RCA

Designer's Handbook

# Solid-State Power Circuits

**Technical Series SP-52**

RCA | Solid State Division | Somerville, NJ 08876

Information furnished by RCA is believed to be accurate and reliable. However, no responsibility is assumed by RCA for its use; nor for any infringements of patents or other rights of third parties which may result from its use. No license is granted by implication or otherwise under any patent or patent rights of RCA.

*Copyright 1971 by RCA Corporation*
*(All rights reserved under Pan-American Copyright Convention)*
*Printed in U.S.A. 9/71*

# Contents


| | PAGE |
|---|---|
| **Introduction** | 1 |
| **Materials, Junctions, and Devices** | 3 |
| Semiconductor Materials | 3 |
| Resistivity | 3 |
| Impurities | 4 |
| P-N Junctions | 6 |
| Current Flow | 7 |
| N-P-N and P-N-P Structures | 8 |
| Types of Devices | 9 |
| **Explanation of Device Ratings** | 16 |
| Basis for Device Ratings | 16 |
| Types of Ratings | 17 |
| **Packaging, Handling, and Mounting** | 22 |
| General Considerations | 22 |
| Hermetically Sealed Packages | 23 |
| Molded-Plastic Packages | 32 |
| **Thermal Factors** | 40 |
| Thermal Impedance | 40 |
| Use of External Heat Sink | 46 |
| Effect of Cyclic Thermal Stresses | 55 |
| **Silicon Rectifiers** | 58 |
| Theory of Operation | 58 |
| Thermal Considerations | 62 |
| Electrical Characteristics | 63 |
| Maximum Ratings | 65 |
| Fast-Recovery Rectifiers | 70 |
| Controlled-Avalanche Rectifiers | 74 |
| Series and Parallel Rectifier Arrangements | 75 |
| High-Voltage Rectifier Assemblies | 76 |
| **Basic Design Considerations for Power Transistors** | 80 |
| General Physical Theory | 80 |
| Special Physical Considerations for Power Transistors | 95 |
| Base-Conductivity Modulation | 95 |
| Current Crowding | 96 |
| Base Widening | 97 |
| Design, Processing, and Packaging | 99 |
| **Physical Basis for Power-Transistor Ratings** | 113 |
| Voltage Ratings | 113 |
| Current and Temperature Ratings | 122 |
| Power-Dissipation Ratings | 123 |
| Second Breakdown | 126 |
| Safe-Area Ratings | 134 |
| Thermal-Cycling Ratings | 145 |

PAGE
**Equivalent-Model Analyses of Power Transistors** . . . . . 150
Small-Signal Analysis of Power Transistors in Linear Service . . . . . 150
Large-Signal Analysis of Power Transistors in Linear Service . . . . . 170
Power Transistors in Switching Service . . . . . 175
Switching Speed . . . . . 175
Switching-Time Reduction Techniques . . . . . 180
Power Dissipation . . . . . 182
Load-Line Analysis . . . . . 183
Analysis of Inductive-Load Switching . . . . . 186

**Thyristors** . . . . . 194
Theory of Operation . . . . . 194
Construction . . . . . 206
Ratings and Limiting Characteristics . . . . . 208
Off-State Voltage Ratings . . . . . 208
Maximum Junction Temperature . . . . . 212
On-State Current Ratings . . . . . 214
Holding and Latching Currents . . . . . 218
Critical Rate of Rise of On-State Current (di/dt). . . 220
Critical Rate of Rise of Off-State Voltage (dv/dt) . . 220
Gate Characteristics . . . . . 223
Effect of Gate Signal on Breakover Voltage . . . . 224
Trigger Level . . . . . 225
Pulse Triggering . . . . . 225
Trigger-Circuit Requirements . . . . . 227
Switching Characteristics . . . . . 230
Turn-on Time . . . . . 231
Turn-off Time (for SCR's) . . . . . 233
Commutating Capability (for Triacs) . . . . . 238

**Power Hybrid Circuits** . . . . . 243
Fabrication . . . . . 243
The HC1000 Linear Power Amplifier . . . . . 246
Operating Considerations . . . . . 247

**DC Power Supplies** . . . . . 249
Rectification . . . . . 249
Filtering . . . . . 253
Capacitive-Load Circuits . . . . . 255
Regulation . . . . . 261

**Power Conversion** . . . . . 300
Transistor Inverters and Converters . . . . . 300
Basic Circuit Configurations . . . . . 301
Transformer Considerations . . . . . 303
Additional Considerations . . . . . 305
Design of Practical Transistor Inverters . . . . 306
SCR Inverter . . . . . 334
Circuit Operation . . . . . 335
Gate-Trigger-Pulse Generator . . . . . 336
Applications . . . . . 337

**Thyristor AC Line-Voltage Controls** . . . . . 339
Types of Thyristor Turn-on Controls . . . . . 339
Phase Control . . . . . 339
Zero-Voltage Switching . . . . . 345

PAGE
Triggering Techniques . . . . . 346
Basic Triggering Configurations . . . . . 347
Triggering Devices . . . . . 349
Isolated Trigger Circuits . . . . . 355
Integrated-Circuit Zero-Voltage Switch . . . . . 358
Three-Phase Triac Controls . . . . . 363
Heating Controls . . . . . 366
General Design Considerations . . . . . 367
Comparison of Heating-Control Techniques . . . . 372
Zero-Voltage-Switched Triac Heater Controls . . . . 373
Incandescent Lighting Controls . . . . . 381
Surge-Current Considerations . . . . . 381
Filament Preheat Circuit . . . . . 383
Lamp Dimmers . . . . . 383
Photocell-Operated On-Off Lamp Controls . . . . . 390
Traffic-Signal-Lamp Controls . . . . . 391
Traffic-Control Flasher . . . . . 392
Motor Controls . . . . . 393
Speed Controls for Universal Motors . . . . . 393
Induction-Motor Reversing Controls . . . . . 399
AC Voltage Regulator . . . . . 400

**Ballast Circuits for Mercury-Arc Lamps** . . . . . 409
Relative Merits of Various Lighting Systems . . . . . 409
Characteristics of Mercury-Arc Lamps . . . . . 411
Conventional Ballasting Methods . . . . . 412
Solid-State Ballasting Circuits . . . . . 413
Design Procedure . . . . . 422
Advantages of Solid-State Mercury-Arc-Lamp Ballasting . . . . . 429

**RF Power Amplifiers** . . . . . 430
High-Frequency Power Transistors . . . . . 430
Design Considerations for RF Power Amplifiers . . . . 439
Matching Networks . . . . . 447
Single-Sideband Systems . . . . . 472
RF Amplifiers for Military Applications . . . . . 488
Military Aircraft Communications . . . . . 489
Sonobuoy Transmitters . . . . . 494
Air-Rescue Beacons . . . . . 497
Miniaturized Low-Power Oscillators . . . . . 498
Mobile and Marine Radio . . . . . 499
Commercial Aircraft Radio . . . . . 505
Community-Antenna Television . . . . . 513

**Microwave Power Amplifiers and Oscillators** . . . . . 529
Transistor Considerations . . . . . 529
Circuit Design Techniques . . . . . 531
Microwave Amplifier Circuits . . . . . 537
Amplifier Chains . . . . . 542
Microwave Oscillator Circuits . . . . . 542
Frequency Multipliers . . . . . 550
System Applications . . . . . 562
Microwave Relay Links . . . . . 562
Air-Traffic-Control Systems . . . . . 564

PAGE

**Audio Power Amplifiers** . . . . . 565
Classes of Operation . . . . . 565
Drive Requirements . . . . . 567
Effect of Operating Conditions on Circuit Design . . . . 569
Basic Circuit Configurations . . . . . 572
Power Output in Class B Audio Amplifiers . . . . . 579
Rating Methods . . . . . 579
Basic Power-Dissipation Relationships . . . . . 585
Ratio of Music Power to Continuous Power . . . . 588
Maximum Effective Series Resistance . . . . . 589
Thermal-Stability Requirements . . . . . 590
Effects of Large Phase Shifts . . . . . 591
Effect of Excessive Drive . . . . . 592
Short-Circuit Protection . . . . . 593
Universal-Amplifier Design Approach . . . . . 595
Universal True-Complementary-Symmetry Circuit . . 596
Universal Quasi-Complementary-Symmetry Circuit . . 600
Bridge-Amplifier Design Approach . . . . . 604
Hybrid-Circuit Design Approach . . . . . 606

**Ultrasonic Power Sources** . . . . . 612
Characteristics of Ultrasonic Transducers . . . . . 612
Ultrasonic Generators . . . . . 615
Ultrasonic Power Amplifiers . . . . . 621

**TV Deflection Systems** . . . . . 623
Scanning Fundamentals . . . . . 623
Sync Pulses . . . . . 624
Sync Separation . . . . . 625
Horizontal-Deflection Systems . . . . . 628
Basic Analysis of Horizontal-System Switching . . . 629
Transistor Horizontal-Deflection Circuit . . . . . 631
SCR Horizontal-Deflection System . . . . . 642
Vertical-Deflection Systems . . . . . 654
Basic Design Approach . . . . . 656
Vertical Circuit that Uses a Conventional Output Stage . . . . . 657
Vertical Circuit that Uses a Complementary-Symmetry Output Stage . . . . . 661
Vertical Circuit that Uses a Quasi-Complementary-Symmetry Output Stage . . . . . 665

**Ignition Systems** . . . . . 667
Basic Considerations for Automotive Systems . . . . . 667
Inductive-Discharge Automotive System . . . . . 669
Transistor Ignition System . . . . . 670
Practical Transistor Inductive-Discharge Circuits . . 673
Capacitive-Discharge Systems . . . . . 677
Basic Circuit Operation . . . . . 678
Economic Considerations . . . . . 679
Technical Considerations . . . . . 679
Component Requirements . . . . . 680
Types of Capacitor-Discharge Systems . . . . . 681
Capacitive-Discharge Automotive Ignition System . . 684

**Index** . . . . . 694

# Introduction

This Handbook is intended as a guide to the designers of solid-state power circuits (from dc to microwave). It includes a brief introduction to solid-state physics and general information on electrical ratings, packaging and mounting techniques, and thermal factors for solid-state power devices. Detailed discussions are provided on the theory of operation, basic design concepts, operating parameters, structures, geometries, and capabilities of power transistors, thyristors, silicon rectifiers, and power hybrid circuits. Specific design criteria and procedures are supplied for circuits that use such devices in the amplification, rectification, conversion, control, and switching of electrical power. Design examples are given, and practical circuits are shown and analyzed.

This Handbook, which is the successor to the popular **RCA Power Circuits** Manual, is a comprehensive, authoritative, up-to-date text on the design of solid-state power circuits. It will be found extremely useful by circuit and system designers, educators, students, radio amateurs, hobbyists, and others interested in solid-state power devices and circuits.

(blank page)

# Materials, Junctions, and Devices

SOLID-STATE devices are small but versatile units that can perfom a great variety of control functions in electronic equipment. Like other electron devices, they have the ability to control almost instantly the movement of charges of electricity. They are used as rectifiers, detectors, amplifiers, oscillators, electronic switches, mixers, and modulators.

In addition, solid-state devices have many important advantages over other types of electron devices. They are very small and light in weight (some are less than an inch long and weigh just a fraction of an ounce). They have no filaments or heaters, and therefore require no heating power or warm-up time. They consume very little power. They are solid in construction, extremely rugged, free from microphonics, and can be made impervious to many severe environmental conditions.

## SEMICONDUCTOR MATERIALS

Unlike other electron devices, which depend for their functioning on the flow of electric charges through a vacuum or a gas, solid-state devices make use of the flow of current in a solid. In general, all materials may be classified into three major categories—conductors, semiconductors, and insulators—depending upon their ability to conduct an electric current. As the name indicates, a semiconductor material has poorer conductivity than a conductor, but better conductivity than an insulator.

The materials most often used in semiconductor devices are germanium and silicon. Germanium has higher electrical conductivity (less resistance to current flow) than silicon, and is used in many low- and medium-power diodes and transistors. Silicon is more suitable for high-power devices than germanium. One reason is that it can be used at much higher temperatures. In general, silicon is preferred over germanium because silicon processing techniques yield more economical devices. As a result, silicon tends to supersede germanium in almost every type of application, including the small-signal area, unless a very low device voltage drop is required.

## RESISTIVITY

The ability of a material to conduct current (conductivity)

is directly proportional to the number of free (loosely held) electrons in the material. Good conductors, such as silver, copper, and aluminum, have large numbers of free electrons; their resistivities are of the order of a few millionths of an ohm-centimeter. Insulators such as glass, rubber, and mica, which have very few loosely held electrons, have resistivities as high as several million ohm-centimeters.

Semiconductor materials lie in the range between these two extremes, as shown in Fig. 1. Pure

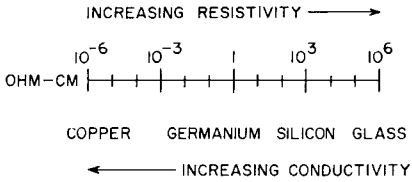



Figure 1. Resistivity of typical conductor, semiconductor, and insulator.

germanium has a resistivity of 60 ohm-centimeters. Pure silicon has a considerably higher resistivity, in the order of 60,000 ohm-centimeters. As used in solid-state devices, however, these materials contain carefully controlled amounts of certain impurities which reduce their resistivity to about 2 ohm-centimeters at room temperature (this resistivity decreases rapidly as the temperature rises).

## IMPURITIES

Carefully prepared semiconductor materials have a crystal structure. In this type of structure, which is called a lattice, the outer or valence electrons of individual atoms are tightly bound to the electrons of adjacent atoms in electron-pair bonds, as shown in Fig. 2. Because such

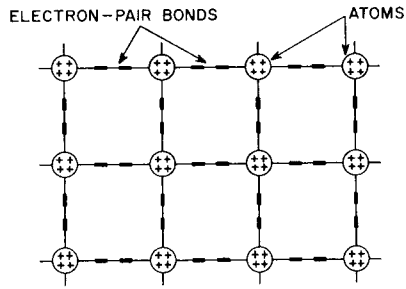



Figure 2. Crystal lattice structure.

a structure has no loosely held electrons, semiconductor materials are poor conductors under normal conditions. In order to separate the electron-pair bonds and provide free electrons for electrical conduction, it would be necessary to apply high temperature or strong electric fields.

Another way to alter the lattice structure and thereby obtain free electrons, however, is to add small amounts of other elements having a different atomic structure. By the addition of almost infinitesimal amounts of such other elements, called **impurities,** the basic electrical properties of pure semiconductor materials can be modified and controlled. The ratio of impurity to the semiconductor material is usually extremely small, in the order of one part in ten million.

When the impurity elements are added to the semiconductor material, impurity atoms take the place of semiconductor atoms in the lattice structure. If the impurity atoms added have the same number of valence electrons as the atoms of the original semiconductor material, they fit neatly into the lattice, forming

the required number of electron-pair bonds with semiconductor atoms. In this case, the electrical properties of the material are essentially unchanged.

When the impurity atom has one more valence electron than the semiconductor atom, however, this extra electron cannot form an electron-pair bond because no adjacent valence electron is available. The excess electron is then held very loosely by the atom, as shown in Fig. 3, and requires only slight excitation to break away. Consequently, the presence of such excess electrons makes the material a better conductor, i.e., its resistance to current flow is reduced.

Impurity elements which are

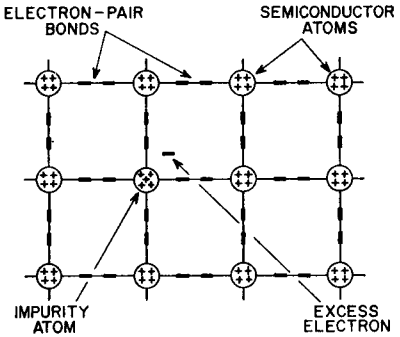



Figure 3. Lattice structure of n-type material.

added to germanium and silicon crystals to provide excess electrons include phosphorus, arsenic, and antimony. When these elements are introduced, the resulting material is called **n-type** because the excess free electrons have a negative charge. (It should be noted, however, that the negative charge of the electrons is balanced by an equivalent positive charge in the center of the impurity atoms. Therefore, the net electrical charge of the semiconductor material is not changed.)

A different effect is produced when an impurity atom having one less valence electron than the semiconductor atom is substituted in the lattice structure. Although all the valence electrons of the impurity atom form electron-pair bonds with electrons of neighboring semiconductor atoms, one of the bonds in the lattice structure cannot be completed because the impurity atom lacks the final valence electron. As a result, a vacancy or **hole** exists in the lattice, as shown in Fig. 4. An electron from an adjacent electron-pair bond may then absorb enough energy to break its bond and move through the lattice to fill the hole. As in the case of excess electrons, the presence of holes encourages the flow of electrons in the semiconductor material; consequently, the conductivity is increased and the resistivity is reduced.

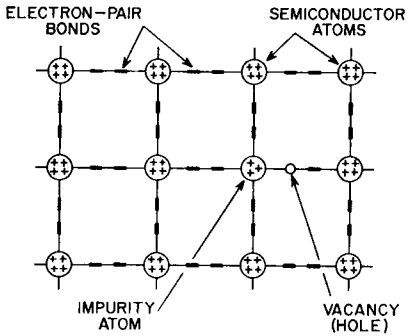



Figure 4. Lattice structure of p-type material.

The vacancy or hole in the crystal structure is considered to have a positive electrical charge because it represents the absence of an electron. (Again, however, the net charge of the crystal is unchanged.) Semiconductor

material which contains these holes or positive charges is called **p-type** material. P-type materials are formed by the addition of boron, aluminum, gallium, or indium.

Although the difference in the chemical composition of n-type and p-type materials is slight, the differences in the electrical characteristics of the two types are substantial, and are very important in the operation of semiconductor devices.

## P-N JUNCTIONS

When n-type and p-type materials are joined together, as shown in Fig. 5, an unusual but

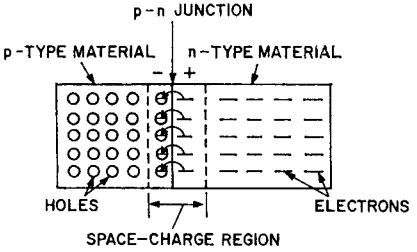



Figure 5. Interaction of holes and electrons at p-n junction.

very important phenomenon occurs at the interface where the two materials meet (called the **p-n junction**). An interaction takes place between the two types of material at the junction as a result of the holes in one material and the excess electrons in the other.

When a p-n junction is formed, some of the free electrons from the n-type material diffuse across the junction and recombine with holes in the lattice structure of the p-type material; similarly, some of the holes in the p-type material diffuse across the junction and recombine with free electrons in the lattice structure of the n-type material. This interaction or diffusion is brought into equilibrium by a small space-charge region (sometimes called the **transition region** or **depletion layer**). The p-type material thus acquires a slight negative charge and the n-type material acquires a slight positive charge.

Thermal energy causes charge carriers (electrons and holes) to diffuse from one side of the p-n junction to the other side; this flow of charge carriers is called **diffusion current.** As a result of the diffusion process, however, a potential gradient builds up across the space-charge region. This potential gradient can be represented, as shown in Fig. 6, by an imaginary battery connected across the p-n junction. (The battery symbol is used merely to illustrate internal effects; the potential it represents is not directly measurable.) The potential gradient causes a flow of charge carriers, referred to as **drift current,** in the opposite direction to the diffusion current.

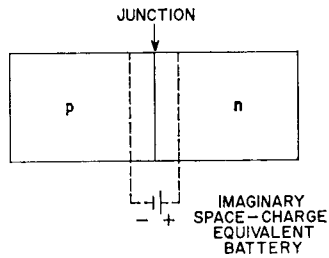



Figure 6. Potential gradient across space-charge region.

Under equilibrium conditions, the diffusion current is exactly balanced by the drift current so that

the net current across the p-n junction is zero. In other words, when no external current or voltage is applied to the p-n junction, the potential gradient forms an **energy barrier** that prevents further diffusion of charge carriers across the junction. In effect, electrons from the n-type material that tend to diffuse across the junction are repelled by the slight negative charge induced in the p-type material by the potential gradient, and holes from the p-type material are repelled by the slight positive charge induced in the n-type material. The potential gradient (or energy barrier, as it is sometimes called), therefore, prevents total interaction between the two types of materials, and thus preserves the differences in their characteristics.

## CURRENT FLOW

When an external battery is connected across a p-n junction, the amount of current flow is determined by the polarity of the applied voltage and its effect on the space-charge region. In Fig. 7(a), the positive terminal of the battery is connected to the n-type material and the negative terminal to the p-type material. In this arrangement, the free electrons in the n-type material are attracted toward the positive terminal of the battery and away from the junction. At the same time, holes from the p-type material are attracted toward the negative terminal of the battery and away from the junction. As a result, the space-charge region at the junction becomes effectively wider, and the potential gradient increases until it approaches the potential of the external battery. Current flow is then extremely small because no voltage difference (electric field) exists across either the p-type or the n-type region. Under these conditions, the p-n junction is said to be **reverse-biased.**

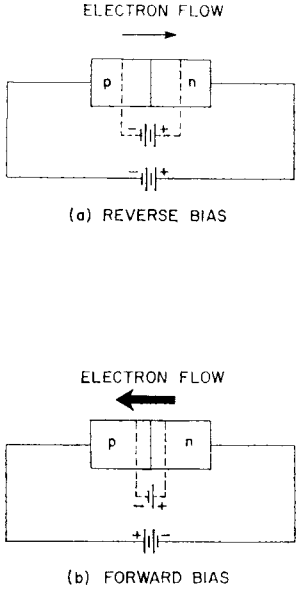



Figure 7. Electron current flow in biased p-n junctions.

In Fig. 7(b), the positive terminal of the external battery is connected to the p-type material and the negative terminal to the n-type material. In this arrangement, electrons in the p-type material near the positive terminal of the battery break their electron-pair bonds and enter the battery, creating new holes. At the same time, electrons from the negative terminal of the battery enter the n-type material and diffuse toward the junction. As a result, the

space-charge region becomes effectively narrower, and the energy barrier decreases to an insignificant value. Excess electrons from the n-type material can then penetrate the space-charge region, flow across the junction, and move by way of the holes in the p-type material toward the positive terminal of the battery. This electron flow continues as long as the external voltage is applied. Under these conditions, the junction is said to be **forward-biased.**

The generalized voltage-current characteristic for a p-n junction in Fig. 8 shows both the reverse-bias and forward-bias regions. In

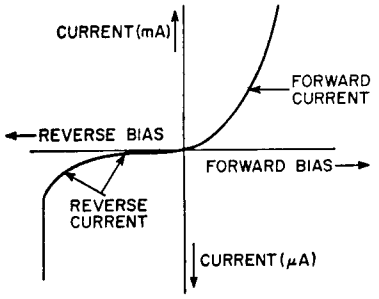



Figure 8. Voltage-current characteristic for a p-n junction.

the forward-bias region, current rises rapidly as the voltage is increased and is relatively high. Current in the reverse-bias region is usually much lower. Excessive voltage (bias) in either direction is avoided in normal applications because excessive currents and the resulting high temperatures may permanently damage the solid-state device.

## N-P-N and P-N-P STRUCTURES

Fig. 7 shows that a p-n junction biased in the reverse direction is equivalent to a high-resistance element (low current for a given applied voltage), while a junction biased in the forward direction is equivalent to a low-resistance element (high current for a given applied voltage). Because the power developed by a given current is greater in a high-resistance element than in a low-resistance element ($P = I^2R$), power gain can be obtained in a structure containing two such resistance elements if the current flow is not materially reduced. A device containing two p-n junctions biased in opposite directions can operate in this fashion.

Such a two-junction device is shown in Fig. 9. The thick end layers are made of the same type of material (n-type in this case), and are separated by a very thin layer of the opposite type of material (p-type in the device shown). By means of the external batteries, the left-hand (n-p) junction is biased in the forward direction to provide a low-resistance input circuit, and the right-hand (p-n) junction is biased in the reverse direction to provide a high-resistance output circuit.

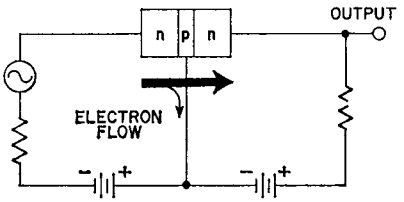



Figure 9. N-P-N structure biased for power gain.

Electrons flow easily from the left-hand n-type region to the center p-type region as a result of the forward biasing. Most of these electrons diffuse through the thin

p-type region, however, and are attracted by the positive potential of the external battery across the right-hand junction. In practical devices, approximately 95 to 99.5 per cent of the electron current reaches the right-hand n-type region. This high percentage of current penetration provides power gain in the high-resistance output circuit and is the basis for transistor amplification capability.

The operation of p-n-p devices is similar to that shown for the n-p-n device, except that the bias-voltage polarities are reversed, and electron-current flow is in the opposite direction. (In general, discussions of semiconductor theory assume that the "holes" in semiconductor material constitute the main charge carriers in p-n-p devices, and discuss "hole currents" for these devices and "electron currents" for n-p-n devices. The direction of hole current flow is considered to be the same as that of conventional current flow, which is assumed to travel through a circuit in a direction from the positive terminal of the external battery back to its negative terminal. This direction is opposite from that of electron current flow, which travels from a negative to a positive terminal.)

## TYPES OF DEVICES

In general, the main active-device functions required in the generation, amplification, rectification, conversion, control, or switching of electrical power can be performed by rectifiers (and other diodes), transistors, thyristors (silicon controlled rectifiers and triacs), or a combination of such devices. A number of other types of devices may also be employed in solid-state power circuits to provide auxiliary functions, such as bias control, temperature compensation, and circuit triggering.

### Diodes

The simplest type of solid-state device is the **diode,** which is represented by the symbol shown in Fig. 10. Structurally, the diode is

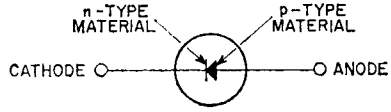



Figure 10. Schematic symbol for a solid-state diode.

basically a p-n junction similar to those shown in Fig. 7. The n-type material which serves as the negative electrode is referred to as the **cathode,** and the p-type material which serves as the positive electrode is referred to as the **anode.** The arrow symbol used for the anode represents the direction of "conventional current flow" mentioned above; electron current flows in a direction opposite to the arrow.

Because the junction diode conducts current more easily in one direction than in the other, it is an effective rectifying device. If an ac signal is applied, as shown in Fig. 11, electron current flows freely during the positive half cycle, but little or no current flows during the negative half cycle.

One of the most widely used types of solid-state diodes is the silicon rectifier. These devices are available in a wide range of

current capabilities, ranging from tenths of an ampere to several hundred amperes, and are capable of operation at voltages as high as 1000 volts or more.

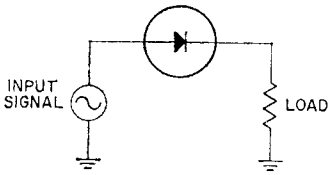



Figure 11. Simple diode rectifying circuit.

Parallel and series arrangements of silicon rectifiers permit even further extension of current and voltage limits. These devices are discussed in detail in the section on **Silicon Rectifiers.**

**Voltage-reference** or **zener diodes** are silicon rectifiers in which the reverse current remains small until the breakdown voltage is reached and then increases rapidly with little further increase in voltage. The schematic symbol for a zener diode is shown in Fig. 12; a typical zener characteristic curve is shown in Fig. 13 in comparison with that of a rectifying diode. The breakdown voltage is a function of the diode material and construction, and can be varied from one volt to several hunderd volts for various current and power ratings, depending on the junction area and the method of cooling. Voltage-reference diodes are useful as stabilizing devices and as reference sources capable of supplying extremely constant current loads.

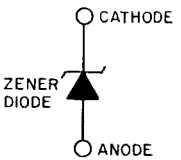



Figure 12. Schematic symbol for a zener diode.

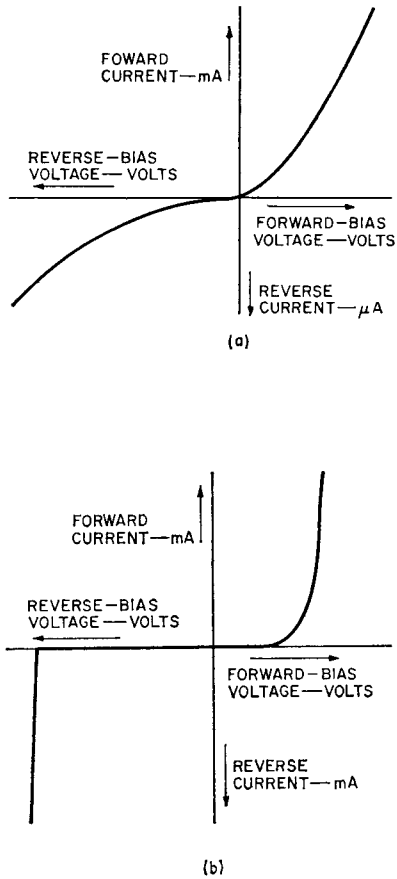



Figure 13. Typical characteristic curves for (a) a rectifying diode and (b) a zener diode.

Current stability can also be obtained by use of a **compensating diode**. Because the forward characteristics of compensating diodes are similar to the transfer characteristics of transistors, these diodes can maintain transistor bias voltages within ±0.015 volt of a desired value despite supply-voltage variations up to 40 per cent and simultaneously compensate for ambient-temperature variations over the range from −20°C to +71°C.

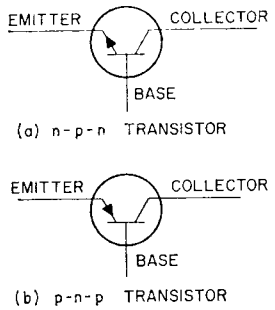



Figure 15. Schematic symbols for transistors.

## Transistors

When a second junction is added to a semiconductor diode to provide power or voltage amplification (as shown in Fig. 9), the resulting device is called a **transistor.** The three regions of the device are called the **emitter,** the **base,** and the **collector,** as shown in Fig. 14. In normal operation, the emitter-to-base junction is biased in the forward direction and the collector-to-base junction in the reverse direction.

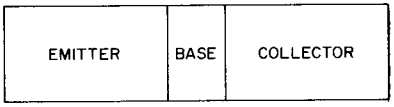



Figure 14. Functional diagram of transistor structure.

Different symbols are used for n-p-n and p-n-p transistors to show the difference in the direction of current flow in the two types of devices. In the n-p-n transistor shown in Fig. 15(a), electrons flow from the emitter to the collector. In the p-n-p transistor shown in Fig. 15(b), electrons flow from the collector to the emitter. In other words, the direction of electron current is always opposite to that of the arrow on the emitter lead. (As in the case of semiconductor diodes, the arrow indicates the direction of "conventional current flow" in the circuit.)

The transistor, which is a three-element device, can be used for a wide variety of control functions, including amplification, oscillation, switching, and frequency conversion. Power-transistor characteristics and ratings are discussed in detail in later sections of this Handbook.

## Thyristors

If two p-type and two n-type semiconductor materials are arranged in a series array that consists of alternate n-type and p-type layers, a device is produced which behaves as a conventional rectifier in the reverse direction and as a series combination of an electronic switch and a rectifier in the forward direction. Conduction in the forward direction can then be controlled or "gated" by operation of the electronic switch. These devices, called **thyristors,** have control characteristics similar to those of thyratron tubes. The **silicon controlled rectifier**

(SCR) and the **triac** are the most popular types of thyristors. Fig. 16 shows the junction diagrams and schematic symbols for the SCR and triac. Such devices are discussed in the section on **Thyristors.**

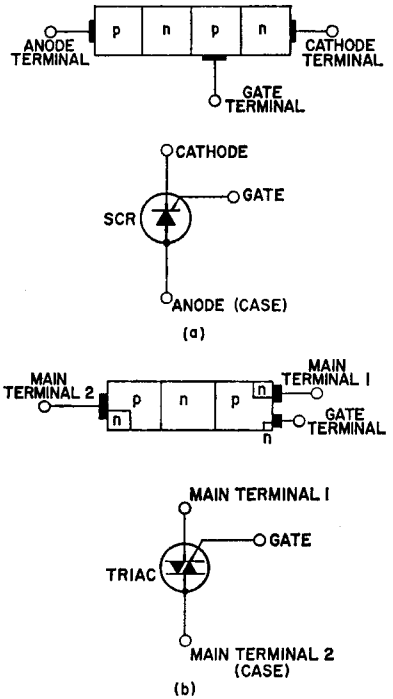



Figure 16. Junction diagrams and schematic symbols for (a) SCR's and (b) triacs.

## Triggering Devices

Although the devices described above are those in most common use in power circuits, several other power devices are used in the circuits in this Handbook and in industry in power application. A **unijunction transistor** is a three-terminal, two-layer device formed by an emitter and a base, as illustrated in Fig. 17. One lead is connected to the emitter

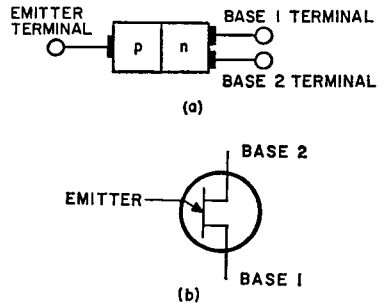



Figure 17. (a) Junction diagram and (b) schematic symbol for a unijunction transistor.

and the other two leads are connected to the base. In normal operation, one base lead is connected to ground while the other is connected to a source of positive bias voltage. When no emitter current flows, the silicon bar used as the base acts as a simple voltage divider and allows a portion of the positive bias voltage to appear at the emitter. If the emitter voltage is less than the portion of the positive bias voltage that appears at the emitter, the emitter is reverse-biased and a small emitter leakage current flows. When the emitter voltage is greater that the positive bias voltage at the emitter, the emitter is forward-biased and an emitter current flows.

Fig. 18 illustrates the use of

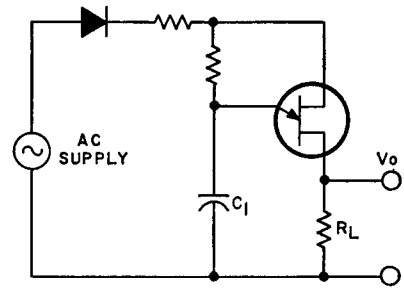



Figure 18. Pulse triggering circuit that uses a unijunction transistor.

the unijunction transistor in a pulse-triggering circuit.

A **diac** is a two-electrode, three-layer bidirectional avalanche diode that can be switched from the OFF state to the ON state for either polarity of applied voltage. Fig. 19 shows the junction diagram and schematic symbol for a diac; Fig. 20 shows the voltage-current characteristic.

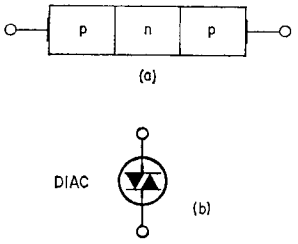



Figure 19. (a) Junction diagram and (b) schematic symbol for a diac.

This three-layer trigger diode is similar in construction to a bipolar transistor, but differs from it in that the doping concentrations at the two junctions are approximately the same and there is no contact made to the base layer. The equal doping levels result in a symmetrical bidirectional switching characteristic, as shown in Fig. 20. When

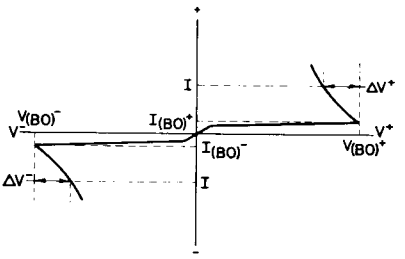



Figure 20. Voltage-current characteristic for a diac.

an increasing positive or negative voltage is applied across the terminals of the diac, a minimum (leakage) current $I_{(BO)}$ flows through the device until the voltage reaches the breakover point $V_{(BO)}$. The reverse-biased junction then undergoes avalanche breakdown and, beyond this point, the device exhibits a negative-resistance characteristic, i.e., current through the device increases substantially with decreasing voltage.

Diacs are primarily used as triggering devices in triac phase-control circuits used for light dimming, universal motor-speed control, heat control, and similar applications. Fig. 21 shows the general circuit diagram for a diac/triac phase-control circuit.

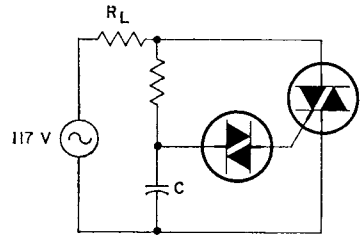



Figure 21. General circuit diagram for a diac/triac phase-control circuit.

A **bilateral switch** is a four-layer, p-n-p-n device in which all layers are accessible as shown in the junction diagram in Fig. 22. Fig. 23 shows the forward-bias anode-to-cathode characteristics of the device. The switch can be turned on by application of a forward bias voltage, or by increasing the anode current through the application of a current to either of the gates. The latter method permits the switch to be turned on even when the

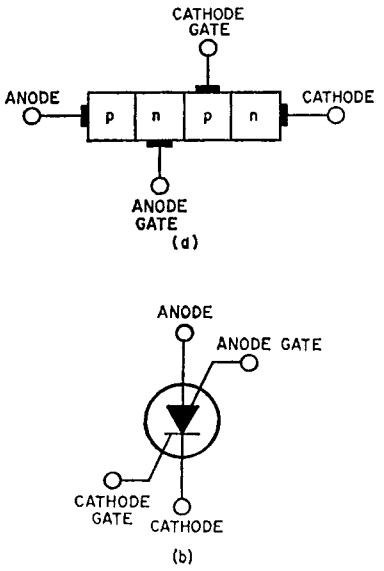



Figure 22. (a) Junction diagram and (b) schematic symbol for a bilateral switch.

junction voltages are well below breakdown. Once the bilateral switch is turned on, it stays on until the anode current decreases below a value called the holding current. To turn the switch off, the anode current may be reduced below the holding value by reverse-biasing the anode, by diverting the current by means of a shunt current path, or by including the switch in an underdamped tuned circuit.

Bilateral switch applications include voltage-level detectors, bistable memory elements, binary counters, shift registers, time delay, pulse, and tone generators, relay drivers, and indicator lamp drivers.

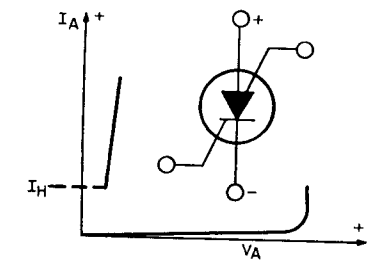



Figure 23. Forward-bias anode-to-cathode characteristics of a bilateral switch.

## Power Hybrid Circuits

Another member of the power-device family consists of a combination of active and passive components interconnected by printed circuitry and packaged as a single **power hybrid circuit.** One fabrication technique for these devices uses a thick-film circuit printed with conductive and resistive inks on a ceramic substrate, and unpackaged transistor and diode chips mounted on the substrate circuit. Another construction uses individually mounted transistors, diodes, and resistors fabricated on individual substrates; these components are fastened to a heat sink, and their terminal pins are soldered to circuit interconnectors. The two construction methods are combined in the power hybrid circuit shown in Fig. 24. In this two-section power amplifier, the low-power input-and-driver section uses thick-film circuitry, diode chips, and transistor chips on a ceramic substrate. The output section consists of two individually mounted power transistors with electrical connections to the thick-film circuit. Regardless of the fabrication technique used, final packaging is accomplished by plastic encapsulation or hermetic sealing.

Power hybrid circuits save

space, weight, and assembly costs in such power applications as amplifiers, switches, regulators, and motor controls. These circuits are described in more detail in a later section of this Handbook.

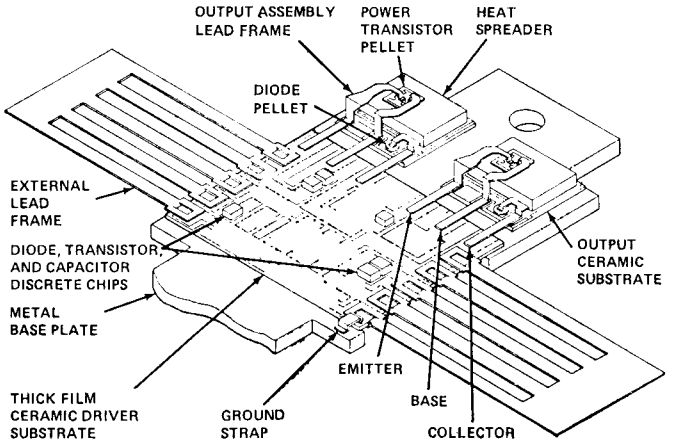



Figure 24. A typical power-hybrid-circuit assembly.

# Explanation of Device Ratings

RATINGS are established for solid-state devices to help circuit and equipment designers use the performance and service capabilities of each type to maximum advantage. They define the limiting conditions within which a device must be maintained to assure satisfactory and reliable operation in equipment applications. A designer must thoroughly understand the constraints imposed by the device ratings if he is to achieve effective, economical, and reliable equipment designs. Reliability and performance considerations dictate that he select devices for which no ratings will be exceeded by any operating conditions of his application, including equipment malfunction. He should also realize, however, that selection of devices that have overly conservative ratings may significantly add to the cost of his equipment.

## BASIS FOR DEVICE RATINGS

Three systems of ratings (the absolute maximum system, the design center system, and the design maximum system) are currently in use in the electronics industry. The ratings given in the technical data for solid-state devices are based on the absolute maximum system. A definition for this system of ratings has been formulated by the Joint Electron Devices Engineering Council (JEDEC) and standardized by the National Electrical Manufacturers Association (NEMA) and the Electronic Industries Association (EIA), as follows:

"Absolute-Maximum ratings are limiting values of operating and environmental conditions applicable to any electron device of a specified type as defined by its published data, and should not be exceeded under the worst probable conditions.

"The device manufacturer chooses these values to provide acceptable serviceability of the device, taking no responsibility for equipment variations, environmental variations, and the effects of changes in operating conditions due to variations in device characteristics.

"The equipment manufacturer should design so that initially and throughout life no absolute-maximum value for the intended service is exceeded with any device under the worst probable operating conditions with respect

to supply-voltage variation, equipment component variation, equipment control adjustment, load variation, signal variation, environmental conditions, and variations in device characteristics."

The rating values specified in the technical data for RCA solid-state devices are determined on the basis of extensive operating and life tests and comparison measurements of critical device parameters. These tests and measurements define the limiting capabilities of a specific device type in relation to the rating factors being considered. The test and measurement conditions simulate, as closely as possible, the worst-case conditions that the device is likely to encounter in actual equipment applications.

Rating tests are expensive, time-consuming, and often destructive. Obviously, therefore, all individual solid-state devices of a given type designation cannot be subjected to these tests. The validity of the ratings is assured, however, by use of stringent processing and fabrication controls and extensive quality checks at each stage in the manufacturing process to assure product uniformity among all devices of a specific type designation and by testing of a statistically significant number of samples.

## TYPES OF RATINGS

Ratings are given for those stress factors that careful study and experience indicate may lead to severe degradation in performance characteristics or eventual failure of a device unless they are constrained within certain limits. Table I lists the critical rating factors used to specify the safe operating capabilities of different types of solid-state devices. In general, these limiting factors include four basic types of ratings.

### Voltage Ratings

A number of voltage ratings are provided for solid-state devices. These ratings are established with respect to a specified electrode (e.g., collector-to-emitter voltage or collector-to-base voltage for transistors) and indicate the maximum potential, for both steady-state and transient operation, that can be safely applied across the two specified electrodes before damage to the crystal occurs. These ratings are specified for particular conditions (e.g., with the third electrode open, or with a specific bias voltage or external resistance, for transistors and thyristors).

Excessive voltage potentials produce high leakage (or reverse) currents in solid-state devices. In silicon rectifiers, the high reverse currents that result from excessive reverse-bias voltages can lead to crystal breakdown and the consequent destruction of the devices. Similar junction breakdown can occur because of the high leakage currents that result from excessive collector-to-emitter or collector-to-base voltages in transistors or excessive off-state voltages in thyristors. Leakage currents flow in solid-state devices at all voltage levels, and device operation is significantly affected by the magnitude of these currents, even at voltages significantly below the breakdown value. For example, in transistors, the collector leakage currents critically

## Table I—Ratings and Limiting Characteristics for Solid-State Devices

| Quantity | Symbol |
|---|---|
| **GENERAL** | |
| Ambient temperature | $T_A$ |
| Case temperature | $T_C$ |
| Junction temperature | $T_J$ |
| Storage temperature | $T_{stg}$ |
| Thermal Resistance | $\theta$ |
| Junction to ambient | $\theta_{J-A}$ |
| Junction to case | $\theta_{J-C}$ |
| Case-to-ambient | $\theta_{C-A}$ |
| Case-to-heat sink | $\theta_{C-S}$ |
| Transient thermal impedance | $\theta_{(t)}$ |
| Junction-to-ambient | $\theta_{J-A(t)}$ |
| Junction-to-case | $\theta_{J-C(t)}$ |
| Delay time | $t_d$ |
| Rise time | $t_r$ |
| Fall time | $t_f$ |
| **SILICON RECTIFIERS** | |
| Forward current: | |
| Total rms value | $I_{F(RMS)}$ |
| DC value, no alternating component | $I_F$ |
| DC value, with alternating component | $I_{F(AV)}$ |
| Instantaneous total | $i_F$ |
| Maximum (peak) total value | $I_{FM}$ |
| Repetitive peak | $I_{FRM}$ |
| Surge (non-repetitive) | $I_{FSM}$ |
| Forward voltage: | |
| Total rms value | $V_{F(RMS)}$ |
| DC value, no alternating component | $V_F$ |
| DC value, with alternating component | $V_{F(AV)}$ |
| Instantaneous total value | $v_F$ |
| Maximum (peak) value | $V_{FM}$ |
| Reverse current: | |
| Total rms value | $I_{R(RMS)}$ |
| DC value, no alternating component | $I_R$ |
| DC value, with alternating component | $I_{R(AV)}$ |
| Instantaneous total value | $i_{RM}$ |
| Reverse recovery time | $t_{rr}$ |
| Reverse voltage: | |
| Total rms value | $V_{R(RMS)}$ |
| DC value, no alternating component | $V_R$ |
| DV value, with alternating component | $V_{R(AV)}$ |
| Instananeous total value | $v_R$ |
| Maximum (peak) total value | $V_{RM}$ |
| Working peak | $V_{RWM}$ |
| Repetitive peak | $V_{RRM}$ |
| Non-repetitive peak | $V_{RSM}$ |
| Reverse breakdown voltage: | |
| DC value, no alternating component | $V_{(BR)R}$ |
| Instantaneous total value | $v_{(BR)R}$ |
| Forward Power Loss: | |
| DC value, no alternating component | $P_F$ |
| DC value, with alternating component | $P_{F(AV)}$ |
| Instantaneous total value | $p_F$ |
| Maximum (peak) total value | $P_{FM}$ |

| Quantity | Symbol |
|---|---|
| **SILICON RECTIFIERS (Cont.)** | |
| Reverse power loss: | |
| DC value, no alternating component | $P_R$ |
| DC value, with alternating component | $P_{R(AV)}$ |
| Instantaneous total value | $p_R$ |
| Maximum (peak) total value | $P_{RM}$ |
| **THYRISTORS AND DIACS** | |
| On-state current: | |
| Total rms value | $I_{T(RMS)}$ |
| DC value, no alternating component | $I_T$ |
| DC value, with alternating component | $I_{T(AV)}$ |
| Instantaneous total value | $i_T$ |
| Maximum (peak) total value | $I_{TM}$ |
| Surge (non-repetitive) | $I_{TSM}$ |
| Overload | $I_{T(OV)}$ |
| Breakover current: | |
| DC value, no alternating component | $I_{(BO)}$ |
| Instantaneous total value | $i_{(BO)}$ |
| Off-state current: | |
| Total rms value | $I_{D(RMS)}$ |
| DC value, no alternating component | $I_D$ |
| DC value, with alternating component | $I_{D(AV)}$ |
| Instantaneous total value | $i_D$ |
| Maximum (peak) total value | $I_{DM}$ |
| Repetitive peak | $I_{DRM}$ |
| Reverse current: | |
| Total rms value | $I_{R(RMS)}$ |
| DC value, no alternating component | $I_R$ |
| DC value, with alternating component | $I_{R(AV)}$ |
| Instantaneous total value | $i_R$ |
| Maximum (peak) total value | $I_{RM}$ |
| Repetitive peak | $I_{RRM}$ |
| Reverse breakdown current: | |
| DC value, no alternating component | $I_{(BR)R}$ |
| Instantaneous total | $i_{(BR)R}$ |
| On-state voltage: | |
| Total rms value | $V_{T(RMS)}$ |
| DC value, no alternating component | $V_T$ |
| DC value, with alternating component | $V_{T(AV)}$ |
| Instantaneous total value | $v_T$ |
| Maximum (peak) total value | $V_{TM}$ |
| Breakover voltage: | |
| DC value, no alternating component | $V_{(BO)}$ |
| Instantaneous total value | $v_{(BO)}$ |
| Off-state voltage: | |
| Total rms value | $V_{D(RMS)}$ |
| DC value, no alternating component | $V_D$ |
| DC value, with alternating component | $V_{D(AV)}$ |
| Instantaneous total value | $v_D$ |
| Maximum (peak) total value | $V_{DM}$ |
| Working peak | $V_{DWM}$ |
| Repetitive peak | $V_{DRM}$ |
| Repetitive peak, with gate open | $V_{DROM}$ |
| Non-repetitive peak | $V_{DSM}$ |
| Non-repetitive peak with gate open | $V_{DSOM}$ |

## Table I—Ratings and Limiting Characteristics for Solid-State Devices (cont'd)

| Quantity | Symbol |
|---|---|
| **THYRISTORS AND DIACS (Cont.)** | |
| Reverse voltage: | |
| Total rms value | $V_{R(RMS)}$ |
| DC value, no alternating component | $V_R$ |
| DC value, with alternating component | $V_{R(AV)}$ |
| Instantaneous total value | $v_R$ |
| Maximum (peak) total value | $V_{RM}$ |
| Working peak | $V_{RWM}$ |
| Repetitive peak, with specified gate-to-cathode resistance | $V_{RRM}$ |
| Repetitive peak, with gate open | $V_{RROM}$ |
| Non-repetitive peak, with specified gate-to-cathode resistance | $V_{RSM}$ |
| Non-repetitive peak, with gate open | $V_{RSOM}$ |
| Reverse breakdown voltage: | |
| DC value, no alternating component | $V_{(BR)R}$ |
| Instantaneous total | $v_{(BR)R}$ |
| Holding current: | |
| DC value, no alternating component | $I_H$ |
| Instantaneous total value | $i_H$ |
| Latching current: | |
| DC value, no alternating component | $I_L$ |
| Instantaneous total value | $i_L$ |
| Gate current: | |
| DC value, no alternating component | $I_G$ |
| DC value, with alternating component | $I_{G(AV)}$ |
| Maximum (peak) total value | $I_{GM}$ |
| Gate trigger current: | |
| DC value, no alternating component | $I_{GT}$ |
| Maximum (peak) total value | $I_{GTM}$ |
| Gate non-trigger current: | |
| DC value, no alternating component | $I_{GD}$ |
| Maximum (peak) total value | $I_{GDM}$ |
| Gate voltage: | |
| DC value, no alternating component | $V_{GT}$ |
| Maximum (peak) total value | $V_{GTM}$ |
| Gate trigger voltage: | |
| DC value, no alternating component | $V_{GT}$ |
| Instantaneous total value | $v_{GT}$ |
| Maximum (peak) total value | $V_{GTM}$ |
| Gate non-trigger voltage: | |
| DC value, no alternating component | $V_{GD}$ |
| Instantaneous total value | $v_{GD}$ |
| Maximum (peak) total value | $V_{GDM}$ |
| Gate power dissipation: | |
| DC value, no alternating component | $P_G$ |
| DC value, with alternating component | $P_{G(AV)}$ |
| Instantaneous total value | $p_G$ |
| Maximum (peak) total value | $P_{GM}$ |

| Quantity | Symbol |
|---|---|
| **POWER TRANSISTORS** | |
| Base current: | |
| DC value, no alternating component | $I_B$ |
| RMS value of alternating component | $I_b$ |
| Instantaneous total value | $i_B$ |
| Collector current: | |
| DC value, no alternating component | $I_C$ |
| RMS value of alternating component | $I_c$ |
| Instantaneous total value | $i_C$ |
| Emitter current: | |
| DC value, no alternating component | $I_E$ |
| RMS value of alternating component | $I_e$ |
| Instantaneous total value | $i_E$ |
| Collector-to-base cutoff current* dc value with emitter open | $I_{CBO}$ |
| Collector-to-emitter cutoff current,* dc value: | |
| With base open | $I_{CEO}$ |
| With specified resistance between base and emitter | $I_{CER}$ |
| With base shorted to emitter | $I_{CES}$ |
| With specified voltage between base and emitter | $I_{CEV}$ |
| With specified circuit between base and emitter | $I_{CEX}$ |
| Emitter-to-base cutoff current,* dc value with collector open | $I_{EBO}$ |
| Power (common-emitter connection): | |
| DC input to base | $P_{BE}$ |
| Instantaneous total input to base | $p_{BE}$ |
| Large-signal output | $P_{OE}$ |
| Total nonreactive input to all terminals | $P_T$ |
| Instantaneous total nonreactive input to all terminals | $p_T$ |
| Emitter-to-base open-circuit dc voltage (floating potential) | $V_{EB(fl)}$ |
| Collector-to-base dc voltage, with emitter open | $V_{CBO}$ |
| Collector-to-emitter dc voltage: | |
| With base open | $V_{CEO}$ |
| With specified resistance between base and emitter | $V_{CER}$ |
| With base shorted to emitter | $V_{CES}$ |
| With specified voltage between base and emitter | $V_{CEV}$ |
| With specified circuit between base and emitter | $V_{CEX}$ |
| Emitter-to-base dc voltage, with collector open | $V_{EBO}$ |
| Second Breakdown | |
| Forward-bias energy level | $I_S/_b$ |
| Reverse-bias energy level | $E_S/_b$ |
| Thermal-cycling capability (number of cycles) | N |

* Cutoff current is also referred to as reverse current or leakage current.

affect biasing levels, and consequently the gain and stability of the over-all circuit. In thyristors, high leakage current levels can cause unwanted switching of device conduction states. In addition to their dependence on voltage, leakage currents also vary with temperature. In the technical data on solid-state devices, therefore, these currents are usually specified for particular voltage and temperature conditions.

## Forward-Current Ratings

If the current in a solid-state device becomes sufficiently large, the semiconductor pellet could be melted by the excessive junction temperatures that result. Maximum current ratings, however, are not usually based on the current-carrying capacity of the semiconductor pellet. Such ratings are usually based on the degradation of specific device performance characteristics that result when the current density exceeds a critical value or on the fusing current of an internal connecting wire.

For devices in which the fusing current of internal connecting wires is not the limiting factor, different parameters are used as the basis for the current ratings of the various types of devices. In silicon rectifiers and thyristors, the maximum on-state current rating is determined on the basis of the maximum permissible forward power dissipation. In power transistors, the current gain is significantly decreased at high current densities. The maximum forward-current rating, therefore, is established on the basis of an arbitrary minimum acceptable gain value.

## Power-Dissipation Ratings

Power is dissipated in the semiconductor material of a solid-state device in the form of heat, which if excessive can cause irreversible changes in the crystal structure or melting of the pellet. This dissipation is equal to the difference between the input power applied to the device and the power delivered to the load circuit. Because of the sensitivity of semiconductor materials to variations in thermal conditions, maximum dissipation ratings are usually given for specific temperature conditions.

In many instances, dissipation ratings for solid-state devices are specified for ambient, case, or mounting-flange temperatures up to 25°C. Such ratings must be reduced linearly for operation of the devices at higher temperatures. Fig. 25 shows a typical power-transistor derating chart that can be used to determine maximum permissible dissipation values at specific temperatures above 25°C. (This chart cannot be assumed to apply to transistor types other than the particular transistors for which it was prepared.) The chart shows the permissible percentage of the maximum dissipation ratings as a function of ambient or case temperature. Individual curves are shown for specific operating temperatures. If the maximum operating temperature of a particular transistor type is some other value, a new curve can be drawn from point A to the desired temperature value on the abscissa, as indicated by the dashed-line curves on the chart.

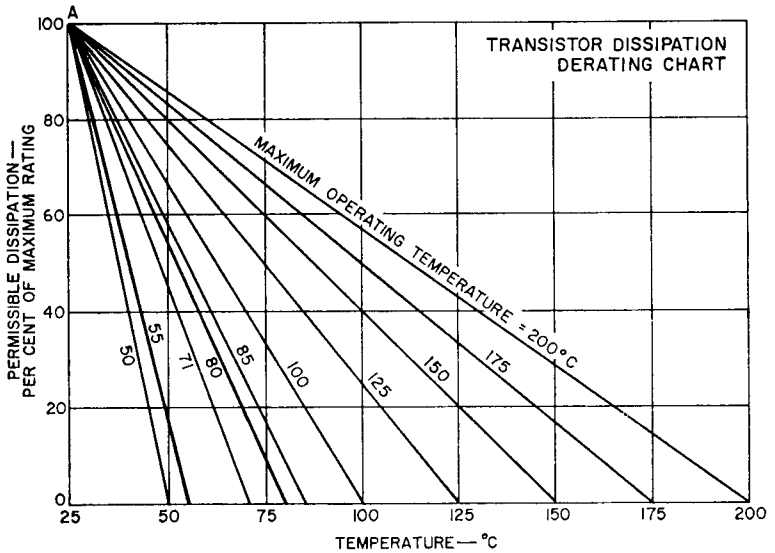



Figure 25. Chart showing maximum permissible percentage of maximum rated dissipation as a function of temperature.

## Junction-Temperature Ratings

The temperature of solid-state devices must be closely controlled not only during operation, but also during storage. For this reason, ratings data for these devices usually include maximum and minimum storage temperatures, as well as maximum operating temperatures. Detailed information on the effect of excessive junction temperatures in solid-state devices is given subsequently in the section on **Thermal Factors** and in the discussions of each particular type of device described in this Handbook.

# Packaging, Handling, and Mounting

SOLID-STATE devices are being designed into an increasing variety of electronic equipment because of their high standards of reliability and performance. However, it is essential that equipment designers be mindful of good engineering practices in the use of these devices to achieve the desired performance.

This section describes device packages and important device mounting and handling recommendations and precautions which should be followed in the interest of assuring the high standards of performance of solid-state devices. (The special considerations involved in the packaging, handling, and mounting of hybrid modules are discussed in a subsequent section on **Power Hybrid Circuits.**)

## GENERAL CONSIDERATIONS

The small size of most solid-state products provides obvious advantages to the designers of electronic equipment. However, it should be recognized that these compact devices usually provide only a relatively small insulation area between adjacent leads and the metal envelope. When these devices are used in moist or contaminated atmospheres, therefore, supplemental protection must be provided to prevent the development of electrically conductive paths across the relatively small insulating surfaces. Specific information on voltage creepage is given in references such as the JEDEC Standard No. 7 "Suggested Standard on Thyristors" and JEDEC Standard No. RS282 "Standards for Silicon Rectifier Diodes and Stacks."

The metal shells of some transistors operate at the collector voltage and of some rectifiers and thyristors at the anode voltage. Therefore, consideration should be given to the possibility of shock hazard if the shells are to operate at voltages appreciably above or below ground potential. In general, in any application in which devices are operated at voltages which may be dangerous to personnel, suitable precautionary measures should be taken to prevent direct contact with these devices.

Devices should not be connected into or disconnected from circuits with the power on be-

cause high transient voltages may cause permanent damage to the devices.

When dip soldering is employed in the assembly of printed-circuit boards, the temperature of the solder should be limited to about 225 to 250°C for a maximum immersion period of 10 seconds. Furthermore, the leads should not be dip-soldered too close to the package case. Fig. 26 shows the areas beyond which dip-soldering should not extend on a silicon-rectifier DO-1 package.

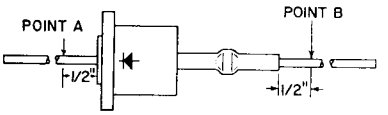



Figure 26. Diagram showing areas beyond which dip-soldering should not extend.

## HERMETICALLY SEALED PACKAGES

Fig. 27 shows photographs of the hermetic packages in which RCA rectifiers, power transistors, and thyristors are supplied. Some information concerning handling and mounting of power devices is given below; before circuit assembly is begun, however, the sections of this Handbook concerning each type of device, the section on **Thermal Factors,** and the technical data for the device to be used should be consulted.

In the following discussion, the information given applies to the package rather than the device unless otherwise specified. In other words, the discussion of handling and mounting of the TO-5 package is understood to cover mounting of transistors, silicon rectifiers, and thyristors in TO-5 packages.

### Packages with Flexible Leads

Some solid-state device packages have flexible leads; these leads are usually soldered to the circuit elements. In all soldering operations, some slack or an expansion elbow should be provided in each lead to prevent excessive tension on the leads. Excessive heat should be avoided during the soldering operation to prevent possible damage to the devices. Some of the heat can be absorbed if the flexible lead of the device is grasped between the case and the soldering point with a pair of pliers.

Although flexible leads can be bent into almost any configuration to fit any mounting requirement, they are not intended to take repeated bending. In particular, repeated bending at the point at which the lead enters the case should be avoided. The leads are not especially brittle at this point, but the sharp edge of the case produces an excessively small radius of curvature in a bend made at the case. Repeated bending with a small radius of curvature at a fixed point will cause fatigue and breakage in almost any material. For this reason, right-angle bends should be made at least 0.020 inch from the case. This practice will avoid sharp bends and maintain sufficient electrical isolation between lead connections and header. A safe bend can be assured if the lead is gripped with pliers close to the case and then bent the requisite amount with the fingers, as shown in Fig. 28. When the leads of a number of devices are to be bent into a particular configuration, it may be advantageous to use a lead-bending fix-

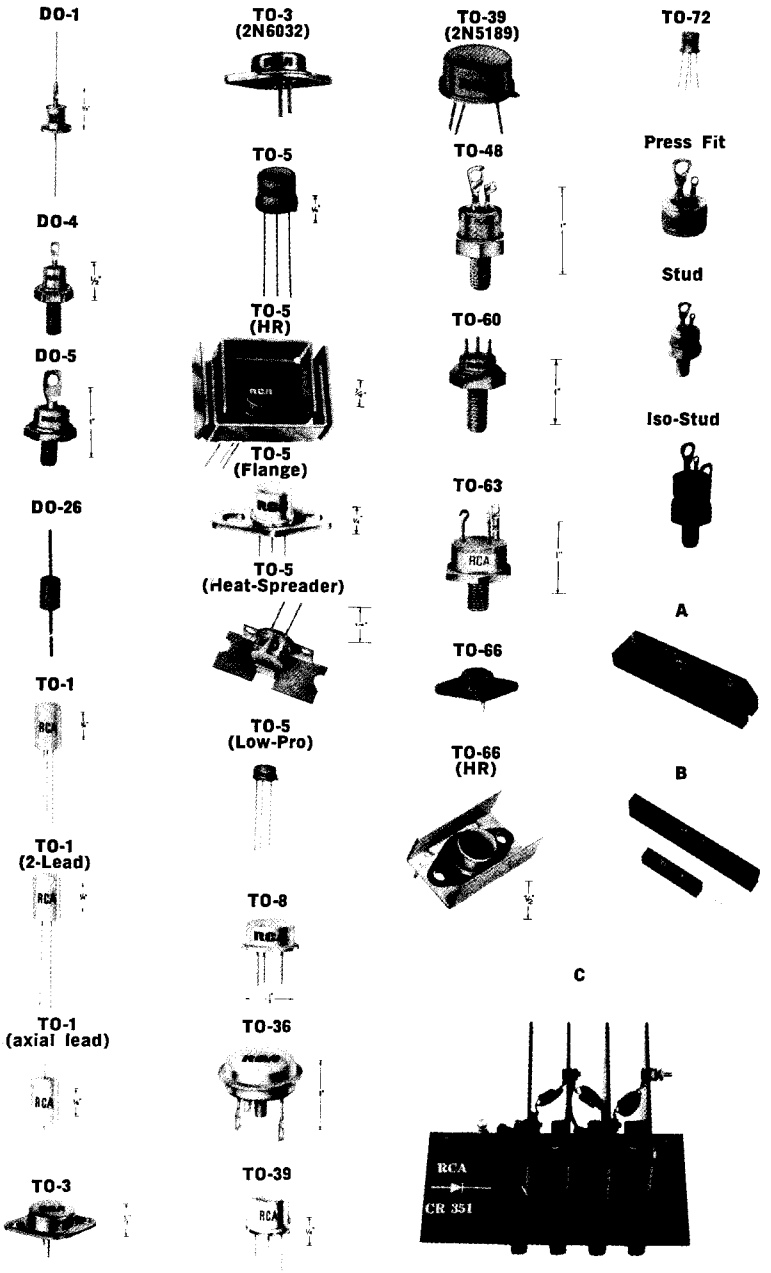



Figure 27. Typical hermetic packages in which RCA rectifiers, power transistors, and thyristors are supplied.

ture to assure that all leads are bent to the same shape and in the correct place the first time, so that there is no need for repeated bending.

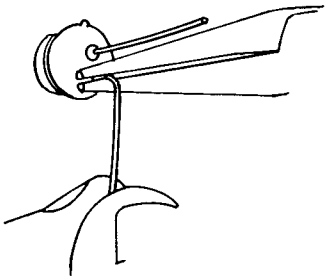


Figure 28. Method of bending leads on a flexible-lead package.

Transistors, thyristors, and rectifiers should be mounted on heat sinks when they are operated at high power levels. Methods of selecting the proper heat sink are discussed in the section on **Thermal Factors.** The most efficient heat-sink method for devices in JEDEC TO-5 and modified TO-5 packages is to provide intimate contact between the heat sink and at least one-half of the base of the device opposite the leads. Fig. 29 shows a variety of ways in which this requirement may be fulfilled. All packages shown can be mounted to the heat sink mechanically, with glue or an epoxy adhesive; soldering, however, is preferable for thyristors and rectifiers. Not only is the solder bond both permanent and most efficient, but the thermal resistance $\theta_{C\text{-}S}$ from the case to the heat sink is easily kept below 1°C per watt under normal soldering conditions. Oven or hot-plate batch-soldering techniques are recommended because of their low cost. Transistors should not be soldered to the heat sink.

Fig. 29(j) shows a combination of a self-jigging bracket and a 60-40 solder preform. If each unit is soldered individually with a flame or electric soldering iron, the heat source should be held on the heat sink and the solder on the unit. Heat should be applied only long enough to permit solder to flow freely.

Fig. 30 shows one method of mounting the flexible-lead TO-1 package, and Fig. 31 shows two suggested methods for attaching the flange-case, axial-lead DO-1 package to a heat sink. The flange of the DO-1 package may also be soldered directly to the heat sink provided that the flange temperature during soldering does not exceed 253°C for a maximum period of 10 seconds. Permanent damage to the device may result if these limits are exceeded.

A good thermal method of isolating the case of a JEDEC TO-5 package from a metal chassis or printed-circuit board is by means of a beryllium oxide washer, as shown in Fig. 32. (WARNING: The crushing, grinding, or abrading of beryllium oxide can produce a dust which may be hazardous if inhaled. Disposal of beryllium oxide should be by burial.) The use of a zinc-oxide-filled silicone compound between the washer and the chassis, together with a moderate amount of pressure from the top of the device, helps to increase thermal conduction. An alternate method is to use a fin-type heat sink, as shown in Fig. 32. Fin-type heat sinks are especially suitable when devices are mounted in Teflon* sockets which provide no thermal conduction to the chassis or printed-circuit board.

* Registered Trademark, E. I. Dupont De-Nemours & Co.

