG. 18-25. Servomultiplier.

**18-14. A-C Motor Servo System.** In Chap. 8 there was some discussion of servomultipliers for analog computers, to produce a product voltage

$$e_{xy} = e_x e_y$$

and an example of such a servomechanism is shown in Fig. 18-25. An input voltage $e_y$ causes motor $M$ to drive a slide-wire potentiometer to

such a position that the voltage picked off equals $e_y$. Mechanically tied to this potentiometer is another whose total voltage is $e_x$. The position of the slider on the second potentiometer is representative of $e_y$, and the voltage picked off by this slider is $e_x e_y = e_{xy}$ as desired.

The control system employs position feedback through the reference potentiometer and derivative feedback from a tachometer generator. A high-pass filter is added to reduce the steady-state error on ramp, or constant-velocity, inputs.

A constant-velocity input would cause a nonvarying or zero-frequency tachometer output, and this is eliminated by the high-pass filter of the type of Fig. 18-20*b*. This action removes the steady-state error inherent in derivative feedback yet leaves the derivative feedback effective at

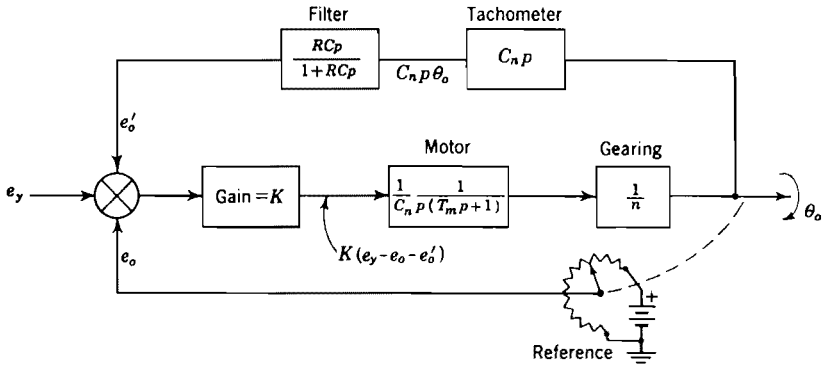



FIG. 18-26. Block diagram of servomultiplier.

higher frequencies for furnishing increased damping. The cutoff frequency of the filter should be considerably below the value of $\omega_n$ of the system.

The d-c input signal $e_y - e_o - e_o'$ is chopped by a 60-cycle vibrator and amplifier to supply one field of a two-phase motor. The other phase is supplied directly from the line. The polarity of the amplifier input determines the direction of motor rotation. Figure 18-26 shows the system redrawn, with transfer functions included for each element. It is then easy to write the system transfer function, using the methods of Sec. 18-10.

**18-15. Stability.** Since most practical servomechanisms are described by differential equations of second order or higher, generally with some complex roots, the possibility of instability of these systems needs to be examined.

From Eq. (18-70)

$$(1 + Y_o)\theta_o = Y_o\theta_i$$

and for steady-state sinusoidal input the output may be assumed $\theta_o = \epsilon^{j\omega t}$ and $Y_o$ transformed from a function of $p$ to a function of $j\omega$, since $p^n\theta_o = (j\omega)^n\theta_o$ under that assumption. If Eq. (18-70) is then written

$$\frac{\theta_o}{\theta_i} = \frac{\mathbf{Y}_o(j\omega)}{1 + \mathbf{Y}_o(j\omega)} \tag{18-89}$$

it reminds us of the feedback-amplifier transfer relation of Chap. 7,

$$\mathbf{A}' = \frac{\mathbf{A}}{1 - \mathbf{A}\beta} \tag{18-90}$$

The two equations are actually of the same form, the positive sign in the denominator of Eq. (18-89) being present because feedback in servomechanisms is assumed inherently negative by reason of $\epsilon = \theta_i - \theta_o$, whereas the sign of $\mathbf{A}\beta$ in the amplifier expression is determined by the angle of $\mathbf{A}\beta$. Essentially, in the servomechanism, analysis starts with an error $\epsilon = \theta_i - \theta_o$ as the input, placing the subtraction outside the analysis, while in the feedback amplifier the analysis starts with an external signal, the subtraction of output from input occurring within the analysis. This difference is merely a matter of custom but accounts for the difference in sign in the denominators of Eqs. (18-89) and (18-90).

Inspection of Eq. (18-89) shows that, if the sinusoidal open-loop transfer function $\mathbf{Y}_o(j\omega)$ approaches a value of $-1 + j0$, the output $\theta_o$ will rise and tend toward infinity, a condition of instability. In fact, if the locus of $\mathbf{Y}_o(j\omega)$ for all $\omega$'s encircles the point $-1, j0$, then oscillation will occur. This result is based on the work of Nyquist and differs from the discussion of Chap. 7 only in that the critical point is moved from the right half plane to the left, because of the change from $-$ to $+$ for $\mathbf{A}\beta$ and $\mathbf{Y}_o(j\omega)$.

The desirability of basing the analysis on the open-loop sinusoidal transfer function $\mathbf{Y}_o(j\omega)$, instead of a closed-loop function, is apparent when it is realized that it may be more easily calculated or measured than the closed-loop function and that in the case of an unstable system the closed-loop function cannot be measured at all.

As an example of an open-loop response plot, similar to the $A\beta$ plot of Chap. 7, take the simple position control of Fig. 18-7, which has an open-loop equation

$$Jp^2\theta_o + Fp\theta_o - K\epsilon = 0$$

from which the open-loop transfer function $Y_o(p) = \theta_o/\epsilon$ is

$$Y_o(p) = \frac{K/F}{p[(J/F)p + 1]} \tag{18-91}$$

For the sinusoidal steady-state $p = j\omega$, and

$$\mathbf{Y}_o(j\omega) = \frac{KJ/F^2}{(j\omega J/F)(j\omega J/F + 1)} = \frac{K\tau/F}{j\omega\tau(j\omega\tau + 1)} \tag{18-92}$$

as the open-loop function, where $\tau = J/F$, a time constant. It may be noted that the numerator of Eq. (18-92) is a constant and proportional

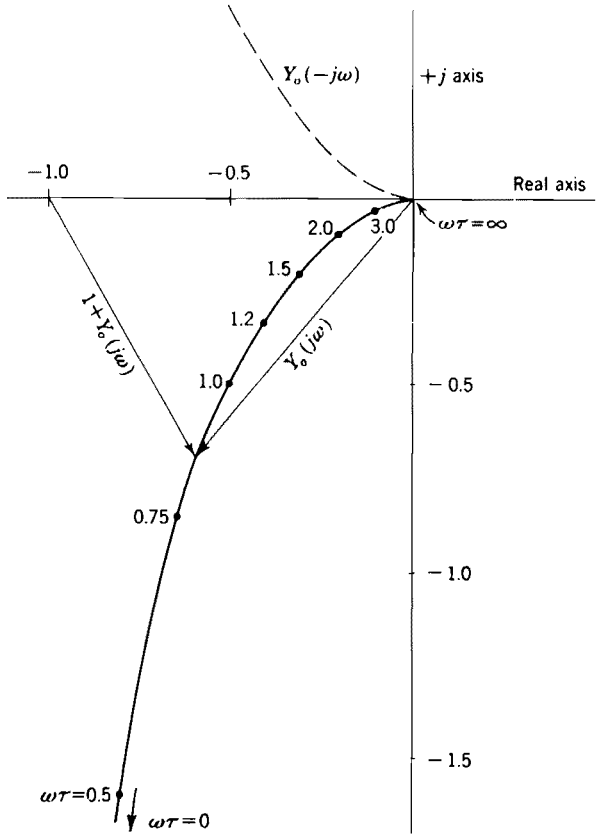



FIG. 18-27. $Y_o(j\omega)$ plot for Eq. (18-92).

to gain for a particular system. Those factors whose magnitude and angle depend on frequency appear in the denominator, and these may be written as $\mathbf{G}(j\omega)$, with $\mathbf{K}'$ employed for the factors independent of frequency, so that

$$\mathbf{Y}_o(j\omega) = \mathbf{K}'\mathbf{G}(j\omega) \tag{18-93}$$

Equation (18-92) is plotted in Fig. 18-27 as a function of $\omega\tau$, with the effective gain $\mathbf{K}' = K\tau/F = 1$. It is apparent that any such system

will be stable, increases in gain $\mathbf{K}'$ serving only to expand the diagram, and cannot make the locus encircle the point $-1, j0$.