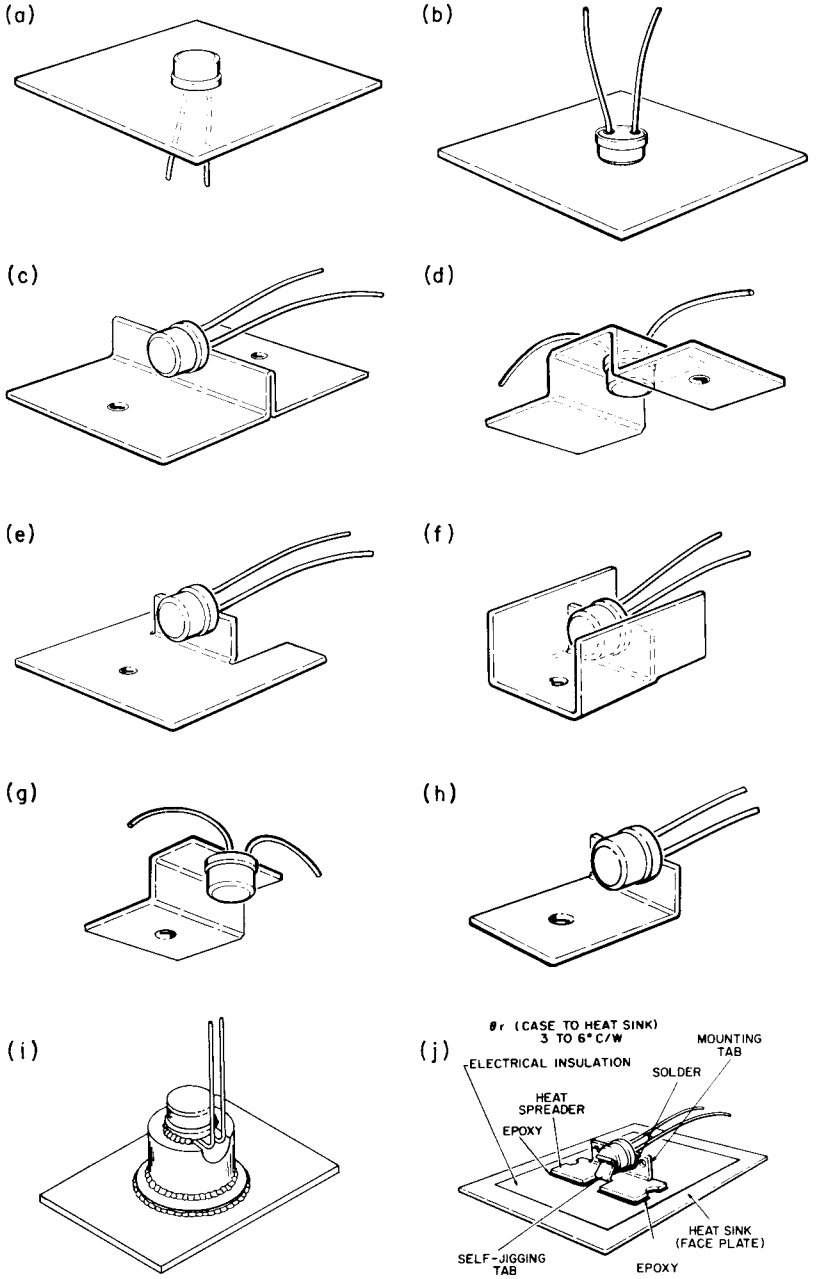
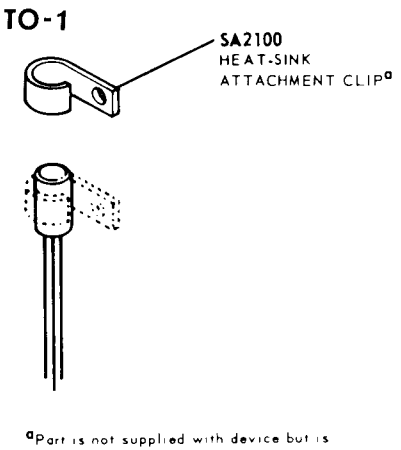



Figure 29. Methods of mounting the JEDEC TO-5 and modified TO-5 packages.




[a]Part is not supplied with device but is available from RCA or an authorized RCA distributor.

Figure 30. Method of mounting the flexible-lead TO-1 package.

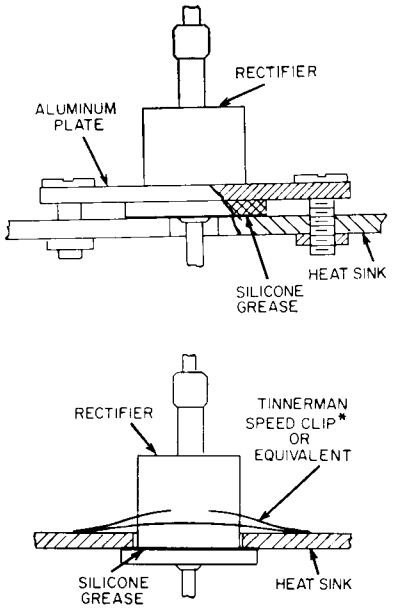



Figure 31. Methods of attaching the flange-case, axial-lead DO-1 package to a heat sink.

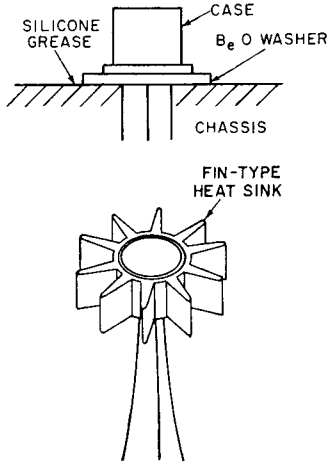



Figure 32. Mounting arrangements for transistors in JEDEC TO-5 packages.

## Packages with Mounting Flanges

The mounting flanges of packages such as the JEDEC-type TO-3 or TO-66 often serve as the collector or anode terminal. In such cases, it is essential that the mounting flange be securely fastened to the heat sink, which may be the equipment chassis. Under no circumstances, however, should the mounting flange be soldered directly to the heat sink or chassis because the heat of the soldering operation could permanently damage the device.

Such devices can be installed in commercially available sockets. Electrical connections may also be made by soldering directly to the terminal pins. Such connections may be soldered to the pins close to the pin seals provided care is taken to conduct excessive heat away from the seals; otherwise, the heat of the soldering operation could crack the pin seals and damage the device.

During operation, the mounting-flange temperature is higher than the ambient temperature by an amount which depends on the heat sink used. The heat sink must provide sufficient thermal conduction to the ambient environment to assure that the temperature of the device mounting flange does not rise above the rated value. The heat sink or chassis may be connected to either the positive or negative supply.

In many applications, the chassis is connected to the voltage-supply terminal. If the recommended mounting hardware shown in the technical data for the specific solid-state device is not available, it is necessary to use either an anodized aluminum insulator having high thermal conductivity or a mica insulator between the mounting-flange and the chassis. If an insulating aluminum washer is required, it should be drilled or punched to provide the two mounting holes for the terminal pins. The burrs should then be removed from the washer and the washer anodized. To insure that the anodized insulating layer is not destroyed during mounting, it is necessary to remove the burrs from the holes in the chassis.

It is also important that an insulating bushing, such as glass-filled nylon, be used between each mounting bolt and the chassis to prevent a short circuit. However, the insulating bushing should not exhibit shrinkage or softening under the operating temperatures encountered. Otherwise the thermal resistance at the interface between the package and the heat sink may increase as a result of decreasing pressure.

Fig. 33 shows methods of mounting flanged packages. Zinc-oxide-filled silicone grease should be used between the device and the heat sink to eliminate surface voids and help conduct heat across the interface. Although glue or epoxy adhesive provides good bonding, a significant amount of thermal resistance may exist at the interface. To minimize this interface resistance, an adhesive material with low thermal resistance, such as Hysol* Epoxy Patch Material No. 6C or Wakefield* Delta Bond No. 152, or their equivalent, should be used.

## Stud Packages

Some high-power solid-state devices are housed in stud packages like the TO-48 or TO-60 shown in Fig. 27. Connection of these packages to the chassis or heat sink should be made at the flat surface of the transistor perpendicular to the threaded stud. A large mating surface should be provided to avoid hot spots and high thermal drop. The hole for the stud should be only as large as necessary for clearance and should contain no burrs or ridges on its perimeter. As mentioned in the discussion of flanged packages, the use of a zinc-oxide-filled silicone grease between the device and the heat sink eliminates surface voids, prevents insulation buildup due to oxidation, and helps conduct heat across the interface. The package can be screwed directly into the heat sink or can be fastened by means of a nut. In either case, care must

* Products of Hyson Corporation, Olean, New York, and Wakefield Engineering, Inc., Wakefield, Massachusetts, respectively.

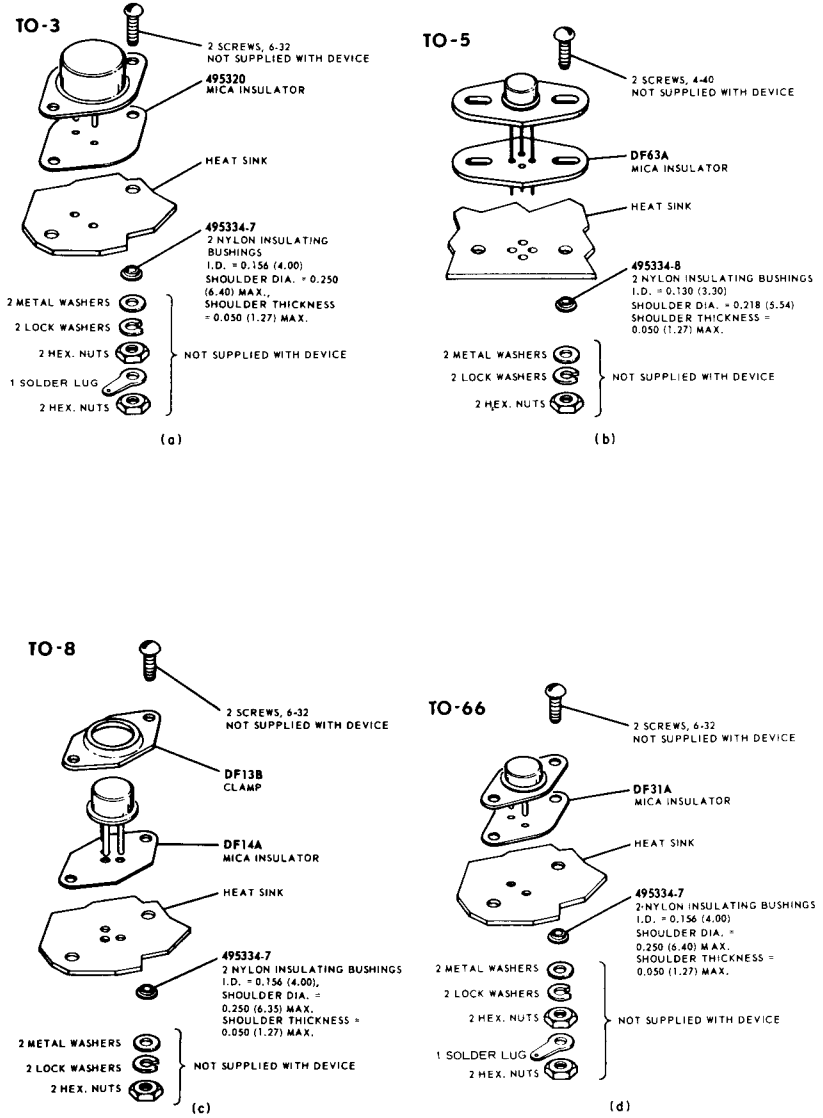



Figure 33. Methods of mounting flanged packages.

be taken to avoid the application of too much torque lest the semiconductor junction be damaged. Maximum limitations are given in the technical data for the particular devices. (CAUTION: Flexible, stranded wire should be used for all connections to the terminals that extend through the glass seals in both stud and press-fit packages. Excessive torque on these terminals may damage the seals and cause a loss in package hermeticity, which leads to premature device failure. These terminals, therefore, should not be bent under any circumstances.)

Although the studs are made of relatively soft copper to provide high thermal conductivity, the threads cannot be relied upon to provide a mating surface. The actual heat transfer must take place on the underside of the hexagonal part of the package. Fig. 34 shows suggested mounting arrangements of the higher-current-type stud packages. Mounting components of the type shown are furnished with each package. With these mounting components, the increase in thermal resistance $\theta_{C\text{-}S}$ from the case to the heat-sink surface can be maintained as low as 0.1°C per watt.

## Press-fit Packages

Press-fit packages are used for some thyristors. Press-fit mounting depends upon an interference fit between the thyristor case and the heat sink. As the thyristor is forced into the heat-sink hole, metal from the heat sink flows into the knurl voids of the thyristor case. The resulting close contact between the heat sink and thyristor case assures low-thermal resistances.

The recommended mounting method shown in Fig. 35 shows press-fit knurl and heat-sink hole dimensions. If these dimensions are maintained, a "worst-case"

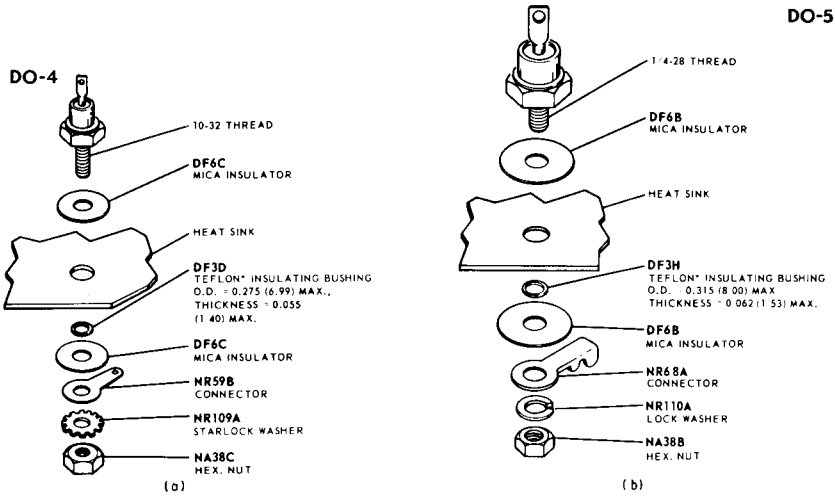



Figure 34. Mounting arrangements for the higher-current-type stud packages
(continued on page 31).

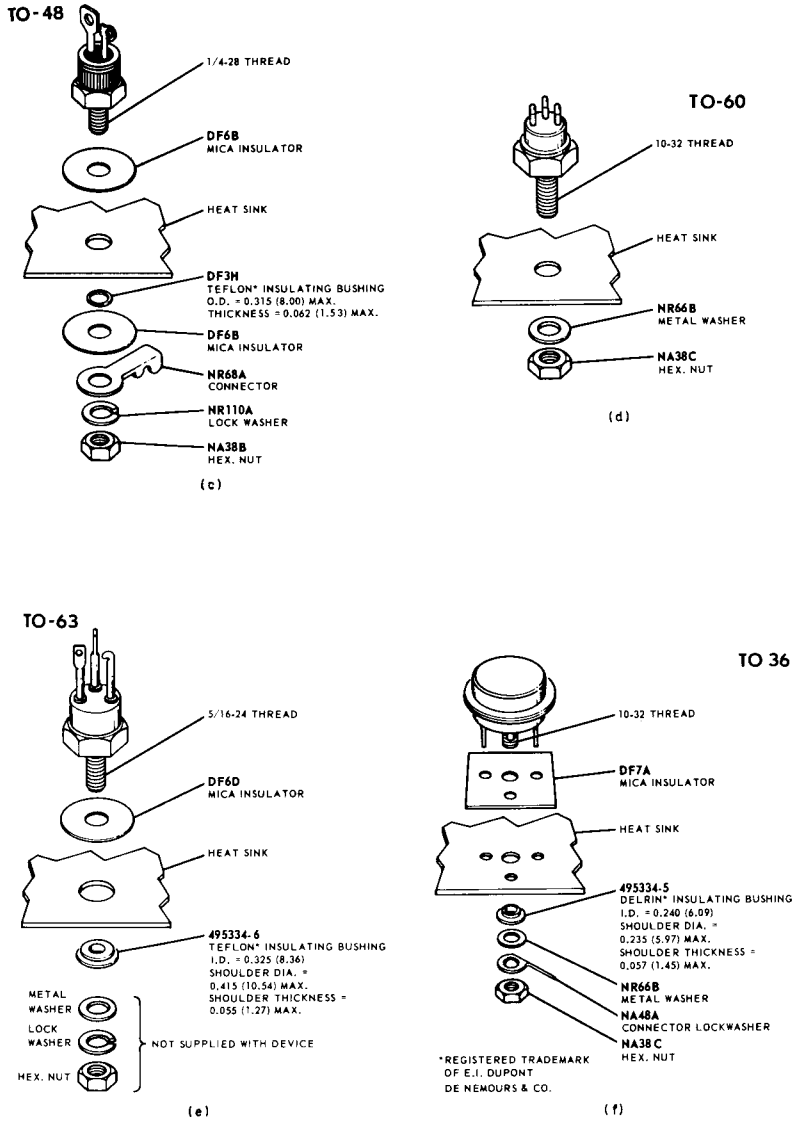



Figure 34. Mounting arrangements for the higher-current-type stud packages (continued from page 30).

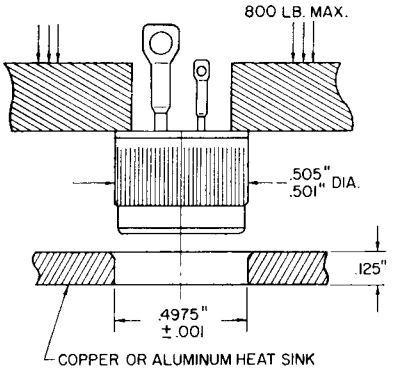



Figure 35. Recommended mounting method for press-fit packages.

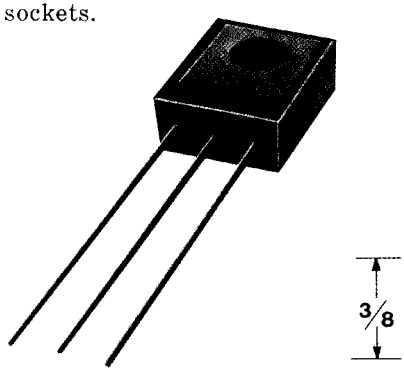
condition of 0.0085-inch interference fit will allow press-fit insertion below the maximum allowable insertion force of 800 pounds. A slight chamfer in the heat-sink hole will help center and guide the press-fit package properly into the heat sink. The insertion tool should be a hollow shaft having an inner diameter of 0.380 ± 0.010 inch and an outer diameter of 0.500 inch. These dimensions provide sufficient clearance for the leads and assure that no direct force is applied to the glass seal of the thyristor. (Refer to CAUTION note shown in section on **Stud Packages.**)

The press-fit package is not restricted to a single mounting arrangement; direct soldering and the use of epoxy adhesives have been successfully employed. The press-fit case is tin-plated to facilitate direct soldering to the heat sink. A 60-40 solder should be used, and heat should be applied only long enough to allow the solder to flow freely.

## MOLDED-PLASTIC PACKAGES

RCA power transistors and thyristors (SCR's and triacs) in molded-silicone-plastic packages are available in a wide range of power-dissipation ratings and a variety of package configurations.

The two most common types of molded-plastic packages include VERSAWATT packages for medium-power applications and high-power plastic packages, both of which are specifically designed for ease of use in many applications. Each basic type offers several different lead options, and the user can select the configuration best suited to his particular application. A third type of package, the TO-5 plastic package, has recently become available; a photograph of this package is shown in Fig. 36. The leads of this package are similar in every way to those on the standard TO-5 package so that the plastic can replace the metallic TO-5 in most sockets.




H-1642

Figure 36. Photograph of the TO-5 plastic package.

Fig. 37(a) through 37(c) shows the options currently available for devices in RCA VERSAWATT

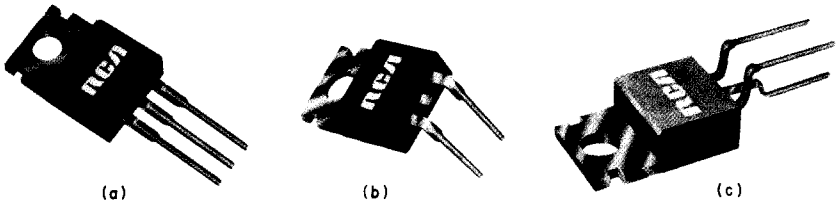



Figure 37. RCA VERSAWATT transistor packages: (a) JEDEC No. TO-220AB in-line lead version; (b) configuration designed for mounting on printed circuit boards; (c) JEDEC No. TO-220AA version, which may be used as replacement for JEDEC No. TO-66 metal packages in JEDEC TO-66 sockets.

packages. The JEDEC Type TO-220AB in-line-lead version, shown in Fig. 37(a) represents the basic style. This configuration features leads that can be formed to meet a variety of specific mounting requirements. Figs. 37(b) and 38 show a package configuration that allows a VERSAWATT package to be mounted on a printed-circuit board with a 0.100-inch grid and a minimum lead spacing of 0.200 inch. Fig. 37(c) shows a JEDEC Type TO-220AA version of the VERSAWATT package. The dimensions of this type of transistor package are such that it can replace the JEDEC TO-66 transistor package in a commercial socket or printed-circuit board without retooling. The pin-connection arrangement of thyristors supplied in TO-220AA packages, however, differs from that of thyristors supplied in conventional TO-66 packages so that some hardware changes are required to effect a replacement. The TO-220AA VERSAWATT package can also be obtained with an integral heat sink.

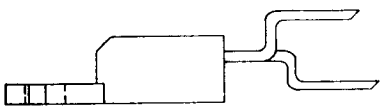


Figure 38. Method of configuring VERSAWATT transistor leads for connection to printed circuit boards and to provide relief in mounting arrangements in which forces are imposed on the package leads.

The RCA molded-plastic high-power packages are also supplied in several configurations for flexibility of application. The JEDEC Type TO-219AB, shown in Fig. 39(a), is the basic high-power plastic package. Fig. 39(b) shows a JEDEC Type TO-219AA version of the high-power plastic package. With the addition of an NR193B top clamp, the TO-219AA package can be used as a direct replacement for the hermetically sealed JEDEC TO-3 package. [The NR193B clamp is shown in the discussion of **Mounting**, Fig. 43(c), later in this section.] The RCA high-power plastic package is also available with an attached header-case lead, as shown in Fig. 39(c). This three-

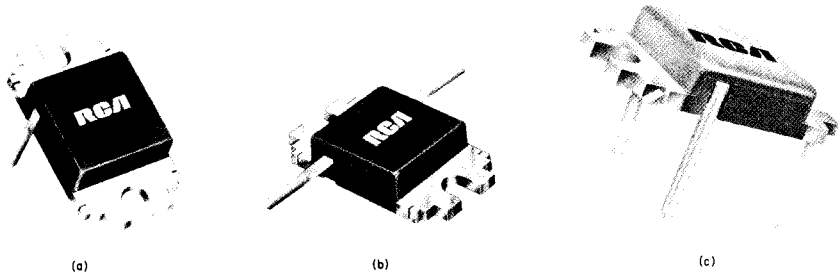



Figure 39. RCA high-power plastic transistor packages: (a) JEDEC No. TO-219AB version which represents the basic configuration; (b) JEDEC No. TO-219AA version which may be used as a replacement for JEDEC TO-3 metal packages in JEDEC TO-3 sockets; (c) configuration designed for mounting on printed circuit boards.

lead package is designed for mounting on a printed-circuit board.

## Lead-Forming Techniques

RCA VERSAWATT plastic packages are both rugged and versatile within the confines of commonly accepted standards for such devices. Although these versatile packages lend themselves to numerous arrangements, provision of a wide variety of lead configurations to conform to the specific requirements of many different mounting arrangements is highly impractical. However, the leads of the VERSAWATT in-line package can be formed to a custom shape, provided that they are not indiscriminately twisted or bent. Although these leads can be formed, they are not flexible in the general sense, nor are they sufficiently rigid for unrestrained wire wrapping.

Before an attempt is made to form the leads of an in-line package to meet the requirements of a specific application, the desired lead configuration should be determined, and a lead-bending fixture should be designed and constructed. The use of a properly designed fixture for this operation eliminates the need for repeated lead bending. When the use of a special bending fixture is not practical, a pair of long-nosed pliers may be used. The pliers should hold the lead firmly between the bending point and the case, but should not touch the case. Fig. 40 illustrates the use of long-nosed pliers for lead bending. Fig. 40(a) shows techniques that should be avoided; Fig. 40(b) shows the correct method.

When the leads of an in-line plastic package are to be formed, whether by use of long-nosed pliers or a special bending fixture, the following precautions must be observed to avoid internal damage to the device:

1. Restrain the lead between the bending point and the plastic case to prevent relative movement between the lead and the case.

2. When the bend is made in the plane of the lead (spreading), bend only the narrow part of the lead.

3. When the bend is made in the plane perpendicular to that of the leads, make the bend at least ⅛ inch from the plastic case.

4. Do not use a lead-bend radius of less than 1/16 inch.

5. Avoid repeated bending of leads.

The leads of the TO-220AB VERSAWATT in-line package are not designed to withstand excessive axial pull. Force in this direction greater than 4 pounds may result in permanent damage to the device. If the mounting arrangement tends to impose axial stress on the leads, some method of strain relief should be devised. Fig. 38 illustrates an acceptable lead-forming method that provides this relief.

Wire wrapping of the leads is permissible, provided that the lead is restrained between the plastic case and the point of the wrapping. Soldering to the leads is also allowed; the maximum soldering temperature, however, must not exceed 275°C and must be applied for not more than 5 seconds at a distance greater than ⅛ inch from the plastic case. When wires are used for connections, care should be exercised to assure that movement of the wire does not cause movement of the lead at the lead-to-plastic junctions.

The leads of the RCA molded-plastic high-power packages are not designed to be reshaped. Simple bending of the leads, however, is permitted to change them from a standard vertical to a standard horizontal configuration, or conversely. Bending of the leads in this manner is restricted to three 90-degree bends; repeated bendings, therefore, should be avoided. Methods of lead forming applicable to the TO-5 plastic package are discussed earlier in this section under the heading Packages with Flexible Leads.

## Mounting

Fig. 41 shows recommended mounting arrangements and sug-

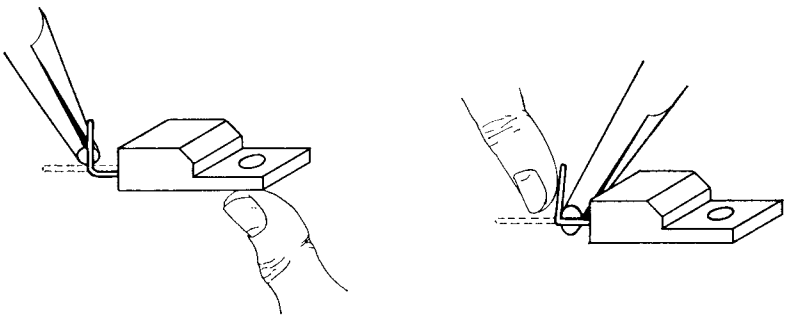
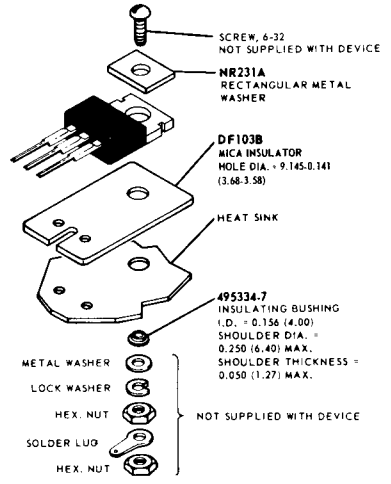
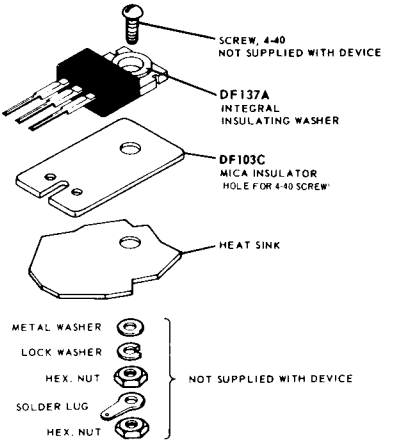
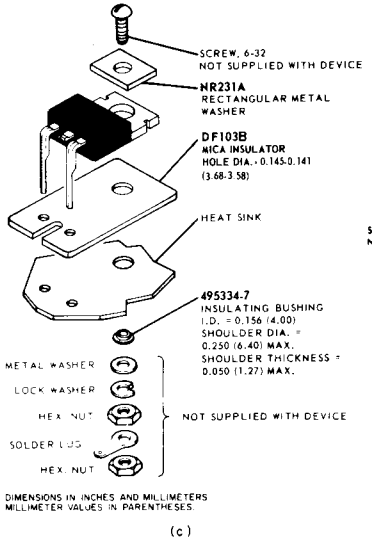



Figure 40. Use of long-nosed pliers for lead bending: (a) incorrect method; (b) correct method.



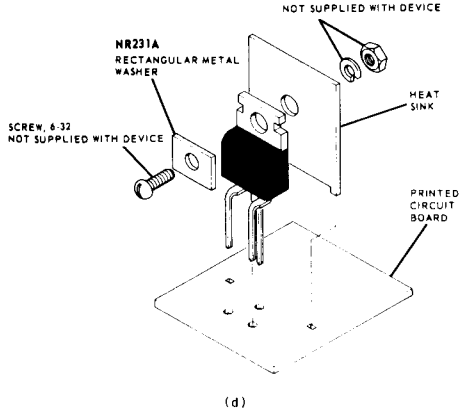




**Figure 41.** Recommended mounting arrangements and suggested hardware for use with VERSAWATT devices.

gested hardware for VERSAWATT devices. The rectangular washer (NR231A) shown in Fig. 41(a) is designed to minimize distortion of the mounting flange when the device is fastened to a heat sink. Excessive distortion of the flange could cause damage to the device. The washer is particularly important when the size of the mounting hole exceeds 0.140 inch (6-32 clearance). Larger holes are needed to accommodate insulating bushings; however, the holes should not be larger than necessary to provide hardware clearance and, in any case, should not exceed a diameter of 0.250 inch. Flange distortion is also possible if excessive torque is used during mounting. A maximum torque of 8-inch-pounds is recommended. The tool used to drive the mounting screw should never come in contact with the plastic body during driving operation. Such contact can result in damage to the plastic body and internal device connections. An excellent method of avoiding this problem is to use a spacer or combination spacer-isolating bushing which raises the screw head or nut above the top surface of the plastic body, as shown in Fig. 42. Suggested materials for these bushings are diallphthalate, fiberglass-filled nylon, or fiberglass-filled polycarbonate. Unfilled nylon should be avoided.

Modification of the flange can also result in flange distortion and should not be attempted. The flange should not be soldered to the heat sink by use of lead-tin solder because the heat required with this type of solder will cause the junction temperature of the device to become excessive.

TO-220AA devices can be mounted in commercially avail-

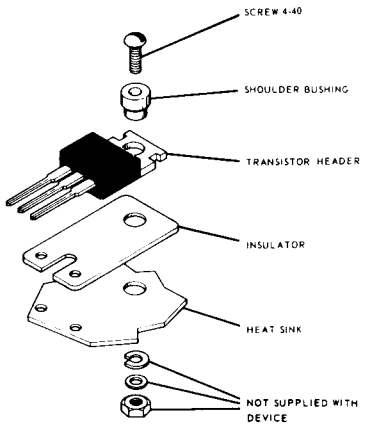
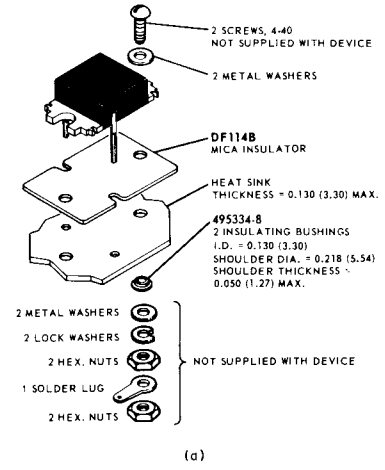


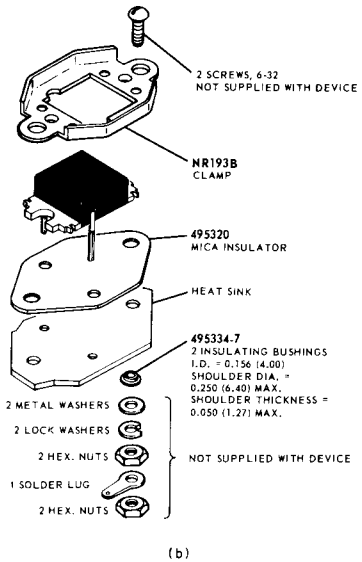

Figure 42. Use of spacer (shoulder bushing) to avoid damage to plastic package in mounting.

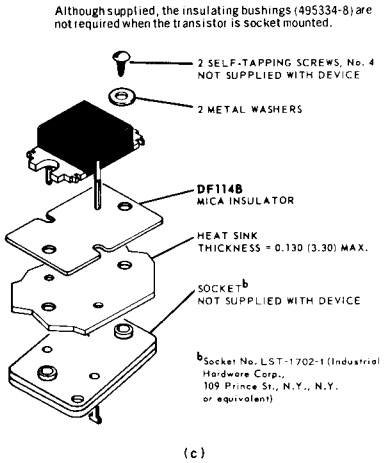
able TO-66 sockets, such as UID Electronics Corp. Socket No. PTS-4 or equivalent. For testing purposes, the TO-220AB in-line package can be mounted in a Jetron Socket No. CD74-104 or equivalent. Regardless of the mounting method, the following precautions should be taken:

1. Use appropriate hardware.
2. Always fasten the devices to the heat sink before the leads are soldered to fixed terminals.
3. Never allow the mounting tool to come in contact with the plastic case.
4. Never exceed a torque of 8 inch-pounds.
5. Avoid oversize mounting holes.
6. Provide strain relief if there is any probability that axial stress will be applied to the leads.
7. Use insulating bushings made of materials that do not have hot-creep problems. Such bushings should be made of diallphthalate, fiberglass-filled nylon, or fiberglass-filled polycarbonate.

Fig. 43 shows the recommended hardware and mounting arrange-



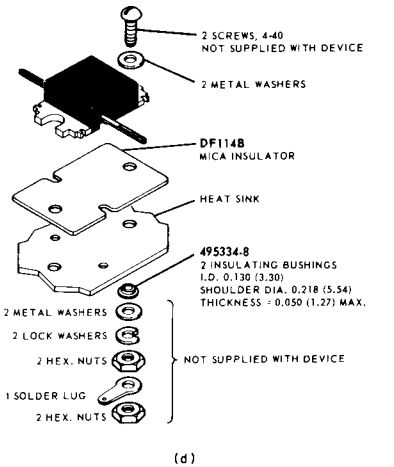




Figure 43. Recommended hardware and mounting arrangements for RCA high-power molded-plastic packages.

ments for RCA high-power molded-plastic devices. These types can be mounted directly in a socket similar to that shown in Fig. 43(b) or they can be mounted in a standard TO-3 socket with the NR193B clamp. The precautions listed for the VERSAWATT packages should also be followed in the mounting of the high-power molded-plastic, and TO-5 plastic packages.

## Cleaning After Mounting

A wide variety of solvents is available for degreasing and removal of flux from device and printed-circuit board after the device has been mounted. The usual practice is to submerge the board in a solvent bath for a specified time. From a reliability standpoint, however, it is extremely important that the solvent, together with other chemicals in the solder-cleaning system (such as flux and solder covers), not adversely affect the life of the device. This consideration applies to all non-hermetic and molded-plastic devices.

It is, of course, impractical to evaluate the effect on long-term device life of all cleaning solvents, which are marketed under a variety of brand names with numerous additives. Chlorinated solvents, gasoline, and other hydrocarbons cause the inner encapsulant to swell and damage the transistor. Alcohols are acceptable solvents and are recommended for flux removal whenever possible. Several examples of suitable alcohols are listed below:

1. methanol
2. ethanol
3. isopropanol
4. blends of the above

When considerations such as solvent flammability are of concern, selected freon-alcohol blends are usable when exposure is limited. Solvents such as those listed below should be safe when used for normal flux removal operations, but care should be taken to assure their suitability in the cleaning procedure:

1. Freon TE
2. Freon TE-35
3. Freon TP-35 (Freon PC)

These solvents may be used for a maximum of 4 hours at 25°C or for a maximum of 1 hour at 50°C.

Care must also be used in the selection of fluxes in the soldering of leads. Rosin or activated-rosin fluxes are recommended; organic fluxes are not.

# Thermal Factors

THE maximum allowable power dissipation in a solid-state device is limited by the temperature of the semiconductor pellet (i.e., the junction temperature). An important factor that assures that the junction temperature remains below the specified maximum value is the ability of the associated thermal circuit to conduct heat away from the device. For this reason, solid-state power devices should be mounted on a good thermal base (usually copper), and means should be provided for the efficient transfer of heat from this base to the surrounding environment.

When a solid-state device is mounted in free air, without a heat sink, the steady-state thermal circuit is defined by the junction-to-free-air thermal resistance given in the published data on the device. Thermal considerations require that there be a free flow of air around the device and that the power dissipation be maintained below that which would cause the junction temperature to rise above the maximum rating. When the device is mounted on a heat sink, however, care must be taken to assure that all portions of the thermal circuit are considered.

Solid-state power devices may also be adversely affected by temperature variations that result from changes in power dissipation during operation or in the temperature of the ambient environment. Such temperature variations produce cyclic mechanical stresses at the interface of the semiconductor pellet and the copper base to which the pellet is attached because of the different thermal-expansion coefficients of these materials. These thermally induced cyclic stresses may eventually lead to a wearout type of failure referred to as **thermal fatigue.**

In this section, the thermal impedances that comprise the basic thermal circuit of a solid-state device are defined, the use and advantages of external heat sinks are described, and the effects of cyclic thermal stresses are analyzed. The basic principles explained are generally applicable to all solid-state power devices regardless of the particular type of device identified in specific examples.

## THERMAL IMPEDANCE

When current flows through a solid-state device, power is dissi-

pated in the semiconductor pellet that is equal to the product of the voltage across the junction and the current through it. As a result, the temperature of the pellet increases. The amount of the increase in temperature depends on the power level and how fast the heat can flow away from the junction through the device structure to the case and the ambient atmosphere. The rate of heat removal depends primarily upon the thermal resistance and capacitance of the materials involved. The temperature of the pellet rises until the rate of heat generated by the power dissipation is equal to the rate of heat flow away from the junction; i.e., until thermal equilibrium has been established.

**Thermal resistance** can be compared to electrical resistance. Just as electrical resistance is the extent to which a material resists the flow of electricity, thermal resistance is the extent to which a material resists the flow of heat. A material that has a low thermal resistance is said to be a good thermal conductor. In general, materials which are good electrical conductors are good thermal conductors, and vice versa.

The methods of rating solid-state power devices under steady-state conditions are indicated by the following definition of thermal resistance: The thermal resistance of a solid-state device is the ratio of the temperature drop to the heat generated through internal power dissipation under steady-state conditions; the temperature drop is measured between the region of heat generation and some reference point.

The over-all thermal resistance of an assembled device is usually expressed as the rise in junction temperature above the case temperature per unit of power dissipated in the device. This information, together with the maximum junction-temperature rating, enables the user to determine the maximum power level at which the device can be safely operated for a given case temperature. Subtraction of the case temperature from the maximum junction temperature indicates the allowable internal temperature rise. If this value is divided by the specified thermal resistance of the device, the maximum allowable power dissipation is determined.

It should be noted that thermal resistance is defined for steady-state conditions. If a uniform temperature over the entire semiconductor junction is assumed, the power dissipation required to raise the junction temperature to a predetermined value, consistent with reliable operation, can be determined. Under conditions of intermittent or switching loads, however, such a design is unnecessarily conservative and expensive. For such conditions, the effect of **thermal capacitance** should also be considered.

## Basic Thermal-Impedance Relationships

The temperature of the semiconductor pellet (i.e., the junction temperature $T_J$) of a solid-state device is related to the temperatures of the various other elements surrounding it by mathematical relationships similar to those that define the properties of an electrical circuit that contains resistance and capacitance. It is convenient, therefore, to describe the thermal properties of a solid-state device in terms of

thermal impedance, thermal resistance, and thermal capacitance.

Fig. 44 shows a layer of thermally conductive material that has a constant cross-sectional

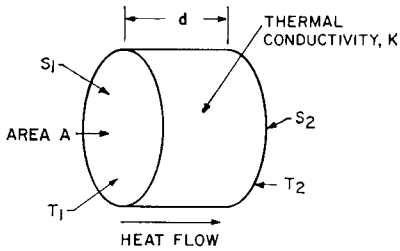



Figure 44. Diagram of a layer of thermally conductive material.

area A, a thickness d, and a thermal conductivity K. (The thermal conductivity K is a basic property of the material itself and is independent of geometry.) If the surface $S_1$ is maintained at a temperature $T_1$ and the surface $S_2$ is maintained at a higher temperature $T_2$, a given quantity of heat Q flows through the layer in a time t. The rate of heat flow P through the layer is expressed by the following relationship:

$$P = Q/t \qquad (1)$$

The thermal conductance $G_T$ of the layer can be determined from the physical dimensions of the layer and the thermal conductivity of the layer material, as follows:

$$G_T = KA/d \qquad (2)$$

The thermal resistance $\theta$ is the reciprocal of thermal conductance $G_T$ and, therefore, is given by

$$\theta = d/KA \qquad (3)$$

The thermal resistance of the layer can be measured experimentally by determination of the time rate of heat flow ($P = Q/t$) and the difference between the temperatures $T_1$ and $T_2$. The following equation defines the thermal resistance, expressed in degrees C per watt, in terms of these quantities:

$$\theta = (T_2 - T_1)/P \qquad (4)$$

The thermal capacitance $C_T$ of a given sample is equal to the product of the specific heat H of the material used in the sample and the mass M of the sample, as follows:

$$C_T = HM \qquad (5)$$

Thermal capacitance may be defined as the quantity of heat absorbed by a sample when its temperature rises 1°C. Therefore, if a given sample absorbs a quantity of heat Q when its temperature is increased from $T_1$ to $T_2$, the thermal capacitance of the sample, expressed in watt-seconds per degree C, can be determined from the following equation:

$$C_T = Q/(T_2 - T_1) \qquad (6)$$

Although both thermal resistance and thermal capacitance vary with temperature, the variation for most materials over the operating range of most solid-state devices is small enough so that it may usually be neglected in thermal calculations.

## Junction-to-Case Thermal Impedance

The thermal properties of a device may be represented by an electrical analog circuit, such as that shown in Fig. 45, which consists of a current generator connected to a series of resistors that have capacitance to ground distributed along their length. The power P dissipated within the crystal of a solid-state device results in a flow of heat outward from the crystal. This flow of

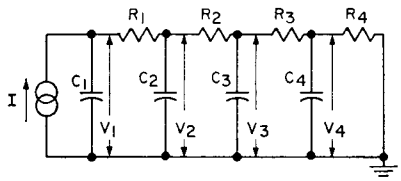



Figure 45. Electrical analog circuit used to describe thermal properties of a solid-state device.

heat (dissipated power P in calories per second or in watts) in a solid-state device is analogous to the flow of charge (electrical current I in coulombs per second or amperes) in such a circuit. Thermal resistances and thermal capacitances of the device are analogous to the electrical resistances and capacitances shown in the circuit. The potential difference or voltage between any two points in the electrical analog circuit is analogous to the temperature difference between the corresponding two points of the device it represents. Table II shows the relationship between various electrical quantities and their corresponding thermal quantities.

Thermal impedance $Z_T$, like electrical impedance Z, is a complex variable because of the time dependence associated with the thermal capacitance $C_T$.

In the electrical or thermal-analog circuit shown in Fig. 45, the thermal resistances closest to the heat source are large because the cross section of the semiconductor element is small (all the heat generated flows through a small area). As shown in Eq. (3), thermal resistance varies inversely with cross-sectional area. In general, thermal resistances become progressively smaller as distance from the semiconductor element increases.

Table II—Comparison of Various Electrical Quantities and Corresponding Thermal Quantities

| Electrical | Thermal |
|---|---|
| Current generator | Heat generator (semiconductor crystal) |
| Resistance R (ohms or volts/ampere) | Thermal Resistance $\theta$ (°C/watt) |
| Capacitance C (ampere-second/volt) | Thermal Capacitance ($C_T$ watt-second/°C) |
| Potential difference $V_1 - V_2$ (volts) | Temperature difference $T_1 - T_2$ (°C) |
| Potential above ground $V - V_G$ (volts) | Temperature above ambient $T - T_A$ (°C) |
| Current I (amperes) | Power dissipation P (watts) |
| Impedance Z (volts/ampere) | Thermal impedance $Z_T$ (°C/watt) |

Eq. (5) indicates that thermal capacitance varies directly with both mass and specific heat. Therefore, the small mass of the semiconductor element of a device causes the thermal capacitance to be smallest at the heat source and to become progressively larger as distance from the heat source increases. The final thermal capacitance in the series must be considered as an infinite capacitance, which electrically is the same as a direct short across the end of the line.

In the electrical-analog circuit shown in Fig. 45, resistance can be determined by application of a steady known current I through the resistors and measurement of the voltage drop E across them. Thus, the resistances are given by

$$R_2 = \frac{V_1 - V_2}{I}, R_3 = \frac{V_2 - V_3}{I}, \ldots \quad (7)$$

In the analogous thermal circuit, thermal resistance is measured by application of a steady known amount of heat, or power P, through the resistors and measurement of the temperature difference $(T_1 - T_2)$ across the thermal resistance $\theta$. Thermal resistances are then given by

$$\Theta_2 = \frac{T_1 - T_2}{P}, \Theta_3 = \frac{T_2 - T_3}{P}, \ldots \quad (8)$$

In the electrical analog circuit, a steady current is essential for accurate measurement because any changes in current are accompanied by charging or discharging of the capacitors, which causes an unknown value of current to flow through the resistors. Because the equation $R = V/I$ is used to solve for resistance, both V and I must be known.

Similarly, in the thermal circuit, there must be a steady heat flow because any charging or discharging of the thermal capacitances produces an unknown variation in the value of P. For example, if the thermal resistance between the junction and the outer case of a solid-state device is to be measured, the arrangement shown in Fig. 46 might be used. The device is mounted on a suitable heat sink, and a steady current I is passed through it. At the same time, measurements are made of the voltage drop V across the device, the temperature $T_J$ at the junction, and the temperature $T_C$ at the case. The power P dissipated as heat within the device, and hence the power that passes out through the thermal resistances, is given by $P = IV$ (watts). The thermal resistance $\theta_{J-C}$ of the device is then given by

$$\Theta_{J-C} = \frac{T_J - T_C}{P} \ (°C/\text{watt}) \quad (9)$$

Such a simple measurement of thermal resistance is applicable to measurement with a constant heat input only. If there is any change or fluctuation in heat-input rate, the change in temperature difference lags behind the change in heat input because some of the heat flows into or out of the thermal capacitances.

If a step function of power is applied to the device (i.e., if the power input at time $t_1$ increases from $P = 0$ to $P = P_1$), the temperature difference between junction and case rises as shown in Fig. 47, and approaches temperature $T_1$ asymptotically. This temperature-rise curve is similar to

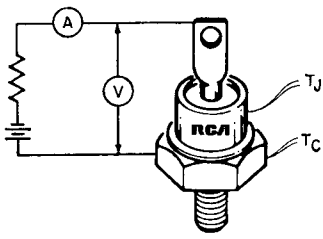



Figure 46. Suitable arrangement for measuring the thermal resistance of a solid-state device.

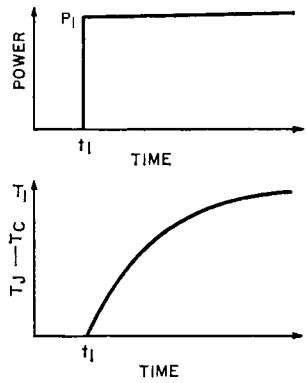



Figure 47. Temperature-rise curve obtained with step function of power.

the voltage-rise curve which would be obtained in the analogous resistance-capacitance electrical circuit.

The exact shape of the curve depends upon the magnitudes of the thermal-resistance and thermal-capacitance components of the device. Fig. 48 shows a typical thermal-response curve for a silicon power transistor. This curve indicates that solid-state devices have multiple thermal time constants.

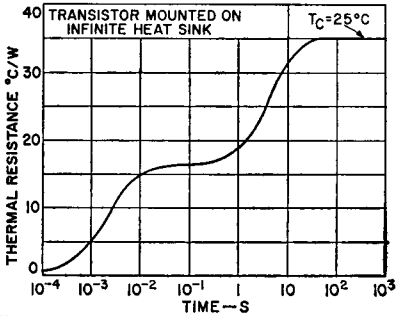



Figure 48. Graphical representation of transient thermal response (i.e., thermal-impedance curve).

## Case-to-Ambient Thermal Resistance

The thermal equivalent circuits for a transistor discussed in the preceding section considered only the thermal paths from junction to case. For power transistors in which the silicon pellet is mounted directly on the header or pedestal, the total internal thermal resistance from junction to case $\theta_{J\text{-}C}$ varies from 50°C per watt to less than 1°C per watt. If the transistor is not mounted on a heat sink, the thermal resistance from case to ambient air $\theta_{C\text{-}A}$ is so large in comparison to that from junction to case that the net over-all thermal resistance from junction to ambient air is primarily the result of the $\theta_{C\text{-}A}$ term. Table III lists values of case-to-air thermal resistance for popular JEDEC cases. Beyond the limit of a few hundred milliwatts, it becomes impractical to increase the size of the case to make the $\theta_{C\text{-}A}$ term comparable to the $\theta_{J\text{-}C}$ term. As a result, most power transistors are designed for use on an external heat sink.

Table III—Case-to-Free-Air Thermal Resistance for Popular JEDEC Cases

| Case | $\theta_{C\text{-}A}$ (°C/W) |
|---|---|
| TO-18 | 300 |
| TO-46 | 300 |
| TO-5 | 150 |
| TO-39 | 150 |
| TO-8 | 75 |
| TO-66 | 60 |
| TO-60 | 70 |
| TO-3 | 30 |
| TO-36 | 25 |

## Case-to-Ambient Thermal Capacitance

The thermal capacitance of the over-all package is also an important factor in the thermal circuit of a solid-state power device. The thermal capacitance of a package is calculated relatively easily. First, the material constituents of the package are determined. The thermal capacitances of these materials are then calculated and added together to obtain the thermal-capacitance value for the over-all package. For example, the RCA copper-button TO-3 package contains about 4 grams of copper and 12 grams of steel. The other constituents of this package have negligible thermal capacitance.

The thermal capacitance of any material is equal to the product of its mass and specific heat. The calculations of the thermal capacitances ($C_{Cu}$ and $C_s$) for the

copper button and the steel case yield the following results:

$$C_{Cu} = 4 \text{ grams} \times 0.093 = 0.37 \text{ cal per } °C$$
$$C_s = 12 \text{ grams} \times 0.105 = 1.2 \text{ cal per } °C$$

The following summation then provides the total thermal capacitance of the over-all package:

$$C_{pkg} = C_{Cu} + C_s = 1.63 \text{ cal per } °C$$

Expressed in terms of watt-seconds per °C, this value becomes

$$C_{pkg} = 1.63 \text{ cal per } °C \times 4.18 \text{ cal per joule} = 6.8 \text{ joules per } °C = 6.8 \text{ watt-seconds per } °C$$

The thermal resistance from case to the ambient air of the copper-button TO-3 package is approximately 20°C per watt. The thermal-cooling time constant for this package is then determined as follows:

$$\Theta_{C-A} \times C_{pkg} = 30°C \text{ per watt} \times 6.8 \text{ watt-seconds per } °C = 204 \text{ seconds}$$

Similar calculations on other types of packages for solid-state devices yield the following values of thermal capacitances and thermal time constants:

| Package | Thermal Capacitance (Joules/°C) | Thermal Time Constant (Seconds) |
|---|---|---|
| TO-5 | 0.58 | 69 |
| TO-66 (no button) | 2.56 | 128 |
| TO-8 | 1.84 | 110 |
| TO-3 (Cu button) | 6.8 | 204 |
| TO-3 (Mod, 2N5575) | 7.8 | 117 |

These values can be used to calculate temperature effects of pulses on devices that are not mounted on heat sinks. The thermal time constants can be used to estimate how long units must be cooled between tests to avoid temperature changes. For example, the application of a 150-watt pulse for 1 second results in a temperature rise in the TO-3 package determined as follows:

$$T_{rise} = 150 \text{ watts}/6.8 \text{ joules } °C = 22°C$$

The time required to cool the package to within 3°C of room temperature can be determined as follows:

$$T = 22e^{-t/204}$$
$$3 = 22e^{-t/204}$$
$$\ln 3/22 = -t/204$$
$$t = 6.1 \text{ minutes}$$

## USE OF EXTERNAL HEAT SINK

In the preceding discussion, the sources of thermal resistance within a solid-state device were explained, and the thermal properties of free-air-mounted devices were described. This section explains the use of heat sinks to increase the power-handling capability of solid-state power devices.

### Effect on Basic Thermal Circuit

The primary purpose of a heat sink is to increase the effective heat-dissipation area. The effect on the thermal equivalent circuit is shown in Fig. 49. From the electrical analog, the effective resistance of the two parallel thermal paths is smaller than that of either of the paths. The effect of the heat sink is to provide an additional low-thermal-resistance path from case to ambient air. The heat-sink thermal resistance actually consists of two

series elements, the thermal resistance from case to heat sink that results from conduction

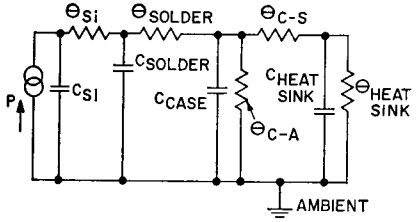
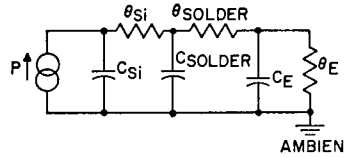



Figure 49. Thermal equivalent circuit for a transistor mounted on a heat sink.

($\theta_{C\text{-}S}$) and the thermal resistance from heat sink to ambient air caused by convection and radiation ($\theta_{S\text{-}A}$).

In practice, the case must be electrically isolated from the heat sink except for grounded-collector circuits. The thermal resistance from case to heat sink, therefore, includes two components. One component is caused by surface irregularities and can be minimized by use of silicone grease compounds; the other component is introduced by the electrical insulating washer required. The thermal capacitance of these two elements is very small and can be neglected.

If the full power-handing capability of a solid-state device, as determined by $\theta_{J\text{-}C}$, is to be realized, there should be no temperature differential between the case and ambient air. This condition can occur only when the thermal resistance of the heat sink is zero, i.e., when the device is mounted on an infinite heat sink. Although an infinite heat sink can never be realized in practice, the greater the ratio $\theta_{J\text{-}C}/\theta_{C\text{-}A}$, the closer is the approximation, and the nearer the maximum power limit defined by $\theta_{J\text{-}C}$ can be approached. When a power transistor is used with a heat sink, the heat loss by convection and radiation through the case is very small compared to the loss through the heat sink. If $\theta_{case}$ and $C_{case}$ are neglected, or at worst combined with $\theta_{heat\ sink}$ and $C_{heat\ sink}$, the thermal equivalent circuit for the transistor can be represented as shown in Fig. 50.




$$C_E = C_{case} + C_{heat\ sink}$$

$$\theta_E = \frac{\theta_{heat\ sink}(\theta_{case})}{\theta_{heat\ sink} + \theta_{case}} \approx \theta_{heat\ sink}$$

Figure 50. Simplified thermal equivalent circuit for a transistor mounted on a heat sink.

## Heat Removal

Heat may be transferred by three basic processes: conduction, convection, and radiation. Each of these processes is used in the removal of heat from silicon power transistors.

**Conduction** is a process of heat transfer in which heat energy is passed from one atom to the next, while the actual atoms involved in the transfer remain in their original positions. If a known amount of power flows through a material, the thermal resistance which may be attributed to conduction is determined by the following equation:

$$\theta_{cond} = d/4.186\ KA\ ^\circ C\ per\ watt \quad (10)$$

where d is the length of the thermal path in centimeters, K is the thermal conductivity in cal/(sec) (cm) (°C), A is the area per-

pendicular to the thermal path t in square centimeters, and the conversion factor 4.186 is given in (watt) (sec)/cal. This equation is merely another form of Eq. (3) in which the conversion factor is used to obtain the result in °C per watt.

**Convection** is a term applied to the transfer of heat by the physical motion of hot material. In forced convection, the medium of heat transfer is moved by a fan. In natural convection, the medium moves because of differences in density. Both forced and natural convection are used for transistor cooling. The following equation defines the thermal resistance of vertical plates freely suspended in free air at ground level:

$$\theta_{conv} = (2300/A)\,(L/T_s - T_{amb})^{\frac{1}{4}} \qquad (11)$$

where A is the total exposed area (twice the area of one side) in square centimeters, $T_s$ is the surface temperature of the heat sink in °C, $T_{amb}$ is the ambient temperature in °C, and L is the height of the heat sink in centimeters.

The third process by which heat may be transferred is **radiation.** The rate of emission from a surface can be found from Stefan's law. In accordance with this law, the equation for radiation thermal resistance may be written as follows:

$$\theta_{rad} = \frac{1793 \times 10^8}{AE\,(T_s^2 - T_{amb}^2)}\,(T_s - T_{amb}) \qquad (12)$$

where A is the total exposed area in square centimeters, E is the emissivity (a function of the surface finish), $T_s$ is the surface temperature in °C, and $T_{amb}$ is the ambient temperature in °C.

## Heat-Sink Requirements

The sources of thermal resistance both internal and external to solid-state devices have been discussed, and the processes which may be used for heat removal have been explained briefly. Solid-state power devices are normally designed to be used with an external heat sink.

Most practical heat sinks used in modern, compact equipment are the result of experiments with heat transfer through convection, radiation, and conduction in a given application. Although there are no set design formulas that provide exact heat-sink specifications for a given application, there are a number of simple rules that reduce the time required to evolve the best design for the job. These simple rules are as follows:

1. The surface area of the heat sink should be as large as possible to provide the greatest possible heat transfer. The area of the surface is dictated by case-temperature requirements and the environment in which the device is to be placed.

2. The heat-sink surface should have an emissivity value near unity for optimum heat transfer by radiation. A value approaching unity can be obtained if the heat-sink surface is painted flat black.

3. The thermal conductivity of the heat-sink material should be such that excessive thermal gradients are not established across the heat sink.

Although these rules are followed in conventional heat-sink systems, the size and cost of such systems often become restrictive

in compact, mass-produced power-control and power-switching applications. The use of mass-produced prepunched parts, direct soldering, and batch-soldering techniques eliminates many of the difficulties associated with heat sinks by making possible the use of a variety of simple, efficient, readily fabricated heat-sink configurations that can be easily incorporated into the mechanical design of equipment.

For most efficient heat sinking, intimate contact should exist between the heat sink and at least one-half of the package base. The package can be mounted on the heat sink mechanically, with glue or epoxy adhesive, or by soldering. (As pointed out in the section on **Packaging, Handling, and Mounting,** soldering is not recommended for transistors.) If mechanical mounting is employed, silicone grease should be used between the device and the heat sink to eliminate surface voids, prevent insulation buildup due to oxidation, and help conduct heat across the interface. Although glue or epoxy adhesive provides good bonding, a significant amount of resistance may exist at the interface resistance; an adhesive material with low thermal resistance, such as Hysol Epoxy Patch Material No. 6C or Wakefield Delta Bond No. 152, or their equivalent, should be used.

## Types of Heat Sinks

Heat sinks are produced in various sizes, shapes, colors, and materials; the manufacturer should be contacted for exact design data. It is convenient for discussion purposes to group heat sinks into three categories as shown below:

1. **Flat vertical-finned types** are normally aluminum extrusions with or without an anodized black finish. They are unexcelled for natural convection cooling and provide reasonable thermal resistance at moderate air-flow rates for forced convection.

2. **Cylindrical or radial vertical-finned types** are normally cast aluminum with an anodized black finish. They are used when maximum cooling in minimum lateral displacement is required, using natural convection.

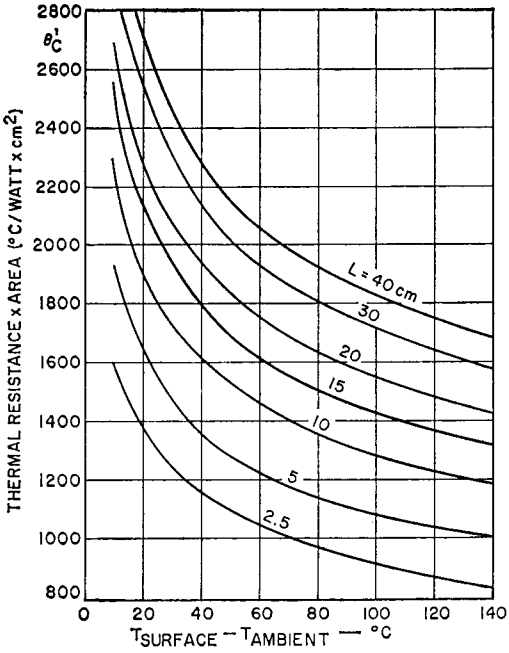
3. **Cylindrical horizontal-finned types** are normally fabricated from sheet-metal rings and have a painted black matte finish. They are used in confined spaces for maximum cooling in minimum displaced volume.

It is also common practice to use the existing mechanical structure or chassis as a heat sink. The design equations and curves for such heat sinks based upon convection and radiation are shown in Figs. 51, 52, and 53.

A useful nomograph which considers heat removal by both convection and radiation is given in Fig. 54. This nomograph applies for natural bright finish on the copper or aluminum.

## Heat-Sink Performance

The performance that may be expected from a commercial heat sink is normally specified by the manufacturer, and the information supplied in the design curves shown in Figs. 51, 52, and 53 provides the basis for the design of flat vertical plates for use as heat

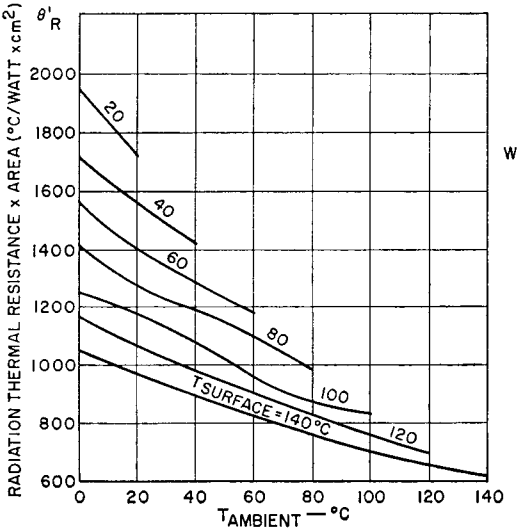



$$\theta_C = \frac{2300}{A} \times \left(\frac{L}{T_s - T_a}\right)^{0.25}$$

$$\theta_C = \frac{\theta'_C}{A} \text{ °C/WATT}$$

WHERE $\theta$ = CONVECTION THERMAL RESISTANCE °C/WATT
A = AREA IN $cm^2$, TOTAL EXPOSED SURFACE
L = HEIGHT IN cm

Figure 51. Convection thermal resistance as a function of temperature drop from the surface of the heat sink to free air for heat sinks of various heights. (Reprinted from **Control Engineering,** October 1956.)




$$\theta_R = \frac{1793 \times 10^8}{Ae\,(T_s^2 + T_a^2)(T_s + T_a)}$$

$$= \frac{\theta'_R}{Ae} \text{ °C/WATT}$$

WHERE A = TOTAL EXPOSED AREA, $cm^2$
e = EMISSIVITY
$T_s$ = SURFACE TEMP, °C
$T_a$ = AMBIENT TEMP, °C
$\theta_R$ = RADIATION THERMAL RESISTANCE

Figure 52. Radiation thermal resistance as a function of ambient temperature for various heat-sink surface temperatures. (Reprinted from **Control Engineering,** October 1956.)

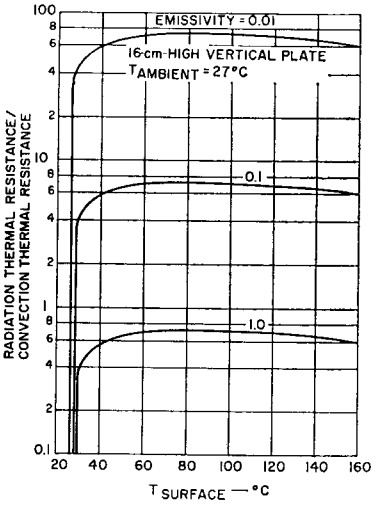



Figure 53. Ratio of radiation thermal resistance to convection thermal resistance as a function of heat-sink surface temperature for various surface emissivities. (Reprinted from **Control Engineering**, October 1967.)

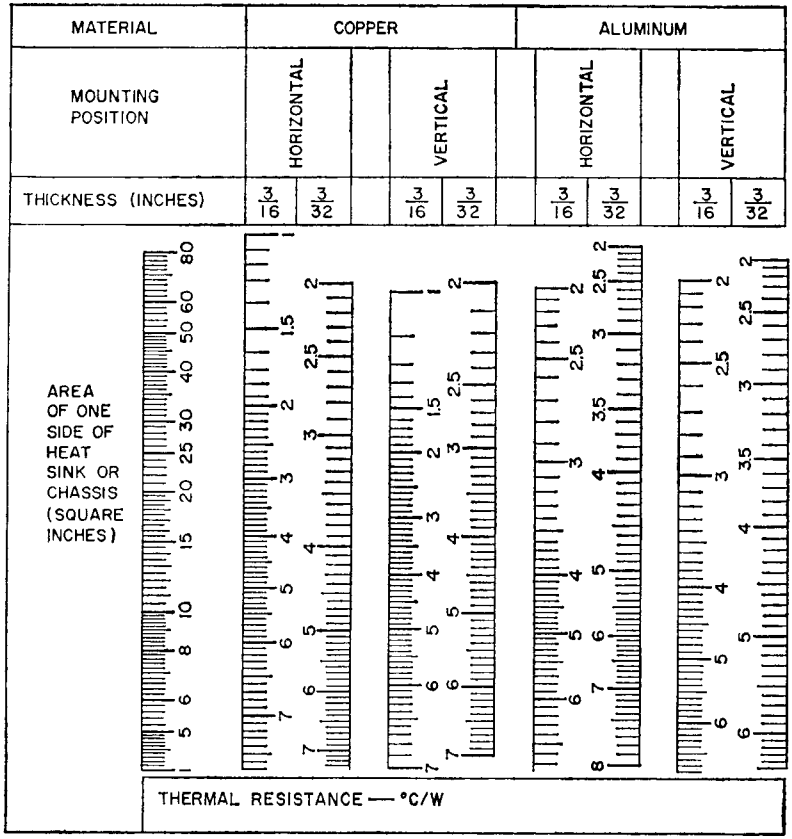
sinks. In all cases, it must be remembered that the heat is dissipated from the heat sink by both convection and radiation. Although surface area is important in the design of vertical-plate heat sinks, other factors such as surface and ambient temperatures, conductivity, emissivity, thickness, shape, and orientation must also be considered. An excessive temperature gradient can be avoided and the conduction thermal resistance in the heat sink can be minimized by use of a high-conductivity material, such as copper or aluminum, for the heat sink. Radiation losses are increased by an increase in surface emissivity, as shown in Fig. 54. Best results are obtained when the heat sink has a black matte finish for which the emissivity is at least 0.9. When free-air convection is used for heat removal, a vertically mounted heat sink provides a thermal resistance that is approximately 30 per cent lower than that obtained with horizontal mounting.

In restricted areas, it may be necessary to use forced-convection cooling to reduce the effective thermal resistance of the heat sink. On the basis of the improved reliability of cooling fans, it can be shown that the over-all reliability of a system may actually be improved by use of forced-convection cooling because the number of components required is reduced.

Economic factors are also important in the selection of heat sinks. It is often more economical to use one heat sink with several properly placed transistors than to use individual heat sinks. It can be shown that the cooling efficiency increases and the unit cost decreases under such conditions.

## Heat-Sink Insulators

As pointed out previously, when solid-state devices are to be mounted on heat sinks, some form of electrical isolation must be provided between the case and the heat sink. Unfortunately, however, good electrical insulators usually are also good thermal insulators. It is difficult, therefore, to provide electrical insulation without introduction of significant thermal resistance between case and heat sink. The best materials for this application are mica, beryllium oxide (Beryllia), and anodized aluminum. A comparison of the properties of these three materials for case-to-heat-sink isolation of the TO-3 package is shown in Table IV. If the area of the seating plane, the thickness of the material, and the thermal conductivity are known, the case-to-heat-sink thermal resistance




INSTRUCTION FOR USE: SELECT THE HEAT-SINK AREA AT LEFT AND DRAW A HORIZONTAL LINE ACROSS THE CHART FROM THIS VALUE. READ THE VALUE OF MAXIMUM THERMAL RESISTANCE DEPENDING ON THE THICKNESS OF THE MATERIAL, TYPE OF MATERIAL, AND MOUNTING POSITION.

Figure 54. Thermal resistance as a function of heat-sink dimensions. (Nomograph reprinted from **Electronic Design**, August 16, 1961.)

$\theta_{C\text{-}S}$ can be readily calculated by use of Eq. (10). In all cases, this calculation should be experimentally verified. Irregularities on the bottom of the transistor seating plane or on the face of the heat sink or insulating washer may result in contact over only a very small area unless a filling compound is used. Although silicone grease has been used for years, recently newer compounds with zinc oxide fillers (e.g., Dow Corning #340 or Wakefield #120) have been found to be even more effective.

Table IV—Comparison of Insulating Washers Used for Electrical Isolation of Transistor TO-3 Case from Heat Sink

| Material | Thickness (inches) | $\theta_{C\text{-}S}$ (°C/W) | Capacitance (pF) |
|---|---|---|---|
| Mica | 0.002 | 0.4 | 90 |
| Anodized Aluminum | 0.016 | 0.35 | 110 |
| Beryllia | 0.063 | 0.25 | 15 |

For small general-purpose transistors, such as the 2N2102, which use a JEDEC TO-5 package, a

good method for thermal isolation of the collector from a metal chassis or printed-circuit board is by means of a beryllium-oxide washer. The use of a zinc-oxide-filled silicone compound between the washer and the chassis, together with a moderate amount of pressure from the top of the transistor, helps to decrease thermal resistance. Fin-type heat sinks, which are commercially available, are also suitable, especially when transistors are mounted in Teflon sockets which provide no thermal conduction to the chassis or printed-circuit board. Fig. 55 illustrates both types of mounting.

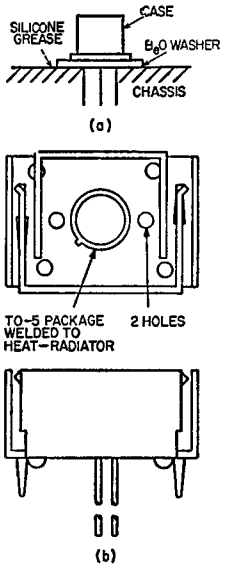


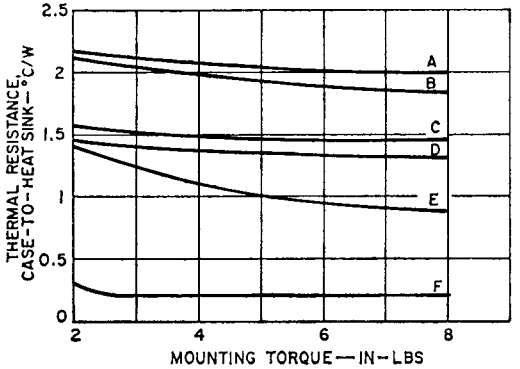

Figure 55. Suggested mounting arrangements for transistors having a JEDEC TO-5 package: (a) without heat sink; (b) with fin-type heat sink.

At frequencies of 100 MHz and higher, the effects of stray capacitances and inductances and of ground paths and feedback coupling have a pronounced effect on the gain and power-output capabilities of transistors. As a result, physical aspects such as mechanical layout, shielding, and heat-sink considerations are important in the design of rf amplifiers and oscillators. In particular, it should be noted that the insulating washer necessary for isolation introduces coupling capacitance from collector to chassis which may seriously limit circuit performance.

## Special Considerations for Plastic-Package Types

Fig. 56 shows a set of curves of typical case-to-heat-sink thermal resistance of the plastic VERSAWATT transistor as a function of mounting torque for several mounting arrangements. Curves A through D show typical case-to-heat-sink thermal resistance for the mounting arrangements shown in Figs. 42(a) through 42(d) in the section on **Packaging, Handling, and Mounting.** Curves E and F are representative of a VERSAWATT transistor mounted over a heat-sink mounting hole that has a diameter of 0.140 inch (No. 6 screw clearance). Curve E shows the wide variation in thermal resistance with torque when the transistor is mounted dry. Curve F shows the effect on contact thermal resistance of a thin layer of Dow Corning No. 340 silicone grease applied between transistor and heat sink. For torques within the recommended range of 4 to 8 inch-pounds, contact thermal resistance is reduced to between 18 and 25 per cent of the dry values.

The curves shown in Fig. 57 represent typical case-to-heat-sink thermal resistance of the high-power molded-plastic transistor package as a function of




| CURVE | MOUNTING ARRANGEMENT FIGURE | HEAT SINK HOLE DIA. (IN.) | MICA THICKNESS (MILS) | THERMAL COMPOUND |
|---|---|---|---|---|
| A | 41(a) | 0.250 | 4 | Dow Corning No. 340 |
| B | 41(b) | 0.113 | 4 | Dow Corning No. 340 |
| C | 41(a) | 0.250 | 2 | Dow Corning No. 340 |
| D | 41(b) | 0.113 | 2 | Dow Corning No. 340 |
| E | — | 0.140 | None | None |
| F | — | 0.140 | None | Dow Corning No. 340 |

Figure 56. Typical case-to-heat-sink thermal resistance as a function of mounting torque for an RCA VERSAWATT transistor.

mounting torque. The thermal resistances shown by curves A and C are representative of the mounting arrangements shown in Fig. 43(a) through 43(d) in the section on **Packaging, Handling, and Mounting.** Curves B and D are typical for mounting without mica over heat-sink mounting holes that have a diameter of 0.113 inch (No. 4 screw clearance). The effect of a thin layer of silicone grease on contact thermal resistance is illustrated by a comparison of curves B and D.

Operation of the transistor with heat-sink temperatures of 100°C or greater results in some shrinkage of the insulating bushing normally used to mount power transistors. The degradation of contact thermal resistance (refer to Figs. 56 and 57) is usually less than 25 per cent if a good thermal compound is used.

During the mounting of RCA molded-plastic solid-state power devices, the following special precautions should be taken to assure efficient heat transfer from case to heat sink:

1. Mounting torque should be between 4 and 8 inch-pounds.
2. The mounting holes should be kept as small as possible.
3. Holes should be drilled or punched clean with no burrs or ridges, and chamfered to a maximum radius of 0.010 inch.
4. The mounting surface should be flat within 0.002 inch/inch.
5. Thermal grease (Dow Corning 340 or equivalent) should always be used (on both sides of

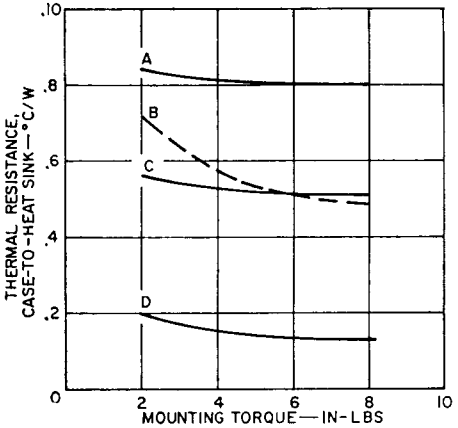
the insulating washer if one is employed).

6. Thin insulating washers should be used (thickness of factory-supplied mica washers ranges from 2 to 4 mils).

7. A lock washer or torque washer should be used, together with materials that have sufficient creep strength to prevent degradation of heat-sink efficiency during life.

which the pellet is attached. The cyclic stresses may eventually result in physical damage to the semiconductor pellet or the mounting interface.

In most solid-state power devices, a small silicon pellet is bonded to a copper header. The coefficient of thermal expansion for silicon ($3 \times 10^{-6}$) is much less than that of copper ($17.5 \times 10^{-6}$). Temperature variations




| CURVE | MOUNTING ARRANGEMENT FIGURE | MICA THICKNESS (MILS) | THERMAL COMPOUND |
|---|---|---|---|
| A | 43(a) thru 43(d) | 4 | Dow Corning No. 340 |
| B | — | None | None |
| C | 43(a) thru 43(d) | 2 | Dow Corning No. 340 |
| D | — | None | Dow Corning No. 340 |

Figure 57. Typical case-to-heat sink thermal resistance as a function of mounting torque for an RCA high-power plastic-package transistor.

## EFFECT OF CYCLIC THERMAL STRESSES

When a solid-state device is alternately heated and allowed to cool, cyclic mechanical stresses are produced within the device because of differences in the thermal expansion of the silicon pellet and the metallic materials to within the transistor, therefore, result in cyclic stresses at the mounting interface of the silicon pellet and the copper header because of the difference in the thermal expansions of these parts. If a hard solder, such as silicon gold, is used to bond the pellet to the header, these stresses are transmitted to the silicon pel-

let. Such stresses often result in pellet fractures. In general, however, lead-tin solder is used to bond the silicon pellet to the copper header. The cyclic thermal stresses then are absorbed by non-elastic deformation of the soft lead solder, and very little stress is transmitted to the pellet.

The continuous flexing that results from cyclic temperature changes may eventually cause fatigue failures in a conventional lead solder system. Such failures are a function of the amount of change in temperature at the mounting interface, the difference in the thermal-expansion coefficients of the silicon pellet and the material to which the pellet is attached, and the maximum dimensions of the mounting interface. Fatigue failures occur whenever the cyclic stresses damage the solder to the point at which the transfer of heat between the pellet and the surface to which it is mounted becomes impaired. This condition, which is indicated by a significant rise in junction-to-case thermal resistance, may exist in only a small portion of the pellet. This portion, however, overheats, and device failure results because of regenerative conditions that lead to thermal runaway.

Thermal-fatigue failures in power transistors are accelerated because of dislocation "pile-ups" that result from impurities in the lead solder. RCA has developed a process that substantially reduces the amount of impurities introduced into the solder. Use of this proprietary "**Controlled Solder Process**" makes it possible to avoid the microcracks that propagate to cause fatigue failure in power transistors and, therefore, greatly increases the thermal-cycling capability of these devices.

The mathematical relationship among the factors that affect fatigue failure in silicon power transistors can be expressed, in terms of the number of thermal cycles to failure N, as follows:

$$N = Ae^{\psi_0/[\Delta T(\alpha_A - \alpha_B)L]} \qquad (13)$$

where A is a constant determined by the mounting system, $\Delta T$ is the change in temperature at the mounting interface, $\alpha_A$ and $\alpha_B$ are the thermal-expansion coefficients of the silicon and the metal under the solder joint, $\psi_0$ is a material constant proportional to the change in temperature $\Delta T$ and the difference in the thermal-expansion coefficients $\alpha_A$ and $\alpha_B$, and L is the maximum length of the solder joint under the pellet.

For a given transistor, the only variable in the thermal-cycling equation that can be controlled by the circuit designer is the change in temperature at the interface of the silicon pellet and the material to which the pellet is mounted. This change in temperature $\Delta T$ is, of course, less than the change in transistor junction temperature $\Delta T_J$, but is greater than the change in case temperature $\Delta T_C$.

RCA has devised a rating chart that relates the thermal-cycling capability of a silicon power transistor to total device dissipation and the change in case temperature. (This rating system is described in the section on **Physical Basis for Power-Transistor Ratings**.)

A circuit designer may use the rating system to define the limiting value to which the change in case temperature must be restricted to assure that a power

transistor is capable of operation at a specified power dissipation over the number of thermal cycles required in a given application. Conversely, if the power dissipation and the change in case temperature are known, the designer may use the rating system to determine whether the thermal-cycling capability of the transistor is adequate for the application.

# Silicon Rectifiers

SILICON rectifiers are essentially cells containing a simple p-n junction. As a result, they have low resistance to current flow in one (forward) direction, but high resistance to current flow in the opposite (reverse) direction. They can be operated at ambient temperatures up to 200°C, current levels as high as hundreds of amperes, and voltage levels greater than 1000 volts. In addition, they can be used in parallel or series arrangements to provide higher current or voltage capabilities.

Because of their high forward-to-reverse current ratios, silicon rectifiers can achieve rectification efficiencies greater than 99 per cent. When properly used, they have excellent life characteristics which are not affected by aging, moisture, or temperature. They are very small and light-weight, and can be made impervious to shock and other severe environmental conditions.

## THEORY OF OPERATION

The operation of a silicon rectifier can be conveniently explained by analysis of the flow of charge carriers across the p-n junction under both forward- and reverse-bias conditions. Alternatively, an analysis of the potential distribution in the junction for each bias condition may be used to predict the behavior of the rectifier.

In a silicon rectifier, the regions adjacent to the metal contacts are heavily doped, one with p-type dopant and the other with n-type dopant, to ensure that nonrectifying ohmic contacts are formed at the silicon-to-metal interfaces. A rectifying junction should exist only within the silicon, at the interface of the n-type and p-type regions. A lightly doped n-type region between the heavily doped n- and p-type regions provides the high blocking-voltage capability required of the rectifier. Because of this lightly doped region, the more heavily doped n-type region adjacent to the metal contact is referred to as the $n^+$ region. The silicon rectifier, therefore, is a p-n-$n^+$ structure.

### Carrier-Flow Analysis

The theory of operation of p-n-$n^+$ silicon junctions can be visualized by use of the diagrams shown in Fig. 58. In these dia-

grams, free electrons are represented by dots and free holes by circles; the movements of electrons and holes are indicated by arrows. Fig. 58(a) shows the junction diagram for the p-n-n + structure.

**Equilibrium Condition**—In equilibrium, as shown in Fig. 58(b), each region of the crystal contains approximately the same number of free electrons or free holes as the amount of **donor** impurities or **acceptor** impurities, respectively. The p-type region contains only holes, the n-type region contains only electrons, and the metal contacts contain both holes and electrons. The nature of the metal-to-semiconductor ohmic contact is such (as explained later) that only electrons can go from the metal into the n-type semiconductor, and only holes can go from the metal to the p-type semiconductor. The resulting behavior under forward- or reverse-bias conditions is as follows:

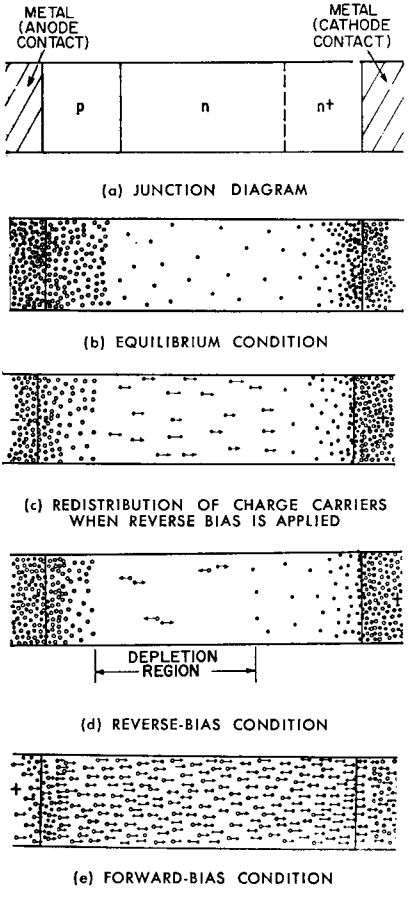



Figure 58. Concentrations of electrons (dots) and holes (circles) in a silicon rectifier.

**Reverse-Bias Condition**—When a reverse bias is applied (positive voltage to the n-type region and negative voltage to the p-type region), a nonequilibrium distribution of holes and electrons occurs because a region around the p-n junction is depleted of free charge carriers. This redistribution occurs because electrons are attracted by the positive voltage applied to the n-type region and holes are attracted by the negative voltage applied to the p-type region so that they are displaced from the equilibrium positions, as shown in Fig. 58(c). The net result is that carriers move away from both sides of the junction to create a **depletion region** or **space-charge region** which can withstand the applied voltage without further current flow, as shown in Fig. 58(d). Only a very small leakage current flows because, as noted above, holes from the metal cannot enter the n-type region and electrons from the metal cannot enter the p-type region. This leakage current can be attributed to thermal generation of electron-hole pairs within the depletion layer, as indicated in Fig. 58(d).

**Forward-Bias Condition**—The junction is forward-biased when a positive voltage is applied to the p-type region and a negative voltage is applied to the n-type region. This bias causes holes and electrons to move toward and across the junction. As a result, the concentration of free charge carriers in the central region of the junction is greatly increased, as shown in Fig. 58(e). Holes from the left metal contact can freely enter the p-type region, and electrons from the opposite metal contact can freely enter the n-type region. An abundant supply of holes and electrons is available, therefore, to replace those that move across the junction.

## Potential-Hill Analysis

The operation of p-n-n$^+$ silicon junctions may also be visualized in terms of the potential-energy diagrams shown in Fig. 59. In these diagrams, the vertical scale represents energy. An increase in electron energy is indicated by the upward direction from the Fermi energy level ($E_F$ line), and an increase in hole energy is indicated by the downward direction from this level. Electrons are always above the $E_F$ line and holes are always below this line, which represents the ground state or zero-energy level for both types of carriers. Both electrons and holes tend to "fall" toward this level unless there is some source of energy to move them away from it.

**Equilibrium Condition**—Thermal energy from the silicon crystal is one source of energy that normally causes some of the carriers to be displaced above and below the Fermi level, as shown in Fig. 59(b).

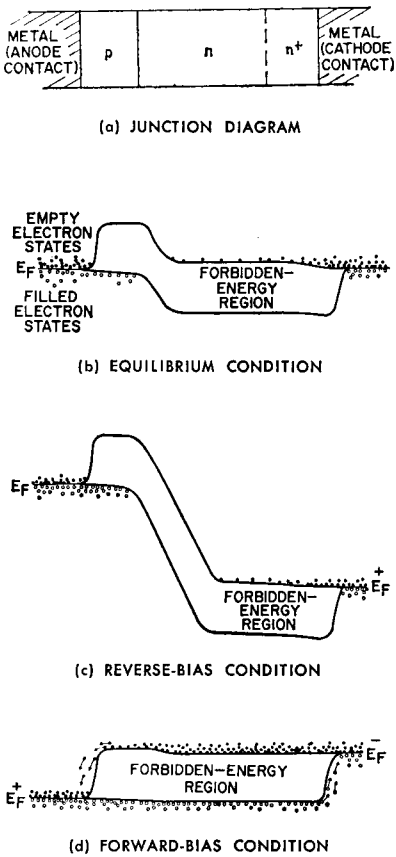



Figure 59. Potential-hill diagrams for various stages of rectifier operation (upward direction indicates increasing electron energy; downward direction indicates increasing hole energy).

In the metal contacts, holes and electrons exist side by side because there is no **forbidden-energy region.** In the semiconductor material, however, the Fermi level lies within a forbidden-energy region which cannot be penetrated by holes or electrons. In the n-type semiconductor, the Fermi level lies near the top of the forbidden-energy region, and there is ample space for free electrons to move about. There are no holes

in the n-type region, however, because more energy is required to force the holes below the forbidden-energy region than can be supplied by the thermal effects. Similarly, in the p-type region, the Fermi level lies close to the bottom of the forbidden-energy region. The holes, therefore, can easily obtain enough thermal energy to get below this region, but electrons cannot obtain enough energy to get above it. As a result, only holes can enter the p-type region from the metal, and only electrons can enter the n-type region from the metal. Holes can freely circulate between the metal and the p-type region, but electrons are excluded. Electrons can freely circulate between the metal and the n-type region, but holes are excluded.

In visualizing the operation of a silicon rectifier by use of the potential-hill diagrams shown in Fig. 59, the following factors must be considered:

1. The shape of the forbidden-energy region is rigid at the metal-to-semiconductor contact. The shape is determined by the **doping level** (or **carrier concentration**), which is extremely high at the contacts and cannot, therefore, be changed by the carriers injected or removed by applied voltage or current.

2. The shape of the forbidden-energy region is flexible at the p-n junction because the carrier concentration at the junction is quite low and can be readily influenced by addition or removal of carriers by means of an applied bias.

The behavior under forward- and reverse-biased conditions may then be explained as follows:

**Reverse-Bias Conditions**—Under reverse-bias conditions, the potential energy of electrons is increased on the negatively biased side of the junction so that the energy at this end is higher, as shown in Fig. 59(c). Although the applied bias is such that it tends to push electrons from the metal into the p-type region and holes from the metal into the n-type region, no current flows because the rigidity of the forbidden-energy region at the contacts prevents such movements of the charge carriers. The applied voltage simply increases the height of the potential hill at the junction because there are no carriers available to move in the direction that the field would cause them to move. On both sides, the carriers have an "uphill" climb to the junction.

**Forward-Bias Conditions**—The application of a positive voltage to the p-type region and a negative voltage to the n-type region raises the electrons to a higher potential energy on the n-type side of the junction, as shown in Fig. 59(d). This bias must alter the shape of the forbidden-energy region so that its ends meet the changed energy levels of the metals. Because the shape is flexible only at the junction, the applied bias causes the profile of the forbidden-energy region to be altered, as shown in Fig. 59(d), to reduce the height of the built-in potential hill. As a result, many electrons now have sufficient thermal energy to get over the hill, and many holes have sufficient thermal energy to get under it. Because the height of the hill is equivalent to about one electron-volt, a forward bias of one volt is sufficient to allow electrons and

holes to move unimpeded across the junction; the current is then limited only by the ohmic resistance of the external circuit.

## THERMAL CONSIDERATIONS

Although silicon rectifiers can operate at high temperatures, the actual pellet of silicon which performs the rectification is quite small and has a very low thermal capacity. During normal operation, the rectifier p-n junction dissipates approximately 1 watt of power for each ampere of forward current. The temperature of the junction rises rapidly during high-current operation. An increase in junction temperature beyond rated capabilities, as a result of either high currents or excessive ambient temperatures, may cause rectifier failure, either directly because of irreversible material damage as a result of the high temperature or indirectly because of the effect of the increased temperature on the reverse-blocking capability of the rectifier, as described later. The heat dissipated in the silicon pellet must be removed rapidly, therefore, so that the temperature of the junction is not allowed to rise above the safe operating value of 200°C. For this reason, the silicon pellet is mounted between heavy copper parts in a symmetrical direct-soldered arrangement that results in uniform distribution of thermal stresses, minimum thermal fluctuations, and low thermal resistance.

Fig. 60 shows a cross-sectional diagram of a typical silicon rectifier. Because of the way in which the rectifier is constructed, there is always a thermal "drop" between the p-n junction and the

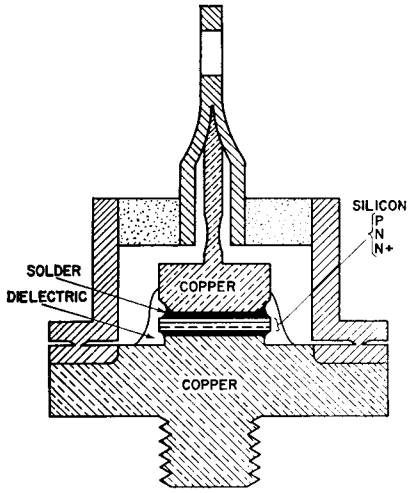



Figure 60. Cross-sectional diagram of a typical silicon rectifier.

outside of the rectifier case. This thermal "drop", which is analogous to the voltage drop across the various components of an electrical circuit, is caused by the **thermal impedances** of the various components of the internal rectifier structure. These impedances include both **thermal resistance** and **thermal capacitance.** The lower side of the silicon pellet is soldered directly to a heavy copper stud that provides a low-thermal-resistance path between the pellet and the rectifier heat sink. The upper side of the pellet is soldered to a heavy copper block which, together with the stud, forms a thermal capacitor.

During long periods of steady-state operation, the thermal capacitance becomes fully charged and does not affect the operation of the rectifier. For this reason, thermal-capacitance values are not included in manufacturers' specifications on silicon rectifiers. It is important, however, that the specifications include the thermal

resistance, expressed in °C per watt, because this value is used, together with the power dissipated by the rectifier, to determine the rise in junction temperature above the case temperature.

The thermal capacitance incorporated into the rectifier structure becomes extremely important when the rectifier junction is subjected to sudden changes in current, such as may occur during a fault condition. This capacitance absorbs heat produced by high-current pulses and allows the heat to flow through the pellet and stud (low-thermal-resistance path) during periods of low current. In this way, fluctuations in junction temperature are held to a minimum. (A more detailed discussion of the thermal impedances of solid-state devices is given in the section on **Thermal Factors.**)

## ELECTRICAL CHARACTERISTICS

Fig. 61 shows the basic current-voltage characteristic for a silicon rectifier. As explained in the charge-carrier and potential-hill analyses given previously, the forward current is many times larger than the reverse current over the normal operating range of the rectifier. The small reverse (leakage) current gradually rises with an increase in reverse voltage. This increase in reverse current eventually leads to junction breakdown, as indicated by an abrupt increase in reverse current at high reverse voltages. Another important feature of the rectifier characteristic is that the forward voltage drop remains small up to the maximum rated current. The basic characteristic curve shown in Fig. 61 serves as a model in the development of the characteristics data given in the manufacturer's specifications on silicon rectifiers.

Characteristics data given for silicon rectifiers are based on the manufacturer's determination of the inherent qualities and traits of the device. These data, which are usually obtained by direct measurements, provide information that a circuit designer needs to predict the performance capabilities of his circuit and form the basis for the ratings that define the safe operating limits for the rectifier.

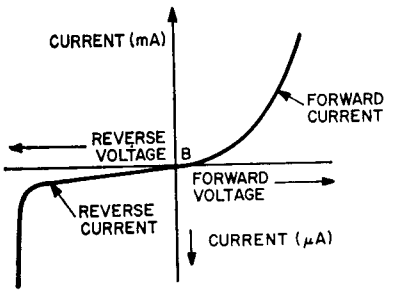



Figure 61. Current-voltage characteristic of a silicon rectifier.

### Forward Voltage Drop

The major source of power loss in a silicon rectifier arises from the forward-conduction voltage drop. This characteristic, therefore, is the basis for many of the rectifier ratings.

A silicon rectifier usually requires a forward voltage of 0.4 to 0.8 volt, depending upon the temperature and impurity concentration of the p-n junction, before a significant amount of current flows through the device. As shown

in Fig. 62, a slight rise in the forward voltage beyond this point causes a sharp increase in the forward current. The slope of the voltage-current characteristic at voltages above this threshold value

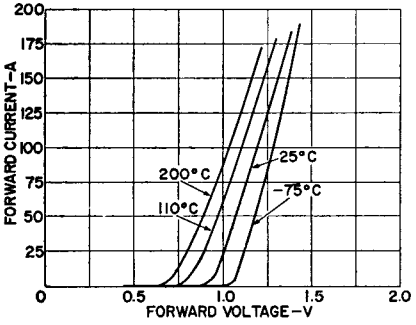



Figure 62. Typical forward characteristics of a silicon rectifier.

represents the **dynamic resistance** of the rectifier. Losses that result from this resistance characteristic increase as the square of the current and thus increase rapidly at high current levels. The dynamic resistance is dependent upon the construction of the rectifier junction and is inversely proportional to the area of the silicon pellet.

Fig. 62 also shows that, at any reasonable current level, the value of forward voltage required to initiate current flow through the rectifier decreases as the temperature of the rectifier junction increases. This voltage-temperature dependence has a compensatory effect in rectifiers operated at high currents, but it is a source of difficulty when rectifiers are operated in parallel.

## Reverse Current

When a reverse-bias voltage is applied across a silicon rectifier, a limited amount of reverse-blocking current flows through the rectifier. This current is in the order of only a few microamperes, as compared to the milliamperes or amperes of forward current produced when the rectifier is forward-biased. Initially, as shown in Fig. 63, the reverse current increases slightly as the blocking voltage increases, but then tends to remain relatively constant, even though the blocking voltage is increased significantly. The figure also indicates that an increase in operating temperature causes a substantial increase in reverse current for a given reverse voltage. Reverse-blocking thermal runaway may occur because of this characteristic if the reverse dissipation becomes so large that, as the junction temperature rises, the losses increase faster than the rate of cooling.

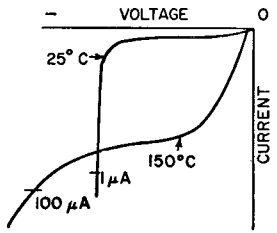



Figure 63. Typical reverse characteristics of a silicon rectifier.

If the reverse blocking voltage is continuously increased, it eventually reaches a value (which varies for different types of silicon rectifiers) at which a very sharp increase in reverse current occurs. This voltage is called the **breakdown** or **avalanche** (or zener) voltage. Although rectifiers can operate safely at the avalanche point, the rectifier may be destroyed as a result of thermal

runaway if the reverse voltage increases beyond this point or if the temperature rises sufficiently (e.g., a rise in temperature from 25°C to 150°C increases the current by a factor of several hundred).

### Reverse Recovery Time

After a silicon rectifier has been operated under forward-bias conditions, some finite time interval (in the order of a few microseconds) must elapse before it can return to the reverse-bias condition. This reverse-recovery time is a direct consequence of the greatly increased concentration of charge carriers in the central region that occurs during forward-bias operation. If the bias is abruptly reversed, some of these carriers abruptly change direction and move out in the reverse direction, and the remainder recombine with opposite-polarity types. Because there is a finite number of these carriers in the central region, and there is no source of additional charge carriers to replace those that are removed, the device will eventually go into the reverse-bias condition. During the removal period, however, the charge carriers constitute a reverse current known as the **reverse-recovery current.**

Fig. 64 shows the current waveform obtained when a sinusoidal voltage is applied across a silicon rectifier. During the positive alternation of the input voltage, the rectifier conducts and accumulates stored charge. When the supply voltage reverses polarity, the reverse recovery current flows through the rectifier until all the stored charge is removed.

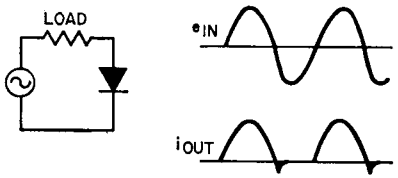



Figure 64. Test circuit and output current waveform obtained when a sinusoidal voltage is applied across a silicon rectifier.

The reverse-recovery time imposes an upper limit on the frequency at which a silicon rectifier may be used. Any attempt to operate the rectifier at frequencies above this limit results in a significant decrease in rectification efficiency and may also cause severe overheating and resultant destruction of the rectifier because of power losses during the recovery period.

## MAXIMUM RATINGS

Ratings for silicon rectifiers are determined by the manufacturer on the basis of extensive testing. These ratings express the manufacturer's judgment of the maximum stress levels to which the rectifiers may be subjected without endangering the operating capability of the unit. The various factors for which silicon rectifiers must be rated include: peak reverse voltage, forward current, surge (or fault) current, operating and storage temperatures, amperes squared-seconds, and mounting torque.

### Peak Reverse Voltage

Peak reverse voltage (PRV) is the rating used by the manufacturer to define the maximum allowable reverse voltage that can be applied across a rectifier. This rating is less than the avalanche

breakdown level on the reverse characteristic. With present-day diffused junctions, the power dissipation at peak reverse voltage is a small percentage of the total losses in the rectifier for operation at the maximum rated current and temperature levels. The reverse dissipation may increase sharply, however, as temperature or blocking voltage is increased to a point beyond that for which the device is capable of reliable operation. It is important, therefore, to operate within ratings.

A transient reverse voltage rating may be assigned when it has been determined that increased voltage stress can be withstood for a short time duration provided that the device returns to normal operating conditions when the overvoltage is removed. This condition is illustrated in Fig. 65.

Peak-reverse voltage ratings for single-junction silicon rectifiers range from 50 to 1500 volts and for multiple-junction silicon-rectifier stacks may be as high as several hundred thousands of volts.

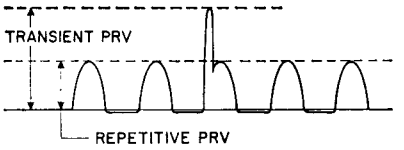



Figure 65. Typical waveform of repetitive and transient reverse voltages applied across a silicon rectifier.

## Forward Current

The current rating assigned to a rectifier is expressed as a maximum value of forward current at a specific case temperature. For these conditions, the power dissipation and internal temperature gradient through the thermal impedance from junction to case are such that the junction is at or near the maximum operating temperature for which the blocking-voltage rating can be maintained. At current levels above this maximum rating, the internal and external leads and terminals of the device may experience excessive temperatures, regardless of the heat sink provided for the pellet itself. The current rating can be described more fully in the form of a curve such as that shown in Fig. 66.

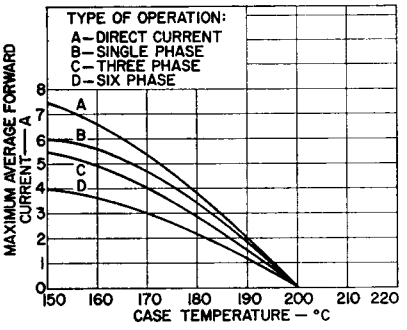



Figure 66. Current rating chart for a 12-ampere silicon rectifier.

Because the current through a rectifier is not normally a smooth flow, current ratings are usually expressed in terms of average current ($I_{avg}$), peak current ($I_{pk}$), and rms current ($I_{rms}$). Each of these currents may be expressed in terms of the other two currents.

The **average current** through a rectifier in half-sine-wave service is related to the peak current by the following equation:

$$I_{avg} = \left[ \frac{\int_0^{\pi} I_{pk} \sin \omega t \, d(\omega t)}{2\pi} \right]$$

$$= I_{pk}/\pi \qquad (14)a$$

or

$$I_{pk} = \pi I_{avg} \qquad (14)b$$

The relationship between the **peak current** and **rms current** of a rectifier in half-sine-wave service can be expressed as follows:

$$I_{rms} = \left[ \frac{\int_0^{\pi} I_{pk}^2 \sin^2 \omega t \, d(\omega t) + \int_{\pi}^{2} 0 \, d(\omega t)}{2\pi} \right]^{\frac{1}{2}} = \tfrac{1}{2} I_{pk} \quad (15a)$$

or

$$I_{pk} = 2 I_{rms} \quad (15b)$$

Table V summarizes the relationships expressed by Eqs. (14) and (15). As discussed later, certain of these relationships are used to determine the power dissipated in a rectifier. The relationships for average, peak, and rms currents are applicable only when the rectifier is used in half-sine-wave service.

Table V—Relationship of $I_{avg}$, $I_{rms}$, and $I_{pk}$

| |
|---|
| $I_{pk} = \pi I_{avg} = 3.14 I_{avg}$ |
| $I_{avg} = (1/\pi) I_{pk} = 0.32 I_{pk}$ |
| $I_{pk} = 2 I_{rms}$ |
| $I_{rms} = \tfrac{1}{2} I_{pk}$ |
| $I_{avg} = (2/\pi) I_{rms} = 0.64 I_{rms}$ |
| $I_{rms} = (\pi/2) I_{avg} = 1.57 I_{avg}$ |

Published data for rectifiers usually list maximum limits for average current and for repetitive peak current. The **maximum average forward-current rating** is the maximum average value of current that is allowed to flow through the rectifier in the forward direction under stated conditions. The **repetitive peak forward-current rating** is the maximum instantaneous value of repetitive forward current permitted under stated conditions. The dual maximum ratings are required because, under certain conditions (e.g., when a highly capacitive load is used), it is possible for the average current to be low and for the peak current to be high enough to cause overheating of the rectifier. The approximate expression for power losses P in a silicon rectifier, given by the following equation, can be used to explain how this type of operation is possible:

$$P_{(watts)} = (V_{dc} I_{dc}) + (I_{rms}^2 R_{dyn}) \quad (16)$$

where the voltage $V_{dc}$ is 0.4 to 0.8 volt depending upon the junction temperature; the direct current $I_{dc}$ is equivalent to the average current $I_{avg}$; the current $I_{rms}$ is the true rms current and, for a fixed average current, increases as the peak current increases; and $R_{dyn}$ is the dynamic resistance of the rectifier over the current range considered.

An analysis of Eq. (16) shows that if the peak current is increased and the conduction time is decreased so that the average current is held constant, the rms current and, therefore, the power dissipated in the rectifier ($I_{rms}^2 R_{dyn}$) are also increased. This behavior explains why the maximum permissible value of average current in multiple-phase circuits is reduced as the number of phases is increased and the conduction period is reduced. Fig. 66 shows the effect of the number of phases on the variation in average current with case temperature.

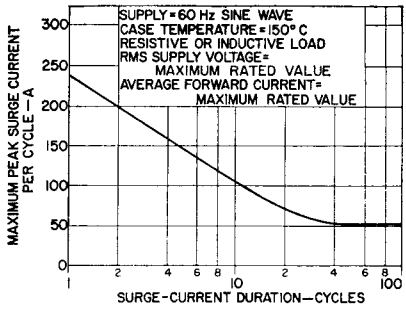
## Surge Current

A third maximum-current limit given in the manufacturer's data

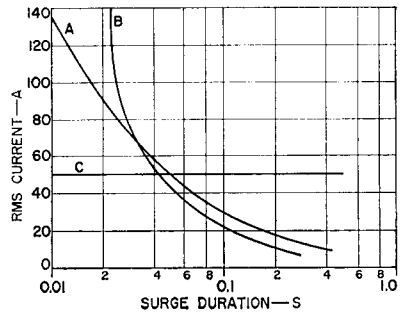
on silicon rectifiers is the surge (or **fault**) current rating. During operation, unusually high surges of current may result from in-rush current at turn-on, load switching, and short circuits. A rectifier can absorb a limited amount of increased dissipation that results from short-duration high surges of current without any effect except a momentary rise in junction temperature. If the surges become too high, however, the temperature of the junction may be raised beyond the maximum capability of the device. The rectifier may then be driven into thermay runaway and, consequently, be destroyed. Fig. 67(a) shows a typical surge-current rating curve for a silicon rectifier.

If the value and duration of anticipated current surges exceed the rating of the rectifier, impedance may be added to the circuit to limit the magnitude of the surge current, or fuses may be used to limit the duration of the surges. In some cases, a rectifier that has an average-current rating higher than that required by the circuit must be used to meet surge requirements of the circuit. This technique eliminates the need for additional circuit impedance elements or special fusing.

If fuses are used to protect the rectifiers, a coordination chart, such as that shown in Fig. 67(b), should be constructed. This chart shows the surge rating of the rectifier (curve A), the opening characteristics of the fuse (curve B), and the maximum surge current available in the circuit (curve C). In the construction of a coordination chart for a particular rectifier, the rms value of the surge




(a)

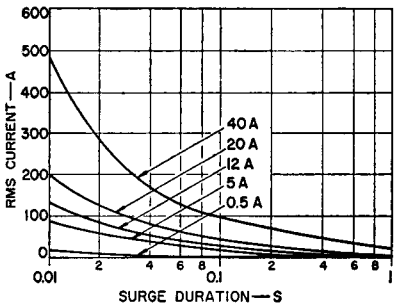



(b)

Figure 67. (a) Peak-surge-current rating chart for a 12-ampere silicon rectifier; (b) coordination chart that relates rectifier surge-current rating (curve A), opening characteristics of circuit fuses (curve B), and maximum available surge current in a circuit (curve C).

current can be obtained from a universal surge-rating chart, such as that shown in Fig. 68. The opening characteristics of the fuse can be obtained from the manufacturer's published data, and the maximum surge current can be calculated.

The coordination chart shown in Fig. 67(b) was prepared for a 12-ampere silicon rectifier operated in half-wave service from a 220-volt rms ac source and protected by a fuse having opening characteristics as shown by curve B. If the total short-




*Note:* The rms current given by this curve is a partial surge rating and should be added to the normal rms current to determine the total surge rating.

Figure 68. Universal surge-current rating chart for RCA silicon rectifiers.

circuit impedance of all the rectifier elements is determined to be 2.25 ohms, the peak surge current $I_s$ for full-wave operation can be calculated as follows:

$$I_s = \frac{220\ V_{rms} \times 1.41}{2.25}$$

$$= 137.6 \text{ amperes}$$

For half-wave service, the peak surge current ($I_s = I_{pk}$) can be converted to rms current by use of the relationships given in Table I, as follows:

$$I_{rms} = \tfrac{1}{2}\ I_{pk}$$

$$= \frac{137.6}{2}, \text{ or } 68.8 \text{ amperes}$$

Curve A of Fig. 67(b), which is merely a reproduction of the 12-ampere curve on the universal rating chart shown in Fig. 68, gives the surge-current rating of the 12-ampere silicon rectifier, but does not consider the normal rms value of current that the rectifier can handle. This normal value of rms current must be subtracted from the total surge current to determine the actual overcurrent of the fault. First, the relationships in Table V are used to convert the average-current rating of the rectifier to the normal rms value, as follows:

$$I_{rms} = 1.57\ I_{avg}$$

$$= 1.57 \times 12, \text{ or } 18.8 \text{ amperes}$$

The overcurrent is then determined from the following calculation:

$$I_{surge} - I_{normal} = 68.8 - 18.8,$$

$$\text{or } 50 \text{ amperes}$$

The 50-ampere fault current is represented on the coordination chart in Fig. 67(b) by the straight-line curve C. The 12-ampere rectifier can sustain a fault current of this magnitude for 51 milliseconds, as indicated by the point of intersection of curves A and C. The fuse, however, opens and interrupts the flow of current in the circuit after 43 milliseconds, as indicated by the point of intersection of curves B and C, and the rectifier is protected.

## Amperes Squared-Seconds ($I^2t$)

The amperes squared-seconds rating of a silicon rectifier provides information on the **maximum subcycle surge current** that the rectifier can sustain when it is used with extremely fast circuit-interrupting devices or is operated in nonsinusoidal rectifier applications. In the manufacturer's published data, the rating is usually given for operation at 60 Hz and is calculated from the

maximum peak surge current that the rectifier can sustain over the period of one cycle (16.67 milliseconds), as follows:

$$I^2t = \left(\frac{\text{one-cycle surge-current rating}}{2}\right)^2 \times 16.67 \times 10^{-3} \quad (17)$$

The peak value of surge current that can be sustained by a 12-ampere silicon rectifier is given by the curve shown in Fig. 67(a) as 240 amperes. The amperes squared-seconds rating for the rectifier is then determined from the following calculation:

$$I^2t = \left(\frac{240}{2}\right)^2 \times 16.67 \times 10^{-3}$$

$$= 240 \text{ amperes squared-seconds}$$

From the value obtained for the $I^2t$ rating, the rms value of the maximum surge current can be calculated for any time between 0.83 millisecond and 8.3 milliseconds (i.e., from 5 to 50 per cent of the period of one cycle). For example, if a square wave of current is to be passed through the 12-ampere rectifier for 3 milliseconds, the maximum current that can be tolerated is determined as follows:

$$I = \sqrt{\frac{I^2t}{t}} \quad (18)$$

$$= \sqrt{\frac{240 \text{ amperes squared-seconds}}{3 \times 10^{-3} \text{ seconds}}}$$

$$= 283 \text{ amperes}$$

If a half-cycle sine wave of current is passed through the rectifier instead of the square wave of current, the peak value of the maximum permissible current is determined by use of the relationship in Table V, as follows:

$$I_{pk} = 2\, I_{rms}$$

$$= 2 \times 283, \text{ or } 566 \text{ amperes}$$

## FAST-RECOVERY RECTIFIERS

In the selection of silicon rectifiers for television high-voltage power supplies, high-speed inverters, switching regulators, and other high-frequency applications, fast recovery characteristics are a major requisite. The actual recovery time of a rectifier is dependent not only upon the structural characteristics of the device, but also upon factors such as the amount of forward current prior to turn-off, the rate of decay of the forward current, the magnitude of the applied reverse voltage, and the junction temperature of the rectifier. A manufacturer's specification for the reverse recovery time of a given rectifier is meaningful, therefore, only if the critical parameters and the actual test circuit used for the measurement are also specified.

### Recovery-Time Test Circuit

Fig. 69(a) shows a circuit recommended by the JEDEC Committee (JC-22) on Power Rectifiers for use in the measurement of rectifier recovery time. In this circuit, capacitor C is charged during the positive alternation of the input ac voltage. During the

negative half-cycle, the silicon controlled rectifier (SCR) is triggered, and capacitor C discharges through inductor L and the rectifier on which the recovery-time measurements is being made. The resultant test-current waveform is shown in Fig. 69(b). Inductor L and capacitor C form a series resonant circuit so that the forward current through the rectifier is very nearly a half sine wave. The peak forward current $I_{FM}$ is specified as $\pi$ times the average rated value of the half-sine-wave current through the rectifier. The rate of decay of the forward current (– di/dt) is specified as the slope of a straight line that passes through the points $I_{FM}/2$ and zero. Regardless of the value of the peak forward current, the di/dt value is specified as 25 amperes per microsecond for high-power stud-mounted rectifiers and 10 amperes per microsecond for low-power lead-mounted rectifiers. For a true half-sine-wave pulse, the increment from $I_{FM}/2$ to zero represents 30 electrical degrees, or one-sixth the total width of the forward-current pulse. The rate of decay of the forward current, therefore, can be expressed in terms of the overall pulse width (PW) as follows:

$$\frac{di}{dt} = \frac{I_{FM}/2}{(PW)/6} = \frac{3I_{FM}}{PW} \qquad (19)$$

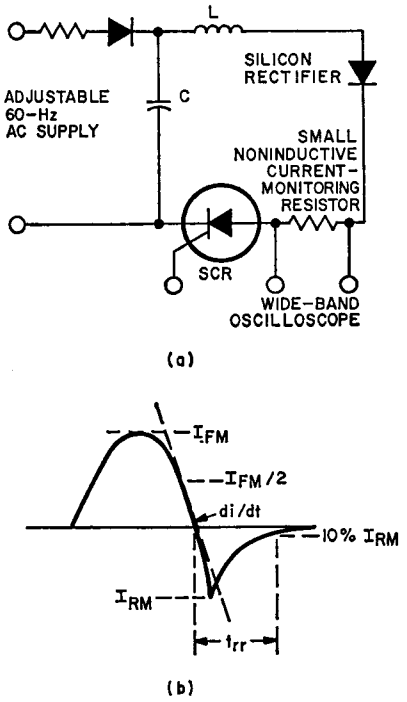



Figure 69. Test circuit and waveform for rectifier reverse-recovery-time measurement.

For stud-mounted rectifiers, the di/dt value is specified as 25 amperes per microseconds. The pulse width, therefore, is defined by the following relationship:

$$PW = 3I_{FM}/25 = 0.12I_{FM} \qquad (20)$$

For lead-mounted rectifiers, the di/dt value is specified as 10 amperes per microsecond, and the expression for the pulse width becomes

$$PW = 3I_{FM}/10 = 0.3I_{FM} \qquad (21)$$

The desired width of the current pulse is obtained by selection of the proper values for L and C in the test circuit. The values of these components are determined from the following relationships:

$$PW = \pi (LC)^{1/2} \qquad (22)$$

$$C = I_{FM} (PW)/\pi V_p \qquad (23)$$

where $V_p$ is the peak voltage across the capacitor.

The relationships expressed above all assume zero circuit losses. Some adjustment of the calculated values of L and C may be required to compensate for these losses.

A typical practical circuit for measurement of the recovery time of a fast-recovery rectifier is shown in Fig. 70. The diode in parallel with the 40216 SCR in this circuit carries the reverse current through the LC circuit so that the reverse recovery characteristics of the rectifier are not affected by the reverse recovery characteristic of the SCR.

Many other circuits have been used for measuring reverse recovery time. Fig. 71 shows one method which has been used as the basis for reverse-recovery data by some manufacturers. In this circuit, the forward-current supply and the associated resistors are adjusted to provide a specified value of forward current. The reverse-current supply is adjusted to supply a specified value of reverse-recovery current when switch S is closed. In some cases, the switch S and the reverse-current supply are replaced by a pulse generator.

Unfortunately, most of the different methods of measuring reverse recovery time yield widely varying results. The values obtained depend on many factors, including the magnitude of forward current, the magnitude of reverse-recovery current, the point on the waveform at which recovery time is measured (usually 10 per cent of peak reverse current), and the rate at which forward current decays toward zero (usually a function of circuit layout, stray capacitance, and inductance).

Correlation of reverse-recovery time measurements between equipments at different locations becomes difficult in circuits which

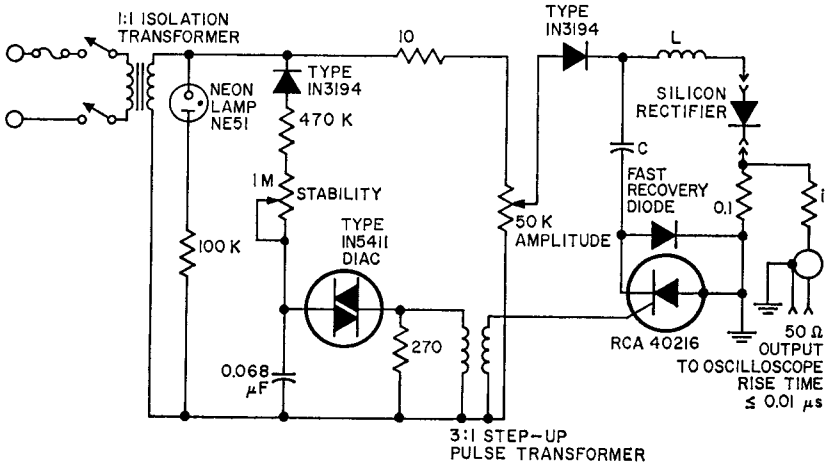



Figure 70. A typical, practical circuit for measurement of the recovery time of a fast-recovery rectifier.

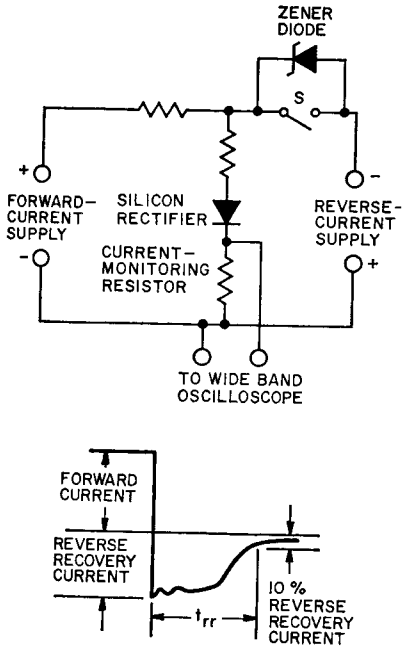



Figure 71. Typical circuit used to measure reverse recovery time of a rectifier.

produce very rapid rates of change in current (such as the circuit shown in Fig. 71). Seemingly minor changes in circuit layout can have a large effect on the measured reverse-recovery time. The circuit shown in Fig. 69, which uses a half-sine-wave test-current pulse, yields results which are readily reproducible, even with widely differing circuit layouts. For this reason it is considered the most meaningful method of evaluating the reverse-recovery characteristic of a rectifier.

## Types of Recovery Characteristics

In a given circuit, three types of recovery characteristics may exist, depending upon the type of rectifier used. Although the exact shape of recovery characteristic is a function of the circuit, the basic form of the characteristic is as shown in Fig. 72. The characteristic shown in Fig. 72(a) is associated with a standard rectifier not designed for fast turn-off characteristics.

Fig. 72(b) shows the recovery characteristic of a rectifier which recovers its blocking-voltage capability suddenly. Although this "snap-off" type of turn-off characteristic is an indication of good high-frequency operation, it can produce undesired effects. If the peak magnitude of reverse current from which the rectifier snaps off is relatively high, an appreciable amount of energy is

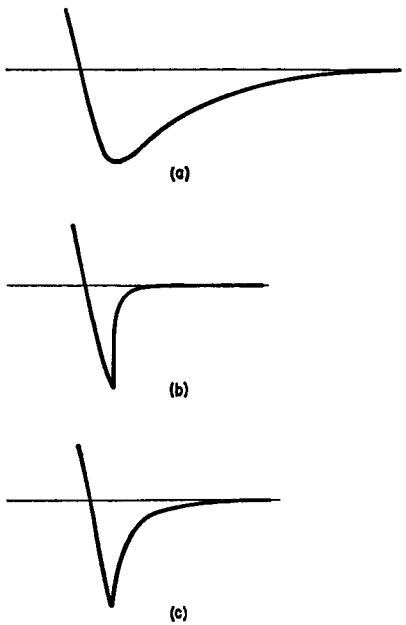



Figure 72. Rectifier reverse-recovery characteristics: (a) for conventional rectifiers; (b) snap-off fast-recovery characterisic; (c) fast-recovery characteristic without abrupt switch off.

contained in the harmonics generated by the snap-off. If sensitive radio or television receiving equipment is in the vicinity, these harmonics can create interference problems. If considerable lead inductance is associated with the rectifier, the snap-off can induce ringing in the lead inductance that may result in circuit malfunction. In some cases, the high di/dt associated with the lead inductance can induce voltages large enough to destroy the rectifier.

The characteristic shown in Fig. 72(c) represents a rectifier which turns off rapidly, but at a somewhat more gradual rate than the rectifier recovery characteristic shown in Fig. 72(b). Reverse-recovery time is only slightly longer than that of the snap-off type, but the generation of harmonics and ringing effects is tremendously reduced. RCA **fast-recovery rectifiers** are designed to have a reverse-recovery characteristic similar to that shown in Fig. 72(c). This type of characteristic is achieved by use of gold-doping to control the lifetime of minority carriers and a junction geometry designed to prevent abrupt decreases in the peak negative current. Fig. 73 shows the reverse-recovery characteristics of a typical RCA fast-recovery rectifier designed for use in an SCR horizontal-deflection system.

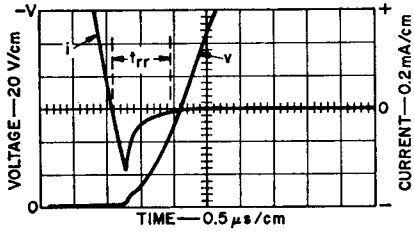



Figure 73. Reverse-recovery characteristics of an RCA-40643 fast-recovery rectifier.

## CONTROLLED-AVALANCHE RECTIFIERS

Controlled-avalanche types are recommended for silicon-rectifier applications in which the ability to withstand high voltage transients is an important design consideration. In controlled-avalanche rectifiers, the voltage at which avalanching occurs is predetermined during manufacture by precise control of the resistivities (i.e., doping-impurity concentrations) in the junction areas and by careful attention to the geometry of the silicon pellet.

In the manufacture of controlled-avalanche rectifiers, special care is taken to assure exceptional regularity of the silicon pellet and an even distribution of impurities in both the n- and p-type regions of the semiconductor junction. In addition, the edges of the silicon p-n junction are shaped to reduce the intensity of localized electric fields at the junction surface. These conditions assure that breakdown will be uniform across the entire junction area rather than concentrated at weak spots close to the junction surface. Such uniform avalanching, when maintained within acceptable limits, is not destructive.

Extensive tests of controlled-avalanche rectifiers permit precise predictions of the behavior of these devices under high-reverse-voltage conditions. The rectifiers are tested to determine their ability to withstand high-voltage transients and are subjected to life tests to determine

their capability for sustained operation in the avalanche region. The slope of the current-voltage curve in the avalanche region defines the dissipation level at the onset of avalanche breakdown and also the maximum dissipation level that the rectifiers are rated to withstand under reverse-bias conditions.

Within the specified ratings, controlled-avalanche rectifiers can safely absorb large bursts of energy, such as may result from abrupt switching of inductive circuits.

## SERIES AND PARALLEL RECTIFIER ARRANGEMENTS

Two or more silicon rectifiers can be used in parallel or series arrangements to extend current and voltage capabilities beyond the limits attainable from a single rectifier. Some basic considerations for multiple connections of silicon rectifiers are discussed in the following paragraphs.

### Parallel Arrangements

When two or more silicon rectifiers are connected in parallel, the current-handling capability of the combined units is substantially greater than that of a single rectifier of the same type. It is often more practical, however, to obtain the greater current capability by use of a multiphase circuit or by selection of a single higher-current rectifier, if available, that can provide the capabilities required.

When rectifiers are to be used in parallel arrangements, the main concern is the forward-voltage characteristics of the rectifiers selected. If the forward-voltage characteristics of the rectifiers are not closely matched, an unbalance in the current division among the rectifiers occurs. The rectifier that has the lower forward-voltage drop receives a larger share of the total current. The higher current causes a greater heating of this rectifier which further reduces the forward voltage drop and thereby causes an additional increase in the current. This regenerative effect can result in destruction of the rectifier and can lead to progressive destruction of all the rectifiers in the parallel array. In parallel operation of silicon rectifiers, therefore, the circuit configuration should assure that the rectifiers receive equal shares of the total current, forward-voltage characteristics of the rectifiers should be closely matched, or a combination of both techniques should be used.

An equal division of current among the rectifiers can be forced by use of resistors or balancing inductors in series with each rectifier. The major disadvantage to the use of series resistors is that they introduce large power losses that reduce rectifier efficiency. The major disadvantage of balancing reactors is the relatively high cost of these components.

The best method to assure equal division of current through parallel rectifiers is to select rectifiers on the basis of the match in their forward-voltage characteristics. This selection can be made more easily when a large number of parallel circuits is to be constructed, because the rectifiers can then be graded into different voltage-drop categories and units from only one category selected for a given parallel circuit. Because the forward voltage drop of a silicon rectifier is dependent

upon the temperature, rectifiers used in a parallel array should be maintained at the same temperature. One technique that may be used to assure that temperature deviations among the rectifiers will be held to a minimum is to mount all the units in the parallel array on the same heat sink.

When silicon rectifiers are connected in parallel arrangements, all contacts should have a low resistance, the wires used should be large enough so that their resistance is negligible, and in high-current arrays the wiring should be arranged so that a minimum unbalance in inductive effects is achieved.

### Series Arrangements

Two or more silicon rectifiers may be connected in series arrangements when voltage requirements exceed the capabilities of a single rectifier. The main concern when rectifiers are to be operated in series is that the reverse voltage be divided equally across each rectifier. The use of resistance-capacitance equalizing networks and the selection of rectifiers that have matched reverse characteristics are the two most common techniques employed to assure equal voltage division. These techniques are discussed in greater detail later in connection with **High-Voltage Rectifier Assemblies.**

A third technique that may be employed when rectifiers are connected in series is the use of transformers that have multiple secondary windings. Each secondary winding is connected across one of the rectifiers in the series array. This technique is practical when only a few rectifiers are to be connected in series. For a large number of rectifiers, the cost and complexity of the multiple-secondary approach become prohibitive.

## HIGH-VOLTAGE RECTIFIER ASSEMBLIES

A series-stack arrangement of rectifier units is used when voltages higher than those obtainable from a single rectifier are required. Several methods have been used to equalize the voltage distribution across series rectifiers for high-voltage assemblies. Among these methods the two most common are RC compensation and selection of matched rectifiers for uncompensated assemblies.

### RC-Compensated Assemblies

In the RC-compensated high-voltage stack, a resistor and a capacitor are placed across each rectifier unit. These resistors and capacitors force an equal division of reverse voltage across each unit in the series string if their values are chosen so that, under all operating conditions, these components, and not the rectifiers, control the distribution of the voltage. The resistors control the voltage division during dc operation. The capacitors control the voltage division during high-frequency operation or when transient voltages are applied. Both the resistor and the capacitor control the voltage division during normal low-frequency operation.

The stray capacitance from the rectifiers to ground, $C_g$ in Fig. 74, tends to cause an unequal distribution of voltage across the rectifiers. The disruptive effect of this stray capacitance is one rea-

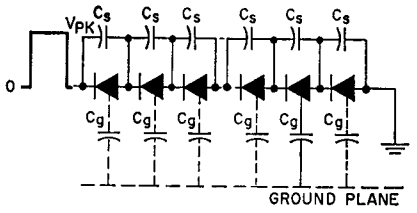



Figure 74. High-voltage rectifier assembly using shunt capacitors ($C_s$) to compensate for stray capacitances ($C_g$) and to equalize the reverse-recovery times of the rectifier units.

son for the use of a shunt capacitor $C_s$ across each rectifier.

The effect of the stray capacitance is greatest during transient conditions. When a step reverse voltage is applied to the rectifier terminal farthest from ground, most of this voltage appears across the first rectifier in the series stack. This condition occurs because the junction capacitance of that rectifier is small and has a large reactance compared to the capacitance to ground of the remainder of the rectifiers in the stack. If a shunt capacitor $C_s$, which is large in comparison to the stray capacitance $C_g$, is connected across each rectifier, an equal voltage distribution can be firmly established among all the rectifiers in the series stack.

The shunt capacitors are also used to equalize reverse-recovery time. As mentioned in the section on **Electrical Characteristics,** reverse recovery of a rectifier is basically the result of two effects: (1) minority carriers are swept out of the junction by reverse current, and (2) minority carriers recombine in the junction area. Of these two effects, the faster one is the sweeping out of minority carriers by reverse current.

In any string of rectifiers, the reverse recovery time of the individual units differs slightly. If the rectifiers are not specially graded for recovery times, then those units which recover first must either block the total reapplied voltage or pass reverse current. When these faster units recover, they stop the flow of reverse current and thereby slow down the recovery time of the remaining units. The shunt capacitors bypass reverse current around the recovered rectifiers and thereby speed up the recovery of the slower rectifiers.

## Uncompensated Rectifier Assemblies

In an uncompensated high-voltage rectifier stack, the characteristics of the series rectifiers are matched to provide an equal division of voltage among individual units in the stack. The characteristics considered for this voltage division are reverse-recovery time, avalanche voltage, and reverse-dissipation capability. The effects of these characteristics are all interrelated and must be considered together.

When rectifiers have been sorted into recovery-time groupings, they can be used more reliably in series arrangement because all of them will recover their blocking ability at about the same time. Any charge which flows into the capacitance $C_g$ is partially supplied while all the rectifiers are still in an "ON" state. Any current that flows during this time passes through units which are in a low-impedance state; the power dissipated across them, therefore, is small. In addition, if all the units have matched recovery-time characteristics, the main mode of reverse recovery results from a sweeping out of minority carriers,

and all the rectifiers recover by the faster recovery method. Any unbalance in recovery time is small, therefore, and the units which recover first have to block excess voltage for only a very short period of time.

A rectifier in an externally uncompensated series stack that recovers before the other rectifiers in the stack immediately begins to block all the voltage; as a result, the blocking-voltage capability of this rectifier may be exceeded sufficiently to cause failure of the device.

In effect, matching of recovery-time and avalanche characteristics of rectifiers performs the same function as the capacitor and resistor in RC-compensated series stacks. The reverse-dissipation capability of the rectifier takes the place of the capacitor in the RC-compensated series stack. The use of matched avalanche characteristics provides the same results as the compensating resistor during dc operation. The match in the reverse-dissipation capability of the rectifiers assures that there is no decrease in reliability of the unit as a result of the avalanche action.

## Packaging

There are generally three methods of packaging high-voltage rectifier assemblies: encapsulation; open exposure to air; and immersion in a special high-voltage atmosphere, such as transformer oil or the newer gaseous insulating mediums. Each of these systems has its advantages and disadvantages.

When encapsulation is used, a high packing density can be achieved. The encapsulant protects the rectifier stack from harmful atmospheres, such as those encountered in high-humidity areas. The encapsulant further protects the stack from accumulations of dust and dirt which can cause leakage paths and upset voltage division among the rectifiers. On the other hand, the encapsulant acts as a barrier to the checking of individual units in the stack so that a rectifier failure cannot be repaired, even if it can be detected; any deterioration of the stack which is detected by an overall check of the total assembly can be corrected only by replacement of the total stack. In addition, the encapsulation of high-current assemblies results in a serious loss of heat-dissipating ability.

For these reasons, it is not advisable to encapsulate very large or expensive assemblies. If a high-voltage, low-current assembly is to be encapsulated, some thought should be given to the use of several encapsulated sections, so that if deterioration does take place in a part of the stack, only that part need be replaced. In general, only those assemblies which can be advantageously replaced as a whole should be encapsulated.

Open assemblies directly exposed to the atmosphere have several advantages. They are fairly easy to install, and cooling can be accomplished by convection and radiation or, for high dissipation, by forced air. All the rectifiers are exposed and can be checked for deterioration. If a rectifier is found to have deteriorated, it can be replaced individually, and it is not necessary to discard the whole stack. Open construction permits efficient use of large heat sinks

and allows the designer to make fuller use of the capability of the rectifier.

The features which make encapsulation attractive are the features which are disadvantages in the open stack. A low packing density is required because of isolation requirements. Harmful atmospheres or high humidity can affect operation. Accumulations of dust and dirt can result in leakage paths which upset the voltage division among the rectifiers, and during high-voltage operation corona may develop if care is not taken.

Oil-immersed assemblies offer some of the advantages of both encapsulated and open assemblies, plus a few additional advantages. When oil is used as an insulating medium, a fairly good packing density can be obtained. The intimate contact between the oil and the rectifier heat sinks aids rectifier cooling. The closed oil system tends to reduce any accumulation of dirt. The dielectric properties of the oil reduce or eliminate corona problems. In addition, the rectifier stacks can be removed from the oil for testing and maintenance.

The disadvantages of oil immersion outweigh the advantages in many applications, e.g., in lower-voltage installations. The use of oil immersion requires an oil tank with high-voltage bushings. The added weight of the system may create a floor-loading problem. When large quantities of power are being dissipated by the rectifiers, additional cooling is required. The use of a heat exchanger to remove the heat from the oil or of a radiator with a fan through which the oil can circulate may be necessary. If care is not taken, the oil can become contaminated with moisture and dirt, with the result that arcing and corona may occur. Although testing and maintenance of an oil-immersed system are possible, removal of the assembly from the oil is often a difficult operation.

# Basic Design Considerations for Power Transistors

IN power transistors, the main design consideration is power-handling capability. This capability is determined by the maximum junction temperature a transistor can withstand and how quickly the heat can be conducted away from the junction. (The thermal circuit for a power transistor is described in the section on **Thermal Factors.**)

In general, the basic physical theory that defines the behavior of any bipolar transistor in relation to charge-carrier interactions, current gain, frequency capabilities, voltage breakdown, and current and temperature ratings is not significantly different for power types. Power transistors, however, must be capable of large current densities and are required to sustain large voltage fields. For power types, therefore, the basic transistor theory must be expanded to include the effect that these conditions have on the physical behavior of the devices. In addition, the physical capabilities of power transistors must be defined in terms of factors, such as second-breakdown energy levels, safe operating area, and thermal-cycling stresses, that are not usually considered for small-signal types. (The various rating factors that define maximum operating limits are explained in the section on **Physical Basis for Power-Transistors Ratings.**)

## GENERAL PHYSICAL THEORY

All bipolar transistors consist of three layers of semiconductor material (usually silicon for power types) referred to as **emitter, base,** and **collector.** The resultant structure forms two back-to-back p-n junctions. The input (emitter-base) junction serves as the source, or injector, of current carriers; the output (base-collector) junction collects the injected current carriers. During normal operation, the emitter-base p-n junction is forward-biased, and the collector-base p-n junction is reverse-biased.

As explained in the section on **Silicon Rectifiers,** a p-n junction biased in the reverse direction is equivalent to a high-resistance element, while a junction biased in the forward direction is equivalent to a low-resistance element. The electric field across the forward-biased junction overcomes the energy barrier at the junction and causes holes to be injected

into the n-type region and electrons to be injected into the p-type region. Because of the large number of free electrons in the n-type region and of holes in the p-type region, the injected holes and electrons are referred to as minority-charge carriers. A forward-biased p-n junction, therefore, is a minority-carrier injector, and the number of minority carriers injected is dependent upon the magnitude of the forward-bias voltage.

## Charge-Carrier Flow

When a symmetrical p-n junction is forward-biased, the lifetime of the injected minority carriers is very short. Because of the many free electrons in the n-type region, a hole injected into this region is not likely to penetrate very far before it meets an electron and is annihilated (i.e., neutralized), as shown in Fig. 75.

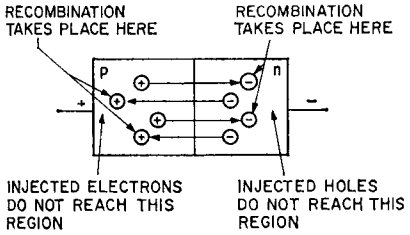



Figure 75. Diagram showing that recombination limits travel of injected carriers in a symmetrical forward-biased p-n junction.

Similarly, any electron injected into the p-type region is usually quickly neutralized by one of the numerous holes in this region. In a symmetrical p-n junction, therefore, injected minority carriers cannot penetrate very far or last very long before they are annihilated.

Fig. 76(a) shows a nonsymmetrical p-n junction in which the n-type region is made very thin and the p-type region is much more heavily doped. When this junction is forward-biased, an injected hole is much less likely to be annihilated by an electron before it crosses to the end of the thin n-type region.

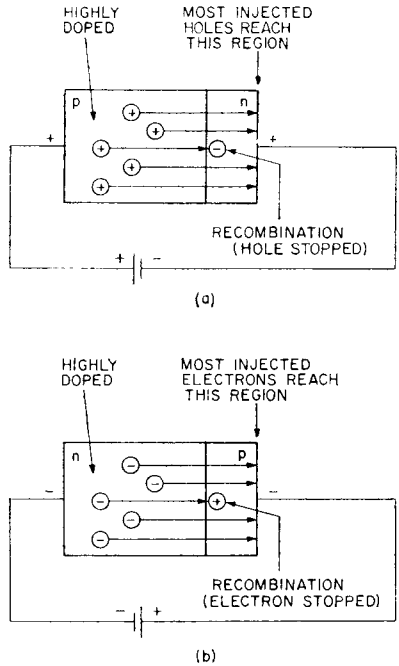



Figure 76. Diagrams showing that (a) hole injection is improved by unequal doping and a thin n-type section and (b) electron injection is improved by unequal doping and a thin p-type section.

Moreover, because of the heavy doping of the p-type region, more holes are injected into the n-type region than there are free electrons in this thin region. Consequently, even though some injected holes are annihilated by free electrons, most of them are able to survive and penetrate the

full width of the n-type region, as shown in Fig. 76(a). Similarly, in a forward-biased p-n junction in which the p-type region is very thin and the n-type region is much more heavily doped, an injected electron is unlikely to meet (and be neutralized by) a hole before it penetrates to the end of the thin p-type region, as shown in Fig. 76(b).

In bipolar transistors, a thin lightly doped semiconductor layer (base region) is sandwiched between two wider (emitter and collector) semiconductor layers that are much more heavily doped with an opposite type of impurity from the dopant used in the thin base layer. The two nonsymmetrical back-to-back p-n junctions that result may form either a p-n-p or an n-p-n transistor. Fig. 77 shows the layer

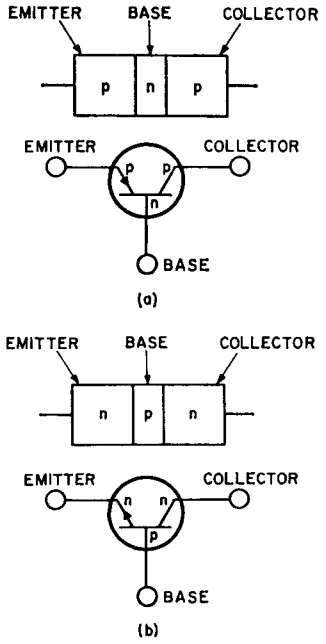



Figure 77. Transistor nomenclature and symbols: (a) p-n-p type; (b) n-p-n type.

structure and the corresponding schematic symbol for each type of transistor.