Since the curve is the locus of $\mathbf{Y}_o(j\omega)$, then by graphically subtracting $-1$ from any $\mathbf{Y}_o(j\omega)$ vector a line for $1 + \mathbf{Y}_o(j\omega)$ can be obtained. This further illustrates the conclusion that the locus must pass through or encircle the critical point to produce instability, since only in that way can $\mathbf{Y}_o(j\omega) = -1 + j0$, or make the vector $1 + \mathbf{Y}_o(j\omega) = 0$.

Resonant peaks can cause the $\mathbf{Y}_o(j\omega)$ locus to approach the $-1, j0$ point at some frequency, the effect being to give a prolonged oscillatory transient error for a sudden change in input.

**18-16. Stability Criteria.** In the design of systems it is usual to develop first the configuration for $\mathbf{G}(j\omega)$, the factor determining the shape of the $\mathbf{Y}_o$ locus. This leaves undetermined the magnitude of the gain to be used, and this is of importance since the closeness of approach of the $\mathbf{Y}_o(j\omega)$ locus to the critical point is a measure of both the relative stability of the system and the speed of its step response. That is, it has been indicated by previous curves that the step response will be rapid but not oscillatory for systems where the sinusoidal response shows only a slight amount of peaking. The peaking effect indicates a close approach of the open-loop locus to the critical point, and a correlation between the open-loop locus and the closed-loop gain amplitude is helpful.

This correlation is also of aid in estimating the performance of higher-order systems which have not been discussed here. In general, if the sinusoidal response of such systems is fairly flat, with only slight peaking at some frequency beyond which the response falls, then satisfactory transient response will be obtained.

Correlation between the amplitude of the open-loop and closed-loop responses is possible directly from the Nyquist plot of the open-loop function $\mathbf{Y}_o(j\omega)$. From Eq. (18-70) the magnitude of the closed-loop relation is

$$\mathbf{Y}(j\omega) = \frac{\theta_o}{\theta_i} = \frac{\mathbf{Y}_o(j\omega)}{1 + \mathbf{Y}_o(j\omega)} = \frac{x + jy}{1 + x + jy} \tag{18-94}$$

and if $M$ is the amplitude and $\phi$ the phase angle of the above as in Sec. 18-9, then

$$M^2 = \frac{x^2 + y^2}{(1 + x)^2 + y^2} \tag{18-95}$$

and

$$M^2 - 1 = \frac{x^2 + y^2}{(1 + x)^2 + y^2} - 1 = -\frac{1 + 2x}{(1 + x)^2 + y^2} \tag{18-96}$$

From the above

$$\frac{M^2}{M^2 - 1} = -\frac{x^2 + y^2}{1 + 2x}$$

$$x + \frac{M^2}{M^2 - 1} = \frac{x^2 + x - y^2}{1 + 2x}$$

and finally

$$y^2 + \left(x + \frac{M^2}{M^2 - 1}\right)^2 = \left(\frac{M^2}{M^2 - 1}\right)^2 \tag{18-97}$$

representing a family of circles in the $Y_o$ plane whose contours are of constant magnitude $M$ for the closed-loop transfer function $Y(j\omega)$. The radii of the circles are

$$r = \frac{M^2}{M^2 - 1}$$

and the centers at the points

$$-\frac{M^2}{M^2 - 1} \qquad 0$$

For $M$ very large, as for an extreme resonant peak, the radius of the circle approaches zero, and the center approaches $-1$, 0. The locus for

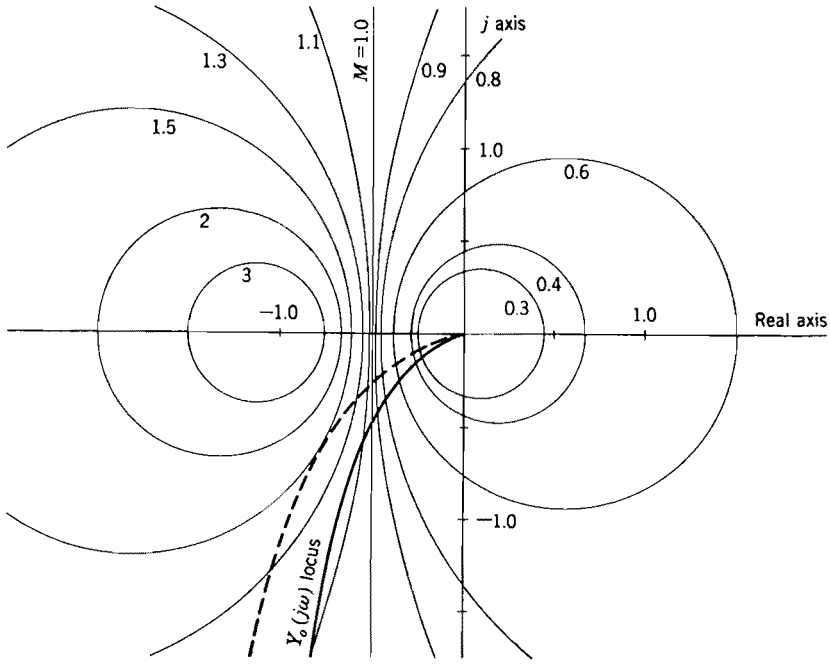



FIG. 18-28. Constant-$M$ circles.

$M = 1$ is a straight line perpendicular to the real axis at $x = -\frac{1}{2}$. Circles for $M > 1$ have centers lying on the negative real axis; those for $M < 1$ have centers on the positive real axis.

The $\mathbf{Y}_o(j\omega)$ locus of Fig. 18-27, with $K' = 1$, is superimposed on such a circle family in Fig. 18-28. As replotted, the locus comes approximately

tangent to a circle for $M = 1.2$ at a frequency given by $\omega\tau \cong 0.85$. This indicates a peak value of $M_p$ of 1.2 for the closed-loop transfer function $\mathbf{Y}(j\omega)$.

Equation (18-93) may be used to calculate the sinusoidal closed-loop transfer function with $K' = 1$ by

$$\mathbf{Y}(j\omega) = \frac{\theta_o}{\theta_i} = \frac{\mathbf{K'G}(j\omega)}{1 + \mathbf{K'G}(j\omega)} \tag{18-98}$$

resulting in Fig. 18-29*a*, confirming that $M_p = 1.2$.