Fig. 78 shows the basic biasing arrangement for an n-p-n bipolar transistor. External batteries bias the emitter-base (n-p) junction in the forward direction to provide a low-resistance input

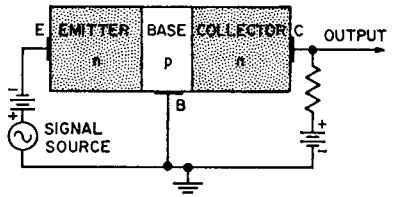



Figure 78. Basic biasing and input-signal connections for an n-p-n transistor.

section, and bias the base-collector (p-n) junction in the reverse direction to provide a high-resistance output section. Electrons flow easily from the n-type emitter region to the p-type base region as a result of the forward biasing. Most of these electrons diffuse through the thin p-type region, however, and are attracted by the positive potential of the external bias supply across the base-collector (p-n) junction. In practical devices, approximately 95 to 99.5 per cent of the injected electrons reach the n-type collector region. This high percentage of current penetration makes possible power gain in the high-resistance output circuit and is the basis for the amplification capability of a transistor.

Fig. 79 shows the electron-energy distribution in an n-p-n transistor for different input-signal conditions. Because electrons are the main current carriers in n-p-n transistors, no potential-energy diagrams of holes are shown for these devices.

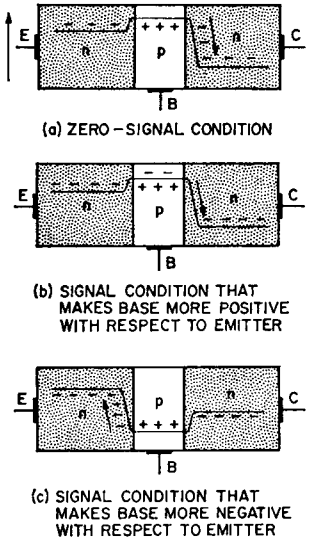



Figure 79. Electron potential-energy diagrams for an n-p-n transistor.

Fig. 79(a) shows the potential-energy condition for a forward-biased n-p-n transistor when no input signal is applied. The forward bias applied between the emitter and the base reduces the potential-energy hill at the emitter-base junction; some electrons, therefore, easily climb this hill and enter the p-type base region. Because the base strip is relatively thin, most of the electrons that enter it will not combine with holes; instead they pass through the strip and readily go down the potential-energy slope at the right p-n junction, as shown by the arrow in Fig. 79(a). The steep potential-energy slope which permits easy entrance to the electrons from the base strip to the n-type collector is produced by application of the reverse bias between the collector and the base.

When the input signal opposes the forward bias on the emitter (makes the base more negative with respect to the emitter), the potential-energy hill between the emitter and the base is increased as shown in Fig. 79(b), and fewer electrons climb the hill to enter the p-type base. The majority of the electrons that enter the p-type base do not recombine with holes, but instead, fall into the collector region of low potential energy.

When the polarity of the applied signal aids the forward bias (makes the base more positive with respect to the emitter), the potential-energy hill between the emitter and the base is decreased and more electrons flow into the p-type region. Most of the electrons do not combine with holes but flow readily to the low potential-energy level of the n-type collector. This condition is shown in Fig. 79(c).

The operation of a p-n-p transistor is essentially identical to that of an n-p-n transistor except that the polarities of the bias voltages are reversed and the main current carriers are holes instead of electrons. Fig. 80 shows the basic biasing and input-signal connections for a p-n-p transistors, and Fig. 81 shows the hole-energy diagrams for different input-signal conditions.

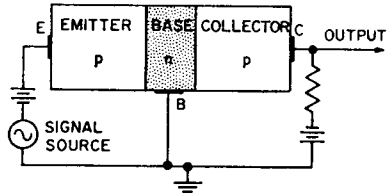



Figure 80. Basic biasing and input-signal connections for a p-n-p transistor.

## Current-Voltage Relationships

The currents in a transistor are directly related to the movement

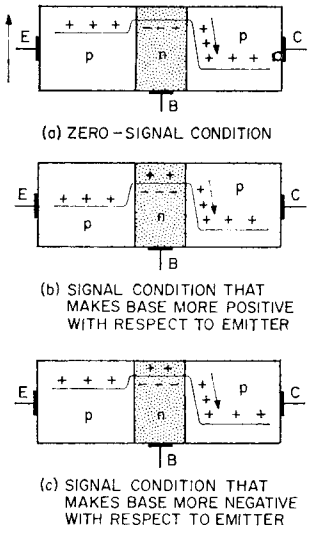



Figure 81. Hole potential-energy diagrams for a p-n-p transistor.

of minority carriers in the base region that results from the application of voltages of the proper polarities to the emitter-base and collector-base junctions. A definite mutual relationship exists between the transistor currents and the voltages applied to the transistor terminals. Graphical representations of the variations in transistor currents with the applied voltages provide an excellent indication of the operation of a transistor under different biasing conditions. Transistor manufacturers usually provide curves of current-voltage characteristics to define the operating characteristics of their devices. Such curves are provided for either common-emitter or common-base transistor connections. Fig. 82 shows the bias-voltage polarities and the current components for both common-emitter and common-base connections of a p-n-p transistor. For an n-p-n transistor, the polarities of the voltages and the directions of the currents are reversed.

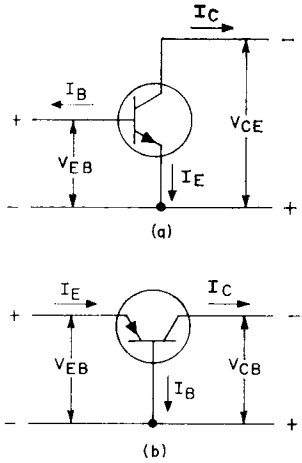



Figure 82. Transistor bias-voltage polarities and current components for (a) the common-emitter connection and (b) the common-base connection.

The common-emitter connection, shown in Fig. 82(a), is the more widely used in practical applications. In this connection, the emitter is the common point between the input (base) and output (collector) circuits, and large current gains are realized by use of a small base current to control a much larger emitter-to-collector current. The common-base connection, shown in Fig. 82(b), differs from the common-emitter connection in that the voltages applied to the transistor are referred to the base rather than to the emitter. Detailed analyses of the common-base and common-emitter connections, and of the common-collector connection, are given in the section on **Equivalent-Model Analyses of Power Transistors**, for both small-signal and large-signal operation.

Fig. 83 shows the input and output current-voltage characteristics of a typical p-n-p transistor in a common-base connection. The input characteristic curves, in Fig. 83(a), show the base current as a function of the emitter-to-base voltage for different values of collector-to-base voltage. For any given value of collector voltage, the base current varies with the emitter-to-base voltage in a manner similar to that of any forward-biased narrow p-n junction. After the initial interval required for the forward-bias voltage to overcome the energy barrier at the junction, the current-voltage relationship is almost linear. The slight effect that the collector voltage has on the base current results because variations in the collector voltage change the effective width of the base.

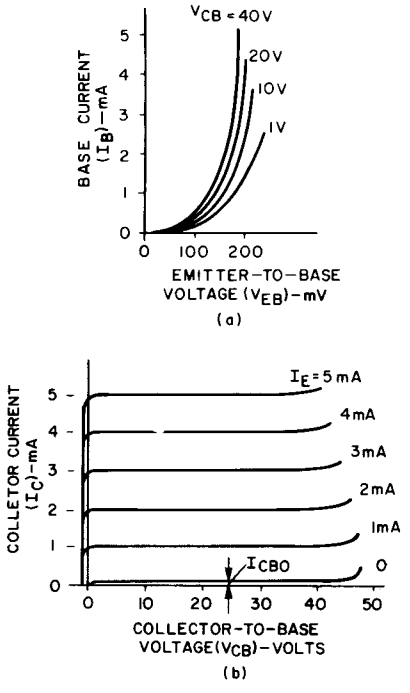



Figure 83. Common-base characteristics of a typical p-n-p transistor.

The effective base width is the distance between the edges of the opposing (collector and emitter) depletion layers. This distance is significantly less than the actual thickness of the n-type base material because both depletion layers penetrate into the base region. Because of the small forward bias applied to the emitter-base junction, the emitter depletion layer is much narrower than that of the collector-base junction, and it does not change appreciably with the small range of variation allowed in the emitter-to-base voltage. The collector-junction depletion layer, however, varies markedly with the much larger changes in collector voltage and causes an attendant change in the effective base width. An increase in collector voltage causes a slight increase in base current (for a constant emitter-to-base voltage) because the hole gradient in the p-n-p transistor increases as the effective base width decreases. Many other transistor characteristics are also affected by changes in collector voltage because of the dependence of the effective base width on the collector depletion layer.

The common-base output characteristics, shown in Fig. 83(b), reveal that the collector current is very nearly equal to the emitter current and is largely independent of collector voltage. This current is made up of the small reverse current $I_{CBO}$, which results from the extraction of thermally generated holes from the base region, and the forward current produced by the diffusion of

holes from the emitter. The diffusion current is almost independent of voltage because all the holes that reach the edge of the collector depletion layer are extracted by the reverse-biased junction. The collector current remains essentially constant, even down to zero voltage, at which point the excess holes are still extracted by the collector. A small forward voltage (less than 1 volt) must be applied to the collector to increase the hole density just outside the junction and oppose the diffusion from the emitter in order to reduce the collector current to zero. The characteristic curves show that the collector current varies extremely rapidly with voltage in this region.

The collector reverse (leakage) current, $I_{CBO}$, is measured with the emitter circuit open. The presence of the second junction, however, still affects the level of this current because the emitter acquires a small negative bias when it is open-circuited. This bias reduces the hole gradient at the collector and causes the reverse current to decrease. This current, therefore, is much smaller with the emitter open than it is when the emitter-base junction is short-circuited. In the characteristic curves shown in Fig. 83(b), the magnitude of the current $I_{CBO}$ is considerably exaggerated. This current is normally in the order of microamperes or less, although it is often increased by excess surface currents in the same way as the reverse current of any p-n junction.

The reverse current increases with collector voltage, and may lead to avalanche breakdown at high voltages, as described for the silicon rectifier. In the same way, all the collector-characteristics curves bend upwards at high enough collector voltages because of multiplication in the collector depletion layer, as will be described subsequently in the section on **Voltage Ratings**. The multiplication factor depends on the junction voltage only; the curvature of the characteristics, therefore, is more noticeable for higher collector currents.

The common-emitter characteristics of a p-n-p transistor, shown in Fig. 84, are similar to those of the common-base connection.

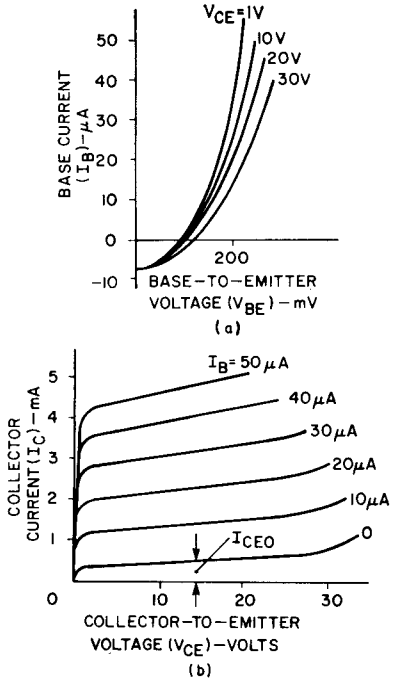



Figure 84. Common-emitter characteristics of a typical p-n-p transistor.

There are, however, several important differences, the most marked being the small magnitude of the base current which

replaces the emitter current in both sets of curves. The base current consists of two components, each of which is of the order of microamperes. These components include the current produced by an inward flow of electrons to replace those lost in the hole injection and diffusion mechanisms, and a current that is largely the result of the extraction of thermally generated holes making up the leakage current of the collector. The total base current, therefore, is just sufficient to make up the losses in the current transfer from emitter to collector. If the base current is increased (by means of the external circuit), more holes can diffuse from emitter to collector, and a considerable increase in the collector current occurs. Physically, this condition is produced by increased hole injection and, therefore, by an increase in the emitter-to-base voltage required to produce an increase in base current. Although changes in base current and voltage must occur together, it is preferable to think of the transistor as a current-controlled device; for this reason, the common-emitter characteristics are usually shown for constant base currents, as indicated in Fig. 84. The curves in this figure confirm that small changes in base current produce much larger changes in collector current.

Another feature of the common-emitter output characteristics is that the common-emitter reverse collector current $I_{CEO}$, measured with zero input current ($I_B = 0$ in this case), is very much larger than the reverse collector current $I_{CBO}$ in the common-base connection. When the base current is zero, the emitter current adjusts itself so that the losses in the hole-injection and diffusion mechanisms are exactly balanced by the supply of excess electrons left in the vicinity of the collector by hole extraction. For this condition, the collector current is equal to the emitter current.

The common-emitter reverse collector current $I_{CEO}$ increases with collector voltage, unlike the common-base reverse collector current $I_{CBO}$. This behavior is another consequence of the variation in the effective base width with collector voltage. The narrower the effective base region, the more efficient is the transfer of current from emitter to collector. The more efficient base transport with the higher collector voltage permits a higher emitter current to flow before the losses are again balanced by the supply of electrons from the vicinity of the collector.

A final point to note about the common-emitter characteristics is that the current falls rapidly to zero for small collector-to-emitter voltages. The collector-to-emitter voltage is normally divided between the two junctions to provide a small forward bias for the emitter-base junction and a much larger reverse bias for the collector-base junction. If the emitter-to-collector voltage is reduced to the "saturation" value, which is a fraction of a volt, the collector reaches zero bias; below this value, it becomes slightly forward-biased. This condition occurs near the knee of the characteristics. Slightly below the saturated condition, holes continue to flow into the collector, in spite of its forward bias. If the

emitter-to-collector voltage is reduced further, the collector current falls rapidly to zero.

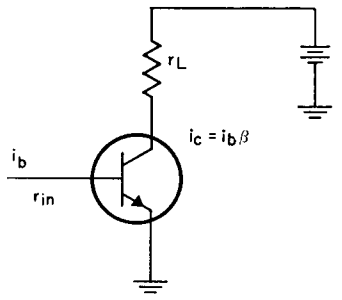
The input characteristics for the common-emitter connection are similar to those for the common-base connection, except for the greatly reduced magnitude of the base current compared with the emitter current. For the base current to be zero, the emitter junction must have a small forward bias; as a result, all the input characteristics cross the current axis. The collector voltage again affects the characteristics by varying the effective base width, and thus the current gain. With a constant emitter hole density ($V_{BE}$ constant), the base current falls as the collector voltage is raised, and the base transport becomes more efficient.

## Physical Basis for Current Gain

As pointed out in the preceding section, power gain in transistor circuits is usually obtained by use of a small control signal to produce larger signal variations in the output current. In an electron-tube circuit, the most common gain parameter is the voltage amplification factor ($\mu$) from control grid to plate. In a transistor circuit, the gain parameter most often specified is the current gain ($\beta$) from the base to the collector. The power gain of a transistor operated in a common-emitter circuit configuration is equal to the square of the current gain times the ratio of the load resistance to the input resistance, as indicated in Fig. 85.

The current gain (or current transfer ratio) of a transistor is expressed by many symbols; the following are some of the most common, together with their particular shades of meaning:

1. beta ($\beta$)—general term for current gain from base to collector (i.e., common-emitter current gain)
2. alpha ($\alpha$)—general term for current gain from emitter to collector (i.e., common-base current gain)
3. $h_{fe}$—ac gain from base to collector (i.e., ac beta)
4. $h_{FE}$—dc gain from base to collector. (i.e., dc beta)




INPUT CURRENT $= i_b$
INPUT VOLTAGE $= i_b\, r_{in}$
OUTPUT CURRENT $= i_c = i_b\beta$
OUTPUT VOLTAGE $= i_c r_L = i_b\beta r_L$
INPUT POWER $= i_b^2\, r_{in}$
OUTPUT POWER $= i_c^2\, r_L = i_b^2\, \beta^2\, r_L$
POWER GAIN $=$ *power output/power input*
$= i_b^2\beta^2 r_L / i_b^2 r_{in}$
$= \beta^2 r_L / r_{in}$

Figure 85. Test circuit and simplified power-gain calculation for a transistor operated in a common-emitter configuration.

In the following paragraphs, a mathematical relationship between $\alpha$ and $\beta$ is derived, and these current-gain parameters are then defined in terms of three basic transistor parameters (the emitter and collector efficiencies and the base transport factor), which are directly dependent

upon the physical properties of the device.

**Relationship Between $\alpha$ and $\beta$**—The **common-emitter current gain** is defined as the ratio of the collector current to the base current (i.e., $\beta = I_C/I_B$). Useful values of $\beta$ are normally greater than 10. Similarly, the **common-base current gain** is defined as the ratio of the collector current to the emitter current (i.e., $\alpha = I_C/I_E$). Values of $\alpha$ usually range from 0.95 to 0.995. Although this gain parameter is slightly less than unity, current gain is realized because of the large difference between input (emitter-base) and output (collector-base) impedances. As explained previously, the input impedance is small because the emitter-base junction is forward-biased, and the output impedance is large because the collector-base junction is reverse-biased.

It is apparent from Kirchoff's Law that the total current at the emitter junction, $I_E$, is the sum of the current that flows into the emitter junction from the base, $I_B$, and the current that flows into the emitter junction from the collector, $I_C$. The following equation expresses this condition:

$$I_E = I_B + I_C \qquad (24)$$

This equation and the definitions of $\alpha$ and $\beta$ are used to derive the following expression of the relationship between the two current gain parameters:

$$\beta = \frac{\alpha}{1 - \alpha} \qquad (25)$$

**Physical Factors that Determine Current Gain**—The common-base current gain $\alpha$ is determined by the emitter efficiency $\gamma$, the base transport factor $\beta_o$, and the collector efficiency $\alpha^*$ (generally very near unity), as follows:

$$\alpha = \gamma\beta_o\alpha^* \qquad (26)$$

The common-emitter current gain $\beta$ is also a function of the emitter and collector efficiencies and the base transport factor. A simple substitution of the expression for $\alpha$ into Eq. (25) results in the following expression for $\beta$ in terms of these parameters:

$$\beta = \frac{\gamma\beta_o\alpha^*}{1 - (\gamma\beta_o\alpha^*)} \qquad (27)$$

**Emitter efficiency:** In a forward-biased p-n-p transistor, holes from the emitter diffuse into the base, and electrons from the base diffuse into the emitter. The total emitter current $I_E$ is the sum of the hole-current component $I_p$ and the electron-current component $I_n$ and, therefore, may be expressed as follows:

$$I_E = I_p + I_n \qquad (28a)$$

The potential collector current, $I_C$, is the difference between the drift currents and is given by

$$I_C = I_p - I_n \qquad (28b)$$

The electrons that diffuse from the base into the emitter originate in the base dc supply and add to the total base current. This electron current $I_n$, however, does not contribute to the collector current and, in effect, represents a loss in current gain that is directly attributable to poor emitter injection efficiency. The loss in current gain can be held to a minimum if the resistivity

of the base is made much greater than that of the emitter so that the number of free electrons in the base available to diffuse into the emitter is substantially smaller than the number of free holes in the emitter available to diffuse into the base.

**Base transport factor:** If high emitter efficiency is assumed, the holes injected from the emitter into the base diffuse to the collector junction. Some of these holes, however, recombine with free electrons in the base and, in effect, are annihilated. Base current must flow to replenish the free electrons used in the recombination process so that the emitter-to-base forward bias is maintained. In other words, charge neutrality must prevail.

The base transport factor $\beta_0$ is an indication of the extent of recombination that takes place in the base region of a transistor. For a high value of the base transport factor $\beta_0$, the lifetime of holes in the base (a function of the property of the material) must be long, or the time necessary for the holes to reach the collector must be short. Any reduction in the time required for the holes to reach the collector requires a decrease in base width or an increase in the accelerating field used to speed the holes through. The relationships of the base transport factor, $\beta_0$, to base width and material properties differs for homogeneous-base and graded-base transistors.

For a homogeneous-base transistor, the base transport factor can be approximated as follows:

$$\beta_0 \approx 1 - \frac{1}{2}\left(\frac{W}{L}\right)^2 \qquad (29a)$$

For transistors that employ a graded base (which has an aiding drift field), the following expression is used to determine the base transport factor:

$$\beta_0 \approx 1 - \frac{1}{4}\left(\frac{W}{L}\right)^2 \qquad (29b)$$

where W is the width of the base and L is the diffusion length of the injected carriers in the base. The value of L may be calculated from the following equation:

$$L = (D\tau)^{1/2} \qquad (30)$$

where D is the diffusion constant of the appropriate carrier in the base material and $\tau$ is the lifetime of the injected carriers in the base (time constant for annihilation of injected carriers).

The value of the base transport factor should be in the order of 0.98 for the transistor to provide useful gain.

## Transit and Delay Times

For all transistors, there is a frequency f at which the output signal cannot properly follow the input signal because of time delays in the transport of the charge carriers. The three principal cut-off frequencies, shown in Fig. 86, may be defined as follows:

1. The **base cut-off frequency** $f_{\alpha b}$ is that frequency at which alpha ($\alpha$) is down 3 dB from the low-frequency value.

2. The **emitter cut-off frequency** $f_{\alpha e}$ is that frequency at which beta ($\beta$) is down 3 dB from the low-frequency value.

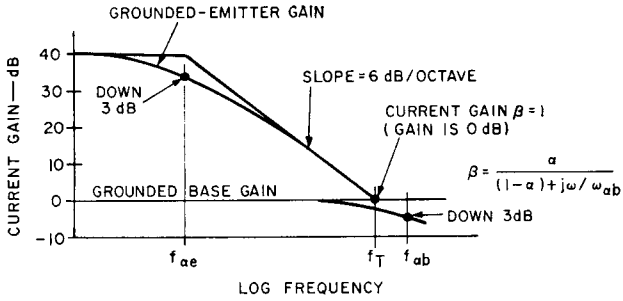



Figure 86. Cut-off frequencies.

3. the frequency $f_T$ is that frequency at which beta theoretically decreases to unity (i.e., 0-dB gain) with a theoretical 6-dB-per-octave fall off. This term, which is a useful figure of merit for transistors, is referred to as the **gain-bandwidth product.**

The frequency $f_T$ is related to the time delays in a transistor by the following expression:

$$f_T \approx \frac{1}{2\tau \Sigma t_d} \qquad (31)$$

where $\Sigma t_d$ is the sum of the emitter-delay time constant $t_e$, the base transit time $t_b$, the collector depletion-layer transit time $t_{xm}$, and the collector-delay time constant $t_c$.

**Emitter-Delay Time**—The emitter delay time is the time required to charge the capacitance, $C_{Te}$, of the emitter-base transition region through the emitter resistance $r_e$. The emitter-delay time constant is expressed as follows:

$$t_e = r_e C_{Te} \qquad (32)$$

The smaller the product of $C_{Te}$ and $r_e$, the smaller the delay time. For a small value of $C_{Te}$, a small emitter area and a high base resistance are needed.

**Base Transit Time**—The base transit time $t_b$ is the time required for the injected carriers to diffuse and drift across the base width W. It is related to the average speed of the carriers and the width of the base. For a homogeneous base, this transit time is determined by the following equation:

$$t_b = \frac{W^2}{D} \qquad (33)$$

The diffusion constant D, in turn, is related to the average carrier velocity (mobility) as follows:

$$D = \frac{\mu kT}{q} \qquad (34)$$

where $\mu$ is the carrier mobility and kT/q is a temperature-dependent constant.

For short base transit time, narrow base width and a high carrier mobility are required.

**Collector Depletion-Layer Transit Time**—The collector depletion-layer transit time, $t_{xm}$, is the time required for the carriers to be swept through the collector depletion layer by the collector field. The collector delay time is related to the width of the depletion layer (xm) and the speed of the carriers (v). The speed, in turn, is governed by the applied field (E), up to a limiting or saturation velocity ($v_{sc}$). The relationship of transit time to the saturation velocity is given by the following equation:

$$t_{xm} = \frac{xm}{v_{sc}} \qquad (35)$$

For silicon, $v_{sc}$ is approximately equal to $8.5 \times 10^6$ centimeters per second at a value of E greater than 10,000 volts per centimeter.

**Collector-Delay Time Constant**—The collector-delay time constant $t_c$ is the time required to charge the capacitance of the collector junction ($C_{Tc}$) through the combined series resistance from the emitter to the collector. Generally, only the collector series resistance ($r_{sc}$) is significant. The collector-delay time constant, therefore, is approximated by the following equation:

$$t_c = r_{sc}\, C_{Tc} \qquad (36)$$

A reduction in the value of the collector resistivity causes an over-all decrease in the collector-delay time constant because $r_{sc}$ is decreased more than $C_{Tc}$ is increased. Also, small collector areas may help reduce $t_c$ if $r_{sc}$ is not grossly affected. The relationship of the collector area A to the series resistance $r_{sc}$, the collector resistivity $\rho_c$, the thickness of the collector d, and the collector-junction capacitance $C_{Tc}$ is shown in the following equations:

$$r_{sc} = \frac{\rho_c d}{A} \qquad (37)$$

and

$$C_{Tc} = A(\rho_c)^{1/2} \qquad (38)$$

## Effect of Temperature on Transistor Characteristics

The characteristics of transistors vary with changes in temperature. In view of the fact that most circuits operate over a wide range of environments, a good circuit design should compensate for such changes so that operation is not adversely affected by the temperature dependence of the transistors.

**Current Gain**—The effect of temperature on the gain of a silicon transistor is dependent upon the level of the collector current, as shown in Fig. 87. At the lower current levels, the current-gain parameter $h_{FE}$ increases with temperature. At higher currents, however, $h_{FE}$ may increase or decrease with a

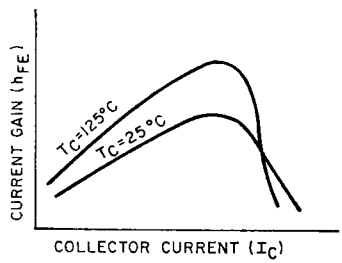



Figure 87. Current gain as a function of collector current at different temperatures.

rise in temperature because it is a complex function of many components.

**Base-to-Emitter Voltage**—Fig. 88 shows the effect of changes in temperature on the base-to-emitter voltage ($V_{BE}$) of silicon transistors. Two factors, the base resistance ($r_{bb}'$) and the height of the potential barrier at the

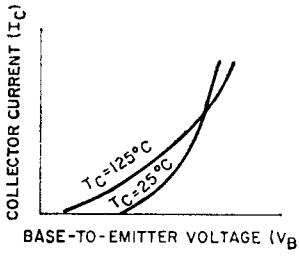



Figure 88. Collector current as a function of base-to-emitter voltage at different temperatures.

base-emitter junction ($V_{BE}'$), influence the behavior of the base-to-emitter voltage. As the temperature rises, material resistivity increases; as a result, the value of the base resistance $r_{bb}'$ becomes greater. The barrier potential $V_{BE}'$ of the base-emitter junction, however, decreases with temperature. The following equation shows the relationship between the base-to-emitter voltage and the two temperature-dependent factors:

$$V_{BE} = I_B r_{bb}' + V_{BE}' = \frac{I_C}{h_{FE}} r_{bb}' + V_{BE}' \qquad (39)$$

As indicated by this equation, the base-to-emitter voltage diminishes with a rise in temperature for low values of collector current, but tends to increase with a rise in temperature for higher values of collector current.

**Collector-to-Emitter Saturation Voltage**—The collector-to-emitter saturation voltage $V_{CE}(sat)$ is affected primarily by collector resistivity ($\rho_C$) and the amount by which the natural gain of the device ($h_{FE}$) exceeds the gain with which the circuit drives the device into saturation. This latter gain is known as the forced gain ($h_{FEf}$).

At lower collector currents, the natural $h_{FE}$ of a transistor increases with temperature, and the IR drop in the transistor is small. The collector-to-emitter saturation voltage, therefore, diminishes with increasing temperature if the circuit continues to maintain the same forced gain. At higher collector currents, however, the IR drop increases, and gain may decrease. This decrease in gain causes the collector-to-emitter saturation voltage to increase and possibly to exceed the room-temperature (25°C) value. Fig. 89 shows the effect of temperature on the collector-to-emitter saturation voltage.

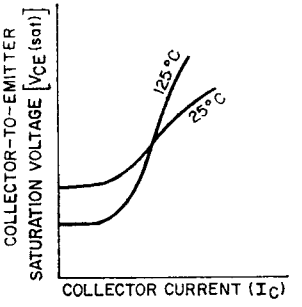



Figure 89. Collector current as a function of collector-to-emitter saturation voltage at different temperatures.

**Collector Leakage Currents**— Reverse collector current $I_R$ is a resultant of three components, as shown by the following equation:

$$I_R = I_D + I_G + I_S \qquad (40)$$

Fig. 90 shows the variations of these components with temperature.

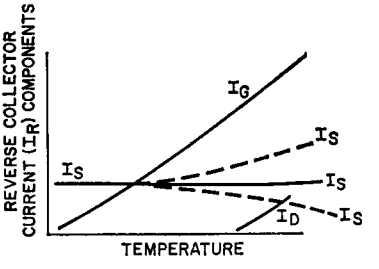



Figure 90. Reverse collector current as a function of temperature.

The diffusion or saturation current $I_D$ is a result of carriers that diffuse to the collector-base junction and are accelerated across the depletion region. This component is small until temperatures near 175°C are reached. The component $I_G$ results from charge-generated carriers that are created by the flow of diffusion carriers across the depletion region. This component increases rapidly with temperature. $I_D$ and $I_G$ are referred to as bulk leakages. The term $I_S$ represents surface leakage which is caused by local inversion, channeling, ions, and moisture. This leakage component is dependent on many factors, and its variations with changes in temperature are difficult to predict.

At low temperatures, either surface or bulk leakage can be the dominant leakage factor, particularly in transistors that employ a mesa structure. At high temperatures, charge-generated carriers and diffusion current are the major causes of leakage in both mesa and planar transistor structures; the current $I_G$, therefore, is the dominant leakage component. Because of the dominance of surface leakage $I_S$ at low temperatures and the fact that this leakage may vary either directly or inversely with temperature, it is not possible to define a constant ratio of the leakage current at low temperatures to that at high temperatures. In view of the fact that power transistors are normally operated at high junction temperatures, it is more meaningful to compare the leakage characteristics of both mesa and planar transistors at high temperatures. The relative reliability of different types of power transistors, which is in no way related to the magnitude of low-temperature leakage current, is also best compared at high temperatures.

Leakage currents are important because they affect biasing in amplifier applications and represent the off condition for transistors used in switching applications. The symbol $I_R$ used in the preceding discussion represents any of several different leakage currents commonly specified by transistor manufacturers. The most basic specification is $I_{CBO}$, which indicates the leakage from collector to base with the emitter open. This leakage is simply the reverse current of the collector-to-base diode.

In addition to the $I_{CBO}$ value, $I_{CEV}$, $I_{CEO}$, and $I_{CER}$ specifications are often given for transistors.

$I_{CEV}$ is the leakage from the collector to emitter with the base-emitter junction reverse-biased. $I_{CER}$ is the leakage current from the collector to the emitter with the base and emitter connected by a specified resistance. $I_{CEO}$ is the leakage current from collector to emitter with the base open. $I_{CEV}$ differs from $I_{CBO}$ only very slightly and in most transistors the two parameters can be considered equal. (This equality is not maintained in symmetrical transistors.) $I_{CEO}$ is simply the product of $I_{CBO}$ at the voltage specified and the $h_{FE}$ of the transistor at a base current equal to $I_{CBO}$. $I_{CEO}$ is of course the largest leakage current normally specified. $I_{CER}$ is intermediate in value between $I_{CEV}$ and $I_{CEO}$.

## SPECIAL PHYSICAL CONSIDERATIONS FOR POWER TRANSISTORS

The physical theory of power transistors is complicated by the large current densities and high collector fields involved. In addition to any voltage-breakdown and thermal effects, and to the physical factors discussed in the preceding sections, physical properties of power transistors must be broadened to include three additional effects, as follows:

(a) base-conductivity modulation,
(b) current crowding,
(c) base widening.

### Base-Conductivity Modulation

Base-conductivity modulation is an effect that results in reduced gain of the transistor because of an increase in the base conductivity induced from the high density of the injected carriers from the emitter. As the conductivity ratios of the emitter and base shift, more base majority carriers are injected into the emitter. In a p-n-p transistor, the increased density of the injected carriers greatly affects the ratio of hole current collected to the total emitter current. The relationship is shown by the following equation:

$$\gamma \approx I_p/(I_n + I_p) \qquad (41)$$

where, as indicated previously, $\gamma$ is the emitter injection efficiency, $I_p$ is the hole current, and $I_n$ is the electron current.

For low-current operation, $I_p$ is related to $I_n$ by the ratio of the emitter majority-carrier impurity concentration to the base majority-carrier impurity concentration. For a typical, well-designed p-n-p transistor, $I_n$ is negligible compared to $I_p$.

For high-current operation, $I_n$ increases to levels that approach those of $I_p$ (because of recombination and the requirement for charge neutrality in the base region). This effect reduces $\gamma$ and substantially decreases gain. (If $I_n = I_p$, then $\gamma = 0.5$ and $\beta = 1$.)

The necessity for charge neutrality in the base requires that the base majority carriers ($n_{base}$) increase with an increase in the injected emitter carriers in accordance with the following equation:

$$n_{base} = \frac{J_E}{qV_{base}} + n_{Bo} \qquad (42)$$

where $J_E$ is the emitter current density, $n_{Bo}$ is the initial base majority concentration of impurity, $V_{base}$ is the diffusion- and drift-dependent base voltage, and q is the electron charge constant. This equation defines the change of base current $I_n$ from the increase in base majority carriers ($n_{base}$) with current density $J_E$ and relates its impact on emitter efficiency $\gamma$.

The rate of recombination is also affected by the presence of increased carriers at higher injection levels which decrease the base transport factor $\beta_o$. If it is assumed that there are no base-recombintaion effects, the shape of the curve shown in Fig. 91 is

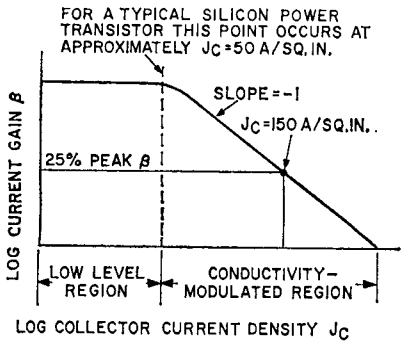



Figure 91. Current gain as a function of collector-current density. (In this diagram, the effects of recombination in the base are neglected.)

governed by the effects of current density on the injection efficiency in accordance with the following equation:

$$\beta = \frac{\sigma_{eo}}{\sigma_{eo}\left(1 + \frac{J_E}{qVn_{Bo}}\right)} \qquad (43)$$

where $\sigma_{eo}$ is the initial conductivity of the emitter, $\sigma_{Bo}$ is the initial conductivity of the base, $J_E$ is the emitter current density, q is the electronic charge constant, V is the carrier velocity of the base, and $n_{Bo}$ is the concentration of initial base impurity doping.

For low-current operation, $\beta$ is equal to $\beta_1$. If $qVn_{Bo}$ is treated as the base modulation constant $K_2$, the value of $\beta$ at a moderate emitter current density $J_E$ is given by

$$\beta = K_2\beta_1/J_E \qquad (44)$$

The graphic representation of this equation (log $\beta$ as a function of log $J_E$), shown in Fig. 91, indicates that $\beta$ varies with a slope of $-1$. For a minimum of conductivity-modulation effects, the initial base concentration $n_{Bo}$ should be as high as practicable with good low-level gain. Also, current density should be low; therefore, the emitter area should be large.

## Current Crowding

The effect of current crowding causes a reduction in the usable gain of a power transistor because of the electrical reduction in the injecting emitter area. This reduction in area results because a diminished forward bias is induced in the central portions of the emitter by the transverse voltage drops in the base. The voltage drops are a result of the flow of base current (through the relatively high base resistance) required to replenish base recombination current and maintain charge neutrality in the base. The current-crowding effect causes the reduced-biased center of the

emitter to stop injecting charge carriers, and the edge becomes the primary injecting area. Simply stated, the emitter edge is biased at a higher emitter-base forward voltage than the center of the emitter.

Physical factors which reduce current crowding are large emitter-periphery-to-area ratios, low base resistance, and wide base widths. (The latter two factors affect the magnitude of the transverse base voltage drop.) Fig. 92 shows a cross section of a typical power transistor under simulated current-crowding conditions.

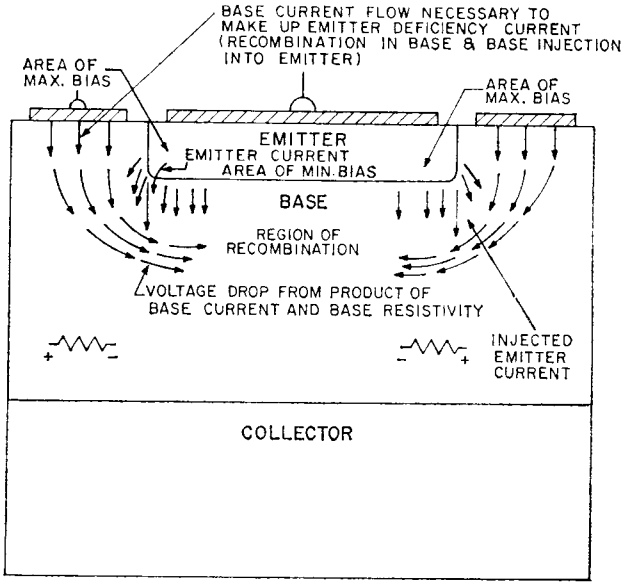



Figure 92. Cross section of a typical n-p-n power transistor under current-crowding conditions.

## Base Widening

Base widening is the effect of the local widening of the electrical base in the areas directly underlying the high injecting regions of the emitter. This condition results when the concentration of injected charge from the emitter exceeds the background doping level of the collector. Fig. 93 shows a cross section of a typical power transistor under simulated base-widening conditions.

As the level of the injected charge exceeds the doped collector impurity level during the transit of injected charges through the collector to the ohmic region, the injected charges dynamically compensate the uncompensated fixed charges (ions) in the collector side of the depletion layer. This compensation makes that local area electrically equivalent to a base region, as shown in Fig. 94. Complete base-widening analysis requires extensive computation,

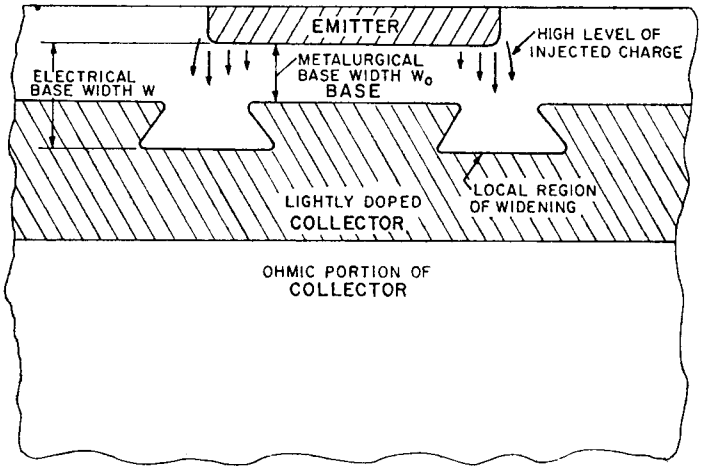



Figure 93. Cross section of a typical power transistor under base-widening conditions.

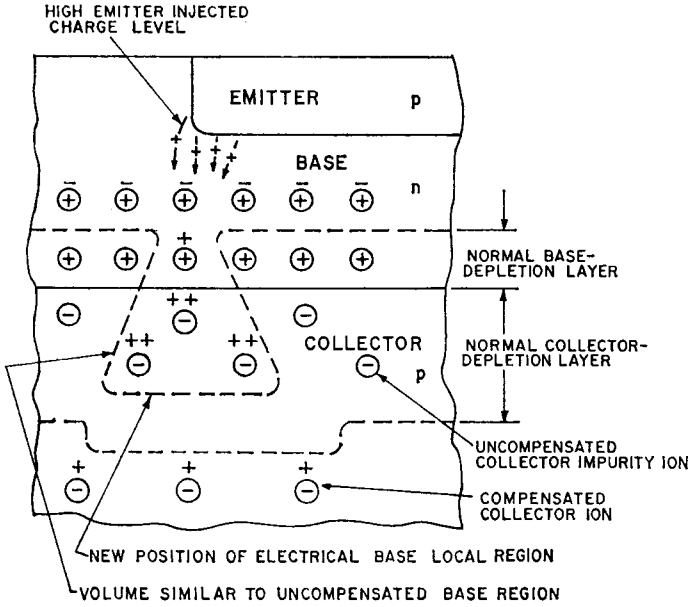



Figure 94. Effect of fixed-impurity compensation (ion compensation).

and the aid of a computer is mandatory for such analyses. Qualitatively, however, transistors with lightly doped collectors suffer more degradation in gain and saturation than transistors with heavily doped collectors, provided that the current density is the same for both types of transistors. The mechanism for reduced gain results from effects on the base transport factor caused by the widened or extended base width and charge pile-up in the vicinity of the collector-base junction.

## DESIGN, PROCESSING, AND PACKAGING

The ultimate aim of all transistor fabrication techniques is the construction of two parallel p-n junctions with controlled spacing between the junctions and controlled impurity levels on both sides of each junction. A variety of structures and geometries have been developed in the course of transistor evolution.

In power transistors, structure refers to the junction depth, the concentration and profile of the impurities (doping), and the spacings of the various layers of the device. Geometry refers to the topography of the transistor. These factors and the method of assembly of the semiconductor pellet into the over-all transistor package have an important bearing on the types of applications in which a power transistor can be used to optimum advantage. The proper choices of trade-offs among these factors determine the gain, frequency, voltage, current, and dissipation capabilities of power transistors.

### Structures

Various structures have been developed to provide different electrical, thermal, or cost properties, with each having certain advantages or compromises to offer. Table VI lists the principal structures available for silicon power transistors, together with some of the advantages and disadvantages of each type.

**Alloy Transistors**—In alloy transistors, impurities are applied directly to the top and bottom surfaces of a carefully prepared slice, or wafer, of silicon. The wafer is then subjected to controlled conditions until the impurity forms an actual metallurgical alloy with the silicon wafer, as shown in Fig. 95.

The alloy fronts form an abrupt (step) junction with the base on both sides. For this type of junction, the base resistivity must be high to support the collector voltage rating. This high resistivity accounts for the extremely wide base width needed to limit punch-through effects. Lifetime must be high in the base to provide adequate gain with the wide base. Relatively low junction-forming temperatures permit preservation of the high initial lifetime of the silicon wafer.

The collector is mounted (soldered) directly to a heat sink to provide the necessary thermal performance. Simple solder contacts are used to make connections with the base (surrounding the emitter) and the emitter directly on top of the deposited impurity.

The principal advantage of the alloy transistor is its extremely rugged junctions which can withstand repeated high-energy power

Table VI - Types of Structures for Silicon Power Transistors

| Structure | Advantages | Disadvantages |
|---|---|---|
| Alloy | Generally rugged | Low speed, high cost |
| Hometaxial-base | Rugged, low cost | Low speed |
| Double-diffused mesa | High speed | Poor saturation resistance |
| Double-diffused planar | High speed, low leakage | Poor saturation resistance |
| Triple-diffused mesa | High speed, low-saturation resistance | Moderate cost, moderate leakage |
| Triple-diffused planar | High speed, low leakage, low saturation resistance | Moderate to high cost |
| Double-diffused epitaxial mesa | High speed, low-saturation resistance | Moderate cost, moderate leakage, less rugged |
| Double-diffused epitaxial planar | High speed, low leakage, low saturation resistance | Higher cost, less rugged |
| Epitaxial-base mesa | Moderate speed, low saturation resistance | Low voltage, moderate leakage |
| Multiple epitaxial-base mesa | Moderate speed, low saturation resistance, rugged, high voltage | Moderate cost |
| Double-diffused multiple-epitaxial mesa | High speed, rugged, low saturation resistance | Moderate cost, moderate leakage |

pulses. This electronic ruggedness results because the very wide base width employed causes the charge carriers to fan out (diffuse) as the carriers travel from the emitter to the collector. In addition, because of the wide base, the transit time of the carriers through the base is relatively long, and lower frequency response and longer switching time result.

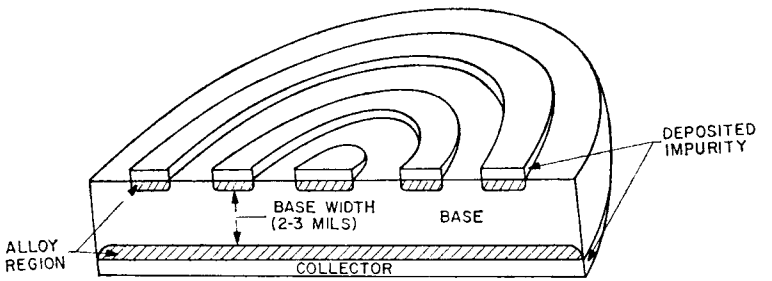



Figure 95. Alloy-transistor structure.

**Hometaxial-Base Transistors**—Hometaxial-base transistors start with a wafer of moderately-high-resistivity silicon on which are deposited several thin layers of impurities. Then, under controlled conditions, the impurities are driven deep into both sides of the silicon wafer. Early in the diffusion, the process is interrupted briefly, and a mesa or raised portion is selectively etched to define an emitter geometry. The process is complete when the deep diffused junctions are separated by a moderately wide (about 1 mil) base region. Fig. 96 shows the development of a typical hometaxial-base transistor, and Fig. 97 shows a typical cross section of a completed single-diffused hometaxial transistor.

The chief advantages of the

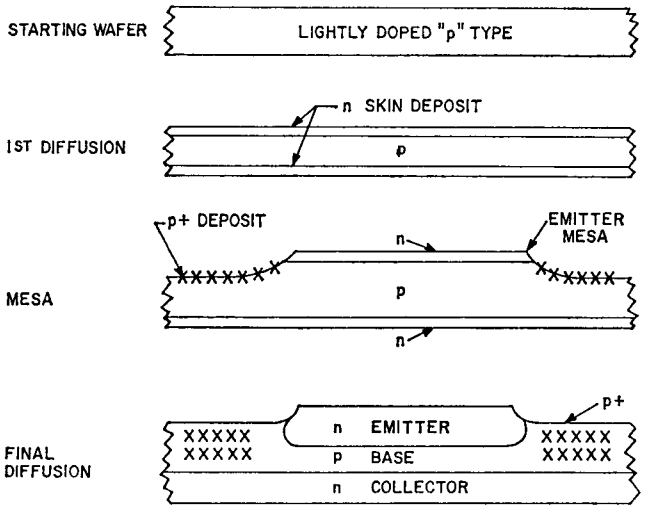



Figure 96. Processing steps in the manufacture of a hometaxial-base transistor.

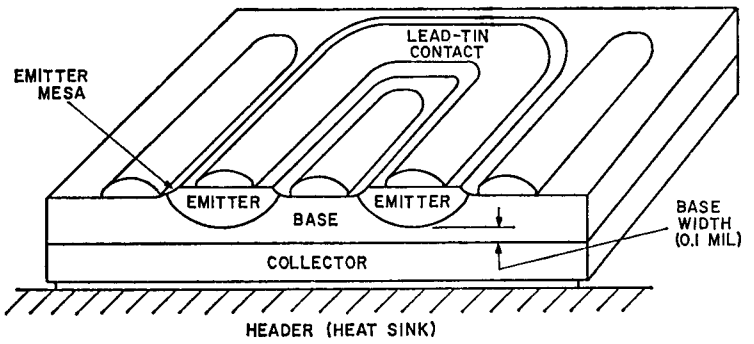



Figure 97. Hometaxial-base (single-diffused) transistor structure.

hometaxial-base transistors are good voltage ratings and excellent electronic ruggedness that permit these transistors to withstand repeated high-energy power pulses. Both advantages result from the very deep graded junctions and the wide base region. The graded junction provides a benefit of either higher voltage ratings with good saturation resistances, or much lower saturation resistances at a given voltage. The electronic ruggedness arises from the moderately wide, undiffused (homogeneous) base region which allows injected charge carriers to fan out and thereby reduce charge-carrier density at the collector junction where heating effects predominate. Another advantage is that the manufacturing cost per unit of power-handling capability is relatively low, primarily as a result of large-batch processing.

Hometaxial-base transistors have a relatively low switching-speed limit because of the moderately wide base spacing, and a low upper voltage limit of about 150 to 200 volts because of punch-through limitations.

**Double-Diffused Transistors**—Double-diffused transistors start with a relatively high-resistivity silicon wafer on which base dopant impurity is deposited. This dopant is then diffused to shallow depths. Then, an oxide ($SiO_2$) is selectively etched to define regions where an emitter impurity is to be deposited and diffused. The oxide acts as an effective mask against the diffusion of most of the usual impurity elements, such as boron or phosphorus. The emitter diffuses more rapidly than the base and, therefore, provides a means to narrow the base width until the desired electrical properties are obtained. The more rapid emitter diffusion results from a much higher impurity level, which enhances the diffusion coefficient as compared to that of the base diffusion. Fig. 98 shows a cross section of a typical double-diffused transistor.

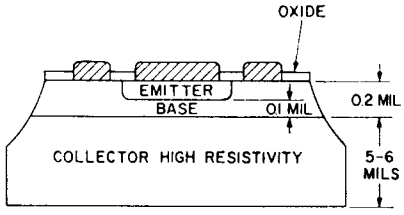



Figure 98. Double-diffused transistor structure.

The double-diffused structure differs from other designs in that the high-resistivity side of the collector-base junction is on the collector side; therefore, the collector voltage can be designed almost independent of the base width. The advantage of the double-diffused transistor is that very narrow non-homogeneous or graded base widths are employed; the frequency responses of these devices, therefore, are orders of magnitude greater than those of earlier types of transistors. Double-diffused transistors, however, have a very high collector saturation resistance and relatively fragile junctions because of the thick high-resistivity collector and narrow graded base width.

**Double-Diffused Planar Transistors**—The double-diffused planar transistor is essentially identical to the double-diffused type with one modification in the planar collector-base junction. The manufacturing process for the

double-diffused planar transistor is similar to that for the double-diffused transistor, with the additional selective masking step for the base impurity. (An oxide similar to the emitter masking step is also used for this mask.) As shown in Fig. 99, the junction terminates at the surface of the silicon wafer instead of on the side. The junction also terminates under a protective oxide layer.

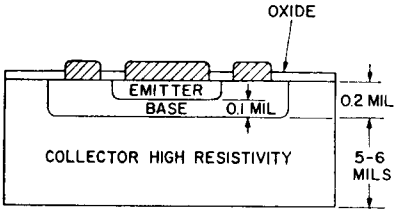



Figure 99. Double-diffused planar transistor structure.

The double-diffused planar transistor features drastically reduced collector leakage currents and better uniformity of device characteristics. The double-diffused planar structure allows the transistor to come very close to the low theoretical limit for silicon junction leakage current.

The disadvantages are similar to those of the double-diffused transistor, in that the double-diffused planar type has a very high collector saturation resistance and relatively fragile junctions. The double-diffused planar transistor has a collector voltage 10 to 20 per cent lower than that of mesa types with the same junction design.

**Triple-Diffused Transistors**—The triple-diffused structure is essentially identical to the double-diffused design except that a third diffusion is performed. The third diffusion, on the opposite side of the silicon wafer, eliminates the major disadvantage of the double-diffused design—high saturation resistance. In the triple-diffused transistor the wafer of silicon is coated with a dopant, followed by a controlled diffusion. Fig. 100 shows a typical cross section of a triple-diffused transistor.

The principal advantage of the triple-diffused structure is that it has low saturation resistance which is of crucial importance in power transistor applications. The saturated switching speeds of this type of transistor are faster than those of the double-diffused design. Both advantages are a result of the thinning down of the high-resistivity section while the bulk of the collector is heavily doped and highly conductive. This technique, however, results in relatively fragile junctions.

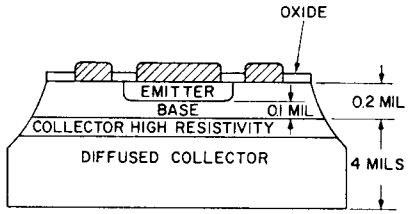



Figure 100. Triple-diffused transistor structure.

**Triple-Diffused Planar Transistors**—The triple-diffused planar transistor, which is similar in structure to the triple-diffused transistor, incorporates a planar collector, as shown in Fig. 101. Critical cross sections of different stages of the manufacturing processes are shown in Fig. 102.

The principal advantages of the triple-diffused planar transistor are very low leakage current, high-speed operation, and low saturation resistance. The

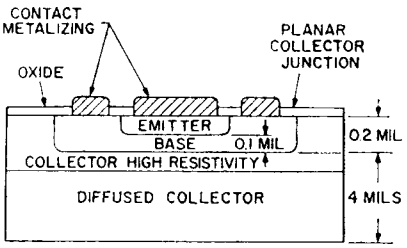



Figure 101. Triple-diffused planar transistor structure.

main disadvantage is that the cost of manufacturing is higher than that of non-planar devices.

collector region is replaced by a heavily doped homogeneous layer referred to as the epitaxial substrate. Because of the difference in doping between the double-diffused epitaxial and triple-diffused structures, some improvements in switching speeds and saturation resistance can be realized. The double-diffused structure, however, has a somewhat poorer reverse "energy profile", so that its capability to withstand inductive or capacitive energy pulses is reduced.

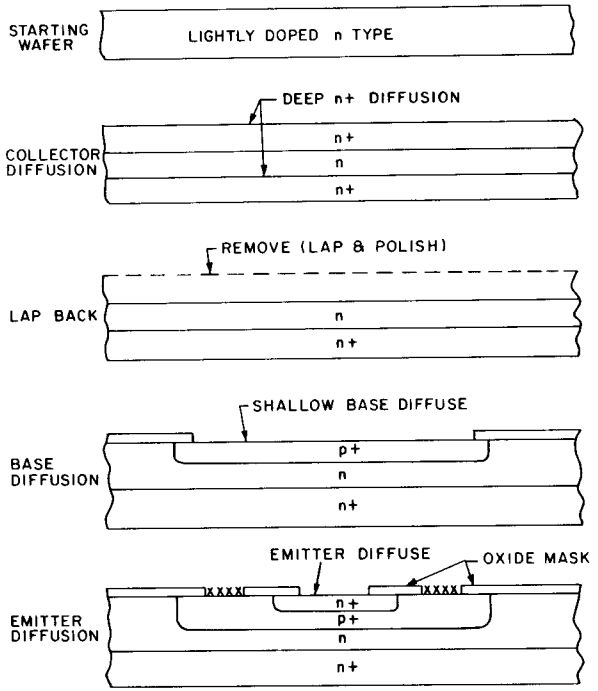



Figure 102. Processing steps in the manufacture of a triple-diffused planar transistor.

**Double-Diffused Epitaxial Transistors**—The double-diffused epitaxial structure is similar in appearance to the triple-diffused design, except that the diffused

Fig. 103 shows a cross section of a typical double-diffused epitaxial transistor, and Fig. 104 shows a planar version of the same kind of transistor.

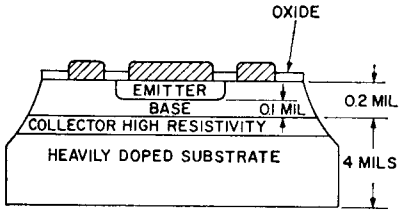



Figure 103. Double-diffused epitaxial transistor structure.

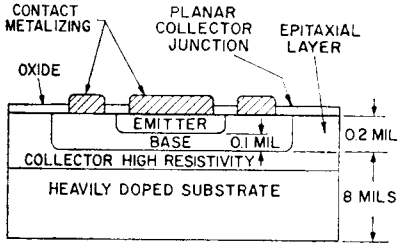



Figure 104. Double-diffused epitaxial planar transistor structure.

**Epitaxial-Base Transistors**—The epitaxial-base structure uses epitaxial layers in the actual formation of the base-collector junction. A single diffusion of the emitter completes this relatively simple design. A layer of impurity (opposite to the substrate impurity) is epitaxially grown on the highly doped substrate. An oxide masking and emitter diffusion into this expitaxial layer completes the construction. Fig. 105 shows a typical cross section of an epitaxial-base power transistor.

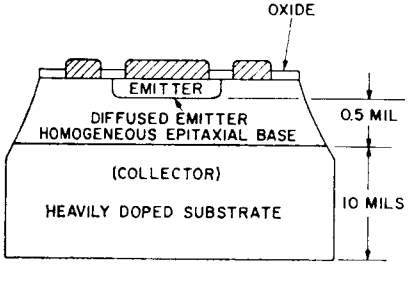



Figure 105. Epitaxial-base transistor structure.

The principal advantage of the epitaxial-base structure, compared to the double-diffused designs, is that it is electronically more rugged (able to withstand energy pulses) as a result of the wider and homogeneous base region. In comparison to the hometaxial structure, the epitaxial-base type has significantly higher frequency response and the ability to carry higher currents for an equivalent emitter area.

The disadvantage of the epitaxial-base design is that it is limited by low voltage ratings imposed by the constraint of the abrupt base-collector junction formed between the heavily doped collector substrate and the epitaxially deposited base layer. The low voltage rating also results from the thin base width necessary for adequate current gain which reduces voltage limits because of punch-through effects. The epitaxial-base transistor also suffers from moderate collector leakage-current levels resulting from step junctions and mesa construction.

**Multiple-Epitaxial-Base Transistors**—The multiple-epitaxial-base structure is similar to the epitaxial-base transistor, but has the added feature of a high-resistivity epitaxial layer for the active collector region. The multiple epitaxial-base transistor is fabricated from a heavily doped silicon wafer on which alternate layers of p-n or n-p high-resistivity silicon are epitaxially grown to create a $\pi$ - $\nu$ or a $\nu$ - $\pi$ base-collector junction. An emitter area is then diffused into the structure. Fig. 106 shows the various stages in the manufac-

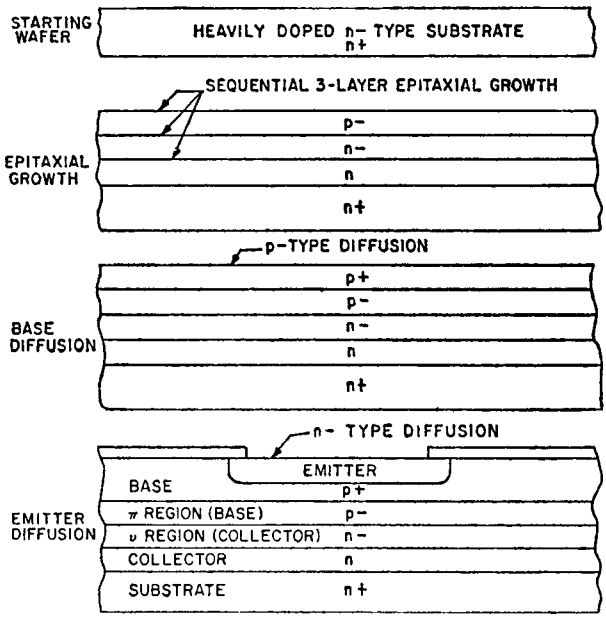



Figure 106. Processing steps in the manufacture of a multiple-epitaxial-base transistor.

ture of the multiple-epitaxial-base transistor structure, and Fig. 107 shows a typical cross section of this type of device.

The principal advantage of the multiple-epitaxial-base structure is that it has high voltage ratings with good current carrying abilities and excellent power-handing capabilities at high voltages (second breakdown). The higher voltage ratings result because the

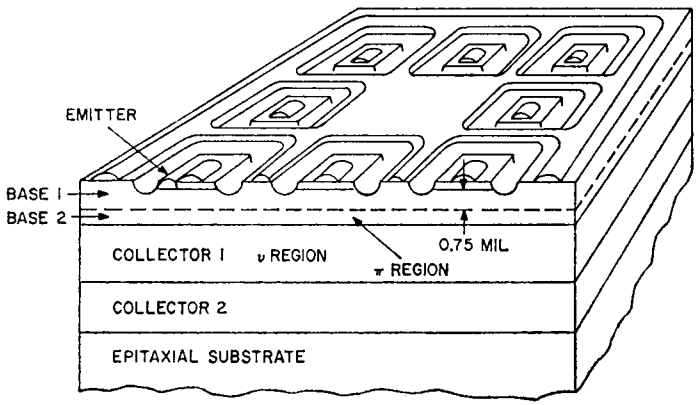



Figure 107. Multiple-epitaxial-base transistor structure.

transistor uses both the base and the collector regions to support the applied collector voltage. The good current-handling characteristic results from the fact that lower collector resistivity can be used for equivalent voltage ratings, as compared to double-diffused epitaxial designs. The lower collector resistivity also minimizes high-current fall-off effects that result from base widening. The excellent second breakdown characteristic results from the moderately wide base width and partial homogeneous base doping, which allows more charge-carrier fan-out (diffusion) and reduced current densities at the collector junction where heating effects predominate.

The principal disadvantage is that the cost of manufacturing the multiple epitaxial-base transistor is relatively high.

**Multiple-Epitaxial Double-Diffused Transistors**—The multiple epitaxial double-diffused structure is almost identical to the double-diffused epitaxial design, with the exception that multiple epitaxial layers are used in the collector region, instead of a single collector layer. The top collector layer is a thin, high-resistivity layer followed by one or more thin, but more heavily doped, layers. These more heavily doped layers are grown sequentially in an epitaxial system onto a thick, heavily doped silicon substrate wafer. Fig. 108 shows the various stages in manufacture of the multiple epitaxial double-diffused structure, and Fig. 109 shows a typical cross section of the completed transistor.

The advantages of the multiple epitaxial double-diffused structure include those of the double-diffused epitaxial design (high

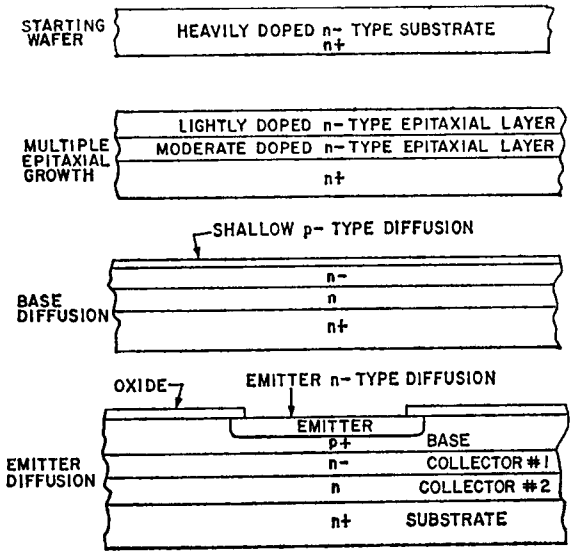



Figure 108. Processing steps in the manufacture of a multiple-epitaxial double-diffused transistor.

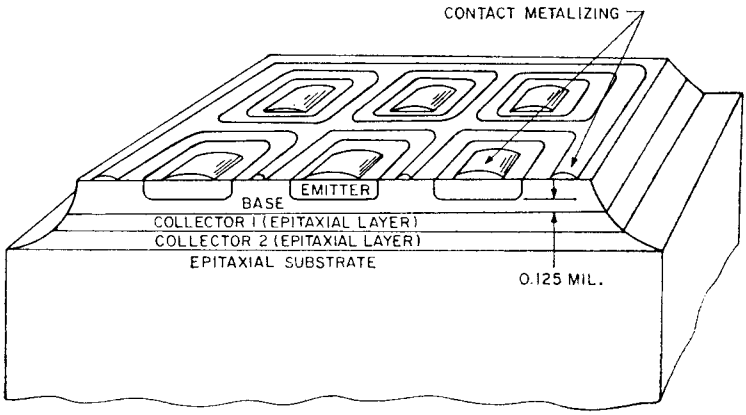



Figure 109. Multiple-epitaxial double-diffused transistor structure.

speed and low saturation), as well as the significant advantages of higher collector-junction voltage ratings, and increased electrical ruggedness. The electrical ruggedness (supplied by the additional collector layers) becomes even more of a factor during power switching with inductive loads in the 100-to-200-volt range where significant inductive energies (reverse second breakdown) may have to be handled by the transistor.

The disadvantages of the multiple epitaxial double-diffused transistor are the moderate-to-high cost per unit and the moderate leakage in the structure.

## Geometries

The topography of a transistor is referred to as its geometry. This transistor geometry, in conjunction with its structure, establishes most of the fundamental transistor electrical, thermal, and economic properties. Proper geometric design of a transistor allows for many compromises, which may result in a variety of advantages and disadvantages from different structures.

The basic premise for most geometry designs for power transistors is to increase current handling per unit area of device. This condition results in lower-cost designs or, as in high-frequency transistors, higher-speed operation as a result of the smaller device areas.

Power-transistor geometries have evolved from the very early inefficient "ring-dot" configurations to the present-day sophisticated "overlay" concepts. Fig. 110 shows some typical geometry milestones in this evolutionary cycle.

The early geometries were characterized by simple shapes, large dimensional tolerances, and poor utilization of active regions. As the state of the art in fine-line mask making and wafer printing improved, the geometries became more involved, with much finer dimensions.

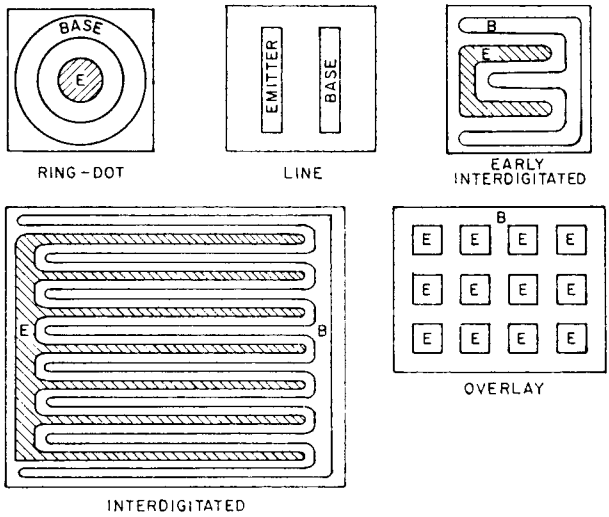



Figure 110. Typical geometries in the development of transistors.

In silicon power transistors, current crowding is the greatest contributor to reductions in current gain. Emitter periphery is the crucial design factor to reduce high current density caused by current crowding. Modern technology, however, has produced the means to yield more than 10 inches of periphery in less than 0.050 square inch of surface area.

Certain device structures have constraints on how fine the emitter geometry can be made. Refinement of emitters is governed by the space needed for emitter and collector mesas and by the thickness of oxide masks needed for deep diffusion, as well as by other factors. Table VII shows progressive geometry refinements of some power-transistor structures.

## Packaging

Three basic operations are involved in the assembly of a transistor pellet into a suitable package. These operations include: (1) attachment of the transistor

Table VII—Geometry Refinements

| Structure | Smallest Practical Emitter Width (nominal) mils | Emitter Mask Tolerance mils |
|---|---|---|
| Alloy | 25 | ±3 |
| Hometaxial-base | 10 | ±1 |
| Triple-diffused mesa (deep junctions) | 2 | ±0.3 |
| Double-diffused epitaxial planar (shallow junctions) | 0.1 | ±0.01 |

pellet to the package, (2) connection of the transistor emitter and base contact to the external leads of the package, and (3) formation of the metal-ohmic transistor contacts.

**Pellet Attachment**—Three technologies predominate for attachment of transistor pellets to the transistor package. Soft solder, hard solder, or polymer adhesive may be used. Soft-solder and hard-solder methods are used almost exclusively for transistors designed to operate at power levels greater than one watt. The three approaches with their advantages and disadvantages are shown in Table VIII. Figs. 111 and 112 show cross sections for soft- and hard-solder methods of pellet attachment.

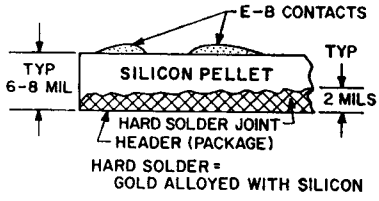



Figure 111. Hard-solder method of pellet attachment.

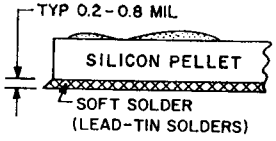



Figure 112. Soft-solder method of pellet attachment.

**Lead Attachment**—Two general methods encompass the technologies used for connecting the ohmic portion of the transistor pellet emitter and base contacts to the external leads of the package: wires (bonds) and soldered contacts (clips). Table IX shows the broad range of connections used for power transistors with some of their advantages and disadvantages. Fig. 113 shows cross sections that illustrate the two methods of lead connection.

**Metal-Ohmic Contacts** — The principal metal-semiconductor contact technology is divided into the following two categories: evaporated-and-alloyed metals and plated-and-sintered metals.

The evaporated-and-alloyed metals consist of aluminum, gold, and silver. Aluminum is used in almost all present designs because of its ease of evaporation and alloying to the semiconductor, relative stability, and good metallurgical strength after wire bonding.

The plated-and-sintered contacts are used when large, heavy solder-coated contacts are needed. Large batch-plating and heat-treating (sintering) processes make these contacts very desirable for low-cost devices.

Table VIII—Attachment Methods

| Attachment Method | Materials | Advantages | Disadvantages |
|---|---|---|---|
| Soft solder | Lead-tin | Low cost, high-dissipation capability | Can be fatigue prone |
| Hard solder | Gold-silicon | High-dissipation capability, strong | High attachment cost for large chips, needs molybdenum block |
| Polymer adhesive | Filled epoxies | Low cost | Poor thermal properties |

Table IX—Methods of Lead Attachment

| Lead Connection | Method of Attachment | Materials Used | Advantages | Disadvantages |
|---|---|---|---|---|
| Thermo-compression bond | High temperature and pressure | Gold-wire ribbon | Very small areas | Costly in large devices |
| Nailhead bond | High temperature and pressure | Gold wire with end balled | Stronger than thermo-compression bond, less costly | Larger contact area required |
| Ultrasonic bond | Ultrasonic weld | Aluminum or gold wire | Avoid gold-alum. problems | Costly in large devices |
| Wire solder | Insert wire in molten solder | Suitable solderable wires | Moderate cost | Large contact area required |
| Clip solder | Pre-set into clips solder | Phosphor-bronze or nickel | Low cost | Large contact areas required |

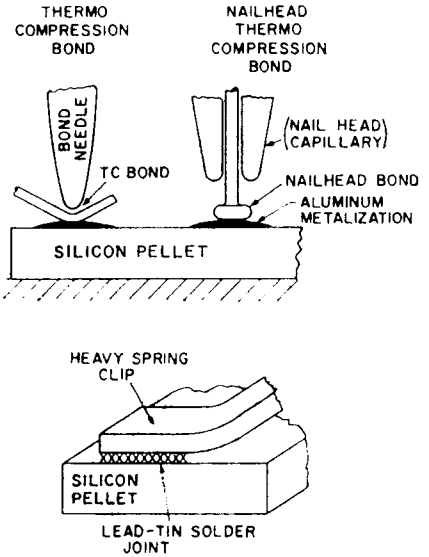



Figure 113. Two general methods of lead attachment.

Fig. 114 shows typical evaporatedand plated techniques.

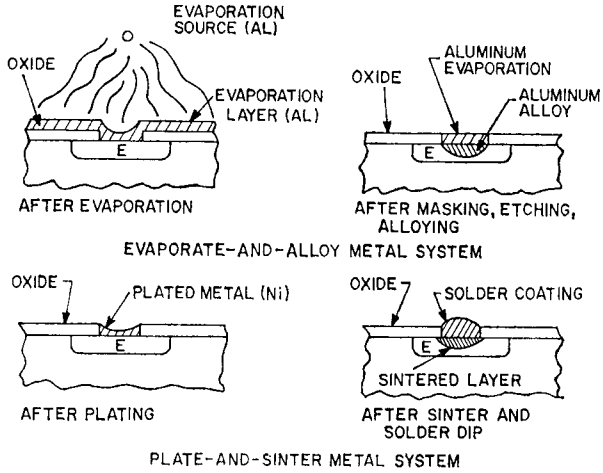



**Figure. 114. Typical evaporated and plated semiconductor contacts.**

# Physical Basis for Power-Transistor Ratings

ALL semiconductor devices undergo irreversible changes if their temperature is increased beyond some critical limit. A number of ratings are given for power transistors, therefore, to assure that this critical temperature limit will not be exceeded on even a very small part of the silicon chip. The ratings for power transistors normally specify the maximum voltages, maximum current, maximum and minimum operating and storage temperatures, and maximum power dissipation that the transistor can safely withstand.

## VOLTAGE RATINGS

Maximum voltage ratings are normally given for both the collector and the emitter junctions of a transistor. A $V_{BEO}$ rating, which indicates the maximum base-to-emitter voltage with the collector open, is usually specified. The collector-junction voltage capability is usually given with respect to the emitter, which is used as the common terminal in most transistor circuits. This capability may be expressed in several ways. A $V_{CEO}$ rating specifies the maximum collector-to-emitter voltage with the base open; a $V_{CER}$ rating for this voltage implies that the base is returned to the emitter through a specified resistor; a $V_{CES}$ rating gives the maximum voltage when the base is shorted to the emitter; and a $V_{CEV}$ rating indicates the maximum voltage when the base is reverse-biased with respect to the emitter by a specified voltage. A $V_{CEX}$ rating may also be given to indicate the maximum collector-to-emitter voltage when a resistor and voltage are both connected between base and emitter.

If a maximum voltage rating is exceeded, the transistor may "break down" and pass current in the reverse direction. The breakdown across the junction is usually not uniform, and the current may be localized in one or more small areas. The small area becomes overheated unless the current is limited to a low value, and the transistor may then be destroyed.

The collector-to-base or emitter-to-base breakdown (avalanche) voltage is a function of the resistivity or impurity doping concentration at the junction of the transistor and of the characteristics of the circuit in which the transistor is used. When there is a breakdown at the junction, a

sudden rise in current (an "avalanche") occurs. In an abruptly changing junction, called a step junction, the avalanche voltage is inversely proportional to the impurity concentration. In a slowly changing junction, called a graded junction, the avalanche voltage is dependent upon the rate of change of the impurity concentration (grade constant) at the physical junction. Fig. 115 shows the two types of junction breakdowns. The basic transistor voltage-breakdown mechanisms and their relationship to external circuits are the basis for the various types of voltage ratings used by transistor manufacturers.

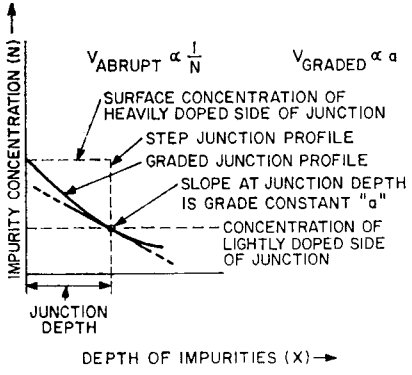



Figure 115. Step-junction and graded-junction breakdown.

## Collector Punch-Through Voltage

The collector voltage can be limited below its avalanche breakdown value if the depletion layer (space-charge region) associated with the applied collector voltage expands through the thin base width and contacts the emitter junction. The doping in the base (under the emitter) and the base width in relation to the magnitude of applied voltage govern whether punch-through occurs before avalanche. Higher doping concentrations and wider bases increase punch-through voltage $V_{PT}$ in accordance with the following relationship:

$$V_{PT} = \frac{qNW^2}{2kE_o} \qquad (45)$$

where q is the electronic charge, N is the doping-level concentration in the base, W is the base width, k is the dielectric constant, and $E_o$ is the permittivity of free space ($kE_o$ is approximately $1 \times 10^{-12}$ farad per centimeter for silicon).

## Common-Base Avalanche Breakdown

Collector-base breakdown of transistors in a common-base connection is caused by avalanche multiplication. When a voltage is applied between collector and base, a depletion layer or space-charge layer is formed at the collector junction and spreads out into both the collector and base regions. Avalanche multiplication takes place in this depletion layer when a high electric field is present. This multiplication effect, which is similar to the "Townsend effect" in gas tubes, is the result of collisions between rapidly accelerating minority carriers that enter the depletion layer and atoms in the crystal lattice. Energy transferred to the atoms as a result of these collisions causes ionization, which releases valence electrons; these electrons are then also accelerated. Avalanche breakdown differs from zener breakdown in

which no multiplication takes place because no free carriers are present in the zener condition. All the carriers of the zener breakdown are formed by stripping of valence electrons in a high-strength field.

The multiplication M that takes place for a given collector-to-base voltage ($V_{CB}$) is given by the following empirical formula for junction transistors:

$$M = \frac{1}{1 - \left(\frac{V_{CB}}{V_A}\right)^n} \qquad (46)$$

where $V_A$ is the true avalanche or "bulk" breakdown and n is the rate of multiplication (i.e., avalanche factor); both terms are constant for a device of a given type. These constants are determined for a particular transistor, as described in the following paragraphs.

For a common-base circuit using constant-current input, the collector current $I_C$ is given by

$$I_C = \alpha M I_E + M I_{CBO} \qquad (47)$$

where $\alpha$ (alpha) is the short-circuit common-base current transfer ratio, $I_E$ is the emitter current, and $I_{CBO}$ is the collector-to-base leakage current. Both $I_E$ and $I_{CBO}$ are multiplied by the multiplication factor M because they cross the depletion layer (the ohmic leakage components of $I_{CBO}$ which do not cross the depletion layer and are not affected by multiplication are not considered here).

If the operating point of a transistor in a common-base circuit is selected so that $I_E$ is much greater than $I_{CBO}$, then Eq. (47) can be simplified as follows:

$$I_C \approx \alpha M I_E \qquad (48)$$

The multiplication factor M is then given by

$$M = \frac{1}{\alpha} \cdot \frac{I_C}{I_E} \qquad (49)$$

Because the collector current $I_C$ is related to the multiplication factor M, which is in turn related to the collector-to-base voltage $V_{CB}$, particular values of M for values of $V_{CB}$ can be obtained from the following re-arrangement of Eq. (46):

$$\log \frac{M - 1}{M} = n \log \frac{V_{CB}}{V_A} \qquad (50)$$

This equation indicates that a log-log plot of $(M - 1)/M$ as a function of $V_{CB}$ is a straight line having a slope n and a value of $V_{CB}$ equal to the true avalanche breakdown $V_A$ when $(M - 1)/M$ is unity, or when M approaches infinity.

Eq. (48) shows that the total gain factor for a transistor in a common-base circuit is reflected by the product $\alpha M$. In addition to the multiplication factor M, therefore, the total gain depends on the short-circuit transfer ratio $\alpha$, which, as indicated by Eq. (26), can be closely approximated by the product of the base transport factor and the emitting efficiency (i.e., $\alpha \approx \beta_0 \gamma$) if the collector efficiency $\alpha^*$ is assumed to be very near unity.

The emitting efficiency $\gamma$ is the ratio of the carriers injected into the base from the emitter to the sum of these carriers plus the carriers injected into the emitter

from the base; it is given by

$$\gamma = 1 - \frac{D_b W N_b}{D_e L_b N_e} \qquad (51)$$

where $D_b$ and $D_e$ are the minority-carrier diffusion constants of the base region and the emitter region, respectively, and $N_b$ and $N_e$ are the carrier concentrations of the base and emitter, respectively.

In a practical transistor, the diffusion length L is much greater than the active base width W, and the emitter is much more heavily doped than the base (i.e., the emitter conductivity $\sigma_e$ is much greater than the base conductivity $\sigma_b$). As a result of the heavier doping, the emitter carrier concentration $N_e$ is much greater than the base carrier concentration $N_b$. Eqs. (30) and (51) indicate that for these conditions ($L \gg W$ and $N_e \ll N_b$) both the transport factor $\beta_o$ and the emitter efficiency $\gamma$ are approximately equal to unity.

Fig. 116 shows the collector characteristics for a transistor

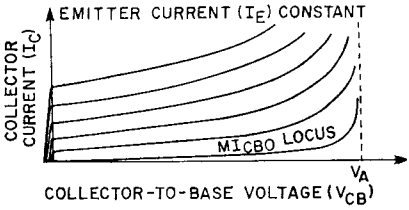



Figure 116. Collector current-voltage characteristics of a transistor operated in a common-base circuit.

operated in a common-base circuit with a constant emitter current. The total gain factor $\beta_o\gamma$ M varies from a value of $\beta_o\gamma$ at low voltages, where $\beta_o\gamma$ is close to unity and M equals unity, to a value approaching infinity when $V_{CB}$ equals $V_A$ (because M approaches infinity at this voltage). Because stable operation can be achieved as long as the total gain remains finite, operation of transistors in common-base circuits is permissible at all voltages up to the collector-base avalanche voltage $V_A$.

## Common-Emitter Avalanche Breakdown

In common-emitter circuits, avalanche breakdown occurs at the collector-to-base voltage at which the common-emitter current-transfer ratio ($\beta$) becomes infinite. $\beta$ can be expressed in terms of the common-base total gain factor $\alpha$M, as follows:

$$\beta = \frac{\alpha M}{1 - \alpha M} \qquad (52)$$

$\beta$ becomes infinite when $\alpha$M equals unity (i.e., when $M = 1/\alpha = 1/(\beta_o\gamma)$.

Avalanche multiplication increases the number of carriers supplied to the collector side of the junction from the depletion layer. The base is then required to supply a similar number of new carriers to the depletion layer to maintain charge neutrality in the layer. At the collector voltage at which the number of carriers supplied to the depletion layer by the base because of multiplication just equals the number of carriers gained by the base through recombination (transport factor $\beta_o$) plus an effective number of opposite-type carriers injected by the base (emitting efficiency $\gamma$),*

* The injection of opposite-type carriers by the base is equivalent to a corresponding gain of similar carriers in the base.

the current transfer ratio becomes infinite because no base current is required to support collector-current flow.

As stated above, $\beta$ becomes infinite when $\alpha M$ equals unity, or when $M = 1/\alpha$. Substitution of this value in Eq. (46) produces the following equation for $\alpha$:

$$\alpha = 1 - \left(\frac{V_{CB}}{V_A}\right)^n \qquad (53)$$

This equation can then be solved for the value of $V_{CB}$ at which $\alpha M$ equals unity (this voltage is represented by $V^{\alpha}_{M=1}$), as follows:

$$V_{\alpha M = 1} = V_A (1 - \alpha)^n \qquad (54)$$

For collector voltages smaller than $V^{\alpha}_{M=1}$, base current $I_B$ flows in the normal direction and $\beta$ is positive. For voltage greater than $V^{\alpha}_{M=1}$, however, the base-current is reversed and $\beta$ is negative. Fig. 117 shows the common-emitter current-transfer ratio $\beta$ and the total common-base gain factor $\alpha M$ as functions of collector-to-base voltage $V_{CB}$.

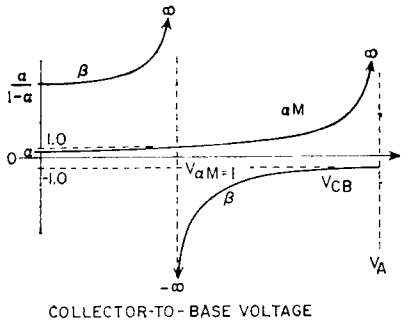



Figure 117. Common-emitter current gain $\beta$ and total common-base current gain $\alpha M$ as a function of collector-to-base voltage.

The collector current $I_C$ of a transistor operating in a common-emitter circuit with a constant-current input is given by

$$I_C = \beta I_B + (\beta + 1) M I_{CBO} \qquad (55)$$

Fig. 118 shows the collector characteristics for a transistor that operates under such conditions.

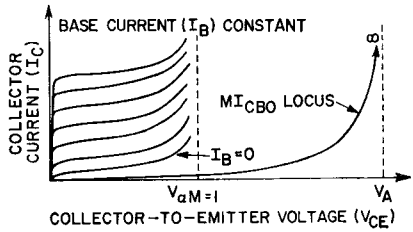



Figure 118. Collector characteristics for a transistor operating in a common-emitter circuit with a constant-current input.

Although the abscissa for these curves is the collector-to-emitter voltage rather than the collector-to-base voltage previously used, no appreciable difference exists between the two set of collector characteristics except at low collector voltages, where multiplication is negligible in any case.

If negative feedback in the form of emitter resistance is applied to a transistor operating in a common-emitter circuit without constant-current input, as shown in Fig. 119, the net effect is an increase in the avalanche breakdown. In the circuit shown in Fig. 119, $R_B$ is the series Thevenin equivalent of all external resistances presented to the transistor base terminal, $R_E$

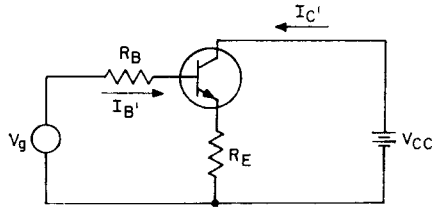



Figure 119. Common-emitter circuit with negative feedback.

is the sum of both external and internal emitter resistances, and $V_g$ is the voltage source or Thevenin voltage at the base terminal.

The base-to-emitter voltage $V_{BE}$ can be assumed to be approximately zero, provided the internal base resistance is small compared to $R_B$. The base current $I_B$ is then given by

$$I_B = \frac{V_g}{R_B} \tag{56}$$

The collector current $I_C'$ for the circuit with external emitter resistance can be determined in terms of initial base current $I_B$, as follows:

$$I_C' = \beta I_B' = \frac{\beta R_B}{R_B + (\beta + 1) R_E} \times I_B \tag{57}$$

Because the input is a finite source voltage, the effect of the external emitter resistance is to reduce the output or collector current. An artificial current ratio $\beta'$ can be introduced to account for the change in $I_C$, as follows:

$$\beta' = \frac{\alpha'}{1 - \alpha'} = \frac{R_B}{R_B + (\beta + 1) R_E} \times \beta \tag{58}$$

Eq. (58) can then be solved to determine an artificial common-base current transfer ratio $\alpha'$, as follows:

$$\alpha' = \alpha \times \frac{R_B}{R_B + R_E} \tag{59}$$

This value of $\alpha'$ is not the true common-base current transfer ratio of the transistor, but it defines the feedback effect which results from the use of external emitter resistance when any type of source other than a pure current source is applied to the transistor in the common-emitter circuit. The term $\alpha'$ can be used to determine the common-emitter avalanche voltage for non-constant-base-current conditions when external emitter resistance is used. Combination of Eqs. (54) and (59) provides the avalanche voltage, as follows:

$$V_{\alpha' M = 1} = V_A \left(1 - \frac{R_B}{R_B + R_E} \alpha\right)^{1/n} \tag{60}$$

The collector characteristics for these conditions are similar to the characteristics shown in Fig. 117, except that the voltage $V_{\alpha M = 1}$ is replaced by the higher voltage $V_{\alpha' M = 1}$ as defined in Eq. (60). If $R_B$ becomes infinite or $R_E$ becomes zero, the condition for constant-base-current operation is reached and $V_{\alpha' M = 1}$ reduces to $V_{\alpha M = 1}$. If $R_E$ becomes infinite or $R_B$ becomes zero, $V_{\alpha M = 1}$ reduces to $V_A$, the common-base avalanche breakdown voltage. Therefore when a source voltage and external emitter resistance are used, the common-emitter avalanche breakdown voltage can vary from a low of $V_{\alpha M = 1}$ to a high of $V_A$, depending upon the ratio of $R_B$ to $R_E$.

## Common-Emitter Voltage Breakdown as a Function of Circuit Conditions

The preceding discussion of common-emitter voltage breakdown considers only forward-bias conditions. Other types of breakdown may occur for circuit input

conditions when no forward bias is applied. Several of these conditions are discussed below.

**Resistive Source**—When a transistor is operated in a common-emitter circuit from a resistive source R, as shown in Fig. 120, the collector current $I_C$ is given by

$$I_C = \frac{MI_{CBO}}{1 - \alpha_N \alpha_i} \left[ 1 + \frac{\alpha_N (1 - \alpha_i)}{(1-\alpha_N) + \frac{kT}{q} \frac{(1-\alpha_N \alpha_i)}{I_{EBO} R}} \right] \quad (61)$$

where $\alpha_N$ is the normal common-base current transfer ratio for the transistor ($\alpha_N = \beta_o \gamma$), $\alpha_i$ is the current transfer ratio for inverted operation, $I_{EBO}$ is the emitter-to-base leakage current, and the term kT/q is equal to 0.026 volt at 25°C.

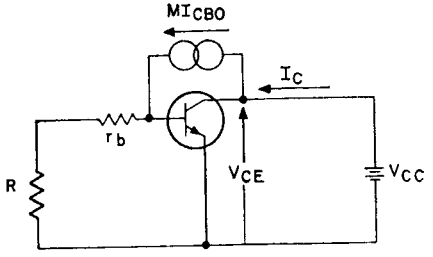



Figure 120. Common-emitter circuit operated from a resistive source.

The total collector leakage current $MI_{CBO}$ divides at the internal base terminal; a portion flows through the internal base resistance $r_b$ and the source resistance R, and the balance flows through the base of the transistor to produce the collector current given by Eq. (61). The voltage produced by the portion of the current that flows through the series combination $r_b$ and R provides a forward bias between the emitter and the base.

If the intrinsic emitter-base diode has a step-function current-voltage characteristic with a threshold voltage $V_d$, rather than an exponential characteristic, and if all leakage current flows through the external base current as long as the forward bias is less than $V_d$, emitter injection takes place when the emitter forward bias equals $V_d$, and collector-to-emitter voltage breakdown occurs. The breakdown condition is given by

$$MI_{CBO} (R + r_b) = V_d \quad (62)$$

Because M is related to $V_{CB}$ and $V_{CE}$, Eq. (62) can be solved for $V_{CE}$ for any given value of $V_{CB}$. The calculated value of $V_{CE}$ is then designated as the collector-to-emitter breakdown voltage with source resistance R, and has the symbol $V_{(BR)CER}$. The value of $V_{(BR)CER}$ is given by

$$V_{(BR)CER} = V_A \left[ 1 - \frac{I_{CBO} (R + r_b)}{V_d} \right]^{1/n} \quad (63)$$

Eq. (63) indicates that $V_{CE}$ is inversely proportional to the logarithm of R. Therefore, the highest breakdown voltage occurs when R is equal to zero. This voltage is designated as the shorted-base breakdown voltage, and has the symbol $V_{(BR)CES}$.

If the base is opened (R approaching infinity), the threshold voltage $V_d$ is reached for any finite value of $MI_{CBO}$, and transistor operation is governed by the common-emitter current-transfer ratio $\beta$. For this condition, the entire leakage current $MI_{CBO}$ must flow through the

transistor base to produce a collector current equal to $(\beta + 1)$ times $MI_{CBO}$.* This lowest value of breakdown voltage occurs at the collector-to-emitter voltage at which $\beta$ becomes infinite, which was previously defined as $V_{\alpha M = 1}$.

The breakdown voltage for all other source-resistance conditions is greater than $V_{\alpha M = 1}$; i.e., when emitter injection starts, the total gain factor ($\alpha$M) is greater than unity, and $\beta$ is negative. Fig. 117 shows that when $V_{CE}$ is greater than $V_{\alpha M = 1}$, $\beta$ increases negatively as voltage decreases. At breakdown, emitter injection occurs, and the collector current increases rapidly. This increasing current causes a decrease in collector voltage as a result of the presence of collector, emitter, and supply resistances. The decreasing collector voltage in turn

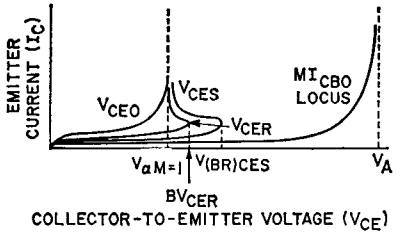



Figure 121. Breakdown characteristics of a transistor in a common-emitter circuit operated from a resistive source.

causes an increase in $\beta$ and collector current, so that the effect becomes cumulative. This effect produces a negative-resistance breakdown-voltage characteristic that becomes asymptotic to $V_{\alpha M = 1}$ when $\beta$ is infinite.

* The intrinsic collector current $I_C''$ is $\beta$ times the intrinsic base current $I_B''$ (for this case $I_B = MI_{CBO}$). The actual measured collector current is the intrinsic collector current plus the leakage current, i.e., $I_C = I_C'' + MI_{CBO}$ and $I_C'' = \beta I_B'' = \beta MI_{CBO}$. Therefore, $I_C = \beta I_B'' + MI_{CBO}$, which reduces to $I_C = (\beta + 1)$.

The source-resistance breakdown characteristics are shown in Fig. 121.

**Reverse-Bias Voltage Source**—When a reverse bias is applied between emitter and base, as shown in Fig. 122, the collector

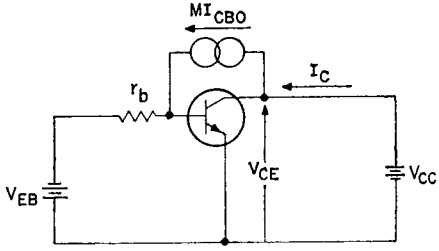



Figure 122. Common-emitter circuit with a reverse bias applied between input terminals.

breakdown voltage can be increased above the value $V_{(BR)CES}$. As in the case of source resistance, no emitter injection takes place as long as the forward emitter bias is less than the threshold voltage $V_d$. Injection occurs when the drop across $r_b$ that results from the current component $MI_{CBO}$ is sufficient to overcome both the base supply voltage $V_{BB}$ and $V_d$. This breakdown condition is given by

$$MI_{CBO}\, r_b = V_D + V_{BB} \qquad (64)$$

An increase in $V_{BB}$ increases the value of both M and $V_{CE}$. Fig. 123 shows a series of breakdown

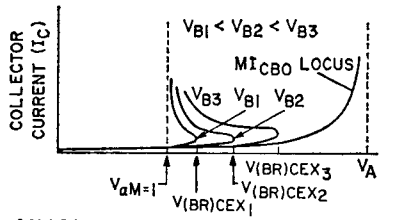



Figure 123. Common-emitter breakdown characteristics for different values of base supply voltage.

curves for different values of $V_{BB}$. Negative resistance occurs when the transistor operates in the region of negative $\beta$, as discussed previously. The peak value of each characteristic is designated by $V_{(BR)CEX}$. The value of $V_{(BR)CEX}$ is given by

$$V_{(BR)CEX} = V_A\left[1 - \frac{I_{CBO}(R + r_b)}{V_d + V_{BB}}\right]^{1/n} \quad (65)$$

## Transistor Operating Regions

The various breakdown voltages discussed up to this point determine the operating regions for general-purpose transistors. In general, transistor characteristics can be divided into two regions of operation, as shown in Fig. 124.

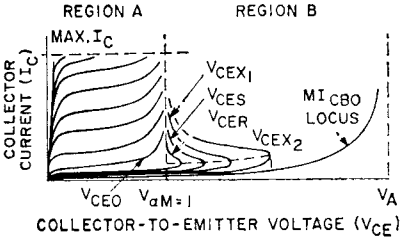



Figure 124. Common-emitter breakdown characteristics showing that a transistor has two operating regions.

The limits of region A, the forward-bias region, are determined by the common-emitter avalanche breakdown voltage $V_{\alpha M=1}$ and the maximum collector-current rating for the transistor. Operation anywhere in this region is permissible provided the peak dissipation ratings for the transistor are not exceeded. There are no restrictions on input-circuit conditions unless the region boundary is set by $V_{\alpha' M=1}$ rather than $V_{\alpha M=1}$; in this case, the conditions discussed previously apply.

The lower limit of region B, the negative-resistance area, is determined by the avalanche breakdown voltage $V_{\alpha M=1}$, and the upper limit by the respective breakdown voltages for particular input conditions, i.e., $V_{(BR)CES}$, $V_{(BR)CER}$, and $V_{(BR)CEX}$.

## Additional Considerations

In the previous discussion of common-emitter avalanche breakdown voltage, the term $V_{\alpha M=1}$ was assumed to be independent of collector current. However, $V_{\alpha M=1}$ is a function of the common-base current transfer ratio $\alpha$ [as shown in Eq. (66)], which varies with $I_C$. It follows, therefore, that $V_{\alpha M=1}$ must change with $I_C$. The current-gain parameters $\beta$ and $\alpha$ vary with $I_C$ differently for abrupt- and graded-junction transistors. Fig. 125 shows the variation of $\beta$ for typical transistors.

Because the minimum value of $V_{\alpha M=1}$ occurs at the peak of the curves shown in Fig. 125, it is possible to construct a locus of points on the $V_C$-$I_C$ curves of a transistor at which the total gain

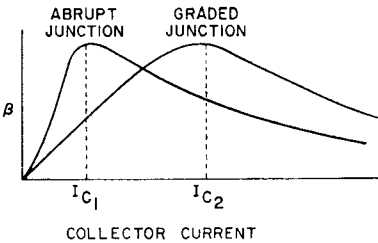



Figure 125. Variation in the current-gain parameter $\beta$ as a function of collector current for both abrupt-junction and graded-junction transistors.

factor $\alpha M$ is equal to unity, as shown in Fig. 126. This locus curve has only a positive-resistance slope for abrupt-junction transistors, but has both positive- and negative-resistance slopes for graded-junction transistors.

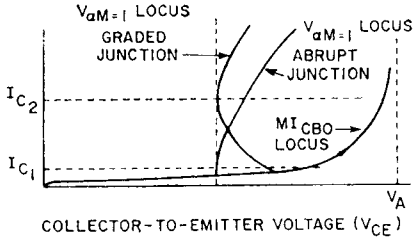



Figure 126. Curves of collector current as a function of collector-to-emitter current for a unity total gain factor in both abrupt-junction and graded-junction transistors.

Because both the forward-bias and reverse-bias curves become asymptotic to $V_{\alpha M=1}$ for common-emitter operation, this variation of $V_{\alpha M=1}$ with $I_C$ modifies all the breakdown curves. It also explains why some forward-bias curves, such as $V_{CEO}$, can have a negative-resistance component for some types of transistors. This effect is observed for most diffused types that have graded junctions; because alloy transistors have abrupt junctions, these types do not normally have negative-resistance forward-biased voltage characteristics.

## CURRENT AND TEMPERATURE RATINGS

The physical mechanisms related to basic transistor action are temperature-sensitive. If the bias is not temperature-compensated, the transistor may develop a regenerative condition, known as **thermal runaway**, in which the thermally generated carrier concentration approaches the impurity carrier concentration. [Experimental data for silicon show that, at temperatures up to 700°K, the thermally generated carrier concentration $n_i$ is determined as follows: $n_i = 3.87 \times 10^{16} \times T \times (3/2) \exp(-1.21/2kT)$.] When this condition becomes extreme, transistor action ceases, the collector-to-emitter voltage $V_{CE}$ collapses to a low value, and the current increases and is limited only by the external circuit.

If there is no current limiting, the increased current can melt the silicon and produce a collector-to-emitter short. This condition can occur as a result of a large-area average temperature effect, or in a small area that produces hot spots or localized thermal runaway. In either case, if the intrinsic temperature of a semiconductor is defined as the temperature at which the thermally generated carrier concentration is equal to the doped impurity concentration, the absolute maximum temperature for transistor action can be established.

The intrinsic temperature of a semiconductor is a function of the impurity concentration, and the limiting intrinsic temperature for a transistor is determined by the most lightly doped region. It must be emphasized, however, that the intrinsic temperature acts only as an upper limit for transistor action. The maximum operating junction temperature and the maximum current rating are established by additional factors such as the efficiency of heat removal, the yield point and melting point of the solder used in fabrication, and the temperature at which permanent changes in the junction properties occur.

The **maximum current rating** of a transistor indicates the highest current at which, in the manufacturer's judgment, the device is useful. This current limit may be established by setting an arbitrary minimum current gain or may be determined by the fusing current of an internal connecting wire. A current that exceeds the rating, therefore, may result in a low current gain or in the destruction of the transistor.

The basic materials in a silicon transistor allow transistor action at temperatures greater than 300°C. Practical transistors, however, are limited to lower temperatures by mounting systems and surface contamination. If the **maximum rated storage** or **operating temperature** is exceeded, irreversible changes in leakage current and in current-gain characteristics of the transistor result.

## POWER-DISSIPATION RATINGS

A transistor is heated by the electrical power dissipated in it. A maximum power rating is given, therefore, to assure that the temperature in all parts of a transistor is maintained below a value that will result in detrimental changes in the device. This rating may be given with respect to case temperature (for transistors mounted on heat sinks) or with respect to "free-air ambient" temperature. Case temperature is measured with a small thermocouple or other low-heat-conducting thermometer attached to the outside of the case or preferably inserted in a very small blind hole in the base so that the measurement is taken as close to the transistor chip as possible. Very short pulses of power do not heat the transistor to the temperature which it would attain if the power level was continued indefinitely. Ratings of maximum power consider this factor and allow higher power dissipation for very short pulses.

The dissipation in a transistor is not uniformly distributed across the semiconductor wafer. At higher voltages, the current concentrations become more severe, and hot spots may be developed within the transistor pellet. As a result, the power-handling capability of a transistor is reduced at high voltages. The power rating of a transistor may be presented most easily by a limiting curve that indicates a peak-power safe operating region. This curve shows power-handling capability as a function of voltage for various time durations.

The factors that determine the boundaries defined by the safe-area curve and the use of this curve are illustrated by the examples given in the following paragraphs. In these examples, the dissipation capability is determined for a 2N3055 silicon power transistor operating at an ambient temperature of 50°C for steady-state conditions and for both repetitive and nonrepetitive transient conditions. It is assumed that the transistor is attached to a 3-inch-by-4.25-inch, vertical-finned, extruded aluminum heat sink having a natural finish. This heat sink has a thermal resistance $\theta_{S\text{-}A}$ of 2.5°C per watt. The transistor is electrically insulated from the heat sink by a 0.002-inch-thick mica washer, which is coated with a zinc-oxide-filled silicone grease. The effective thermal resistance $\theta_{C\text{-}S}$ of the washer and silicone grease is 0.5°C per watt.

## Steady-State Operation

The maximum dissipation capability of a transistor under steady-state conditions depends on the sum of the series thermal resistances from the transistor junction to ambient air, the maximum junction temperature $T_J(max)$, and the ambient temperature $T_{amb}$ at which the transistor is operated. The sum of the series thermal resistances can be determined from the following relationship:

$$\Theta_{J-A} = \Theta_{J-C} + \Theta_{C-S} + \Theta_{S-A} \quad (67)$$

The maximum value of the junction-to-case thermal resistance $\theta_{J-C}$ for the 2N3055 transistor, as given in the manufacturer's specifications, is 1.5°C per watt. For the thermal system specified, the sum of the series thermal resistance can then be determined, as follows:

$$\Theta_{J-A} = 1.5 + 0.5 + 2.5$$
$$= 4.5 \text{ °C per watt}$$

The maximum junction temperature of the 2N3055, as specified in the manufacturer's rating, is 200°C. For operation at an ambient temperature of 50°C, the maximum dissipation capability of the 2N3055 transistor, under steady-state conditions, is calculated as follows:

$$P_{ss}(max) = \frac{T_J(max) - T_{amb}}{\Theta_{J-A}} \quad (68)$$
$$= (200-50)/4.5$$
$$= 33.3 \text{ watts}$$

The case temperature of the 2N3055 transistor that results from the maximum steady-state dissipation is calculated from the following equation:

$$T_C = P_{ss}(\theta_{C-A}) + T_{amb} \quad (69)$$

The term $\theta_{C-A}$ represents the total thermal resistance from case to ambient air (i.e., $\theta_{C-A} = \theta_{C-S} + \theta_{S-A} = 3$°C per watt). The case temperature, then, can be calculated as follows:

$$T_C = 33.3\ (3) + 50 = 150°C$$

## Single-Pulse Operation

When a transistor is operated in response to a single, nonrepetitive, short-duration pulse of power, the maximum allowable power dissipation during this transient period is substantially greater than the steady-state dissipation capability of the transistor. In the following calculations, the dissipation capability of the 2N3055 transistor operated in the specified thermal system from a single 1-millisecond pulse of power is determined.

Before the maximum dissipation capability of the transistor can be determined, the transient thermal resistance of the transistor must be known. The transient thermal resistance is usually obtained from the transistor maximum-operating-area curve, shown in Fig. 127,

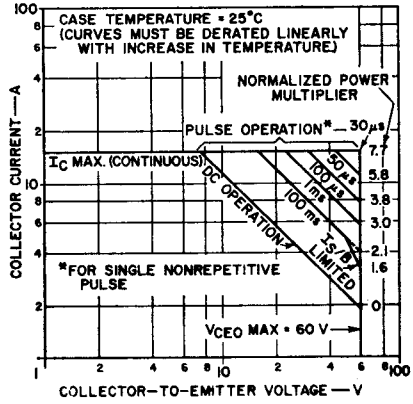



Figure 127. Maximum-operating-area curves for the 2N3055 silicon power transistor.

in the form of a normalized power multiplier M. For a 1-millisecond pulse, the normalized power multiplier for the 2N3055, given for a case temperature of 25°C, is 3. At higher case temperatures, the power multiplier M must be linearly derated so that it is reduced to zero at the maximum allowable junction temperature (200°C) for the transistor. The temperature derating factor TDF is determined as follows:

$$TDF = 1 - \frac{T_C - 25°C}{T_J\,(max) - 25°C} \quad (70)$$

For the system specified, the thermal capacitance of the heat sink is so large that the temperature of the heat sink does not change during the 1-millisecond duration of the pulse. The case temperature $T_C$, therefore, is essentially the same as the ambient temperature (50°C).

The data given above, together with the maximum steady-state dissipation rating of the transistor ($P_{max} = 115$ watts at a case temperature of 25°C), are used to determine the maximum allowable dissipation $P_{sp}$ for the 2N3055 transistor during the 1-millisecond period of the pulse, as follows:

$$\begin{aligned} P_{sp} &= M\,(TDF)\,(P_{max}) \quad (71)\\ &= M\left(1 - \frac{T_C - 25}{T_J\,(max) - 25}\right)(P_{max})\\ &= 3\left(1 - \frac{50 - 25}{200 - 25}\right)(115)\\ &= 296 \text{ watts} \end{aligned}$$

## Repetitive-Pulse Operation

When a transistor is operated in a repetitive pulse mode, the previous analysis must be modified to take into account the rise in case temperature caused by the average power dissipation. In the following example, it is assumed that the 2N3055 transistor operates in response to a train of 1-millisecond power pulses at a repetition rate of 100 Hz. The calculations are based on the same thermal system as that specified for steady-state and single-pulse calculations.

For repetitive-pulse operation, the average power dissipation $P_{avg}$ in the transistor is determined by the following relationship:

$$P_{avg} = P_{pk}(d) \quad (72)$$

where $P_{pk}$ is the peak pulse power dissipated in the transistor and d is the duty factor.

The effective case temperature that results from this average dissipation is determined from the following expression:

$$\begin{aligned} T_C(eff) &= T_{amb} + P_{avg}\,\theta_{J-A}\\ &= T_C + P_{avg}\,\theta_{J-C} \quad (73) \end{aligned}$$

Substitution of Eq. (72) into Eq. (73) yields the following result:

$$T_C(eff) = T_{amb} + P_{pk}(d)\,\theta_{J-A} \quad (74)$$

If the effective case temperature $T_C(eff)$, as defined by Eq. (74), is substituted for the case temperature $T_C$ in Eq. (71), the following expression is obtained for the maximum allowable power dissipation $P_{rp}$ for repetitive pulses:

$$P_{rp} = \frac{M\,[T_J(max) - T_{amb}]\,P_{max}}{T_J(max) - 25 + MdP_{max}\,\theta_{J-A}} \quad (75)$$

On the basis of the definition given in the section on **Thermal Factors**, junction-to-case thermal resistance may be expressed by the following equation:

$$\theta_{J-C} = \frac{T_J(max) - 25}{P_{max}} \quad (76)$$

If the relationship expressed by Eq. (76) is used, Eq. (75) can be simplified to the following form:

$$P_{rp} = M\left(\frac{T_J(max) - T_{amb}}{\theta_{J-C} + M\ d\ \theta_{J-A}}\right) \quad (77)$$

For the numerical example considered, the following values were previously assumed:

| | |
|---|---|
| d | = (1 ms/10 ms) × 100 |
| | = 10 per cent |
| $T_J(max)$ | = 200°C |
| $T_{amb}$ | = 50°C |
| $\theta_{J-C}$ | = 1.5°C/watt |
| M | = 3 |
| $\theta_{J-A}$ | = 4.5°C/watt |

When these values are substituted in Eq. (77), the maximum allowable power dissipation in the 2N3055 under repetitive pulse conditions is calculated as follows:

$$P_{rp} = \frac{3(200 - 50)}{1.5 + 3(0.1)4.5} = 158 \text{ watts}$$

Repetitive pulsing with a 10-percent duty factor reduces the peak-power capability in this case to about 50 per cent of the single-pulse peak-power capability.

When a transistor is to be subjected to irregularly shaped repetitive pulses, the following procedure may be used to obtain a conservative design:

1. The maximum allowable average power for the irregular pulse is calculated on the basis of the pulse width $T_p$, the period between the leading edges of successive pulses t, and the maximum steady-state dissipation $P_{ss}$ (max), as follows:

$$P_{avg} = P_{ss}(max)\ (T_p/t) \quad (78)$$

2. The ratio of peak power to average power ($N = P_{pk}/P_{avg}$) and the average case temperature ($T_C = P_{ss}\ \theta_{S-A} + T_{amb}$) are then used to determine the effective pulse width ($T_p' = T_p/N$).

3. The maximum power capability is then calculated from the following equation:

$$P_{tr} = \frac{T_J(max) - T_C(avg)}{\theta_{J-C}'} \quad (79)$$

where $\theta_{J-C}'$ is the effective junction-to-case thermal resistance as determined from the manufacturer's specifications for the effective pulse width $T_p'$.

## SECOND BREAKDOWN

Second breakdown (S/b) is a potentially destructive phenomenon that occurs in all bipolar (n-p-n and p-n-p) transistors. This phenomenon results when the energy absorbed by a transistor exceeds a critical level, and causes localized hot spots within the transistor pellet. The start of second breakdown is characterized by an abrupt decrease in collector-to-emitter voltage with a small dynamic resistance in the second-breakdown region, as shown in Fig. 128.

Although second breakdown may result from several modes of transistor operation, this phenomenon can be broadly categorized into two major classes: (1) forward-biased emitter-to-base second breakdown, which occurs when the transistor operates in the active region, and (2) reverse-biased emitter-to-base second breakdown, which occurs during the cutoff mode of transistor operation.

### Forward-Bias Second Breakdown

Fig. 129 shows a cross section of a typical silicon diffused-junction power-transistor pellet. When the transistor is heavily forward-biased into the active region, a

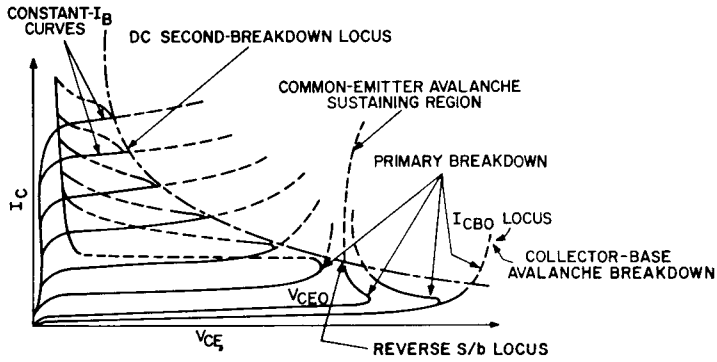



Figure 128. Transistor collector characteristics.

transverse electric field is produced in the base region, and a space-charge layer is formed at the base-to-collector junction. As current flows from emitter to collector, the transverse field focuses the current into a narrow region below the emitter edge. When the current flows through the space-charge layer, a significant amount of heat is generated. With the current focused into a small area, the heating effect is localized in this area, and hot spots (circled areas in Fig. 129) may be formed within the silicon pellet. If unchecked, these hot spots initiate a regenerative cycle of high-density current which results in forward-biased second breakdown in the transistor.

The carrier concentration in the pellet and, consequently, the severity of the hot spots are determined mainly by the magnitude of the transverse base field and by the applied collector voltage, which determines the intensity of the electric field across the space-charge layer formed at the base-to-collector junction. The magnitude of the transverse base field depends on the width of the transistor base, the base resistance, and the spreading of the space-charge layer that results from the application of collector voltage. The transverse base field increases with increases in base current that result from higher injection levels and reduced current gain at the higher injection levels.

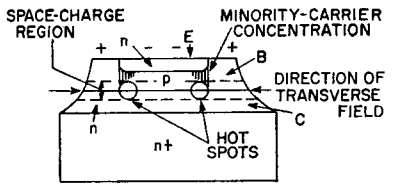



Figure 129. Cross-section of power transistor under forward-bias conditions.

The severity of the hot spots is inversely proportional to the width of the transistor base and directly proportional to the magnitude of the applied collector voltage. As a result, the current level $I_S/_b$ at which second breakdown occurs decreases rapidly with an increase in applied collector voltage. The $I_S/_b$ capability is also decreased as the transistor frequency capability is increased (i.e., as the width of the base is decreased). Figs. 130 and 131 show variation in $I_S/_b$ capability as a function of frequency capability $f_T$ and collector voltage

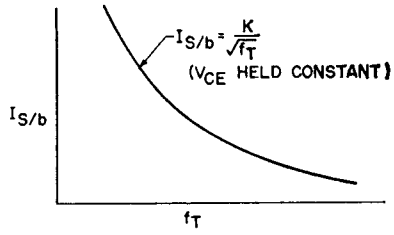



Figure 130. Variation in forward-bias second-breakdown energy level as a function of transistor frequency capability.

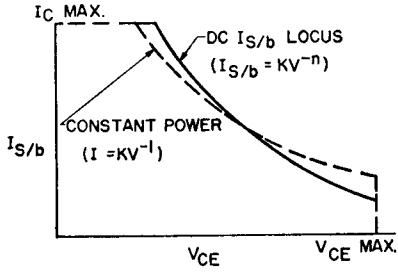



Figure 131. Variation in forward-bias second-breakdown energy level as a function of collector voltage.

$V_{CE}$, respectively. The curves shown in these figures are graphical representations of the following empirical relationships:

1. Frequency relationship:

$$I_{S/b} = K_1/\sqrt{f_T} \qquad (80)$$

2. Voltage relationship:

$$I_{S/b} = K_2/V_{CE}^{n} \qquad (81)$$

where $K_1$ and $K_2$ are constants determined by the device being considered and n is a constant that ranges from 1.5 to 4 depending upon the construction of the transistor (i.e., graded or abrupt junctions) and other factors.

Eq. (80) suggests that the circuit designer should select the transistor that has the lowest frequency capability, consistent with circuit requirements, to achieve the maximum resistance to second breakdown. Eq. (81) indicates that supply voltage should be as small as possible and that high-voltage transients should be restricted as much as possible to achieve second-breakdown protection.

The thermal capacitance of the transistor pellet results in a localized thermal time constant that restricts instantaneous formation of hot spots. As a result, the $I_{S/b}$ capability of a transistor is increased when the time duration of applied current and voltage is very short. Fig. 132 shows the variation in $I_{S/b}$ capability with pulse width.

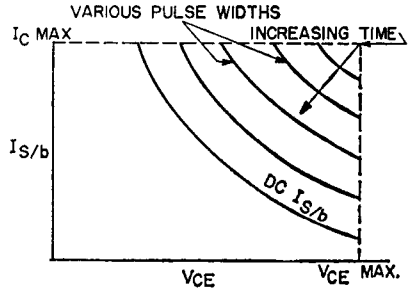



Figure 132. Variation in forward-bias second-breakdown energy level as a function of pulse width.

## Reverse-Bias Second Breakdown

When current flows through a power transistor in which the emitter-to-base junction is reverse-biased, the direction of the transverse base field is opposite from that produced in a forward-biased transistor. As a result, the emitter current is focused into a small region at or near the center of the emitter. Because the current is crowded into a smaller region under reverse-bias conditions, reverse-bias second breakdown can be encountered at substantially

lower energy levels than those at which forward-bias second breakdown occurs. Fig. 133 shows a cross section of a typical power-transistor pellet under reverse-bias conditions.

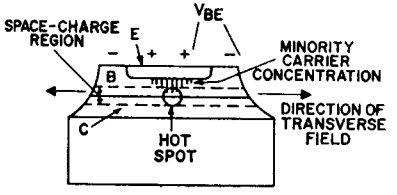



Figure 133. Cross-section of power transistor under reverse-bias conditions.

The resistance of a transistor to reverse-bias second breakdown is reduced by any design alteration that increases current density or prevents spreading of emitter current. In power transistors that have a narrow base, an accelerating base field, or insufficient emitter size for their operating current, second breakdown generally occurs at lower energy levels than in transistors without these factors.

Reverse-bias second breakdown is usually described in terms of energy because voltage, current, and pulse duration are interdependent when the emitter-to-base junction is reverse-biased. The transverse base field in the transistor depends, to a large extent, on the turn-off base current and turn-off voltage. It is natural, therefore, to assume that the energy level at which second breakdown takes place is a strong function of the input-circuit turn-off voltage $V_{BE}$ and the series resistance $R_{BE}$. Fig. 134 shows the variation in second-breakdown energy level as a function of $V_{BE}$ and $R_{BE}$ for a typical power transistor. As shown in this figure, if $R_{BE}$ is increased and $V_{BE}$ is decreased, the turn-off base current and the transverse base field are both reduced, and the second-breakdown energy level is raised. Because turn-off time is an inverse function of base currents, the circuit designer must compromise between transistor turn-off speed and second-breakdown considerations.

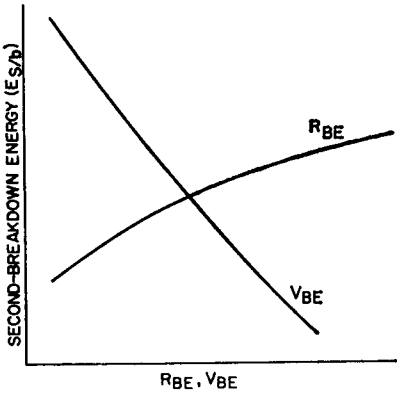



Figure 134. Variation in reverse-bias second-breakdown energy level as a function of $V_{BE}$ and $R_{BE}$.

## Second-Breakdown Evaluation Techniques

The ability of a power transistor to withstand second breakdown under either forward or reverse-bias conditions can be verified easily by testing the devices to destruction. The establishment of meaningful second-breakdown limits on power transistors, however, requires the use of nondestructive verification tests. Such nondestructive tests are described briefly in the following paragraphs.

**Forward-Bias $I_S/_b$ Test**—Fig. 135 shows the block diagram of a typical nondestructive $I_S/_b$ test set. In this test set, the transistor under test is in series with a pass transistor. The transistor under

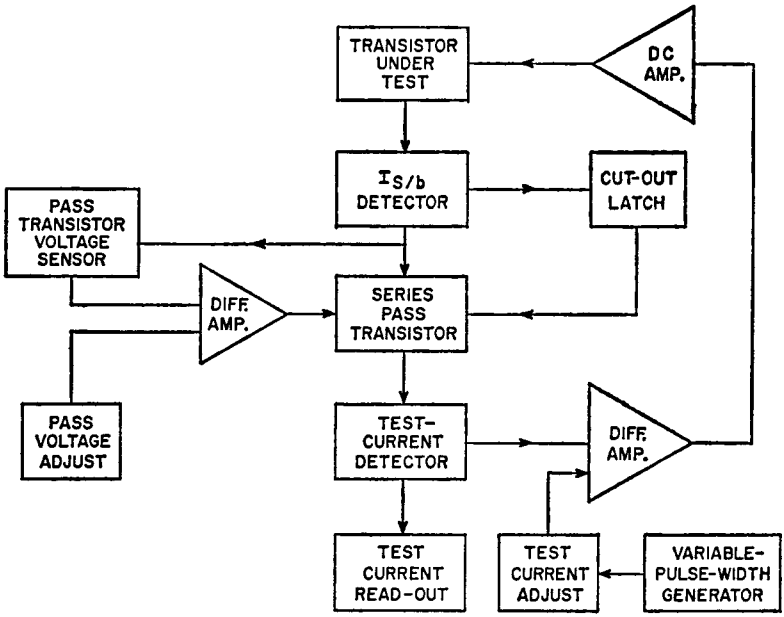



Figure 135. Block diagram of typical $I_s/_h$ test set.

test is driven by a differential amplifier to provide a preselected value of test current at a level independent of transistor current gain. The pass transistor is operated below saturation so that fast turn-off is possible. A second differential amplifier senses the voltage across the pass transistor and a 1-ohm resistor in series with this transistor. This voltage is maintained at a constant value throughout the test to improve the accuracy of the $I_S/_b$ test voltage. The circuit is arranged so that only the collector current of the transistor under test passes through the 1-ohm resistor. The voltage across this resistor, therefore, is an accurate indication of collector current.

The onset of second breakdown is detected by use of the primary winding of a pulse transformer which is placed in series with the collector of the transistor under test. During second breakdown, the rapid rate of rise of collector current induces a voltage $L(d_i/d_t)$ in the transformer which is coupled to the input circuit of the series pass transistor. This voltage turns off the series pass transistor in one microsecond. The primary inductance of the transformer also limits the immediate current rise. The voltage developed across this primary inductance is of a polarity that immediately reduces the voltage across the transistor under test. These characteristics, together with the protective cut-out circuit, prevent transistor degradation during the $I_S/_b$ tests. The complete cut-out time of the actual test set is approximately one microsecond, which is sufficient to prevent degradation of

even the highest-speed power transistor available in the industry. The pulse width of voltage and current applied to the transistor under test can be varied from 0.5 millisecond to several seconds. For dc $I_S/_b$ tests, a pulse width of 1 to 2 seconds is required because the thermal time constant of the power transistor pellet and mounting block may be several tenths of a second.

**Forward-Bias Capacitance-Discharge Test**—Fig. 136 shows the schematic diagram of a typical test circuit used to determine forward-bias second-breakdown energy level. This test circuit operates on the principle of a charged capacitor that discharges its energy ($E_C = CV_C^2/2$) into a power transistor at a constant current rate. As long as the initial voltage across the capacitor is less than the $V_{CER}(sus)$ rating of the power transistor, the collector current of the transistor is approximately constant; as a result, collector voltage decays linearly. The capacitor is charged by constant current $V_{CC}/R$ to a value $V_C$, at which time switch $S_1$ is closed. A constant current $I_C$, as determined by the combination of $V_{BB}$, $R_B$, and $R_E$, then flows through the transistor until the capacitor is completely discharged.

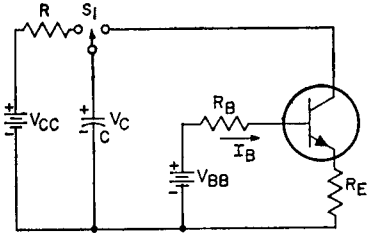



Figure 136. Capacitive-discharge test circuit used to determine forward-bias second-breakdown energy levels in transistors.

If the drop across $R_E$ is neglected, the energy absorbed by the transistor is given by the following equation:

$$E_C = C V_C^2/2 \qquad (82)$$

Current and voltage waveforms of the 2N3442 silicon power transistor tested in the capacitance discharge test set are shown in Fig. 137.

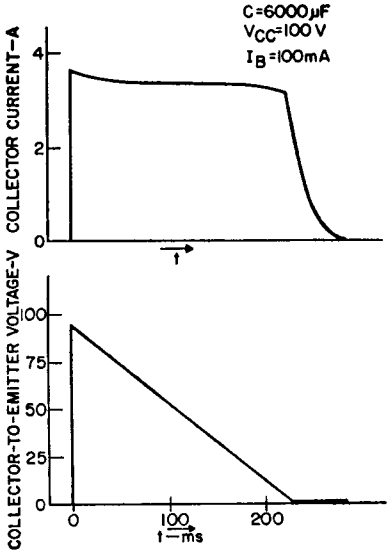



Figure 137. Test waveforms for a 2N3442 transistor subjected to capacitive-discharge second-breakdown energy test.

The capacitance discharge test is useful because it closely approximates actual circuit conditions, such as series-regulator load shorting, inverter and audio-amplifier initial turn-on, and other effects in capacitive-loaded circuits.

**Reverse-Bias $E_S/_b$ Test**—One practical way to measure reverse-bias second-breakdown energy levels in a power transistor is by use of a series inductive output circuit. If current is passed through the inductance while the

transistor under test is operated in the saturation region and a reverse bias is then abruptly applied to the transistor, a voltage is induced across the series inductance. This induced voltage rises to the transistor reverse sustaining breakdown voltage $V_{CEX}(sus)$ provided that the following limitation is imposed on transistor turn-off time:

$$t_{off} \ll L\, I_C(max)/V_{CEX}\,(sus) \quad (83)$$

The energy that the transistor absorbs during the $E_S/_b$ test can be determined from the following equation:

$$E_{S/b} = \frac{L\, I_C(max)^2}{2} \left[\frac{V_{CEX}\,(sus)}{V_{CEX}(sus) - V_{CC}}\right] \quad (84)$$

Voltage and current waveforms for a 2N3442 silicon power transistor subjected to the $E_S/_b$ test are shown in Fig. 138.

Fig. 139 shows a block diagram of a test circuit used for nondestructive $E_S/_b$ evaluations. The main feature of this test circuit is the unique second-breakdown detection technique. This circuit provides the rapid second-breakdown detection necessary to prevent damage to the transistor by an immediate sensing of a large-amplitude rf noise voltage developed at the base of the transistor under test at the onset of second breakdown.

Basically, the test circuit turns on the transistor under test and causes it to operate into a variable series inductance from a constant-current supply. When the input drive is removed, a reverse bias, provided by $V_{BE}$ and $R_{BE}$, is applied to the base of the transistor under test. The energy stored in the series inductance and the constant-current supply is then absorbed by the transistor. If this energy is sufficient to drive the transistor into reverse-

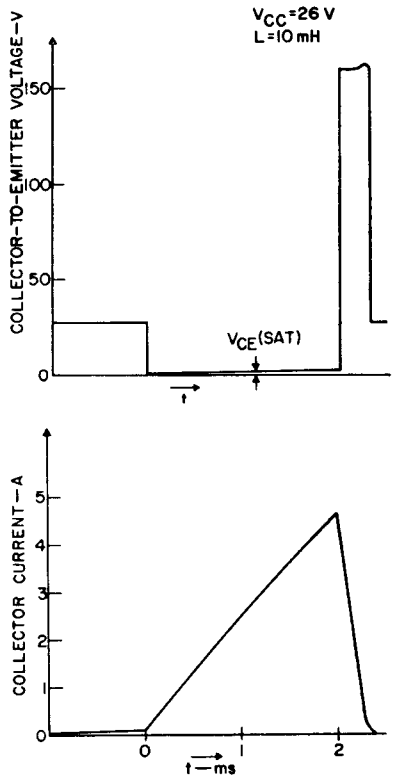



Figure 138. Test waveforms for a 2N3442 transistor subjected to $E_S/_b$ inductive energy test.

bias second breakdown, the detector triggers a protective clamp circuit in shunt with the transistor under test. The reaction time of the test circuit is in the order of only a few hundred nanoseconds.

**Reverse-Bias Capacitance-Discharge Test**—The reverse-bias capacitance-discharge test is similar to the forward-bias capacitance-discharge test described previously, except that the capacitor

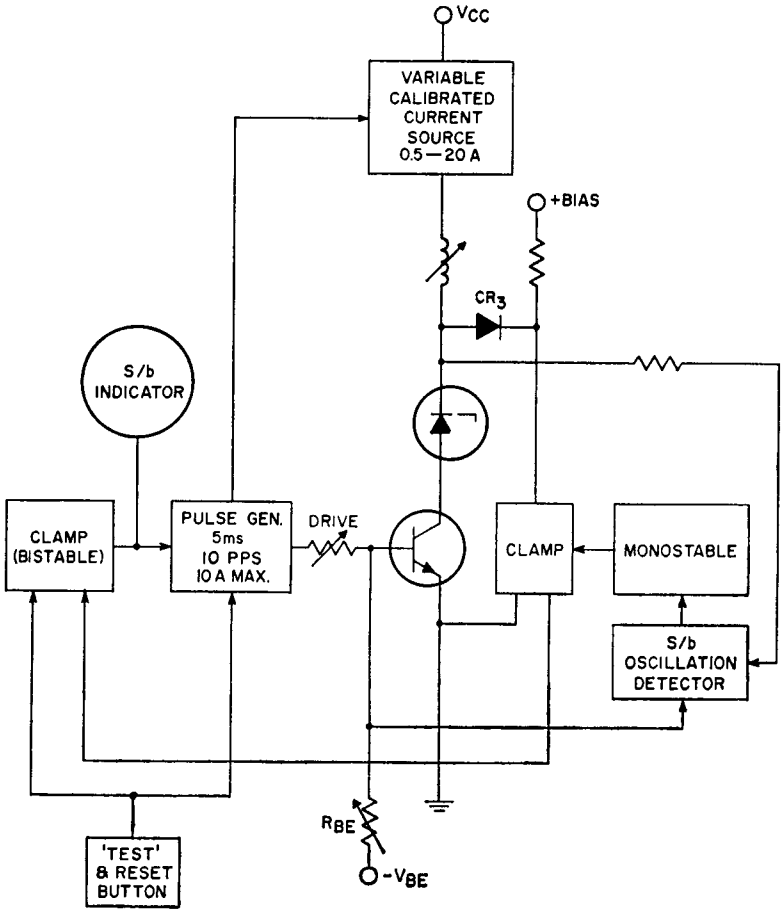



Figure 139. Block diagram of $E_{s/b}$ test set.

is charged to a value greater than the reverse sustaining breakdown voltage of the transistor under test. The base of the transistor is reverse-biased, as shown in Fig. 140.

The value of the series resistance is adjusted so that the current through the transistor is limited to a value below the maximum collector-current rating. The energy absorbed by the transistor in this reverse-biased condition when switch $S_1$ is moved from position 1 to position 2 is given by the following equations:

$$E_{C_{S/b}} = \int_0^\infty (V_{CEX}/R)\,(V_{CC} - V_{CEX})\,\epsilon^{-t/RC}dt \qquad (85)$$

$$= C\,[V_{CEX}(sus)\,V_{CC} - V_{CEX}(sus)^2] \qquad (86)$$

These equations assume that the $V_{CEX}(sus)$ breakdown value of the transistor does not change with current. Variations in $V_{CEX}(sus)$

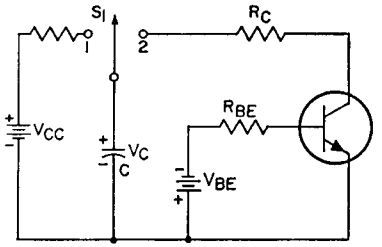



Figure 140. Capacitive-discharge test circuit used to determine reverse-bias second-breakdown energy levels in transistors.

with collector current usually do not exceed 10 per cent in most power transistors.

## SAFE-AREA RATINGS

Safe-area ratings are given for power transistors so that the circuit designer can select the proper type for his application and can determine the best trade-offs between desired circuit performance and the actual capabilities of the device. These ratings must include forward-bias second-breakdown ratings for both dc and pulsed operation, reverse-bias second-breakdown ratings for both inductive and capacitive loads, and thermal ratings for both steady-state and transient conditions. Thermal ratings and forward-bias second-breakdown ratings can be readily combined into a single rating system. A separate rating system is necessary, however, for reverse-bias conditions.

### Forward-Bias Safe-Area Ratings

Forward-bias safe-area ratings are displayed on a voltage-current chart which shows rating curves for dc operation and for pulsed operation of various time durations. Eq. (82) in the section on **Second Breakdown** defines the forward-bias second-breakdown energy level in a power transistor. On a log-log graph of the voltage-current curves, the locus of this equation results in a linear derating curve that has a slope equal to the junction constant n. If the safe operating area of a power transistor is limited within any portion of the voltage-current characteristics by thermal factors (thermal impedance, maximum junction temperature, or operating case temperature), this limiting is defined by a constant-power hyperbola ($I = KV^{-1}$) which can be represented on the log-log voltage-current curve by a straight line that has a slope of $-1$.

As pointed out in the section on **Second Breakdown,** the energy level at which second breakdown occurs in a power transistor increases as the time duration of the applied voltage and current decreases. The power-handling capability of the transistor also increases with a decrease in pulse duration because the thermal mass of the power-transistor chip and associated mounting parts imparts an inherent thermal delay to a rise in junction temperature. Fig. 141 shows a curve of normalized thermal resistance $N_R$ as a function of the time duration of applied power for the RCA-2N3442 silicon power transistor. The two

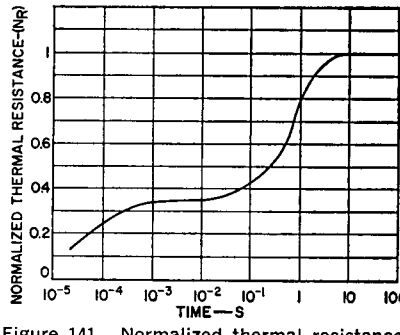



Figure 141. Normalized thermal resistance as a function of the duration of applied power for the 2N3442 silicon power transistor.

horizontal regions of the curve show that the 2N3442, as is typical of most power transistors, has two major thermal time constants. The shorter time constant results from the thermal resistance and capacitance of the silicon chip and its interface with the transistor case. The longer time constant is caused by the thermal impedances of the case header and the internal case parts, such as copper, molybdenum blocks, and beryllia.

For a given case temperature $T_C$, maximum junction temperature $T_J(\max)$, and junction-to-case thermal-resistance rating $\theta_{J-C}$, together with the value determined for the normalized thermal impedance $N_R$, the following equation can be used to calculate maximum power dissipation as a function of the duration of an applied pulse of power:

$$P_{diss} = \frac{T_J(\max) - T_C}{\theta_{J-C}(N_R)} \qquad (87)$$

The absence of a $V_{CE}$ term from the equation for power dissipation under pulsed conditions indicates that this equation also defines a constant-power curve which can be represented on a log-log voltage-current curve by a straight line that has a slope of −1.

Fig. 142 shows a forward-bias safe-area rating chart for a typical high-speed silicon power transistor, the RCA-2N3585, which has a gain-bandwidth product $f_T$ in the order of 20 MHz. Fig. 143 shows a similar rating chart for a lower-frequency silicon power transistor, the RCA-2N3442, for which the $f_T$ is typically 1 MHz. The boundaries defined by the curves in the safe-area charts indicate, for both continuous-wave and nonrepetitive-pulse operation, the maximum current ratings, the maximum collector-to-emitter forward-bias avalanche breakdown-voltage rating [$V^{\alpha}{}_M = 1$, which is usually approximated by $V_{CEO}(\text{sus})$], and the thermal and second-breakdown ratings of the transistors.

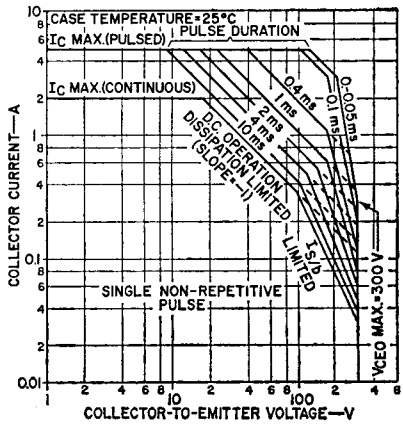



Figure 142. Safe-area rating chart for the 2N3585 silicon power transistor.

As shown in Fig. 142, the thermal (dissipation) limiting of the 2N3585 ceases when the collector-to-emitter voltage rises above 100 volts during dc operation. Beyond this point, the safe

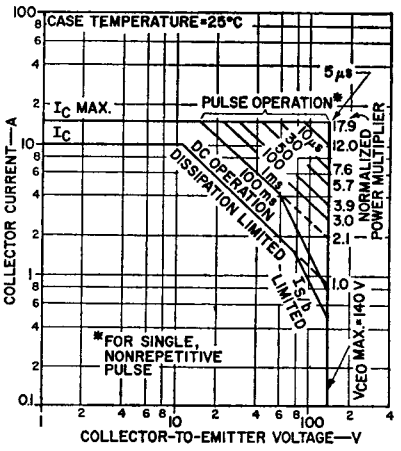



Figure 143. Safe-area rating chart for the 2N3442 silicon power transistor.

operating area of the transistor is limited by the second-breakdown ratings. During pulsed operation, the thermal limiting extends to higher values of collector-to-emitter voltage before the second-breakdown region is reached, and as the pulse duration decreases, the thermal-limited region increases. A comparison of Figs. 142 and 143 shows that the second-breakdown ratings of the 2N3442 are substantially higher than those of the 2N3585, as can be predicted from Eq. (80) on the basis of the lower $f_T$ for the 2N3442.

If a transistor is to be operated at a pulse duration that differs from those shown on the safe-area chart, the boundaries provided by the safe-area curve for the next higher pulse duration must be used, or the transistor manufacturer should be consulted. Moreover, as indicated in Figs. 142 and 143, safe-area ratings are normally given for single non-repetitive pulse operation at a case temperature of 25°C and must be derated for operation at higher case temperatures and under repetitive-pulse or continuous-wave conditions.

Fig. 144 shows temperature derating for the 2N3585 safe-area

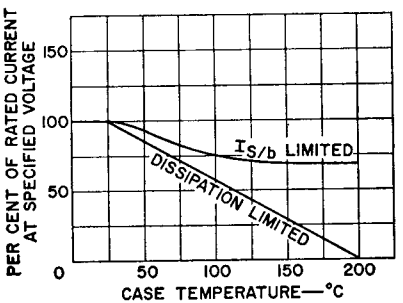



Figure 144. Safe-area temperature-derating curves for the 2N3585 silicon power transistor.

chart of Fig. 142. These curves show that thermal ratings are affected far more by increases in case temperature than are second-breakdown ratings. The thermal (dissipation-limited) derating curve is a graphic representation of Eq. (70) for the temperature derating factor given in the section on **Power-Dissipation Ratings.** This curve, as expected, decreases linearly to zero at the maximum junction temperature of the transistor [$T_J$(max) = 200°C]. The second-breakdown ($I_{S/b}$-limited) temperature derating curve, however, is less severe because formation of the high current concentrations that cause second breakdown is less likely as the temperature increases.