If an increased gain peak is considered desirable to improve the transient response, the $\mathbf{Y}_o(j\omega)$ locus might be expanded by lengthening each $\mathbf{Y}_o$ vector by a constant factor, keeping its angle constant, to produce

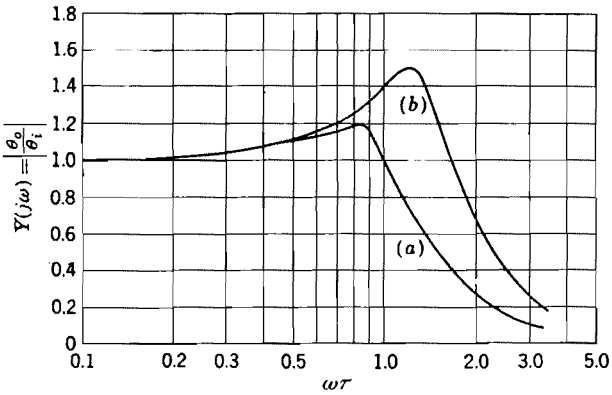



FIG. 18-29. Closed-loop gain magnitude (*a*) for solid curve of Fig. 18-28 with $K' = 1$, (*b*) for dashed curve with $K' = 2$.

the dashed locus as shown in Fig. 18-28. This would increase the amplitude at the peak to $M_p = 1.5$, at a frequency given by $\omega\tau \cong 1.2$. By Eq. (18-63) it is found that $\xi = 0.356$ for $M_p = 1.5$, and if it is noted that $K' = K_T/F = KJ/F^2 = (\frac{1}{2}\xi)^2$, it may be reasoned that, for $M_p = 1.5$, then $K' = 1.98$, and this is the value of gain required for the increased amplitude peak. The closed-loop transfer function as calculated by Eq. (18-98) with $K' = 1.98$ appears in Fig. 18-29*b*, confirming that $M_p = 1.5$.

An approximate means of determining stability directly from open-loop sinusoidal gain and phase-angle data is available in plots of decibel magnitude and of phase angle as functions of logarithmic frequency, using the corner-frequency plotting techniques of Chap. 5. Choosing the system described by Eq. (18-92) as an example,

$$\mathbf{Y}_o(j\omega) = \mathbf{K'G}(j\omega) = \frac{K'}{j\omega\tau(j\omega\tau + 1)} \tag{18-99}$$

The magnitude in decibels is

$$\text{db} = 20 \log \frac{1}{\omega\tau} + 20 \log \left| \frac{1}{\omega\tau + 1} \right|$$
$$= -20 \log (\omega\tau) - 20 \log \sqrt{\omega^2\tau^2 + 1} \qquad (18\text{-}100)$$

so that a curve of decibels against $\omega\tau$ may be plotted for each factor separately and the ordinates added to give Fig. 18-30.

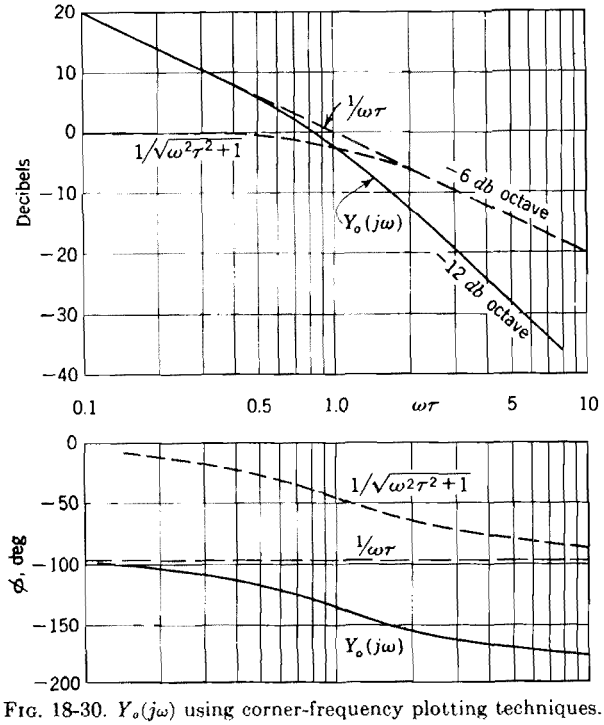



FIG. 18-30. $Y_o(j\omega)$ using corner-frequency plotting techniques.

It can be seen that the first factor of Eq. (18-99) gives a gain reduction of $-6$ db per octave. It also produces a constant phase angle of $-90°$. At the higher frequencies where $\omega\tau \gg 1$ the second term also produces a slope of $-6$ db per octave and a phase angle of $-90°$. The resultant of both factors at $\omega\tau \gg 1$ is a slope of $-12$ db per octave and a phase angle approximating $-180°$. Terms of quadratic form will, of course, introduce bumps, or peaks.

It is possible to conclude that when a decibel magnitude curve has a slope of $-6$ db per octave there will be associated with it a phase angle of $-90°$, when the slope is $-12$ db per octave the phase angle will be approximately $-180°$, and so on.

From Fig. 18-27 it may be noted that if, at the frequency of unity gain, the phase angle is less than $-180°$, then the Nyquist plot will not encircle the $-1, j0$ point and the system is stable. From the log plot of Fig. 18-30 it is possible to estimate the phase angle at unity gain (zero decibels) since if the slope equals or exceeds $-12$ db per octave the phase angle at that point will equal or exceed $-180°$. Thus asymptotes drawn to the decibel plot at the zero-decibel gain point help to give a quick appraisal of the likelihood of instability.

Since an experimental plot cannot be made for an unstable system, then this method serves largely in design and really allows an estimate of the closeness with which instability is approached or exceeded. In design synthesis of a control system it is desirable that its open-loop attenuation vs. frequency plot not fall off as fast as $-12$ db per octave in the region around the zero-decibel gain point. In fact it is usually desirable that the slope not exceed $-6$ db per octave in that region. Such a system will be stable in closed loop, and if the $-6$ db per octave slope continues well beyond the zero-decibel gain point, the step response will approximate an exponential and the system will approximate one of first order.

**18-17. Gain and Phase Margin.** A better means of determining the closeness of approach to instability is available in the Nyquist plot of an open-loop sinusoidal transfer function, as shown in Fig. 18-31.

The *gain margin* is defined as the magnitude of the $1 + \mathbf{Y}_o(j\omega)$ locus when the phase shift is $-180°$ (along the negative real axis). This then represents the decibel rise in gain which may be tolerated without oscillation and is indicated in Fig. 18-31.

The *phase margin* is $180°$ minus the angle of $\mathbf{Y}_o(j\omega)$ at the frequency for which $|Y_o(j\omega)|$ is unity. This is found in Fig. 18-31 at the point at which the locus crosses the unit circle.

Recommended values for phase and gain margin vary with the system and with the application. If the slope of the decibel attenuation plot is too negative in a given region, then compensating or shaping circuits of the form of the networks of Sec. 18-11 must be included in order to reduce the slope and phase angle to an acceptable value. Essentially this is accomplished by these circuits by changing the gain and phase angle at particular frequencies of interest.

It is usually possible to include such lead or lag networks directly in the amplifiers, and thus the analysis may be altered only by the inclusion of another transfer function. It should be noted, however, that a lead network will reduce the zero-frequency gain by a factor $\alpha$, and the internal gain must be increased by this amount if the steady-state error is to be kept small. A lag network will reduce the high-frequency gain by a similar factor.

**18-18. Example: Further Nyquist Plots.** The use of the Nyquist plot for stability analysis may be further demonstrated by use of the derivative and position control of Fig. 18-9. The effect of introducing the derivative-feedback loop may be found by considering the diagram of Fig. 18-32 as that of an open-loop system, thus incorporating the position

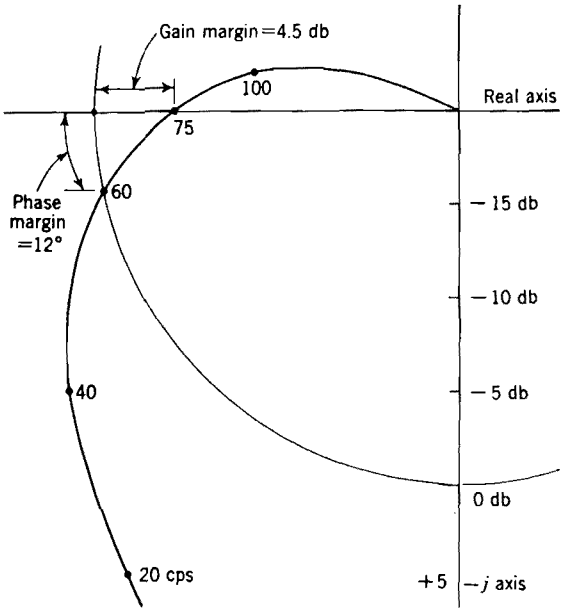



FIG. 18-31. Phase and gain margin.