Because the thermal and second-breakdown deratings are different, it may be necessary to use both curves to determine the proper derating factor for a voltage-current point that occurs near the breakpoint of the thermal-limited and second-breakdown-limited regions on the safe-area curve. For this condition, a derating factor is read from each derating curve. For one of the readings, however, either the thermal-limited section of the safe-area curve must be extrapolated upward in voltage or the second-breakdown-limited section must be extrapolated downward in voltage, depending upon which side of the voltage breakpoint the voltage-current point is located. The smaller of the collector-current values obtained from the thermal and second-breakdown deratings must be used as the safe rating.

The procedure used to derate a voltage-current point under repetitive-pulse or continuous-wave operation was described

previously in the section on **Power-Dissipation Ratings**. Basically, this derating requires the use of an artificially calculated case temperature ($T_{C(eff)} = T_C + \theta_{J-C} P_{avg}$) with the single-pulse safe-area ratings and the temperature derating curves. This calculated case temperature accounts for the rise in the operating case temperature that results from the transistor thermal resistance $\theta_{J-C}$ and the average power of the periodic waveform. The value obtained for $T_{C(eff)}$ is used as the case-temperature value on the temperature derating chart to obtain the repetitive-pulse derating factor for the safe-area curves.

The preceding discussion covers nonrepetitive and repetitive rectangular-pulse operation only. The following steps must be used to resolve all other voltage-current waveforms into equivalent rectangular pulses before the derating procedure described can be used:

1. Plot the actual voltage-current load line on the appropriate transistor safe-area chart.

2. Select the voltage-current point on the load line that makes the greatest excursion into the safe-area region.

3. Estimate the total energy content of the actual voltage-current waveform. This value can be most easily estimated by graphical integration of the waveforms.

4. Determine an effective pulse duration $t_p$(eff) by dividing the total energy in the waveforms by the voltage-current product at the point selected in Step 2.

The voltage, current, and effective pulse duration computed above define a rectangular pulse equivalent to the actual waveforms.

## Safe-Area Design Analysis

Two examples of the use of the forward-bias safe-area system just developed are given below. The first example requires the analysis of a typical power inverter under initial turn-on conditions; the second example applies to the analysis of a direct-coupled audio power amplifier operating at low frequency into an "inductive speaker" load.

**Example 1: Inverter Initial Turn-On Analysis**—Fig. 145 shows a typical high-speed, high-

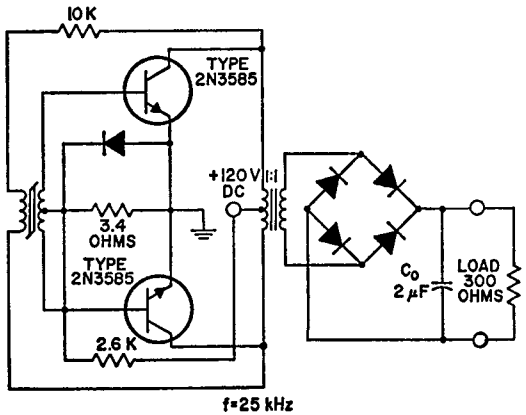



Figure 145. High-speed inverter using RCA 2N3585 transistors.

voltage, 100-watt, two-transformer inverter. Fig. 146 shows the collector voltage, collector current, and peak power of the inverter as functions of time for the initial turn-on condition when the output capacitor $C_o$ is uncharged. It is assumed that a circuit designer wants to determine the ability of the 2N3585 to operate safely in the circuit and, if safe operation is shown to be feasible, the maximum permissible case temperature.

The analysis begins with the plotting of the resistive load line on the 2N3585 safe-area curve, as shown in Fig. 147. Because the circuit being analyzed is a switching circuit, each pulse has a different load line associated

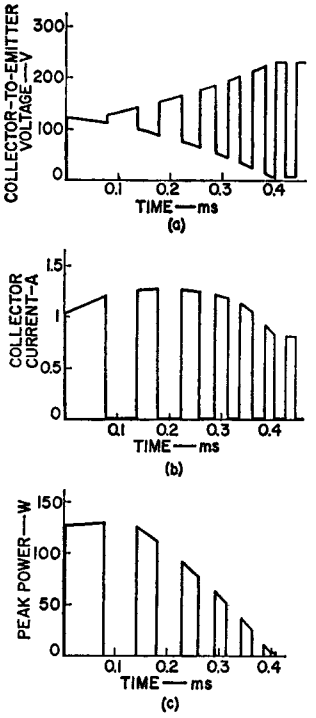



Figure 146. Waveforms for inverter circuit shown in Fig. 145.

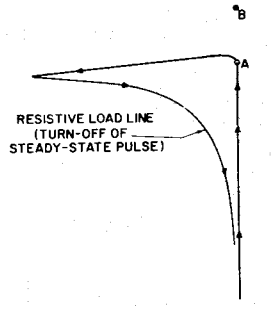



Figure 147. Inverter turn-on-load line on 2N3585 safe-area curve.

with it. Initially, the pulses follow a highly capacitive load line, but become more resistive as the output capacitor charges. This change of pulse character presents no problem in this analysis because it is being performed for initial turn-on only. For this reason, the designer need only plot the turn-on load line of the first pulse and the locus of peak voltage and current of the remaining pulses, as shown in Fig. 147. Point A ($V_{CE}$ = 120 volts and $I_C$ = 1.2 amperes) is the point of greatest excursion of the locus of peak voltage and current into the safe-area region.

The total energy required during turn-on is determined by graphical integration, shown in Fig. 146(c). The result of the integration indicates that the energy required during the 0.4-millisecond turn-on is approximately 20 millijoules. The peak power at point A (the product of the voltage and current coordinates at that point) is 140 watts, and the effective pulse duration $t_o$(eff) is 20 millijoules divided by 140 watts, or 0.14 millisecond. Because it is assumed that the initial circuit turn-on is a nonrepetitive-pulse operation in

this example, the 120-volt, 1.2-ampere, 0.14-millisecond, nonrepetitive equivalent rectangular pulse can be applied directly to the safe-area curve of Fig. 147. The result is the definition of point B ($V_{CE}$ = 120 volts, $I_C$ = 3.4 amperes), the equivalent safe-area nonrepetitive peak pulse for t = 0.14 millisecond. The position of point B indicates that the transistor will operate safely in the inverter at a case temperature of 25°C.

For determination of the maximum case temperature at which the 2N3585 will continue to perform satisfactorily, the temperature-derating factor must be calculated. This factor, the ratio of collector current at point A to collector current at point B, is 1.2/3.4 = 0.35, or 35 per cent. Because point A is in the dissipation portion of the safe-area rating curve, the dissipation-limited curve in Fig. 146 is used to find the maximum case temperature. For a 35-per-cent derating factor, this temperature is found to be 130°C. Thus, it is possible to turn on this inverter safely at case temperatures up to 130°C.

If the circuit is to be keyed on and off at some set repetition rate, a repetitive analysis which takes into account the effective case temperature $T_C$(eff) must be performed and a new and lower value of maximum case temperature must be determined.

**Example 2: Analysis of a Direct-Coupled Audio Power Amplifier at Low Frequency**— A quasi-complementary 70-watt power amplifier is shown in Fig. 148. This amplifier is terminated in a resistance-inductance series

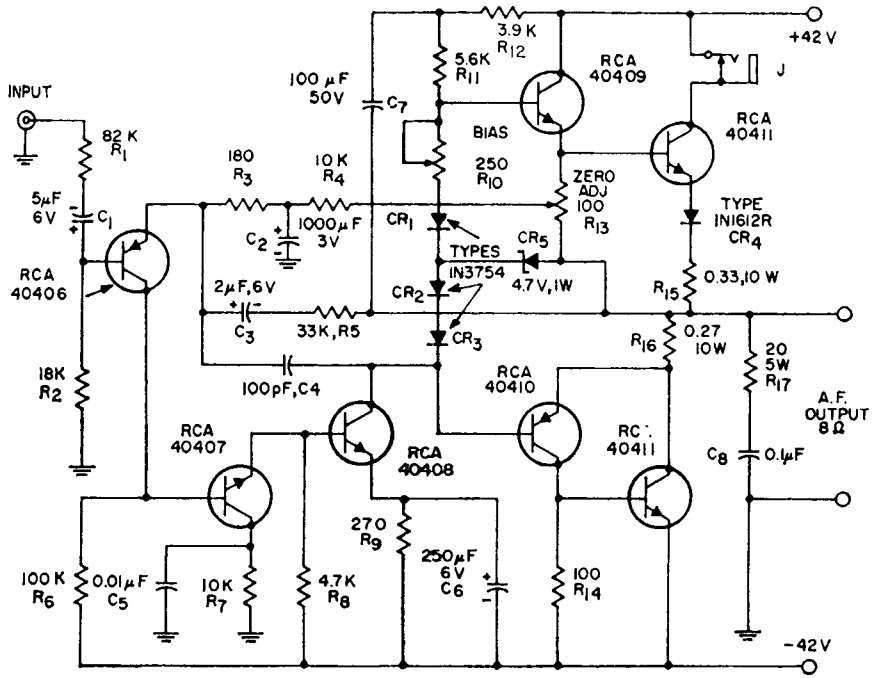



Figure 148. 70-watt, silicon-transistor audio amplifier.

load circuit in which the inductance is 40 millihenries and resistance is 1 ohm, to simulate a worst-case speaker impedance. The amplifier is driven at 20 Hz. Voltage, current, and power waveforms as functions of time are shown in Fig. 149.

It is assumed that the circuit designer wants to determine the ability of the RCA-40411 output transistor to operate safely in the circuit with the frequency and load specified and with a maximum case temperature of 70°C. The circuit designer may also want to analyze the circuit at other load and frequency conditions. The safe-area curve for the 40411 (a 15-ampere silicon power transistor) and the load line for a single cycle are given in Fig. 150. Point A ($V_{CE}$ = 55 volts, $I_C$ = 4.1 amperes, P = 225 watts) represents the point of maximum excursion of the load line into the safe-area region. Graphical integration of Fig. 149(c) yields an equivalent energy of 2.1 joules. When the equivalent energy is divided by the power at point A, the effective pulse duration is found to be 9.3 milliseconds. Thus, a rectangular pulse of 55 volts, 4.1 amperes and 9.3 milliseconds duration is equivalent to the actual circuit waveforms. Point B ($V_{CE}$ = 55 volts, $I_C$ = 8.2 amperes) is the safe-area value for this single, nonrepetitive equivalent pulse at a case temperature of 25°C. Derating to obtain the safe-area point for higher case temperature as well as for repetitive-pulse conditions must then be performed. Because there is a continuous 20-Hz sine-wave input with a period of 50 milliseconds to the amplifier, the duty cycle of this equivalent pulse is 9.3 milliseconds divided by 50 milliseconds, or 18.6 per cent. The average power calculated by dividing the energy per pulse (2.1 joules) by the total period (50 milliseconds) is 42 watts. When these values are substituted in Eq. (73), the effective case temperature is computed as follows:

$$T_c(\text{eff}) = 70°C + 42\ W \quad (1.1)$$
$$= 70°C + 46°C = 116°C$$

where 1.1°C per watt is the junction-to-case thermal resistance ($\theta_{J\text{-}C}$) for the 40411.

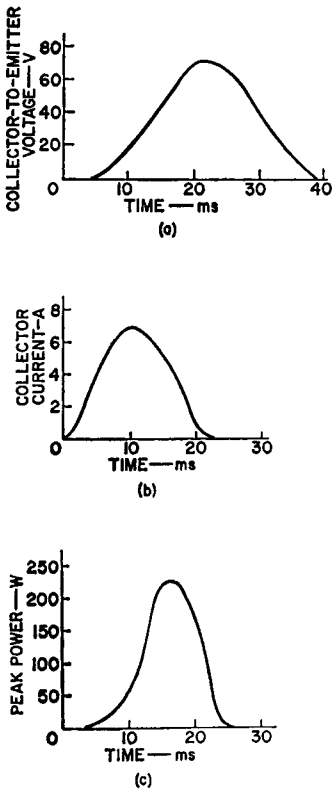



Figure 149. Audio power-amplifier waveforms.

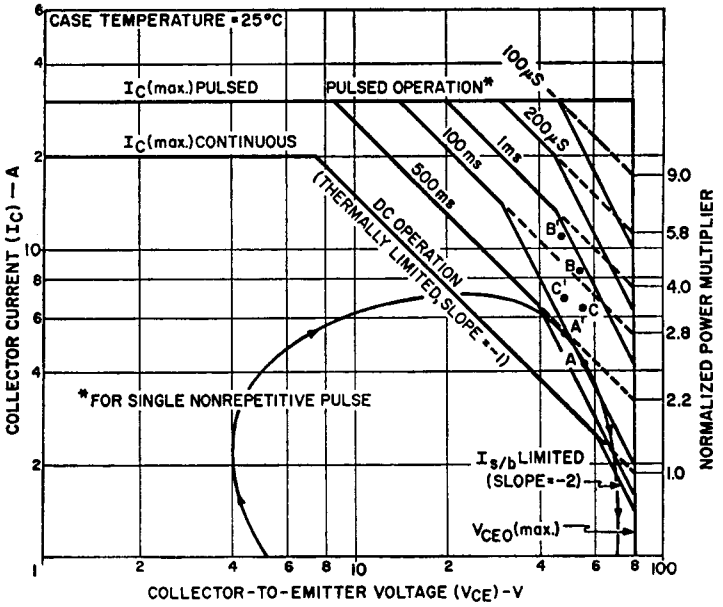



Figure 150. Safe-area rating chart for an RCA 40411 transistor.

The temperature derating curves for the 40411 are given in Fig. 151. Because point B falls within the second-breakdown-limited section of the safe-area curve, the derating factor is read from the $I_{S/b}$-limited curve of Fig. 151. For a $T_C$(eff) of 116°C, the derating factor is 75 per cent. When the current is derated at point B, point C ($V_{CE}$ = 55 volts, $I_C$ = 6.2 amperes) is obtained. If second-breakdown limitation were the only consideration in determining the ability of the 40411 to operate satisfactorily in the circuit, the location of point C would indicate that under the specified conditions the transistor would perform as desired. However, because the load line in Fig. 150 also comes close to the dissipation-limiting curve at point A' ($V_{CE}$ = 45 volts, $I_C$ = 5.4 amperes), it is necessary to consider transistor dissipation limitations. Point B' ($V_{CE}$ = 45 volts, $I_C$ = 12.5 amperes) is the single-pulse 25°C case-temperature equivalent. If the value of 116°C is used for effective case temperature, as determined earlier, a temperature-derating factor of 50 per cent is read from the dissipation-limited curve of Fig. 151. This factor yields a point C' ($V_{CE}$ = 45 volts, $I_C$ = 6.3 amperes) which is above the point A' that represents expected circuit operating limits. This result indicates that dissipation limitations will not adversely affect transistor performance in the circuit. The

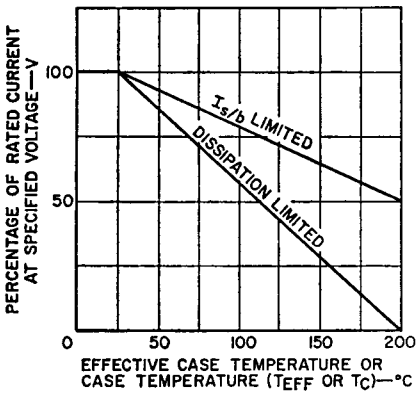



Figure 151. Temperature derating for the RCA 40411 transistor.

greater distance between points A and C than between points A′ and C′ in Fig. 150 indicates that there is a greater margin of safety in the second-breakdown region than in the dissipation region.

## Reverse-Bias Safe-Area Ratings

Power transistors are required to absorb energy under reverse-bias conditions in a wide variety of switching circuits including solenoid drivers, power inverters, switching regulators, magnetic deflection circuits, transformer-coupled power amplifiers, and motor and lighting controls. A characteristic of these circuits is the presence of series inductance such as transformer leakage inductance in inverters and power amplifiers, solenoid inductance, motor armature and field inductance, and regulator low-pass filter inductance. The best means for determining the reverse-bias safe-operation rating for these circuits makes use of a series inductance L (without diode clamp), a turn-off circuit of series resistance $R_{BE}$, and a series voltage $V_{BE}$. If the transistor under test is driven into saturation with a collector current of $I_C$(peak) and the forward base drive is abruptly removed, the test transistor turns off through the turn-off circuit and absorbs an amount of energy equal to the second-breakdown energy $E_{S/b}$ given by Eq. (84).

As explained earlier, the second-breakdown energy is a function of $R_{BE}$, $V_{BE}$, and series inductance. Therefore, because it is possible to resolve all of the circuits mentioned above into a simple series-inductive switch with a turn-off series resistance of $R_{BE}$ and a base-to-emitter voltage of $V_{BE}$, it follows that the development of a set of curves defining a minimum energy rating as a function of L, $R_{BE}$, and $V_{BE}$ for this representative circuit will provide an adequate basis for determining the reverse-bias safe-operation rating of any of the more specialized circuits represented.

A set of curves used to define reverse-bias safe operation for the 2N3585 is given in Fig. 152. These curves, expressed in terms of peak current $I_{pk}$, can be readily converted to energy E through the use of the following relationship:

$$E = \tfrac{1}{2} L I_{pk}^2 \qquad (88)$$

The temperature-derating factor for the reverse-bias condition is determined in the same manner as that for the forward-bias second-breakdown condition. That is, the $I_{S/b}$-limited portion of the derating curve in Fig. 144 is used.

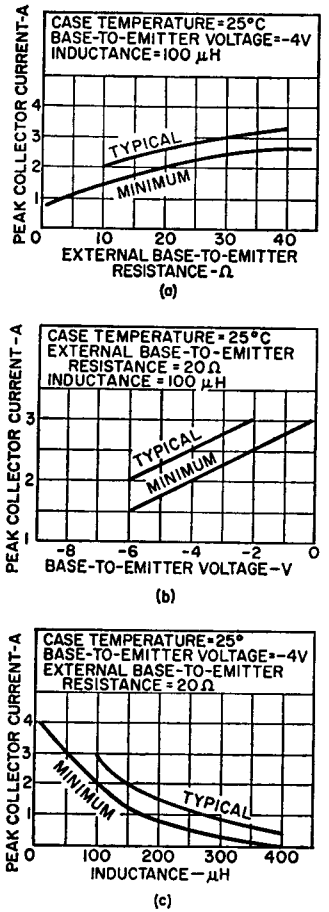



Figure 152. Reverse-bias energy of the RCA 2N3585 transistor.

The use of the reverse-bias second-breakdown rating curves of Fig. 152 is illustrated below by analysis of the inverter circuit shown in Fig. 145. The analysis assumes that the inverter is in the turn-off condition.

## Analysis of the Inverter in the Turn-Off Condition

The leakage inductance in the primary of the output transformer in the inverter shown in Fig. 145 was measured and found to be 5 microhenries. However, in order that the analysis represent the worst case, the maximum transformer leakage inductance was estimated at 100 microhenries. For use of the rating curves, an effective value of series inductance $L_{eff}$, an equivalent input series resistance $R_{BE}$, and a turn-off voltage $V_{BE}$ must also be determined.

Because the inverter operates from a constant voltage, the turn-off or second-breakdown energy $E_{S/b}$ is given by Eq. (84). However, the rating curves in Fig. 152 are based on measurements made in the $E_{S/b}$ test set with constant current drive, and on results calculated by use of Eq. (88). Therefore, an effective series inductance for the circuit is obtained by setting Eq. (84) equal to Eq. (88) and solving in terms of the inductance L or, in this case, $L_{eff}$.

$$L_{eff} = L\left[\frac{V_{CEX}(sus)}{V_{CEX}(sus) - V_{CC}}\right] \quad (89)$$

Eq. (89) is valid for all circuits that operate from a constant supply voltage. For the inverter circuit of Fig. 145, L = 10 microhenries, $V_{CC}$ = 240 volts, and $V_{CEX}$ = 400 volts (from 2N3585 published data). The value of 240 volts for the constant voltage $V_{CC}$ is composed of the sum of the supply voltage and the voltage induced across the primary of the transformer as one of the transistors in the circuit is turning off. When these values are substituted in Eq. (89), the computed effective series inductance is found to be 25 microhenries. $R_{BE}$ is calculated at approximately 0.5 ohm,

the equivalent series resistance of the diode in shunt with the 3.4-ohm resistor; $V_{BE}$ is measured at approximately 5 volts.

A set of rating curves for the 2N3585 is shown in Fig. 152. In the circuit used to obtain the curves, $R_{BE} = 20$ ohms, $L = 100$ microhenries, and $V_{BE} = -4$ volts. The minimum peak current for these values of $V_{BE}$ and $R_{BE}$ is given in Figs. 152(a) and 252(b) as 2 amperes.

To permit the application of the curves of Fig. 152 to the $R_{BE}$ and $V_{BE}$ of the inverter circuit, translation ratios must be calculated from the slope of the curves in Figs. 152(a) and 152(b).

For $R_{BE}$, the translation ratio is determined from the value of minimum peak current at an $R_{BE}$ of 0.5 ohm divided by the value of minimum peak current at 20 ohms; both values are taken from Fig. 152(a). The result is as follows:

$$R_{BE}\ (\text{trans}) = \frac{0.7}{2.0} = 0.35$$

For $V_{BE}$, the translation ratio is determined from the value of minimum peak current at a $V_{BE}$ of $-5$ volts divided by the value of minimum peak current at $-4$ volts; both values are taken from Fig. 152(b). This ratio is given by

$$V_{BE}\ (\text{trans}) = \frac{1.5}{2.0} = 0.75$$

The minimum peak current for the series inductance of 25 microhenries is determined from Fig. 152(c) as 3.4 amperes. The equivalent minimum peak current for the inverter circuit is obtained by translating this value as follows:

$$\begin{aligned} I_{pk} &= (3.4\text{A})\ (0.35)\ (0.75) \\ &= 0.89 \text{ amperes} \end{aligned}$$

This peak current can then be converted to the second-breakdown energy of the inverter circuit, as follows:

$$\begin{aligned} E_{S/b} &= \tfrac{1}{2} LI_p^2 \\ &= \tfrac{1}{2} (25\ \mu\text{H})\ (0.89\ \text{A})^2 \\ &= 10 \text{ microjoules} \end{aligned}$$

These values of $I_{pk}$ and $E_{S/b}$ are calculated safe-operation-area ratings. The actual peak current necessary for inverter turn-off is 0.8 ampere. For this current, the turn-off or second-breakdown energy is given by

$$\begin{aligned} E_{S/b} &= \tfrac{1}{2} L_{eff} I_{pk}^2 \\ &= (\tfrac{1}{2})\ (25\ \mu\text{H})\ (0.8\ \text{A})^2 \\ &= 8 \text{ microjoules} \end{aligned}$$

The average-power contribution to this turn-off energy is determined by dividing the energy by the pulse period, as follows:

$$P_{avg} = 8\ \mu\text{J}/40\ \mu\text{s} = 0.2\ \text{W}.$$

Substitution of this value in Eq. 73 yields

$$\begin{aligned} T_C\ (\text{eff}) &= T_C + P_{avg}\Theta_{J-C} \\ &= T_C + (0.2\ \text{W})\ (5^\circ\ \text{C/W}) \\ &= T_C + 1^\circ\text{C} \end{aligned}$$

where 5°C per watt is the junction-to-case thermal resistance for the 2N3585.

The final problem is to determine the maximum case temperature at which the inverter can safely turn off on a continuous basis. The temperature derating factor is calculated by dividing the actual peak current $I_{pk}(\text{act})$ by the maximum safe-area operation value of $I_{pk}(\text{S.A.})$, as follows:

$$\frac{I_{pk}(\text{act})}{I_{pk}(\text{S.A.})} = \frac{0.8}{0.89} = 0.9, \text{ or } 90\%$$

From Fig. 144, the temperature-derating curve for the transistor of interest, a 90-per-cent derating

factor indicates a $T_C$(max) of 60°C on the $I_{S/b}$-limiting curve. Thus the inverter can be safely turned off at a case temperature of 60°C −1°C = 59°C.

As a final check, the actual total energy absorbed by the circuit under reverse-bias conditions should be compared with the locus of peak pulse power (derated to the 59°C case temperature and expressed in terms of energy) on the forward-bias safe-area chart. If the total energy absorbed exceeds the forward-bias energy, additional derating on a thermal basis should be performed.

## THERMAL-CYCLING RATINGS

Significant temperature variations occur in power transistors because of changes in ambient temperature and in the power dissipation during operation. As explained in the section on **Thermal Factors**, these variations in temperature result in cyclic mechanical stresses at the interface of the semiconductor pellet and the metal header to which the pellet is bonded because of the difference in the thermal expansions of these parts. These stresses are a function of the difference in the coefficients of thermal expansion of the semiconductor and metallic materials, of the change in temperature at the interface, and of the dimensions of the interface.

Power transistors are subjected to thermal-cycling stresses in all practical applications. Table X lists examples of the thermal cycling that a power transistor may be required to withstand in several typical applications. These data show that the thermal-cycling requirements

Table X—Thermal-Cycling Requirements for Typical Applications of Power Transistors

| Application | Circuit | $P_T$ (W) | $\triangle T_C$ (°C) | Minimum Equipment Life Required (years) | Typical Thermal-Cycling Rating Required (cycles) |
|---|---|---|---|---|---|
| Auto radio audio output | Class A | 8 | 75 | 5 | 5,000 |
| | Class AB | 2 | 45 | 5 | 5,000 |
| Power supply | Series regulator | 50 | 65 | 5 | 5,000 |
| | Switching regulator | 15 | 65 | 5 | 5,000 |
| Hi-Fi audio amplifier | Class AB | 35 | 50 | 5 | 5,000 |
| Computer power supply | Series regulator | 50 | 65 | 10 | 10,000 |
| Computer peripheral equip. | Solenoid driver | 5 | 5 | 10 | $1.3 \times 10^8$ |
| Television | Vertical output | 10 | 75 | 5 | 5,000 |
| | Audio output | 8 | 75 | 5 | 5,000 |
| Sonar modulator | Linear amplifier | 100 | 55 | 10 | $144 \times 10^3$ |

may be very severe even in some of the more common types of applications. The cyclic stresses may eventually lead to transistor failures. This type of failure may be considered simply as fatigue wearout that results from continuous flexing of materials during thermal cycling.

## Improvement of Thermal-Cycling Capability

The thermal-cycling stresses set up at the interface of two dissimilar materials because of the difference in the coefficients of thermal expansion of the materials can be reduced by insertion of a material that has an intermediate expansion coefficient between them. Fig. 153(a) illustrates the use of a molybdenum slab as an expansion matcher in a silicon power transistor to reduce the cyclic thermal stresses between the silicon pellet and the copper header. Use of this technique can result in significant improvement in the thermal-cycling capability of power transistors.

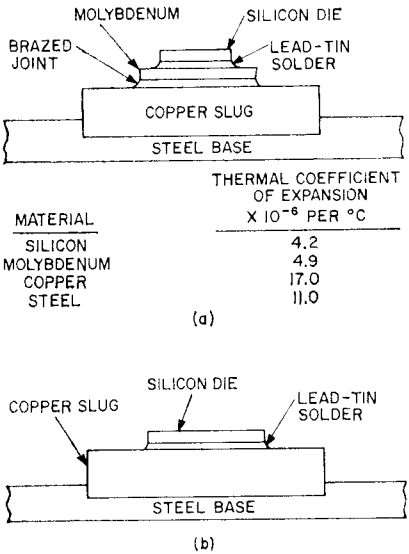



Figure 153. (a) Cross section of a transistor that uses a molybdenum expansion matcher between pellet and header; (b) cross section of a transistor in which pellet is soldered directly to copper.

Use of silicon-gold eutectic bonding to attach the semiconductor pellet to the header results in a pellet-to-header joint that can withstand a very large number of thermal cycles. When this type of hard-solder bonding is used, however, the stress generated because of a thermal mismatch is transmitted to the pellet, which in most power transistors is made of silicon. Because silicon is relatively weak in tensile strength and is highly "notch sensitive," the cyclic thermal stresses may result in the propagation of cracks in the silicon pellet unless either the pellet is very small or an expansion matcher is used.

In most silicon power transistors, lead solder is used to bond the pellet to the header, as shown in Fig. 153(b). The cyclic thermal stresses produced at the mounting interface are then absorbed by non-elastic deformation of the soft solder material, and very little stress is transmitted to the pellet. The continuous flexing of the solder, however, may eventually lead to fatigue failure in this material. Any impurities in the solder results in dislocation pile-ups that accelerate the failure. RCA has developed a process that significantly reduces the impurities introduced into the lead solder. Use of this proprietary "controlled solder process" makes it possible to avoid microcracks that propagate to cause fatigue failures in

Table XI—Thermal-Fatigue Performance of some Typical RCA Power Transistors

| Type | Pellet Size Mils x Mils | | Mounting Material | Material to which Die is Attached | Controlled Solder Process | Change in Case Temp. °C | Power Dissipation Watts | No. of Cycles to 10% Failure |
|---|---|---|---|---|---|---|---|---|
| 2N3773** | 250 | 250 | Lead | Copper | No | 42 | 85 | 1,000 |
| 2N3773 | 250 | 250 | Lead | Molybdeum | No | 42 | 85 | 9,600 |
| 2N3772 | 250 | 250 | Lead | Copper | Yes | 90 | 16 | 34,500** |
| 2N3055 | 180 | 180 | Lead | Copper | No | 65 | 50 | 3,500 |
| 2N3055 | 180 | 180 | Lead | Copper | Yes | 90 | 6.7 | 40,000*** |
| 2N6032 | 230 | 230 | Silicon Gold | Molybdeum | No | 53 | 105 | 12,793*** |
| 2N5298 | 130 | 130 | Lead | Copper | No | 50 | 18 | 10,000 |
| 2N5240 | 130 | 130 | Lead | Copper | Yes | 42 | 51 | 8,500*** |
| 2N5039 | 145 | 183 | Lead | Copper | Yes | 73 | 59 | 10,000*** |

* Early design.
** Test still operating.
*** Test terminated—less than 10% failure.

power transistors and, therefore, greatly increases the thermal-cycling capability of these devices. Table XI lists the results of thermal-cycling tests on several RCA power transistors that employ different pellet-to-header mounting system. These results confirm the effectiveness of the "controlled solder process."

## Thermal-Cycling Rating Chart

An equipment manufacturer should make certain that power-transistor circuits are designed so that cyclic thermal stresses are mild enough to assure that no transistor fatigue failures occur during the required operating life of his equipment. RCA has developed a rating system, the first of this type in the industry, that shows the number of thermal cycles that the transistor is rated to withstand as a function of the total transistor power dissipation and the change in case temperature.

Fig. 154 shows a typical thermal-cycling rating chart. This chart is provided in the form of a log-log presentation in which total transistor power dissipation is denoted by the ordinate and the thermal-cycling capability (number of cycles to failure) is indicated by the abscissa. Rating curves are shown for various magnitudes of change in case temperature. Use of this chart makes it possible for a circuit designer to avoid transistor thermal-fatigue failures during the operating life of his equipment. In general, power dissipation is a fixed system requirement. The designer also knows the number of thermal cycles that a power transistor will be subjected to during the minimum required life of the equipment. For these conditions, the chart indicates the maximum allowable change in case temperature. (If the rating point does not lie exactly on one of the rating curves, the allowable change in case temperature can be approximated by linear interpolation.) The designer can then determine the minimum size of heat

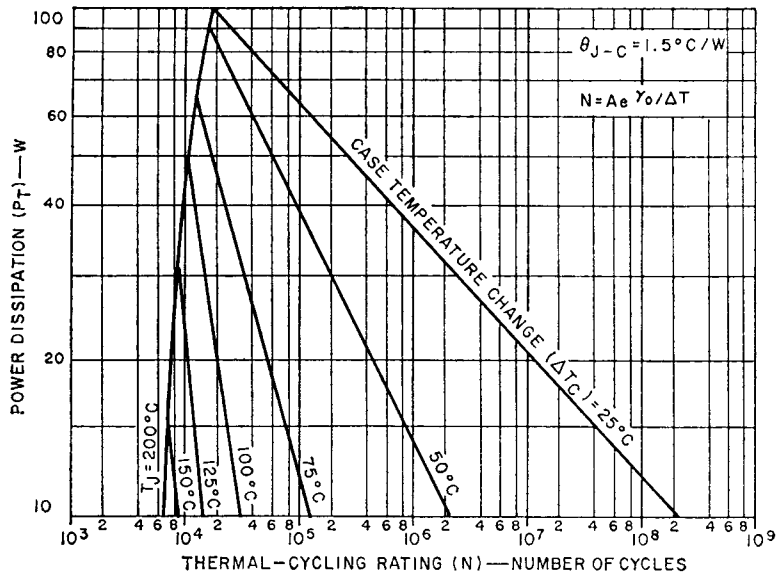



Figure 154. Thermal-cycling rating chart.

sink required to restrict the change in case temperature within this maximum value.

Thermal-cycling ratings are included in the technical data for all RCA silicon power transistors announced since January 1, 1971. Similar ratings are being added for earlier power transistors as sufficient data are accumulated.

## Thermal-Fatigue Testing

The RCA thermal-cycling ratings allows a circuit designer to use power transistors with assurance that fatigue failures of these devices will not occur during the minimum required life of his equipment. These ratings provide valid indications of the thermal-cycling capability of power transistors for all types of operating conditions. On the basis of these ratings, limiting conditions can be established during circuit design so that the possibility of transistor thermal-fatigue failures are avoided.

Obviously, all individual power transistors cannot be tested to determine their thermal-cycling capability because such tests are expensive, time consuming, and destructive. The validity of the RCA thermal-cycling ratings results from the application of stringent process controls at each step in the manufacture of power transistors and from the testing of a statistically significant number of samples. Thermal-cycling ratings for power transistors provide the same type of assurance that a device will not fail when operated within ratings as that provided by the more familiar voltage, current, and second-breakdown ratings.

During thermal-fatigue testing

of power transistors, the operating power for the device is usually equivalent to that expected to be applied during normal operation. The transistor is operated until the rise in case temperature is equal to the maximum value anticipated in the intended application. The case temperature is then reduced to the initial value by use of forced-air or water cooling. The cycle is repeated until failure occurs, as indicated by a significant increase in the transistor thermal resistance. The transistor heat sink and the timing of the temperature cycling are selected to simulate as closely as possible the actual conditions that the transistor will be subjected to in the actual application.

# Equivalent-Model Analyses of Power Transistors

POWER transistors are designed primarily for use in applications, such as drivers and power-output stages in linear systems and high-level switching, that demand large signal swings and high power outputs (usually greater than 1 watt). Power transistors, however, are not limited to such applications and, if desired, may be used successfully to provide low-level amplification. When used in such applications, their operation may be analyzed by use of small-signal models and concepts that are almost identical to those used in the analysis of small-signal transistors. In large-signal linear or switching applications, power transistors are operated under conditions for which the small-signal models do not provide a valid representation, and other models and concepts must be devised to analyze the operation of the transistors.

## SMALL-SIGNAL ANALYSIS OF POWER TRANSISTORS IN LINEAR SERVICE

Silicon power transistors may be used for a wide variety of circuit applications in which the output signal is proportional to an applied input signal. Such applications are lumped into a broad category referred to as linear service. In most linear-service applications, the terminal voltage and current of a transistor are small compared to the levels of voltage and current established by dc bias conditions. In this mode of operation, the transistor can be conveniently analyzed by means of a small-signal equivalent circuit.

### Small-Signal Equivalent Circuits

In a small-signal ac analysis, a transistor can be represented as a three-terminal linear network. From elementary circuit theory, it is known that a three-terminal linear network may be considered as a linear two-port configuration, as shown in Fig. 155. The terminal parameters for the two-port network are defined by six sets of equations which differ only in the

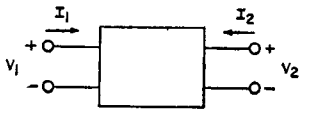



Figure 155. Linear two-port network.

parameters that are selected as dependent variables. Three sets of these equations have found some historical use in the analysis of transistors in common-emitter circuit configurations. These three sets of equations are listed below for both the general two-port network and a transistor used in a common-emitter (CE) circuit configuration:

1. **z-parameter equations:**

General $V_1 = z_{11}I_1 + z_{12}I_2$ (90)

CE $V_i = z_{ie}i_i + z_{re}i_o$ (91)

General $V_2 = z_{21}I_1 + z_{22}I_2$ (92)

CE $V_o = z_{fe}i_i + z_{oe}i_o$ (93)

2. **y-parameter equations:**

General $I_1 = y_{11}V_1 + y_{12}V_2$ (94)

CE $i_i = y_{ie}V_i + y_{re}V_o$ (95)

General $I_2 = y_{21}V_1 + y_{22}V_2$ (96)

CE $i_o = y_{fe}V_i + y_{oe}V_o$ (97)

3. **h-parameter equations:**

General $V_1 = h_{11}I_1 + h_{12}V_2$ (98)

CE $V_i = h_{ie}i_i + h_{re}V_o$ (99)

General $I_2 = h_{21}I_1 + h_{22}V_2$ (100)

CE $i_o = h_{fe}i_i + h_{oe}V_o$ (101)

In the double-subscript notation used in the common-emitter equations, the first subscript denotes the parameter function within the equivalent circuit (i.e., "i" denotes input, "o" denotes output, "f" denotes forward from input to output, and "r" denotes reverse from output to input), and the second subscript denotes the terminal common to both the input and output loops for the transistor configuration being employed. The equations shown are for the common-emitter transistor configuration; equivalent equations, however, can be written for the common-base and common-collector configurations.

The y terms defined by the equations are **short-circuit admittance parameters,** the z terms are **open-circuit impedance parameters,** and the h terms are the **hybrid-circuit parameters,** which may be defined as follows:

short-circuit input impedance

$$h_{ie} = \left.\frac{V_i}{i_i}\right|_{V_o=0} \quad (102)$$

short-circuit forward-current transfer ratio

$$h_{fe} = \left.\frac{i_o}{i_i}\right|_{V_o=0} \quad (103)$$

open-circuit reverse-voltage transfer ratio

$$h_{re} = \left.\frac{V_i}{V_o}\right|_{i_i=0} \quad (104)$$

open-circuit output conductance

$$h_{oe} = \left.\frac{i_o}{V_o}\right|_{i_i=0} \quad (105)$$

The equivalent circuit for these equations is shown in Fig. 156.

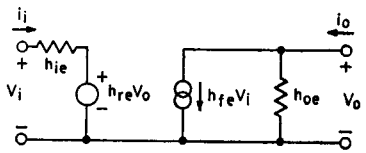



Figure 156. h-parameter equivalent circuit for a transistor used in a common-emitter configuration.

The techniques used to measure small-signal z, y, and h parameters are more or less implicit in their definitions. In general, the parameters most commonly measured are the short-circuit admittance parameters.

The z, y, and h parameters are in general complex quantities and, therefore, are frequency-dependent. At frequencies above about 100 MHz, these parameters are

extremely difficult to measure because it is hard to produce true open-circuit or short-circuit conditions. For this reason, another set of parameters, known as the "S" or **scattering parameters,** is used to derive an alternative model for transistors which can be employed for high-frequency design. The following equations define "S" parameters on the basis of the signal-flow diagram of the transistor shown in Fig. 157:

$$E_{r1} = S_{11}E_{i1} + S_{12}E_{i2} \qquad (106)$$

$$E_{r2} = S_{21}E_{i1} + S_{22}E_{i2} \qquad (107)$$

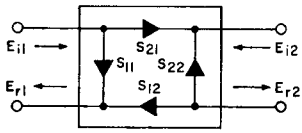



Figure 157. Scattering parameters for linear two-port network.

If Eqs. (106) and (107) are solved for the scattering parameters, the following results are obtained:

**voltage-reflection coefficient** at port 1 with port 2 terminated in a matched load

$$S_{11} = \left.\frac{E_{r1}}{E_{i1}}\right|_{E_{i2}=0} \qquad (108)$$

**reverse-transmission** voltage **ratio** with port 1 terminated in a matched load

$$S_{12} = \left.\frac{E_{r1}}{E_{i2}}\right|_{E_{i1}=0} \qquad (109)$$

**forward-transmission voltage ratio** with port 2 terminated in a matched load

$$S_{21} = \left.\frac{E_{r2}}{E_{i1}}\right|_{E_{i2}=0} \qquad (110)$$

**voltage-reflection coefficient** at port 2 with port 1 terminated in a matched load

$$S_{22} = \left.\frac{E_{r2}}{E_{i2}}\right|_{E_{i1}=0} \qquad (111)$$

It should be noted that the scattering parameters are complex numbers that have both magnitude and phase. The chief advantage of the scattering parameters is that they are measured under matched-termination conditions which are easier to obtain than a true open or short circuit.

It sometimes becomes convenient or necessary to convert one type of network parameter to another. This conversion is readily achieved if any pair of the parameter equations is solved for a different independent variable, and the variables are then equated. The results of such an analysis are shown in Table XII.

When the transistor is connected in a practical circuit, such as that shown in Fig. 158, the source and load impedances affect the terminal properties of the network. This effect can best be

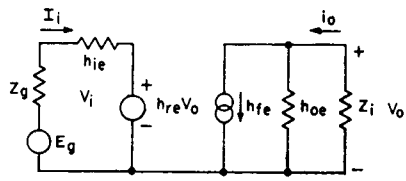



Figure 158. Terminated h-parameter equivalent circuit for a transistor used in a common-emitter configuration.

shown by calculation of the network input impedance from the h-parameter equations [Eqs. (99) and (101)] given previously. For convenience of reference, these equations are repeated below:

$$V_i = h_{ie} i_i + h_{re} V_o \qquad (99)$$

$$i_o = h_{fe} i_i + h_{oe} V_o \qquad (101)$$

Table XII—Parameter Equivalencies

| | z | | y | | h | | S | |
|---|---|---|---|---|---|---|---|---|
| z | $z_{11}$ | $z_{12}$ | $\frac{y_{22}}{\Delta y}$ | $\frac{-y_{12}}{\Delta y}$ | $\frac{\Delta h}{h_{22}}$ | $\frac{h_{12}}{h_{22}}$ | $\frac{1+S_{11}-S_{22}-\Delta S}{1-S_{11}-S_{22}+\Delta S}$ | $\frac{2S_{12}}{1-S_{11}-S_{22}+\Delta S}$ |
| | $z_{21}$ | $z_{22}$ | $\frac{-y_{21}}{\Delta y}$ | $\frac{y_{11}}{\Delta y}$ | $\frac{-h_{21}}{h_{22}}$ | $\frac{1}{h_{22}}$ | $\frac{2S_{21}}{1-S_{11}-S_{22}+\Delta S}$ | $\frac{1-S_{11}+S_{12}-\Delta S}{1-S_{11}-S_{22}+\Delta S}$ |
| y | $\frac{z_{22}}{\Delta z}$ | $\frac{-z_{12}}{\Delta z}$ | $y_{11}$ | $y_{12}$ | $\frac{1}{h_{11}}$ | $\frac{-h_{12}}{h_{11}}$ | $\frac{1-S_{11}+S_{22}-\Delta S}{1+S_{11}+S_{22}+\Delta S}$ | $\frac{-2S_{12}}{1+S_{11}+S_{22}+\Delta S}$ |
| | $\frac{-z_{21}}{\Delta z}$ | $\frac{z_{11}}{\Delta z}$ | $y_{21}$ | $y_{22}$ | $\frac{h_{21}}{h_{11}}$ | $\frac{\Delta h}{h_{11}}$ | $\frac{-S_{21}}{1+S_{11}+S_{22}+\Delta S}$ | $\frac{1+S_{11}-S_{22}-\Delta S}{1+S_{11}+S_{22}+\Delta S}$ |
| h | $\frac{\Delta z}{z_{22}}$ | $\frac{z_{12}}{z_{22}}$ | $\frac{1}{y_{11}}$ | $\frac{-y_{12}}{y_{11}}$ | $h_{11}$ | $h_{12}$ | $\frac{1-S_{11}+S_{22}-\Delta S}{1-S_{11}-S_{22}+\Delta S}$ | $\frac{2S_{12}}{1-S_{11}+S_{22}-\Delta S}$ |
| | $\frac{-z_{21}}{z_{22}}$ | $\frac{1}{z_{22}}$ | $\frac{y_{21}}{y_{11}}$ | $\frac{\Delta y}{y_{11}}$ | $h_{21}$ | $h_{22}$ | $\frac{-2S_{21}}{1-S_{11}-S_{22}-\Delta S}$ | $\frac{1-S_{11}-S_{12}+\Delta S}{1-S_{11}+S_{22}-\Delta S}$ |
| S | A | C | E | G | I | K | $S_{11}$ | $S_{12}$ |
| | B | D | F | H | J | L | $S_{21}$ | $S_{22}$ |

$$A=\frac{\Delta z+z_{11}-z_{22}-1}{\Delta z+z_{11}+z_{22}+1} \qquad C=\frac{2z_{12}}{\Delta z+z_{11}+z_{22}+1} \qquad E=\frac{(1-y_{11})(1+y_{22})-y_{12}y_{21}}{1+y_{11}+y_{22}+\Delta y}$$

$$B=\frac{2z_{21}}{\Delta z+z_{11}+z_{22}+1} \qquad D=\frac{\Delta z-z_{11}+z_{22}-1}{\Delta z+z_{11}+z_{22}+1} \qquad F=\frac{-2y_{21}}{1+y_{11}+y_{22}+\Delta y}$$

$$G=\frac{-2y_{12}}{1+y_{11}+y_{22}+\Delta y} \qquad I=\frac{\Delta h+h_{11}-h_{22}-1}{\Delta h+h_{11}+h_{22}+1} \qquad K=\frac{2h_{12}}{\Delta h+h_{11}+h_{22}+1}$$

$$H=\frac{(1+y_{11})(1-y_{22})-y_{12}y_{21}}{1+y_{11}+y_{22}+\Delta y} \qquad J=\frac{-2h_{21}}{\Delta h+h_{11}+h_{22}+1} \qquad L=\frac{-\Delta h+h_{11}-h_{22}+1}{\Delta h+h_{11}+h_{22}+1}$$

$$\Delta z = z_{11}z_{22}-z_{12}z_{21} \qquad \Delta h = h_{11}h_{22}-h_{12}h_{21}$$
$$\Delta y = y_{11}y_{22}-y_{12}y_{21} \qquad \Delta S = S_{11}S_{22}-S_{12}S_{21}$$

For the terminated network shown in Fig. 158, $V_o$ may be expressed by the following relationship:

$$V_o = -i_o z_L = -i_o / y_L \qquad (112)$$

Eq. 101) then becomes

$$i_o = h_{fe} i_i + h_{oe} (-i_o / y_L) \qquad (113)$$

or

$$i_o = \frac{h_{fe}(y_L) i_i}{y_L + h_{oe}} \qquad (114)$$

Substitution of Eq. (112) in Eq. (99) yields the following result:

$$V_i = \left( h_{ie} + \frac{h_{fe}\, h_{re}}{y_L + h_{oe}} \right) i_i \qquad (115)$$

If both sides of Eq. (115) are divided by $i_i$, the following equation for the input impedance $Z_{in}$ is obtained:

$$Z_{in} = \frac{V_i}{i_i} = h_{ie} + \frac{h_{fe}\, h_{re}}{y_L + h_{oe}} \qquad (116)$$

In a similar manner, the equation for the output impedance $Z_{out}$ may be derived to obtain the following relationship:

$$Z_{out} = \frac{V_o}{i_o} = \frac{Z_g + h_{ie}}{Z_g h_{oe} + h_{ie} h_{oe} - h_{re} h_{fe}} \qquad (117)$$

Eq. (114) is rewritten to obtain the following expression for the current gain $K_i$ of the circuit shown in Fig. 158:

$$K_i = \frac{i_o}{i_i} = \frac{h_{fe}\, y_L}{h_{oe} + y_L} \qquad (118)$$

The voltage gain $K_v$ of the circuit, as determined from Eqs. (112), (114) and (115), may be expressed by the following relationship:

$$K_v = \frac{V_o}{V_i} = \frac{-h_{fe}/h_{ie}}{y_L + h_{oe}\left(1 - \frac{h_{fe}\, h_{re}}{h_{ie}\, h_{oe}}\right)} \qquad (119)$$

The power gain PG, which is the product of the current and voltage gains, may be expressed as follows:

$$PG = K_i K_v \qquad (120)$$

This type of analysis can also be applied to the z and y parameters; the general results for all three circuits are shown in Table XIII.

In some cases, information is supplied for the common-emitter configuration, and it is desired to find the parameters for the common-base or the common-collector configuration. This conversion for the hybrid parameters is shown in Table XIV.

**Common-Emitter Equivalent Circuit**—The hybrid-pi small-signal circuit has become popular in transistor analyses because it offers a reasonable compromise between the "black-box" two-port representation (y, z, or h) and the complex equations derived from semiconductor physics. In addition, the hybrid-pi equivalent circuit represents a transistor by parameters which are independent of the operating frequency and which can be related to physical processes that occur within the transistor. The complete hybrid-pi equivalent circuit for a transistor in a common-emitter configuration is shown in Fig. 159. The discrete components have not been

combined into equivalent components so that the association of the components with transistor physical processes can be readily shown.

When a transistor is connected for normal operation (i.e., emitter-to-base junction forward-biased and collector-to-base junction reverse-biased) and a small increment of base-to-emitter voltage is applied, this incremental

## Table XIII—Network Terminal Properties

| Parameter | $Z_{in}$ | $Z_{out}$ | $K_i$ | $K_v$ |
|---|---|---|---|---|
| z | $\frac{\Delta z + z_{11} z_L}{z_{22} + z_L}$ | $\frac{\Delta z + z_{22} z_g}{z_{11} + z_g}$ | $\frac{z_{21}}{z_{22} + z_L}$ | $\frac{z_{21} + z_L}{\Delta z + z_{11} z_L}$ |
| y | $\frac{y_{22} + y_L}{\Delta y + y_{11} y_L}$ | $\frac{y_{11} + y_g}{\Delta y + y_{22} y_g}$ | $\frac{-y_{21} y_L}{\Delta y + y_{11} y_L}$ | $\frac{-y_{21}}{y_{22} + y_L}$ |
| h | $\frac{\Delta h + h_{11} y_L}{h_{22} + y_L}$ | $\frac{h_{11} + z_g}{\Delta h + h_{22} z_g}$ | $\frac{-h_{21} y_L}{h_{22} + y_L}$ | $\frac{-h_{21} z_L}{h_{11} + \Delta h z_L}$ |

$\Delta z = z_{11} z_{22} - z_{12} z_{21}$
$\Delta y = y_{11} y_{22} - y_{12} y_{21}$
$\Delta h = h_{11} h_{22} - h_{12} h_{21}$

## Table XIV—Hybrid Parameter Relationships

| | Common-Emitter Circuit | Common-Base Circuit | Common-Collector Circuit |
|---|---|---|---|
| $h_{11}$ | $h_{ie}$ | $h_{ib} = \frac{h_{ie}}{1 + h_{fe}}$ | $h_{ic} = h_{ie}$ |
| $h_{12}$ | $h_{re}$ | $h_{rb} = \frac{h_{ie} h_{oe}}{1 + h_{fe}} - h_{re}$ | $h_{rc} = 1 - h_{re}$ |
| $h_{21}$ | $h_{fe}$ | $h_{fb} = \frac{-h_{fe}}{1 + h_{fe}}$ | $h_{fc} = -(1+h_{fe})$ |
| $h_{22}$ | $h_{oe}$ | $h_{ob} = \frac{h_{oe}}{1 + h_{fe}}$ | $h_{oc} = h_{oe}$ |

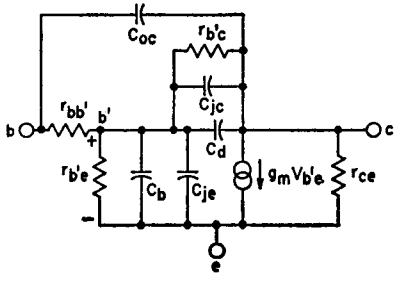



Figure 159. Complete hybrid-pi equivalent circuit for a transistor used in a common-emitter configuration.

change in base-to-emitter voltage produces two components of base current. One component is produced by the incremental increase of charge recombination in the base that is caused by the increase of excess charge stored in the base. This component $i_{b1}$ can be expressed as follows:

$$i_{b1} = g_{b'e} V_{be} = (1/r_{b'e}) V_{be} \quad (121)$$

The other base-current component results from the incremental change in excess majority carriers that is required to maintain electrical neutrality over the increased minority carriers stored in the base. This component $i_{b2}$ can be represented by the following relationship.

$$i_{b2} = C_b (dV_{be}/dt) \quad (122)$$

The total base current $i_b$ then is defined by the following equation:

$$_b = (1/r_{b'e}) V_{be} + C_b (dV_{be}/dt) \quad (123)$$

where $C_b$ is the base charging capacitance.

Eq. (123) provides the basis for a first-order equivalent-circuit approximation of the transistor. This equivalent circuit, shown in Fig. 160, represents the basic gain mechanism in a transistor.

For large voltage gains, a second-order effect requires modification of the basic circuit shown in Fig. 160. Under conditions of large voltage gain the load impedance is large. The width of the collector-junction depletion layer is voltage-dependent. As a result, the effective width of the transistor base varies with the output voltage. This phenomenon is known as "base-width modulation." These effects are accounted for in a transistor equivalent circuit by the addition of two feedback elements, $C_d$ and $r_{b'c}$, and one output shunt element $r_{ce}$ which account for the transient and incremental change of collector current $I_C$ because of base-width modulation. The resulting equivalent circuit is shown in Fig. 161.

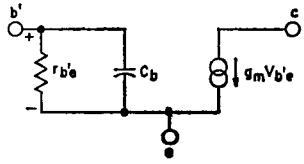



Figure 160. Elementary small-signal transistor equivalent circuit.

The equivalent circuit is essentially complete from the standpoint of physical mechanisms within the transistor. It is known, however, that the emitter and collector junctions have space-charge-layer capacitances, $C_{je}$ and $C_{jc}$, which must also be included in

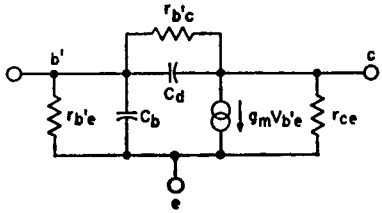



Figure 161. Elementary small-signal transistor equivalent circuit which includes components to represent effects of base-width modulation.

the equivalent circuit. In addition, the transverse majority-carrier base current produces a voltage drop in the transistor base which can be represented by addition of the spreading resistance $r_{bb'}$. Finally, because of the transverse voltage drop across $r_{bb'}$ it frequently becomes convenient or necessary to divide the collector junction capacitance into two components, the capacitance $C_{jc}$ mentioned previously and a capacitance $C_{oc}$ which is defined as the overlap capacitance of the collector-to-base junction. This division is desirable because the capacitance $C_{oc}$ is not charged through $r_{bb'}$.

When all the elements discussed above are combined into one circuit, the complete equivalent of the transistor, as shown previously in Fig. 159, is obtained. This equivalent circuit is rather cumbersome, but when parallel elements are combined the equivalent circuit is reduced to the more conventional form shown in Fig. 162.

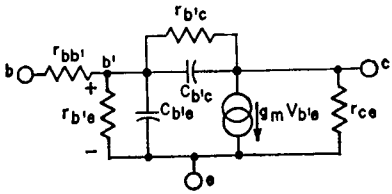



Figure 162. Conventional form of the complete hybrid-pi equivalent circuit shown in Fig. 159.

Even this equivalent circuit may seem too cumbersome for circuit analysis. In practice, however, certain circuit elements are dominant over a portion of the frequency spectrum, while other elements may have negligible effect on transistor behavior. It is permissible, therefore, to make further simplifications in the equivalent circuit which are applicable over a limited frequency range.

As with vacuum-tube amplifiers, it is convenient to use low-, medium-, and high-frequency equivalent circuits in the analysis of a transistor. The range of applicability is determined by the parameter values of the transistor represented by the equivalent-circuit model. A numerical example is necessary, therefore, to assess the relative importance of the various elements. In a small-signal analysis of a transistor, it is important to realize that, although the small-signal response depends on frequency, this response is also affected by the operating point and the temperature. In the following paragraphs, the response of a transistor at low, medium, and high frequencies is considered; changes in response because of variations in operating point and temperature are discussed later.

The numerical values for the hybrid-pi parameters of the RCA-2N2102 triple-diffused silicon planar transistor are given in the equivalent circuit shown in Fig. 163. As the first step in the estimation of the frequency response for this circuit, it is necessary to

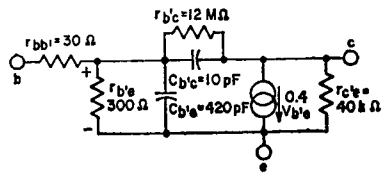



Figure 163. Hybrid-pi equivalent circuit for the 2N2102 triple-diffused silicon planar transistor.

calculate the frequency $f_1$ at which the resistance $r_{b'e}$ is equal to the reactance of the capacitance $C_{b'e}$. This frequency may be calculated as follows:

$$f_1 = \frac{1}{2\pi\, r_{b'e}\, C_{b'e}} \qquad (124)$$

$$= \frac{1}{6.28 \times 300 \times 420 \times 10^{-12}}$$

$$= 1.3 \text{ MHz}$$

For frequencies much below $f_1$, the resistance $r_{b'e}$ is dominant, and the effect of the capacitance $C_{b'e}$ is negligible. If a similar calculation is performed for the feedback elements $r_{b'c}$ and $C_{b'c}$, a frequency $f_2$ is determined, as follows:

$$f_2 = \frac{1}{2\pi\, r_{b'c}\, C_{b'c}} \qquad (125)$$

$$= \frac{1}{6.28 \times 12 \times 10^6 \times 10 \times 10^{-12}}$$

$$= 1.3 \text{ kHz}$$

For frequencies below $f_2$, the effect of the feedback capacitance $C_{b'c}$ is negligible.

The preceding calculations are based upon design-center measured values; individual elements, therefore, may vary somewhat from one transistor to another. Because the frequencies $f_1$ and $f_2$ differ by three orders of magnitude, several factors concerning the transistor operation become apparent. At frequencies much below $f_2$, both capacitances $C_{b'e}$ and $C_{b'c}$ may be neglected, and the low-frequency equivalent circuit shown in Fig. 164 is applicable. At frequencies above $f_2$ but much below $f_1$, $r_{b'e}$ and $C_{b'c}$ are negligible, and the middle-frequency equivalent circuit shown in Fig. 165 is useful. Finally, at frequencies much above $f_1$, both $r_{b'e}$ and $r_{b'c}$ are effectively bypassed by the shunt capacitances. For such frequencies, the high-frequency equivalent circuit shown in Fig. 166 shown be used to represent the transistor.

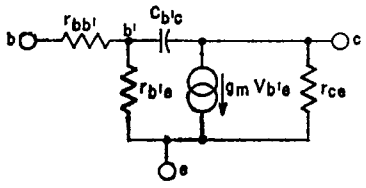



Figure 164. Simplified low-frequency hybrid-pi transistor equivalent circuit.

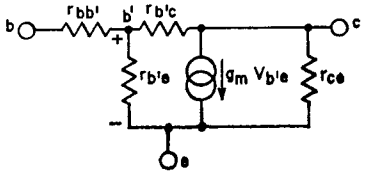



Figure 165. Simplified medium-frequency hybrid-pi transistor equivalent circuit.

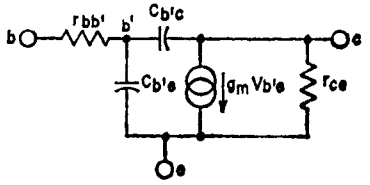



Figure 166. Simplified high-frequency hybrid-pi transistor equivalent circuit.

In practice, the mid-frequency model shown in Fig. 165 is not very useful. A more useful equivalent circuit is produced if both capacitors $C_{b'e}$ and $C_{b'c}$ are retained in the circuit. Such an equivalent circuit, shown in Fig. 167, is useful over the middle- and high-frequency ranges.

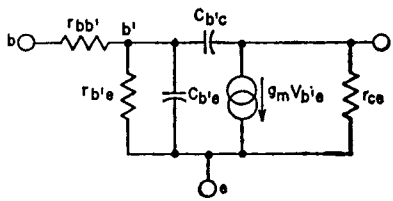



Figure 167. Simplified hybrid-pi transistor equivalent circuit for use at medium and high frequencies.

It is apparent from the foregoing discussion that the hybrid-pi equivalent circuit offers many ad-

vantages as a broad-band transistor model. If hybrid data are not specified by the transistor manufacturer, the parameters at a given operating point can be determined from known information or from simple low-frequency measurements, as follows:

The **transconductance** $g_m$ can be calculated from the following physical relationship:

$$g_m = \frac{q}{kT} \mid I_C \mid \text{mhos} \qquad (126)$$

where

$q$ = electronic charge = $1.6 \times 10^{-19}$ coulomb
$k$ = Boltzmann's constant = $1.38 \times 10^{-23}$ watt-sec/°K
$T$ = absolute temperature in °K
$I_C$ = collector current in amperes.

At room temperature, T = 290°K. Eq. (126) can then be rewritten as follows:

$$g_m = 0.04 \mid I_C \mid \text{mhos} \qquad (127)$$

where the collector current $I_C$ is given in milliamperes.

The **base-spreading resistance** $r_{b'e}$ is calculated on the basis of the measured value for the short-circuit current gain at low frequencies. Under the conditions required to measure the short-circuit current gain of the low-frequency equivalent-circuit model shown in Fig. 164, the relationship between the resistance $r_{b'e}$ and the low-frequency current gain may be expressed as follows:

$$h_{fe}|_{flo} = i_o/i_i = g_m r_{b'e} \qquad (128)$$

When this equation is solved for the resistance $r_{b'e}$, the following result is obtained:

$$r_{b'e} = \frac{h_{fe}|_{flo}}{g_m} \qquad (129)$$

At low frequencies, the **distributed base resistance** $r_{bb'}$ can be determined from the calculated value for the resistance $r_{b'e}$ and the measured value for the short-circuit input impedance. Under the conditions required for the masurement, the short-circuit input impedance for the low-frequency equivalent-circuit model can be expressed in terms of the resistances $r_{bb'}$ and $r_{b'e}$ as follows:

$$h_{ie}|_{flo} = V_i/i_i = r_{bb'} + r_{b'e} \qquad (130)$$

The low-frequency value for the resistance $r_{bb'}$, therefore, may be determined from the following relationship:

$$r_{bb'} = h_e|_{flo} - r_{b'e} \qquad (131)$$

It should be realized that the resistance $r_{bb'}$ is a distributed component and is, therefore, a function of frequency. At high frequencies, the following equation should be used to calculate this resistance:

$$1/r_{bb'} = R_e(y_{ie}) \qquad (132)$$

Eq. (84) is valid for the following condition:

$$\omega(C_{b'c} + C_{b'e}) > > 1/r_{bb'} \qquad (133)$$

In the calculation of the **collector-to-base feedback resistance** $r_{b'c}$, it is first necessary to measure the open-circuit reverse-voltage transfer ratio at low frequencies. For the low-frequency equivalent-circuit model shown in

Fig. 164, the open-circuit reverse-voltage transfer ratio can be related to the resistance $r_{b'c}$ as follows:

$$h_{fe}|_{flo} \simeq \frac{V_i}{V_o} = \frac{r_{b'e}}{r_{b'e} + r_{b'c}} \quad (134)$$

From practical considerations, it is known that $r_{b'c}$ is much greater than $r_{b'e}$. Eq. (134), therefore, may be rewritten as follows:

$$h_{re}|_{flo} = r_{b'e}/r_{b'c} \quad (135)$$

or

$$r_{b'c} = \frac{r_{b'e}}{h_{re}|_{flo}} \quad (136)$$

In most practical applications of power transistors, the **output shunt resistance** $r_{ce}$ is so large that its effect is negligible. In those cases for which the effect of this resistance is significant, however, the following approximation is valid:

$$h_{oe}|_{flo} \approx 1/r_{ce} \quad (137)$$

or

$$r_{ce} \approx \frac{1}{h_{oe}|_{flo}} \quad (138)$$

Before the **collector-to-base feedback capacitance** $C_{b'e}$ can be calculated, the **common-base output capacitance** $C_{ob}$ must be measured. Fig. 159 shows that the feedback capacitance is shunted by the **diode overlap capacitance** $C_{oc}$ and the **header capacitance** $C_H$. The capacitance $C_{b'c}$, therefore, can be determined from the following relationship:

$$C_{ob} = C_{b'c} + C_{oc} + C_H \quad (139)$$

$$C_{b'c} = C_{ob} - C_{oc} - C_H \quad (140)$$

For small-signal conditions, it is frequently possible to neglect the $C_{oc}$ and $C_H$ terms so that, as a first approximation, the capacitance $C_{b'c}$ may be determined as follows:

$$C_{b'c} \approx C_{ob} \quad (141)$$

The remaining capacitance $C_{b'e}$ may be determined by use of the high-frequency equivalent circuit shown in Fig. 167. For this circuit, the short-circuit current gain is determined from the following relationship:

$$h_{fe}|_{fhi} = \frac{i_o}{i_i} = \frac{g_m}{\omega(C_{b'e} + C_{b'c})} \quad (142)$$

If the gain-bandwidth product $f_T$ is defined as the frequency at which $h_{fe} = 1$, the following frequency relationships become apparent:

$$h_{fef} = f_T \quad (143)$$

$$h_{fe\omega} = \omega_T \quad (144)$$

$$\omega = \omega_T/h_{fe} \quad (145)$$

If the relationship expressed by Eq. (144) is substituted into Eq. (142), the following result is obtained:

$$\omega_T = g_m/(C_{b'e} + C_{b'c}) \quad (146)$$

$$C_{b'e} = (g_m/\omega_T) - C_{b'c} \quad (147)$$

Eqs. (126) through (147) demonstrate how the hybrid-pi parameters may be calculated for a known operating condition. For convenience and reference, these relations are summarized in Table XV.

Table XV—Hybrid-Pi Parameters

| Parameter | | Relationship |
|---|---|---|
| $g_m$ | = | $\frac{q}{KT}\lvert I_c\rvert$ |
| $r_{b'e}$ | = | $\frac{h_{fe}\lvert f_{lo}}{g_m}$ |
| $r_{bb'}\lvert f_{lo}$ | = | $h_{ie}\lvert f_{lo} - r_{b'e}$ |
| $r_{bb'}\lvert f_{hi}$ | = | $\frac{1}{R_e}(y_{ie})$ |
| For $\omega (C_{b'e} + C_{b'c}) >> \frac{1}{r_{bb'}}$ | | |
| $r_{b'c}$ | ≅ | $\frac{r_{b'e}}{h_{re}\lvert f_{lo}}$ |
| $r_{ce}$ | ≅ | $h_{oe}\lvert f_{lo}$ |
| $C_{b'c}$ | = | $C_{ob} - C_{oc} - C_H \cong C_{ob}$ |
| $C_{b'e}$ | = | $\frac{g_m}{\omega_T} - C_{b'c}$ |

**Equivalent Input Circuits**—Because of the differences in the magnitude of the elements of the hybrid-pi equivalent circuit, it becomes convenient to use this circuit to derive an equivalent input circuit for a transistor. If $R_b$ and C in the amplifier circuit shown in Fig. 168 are both very large, their effects are negligible, and the equivalent circuit is as shown in Fig. 169. The node equations for this circuit can be written as follows:

$$V_{b'e}[(1/r_{b'e}) + sC_{b'e} + sC_{b'c}] - V_o(sC_{b'c}) = I_i \quad (148)$$

$$V_{b'e}(g_m - sC_{b'c}) + V_o[(1/r_{ce}) + (1/R_L) + sC_{b'c}] = 0 \quad (149)$$

where $s = j\omega C$.

When conductance terms are substituted for the resistance terms, the node equations become

$$V_{b'e}[g_{b'e} + s(C_{b'e} + C_{b'c})] - V_o(sC_{b'c}) = I_i \quad (150)$$

$$V_{b'e}(g_m - sC_{b'c}) + V_o(g_{ce} + G_L + sC_{b'c}) = 0 \quad (151)$$

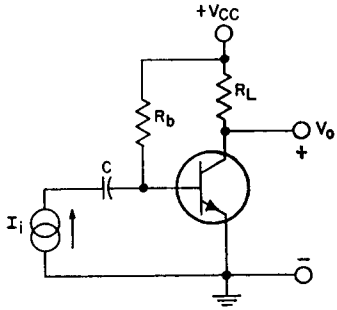



Figure 168. Common-emitter transistor amplifier.