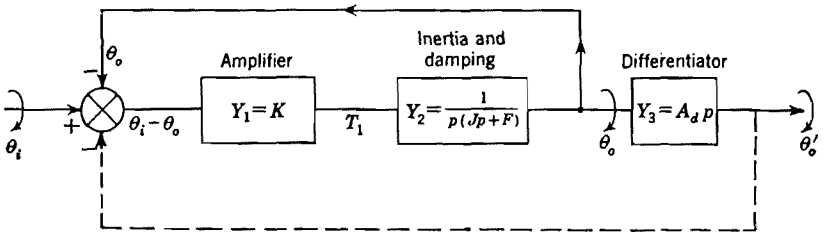



FIG. 18-32. Derivative control as an open loop.

feedback as an internal part of the system. This makes $\theta_i$ the input and $\theta_o'$ the output of the open loop. When the loop is finally closed as indicated by the dashed line, a system equivalent to Fig. 18-9 will be obtained.

The open-loop function, before closure by the dashed line, is obtainable from

$$\theta_o' = Y_1 Y_2 Y_3(\theta_i - \theta_o) \tag{18-101}$$

but

$$\theta_o = \frac{\theta_o'}{Y_3}$$

so that $$Y_o(p) = \frac{\theta_o'}{\theta_i} = \frac{Y_1Y_2Y_3}{1 + Y_1Y_2} = \frac{Y_3}{1/Y_1Y_2 + 1} \tag{18-102}$$

Inserting the individual transfer functions,

$$Y_o(p) = \frac{A_d p}{(J/K)p^2 + (F/K)p + 1}$$

and by use of

$$\omega_n = \sqrt{\frac{K}{J}} \qquad \xi = \frac{F}{2\sqrt{KJ}}$$

this becomes

$$Y_o(p) = \frac{A_d p}{p^2/\omega_n^2 + 2\xi p/\omega_n + 1} \tag{18-103}$$

Translating this to sinusoidal steady state by use of $p = j\omega$,

$$\mathbf{Y}_o(j\omega) = \frac{j\omega A_d}{1 - (\omega/\omega_n)^2 + j(2\xi\omega/\omega_n)} \tag{18-104}$$

For an underdamped case let $\xi = 0.3$, so that after altering the numerator slightly

$$\mathbf{Y}_o(j\omega) = \mathbf{K}'\mathbf{G}(j\omega) = \frac{j(\omega/\omega_n)(\omega_n A_d)}{1 - (\omega/\omega_n)^2 + j0.6(\omega/\omega_n)}$$

where the open-loop gain is

$$\mathbf{K}' = \omega_n A_d \tag{18-105}$$

and the function varying with frequency is

$$\mathbf{G}(j\omega) = \frac{j(\omega/\omega_n)}{1 - (\omega/\omega_n)^2 + j0.6(\omega/\omega_n)} \tag{18-106}$$

The closed-loop transfer function is

$$\mathbf{Y}(j\omega) = \frac{\mathbf{K}'\mathbf{G}(j\omega)}{1 + \mathbf{K}'\mathbf{G}(j\omega)} = \frac{\mathbf{G}(j\omega)}{1/\mathbf{K}' + \mathbf{G}(j\omega)} \tag{18-107}$$

In this form the critical point is at $-1/\mathbf{K}'$, $j0$ for a $\mathbf{G}(j\omega)$ plot, and this allows a study of the effect of open-loop gain on stability. Only one $\mathbf{G}(j\omega)$ plot is needed and the critical point moves over the plot on the real axis, in accordance with the effective gain $\mathbf{K}'$.

The function $G(j\omega)$ may be plotted in the complex plane as in Fig. 18-33. For $\omega/\omega_n = 0$ the curve lies at the origin, increasing in a clockwise direction with $\omega/\omega_n$. It follows the circular path, crossing the real axis at 1.67 for $\omega/\omega_n = 1$, returning toward the origin for large values of $\omega/\omega_n$, and reaching the origin for $\omega/\omega_n = \infty$. For negative values of $\omega$ the curve leaves the origin at $-\infty$ and retraces the circle in a counter-

clockwise direction. The complete plot for all $\omega$'s from $\infty$ to $-\infty$ produces two revolutions of the circle.

With the critical point at $-1/K' + j0$, or at $-1/\omega_n A_d + j0$ as indicated above, the system may be stable or unstable depending on the value of $A_d$. Of course, if $A_d$ is zero or positive, the curve cannot encircle a critical point, which then will lie on the negative real axis, and the system is stable. However, if $A_d$ becomes negative, the critical point will move to the positive real axis and stability is dependent on the magnitude of $A_d$. For $-0.6/\omega_n < A_d < 0$ the critical point lies outside the

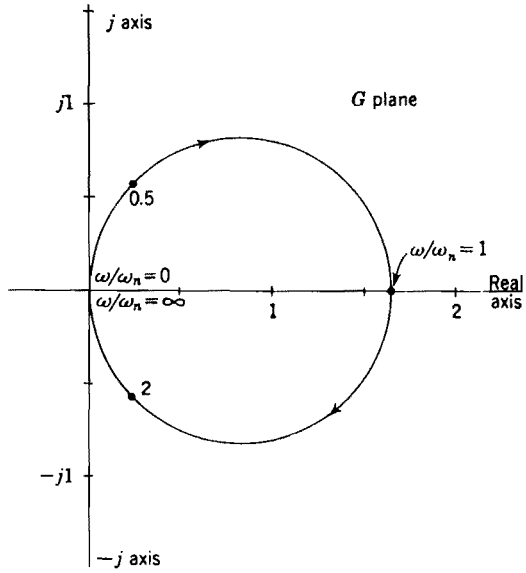



FIG. 18-33. $G(j\omega)$ plot.

circle, and the system is stable. For $A_d < -0.6/\omega_n$ the critical point will move inside the circle and will be encircled twice and the system will be unstable.

In general it can be seen that the critical point is determined by the damping. That is, for $A_d > -2\xi/\omega_n$ the system will be stable, whereas for $A_d < -2\xi/\omega_n$ the system will be unstable. For $A_d = 2\xi/\omega_n$ the system will oscillate after any initial disturbance so that this point also may be considered unstable.

**18-19. Other Systems.** The systems discussed, while typical of the fundamental principles, are somewhat less complicated than are frequently encountered in practice. Higher-order systems, with their more complicated Nyquist plots and stability questions, have not been discussed, nor has the whole class of servomechanisms operating with modulated a-c

signals. Since the literature on servomechanisms is so comprehensive, it seemed desirable here to restrict the discussion to fundamental principles, leaving the discussion of more elaborate systems to specialized texts on the subject.

## PROBLEMS

**18-1.** The equation of motion of a system having negligible inertia is

$$a_1 p\theta + a_o\theta = f(t)$$

Solve for the response to a step-function input where $f(t) = 0$ for $t < 0$, $f(t) = A$ for $t > 0$. Assume $\theta = 0$ at $t = 0$, and plot the result in terms of $\theta/A$ as a function of $t$.

**18-2.** A second-order system has

$$mp^2 s + cps + ks = f(t)$$

with $m = 2$ lb weight, $k = 4$ lb per ft, $c = 0.5$ lb per ft. Find the natural frequency, damping factor, and value of $c$ needed for critical damping. Also find the time needed to the first maximum of the response if $\xi = 0.3$.

**18-3.** A servomechanism has a transfer function

$$\mathbf{Y}_o(j\omega) = \frac{K'}{j\omega\tau(3j\omega\tau + 1)(j\omega\tau + 1)}$$

Using $M$ circles, determine the maximum value of $K'$ which is allowable if the gain curve is to have a peak value of 1.4. What is the value of $\omega\tau$ at this peak?

**18-4.** Steady-state frequency response of a system obtained on open loop is as in

| *Frequency* | *Output, db* | *Frequency* | *Output, db* |
|---|---|---|---|
| 0.5 | 21.5 | 7.0 | −10.0 |
| 1.0 | 15.0 | 10 | −16.0 |
| 2.0 | 8.0 | 15 | −23.5 |
| 3.0 | 2.5 | 20.0 | −28.0 |
| 5.0 | −5.0 | 30 | −35.5 |

the accompanying table. What reduction of gain is necessary to give a gain margin of 10 db?

**18-5.** Make an open-loop plot for

$$\mathbf{Y}_o(j\omega) = \frac{4}{j\omega(1 + 0.04j\omega)}$$

and determine whether the system is stable.

**18-6.** Make an open-loop plot for

$$\mathbf{Y}_o(j\omega) = \frac{100}{j\omega(15 - j\omega)}$$

**18-7.** Make an open-loop plot for

$$\mathbf{Y}_o(j\omega) = \frac{-\omega^2 X}{(1 + 2j\omega)(1 - 0.5j\omega - \omega^2)}$$

and determine the conditions which $X$ must fulfill to lead to a stable system.

**18-8.** The system of Fig. 18-34 has an equation of motion given by

$$C(\theta_i - \theta_o) = Fp\theta_o + T_L$$

Using the method which led to Eq. (18-22), write the solution for $\theta_o$ as a power series in $p$, and solve for the steady-state output for inputs of $\theta_i = \alpha_1 t$ and $\theta_i = \beta_1 t^2$, with $T_L$ a constant load.

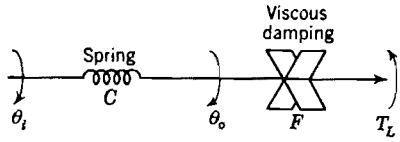



FIG. 18-34

**18-9.** The system of Fig. 18-7 has $J = 0.75$ slug-in.$^2$ and $F = 0.15$ in.-lb per radian per sec. (*a*) Plot the open-loop transfer function in the $Y_o$ plane. (*b*) From this plot determine the value of gain $K'$ for $M_p = 1.25$, and find the frequency at which this gain occurs. (*c*) What is the damping ratio for the system?

**18-10.** To the amplifier of Prob. 18-9 is added integral control, which can be done by changing the amplifier transfer function to $Y = K(1 + A_i p^{-1})$, where $A_i = 2$ sec$^{-1}$. (*a*) Plot the open-loop transfer function in the $Y_o$ plane. (*b*) What is the gain $K'$ for minimum $M_p$?

**18-11.** A feedback control system has $\mathbf{K'G}(j\omega)$ given by

$$\mathbf{K'}G(j\omega) = \frac{K'}{(j\omega\tau_1 + 1)(j\omega\tau_2 + 1)}$$

(*a*) For $\tau_1 = \tau_2 = 1$ sec plot the Nyquist locus, and find the value of $M_p$. (*b*) Repeat for $\tau_1 = 10$, $\tau_2 = 1$ sec.

## REFERENCES

1. Minorsky, N.: Directional Stability of Automatically Steered Bodies, *J. Am. Soc. Naval Engrs.*, **34**: 280 (1922).
2. Hazen, H. L.: Theory of Servomechanisms, *J. Franklin Inst.*, **218**: 279 (1934).
3. Hazen, H. L.: Design and Test of a High Performance Servomechanism, *J. Franklin Inst.*, **218**: 543 (1934).
4. Nyquist, H.: Regeneration Theory, *Bell System Tech. J.*, **11**: 126 (1932).
5. Smith, E. S.: "Automatic Control Engineering," McGraw-Hill Book Company, Inc., New York, 1944.
6. Eckman, D. P.: "Principles of Industrial Process Control," John Wiley & Sons, Inc., New York, 1945.
7. Brown, G. S., and D. P. Campbell: "Principles of Servomechanisms," John Wiley & Sons, Inc., New York, 1948.
8. James, H. M., N. B. Nichols, and R. S. Phillips: "Theory of Servomechanisms," McGraw-Hill Book Company, Inc., New York, 1947.
9. Lauer, H., R. Lesnick, and L. E. Matson: "Servomechanism Fundamentals," McGraw-Hill Book Company, Inc., New York, 1947.
10. Porter, A.: "Introduction to Servomechanisms," John Wiley & Sons, Inc., New York, 1950.
11. Ahrendt, W. R., and J. F. Taplin: "Automatic Feedback Control," McGraw-Hill Book Company, Inc., New York, 1951.

12. Trimmer, J. D.: "Response of Physical Systems," John Wiley & Sons, Inc., New York, 1950.
13. Nixon, F. E.: "Principles of Automatic Control," Prentice-Hall, Inc., Englewood Cliffs, N.J., 1953.
14. Fett, G. H.: "Feedback Control Systems," Prentice-Hall, Inc., Englewood Cliffs, N.J., 1954.
15. Chestnut, H., and R. W. Mayer: "Servomechanisms and Regulating Systems Design," John Wiley & Sons, Inc., New York, 1951.
16. Thaler, G. J., and R. G. Brown: "Servomechanism Analysis," McGraw-Hill Book Company, Inc., New York, 1953.

# INDEX


Acceptor, 415
Access time, 314
Acoustic delay line, 314
Addition logic, 322
Address, memory, 315
Alternating current, control of, 555
  motor, 607
    servo system, 641
  regulator, 360
*Amplification factor*, 25, 38
  collector-to-base short circuit, 427
  gas, 473
  short circuit, 425
Amplifier, a-c operation of, 84
  bandwidth, 92
    alteration by *RC* networks, 99
    factors affecting, 103
    for recurrent pulses, 128
  basic types, 64
  cascaded, 65, 450
  cathode follower, 65
    analysis of, 73
    bias for, 80
    design of, 75
    input admittance of, 74
    modified, 81
    output impedance of, 74
  chopper-modulated, 220
  Class A power, 136
    ideal, 144
    load-line analysis for, 142
    minimum distortion in, 156
    *optimum power output of*, 153
    plate-circuit efficiency of, 146
    power relations in, 145
    power sensitivity, 145
    reactive load on, 142
    theoretical efficiency of, 147
  Class A push-pull, 158
    composite tube for, 163
Amplifier, Class A push-pull, equivalent circuit for, 162
    graphical analysis for, 164
    phase inverters for, 178
  Class AB, 167
  Class B, 168
    design of, 171
    grid power required, 176
  Class C, analysis, 368
    definition, 55
    frequency multiplier, 377
    power gain, 374
    resonant load for, 375
    waveforms in, 367
  coupling methods for, 65
  decibel-gain plots, 96
  decoupling in, 196
  difference, 212
  direct-coupled, 206
    balanced, push-pull, 211
    bridge-balanced, 214
    cathode-coupled, 210
    chopper-modulated, 223
    chopper-stabilized, 223
    difference, 212
    drift compensation for, 213
      Miller, 213
      series, 214
    Loftin-White, 207
    residual drift in, 219
    resistance-coupled, 208
    varistor modulated, 223
  distortion, determination of, 148
  electrometer, 217
  feedback, 184
    basic circuit, 191
    current, 191
    distortion in, 187
    noise in, 189
    oscillation conditions for, 198