Eqs. (150) and (151) may be further simplified by practical considerations. In the second term of Eq. (150), the $sC_{b'c}$ element represents feedback from output to input which is necessary and critical. In Eq. (151), the $sC_{b'c}$ element in

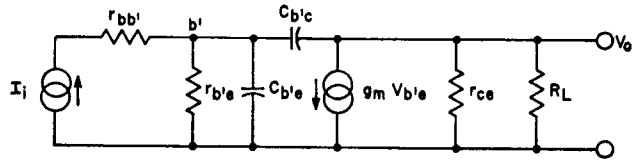



Figure 169. Equivalent circuit for common-emitter amplifier.

the first term represents signal fed forward from the input to the output, which is normally negligible compared to the signal fed forward by the $g_m V_{b'e}$ generator. In the second term of Eq. (151), the $sC_{b'c}$ term represents the loading of the output node, which is normally negligible. In addition, except for very large load resistances, $g_{ce}$ is much larger than $G_L$. Under such conditions, the $g_{ce}$ term can be neglected. When these facts are taken into consideration, Eqs. (150) and (151) may be rewritten in the following form:

$$V_{b'c}\,[g_{b'e} + s(C_{b'e} + C_{b'c})] - V_o\,(sC_{b'c}) = I_i \qquad (152)$$

$$V_{b'e}\,(g_m) + V_o\,G_L = 0 \qquad (153)$$

Eqs. (150) and (151) can be solved for the internal input admittance $I_i/V_{b'e}$ as follows:

$$V_o = (g_m/V_{b'e})/G_L \qquad (154)$$

$$V_{b'e}\,[g_{b'e} + s(C_{b'e} + C_{b'c})] + [(V_{b'e}\,g_m)/G_L]\,(sC_{b'c}) = I_i \qquad (155)$$

$$V_{b'e}\,\{g_{b'e} + s[C_{b'e} + C_{b'c}\,(1 + g_m\,R_L)]\} = I_i \qquad (156)$$

$$y = \frac{I_i}{V_{b'e}} = g_{b'e} + s\,[C_{b'e} + C_{b'c}\,(1 + g_m R_L)] \qquad (156)$$

Eq. (156) represents an equivalent input circuit that consists of a resistor $r_{b'e}$ in shunt with an equivalent capacitance expressed by the following equation:

$$C_{eq} = C_{b'e} + C_{b'c}\,(1 + g_m\,R_L) \qquad (157)$$

When the series spreading resistance $r_{bb'}$ is added, the total input circuit can be represented as shown in Fig. 170. Eq. (156) was derived for the case of a purely

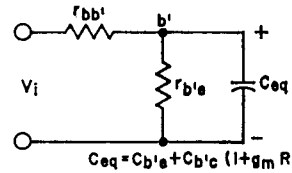



Figure 170. Equivalent input circuit for common-emitter amplifier.

resistive load; this equation, however, can be generalized so that it is applicable to a complex load impedance $Z_L$ as follows:

$$y = g_{b'e} + s\,[C_{b'e} + C_{b'c}\,(1 + g_m\,Z_L)] \qquad (158)$$

A circuit representation of the generalized input impedance defined by Eq. (158) is shown in Fig. 171. The equivalent admit-

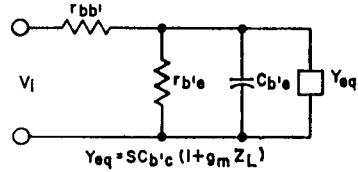



Figure 171. General equivalent input circuit for a common-emitter amplifier.

tance for this circuit is expressed by the following equation:

$$y_{eq} = sC_{b'c}\,(1 + g_m\,Z_L) \qquad (159)$$

Eqs. (156) and (158) show that the small feedback capacitance $C_{b'c}$ can have a significant effect on the frequency response of the amplifier because of the large equivalent capacitance reflected at the input. For example, if the component values for the 2N2102 power transistor are used and the load resistance $R_L$ is assumed to be

1000 ohms, the reflected capacitance may be calculated from Eq. (157) as follows:

$$C_{eq} = C_{b'e} + C_{b'c}(1 + g_m R_L)$$
$$= 420\,pF + 10\,pF\,(1 + 0.4 \times 10^3)$$
$$= 420\,pF + 4010\,pF = 4430\,pF$$

This calculation shows that the small 10-picofarad collector-to-base capacitance $C_{b'c}$ is reflected back to the input terminals as a 4010-picofarad capacitance which overshadows the $C_{b'e}$ term by a factor of 10. This reflected capacitance is analogous to the Miller effect in vacuum-tube triodes where the grid-to-plate capacitance reflected back to the input terminals is multiplied by $(1 + g_m R_L)$.

**Common-Base Equivalent Circuit**—Historically, the common-base circuit was the first configuration used for transistors. This circuit configuration, however, offers low input impedance and less than unity current gain and is no longer used except in certain specific applications.

Although the hybrid-pi equivalent circuit is still applicable for the common-base configuration, it becomes difficult to take into account the effects of base-width modulation which now produce coupling from the output to the common terminal. For small voltage gain, the base-width modulation effects are negligible, and the hybrid-pi common-emitter equivalent circuit shown in Fig. 162 can be redrawn for the common-base configuration as shown in Fig. 172. By suitable manipulation of the circuit equations, it can be shown that this equivalent circuit reduces to the common-base "T" equivalent circuit shown in Fig. 173. In this circuit, the resistance $r_e$ and

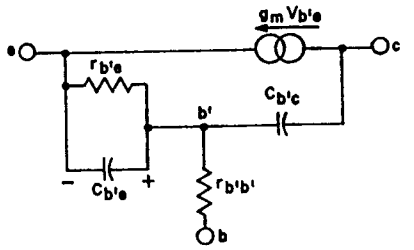



Figure 172. Hybrid-pi equivalent circuit for a transistor used in a common-base configuration (effects of base-width modulation are neglected).

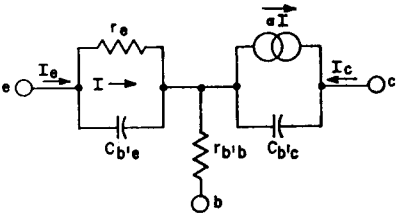



Figure 173. "T" equivalent circuit for a transistor used in a common-base configuration when the effects of base-width modulation are neglected.

the current-gain parameter $\alpha$ are defined as follows:

$$r_e = r_{b'e} \parallel \frac{1}{g_m} = \frac{1}{(1/r_{b'e}) + g_m} \approx \frac{1}{g_m} \qquad (160)$$

$$\alpha = \frac{g_m}{g_m + (1/r_{b'e})} \approx 1 \qquad (161)$$

**Common-Collector Equivalent Circuit** — The common-collector configuration is sometimes used because it offers high input impedance and current gain. It is possible to draw the hybrid-pi equivalent circuit for the common-collector configuration as shown in Fig. 174. It can be shown that if the base-modulation, base-spreading, output-loading, and input-loading terms are neglected, the

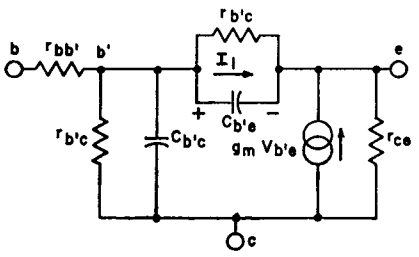



Figure 174. Hybrid-pi equivalent circuit for a transistor used in a common-collector configuration.

common-collector equivalent circuit can be simplified to the form shown in Fig. 175, in which the dependent generator is converted into a current-dependent source.

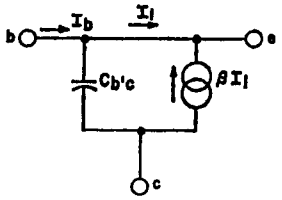



Figure 175. Simplified form of hybrid-pi equivalent circuit shown in Fig. 174.

## Network Properties

In transistor circuit design, one of the first decisions to be made is the type of circuit configuration (common-base, common-collector, or common-emitter) to be used. In the preceding section, the equivalent circuits for each of these configurations were described, and the way that the network terminal properties ($Z_{in}$, $Z_{out}$, $K_i$, and $K_v$, as defined in Table XIII) are affected by the source and load impedances was discussed. The terminal properties of any configuration may be calculated as a function of $Z_L$ and $Z_g$ by use of the equations shown in Table XIII. The calculations are laborious, and the results are readily available in the literature. The general qualitative results are shown in Table XVI. Because the common-emitter configuration provides the highest power gain ($K_iK_v$), this type of configuration is normally used unless the impedance properties of one of the other two configurations are required.

## Frequency Considerations

In the hybrid-pi model of a power transistor, the two capacitances $C_{b'e}$ and $C_{b'c}$ define two critical frequencies $f_1$ and $f_2$ that determine the significant elements in the model at a particular frequency. These two capacitances also affect the short-circuit current gain. For the medium- and high-frequency equivalent circuit shown in Fig. 176, the short-circuit current gain $h_{fe}$ can be determined from the following equations:

Table XVI—Qualitative Comparison of Transistor Circuit Configurations

| Property / Terminal | Common-Base Circuit | Common-Collector Circuit | Common-Emitter Circuit |
|---|---|---|---|
| $Z_{in}$ | low | high | moderate |
| $Z_{out}$ | extremely high | moderate | high |
| $K_i$ | low ($< 1$) | high | high |
| $K_v$ | high | low ($< 1$) | high |

$$V_{b'e} = \frac{i_i}{y_i|_{s.c.}} = \frac{i_i}{(1/r_{b'e})+s(C_{b'e}+C_{b'c})} \quad (162)$$

$$i_o = g_m V_{b'e} \quad (163)$$

$$h_{fe} = \frac{i_o}{i_i} = \frac{g_m}{(1/r_{b'e}) + s(C_{b'e} + C_{b'c})} \quad (164)$$

The s term ($s = j\omega C$) in Eq. (164) shows that $h_{fe}$ is a complex number that has both magnitude

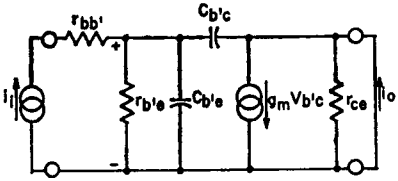



Figure 176. Medium- and high-frequency common-emitter hybrid-pi transistor model used for calculation of $h_{fe}$.

and phase. The short-circuit current ratio $h_{fe}$ has a low-frequency value of $(g_m r_{b'e})$, and a single breakpoint occurs at a frequency defined by the following equations:

$$(1/r_{b'e}) = 2\pi f_\beta (C_{b'e} + C_{b'c}) \quad (165)$$

$$f_\beta = \frac{1}{2\pi r_{b'e} (C_{b'e} + C_{b'c})} \quad (166)$$

The beta-cutoff frequency $f_\beta$ is the frequency at which the short-circuit current ratio $h_{fe}$ (or $\beta$) is reduced to 0.707 of its low-frequency value. At this frequency, the phase angle of $h_{fe}$ is $-45°$. This information is used to determine the variations in the magnitude and phase of $h_{fe}$ for the 2N2102 power transistor as a function of frequency shown in Figs. 177 and 178. For frequencies much above $f_\beta$, the response curve becomes asymptotic to a line that has a slope of $-1$ on the log-log

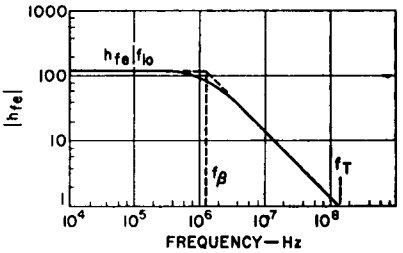



Figure 177. Magnitude of $h_{fe}$ as a function of frequency for the 2N2102 silicon power transistor.

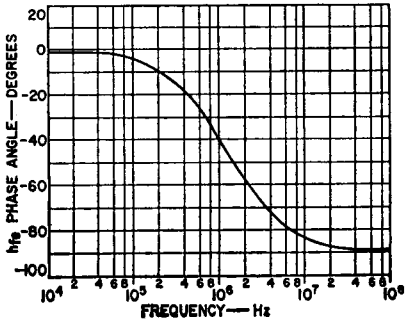



Figure 178. Phase of $h_{fe}$ as a function of frequency for the 2N2102 silicon power transistor.

scale shown. Extrapolation of this asymptote to the frequency at which $h_{fe} = 1$ defines another critical frequency $f_T$, which may be expressed by the following equation:

$$f_T = 2\pi \frac{g_m}{(C_{b'e} + C_{b'c})} \quad (167)$$

The term $f_T$ is called the gain-bandwidth product. At any frequency along the $-1$ asymptote (i.e., $f > 3\ f\beta$), Eq. (167) can be rewritten in terms of the operating frequency, as follows:

$$|\ h_{fe}\ |\ \omega = \omega_T \quad (168)$$

$$|\ h_{fe}\ |\ f = f_T \quad (169)$$

For the hybrid-pi model for the common-base configuration shown in Fig. 179, the common-base short-circuit current gain $h_{fb}$ can

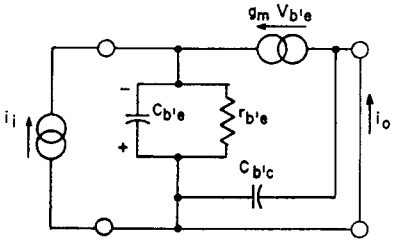



Figure 179. High-frequency common-base hybrid-pi transistor model used for calculation of $h_{fb}$ (transverse voltage drops in the base are neglected).

be calculated, for the case when $r_{bb'} = 0$, from the following equation:

$$h_{fb} = \frac{i_o}{i_i} = \frac{g_m}{(1/r_{b'e}) + sC_{b'e}} \quad (170)$$

The s term in Eq. (170) indicates that $h_{fb}$ is a complex number that has both magnitude and phase. This gain parameter has a low-frequency value of $(g_m r_{b'e})$ and a single breakpoint. This breakpoint occurs at a frequency $f_\alpha$, referred to as the alpha cutoff frequency, at which $h_{fb}$ is reduced to 0.707 of its low-frequency value. The frequency $f_\alpha$ is determined as follows:

$$f_\alpha = \frac{1}{2\pi\, r_{b'e}\, C_{b'e}} \quad (171)$$

## Variation of Small-Signal Parameters

During the discussion of the small-signal equivalent circuits, it was mentioned that the transistor parameters also depend on the operating conditions, or bias and temperature. This section provides a qualitative analysis of this dependence. The hybrid-pi equivalent circuit is particularly useful for this discussion because the parameters can be easily related to physical mechanisms within the transistor (as given in Table XV), which in turn can be related to the properties of the semiconductor material. It should be noted that the hybrid-pi equivalent circuit was derived as a linear, one-dimensional model, based upon an assumption of low-level injection of minority carriers. In particular, the model does not take into account transverse voltage drops, which have been shown to limit the useful safe area of operation of some transistors because of second breakdown. Over a major portion of the normal operating region, however, the model is quite adequate for the analysis.

**Bias Dependence**—The general hybrid-pi equivalent circuit shown in Fig. 159 includes three parameters that are associated with the injection of minority carriers and the basic gain mechanism in transistors. These parameters are $g_m$, $r_{b'e}$ and $C_b$. Table XV shows that the transconductance $g_m$ may be defined as follows:

$$g_m = \frac{q}{kT} \mid I_C \mid \quad (172)$$

Eq. (172) indicates that $g_m$ is directly dependent upon the collector current and is independent of the collector voltage $V_C$. Table XV also shows that $r_{b'e}$ may be determined from the following relationship:

$$r_{b'e} = \frac{h_{fe}|_{flo}}{g_m} = \frac{h_{fe}|_{flo}}{\frac{q}{kT}|I_C|} \qquad (173)$$

This equation shows that $r_{b'e}$ is inversely proportional to the collector current. The base charging capacitance $C_b$ is added to the equivalent-circuit model to account for the incremental change in the excess majority carriers stored in the base. This capacitance is defined by the following equation:

$$i_{b2} = (C_b \, d \, V_{be})/dt \qquad (174)$$

Eq. (174) may be rewritten in the following form:

$$dQ_b = i_{b2}dt = C_b dV_{be} \qquad (175)$$

or

$$C_b = dQ_b/d\,V_{be} \qquad (176)$$

where $dQ_b$ is the incremental change in the excess charge stored in the base.

Because electrical neutrality is preserved in the base, $dQ_b$ must also be equal to the incremental change in the excess minority carriers stored in the base region. During normal operation, the excess minority charge injected at the collector is very small compared to that injected at the emitter, and may be neglected. Under these conditions, the minority-carrier concentration in the base region of a uniform-base n-p-n transistor becomes as shown in Fig. 180. The total base charge can be found by integration of Eq. (175), and for this simple distribution is represented by the area inside the triangle. The equation for the total base charge, therefore, may be written as follows:

$$Q_b = (q\,n'_{b(o)}\,WA)/2 \qquad (177)$$

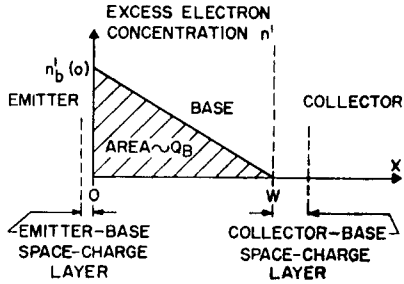



Figure 180. Excess minority-carrier concentration in the base region of a uniform-base n-p-n transistor.

where q is the electronic charge, $n'_{b(o)}$ is the concentration of excess electrons at the edge of the base region (X = 0), W is the base width, and A is the area perpendicular to the direction of electron injection.

An analysis of the diffusion-current mechanism in transistors indicates that the collector current may be determined from the following equation:

$$I_C = qD_nA\left[\frac{n'_{b(o)}}{W}\right] \qquad (178)$$

where $D_n$ is the diffusion constant for minority-carrier electrons.

In charge-control theory, a term known as the average base-charge replacement time, $T_F$, is introduced. This term is defined as follows:

$$T_F = Q_b/I_C = W^2/2D_n \qquad (179)$$

Eqs. (176) through (179) are combined to obtain the following equation, which can be used to determine the variation in $C_b$ with incremental changes in $Q_b$:

$$C_b = \frac{dQ_b}{dI_C}\left(\frac{d\,I_C}{d\,V_{be}}\right) = \left(\frac{W^2}{2\,D_n}\right)g_m = T_F g_m \qquad (180)$$

It has been shown previously, by Eq. (172), that $g_m$ is directly dependent upon $I_C$; Eq. 180), therefore, shows that the same dependence applies for $Q_b$.

The relationships expressed by Eqs. (172) through (180) can be used to predict the effect of collector voltage on each of the three parameters (i.e., $g_m$, $r_b'_e$, and $C_b$) being considered. Eqs. (172) and (173) show that $g_m$ and $r_b'_e$ are independent of $V_C$. The base width W is inversely dependent upon $V_C$ because, as the reverse bias across the collector-to-base space-charge layer is increased, the effective base width is decreased. Eqs. (179) and (180) show that $C_b$ is directly proportional to $T_F$ and, therefore, is inversely proportional to the square of $V_C$.

The two junction capacitances $C_{je}$ and $C_{jc}$ in the complete hybrid-pi equivalent circuit, shown in Fig. 159, are included to explain the voltage-dependent charge associated with the dipole space-charge layer at each junction. These capacitances are practically independent of current and vary with voltage according to the following equation:

$$C_J = K/(V'/n) \qquad (181)$$

where K is the material constant, V' is the voltage across the space-charge layer, and n is the junction constant (for an abrupt junction n = 2; for a graded junction n = 3).

The $r_b'_c$, $r_{ce}$, and $C_d$ terms in the hybrid-pi equivalent circuit are all associated with base-width modulation effects and are not too significant in normal applications.

The $r_{bb}'$ term is included in the hybrid-pi equivalent circuit to account for the voltage drops in the base caused by transverse majority-carrier currents. At high collector currents, this transverse voltage tends to concentrate the emitter injection current at the emitter edge, an effect termed "current crowding". The net result is that the effective length of the path for majority carriers is shortened, the transverse voltage drop is reduced, and therefore $r_{bb}'$ is decreased. At high collector voltages, the collector-to-base space-charge layer is increased; as a result, the base width W is reduced. If $r_{bb}'$ is considered to be a bulk resistance, then, from Fig. 181 and a knowledge of resistance effects, $r_{bb}'$ can be expressed as follows:

$$r_{bb}' = P_b X/A = P_b X/WY \qquad (182)$$

where Y is the transistor dimension perpendicular to the flow of base current in the plane of A. Since W is reduced as $V_C$ increases, $r_{bb}'$ should increase with increased $V_C$.

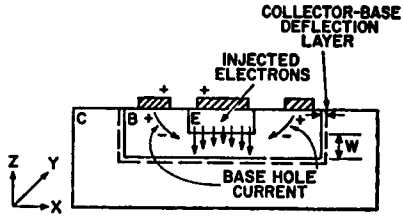



Figure 181. Model of n-p-n transistor showing transverse base field and current-crowding effect.

**Temperature Dependence**—The parameters shown in the hybrid-pi equivalent circuit are also temperature-dependent. Eq. (160) shows that $g_m$ is inversely proportional to temperature. It can be determined from Eq. (173) that $r_b'_e$ is proportional to $h_{fe}|f_{lo}(T)$. Because the parameter $h_{fe}|f_{lo}$ in-

Table XVII—Summary of Dependence of Model Parameters on Operating Condition. (Except where in parentheses, statements refer to low-level injection conditions and voltages low enough for no avalanche multiplication effects.)

| | Increasing $I_C$ | Increasing $V_C$ | Increasing T |
|---|---|---|---|
| $g_m$ | $\propto \lvert I_C \rvert$ (Increases less rapidly at high currents.) | (slight increase) | $\propto 1/T$ |
| $h_{fe\lvert LO}$ | Constant at moderate currents. Falls at low currents. (Falls at high currents.) | Increases steadily (Increases more rapidly at higher voltages; becomes infinite at sustaining voltage.) | Increases steadily. |
| $r_{b'e}$ | Almost $\propto 1/\lvert I_C \rvert$ if $\beta_0$ shows little variation. | Increases steadily. (Becomes infinite at sustaining voltage.) | Increases more rapidly than $\propto T$. |
| $r_{b'c}$ | $\propto 1/\lvert I_C \rvert$ (Falls more rapidly at higher currents.) | Increases at low voltages. [Falls at higher voltages (main effect) due to onset of avalanche multiplication.] | Approximately same as $\beta_0$. |
| $r_{ce}$ | $\propto 1/\lvert I_C \rvert$ | Increases at low voltages; at higher voltages (main effect) *falls* due to reduction of base width (and onset of avalanche multiplication). | Insensitive |
| $r_{bb'}$ | Constant (falls at high currents). | Increases steadily. | $\propto T^{m+2}$; for Si and Ge $-0.3 \leq m \leq +0.7$ $\propto 1/D_bT$ |
| $C_b$ | $\propto \lvert I_C \rvert$ (At higher currents: uniform base, less rapid increase; graded base, more rapid increase.) | Reduces steadily. | $\propto T^m$; where $+0.3 < m < +0.7$ for all Si and for Ge p-n-p $m \approx -0.3$ for Ge n-p-n |
| $C_{je}$ | Weak increase | Insensitive | Insensitive |
| $C_{jc}$ | Insensitive | $\propto \lvert V_{CB} \rvert^{-n}$ approx. $n = \frac{1}{2}$ to $\frac{1}{3}$ | Insensitive |
| $C_d$ | $\propto \lvert I_C \rvert$ | Decreases | $\propto T^{m+1}$ (see $C_b$) |
| $f_T$ | Increases (steadily for uniform base; reaches peak and then falls at high currents for graded base). | Increases somewhat due to reduction in base width. | 1. Reduces as $T^{-1}$ if $C_{je}$ and $C_{b'c}$ are comparable with $C_b$, (except at high currents). 2. Approximately $\propto D_b$ if $(C_{je}+C_{b'c}) \ll C_b$. |

creases with temperature over the normal range of interest, $r_{b'e}$ also increases with temperature, but at a rate faster than that at which the temperature changes. Eqs. (172), (179), and (180) can be solved simultaneously to show that $C_b$ is proportional to $1/D_nT$. When the temperature variation $h_{fe}D_n$ is considered, $r_{b'e}$ is proportional to $T^m$, where m is a material constant. The temperature dependence of $r_{bb'}$ may be inferred from Eq. (182) with the knowledge that, for the normal resistivities used, $\rho$ increases with temperature, so that $r_{bb'}$ also increases. The base-modulation terms show very little temperature dependence and so are not discussed here.

It is apparent that the parameter variations with temperature and bias involve many interrelated mechanisms, and only those of an elementary nature have been discussed. The discussion is intended to serve as a starting point and to indicate a direction in which the parameter may be expected to travel based upon the assumptions of the equivalent-circuit model. Table XVII summarizes these discussions, and Figs. 182 and 183 present measured data on the RCA-2N2102 silicon power transistor.

## LARGE-SIGNAL ANALYSIS OF POWER TRANSISTORS IN LINEAR SERVICE

The previous discussions of device theory and construction have described in detail the physical configuration of junction transistors. A junction transistor may be considered as two p-n junction diodes coupled back to back with a common region (the base) of the same conductivity material between them. For the purposes of this discussion, the most important feature of the base region is that it is very thin. It was also pointed out that excess carriers could be injected or extracted at either p-n junction by the application of an external voltage across the junction. Because the base region is so thin, it is possible for excess minority carriers, which are injected into the base by a forward-biased emitter-to-base junction, to be transported across the base without recombining with majority carriers. When they reach the collector

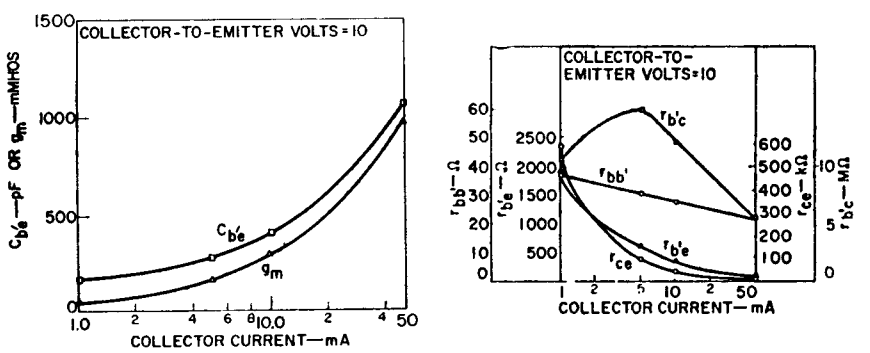



Figure 182. Hybrid-pi parameters as a function of collector current for the 2N2102 silicon power transistor.

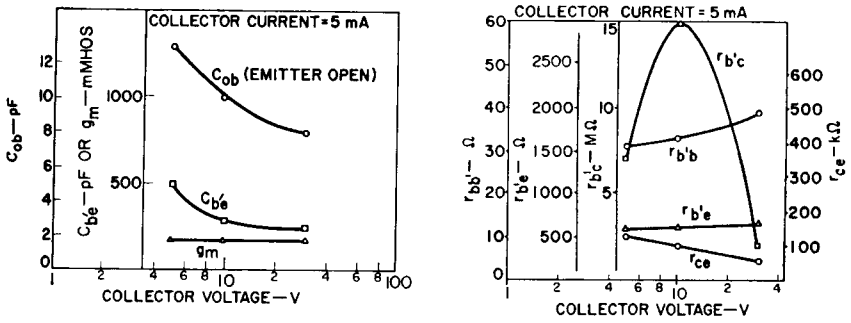



Figure 183. Hybrid-pi parameters as a function of collector voltage for the 2N2102 silicon power transistor.

junction, the excess minority carriers are swept into the collector region by the electric field that exists in the space-charge layer. The base current that flows during this transition consists primarily of majority carriers used for recombination in the base region and a small amount needed to supply the majority carriers injected or extracted across the two junctions. Both components of base current are small, however, so that the total base current is small in comparison to the emitter or collector current. This discussion suggests that the emitter and collector currents can be resolved into two independent components.

## Large-Signal Equivalent Circuits

For large-signal linear service, the nonlinear characteristics of a transistor can be represented in an equivalent-circuit model if it is assumed that the mechanisms which control the flow and distribution of carriers in the electrically neutral regions (outside of the space-charge layers) are linear; the only nonlinear properties then are those associated with the Boltzmann distribution of carriers which occurs at the edges of the space-charge region. This assumption is the basis for the Ebers and Moll equations which may be used to describe the large-signal properties of a transistor over the entire range of operation (i.e., from cutoff to saturation).

In the following analysis, the Ebers-Moll equations for a p-n-p transistor are derived. The terminal currents and voltages assumed for this derivation are shown in Fig. 184. The emitter and collector terminal currents can both be resolved into forward and reverse components. The forward components, which are controlled by the emitter-to-base junction voltage, are independent of the collector-to-base voltage and can be determined from the following equations:

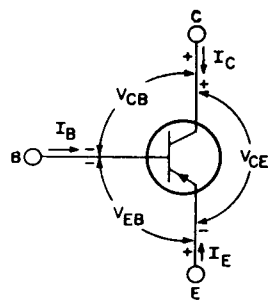



Figure 184. Terminal parameters for a p-n-p transistor.

$$I_{EF} = I_{ES}\,[\exp\,(qV_{EB}/kT) - 1] \quad (183)$$

$$I_{CF} = -\alpha_F I_{EF} \quad (184)$$

where the subscript F denotes forward operation, $\alpha_F$ is the common-base forward-current gain with a short-circuit output, and $I_{ES}$ is the equivalent emitter-diode saturation current.

In a similar manner, the reverse components of current, which are controlled by the collector-to-base voltage, are defined as follows:

$$I_{CR} = I_{CS}\,[\exp\,(qV_{CB}/kT) - 1] \quad (185)$$

$$I_{ER} = -\alpha_R I_{CR} \quad (186)$$

where the subscript R denotes reverse operation and the remaining symbols are obvious from previous definitions. The total emitter and collector currents are the sum of the forward and reverse components and may be expressed by the following equations:

$$I_E = I_F + I_{ER} = I_{ES}\,[\exp\,(qV_{EB}/kT) - 1] - \alpha_R I_{CS}\,[\exp\,(qV_{CB}/kT) - 1] \quad (187)$$

$$I_C = I_{CF} + I_{CR} = -\alpha_F I_{ES}\,[\exp\,(qV_{EB}/kT) - 1] + I_{CS}\,[\exp\,(qV_{CB}/kT) - 1] \quad (188)$$

These two equations, which are known as the Ebers and Moll equations, represent the static current-voltage characteristics of the idealized transistor model shown in Fig. 185. This model represents the processes of injection of carriers by the ideal diodes having the equivalent diode saturation currents $I_{ES}$ and $I_{CS}$; the processes of minority-carrier transport are associated with the two current-actuated current generators.

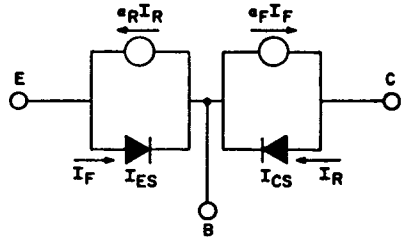



Figure 185. Ebers-Moll model of a p-n-p transistor controlled by diode currents.

By an identical process, it is possible to arrive at the Ebers-Moll equations for an n-p-n transistor. These equations may be expressed as follows:

$$I_E = I_{EF} + I_{ER} = -I_{ES}\,[\exp\,(-qV_{EB}/kT) - 1] + \alpha_R I_{CS}\,[\exp\,(-qV_{CB}/kT) - 1] \quad (189)$$

$$I_C = I_{CF} + I_{CR} = \alpha_F I_{ES}\,[\exp\,(-qV_{EB}/kT) - 1] - I_{CS}\,[\exp\,(-qV_{CB}/kT) - 1] \quad (190)$$

Figs. 186 and 187 show the Ebers-Moll models for the n-p-n transistor.

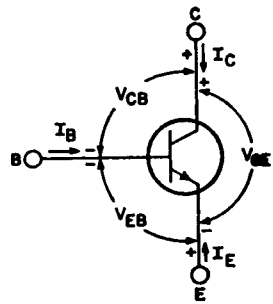



Figure 186. Terminal parameters for an n-p-n transistor.

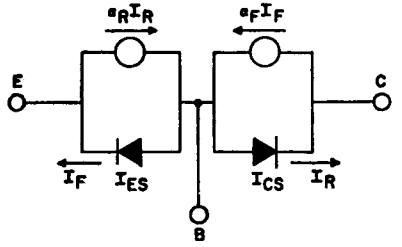



Figure 187. Ebers-Moll model of an n-p-n transistor controlled by diode currents.

## Large-Signal Characteristics

Fig. 188 shows the typical input characteristic (i.e., base current $I_B$ as a function of base-to-emitter voltage $V_{BE}$) for a 2N3053 silicon power transistor. From the hybrid-pi model, it is apparent

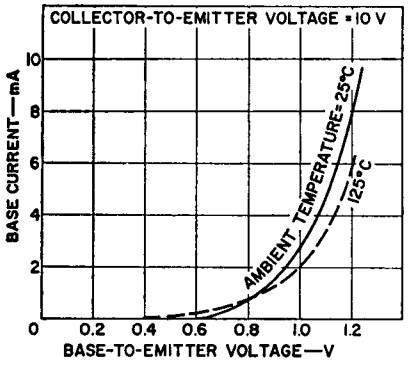



Figure 188. Typical input characteristics for the 2N3053 silicon power transistor.

that the resistance defined by this characteristic may be determined from the following calculation:

$$r_B = d V_{BE}/dI_B = r_{b'e} + r_{bb'} \qquad (191)$$

Eq. (191) may be rewritten as follows:

$$r_B = \frac{h_{fe}|_{f_{lo}}}{g_m} + r_{bb'} \qquad (192)$$

or may be expressed by

$$r_B = \frac{h_{fe}|_{f_{lo}}}{0.04|I_C|} + r_{bb'} \qquad (193)$$

For small values of emitter current, the base input resistance is determined primarily by the first term of the equations, and $r_{bb'}$ is negligible. At high currents, the opposite is true. At extremely high currents, the current crowding that occurs further increases $r_{bb'}$. The base input characteristic is then caused to "lean over" because of the large $I_b r_{bb'}$ voltage drop.

Fig. 189 shows typical output characteristics for a 2N3053 transistor. These characteristics display the static or dc value of

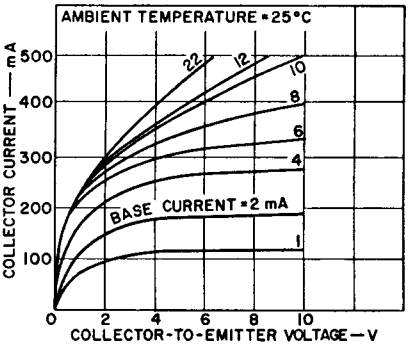



Figure 189. Typical output characteristics for the 2N3053 silicon power transistor.

$h_{FE}$, which is not the same as the low-frequency small-signal $h_{fe}|_{f_{lo}}$ used in previous calculations. The dc $h_{FE}$ is useful primarily for establishment of dc bias conditions.

Fig. 190 shows the dc $h_{FE}$ as a function of $I_C$ for a typical 2N3053 transistor. The shape of the curve may be explained by analysis of the following equations:

$$h_{FE} = \alpha/(1-\alpha) \qquad (194)$$

$$\alpha = \beta_o \gamma \mu \qquad (195)$$

where $\alpha$ is the $h_{FB}$ or common-base short-circuit current gain, $\beta_o$ is the transport factor, $\gamma$ is the

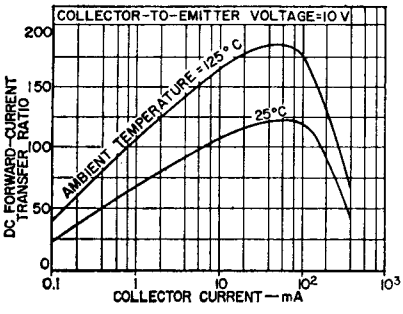



Figure 190. Typical dc forward-current transfer ratio ($h_{FE}$) as a function of collector current for the 2N3053 silicon power transistor.

emitter efficiency, and $\mu$ is the collector multiplication factor. For silicon transistors at low current, $\gamma$ is small because the recombination current in the emitter space-charge layer is high in comparison to the emitter current. As the current is increased, $\gamma$ increases because the recombination current remains constant. $\beta$ also increases because higher accelerating fields are produced as the recombination sites become filled. The combination of these two effects cause $h_{FE}$ to increase.

At high currents, $h_{FE}$ decreases because the increased carrier density near the emitter junction increases the rate of recombination. In addition, the flow of base current creates a radial transverse field which tends to forward-bias the emitter edge more than the center; this effect increases the current density and, as a result, decreases the emitter efficiency.

An increase in the collector-to-emitter voltage $V_{CE}$ at which $h_{FE}$ is measured results in an increase in $h_{FE}$. This increase results because of the reduced base width W produced by the increased $V_{CE}$. The reduced W increases the transport factor $\beta_0$ and, therefore, increases $h_{FE}$.

The effect of temperature on $h_{FE}$ is also shown in Fig. 190. This effect is the net sum of many individual temperature effects which are complex and are not necessarily consistent from one transistor to another. The best information of this type is empirical and is normally given in the transistor manufacturer's specifications.

The transconductance curve for a typical 2N3053 transistor is shown in Fig. 191. If leakage current is considered, the equation for $V_{BE}$ is as follows:

$$V_{BE} = (I_B - I_{CBO})\, r_{bb}' + V_{b'e} \quad (196)$$

$$= (I_C/h_{FE})\, r_{bb}' - I_{CBO}\, r_{bb}' + V_{b'e} \quad (197)$$

If variation of these parameters with temperature is taken into account, the following effects are noted: $r_{bb}'$ increases with temperature because of an increase in resistivity; $h_{FE}$ may increase or decrease with temperature depending upon the current level; $I_{CBO}$ increases with temperature; and $V_{b'e}$ decreases with temperature

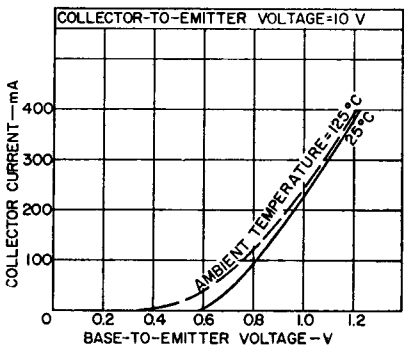



Figure 191. Typical transfer characteristics for the 2N3053 silicon power transistor.

because of the decrease in barrier height. On the basis of these effects, it can be predicted that, at low current levels, $V_{BE}$ decreases with increasing temperature and, at high current levels, the trend reverses, and $V_{BE}$ increases with temperature as shown in Fig. 192.

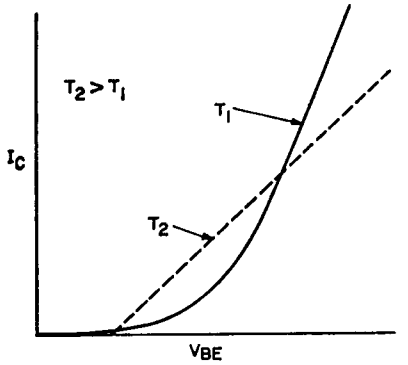



Figure 192. Typical transconductance characteristics for a silicon power transistor at high currents.

It should be noted that the input characteristic is highly nonlinear at low currents, but the output characteristic has lower $h_{FE}$ at such low currents. These two effects tend to be compensatory and help produce linear amplification in many cases.

## POWER TRANSISTORS IN SWITCHING SERVICE

An important application of power transistors is power switching. Large amounts of power, at high currents and voltages, can be switched with small losses by use of a power transistor that is alternatively driven from cutoff to saturation by means of a base control signal. The two most important considerations in such switching applications are the speed at which the transistor can change states between saturation and cutoff and the power dissipation.

### Switching Speed

Fig. 193 shows a typical test circuit used to measure the switch-

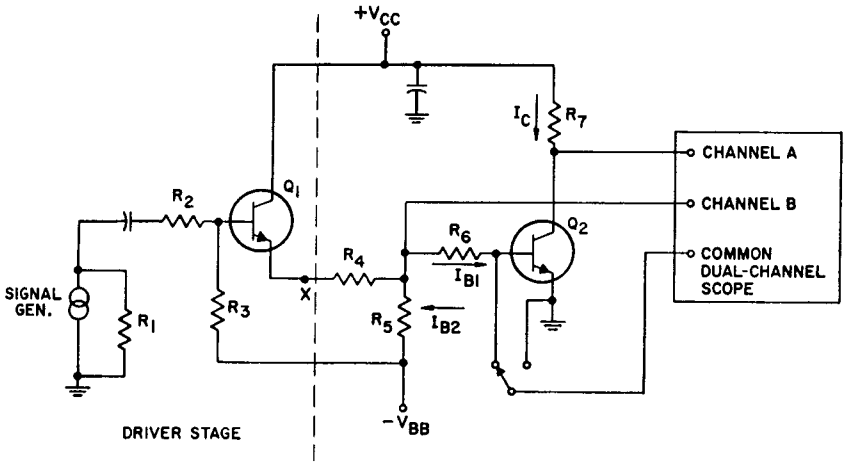



Figure 193. Test circuit used to measure switching times in power transistors.

ing times of a power transistor, and Fig. 194 shows the base and collector current waveforms that result from the changes in the conducting state of the transistor. In the test circuit, transistor $Q_2$ is the transistor on

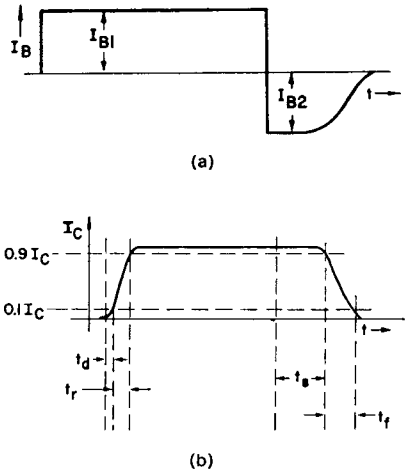



Figure 194. (a) Base-current waveform and (b) collector-current waveform that results from change in conducting state of a power transistor.

which the switching-time measurements are taken. Transistor $Q_1$, which is used in a driver stage, has a faster switching speed than transistor $Q_2$. This driver transistor amplifies the input drive signal from a conventional pulse generator. Resistors $R_1$, $R_2$, and $R_3$, together with the input impedance of transistor $Q_1$, combine to present a matched impedance to the pulse generator. Resistor $R_3$, which is usually much larger than resistors $R_1$ and $R_2$, merely provides reverse bias to increase the switching speed of transistor $Q_1$. Point X on the test-circuit schematic represents the origin of the input pulse to the transistor under test. This input consists of only a positive pulse of voltage. If a generator capable of supplying the required pulse is available, it could be attached to the test circuit directly at point X, and the driver stage could then be omitted.

The input pulse at point X, resistors $R_4$, $R_5$, and $R_6$, and the voltage $V_{BB}$ can be adjusted to provide the required turn-on current $I_{B1}$ and the required turn-off current $I_{B2}$. The value selected for $V_{BB}$, however, must be less than the collector-to-base breakdown voltage of the transistor to avoid breakdown of the emitter junction when the transistor is turned off. The value of resistor $R_7$ is chosen, together with $V_{CC}$, to provide the required $I_C$. Usually, switching times are given for $I_{B1} = I_{B2}$, and a condition of forced beta = 10 is most common.

When the common point of the oscilloscope is connected to the base of transistor $Q_2$, the turn-on current $I_{B1}$ and the turn-off current $I_{B2}$ can be sampled on channel B. The waveform appears as shown in Fig. 194(a). When the common point of the oscilloscope is connected to ground, the collector current $I_C$ can be viewed. Fig. 194(b) shows the collector waveform. (It is, of course, inverted.) The waveforms for $I_B$ and $I_C$ shown in Fig. 194 indicate that switching times consist of four components, as follows:

1. **Delay Time** ($t_d$) represents the time for the collector current to go from an off state to 10 per cent of its final on value after the turn-on base pulse is applied.

2. **Rise Time** ($t_r$) represents the time for the collector current to go from 10 per cent to its final on value to 90 per cent of its final on value.

3. **Storage Time** ($t_s$) represents the time for the collector current to go from its final on value to

90 per cent of its final on value after the turn-off base pulse is applied.

4. **Fall Time** ($t_f$) represents the time for the collector current to go from 90 per cent of its on value to 10 per cent of its on value.

**Turn-on time** $T_{on}$ and **turn-off time** $T_{off}$ are defined by the following equations in terms of the preceding four components:

$$T_{on} = t_d + t_r \tag{198}$$

$$T_{off} = t_s + t_f \tag{199}$$

If the correlation of switching times is to be accurate, exact circuits must be used. These circuits must include all voltages and all resistors. One reason for this requirement is obvious, if only $t_s$ and $t_f$ are considered. These switching times depend primarily on $I_{B2}$, but the shape of $I_{B2}$ depends upon how closely $I_{B2}$ approaches a current generator. As previously stated, the approximation of $I_{B2}$ as a current generator is limited by the $BV_{EBO}$ breakdown rating. Hence, a specification of $I_{B1}$ and $I_{B2}$ does not totally define the $I_{B2}$ source impedance and, therefore, does not guarantee equivalent switching times.

**Qualitative Description of Switching Times**—All switching times can be explained in terms of transistor physics; accurate estimates of switching times from parameter measurements, however, are more difficult.

Delay time $t_d$ arises from a stored charge at the emitter and collector junctions. The emitter is reverse-biased by $V_{BE}(off)$; as a result, charge is stored in the depletion layer around the emitter. Similarly, the collector is reverse-biased by $V_{CC} + V_{BE}(off)$ and has an associated capacitance. The depletion layers, and thus the capacitance, change when $V_{BE}(off)$ switches to $V_{BE}(on)$. The time required to effect this change in charge is termed $t_d$ and can be considered as the time required to charge the junction capacitances to new values.

The rise time $t_r$ involves the same capacitances discussed above, plus other factors. The junction capacitances are still changing because the collector voltage is decreasing and $V_{BE}(on)$ is increasing. The rise time is also affected by the base transit time and the charging of the collector capacitance through the collector series resistance.

The storage time $t_s$ depends upon the time required for minority carriers in the base and collector to recombine and produce a charge distribution that exists when the transistor is just ready to come out of saturation.

Fall time $t_f$ can be considered as the reverse process of rise time; thus, the charges in the emitter and collector junction capacitances are again important.

**Quantitative Relationships for Switching Times**—The concept of a transistor as a charge-controlled device is useful for prediction of switching-time phenomena. This concept views the transistor as a device in which the terminal currents ($I_C$, $I_B$, and $I_E$) are controlled by the charge in the base.

Transistor theory predicts that $I_C$, $I_B$, and $I_E$ are linearly related to the base charge. As a result, three separate time constants can be defined:

1. Emitter time constant,
$$\tau_E = (Q_B/I_E) \tag{200}$$

2. Base time constant,
$$\tau_B = (Q_B/I_B) \quad (201)$$

3. Collector time constant,
$$\tau_C = (Q_B/I_C) \quad (202)$$

The time constant $\tau_B$ represents the effective minority-carrier lifetime in the base. This time constant is related to the time constant $\tau_E$ as follows:

1. For a uniform base,
$$\tau_E = \tau_B (1-\alpha) \approx 1.2/\omega \quad (203)$$

2. For a graded base,
$$\tau_E = \tau_B (1-\alpha) \approx 0.6/\omega_b \quad (204)$$

where $\omega_b$ is the base cutoff frequency (i.e., the frequency at which the base transport factor is 0.707 of its original value).

The time constant $\tau_C$ can be defined in terms of $\tau_E$ by the following equation:

$$\tau_C = \tau_E/\alpha \quad (205)$$

The basic equation of charge continuity for transistors may be written as follows:

$$I_B = (dQ_B/dt) + (Q_B/\tau_B) \quad (206)$$

where $I_B$ is the base current, $dQ_B/dt$ is equal to the change in base charge, and $Q_B/\tau_B$ can be interpreted as the recombination rate.

Eq. (206) can be integrated with respect to time to obtain the following result:

$$\int_0^t I_B dt = \int_{Qo}^{Qt} dQ_B + \int_0^t (Q_B dt/\tau_B) \quad (207)$$

Eq. (207) states that the charge delivered to the base is equal to the change in the base charge necessary to establish a new current level, plus the charge needed to maintain $Q_B$ against recombination.

The total base-charge variation can be expressed by the following equation, which is derived from the charge-continuity equation [Eq. (206)]:

$$\int_0^t I_B dt = \int_{V_{BE1}}^{V_{BE2}} C_{Te}\, dV_{BE} + \int_{V_{CB1}}^{V_{BC2}} C_{Tc}\, dV_{CB} + \int_{Q_{B1}}^{Q_{B2}} dQ_B + \int_0^t (Q_B dt/\tau_B) \quad (208)$$

where $C_{Te}$ and $C_{Tc}$ are emitter and collector transition capacitances.

Because $C_{Te}$ and $C_{Tc}$ depend on $V_{BE}$ and $V_{CB}$, respectively, the following assumptions can be made:

$$C_{Te} = C'_{Te}\, V_{BE}^{-\frac{1}{2}} \quad (209)$$
(step-junction assumption)

$$C_{Tc} = C'_{Tc}\, V_{CB}^{-\frac{1}{2}} \quad (210)$$
(reasonable assumption for step or graded junction)

where the prime capacitances indicate a measurement of $C_{Te}$ and $C_{Tc}$ at a total voltage of 1 volt. (This total voltage includes the junction voltage and the junction contact potential.)

The notations in Eqs. (209) and (210) are used and the indicated integrations in Eq. (208) are performed to obtain the following expression:

$$t_d = (2/I_{B1}) \{C'_{Te} [V_{BE(off)}]^{\frac{1}{2}} + C'_{Tc} [(V_{CE} + V_{BE(off)})^{\frac{1}{2}} - V_{CC}^{\frac{1}{2}}]\} \quad (211)$$

The important conclusions that can be drawn from Eq. (211) are that $t_d$ increases with $V_{BE}(off)$, decreases with increased $I_{B1}$, and depends primarily on the emitter transition capacitance when $V_{CC} >> V_{BE}(off)$.

Rise time starts at the edge of conduction and ends just short of saturation. The charge equation is given by

$$I_B = \frac{dQ_{Te}}{dt} + \frac{dQ_{Tc}}{dt} + \frac{dQ_B}{dt} + \frac{Q_B}{\tau_B} \tag{212}$$

After a number of simplifying assumptions are made the following equation can be obtained:

$$t_r \approx h_{FE}\left(\frac{1}{\omega_T} + 1.7\, R_L\, C_{Tc}\right)$$

$$\ln \frac{h_{FE}\, I_{B1}}{h_{FE} I_{B1} - 0.9\, I_C} \tag{213}$$

where $\omega_T$ is the current-gain bandwidth at the edge of saturation. All terms, except $C_{Tc}$, are given at the edge of saturation. The equation indicates that rise time is reduced for small ratios of $I_C/I_B$. In addition, transistors that have a high $f_T$ (cutoff frequency) with $R_L$ as a load produce the lowest rise time.

Storage time is determined by the length of time required to remove excess base charge over that required to maintain $I_C$. Additionally, charge must be removed from the collector. The collector charge depends on the forward bias at the collector and on the collector resistivity.

The charge-control equation is as follows:

$$I_B = \frac{Q_B}{\tau_B} + \frac{Q_{BS}}{\tau_s} + \frac{dQ_B}{dt} + \frac{dQ_{BS}}{dt} \tag{214}$$

where $I_B$ is the base current, $Q_B$ is the base charge required to maintain $I_C$, $Q_{BS}$ is the excess stored base charge, $dQ_B/dt$ is the current that results from a change in $Q_B$, and $dQ_{BS}/dt$ is the current that results from a change in $Q_{BS}$.

If $Q_B$ is constant and is equal to $\tau_C I_C$, the derivative $dQ_B/dt$ is zero because $Q_B$ is the charge that is required to maintain $I_C$ and is not changed.

If $I_{B2}$ is substituted for $I_B$ in Eq. (211), the equation may be rewritten in the following form:

$$I_{B2} = \frac{\tau_C\, I_C}{\tau_B} + \frac{Q_{BS}}{\tau_S} + \frac{dQ_{BS}}{dt} \tag{215}$$

If Eq. (215) is integrated with respect to time, the following equation is obtained:

$$\int_{I_{B1}-I_{B2}}^{0} \frac{dI_B}{I_{B2} - I_C\,(\tau_C/\tau_B) - (Q_{BS}/\tau_S)} = \int_0^{t_s} \frac{dt}{(dQ_{BS}/dI_B)} \tag{216}$$

With the additional information that $\tau_C/\tau_B$ is equal to $1/h_{FE}$ and that $dQ_{BS}/dI_B$ is equal to $\tau_S$, the final expression for $t_s$ becomes

$$t_S = \tau_S \ln \frac{I_{B1} - I_{B2}}{(I_C/h_{FE}) - I_{B2}} \tag{217}$$

where $I_{B2}$ is negative.

The term $\tau_S$ in Eq. (217) is the storage time constant, and its value depends on the structure of the transistor. The equations used to determine $\tau_S$ for different transistor types are tabulated below:

1. For alloy-type transistors having no collector storage charge,

$$\tau_S = 2.4/\omega_b(1-\alpha) \tag{218}$$

where $\omega_b$ is the base cutoff frequency and $\alpha$ is the grounded-base current gain, both measured at the edge of saturation.

2. For diffused-base alloy transistors,

$$\tau_S = 3/\omega_b(1-\alpha) \quad (219)$$

3. For graded-base mesa or planar transistors,

$$\tau_S = (0.6/\omega_b) + (\tau_{mc}/2) \quad (220)$$

where $\tau_{mc}$ is the minority-carrier lifetime in the collector.

From the equation for $\tau_S$, it is apparent that an increase in $I_{B2}$ causes a decrease in $\tau_s$, an increase in $I_{B1}$ (more overdrive) causes an increase in $\tau_s$, and the limiting value for $t_s$ is approximately $0.7\tau_S$.

The equation for fall time $t_f$ can be derived in a manner similar to that used to derive the rise-time equation. The same charges must be moved, but the limits are different. The result is as follows:

$$t_f = h_{FE}\left(\frac{1}{\omega_T} + 1.7\ R_L\ C_{Tc}\right)$$

$$\ln\frac{I_C - h_{FE}\ I_{B2}}{0.1\ I_C - h_{FE}I_{B2}} \quad (221)$$

Because $I_{B2}$ is negative, Eq. (221) indicates that $t_f$ increases with $h_{FE}$ but decreases as the ratio $I_C/I_{B2}$ decreases.

## Switching-Time Reduction Techniques

From the previous discussion, it is obvious that a proper choice of the ratio $I_C/I_B$ can minimize switching time for a given transistor. In addition, some circuit techniques are available that further improve switching speed. Two common techniques for this purpose are the use of "collector-catcher" circuits and of speed-up capacitors.

A typical "collector-catcher" circuit is shown in Fig. 195. Improvement in switching speed is obtained because the transistor is not allowed to go into saturation; storage time, therefore, is drastically reduced. With no input pulse applied to the circuit, the transistor is initially biased off by $-V_{BB}$. A positive pulse of voltage turns the transistor on and the collector voltage begins to drop from $+V_{CC}$ toward $+V_R$. If $V_R$ is greater than the sum of the voltage drop across CR and $V_{CE}$ just out of saturation, then the collector is clamped at some value of voltage which maintains the transistor out of saturation.

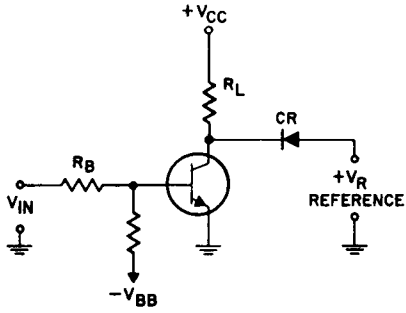



Figure 195. "Collector-catcher" circuit.

The difficulty with this circuit is that the maximum $I_C$ is essentially beta-dependent and, because beta varies with temperature, this circuit is not very stable. Burn-out of the diode or transistor is possible.

More practical circuits are shown in Figs. 196(a) and 196(b). In these circuits, the battery $V_B$ and the diode keep the

transistor out of saturation as before. However, the feedback arrangement tends to keep $I_C$ constant by automatic variation of the base drive. For example, if beta increases and $I_C$ tends to rise, the collector voltage begins to decrease, but the base drive is decreased and the circuit is returned to near its original condition. The only disadvantage of this circuit is the isolated battery $V_B$.

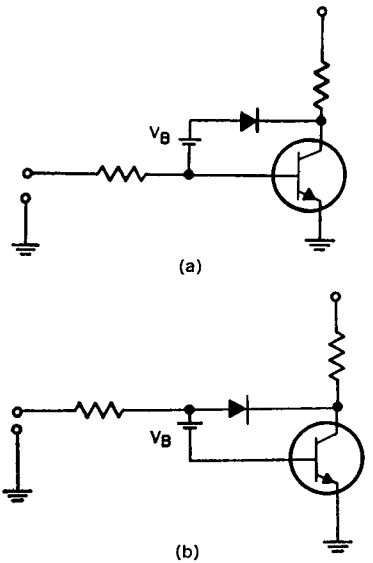



Figure 196. "Collector-catcher" circuits in which feedback is used.

A practical circuit is shown in Fig. 197. Operation is similar to that described above. The drop across $CR_2$ acts as the $V_B$ supply. The diode pair $CR_1$ constitutes the feedback arrangement. The function of $CR_3$ is to keep the transistor turned off under conditions of low $V_{BE}$ voltage (to reduce delay time). C acts as a speed-up capacitor. The use of the proper speed-up capacitor effectively increases turn-on drive and turn-off

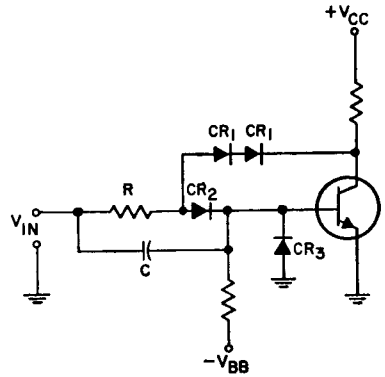



Figure 197. Practical "collector-catcher" circuit.

drive without supplying large amounts of "on" base current. As a result, faster rise times and faster fall times are achieved without the penalty of long storage time.

An example of a circuit that uses a speed-up capacitor is shown in Fig. 198. Waveforms for this circuit are shown in Fig. 199.

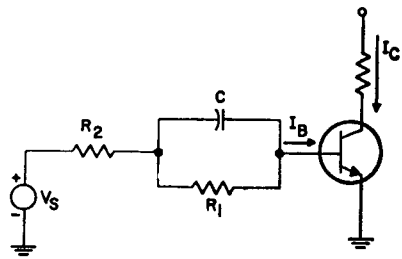



Figure 198. Circuit that uses speed-up capacitor to reduce transistor switching times.

The optimum value of C for fastest response can be found experimentally. If $V_S$ is large compared to $V_{BE}$ and $R_2$ can be neglected, then the charge stored on the capacitor while the transistor is ON is $V_C$. This charge should

equal the stored base charge for best response. Practical values for $R_2$, $V_S$, and $V_{BE}$ will modify this relation.

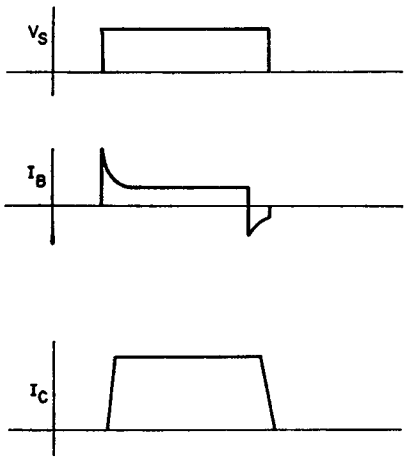



Figure 199. Waveforms for circuit shown in Fig. 198; (a) source voltage $V_S$; (b) base current $I_B$; (c) collector current $I_C$.

## Power Dissipation

An important consideration for transistors being used in a switching mode is power dissipation. Dissipation can be handled in four parts: (1) average dissipation; (2) saturation dissipation; (3) cutoff dissipation; (4) switching transient dissipation.

Average dissipation is the average power that must be handled over a complete cycle of operation and basically determines the heat-sink requirements for a given transistor case temperature. It is the average of saturation, turn-off transient, cutoff, and turn-on transient dissipations for the case where a transistor is continuously turning on and off. Under such conditions, the average power can be expressed by the following general equation:

$$P_{avg} = (1/T)\int_0^{t_1} V_{CE(sat)}\, I_{C(sat)}\,dt + (1/T)\int_{t_1}^{t_2} V_{CE}(t)_{off}\, I_C(t)_{off}\,dt + (1/T)\int_{t_2}^{t_3} V_{CE(off)}\, I_{C(off)}\,dt + (1/T)\int_{t_3}^{T} V_{CE}(t)_{on}\, I_C(t)_{on}\,dt \quad (222)$$

where $T$ = total switching period
$t_1$ = time interval the transistor is ON
$t_2$-$t_1$ = turn-off time
$t_3$-$t_2$ = time interval the transistor is OFF
$T$-$t_3$ = turn-on time
$V_{CE}(off)$ = $V_{CE}$ while transistor is OFF
$I_C(off)$ = current flowing while transistor is OFF [$I_{CEX}$ at $V_{CE}$ (off)]
$V_{CE}(t)on$ = $V_{CE}$ as a function of time while transistor is turning on
$I_C(t)on$ = $I_C$ as a function of time while transistor is turning on
$V_{CE}(sat)$ = $V_{CE}$ while transistor is ON
$I_C(sat)$ = $I_C$ while transistor is ON
$V_{CE}(t)off$ = $V_{CE}$ as a function of time while transistor is turning off
$I_C(t)off$ = $I_C$ as a function of time while transistor is turning off

Base power is usually very small and is neglected in the equation for average power. If required, however, an $I_B V_{BE}(t)$ term can be included for each interval of time. It should be noted that the four

integrals in Eq. (221) are, respectively: (1) saturation dissipation, (2) turn-off dissipation, (3) cut-off dissipation, and (4) turn-on dissipation.

Evaluation of the four integrals usually requires the use of simplifying assumptions or a graphical approach. If a particular circuit has been constructed and average power is to be calculated, two approaches are available:

(1) The instantaneous voltage-current characteristic can be plotted, and the integrations performed graphically.

(2) The heat-sink temperature can be measured under normal operation. A controlled source of power (for example, dc power) can then be applied to the transistor until the steady-state temperature of the heat sink is equal to the normal operating temperature. The amount of power necessary to establish this heat-sink temperature is the average power.

If average power is to be calculated for a tentative design, some simplifying assumptions must be made. For example, if a particular transistor is used to switch a resistive load with equal turn-on and turn-off base currents (the idealized waveforms are shown in Fig. 200), the following assumptions can be made:

1. base power is negligible,
2. turn-on and turn-off, switching times are equal,
3. collector voltage and current vary linearly with time during the switching transient.

On the basis of these assumptions, $P_{avg}$ can be determined as follows:

$$P_{avg} = \underbrace{\frac{1}{\tau}\int_0^{t_1} I_P\, V_{CE(sat)}\, dt}_{\text{ON}} + \underbrace{\frac{1}{T}\int_{t_1}^{t_2} I_p \left(1 - \frac{t}{T_{sw}}\right) V_P \left(\frac{t}{T_{sw}}\right) dt}_{\text{TURN-OFF}} + \underbrace{\frac{1}{T}\int_{t_2}^{t_3} I_{CEX}\, V_p dt}_{\text{OFF}} + \underbrace{\frac{1}{T}\int_{t_3}^{t_4} I_p \left(\frac{t}{T_{sw}}\right) V_P \left(1 - \frac{t}{T_{sw}}\right) dt}_{\text{TURN-ON}}$$

$$= \frac{I_P\, V_{CE(sat)}\, T_{on}}{\tau} + \frac{I_{CEX}\, V_P\, T_{off}}{\tau} + \frac{I_P V_P\, T_{sw}}{3\tau} \qquad (223)$$

From Eq. (223), it is obvious that the average power dissipated in the transistor can be reduced, and the efficiency can therefore be increased, by use of a transistor that has the following characteristics: low $V_{CE}(sat)$, low $I_{CBO}$, and fast switching characteristics (i.e., minimum $T_{SW}$).

## Load-Line Analysis

An analysis of the transistor load line is an important consideration in achievement of maximum reliability in a high-power switch. In general, the load is a combination of resistive and reactive elements. It is almost never purely resistive, and for "worst-case" analysis can be assumed to be purely inductive.

A simple test circuit for observation of a load line is shown in Fig. 201. The current-sensing resistor in the collector circuit should be non-inductive and should have a resistance value much

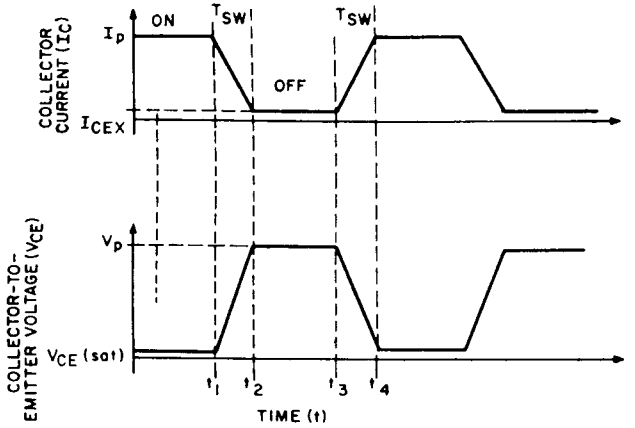



Figure 200. Idealized collector current and voltage waveforms for a transistor that has equal turn-on and turn-off currents.

smaller than any other impedance in series with the transistor. A typical load line ($V_{CE}$ as a function of $I_C$) for this circuit is shown in Fig. 202(a). If the inductance is reduced so that $L(di/dt) < V_{CEX(sus)}$, the circuit has a load line as illustrated in Fig. 202(b). [The $V_{CEX(sus)}$ curve in both cases should be determined under the bias conditions of the circuit.] However, in most applications, and certainly for the worst-case design, the load line of Fig. 202(a) is applicable.

Fig. 203 shows typical voltage and current curves as a function of time for this switch, with the load line indicated in Fig. 202(a). From these curves, the peak and average power dissipation, voltage limitations, and secondary-voltage-breakdown energy can be determined, as follows:

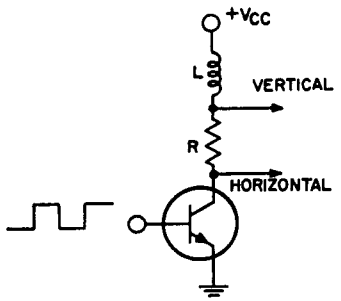



Figure 201. Test circuit used to determine transistor load lines.

$$P_{pk} = V_{CEX}\ (sus)\ I_{pk} \qquad (224)$$

$$P_{avg} = \underbrace{\frac{1}{\tau}\int_{}^{t_1} I_{CEX}\ V_{CC}\ dt}_{ON}$$

$$+ \underbrace{\frac{1}{\tau}\int_{t_1}^{t_2} \frac{I_{pk}}{2}\left(\frac{I_{pk}\ R_{sat}}{2}\right) dt}_{TURN\text{-}ON}$$

$$+ \underbrace{\frac{1}{\tau}\int_{t_2}^{t_3} I_{pk}\ (I_{pk}R_{sat})\ dt}_{OFF}$$

$$+ \underbrace{\frac{1}{\tau}\int_{t_3}^{t_4} \frac{I_{pk}}{2}\ (V_{CEX\ (sus)})\ dt}_{TURN\text{-}OFF}$$

$$\approx \frac{1}{\tau}\int_{t_3}^{t_4} \frac{I_{pk}}{2}\ (V_{CEX(sus)})\ dt \qquad (225)$$

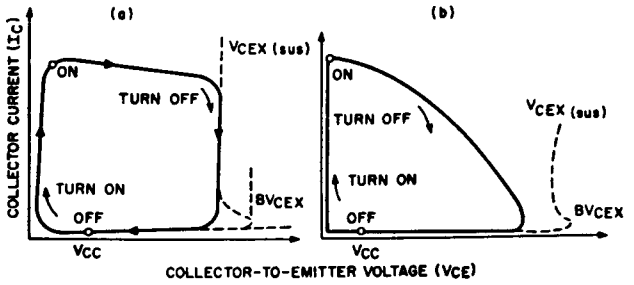



Figure 202. Inductive load line for the test circuit shown in Fig. 201; (a) typical load line; (b) load line when $L(di/dt) < V_{CEX(sus)}$.

Because $I_{CEX}$ and $R_{sat}$ are small, the following approximation is valid:

$$t_4 - t_3 = t_{off} = \int_{I_{CEX}}^{I_P} \frac{L\,di}{V_{CEX(sus)} - V_{CC}} \approx \frac{LI_P}{V_{CEX(sus)} - V_{CC}} \qquad (226)$$

The average power dissipated in the switch shown in Fig. 199 is then given by

$$P_{avg} = \frac{t_{off}}{\tau}\left(\frac{I_P}{2}\right) V_{CEX(sus)} = \frac{1}{2}\left(\frac{LI_P^2}{\tau}\right)\left[\frac{V_{CEX(sus)}}{V_{CEX(sus)} - V_{CC}}\right] \text{watts} \qquad (227)$$

The turn-off energy is expressed by the following equation:

$$Et_{off} = \frac{1}{2}(LI_p^2)\left[\frac{V_{CEX(sus)}}{V_{CEX(sus)} - V_{CC}}\right] \qquad (228)$$

This energy must not exceed the $E_{S/b}$ rating of the transistor.

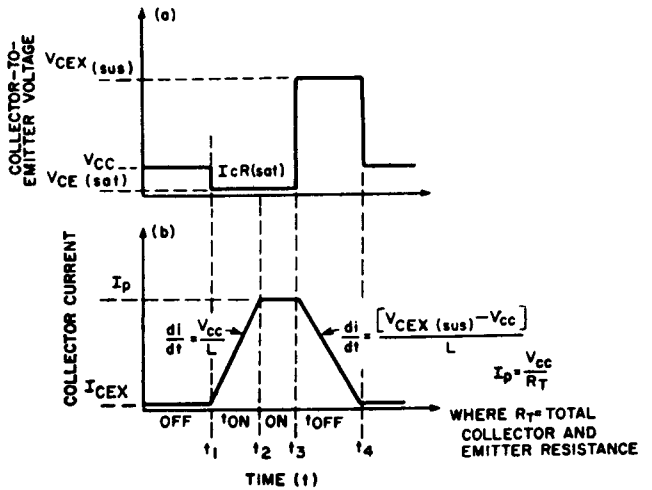



Figure 203. Voltage and current waveforms for the switching circuit shown in Fig. 201.

Preceding discussions have been directed toward the calculation of average power. Although this calculation is important, the peak power during switching is usually most critical. Whether the transistor can handle the peak power pulses during switching can be determined from safe-area-of-operation curves.

The general procedure for determination of the safe area of operation is as follows:

1. The voltage and current are plotted as a function of time.

2. An average junction temperature is calculated on the basis of average power and known case temperature.

3. The safe-area curves are derated on the basis of an effective cause temperature which is equal to true case temperature plus the average junction-to-case temperature.

4. The load line of voltage as a function of current is plotted on the derated safe-area curve, and it is determined that this load line does not remain in any area longer than the safe-area curve indicates.

5. If the voltage at any point reaches the sustaining region, all power dissipated in the sustaining region is considered as reverse-bias second-breakdown energy. This energy should not be allowed to exceed the $E_S/_b$ rating. (See section on **Safe-Area Ratings** for clarification of $E_S/_b$ and $I_S/_b$ safe-area curves.)

## Analysis of Inductive-Load Switching

Inductive switching requires rapid transfer of energy from the switched inductance to the switching mechanism, which may be a relay, a transistor, a commutating diode, or some other device. Often, it is necessary to calculate accurately the energy that will be dissipated in the switching device. This type of calculation is especially important when the switching element is a semiconductor device that may not be able to handle the amount of energy involved safely.

Most inductive switching circuits can be represented by the basic equivalent circuit shown in Fig. 204(a). This circuit shows the situation at the instant of turn-off. Variation of this basic circuit for three methods of turn-off are shown in Figs. 204(b), 204(c), and 204(d).

In the circuits of Fig. 204, $V_{CC}$ is the voltage source in series with the turn-off device and its inductive load, R and L are the series resistance and inductance, $I_o$ is the current flowing at the instant of turn-off, and $V_s$ represents the breakdown voltage of the turn-off device.

The energy dissipated in the turn-off device depends on $V_s$, $V_{CC}$, $I_o$, R, and L. For purposes of analysis, $V_s$ is assumed to remain constant during the turn-off transient [a reasonable assumption for the circuits of Figs. 204(b), 204(c), and 204(d)].

**Five Theoretical Cases**—The five cases to be analyzed are shown in Fig. 205. The energy equations for each case and corresponding current and voltage waveforms for the switching device are also shown.

**In case 1**, shown in Fig. 205(a), the energy is given by the famil-

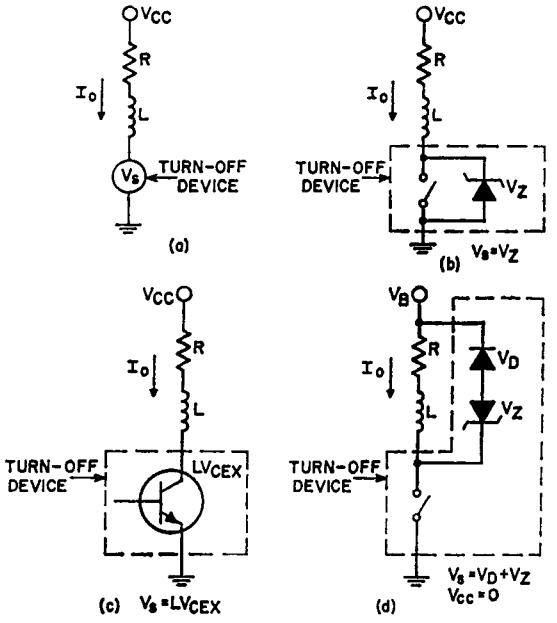



Figure 204. Basic switching circuits with inductive loads.

iar expression for energy stored in an inductor, as follows:

$$E_1 = (1/2)\, L\, I_0^2 \qquad (229)$$

where $E_1$ is the energy for case 1, L is the circuit inductance, and $I_0$ is the initial current.

**Case 2,** shown in Fig. 205(b), differs from case 1 in that energy is supplied to the switching device by the battery during turnoff. For case 2, the turnoff energy, $E_2$, is given by the following expression:

$$E_2 = (1/2)\, L\, I_0^2 \left(\frac{V_s}{V_s - V_{CC}}\right) \qquad (230)$$

To show that case 2 is a modification of case 1, a multiplying factor $k_2$ can be defined as follows:

$$E_2 = (1/2)\, L\, I_0^2\, k_2 \qquad (231)$$

where the factor $k_2$ is given by

$$k_2 = \frac{V_s}{V_s - V_{CC}} = \frac{1}{1 - \rho} \qquad (232)$$

Fig. 206 shows a curve of $k_2$ as a function of the ratio $V_{CC}/V_s$.

**Case 5,** shown in Fig. 205(e), causes an energy dissipation $E_3$ given by the following expression:

$$E_3 = I_0\, V_s \frac{L}{R} \left[\frac{\ln\left(\dfrac{1}{1 + (I_0 R)/(V_s - V_{CC})}\right)}{(I_0 R)/(V_s - V_{CC})} + 1\right] \qquad (234)$$

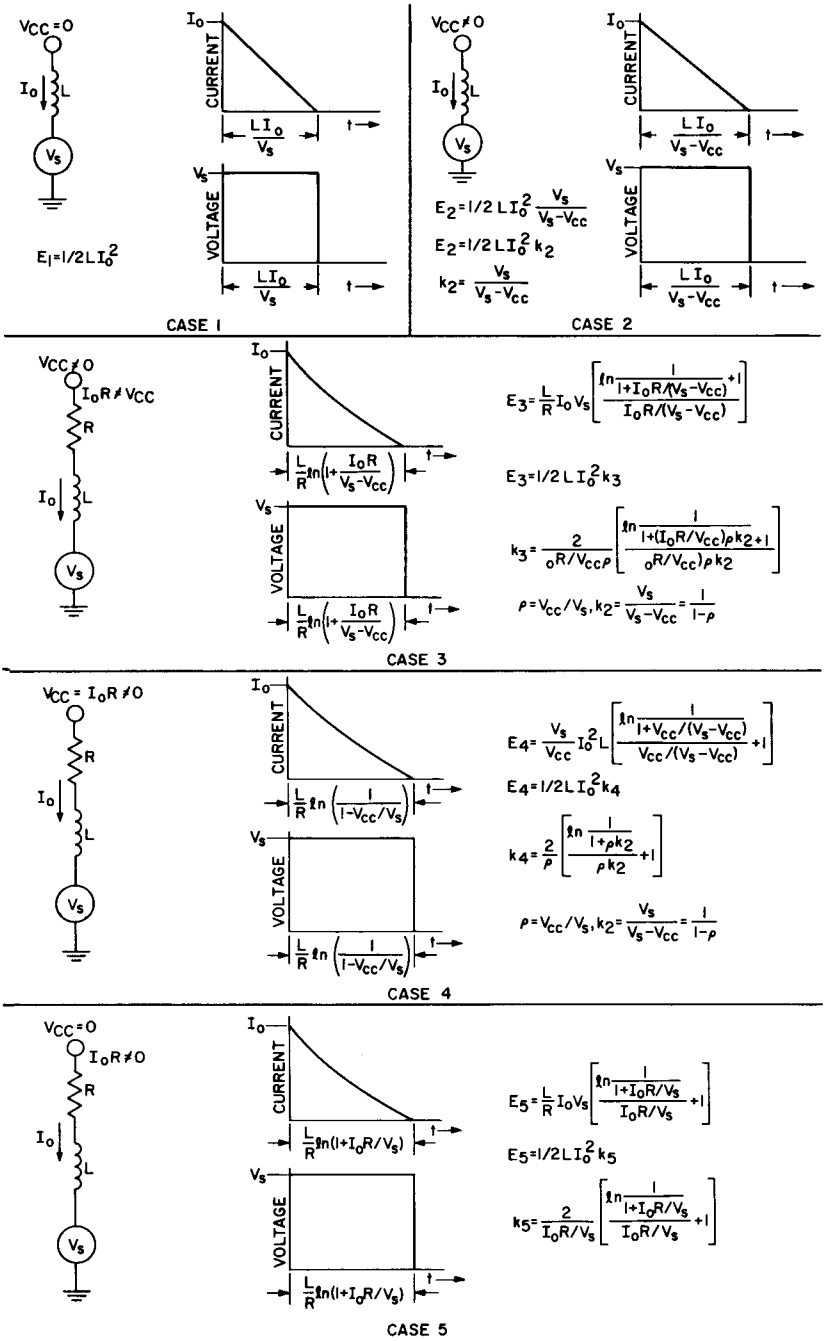



Figure 205. Equivalent circuits, energy equations, and voltage and current waveforms for the five basic circuit configurations.

Eq. (234) can be rearranged, as was Eq. (230), to show that it is a modification of case 1. This rearrangement introduces a new multiplying factor $k_3$, which is a function of both the voltage ratio $V_{CC}/V_s$ and the ratio $I_oR/V_{CC}$. Eq. 235 then becomes:

$$E_3 = (1/2)\ L\ I_0^2\ k_3 \qquad (235)$$

where the factor $k_3$ is defined as follows:

$$k_3 = f\ (I_0R/V_{CC}\ ,\ V_{CC}/V_s) = \frac{2}{(I_0R/V_{CC})_p}\left[\frac{\ln\left(\frac{1}{1+(I_0R/V_{CC})_p\ k_2}\right)}{(I_0R/V_{CC})_p\ k_2}+1\right] \qquad (236)$$

Fig. 206 shows curves of $k_3$ as a function of the ratio $V_{CC}/V_s$ for several values of $I_oR/V_{CC}$. When the ratio $I_oR/V_{CC}$ becomes less than 0.01, $k_3$ can be approximated by $k_2$. Case 3 then reduces to case 2.

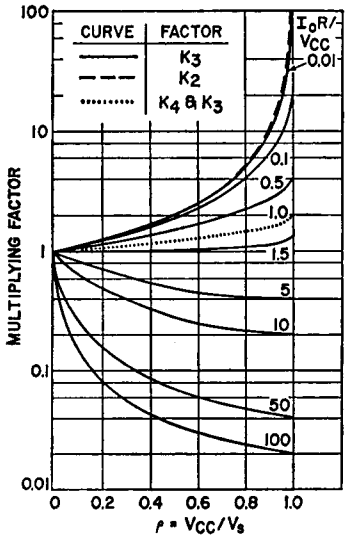



Figure 206. Multiplying factors ($K_2$, $K_3$, $K_4$) as a function of the ratio $V_{CC}/V_s$.

**Case 4**, shown in Fig. 205(d), is a special form of case 3. The condition $I_oR = V_{CC}$ arises when there is negligible voltage drop across the switching device, in the on state, and the current is limited only by the circuit resistance. The energy dissipation $E_4$ is then given by:

$$E_4 = \frac{V_s}{V_{CC}} I_0^2\ L \left[\frac{\ln\left(\frac{1}{1+V_{CC}/(V_s-V_{CC})}\right)}{V_{CC}/(V_s-V_{CC})}+1\right] \qquad (237)$$

Again, it is convenient to define a multiplying factor $k_4$ which is a function of the ratio $V_{CC}/V_s$, as follows:

$$E_4 = (1/2)\ L\ I_0^2\ k_4 \qquad (238)$$

where

$$k_4 = f\ (\rho) = \frac{2}{\rho}\left[\frac{\ln\frac{1}{(1+\rho\ k^2)}}{\rho\ k_2}+1\right] \qquad (239)$$

Fig. 206 includes a curve of $k_4$ plotted as a function of $\rho$. This curve corresponds to the curve of $k_3$ for $I_oR/V_{CC} = 1$. As $V_{CC}$ approaches $V_s$, the multiplying factor approaches its maximum value of two. Therefore, for any inductive switch in which $V_{CC} = I_oR$, the maximum energy that must be handled by the switching device is $L\ I_o^2$, or twice the energy of $(1/2)\ L\ I_o^2$ for case 1.

**Case 5**, shown in Fig. 205(e), can also be regarded as a special form of case 3. A practical example of case 5 is shown in Fig. 204(d). Note that there is no volt-

age source in series with the inductance and the commutating diodes; therefore $V_{CC} = 0$. For this case, the energy $E_5$ is given by

$$E_5 = V_s I_0 \frac{L}{R} \times \left[ \frac{\ln \frac{1}{1 + (I_0R/V_s)}}{(I_0R/V_s)} + 1 \right] \quad (240)$$

Because $V_{CC} = 0$, any $I_0R$ term gives an infinite $I_0R/V_{CC}$ ratio. Therefore, case 5 cannot be plotted by the method used in Fig. 204. However, a new multiplying factor $k_5$ can be defined in terms of $I_0R/V_s$ instead of $I_0R/V_{CC}$. When this new independent variable is used, Eq. 240 becomes

$$E_5 = (1/2)\ L\ I_0^2\ k_5 \quad (241)$$

where

$$k_5 = f\ (I_0R/V_s) = \frac{2}{(I_0R/V_s)} \left[ \frac{\ln \frac{1}{1 + (I_0R/V_s)}}{(I_0R/V_s)} + 1 \right] \quad (242)$$

Fig. 207 shows a curve of $k_5$ as a function of the ratio $I_0R/V_s$. (Because $k_5$ can be defined as a function of $I_0R/V_s$, obviously $k_5$ could also have been defined in terms of $\rho$ and $I_0R/V_s$, instead of $\rho$ and $I_0R/V_{CC}$. The choice was arbitrary, and was made on the basis that $I_0R/V_{CC}$ is usually a more meaningful ratio from a circuit standpoint.)

The preceding analyses show that energy dissipated in an inductive switch can, in general, be considered as a modification of the simple $(1/2)\ LI_0^2$ relationship. Each specific case requires a different multiplying factor k. The

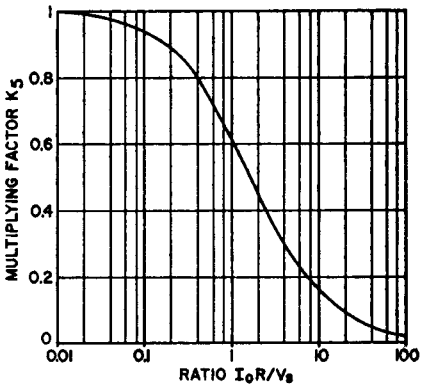



Figure 207. Multiplying factor $k_5$ as a function of the ratio $I_0R/V_s$.

constants $k_2$, $k_3$, $k_4$ and $k_5$ can be easily determined from Figs. 206 and 207 for most practical circuits.

**Four Practical Examples**—The following examples illustrate the use of the equations given for the five theoretical cases.

**Example 1:** The two circuits shown in Fig. 208 are identical except that $R_2$ in Fig. 208(b) is greater than $R_1$ in Fig. 208(a).

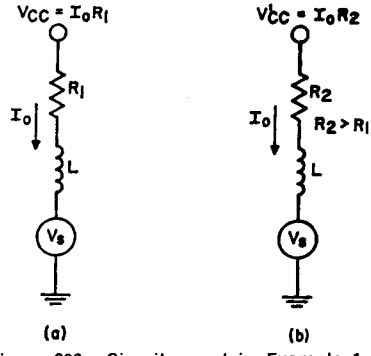



Figure 208. Circuits used in Example 1 to demonstrate the use of case-4 equations.

The problem is to decide in which circuit the switching device must handle the most energy. Because $V_{CC} = I_0R$, case 4 applies for both circuits. $I_0$ and $V_s$ remain

constant; therefore, $(1/2)\ L\ I_o^2$ is the same for both circuits.

To determine the switching energy, it is necessary first to determine how $k_4$ varies. Fig. 204 indicates that as $V_{CC}/V_s$ increases (i.e., as $V_{CC}$ increases), $k_4$ also increases. Because $E_4 = LI_o^2k_4/2$, an increase in $V_{CC}$ increases the energy requirement of the device. Thus, for a given circuit with a specified $I_o$, such as shown in Fig. 208, minimum energy is obtained with the lower resistance in series with the inductor (because a lower $V_{CC}$ is needed to establish $I_o$).

**Example 2:** The circuits shown in Fig. 209 demonstrate the difference between cases 3 and 4.

The circuit of Fig. 209(a) represents a turn-off condition after the circuit has reached steady state (because $I_oR = V_{CC}$). Therefore, this circuit corresponds to case 4.

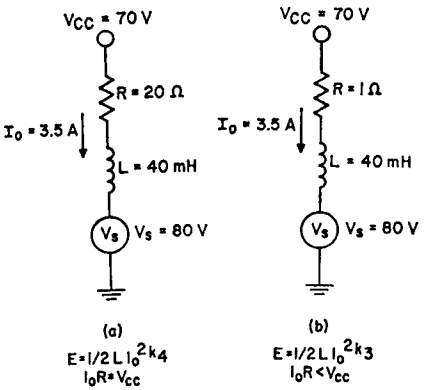



Figure 209. Circuits used in Example 2 to show the difference between cases 3 and 4.

The circuit of Fig. 209(b) represents a turn-off condition before the circuit has reached steady state (because $I_oR < V_{CC}$). Therefore, this circuit corresponds to case 3.

The energy requirement for the circuit of Fig. 209(a) is calculated as follows: The ratio $V_{CC}/V_s = 70/80 = 0.875$. The corresponding $k_4$ multiplying factor (from Fig. 206) is approximately 1.6. The energy $E_4$ is then given by

$$\begin{aligned} E_4 &= (1/2)\ L\ I_o^2\ k_4 \\ &= (1/2)\ (40 \times 10^{-3})\ (3.5)^2\ (1.6) \\ &= 400 \text{ millijoules} \end{aligned}$$

[Calculation of $E_4$ directly from Eq. (237) yields a value of 393 millijoules.]

To find the energy requirement for the circuit of Fig. 209(b) both the $V_{CC}/V_s$ ratio and the $I_oR/V_{CC}$ ratio must be used. The $V_{CC}/V_s$ ratio is again 0.875. The ratio $I_oR/V_{CC} = 3.5/70 = 0.05$. From Fig. 206, the interpolated value for $k_3$ is approximately 6.5. From Eq. (235), tht energy for this circuit is determined as follows:

$$\begin{aligned} E_3 &= (1/2)\ L\ I_o\ k_3 \\ &= 1/2\ (40 \times 10^{-3})\ (3.5)^2\ (6.5) \\ &= 1590 \text{ millijoules} \end{aligned}$$

[Calculation of $E_3$ directly from Eq. (234) yields a value of 1570 millijoules.]

It should be noted that the energy requirement for the circuit of Fig. 209(b) is considerably larger than $L\ I_o^2$. The expression $L\ I_o^2$ only represents the maximum that occurs when $I_oR = V_{CC}$ (as in case 4).

**Example 3:** The circuit shown in Fig. 210 is an example of case 5 because there is no battery voltage in series with L and the diodes. When the switch is opened, current circulates in the loop consisting of R, L, and the diodes, until the energy stored in L is dissipated.

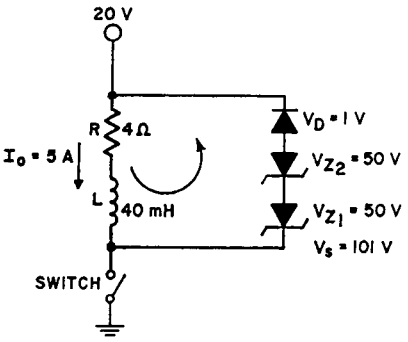



Figure 210. Circuits used in Example 3 to demonstrate the use of case-5 equations.

The problem is to find the energy dissipated in each of the diodes. It is assumed that the forward diode voltage $V_D$ remains constant at 1 volt during the switching transient.

The multiplying factor $k_5$ can be obtained from $I_0R/V_s$ and Fig. 207. The ratio $I_0R/V_s = 20/101$, or approximately 0.2; from Fig. 207, the corresponding value of $k_5$ is 0.88. Therefore, the energy dissipated by the diode combination is given by

$$E_5 = (1/2)\ L\ I_0^2 k_5$$
$$= 1/2\ (40 \times 10^{-3})\ (5)^2\ (0.88)$$
$$= 445 \text{ millijoules}$$

Because the same current flows through each of the diodes, the energy divides in proportion to the voltage drop. Therefore, the energy dissipated in each diode is given by

$$E_{Z1} = \frac{V_{Z1}}{V_{Z1} + V_{Z2} + V_D} \times E_5$$

$$= \frac{50}{50 + 50 + 1} \times 445 \text{ mJ} = 220 \text{ mJ}$$

$$E_{Z2} = \frac{V_{Z2}}{V_{Z1} + V_{Z2} + V_D} \times E_5$$

$$= \frac{50}{50 + 50 + 1} \times 445 \text{ mJ} = 220 \text{ mJ}$$

$$E_D = \frac{V_D}{V_{Z1} + V_{Z2} + V_D} \times E_5$$

$$= \frac{1}{50 + 50 + 1} \times 445 \text{ mJ} = 5 \text{ mJ}$$

**Example 4:** Fig. 211(a) shows a typical power-inverter circuit. The problem is to find the energy requirement for transistor $Q_1$. At the instant that $Q_1$ turns off, the circuit can be represented by the equivalent circuit shown in Fig. 211(b).

The following parameter values are assumed for this example:

$L_L$ = 100 microhenries
$I_0$ = 15 amperes
$R_s$ = 0.2 ohm
$V_E$ = 40 volts
$LV_{CEX}$ = 100 volts

The base-turnoff voltage $V_{BE(off)}$ and the base-circuit resistance $R_B$ establish the value of 100 volts for the breakdown voltage $LV_{CEX}$ of $Q_1$ (or $V_s$). The induced voltage from the $Q_2$ side of the inverter approaches $V_E$ and is assumed to be equal to $V_E$. $R_s$ represents the cumulative series resistance in the collector circuit, and $L_L$ represents all leakage and uncommutated inductance.

The circuit of Fig. 211(b) is similar to case 3. The equivalent circuit is shown in Fig. 211(c).

First the ratios $V_{CC}/V_s$ and $I_0R/V_{CC}$ are calculated as follows:

$$V_{CC}/V_s = 80/100 = 0.8$$

$$I_0R/V_{CC} = \frac{(15)\ (0.2)}{80}$$

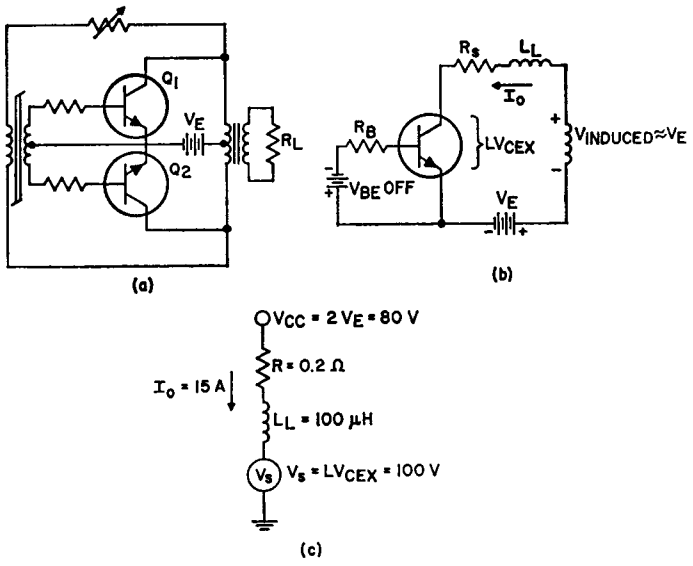



Figure 211. Typical power inverter circuit and equivalent circuits used in Example 4 as an application of case-3 equations.

$$I_0R/V_{CC} = \frac{3.0}{80} = 0.0375$$

Then interpolating from Fig. 206, the value of $K_3$, which is 4.5, is obtained. The energy $E_3$ for the circuit is then given by

$$\begin{aligned} E_3 &= (1/2)\ L\ I_0^2\ k_5 \\ &= (1/2)\ (100 \times 10\text{-}^6)\ (15)^2\ (4.5) \\ &= (50 \times 10\text{-}^6)\ (225)\ (4.5) \\ &= 51 \text{ millijoules} \end{aligned}$$

[Calculating $E_3$ directly from Eq. (234) gives 46 millijoules.]

The energy requirement for the transistor $Q_1$ in Fig. 211 (a), therefore, is almost five times the value that would be estimated by use of the simple energy relation (1/2) $L\ I_0^2$, and it has been determined that $Q_1$ must handle approximately 50 millijoules in the $LV_{CEX}$ mode for each cycle of inverter operation.

This type of calculation provides a conservative estimate of the $LV_{CEX}$-mode energy requirement, because the finite switching times of $Q_1$ allow some of the energy to be dissipated in the active region.

# Thyristors

THYRISTORS are solid-state devices that have characteristics similar to those of thyratron tubes; more specifically, they are solid-state switches whose bistable state depends on the regenerative feedback associated with a p-n-p-n structure. Basically, this group includes any bistable semiconductor device that has three or more junctions (i.e., four or more semiconductor layers) and can be switched from a high-impendance (off) state to a conducting (on) state, and from the conducting (on) state to the high-impedance (off) state, within at least one quadrant of the principal-voltage characteristics.

There are several types of thyristors, which differ primarily in number of electrode terminals and operating characteristics in the third (negative) quadrant of the voltage-current characteristics, as shown in Table XVIII. Reverse-blocking triode thyristors, commonly called **silicon controlled rectifiers** (SCR's), and bidirectional triode thyristors, referred to as **triacs**, are the most popular types.

Table XVIII—Different Types of Thyristors

| No. of Terminals | Third-Quadrant Operation | | |
|---|---|---|---|
| | **Blocking** | **Conducting** | **Switching** |
| 2 | Reverse-blocking diode thyristor | Reverse-conducting diode thyristor | Bidirectional diode thyristor |
| 3 | Reverse-blocking triode thyristor | Reverse-conducting triode thyristor | Bidirectional triode thyristor |

## THEORY OF OPERATION

SCR's and triacs have unique characteristics and capabilities that make them particularly useful for power switching and control applications in which low cost, small package size, device reliability, and circuit simplicity are important requirements. These devices differ primarily in that the SCR is a unidirectional device that is used for both dc and ac functions and the triac is a bidirectional device that is used mainly for ac functions. The following analyses of the switching transitions that occur in these devices provide useful informa-

tion concerning the operation and possible applications of SCR's and triacs.

## Equivalent-Model Analysis

Just as a transistor may be considered as basically a solid-state diode with a third semiconductor layer added to form two back-to-back diode junctions, the SCR may be considered as a transistor with an additional semiconductor region, and the triac as an SCR with one more semiconductor region. Simple models of the "lower order" devices may be analyzed, therefore, to show the effect of the additional semiconductor regions on the operation of the devices.

**Two-Transistor Analogy of an SCR**—An SCR is basically a four-layer p-n-p-n unidirectional device designed to provide bistable switching when operated in the forward-bias mode. The device has three electrodes, referred to as the **cathode**, the **anode**, and the **gate**. The gate is the control electrode for the device. For forward-bias operation, the anode potential must be positive with respect to the cathode. During normal operation, the SCR is turned on by application of a positive voltage to the gate electrode. The SCR then remains on, even though the gate voltage is removed or made negative, until the anode-to-cathode voltage is reduced to a value below that required to sustain regeneration, or forward current. Faster turn-off can be achieved by a reversal of the forward-current flow.

As shown in Fig. 212, the basic p-n-p-n SCR structure is analogous to a pair of complementary n-p-n and p-n-p bipolar transistors. Fig. 212(a) shows the schematic symbols for an SCR and equivalent connection of the complementary pair of transistors, and Fig. 212(b) shows the equivalent relationship of the p-n-p-n SCR structure and the interconnected transistor structures. The n-p-n and p-n-p transistors in the equivalent model are interconnected so that regenerative action occurs when a proper gating signal is applied to the base of the n-p-n transistor.

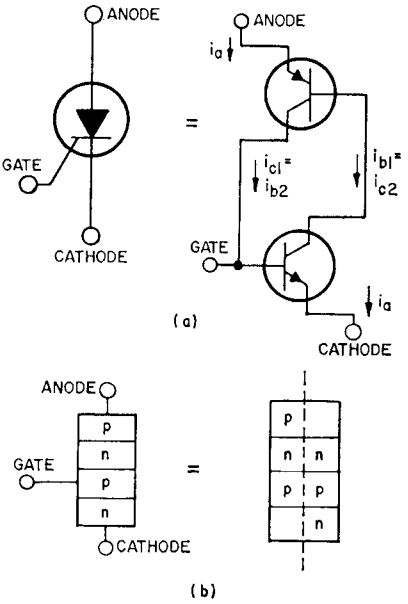



Figure 212. Two-transistor analogy of an SCR: (a) schematic symbols of an SCR and the equivalent two-transistor model; (b) structure of an SCR and of the equivalent two-transistor model.

When the two-transistor model is connected in a circuit to simulate normal SCR operation, the emitter of the p-n-p transistor is returned to the positive terminal of a dc supply through a limiting resistor $R_2$, and the emitter of the n-p-n transistor $Q_2$ is returned to the negative terminal

of the dc supply to provide a complete electrical path, as shown in Fig. 213. When the model is in the off state, the initial value of principal-current flow is zero. If a positive pulse is then applied to the base of the

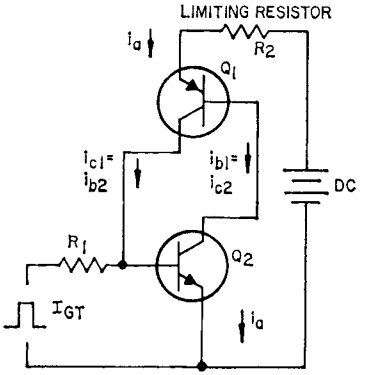



Figure 213. Two-transistor model connected to show a complete electrical path.

n-p-n transistor, the transistor turns on and forces the collector (which is also the base of the n-p-n transistor) to a low potential; as a result, a current $I_a$ begins to flow. Because the p-n-p transistor $Q_1$ is then in the active state, its collector current flows into the base of the n-p-n transistor ($I_{c1} = I_{b2}$) and sets up the conditions for regeneration. If the external gate drive is removed, the model remains in the on state as a result of the division of currents associated with the two transistors, provided that sufficient principal current ($I_a$) is available.

Theoretically, the model shown in Fig. 213 remains in the on state until the principal current flow is reduced to zero. Actually, turn-off occurs at some value of current greater than zero. This effect can be explained by observation of the division of currents as the value of the limiting resistor is gradually increased. As the principal current is gradually reduced to the zero current level, the division of currents within the model can no longer sustain the required regeneration, and the model reverts to the blocking state.

The two-transistor model illustrates three features of thyristors: (1) a gate trigger current is required to initiate regeneration, (2) a minimum principal current (referred to as "latching current") must be available to sustain regeneration, and (3) reduction of principal-current flow results in turn-off at some level of current flow (referred to as "holding current") that is slightly greater than zero.

Fig. 214 shows the effects of a resistive termination at the base of the n-p-n transistor on the latching and holding currents. The collector current through the p-n-p transistor must be increased to supply both the base current for the n-p-n transistor and the shunt current through the terminating resistor. Because the principal-current flow must

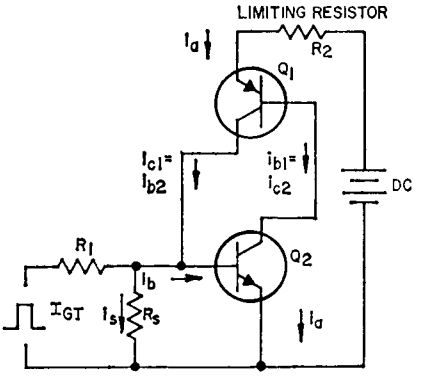



Figure 214. Two-transistor model of SCR with resistive termination of the n-p-n transistor base.

be increased to supply this increased collector current, latching- and holding current requirements also increase. The use of the two-transistor model provides a more concise meaning to the mechanics of thyristors. In thyristor fabrication, it is generally good practice to use a low-beta p-n-p unit and to include internal resistance termination for the base of the n-p-n unit. Termination of the n-p-n provides immunity from "false" (non-gated) turn-on, and the use of the low-beta p-n-p units permits a wider base region to be used to support the high voltage encountered in thyristor applications.

**Two-SCR Analogy of a Triac**— A triac is a bidirectional device designed to provide bilateral switching characteristics for either polarity of applied voltage. The three electrodes of this n-p-n-p-n device are referred to as **main terminal No. 1, main terminal No. 2,** and the **gate.** The gate is specially designed so that either positive or negative gate voltage can trigger the triac into conduction for either polarity of the voltage across the main terminals. As with the SCR, however, once the triac is turned on, the gate has no further control. The device remains in the on state until the voltage across the main terminals is reduced below the value required to sustain conduction. Unlike the SCR, however, the triac cannot be turned off by a reversal of the polarity of the voltage across the main terminals. A reversal of this voltage merely causes current to flow in the opposite direction.

Functionally, a triac may be considered as two parallel SCR (p-n-p-n) structures oriented in opposite directions, as shown in Fig. 215. The same approach used to explain gating, latching, and holding currents in the SCR can be extended to include the two-SCR model of a triac.

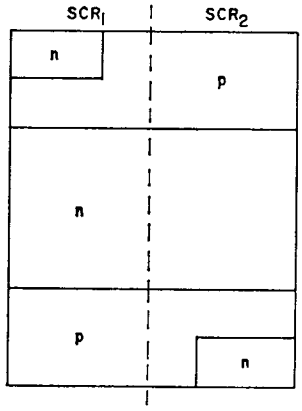



Figure 215. Diagram of a triac structure which shows that this device is basically two SCR's structures in an inverse parallel arrangement.

In triacs, the gate-trigger-pulse polarity is usually measured with respect to main terminal No. 1, which is comparable to the cathode terminal of an SCR. The triac can be triggered by a gate-trigger pulse which is either positive or negative with respect to main terminal No. 1 when main terminal No. 2 is either positive or negative with respect to main terminal No. 1. The triac, therefore, can be triggered in any of four operating modes, as summarized in Table XIX. The quadrant designations refer to the operating quadrant on the principal voltage-current characteristics (either I or III), and the polarity symbol represents the gate-to-main-terminal-No. 1 voltage. Fig. 216 shows the flow of current in a triac for each of the four triggering modes.

Table XIX—Triac Triggering Modes

| Main-Terminal-No.2-to-Main-Terminal-No.1 Voltage | Gate-to-Main-Terminal-No.1 Voltage | Operating Quadrant* |
|---|---|---|
| Positive | Positive | I (+) |
| Positive | Negative | I (−) |
| Negative | Positive | III (+) |
| Negative | Negative | III (−) |

* Positive (+) and negative (−) signs indicate polarity of gate trigger pulse.

The gate-trigger requirements of the triac are different in each operating mode. The I(+) mode (gate positive with respect to main terminal No. 1 and main terminal No. 2 positive with respect to main terminal No. 1), which is comparable to equivalent SCR operation, is usually the most sensitive. The smallest gate current is required to trigger the triac in this mode. The other three operating modes require slightly higher gate-trigger currents. For RCA triacs, the maximum trigger-current rating in the published data is the largest value of gate current that is required to trigger the selected device in any operating mode.

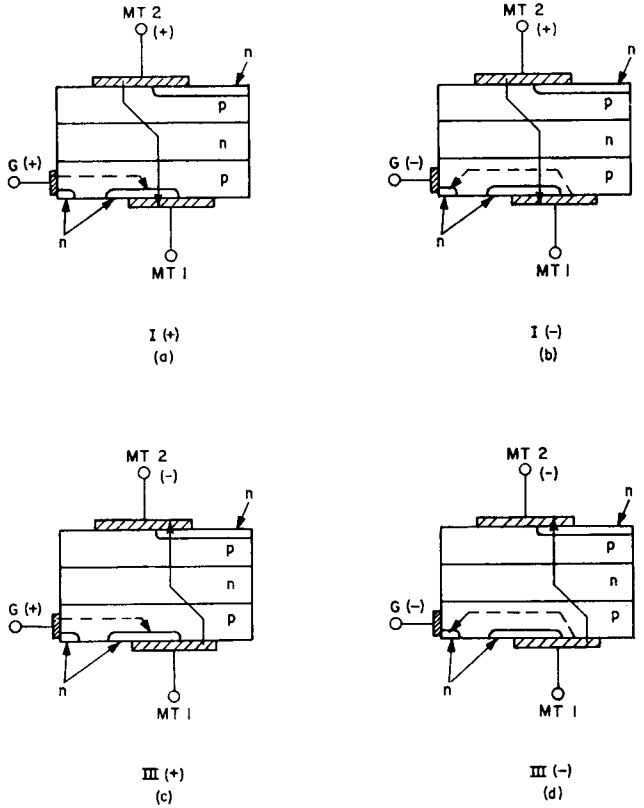



Figure 216. Current flow in the four triggering modes of a triac: (a) Mode I(+); (b) Mode I(−); (c) Mode III(+); (d) Mode III(−).

## Potential-Energy Analysis

The bistable action of thyristors can be explained by analysis of the potential-energy conditions in an SCR structure for each switching transition. This analysis can be directly related to either operating quadrant of a triac because, as mentioned previously, the triac is essentially equivalent to two SCR's in an inverse parallel connection.

The electron-hole interactions that make possible the switching transitions in p-n-p-n semiconductor structures are represented graphically by the potential-energy diagrams shown in Figs. 217, 218, and 219. These diagrams show the potential energies of holes and electrons as a function of distance through the crystal. The upward direction indicates increasing levels of electron energy, and the downward direction indicates increasing levels of hole energy. The dots in the diagrams represent free electrons, and the circles represent free holes.

**Basic Energy States**—The electrons in a solid can occupy only specific energy levels or electron states. Each existing state can be occupied by only one electron. The **Fermi energy level $E_F$** is the dividing line above which most of the existing electron states are empty and below which most states are full. Conduction in a solid occurs only by movement of free charge carriers, i.e., free electrons or free holes. A free electron is an electron which is at an energy level for which most of the existing states are empty, and a free hole is an empty state at an energy level for which most of the existing states are filled. Free electrons exist therefore, only at energy levels above $E_F$, and free holes exist only at levels below $E_F$. Because electrons and holes tend to seek the lowest available energy levels, they both move toward $E_F$, which is the zero-energy level for both types of charge carriers. On the potential-hill diagrams, electrons always tend to "fall", and holes always tend to "rise". If the charge carriers were not affected by outside influences, therefore, all free electrons and holes would eventually reach the Fermi level and disappear. A distribution of free electrons above $E_F$, however, is maintained by the thermal energy of the lattice which constantly agitates the electrons to non-zero energy levels.

In the metal contact regions, there is a continuous distribution of electron states about the Fermi energy level so that free holes and free electrons exist simultaneously side by side. In the semiconductor regions, there is a band of energy, called the **forbidden-energy region,** in which no electron states exist. As a free carrier tries to move through the system of metal to semiconductor to metal, it finds that it can move freely through the metal, but when it reaches the semiconductor it encounters an obstacle, the forbidden-energy region, which it must go over or under depending upon whether it is a free electron or a free hole. The carrier must obtain sufficient energy so that it is displaced far enough from the Fermi level to go over or under the forbidden-energy region. If sufficient energy, such as thermal agitation or an applied voltage, is not available, the carrier is reflected back to its origin.

In silicon crystal, the forbidden

region is wide enough so that, at ordinary temperatures, there is not sufficient thermal energy available to distribute carriers both above and below the band. If the Fermi energy level is close to the top of the band, thermal energy is sufficient to lift electrons into states on top of the forbidden region, but is not sufficient to push holes into states below this region. As a result, the material contains many free electrons, but very few free holes, and is referred to as an n-type semiconductor because it contains mostly negative-charge carriers.

Similarly, a p-type region in the semiconductor, which contains mostly positive-charge carriers, results when the Fermi level is close to the bottom of the forbidden band. For this condition, thermal energy is sufficient to excite holes into states below the band, but is not sufficient to excite electrons into states above the band.

The position of the Fermi level in the forbidden region is determined by the carrier concentration. This concentration, in turn, is determined by both the dopant concentration and the concentration of injected carriers.

At the metal-to-semiconductor interfaces, the dopant concentration is very high. At such interfaces, the carrier concentration cannot be changed significantly by injected carriers, and the position of the Fermi level in the forbidden region is firmly fixed. In the inner semiconductor regions and near the junctions, the dopant concentration is relatively low so that the total carrier concentration and, therefore, the position of the Fermi level in the forbidden region can be changed by injection of carriers from surrounding regions. These factors make the forbidden-energy region appear flexible within the body of the semiconductor but rigid at the metal-to-semiconductor contacts. This rigidity of the potential hill at the contacts prevents electrons in the metal from entering the p-type semiconductor, but allows holes to circulate freely between the metal and the p-type region; similarly, electrons can circulate freely between metal and the n-type region but holes cannot cross the metal-to-n-type-semiconductor interface.

**Forward-Blocking State**—The sequence of diagrams in Fig. 217 illustrates the transition of the thyristor from the equilibrium (zero-bias) condition to the forward-blocking state. In the equilibrium condition, the concentration of charge carriers (electrons and holes) is determined primarily by dopant concentrations. For this condition, which is represented by the potential-hill diagram shown in Fig. 217(b), there is approximately one free carrier for each dopant atom.

When the cathode side of the thyristor is biased negatively with respect to the anode side, the potential energy of the electrons is increased in the cathode region and that of the holes is increased in the anode region. Because of the difference in energy level from cathode to anode, the shape of the forbidden-energy region is altered in the most lightly doped section (i.e., the n-type base) so that the height of the potential hill of the central junction is increased. As shown in Fig. 217(c), any electrons that exist in this region "fall down" the resultant hill, and any holes in this region "rise" to the top of the hill. In

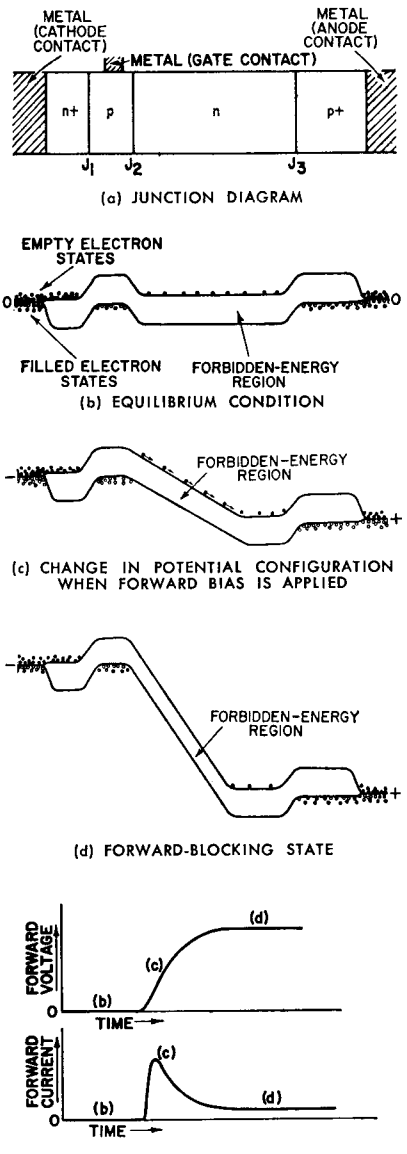


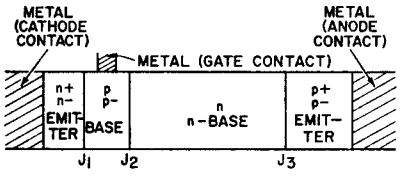

Figure 217. Potential-hill diagrams for various stages of thyristor transition from equilibrium condition to forward-blocking condition (electron energy increases upward, hole energy increases downward).

this way, all free charge carriers are removed, and the hill becomes a depletion region, as shown in Fig. 217(d).

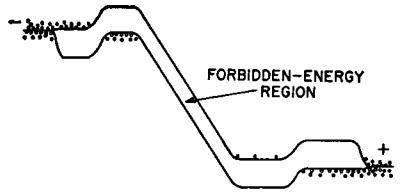
The movement of charge carriers with an increase in the forward voltage results in a charging, or displacement, current similar to the current ($i = Cdv/dt$) that charges a capacitor. This displacement current ceases when the forward voltage reaches a steady value because there are no additional carriers for the field to move. Although there are many electrons available on the cathode side of the thyristor and many holes available on the anode side, these carriers cannot enter the depletion region because they do not have sufficient energy to "climb" the 0.8- to 1.0-volt potential hills at junctions $J_1$ and $J_3$.

The current and voltage waveforms during the transition from the equilibrium to the forward-blocking state are shown in Fig. 217(e).

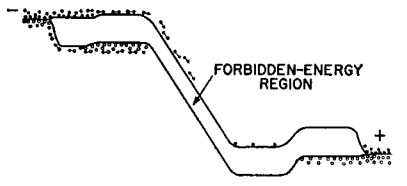
**Forward-Conducting State**—The transition in a thyristor from the forward-blocking state to the forward-conducting state is illustrated by the potential-hill diagrams shown in Fig. 218. When a thyristor is in the forward-blocking state, shown in Fig. 218(b), application of a positive bias to the gate causes the potential energy of electrons in this region to be reduced so that the height of the potential hill at junction $J_1$ is decreased, as shown in Fig. 30(c). A positive gate bias of 0.8 to 1.0 volt reduces the barrier of $J_1$ sufficiently so that electrons from the n-type emitter can move across the p-type base into the depletion region. The electric field then sweeps them across this region, as indicated in Fig. 218(c).



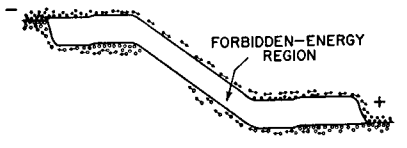

(a) JUNCTION DIAGRAM




(b) FORWARD-BLOCKING STATE

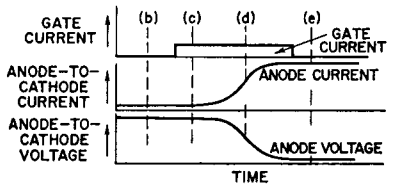



(c) CHANGE IN POTENTIAL CONFIGURATION WHEN POSITIVE BIAS IS APPLIED TO GATE




(d) POTENTIAL CONFIGURATION AS FORWARD BIAS BEGINS TO INCREASE


(e) FORWARD-CONDUCTING STATE




(f) CURRENT AND VOLTAGE WAVEFORMS DURING TURN-ON TRANSITION

Figure 218. Potential-hill diagrams for various stages of thyristor transition from forward-blocking state to forward-conducting state.

Electrons accumulate in the "well" at the bottom of the depletion region until their combined negative charge increases the potential electron energy sufficiently to cause the potential hill at junction $J_3$ to disappear. Holes can then move from the p-type emitter across the n-type base into the depletion region. These holes then immediately "climb" the potential hill at $J_2$, as shown in Fig. 218(d).

The increased supply of holes to the p-type base further depresses the potential hill at $J_1$ so that the n-type emitter can inject an even greater number of electrons into the depletion layer. This action, in turn, increases the injection of holes from the p-type emitter. As a result of these regenerative effects, the current through the thyristor increases rapidly, and the depletion region collapses to complete the transition to the forward-conducting state. Fig. 218(e) illustrates this condition. In this state, the concentrations of both holes and electrons are greatly increased over the equilibrium concentrations. The thyristor can be sustained in the forward-conducting state by an anode-to-cathode forward-voltage drop of approximately 1 volt, and the thyristor current is limited only by the impedance of the external circuit.

The current and voltage waveforms during the transition from the forward-blocking to the forward-conducting state are shown in Fig. 218(f).

**Turn-Off**—The transition in the thyristor from the forward-conducting state back to the forward-blocking state is illustrated in Fig. 219. This transition is accomplished either by momentary reduction of the anode current to

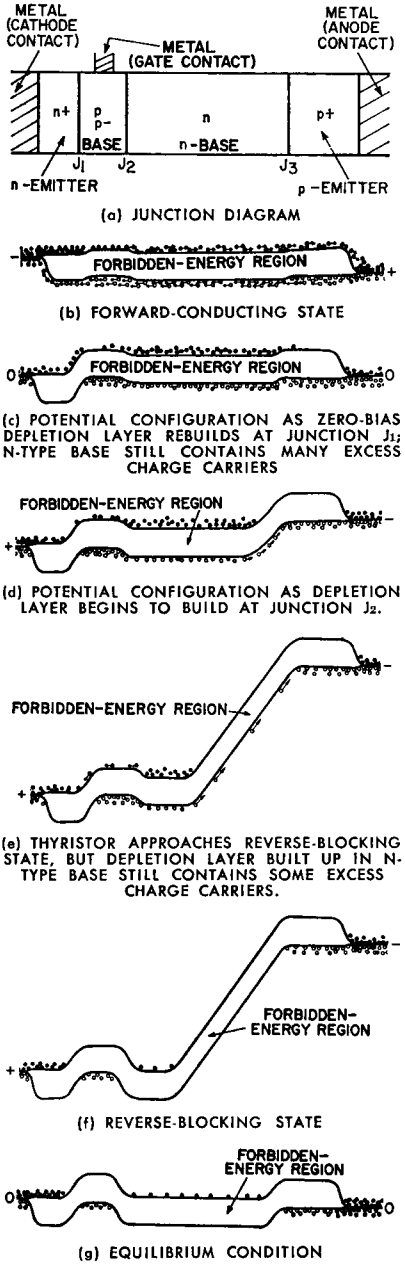



Figure 219. Potential-hill diagrams for various stages of thyristor transition from forward-conducting state to turn-off.

zero, or by momentary reversal of the anode-to-cathode voltage.

In the conducting state, carrier concentrations far in excess of the equilibrium level are injected into the n- and p-type regions. These excess carriers remain for a finite time after the anode current is reduced to zero. If the forward bias is re-applied before these excess carriers are removed, the device simply returns to the conducting state and does not switch to the blocking condition. After the excess carriers are removed and the device is returned to equilibrium, the potential hills rebuild, and the device can return to the forward-blocking state, as shown in Fig. 219.

The removal of excess carriers can be accomplished if the anode current is reduced to zero until the excess carriers recombine or move out of the depletion region. This removal corresponds to a direct transition from the conditions shown in Fig. 219(c) to those shown in Fig. 219(g). The potential hill at junction $J_1$ rebuilds first because it is in the more heavily doped region of the device, but the hills at $J_2$ and $J_3$ also rebuild as the excess carriers disappear during the zero-anode-current condition.

For SCR's, a more rapid removal of the excess carriers can be accomplished by a momentary reversal of the anode-to-cathode voltage. This transition is shown in Figs. 219(d) through 219(f). As the reverse voltage increases, carriers are pulled out of the device in the direction opposite to that in which they were injected so that a substantial reverse current results.

The removal of carriers is aided as a potential hill and a depletion region begin to build at junction

$J_3$, as shown in Fig. 219(d). As the remaining quantity of excess carriers is reduced, the reverse current decreases, and reverse voltage builds up. At the stage shown in Fig. 219(e), the reverse depletion region has built up, but the undepleted n-type base region still contains some excess carriers which prevent the potential hill at $J_2$ from rebuilding, and which continue to flow out as reverse current. At the stage shown in Fig. 219(f), the excess carriers have all been removed, and device has reached its steady-state reverse-blocking condition. In Fig. 219(g), the reverse bias has been removed, all regions return to the equilibrium zero-bias carrier concentrations, and the device is ready for return to the forward-blocking condition.

Current and voltage waveforms corresponding to the various condictions described in Figs. 217, 218, and 219 are shown in Fig. 220.

## Voltage-Current Characteristics

The principal voltage-current characteristics of SCR's and triacs indicate that these devices are ideal for power switching applications. When the voltage across the main terminals of either type of thyristor is below the breakover point, the current through the device is extremely small, and the thyristor is effectively an open switch. When the voltage across the main terminals increases to a value exceeding the breakover point, the thyristor switches to its high-conduction state and is effectively a closed switch. The thyristor remains in the on state until the current through the main terminals drops below a value which is called the **holding current.** When the source voltage of the main-terminal circuit cannot support a current equal to the holding current, the thyristor reverts back to the high-impedance off state.

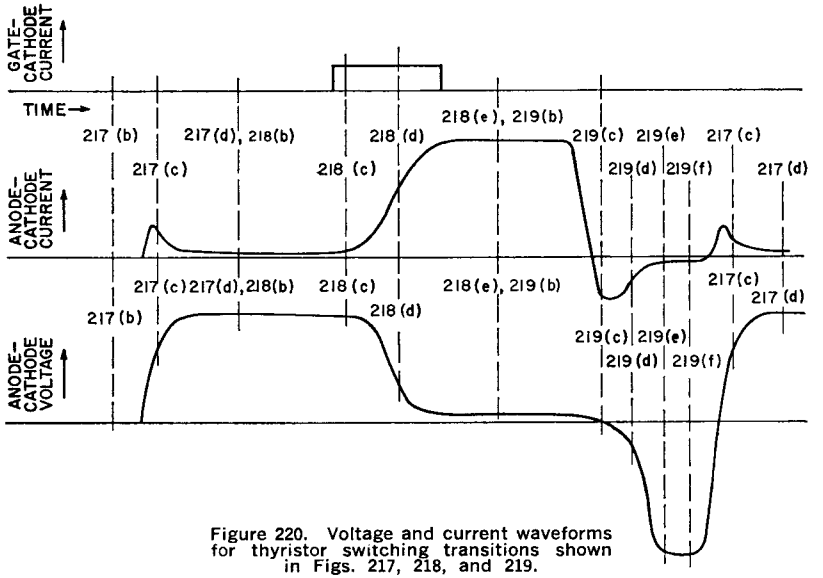



Figure 220. Voltage and current waveforms for thyristor switching transitions shown in Figs. 217, 218, and 219.

**SCR Characteristic**—Fig. 221 shows the principle voltage-current characteristic curve for an SCR. This curve shows that the operation of an SCR under reverse-bias conditions (anode negative with respect to cathode) is very similar to that of reverse-biased silicon rectifiers or other solid-state diodes. In this bias mode, the SCR exhibits a very high internal impedance, and only a slight amount of reverse current, called the **reverse blocking current,** flows through the p-n-p-n structure. This current is very small until the reverse voltage exceeds the reverse breakdown voltage; beyond this point, however, the reverse current increases rapidly. The value of the reverse breakdown voltage differs for individual SCR types.

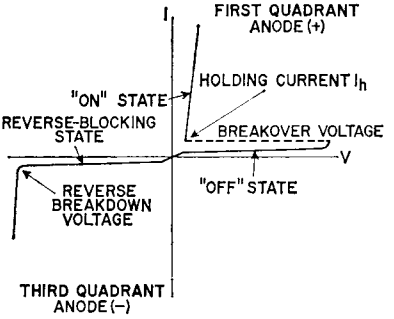



Figure 221. Principal voltage-current characteristic for an SCR.

During forward-bias operation (anode positive with respect to cathode), the p-n-p-n structure of the SCR is electrically bistable and may exhibit either a very high impedance (forward-blocking or off state) or a very low impedance (forward-conducting or on state). In the forward-blocking state, a small forward current, called the forward on-state current, flows through the SCR. The magnitude of this current is approximately the same as that of the reverse-blocking current that flows under reverse-bias conditions. As the forward bias is increased, a voltage point is reached at which the forward current increases rapidly, and the SCR switches to the on state. This value of voltage is called the **forward breakover voltage.**

When the forward voltage exceeds the breakover value, the voltage drop across the SCR abruptly decreases to a very low value, referred to as the **forward on-state voltage.** When an SCR is in the on state, the forward current is limited primarily by the impedance of the external circuit. Increases in forward current are accompanied by only slight increases in forward voltage when the SCR is in the state of high forward conduction.

**Triac Characteristic**—A triac exhibits the forward-blocking, forward-conducting voltage-current characteristic of a p-n-p-n structure for either direction of applied voltage, as shown in Fig. 222. This bidirectional switching capability results because, as mentioned previously, a triac consists essentially of two p-n-p-n devices of opposite orientation built into the same crystal. The

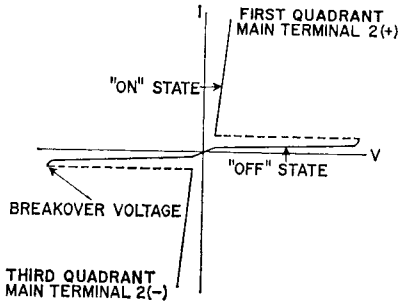



Figure 222. Principal voltage-current characteristic for a triac.

device, therefore, operates basically as two SCR's connected in parallel, but with the anode and cathode of one SCR connected to the cathode and anode, respectively, of the other SCR. As a result, the operating characteristics of the triac in the first and third quadrants of the voltage-current characteristics are the same, except for the direction of current flow and applied voltage. The triac characteristics in these quadrants are essentially identical to those of an SCR operated in the first quadrant. For the triac, however, the high-impedance state in the third quadrant is referred to as the off state rather than as the reverse-blocking state. Because of the symmetrical construction of the triac, the terms forward and reverse are not used in reference to this device.

## CONSTRUCTION

Construction details for typical RCA thyristors are shown in Figs. 223 through 226. Fig. 223 shows details for the 2-lead TO-5 package. This compact package is designed for applications in which mounting space is limited and can be attached to a wide variety of heat sinks with sizes and shapes to fit the available space. (Various types of thyristor heat-sink arrangements are described in the section on **Packaging, Handling, and Mounting** given earlier in this Handbook.) This package is used at current levels up to 7 amperes.

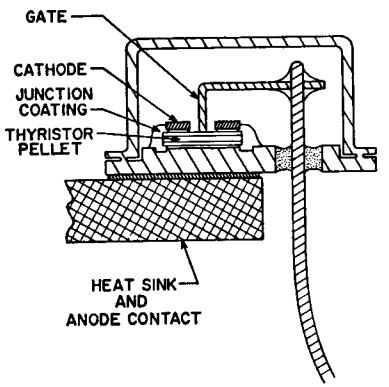



Figure 223. Cross-section of RCA two-lead TO-5 thyristor package.

In higher-current applications the TO-66, TO-3, and press-fit and stud-mounted TO-48 packages are used. Internal construction details of the press-fit package are shown in Fig. 224.

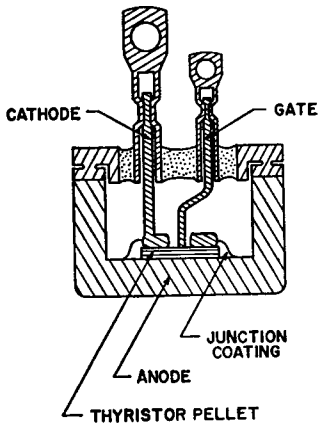



Figure 224. Cross-section of RCA press-fit thyristor package.

Construction details of a typical SCR pellet are shown in Fig. 225. The shorted-emitter construction used in RCA SCR's can be recognized by the metallic cathode electrode in direct contact with the p-type base layer around the periphery of the pellet. The gate, at the center of the pellet, also makes direct metallic contact to the p-type base so that the portion of this layer under the n-type emitter acts as an ohmic path for current flow between gate and cathode.

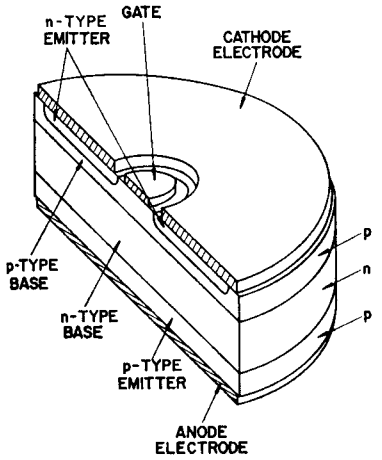



Figure 225. Cross-section of a typical SCR pellet.

Because this ohmic path is in parallel with the n-type emitter junction, current preferentially takes the ohmic path until the IR drop in this path reaches the junction threshold voltage of about 0.8 volt. When the gate voltage exceeds this value, the junction current increases rapidly, and injection of electrons by the n-type emitter reaches a level high enough to turn on the device.

In addition to providing a precisely controlled gate current, the shorted-emitter construction also improves the high-temperature and dv/dt (maximum allowable rate of rise of off-state voltage) capabilities of the device. The junction depletion layer acts as a parallel-plate capacitor which must be charged when blocking voltage is applied. Because the charging, or displacement, current ($i = Cdv/dt$) into this capacitor varies as the rate of rise of forward voltage (dv/dt), a very high dv/dt can result in a high current between anode and cathode. If this current crosses the n-type emitter junction and is of the same order of magnitude as the gate current, it can trigger the device into the conducting state. Such unwanted triggering is minimized by the shorted-emitter construction because the peripheral contact of the p-type base to the cathode electrode provides a large-area parallel path by which the dv/dt current can reach the cathode electrode without crossing the n-type emitter junction.

The center-gate construction of the SCR pellet provides fast turn-on and high di/dt capabilities. In an SCR, conduction is initiated in the cathode region immediately adjacent to the gate contact and must then propagate to the more remote regions of the cathode. Switching losses are influenced by the rate of propagation of conduction and the distance conduction must propagate from the gate. With a central gate, all regions of the cathode are in close proximity to the initially conducting region so that propagation distance is significantly decreased; as a result, switching losses are minimized.

Construction of a typical RCA triac pellet is shown in Fig. 226. In this device, the main-terminal-No. 1 electrode makes ohmic contact to a p-type emitter as well as to an n-type emitter. Similarly, the main-terminal-No. 2 electrode also makes ohmic contact to both types of emitter, but the p-type emitter of the main-terminal-No. 2 side is located opposite the n-type emitter of the main-terminal-No. 1 side, and the main-terminal-No. 2 n-type emitter is opposite the main-terminal-No. 1 p-type emitter. The net result is two four-layer switches in parallel, but oriented in opposite directions, in one silicon pellet. This type of

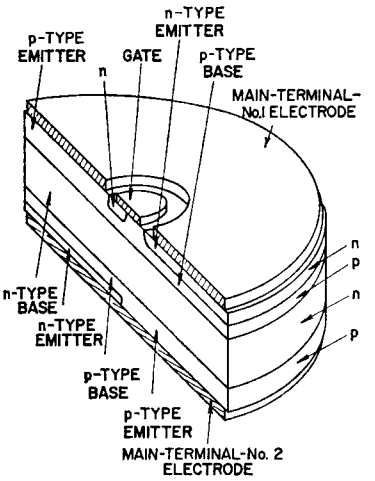



Figure 226. Cross-section of a typical triac pellet.

construction makes it possible for a triac either to block or to conduct current in either direction between main terminal No. 1 and main terminal No. 2.

The gate electrode also makes contact to both n- and p-type regions. As a result, the device can be triggered by either positive or negative gate signals, for either polarity of voltage between the main-terminal electrodes. When the triac is triggered by a positive gate signal, conduction is initiated, as in the SCR, by injection of electrons from the main-terminal-No. 1 n-type emitter, and the gate n-type region is passive. The gate n-type region becomes active when the triac is triggered by a negative gate signal, because it then acts as the n-type emitter of a grounded-base n-p-n transistor. Electrons injected from this region enter the n-type base and cause a forward bias on one of the p-type emitters, depending on which is at the positive end of the voltage between the main-terminal electrodes.

As shown in Figs. 225 and 226, the cathode of an SCR and the main terminal No. 1 of a triac are fully covered by a relatively heavy metallic electrode. This electrode provides a low-resistance path to distribute current evenly over the cathode or main-terminal-No. 1 area and serves as a thermal capacitor to absorb heat generated by high surge or overload currents. Junction-temperature excursions that result from such conditions are, therefore, held to a minimum.

## RATINGS AND LIMITING CHARACTERISTICS

Thyristors must be operated within the maximum ratings specified by the manufacturer to assure best results in terms of performance, life, and reliability. These ratings define limiting values, determined on the basis of extensive tests, that represent the best judgment of the manufacturer of the safe operating capability of the device. The manufacturer also specifies a number of device characteristics, which are directly measurable properties that define the inherent qualities and traits of the thyristor. Some of these characteristics are important factors in the determination of the maximum ratings and in the prediction of the performance, life, and reliability that the thyristor can provide in a given application.

### Off-State Voltage Ratings

All thyristor structures consist of one relatively wide, lightly doped base region between two more heavily doped regions of opposite impurity type. This lightly doped base region sup-

ports the main blocking voltage in both directions. The choice of design parameters (width and doping) for this base has the most fundamental effect on all thyristor electrical properties because blocking voltages, on-state voltage, power dissipation, and switching speeds all stem from the design of this region.

Voltage breakdown is the most important design criterion for a thyristor. Voltage initiated turn-on can occur as a result of avalanche breakdown or voltage punch-through. (An excessive rate of rise of off-state voltage can also initiate thyristor turn-on, as explained subsequently in the discussion on **Critical Rate of Rise of Off-State Voltage.**)

**Avalanche breakdown** occurs when the electric field in the depletion region reaches the critical field at which carriers traveling in the field gain sufficient energy between collisions to generate additional carriers when collision occurs. Avalanche voltage increases with lighter doping. Fig. 227(a) illustrates the avalanche-breakdown condition.

Fig. 227(b) provides a diagrammatic representation of **voltage punch-through.** Punch-through may occur with an excess blocking voltage of a polarity for which turn-on is possible. The blocking voltage causes the depletion region to spread to such an extent that it encompasses the p-type emitter. When this spreading occurs, a free flow of holes into the depletion region is possible, and turn-on results. In contrast to avalanche breakdown, punch-through voltage is increased by heavier doping and wider base widths.

The conflicting demands on impurity concentration imposed by voltage-breakdown and punch-through considerations suggests that some optimum range of base width and doping should be selected for any specific design-voltage objective. Fig. 228 shows a basic design chart that may be used to determine this optimum range. In this chart, the impurity concentration is shown on the horizontal axis, and the breakdown voltage is shown on the vertical axis. Optimum ranges of material resistivity and base width are shown for various punch-through values. The maximum voltage for any range chosen is limited by avalanche considerations. The chart shows that avalanche breakdown voltage increases with a reduction in impurity concentration, but that punch-through voltage is decreased for this condition.

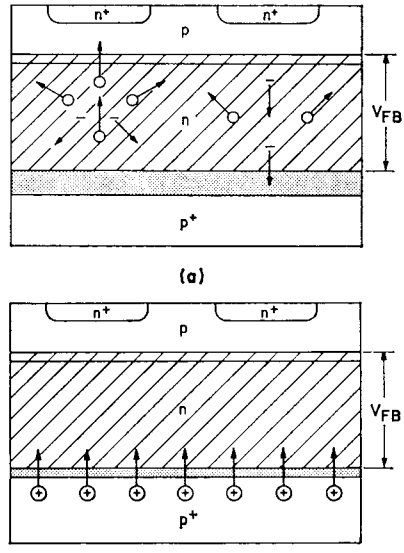



Figure 227. (a) Avalanche-breakdown and (b) voltage punch-through conditions.

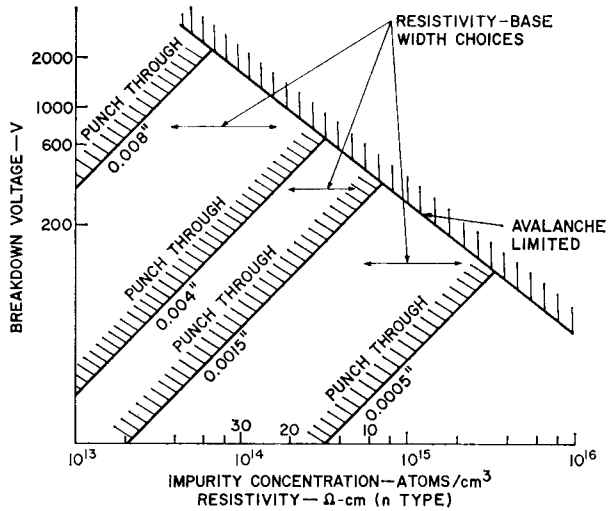



Figure 228. Basic design chart used to determine optimum range of base width and doping.

The important point illustrated by the chart is that for a practical range of base-material resistivity, a specific base width is required for each voltage class. Thus, a small SCR intended for logic application below 200 volts might have an acceptable base width of 1.5 mils