Amplifier, feedback, positive, 196
stability of, 197
conditions for, 201
gain, 185
voltage, 191
frequency ranges of, 88
gain, 43
decibel, 56
gain-bandwidth product, 104
table, 105
grounded-cathode, 64
analysis of, 69
frequency response of, 87
gain, 90
input admittance, 66
phase response, 87
grounded-grid, 65
analysis of, 82
impedance of, input, 83
output, 83
high-frequency-compensated, 108
light, 484
load-line analysis for, 142
low-frequency-compensated, 107
noise in, 130
operating conditions, 54
operational, 226
in analog computer, 240
positive feedback in, 231
power, low-frequency, 136
pulse response of, 114
*RC*-coupled, 87
pulse response of, 114
square-wave testing of, 120
relation of pulse to sinusoidal response in, 118
self-bias for, 51
series-compensated, 111
shunt-compensated, 108
small signal, 64
square-wave response of, 120
transistor (*see* Transistor)
Amplitude distortion, 53, 148, 151
Analog computer, 231
diode nonlinear resistance for, 246
direct, 232
scale factors for, 238
function generators for, 243
functional, 232
limiting circuits for, 247
multiplier for, 248
quarter-square, 250
Analog computer, nonlinear elements for, 246
operational, 240
Analog relations, table, 238
*And* circuit, 321
Angstrom unit, 467
Anode, 22
Arc, 29
Arc-back, 496
Arc discharge, 488
Arc drop, 495
Asynchronous computer, 310
Atom, 1
Average tube current, 353

Bandwidth, alteration, 99
definition, 92
factors affecting, 103
for recurrent pulses, 128
Bardeen, J., 424
Barium titanate, 387
Barkhausen criterion, 378
Barrier layer, 420
Base, 424
Battery as voltage reference, 362
Beam tube, 50
Bel, 56
Bell, Alexander Graham, 56
Bessel function, 391
values of, 394
Bias, cathode follower, 80
self-, 51
in transistor amplifier, 448
Bias control of thyratron, 533
Bias-phase control, 537
Binary number system, 305
Bit, definition, 308
Black body, 403
Bleeder resistor, 346
Blocking oscillator, 276
Bootstrap sweep, 284
Brattain, W. H., 424
Breakdown, gas, 29
potential, 489
Bridge, phase-shifting, 539
Bridge rectifier, 502
Buffer, 382

Candle, 468
Capacitor-input filter, 350

Cascaded amplifier, vacuum-tube, 65
transistor, 450
Cathode, equipotential, 17
filamentary, 17
heat-shielded, 17
mercury-pool, 491
oxide-coated, 16
photoemissive, 469
sheath, 489
spot, 491
thermionic, in gas discharge, 490
thoriated-tungsten, 16
tungsten, 15
virtual, 23
Cathode coupling, in direct-coupled amplifier, 210
in inverter, 179
Cathode-drift compensation, 213
Cathode follower (*see* Amplifier)
Cathode-ray tube, 57
oscillograph, 57
photography of, 61
screen materials, table, 60
sweep voltage for, 59
Charge density, 5
Choke, 343
swinging, 349
Chopper, 220
Clamping circuits, 270
in television, 271
Class A, definition, 54
(*See also* Amplifier)
Class B, definition, 54
(*See also* Amplifier)
Class C, definition, 55
(*See also* Amplifier)
Clipping, diode, 265
triode, 268
Clock, 310
Closed-loop system, 613
Collector, 424
Color temperature, 468
Colpitts oscillator, 382
Composite tube, 163
Conductors, physical constants for, 405
Constant-amplitude input, 617
Constant-current curves, 369
Constant-$M$ circles, 646
Constant-velocity input, 618
Control, continuous, 615
on-off, 615
proportional, 615
Copper-oxide rectifier, 462
Critical input inductance, 348
Crystal, quartz, 384
as transducer, 387
$X$- and $Y$-cut, 383
Cutoff, definition, 26
extended, 26
remote, in pentode, 51
voltage, 26
Cyclic storage, 313

Damping ratio, 623
Dark current, 483
D-c motor, armature-current limiting, 603
auxiliary controls, 605
constant horsepower, 590
constant torque, 590
dynamic braking, 606
field weakening in, 603
$IR$-drop compensation in, 601
relations, 588
speed-torque, 595
speed coefficient, 595
speed regulation, 598
thyratron supply for, 590
torque coefficient, 595
transfer function, 639
D-c restorer, 270
Decade-counter tube, 304
Decibel, 56
feedback, definition, 189
Decibel-gain plots, 96
Decimal-binary conversion, 312
table, 306
Decoupling, 196
De Forest, Lee, 3
Deionization, 531
Depth of penetration, 392
Derivative control, 625
Dielectric heating, 399
power required for, 403
Differentiating circuit, 262
Digital computer, 232, 308
functions of, 309
input and output, 311
memory, 313
access time, 314
acoustic delay lines for, 314
address, 315
electrostatic, 315

Digital computer, memory, magnetic drum, 315
magnetic tape, 314
matrix, core, 316
relay, 314
Diode, temperature-saturated, 362
as voltage reference, 363
thermionic, 21
Diode limiter, 247
Diode line, 172
Diode modulator, 223
Diode rectifier, 328
full-wave, 331
half-wave, 328
Direct-coupled amplifier (*see* Amplifier)
Distortion, 52, 53
amplitude, determination of, 148
total, 151
frequency, 53
intermodulation, 54, 152
nonlinear, 53
phase, 54
Distortion rule, 158
Donor, 414
Dushman's equation, 14
Dynamic braking, 606
Dynamic resistance, 255
Dynode, 479

Eccles-Jordan circuit, 296
Edison, Thomas A., 3
Efficiency of rectification, 330
Einstein, Albert, 19
Electric current, definition, 2
in space, 7
Electric field intensity, 5
Electromagnetic fields in metals, 389
Electromechanical coupling factor, 388
Electrometer, 217
Electron, 1
$e/m$ ratio, 2
emission, 13
force on, in electric field, 5
in magnetic field, 9
free, 11
in metals, 11
primary, 20
secondary, 20
work done on, 6
Electron coupling, 382
Electron-hole pair, 414
Electron multiplier, 479
Electron volt, 6
Electron wavelength, 2
Electronic motor control, 588
Electrostatic storage, 315
Emission, high-field, 20
secondary, 20
in electron multiplier, 479
thermionic, 13
Emission constants, table, 15, 17
Emission efficiency, 16
Emitter, 424
Energy loss in tube, 27
Energy-storage welding, 583
ENIAC, 308
Erasable storage, 313
Error-rate control, 627
Even function, 124
Events per unit time, 574
Excitron, 492
Extended cutoff, 26

Feedback, circuits for, 194
current, 191
in decibels, 189
distortion with, 187
factor, 186
negative, 185
noise with, 189
in oscillators, 377
positive, 196
use of, 231
principles of, 184
in scaling circuits, 300
stability with, 197
voltage, 191
Fermi characteristic energy, 12
Fermi-Dirac distribution, 11
Field weakening, 603
Filter, 331
capacitor input, 350
critical input inductor for, 348
inductor-input, 345
L-section, 345
series-inductor, 340
shunt-capacitor, 334
Fleming, J. A., 3
Fleming valve, 3
Flip-flop, 296
scaling circuits, 298
Follow-up circuit, 578

Foot-candle, 469
Forbidden energy gap, 415
Fourier analysis, of full-wave output, 332
  of half-wave output, 330
Fourier integral, 128
Fourier series, 123
Frequency distortion, 53
Frequency multiplier, 377
Function generator, 243

Gain, current, 438
  decibel, 56
  definition, 45
  magnitude, 43
Gain-bandwidth product, 104
Gain margin, 649
Gain plots in decibels, 96
Gas diode, parallel operation of, 503
Gas ionization, 27
Gating circuit, 316
Germanium, resistivity of, 419
  in transistors, 417
Glow discharge, 29
Grid, 25
  suppressor, 48
Grounded-grid amplifier (*see* Amplifier)

Half-adder, 323
Hartley oscillator, 381
Harvard Mark I computer, 308
Heat control, 582
Heating, dielectric, 399
  high-frequency, 389
  induction, 396, 405
  power calculations for, 405
High-field emission, 13, 20
High-frequency compensation, 108
High-frequency heating, 389
Hold time, 575
Holes, 4, 414
  mobility of, 415
Hull, A. W., 3, 47, 528
Hum, cathode, 131
  in plate supply, 131

Ideal transformer, 137
Ignitron, 3, 528, 546
  control of a-c by, 555
  controlled rectifier, 550
Ignitron, firing circuits for, 548
  inverse parallel operation of, 556
  per cent duty of, 559
  ratings, 558
  reactor firing for, 549
  trailing-tube circuit for, 578
Image converter, 484
Impedance, operational, 227
Induction heating, 396
  conductor physical constants for, 405
  power calculations for, 405
Inductor-input filter, 345
Input admittance, cathode follower, 74
  grounded-grid amplifier, 83
  triode, 66
Integral control, 628
  for temperature regulator, 630
Integrating circuit, 264
Intermodulation distortion, 54, 152
Intrinsic conductivity, 414
Inverse parallel operation, 556
Inverter, thyratron, 564
  transistor, 567
Ionization, 27
  potential, table, 28
*IR*-drop compensation, 601

Junction, *P-N*, 420
  rectifier, 422
  transistor, 425
  Zener level in, 422

L-section filter, 345
Lag circuit, 102
Lagrange's equation, 232
Langmuir, I., 16, 528
Langmuir-Child law, 24
Laplace transform, 200
Lattice, 413
Lead circuit, 103
Light, 467
  amplifier, 484
  units of, 468
Load line, 41, 142
  a-c, 42
  for push-pull, 164
  for reactive load, 142
Loftin-White circuit, 207
Logic, rules of, for addition, 323
Logical circuits, 321

Low-frequency compensation, 107
Lumen, 469
Luminous light flux, 467
Luminous sensitivity, 468

McLachlan, N. W., 398
Magnetic core matrix, 316
Magnetic drum storage, 315
Magnetic tape, 314
Magnetostriction, 386
Matrix, diode switching, 319
  magnetic core, 316
Memory (*see* Digital computer)
Mercury-vapor tube, 494
Microphonic noise, 131
Miller drift compensation, 213
Mobility, 415
Monoformer, 244
Motor control, a-c, 607
  d-c motor relations for, 588
Multiplier, servo-, 641
Multiplying circuits, 248
  quarter-square, 250
Multivibrator, 290
  effect of bias on, 293
  flip-flop, 296
  free-running, 290
  one-shot, 294
  synchronization of, 293
  trigger, 296
    scaling circuits, 298
  types of, 290

Network, four-terminal, 430
  parameters of, 431
Neutron, 1
Newton's law, 6
Noise in amplifier, 130
Noncyclic storage, 313
Nonerasable storage, 313
Nonlinear circuit elements, 246
Nonlinear distortion, 53
Nonvolatile storage, 313
*Not* circuit, 321
Nucleus, 1
Nyquist criterion, 202

Odd function, 124
On-off control, 615
Open-loop system, 613
*Or* circuit, 321
Oscillator, Barkhausen criterion for, 378
  basic feedback circuits for, 381
  blocking, 276
  Colpitts, 382
  crystal control for, 382
  electron-coupled, 382
  feedback in, 377
  Hartley, 381
  as high-voltage source, 407
  magnetostrictive, 386
  negative-resistance, 458
  phase-shift, 384
  *RC*, 385
  ringing, 272
  transistor, 456
  ultrasonic, 386
Oscillograph, 57
  block diagram for, 60
Oxide-coated cathode, 16

Parallel computer operation, 311
Parameters of network, 431
Partition noise, 130
Peak inverse voltage, 331
Peak tube current, 353
Peaking transformer, 544
Pentode, 48
  remote cutoff, 51
  variable-$\mu$, 51
Per cent duty, 559
Permeability, 390
*Permittivity*, 390
Persistence of cathode-ray screen, 60
Perveance, 26
Phase distortion, 54
Phase inverter, 178
Phase margin, 649
Phase-selective circuits, 609
Phase-shift control, 536
Phase-shift oscillator, 384
Phase-shifting bridge, 539
Photoconductive cell, 467
Photoemission, 13
Photoemissive cathodes, 469
Photoemissive cell, 467
  gas-filled, 472
Photoemissive vacuum, 470
Photoformer, 244
Photography of cathode-ray screens, 61

Photon, 2, 19
Phototransistor, 467, 482
Phototube, 470
  gas-filled, 472
  modulated light with, 476
  relays, 473
Photovoltaic cell, 467, 481
Pickup, electric and magnetic, 131
Piezoelectric driver, 387
Plasma, 489
  gas, 29
Plate-circuit efficiency, 146
Plate dissipation, 146
Point-contact rectifier, 422
Point-contact transistor, 424
  instability in, 455
Poisson's equation, 24
Polyphase welding system, 585
Power gain, 374
  of transistor, 438
Power sensitivity, 145
Programming, 308
Proportional control, 615
Proton, 1
Pulse, analysis of, 126
  recurrent rectangular, 126
  repetition rate, 127
  square-wave, 125
Pulse control of thyratron, 543
Pulse response, of *RC* amplifier, 114
  relation to sinusoidal response, 118

Q, definition, 375
Q point, 39
Quantum, 19
Quiescent point, 39

Ramp input, 618
*RC* circuit response, 257
*RC* coupling, pulse response of, 114
*RC* differentiating circuit, 262
*RC* integrating circuit, 264
*RC* network, bandwidth alteration by, 99
*RC* oscillator, 385
*RC* sweep generator, 278
  linearization of, 283
Reactor firing, 549
Rectifier, bridge, 502
  circuit design of, 352
  controlled, 550
Rectifier, diode, with series-inductor filter, 340
  with shunt-capacitor filter, 334
  double-Y, 520
  effect of transformer reactance, 515
  full-wave, diode, 331
    gas, 502
  half-wave, diode, 328
    gas, 497
  junction, 422
  large-area, 462
  *m*-phase, 507
    with inductance filter, 512
  performance, 524
  point-contact, 422
  polyphase, branched connections in, 523
  ripple factor in, 333
  selenium, 462
  three-phase, 504
    bridge, 519
  transformer utility factor, 509
  voltage-multiplying, 352
  voltage regulation of, 525
Recurrent rectangular pulse, 126
Reference, voltage, 362
Regenerative braking, 606
Regulator, a-c, 360
  definition, 614
  temperature, 630
  VR tube, 354
Relay, contact, 570
  phototube, 473
  time delay, a-c operated, 572
    d-c operated, 570
Repetition rate, 127
Reset control, 628
Resistance, linear, 33
  nonlinear, 33
  static, 32
  variational, 32
Resistance welding, 575
  automatic sequence timer for, 576
  energy-storage, 583
  heat control for, 582
  polyphase, 585
  synchronous, 576
  timing, 579
Resistive circuit elements, 255
Resolving time, 298
Resonant load circuit, 375
Ringing circuit, 272

Ripple factor, 333
*RL* circuit response, 257

Saturable reactor, 541
  motor control by, 600
Scale factor in analog computation, 238
Scale-of-two counter, 564
Scaling circuits, 298
  decade, 300
  resolving time of, 298
Screen-grid tube (*see* Tetrode)
Secondary emission, 13
  coefficient, 20
Selenium rectifier, 462
Self-maintaining discharge, 29
Semiconductor, 3, 412
  electron arrangement in, 413
  electron-hole pair in, 414
  energy gap in, 415
    table of values, 417
  energy levels in, 415
  holes in, 414
  intrinsic conductivity of, 414
  intrinsic resistivity, equation for, 417
  *N*-type, 414
  *P-N* junction, 420
  *P*-type, 415
  thermistor, 463
Series-compensated amplifier, 111
Series-computer operation, 311
Series-inductor filter, 340
Servomechanism, 613
  a-c motor in, 641
  constant-*M* circles for analysis of, 646
  control for, derivative, 625
    error-rate, 627
    integral, 628
  damping ratio of, 623
  first-order, 615
  gain margin of, 649
  lag network for, 638
  lead network for, 637
  phase margin of, 649
  ramp input for, 618
  second-order, 620
    response of, sinusoidal, 631
      transient, 622
  stability of, 642
    criteria for, 645
  transfer function for, 634
  undamped natural frequency of, 623
Servomultiplier, 641
Shells, 413
Shot noise, 130
Shunt-capacitor filter, 334
Shunt-compensated amplifier, 108
Silicon in transistor, 417
Skin effect, 389
  in metals, 392
Slepian, J., 528, 546
Snooperscope, 484
Space charge, 21
  Langmuir-Child law of, 24
Speed coefficient, 595
Speed control by tachometer, 606
Square wave, Fourier series of, 125
  frequency spectrum of, 129
  table of harmonics for, 126
Square-wave testing, 120
  of transformers, 141
Squeeze time, 575
Stability criteria for servomechanism, 645
Static resistance, 255
Stefan-Boltzmann law, 363, 403
Step function, 616
Storage (*see* Digital computer)
Sweep generator, 277
  bootstrap, 284
  constant-current, 284
  linearization of, 283
  for magnetic deflection, 287
  *RC*, 278
  synchronization of, 281
  thyratron, 280
  vacuum-tube, 282
Sweep voltage, 59
Swinging choke, 349
Switching circuits, 255
Switching matrices, diode, 319
Synchronous computer, 310
Synchronous weld control, 577
Synchrotransmitter, 639

Taylor's series, 148
Temperature regulator, 630
Temperature saturation, 14
Tetrode, gas, 533
  transistor, 461
  vacuum, 47
Thermal emission, 13
Thermal materials, 15
Thermal noise, 130

Thermionic emission materials, 15
Thermistor, 463
Thoriated tungsten, 16
Three-halves power law, 24
Thyratron, 3, 528
  control of, bias, 533
    bias-phase, 537
    phase-shift, 536
    by pulses, 543
  control of alternating current by, 555
  controlled rectifier, 550
  counting circuit, 564
  critical grid characteristic, 530
  with d-c potential, 561
  deionization time, 531
  frequency changer, 566
  grid current, 532
  inverter, 564
  ionization time, 531
  motor control, armature-current limiting, 603
    auxiliary controls for, 605
    *IR*-drop compensator for, 601
    regenerative braking in, 606
    speed control, 598
      bias-phase, 599
      tachometer, 606
  shield grid, 533
  supply for d-c motor, 590
  tetrode, 533
Thyrite, 246
Thyrite bridge, 363
Time base, 286
Time base generation, 277
Time constant, 259
  of servomechanism, 617
Time-interval measurement, 573
Timer, automatic sequence, 576
Torque amplifier, hydraulic, 640
Torque coefficient, 595
Trailing-tube circuit, 578
Transconductance, grid-plate, 38
Transducer, 386
  ultrasonic, 389
Transfer function, closed-loop, 635
  definition, 634
  hydraulic torque amplifier, 641
  lead and lag networks, 637
  multiple-loop, 636
  open-loop, 635
  of physical elements, 639
  synchrotransmitter, 639
Transformer, ideal, 137
  internal capacities of, 137
  output, 139
  peaking, 544
  reactance, effect of, 515
  square-wave testing of, 141
  utility factor, 509
  zig-zag, 523
Transistor, 412
  amplifier, bias stabilization of, 448
    cascaded, 450
    complementary symmetry in, 454
    equivalent circuit, common-base, 434
      common-collector, 442
      common-emitter, 439
    frequency response of, 451
    large-signal analysis of, 447
    push-pull, 454
    summary, 445
  current-amplification factor, 425
  cutoff frequency, 453
  germanium in, 417
  inverter, 567
  junction, 425
  oscillator, 456
    negative resistance, 458
  point-contact, 424
    instability in, 455
  power gain of, 438
  silicon in, 417
  tetrode, 461
  trigger, 458
  versus vacuum tubes, 428
Trigger circuit, 296
  resolving time for, 298
  transistor, 458
Triode, 3, 25
  gas, counting circuit, 564
    switching, 561
  vacuum, amplification factor, 38
    constant-current curves, 38
    current-source equivalent for, 45
    dynamic-transfer characteristic, 39
    gain, definition, 43
    graphical analysis for, 40
    load line for, 41
    notation, 34
    parameters, 6
      measurement of, 46
    plate characteristics, 36
    *Q* point for, 39
    shot noise in, 130

Triode, vacuum, transfer characteristics, 37
voltage-source equivalent, 43
Tube drop, 495
measurement of, 501
Tungsten properties, 15

Ultrasonic oscillator, 386
Undamped natural frequency, 623
Utility factor, 509

Vacuum tube versus transistor, 428
Vacuum-tube electrometer, 217
Vacuum-tube gating circuits, 316
Vacuum-tube regulator, 355
Valence electrons, 413
Variable-$\mu$ tube, 51
Varistor, 223
Volatile storage, 313
Voltage doubler, 352
Voltage gain, 45
Voltage reference, 362
Voltage regulation of rectifier, 525
Voltage regulator, vacuum-tube, 355
VR tube, 354
Voltage-sensitive lamp bridge, 363
VR tube, 363

Waveform, distortion of, 52
Wehnelt, A., 16
Weld time, 575
Williams tube, 315
Work function, 14

Zener level, 422

ANDHRA PRADESH
LIBRARY
College of Basic Courses,
Rajendranagar.
AGRICULTURAL UNIVERSITY

APAU CENTRAL LIBRARY
Acc: No. 12463
Date: 13.8.85

ANDHRA PRADESH
LIBRARY
College of Basic Courses,
Rajendranagar.
AGRICULTURAL UNIVERSITY

CHECKED 2000

12469

**PROF. JAYASHANKAR TELANGANA STATE**
**Agricultural University**
University Library, Rajendranagar, Hyd- 30.

*Accn. No.* 12469 *Call No.* 621.381 N57 IRD

The book should be returned on or before the last date stamped below. Otherwise fines as per the Library rules be charged.

**PROF. JAYASHANKAR TELANGANA STATE**
**Agricultural University**
University Library, Rajendranagar,
Hyderabad - 500 030.

ACC. No. 12469
Call No. 621.381 N157 RYD
Author Ryder, J.D
Title Engineering Electronics

| Borrower's No. | Due Date | Borrower's No. | Due Date |
|---|---|---|---|
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |

| Borrower's No. | Due Date | Borrower's No. | Due Date |
|---|---|---|---|
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |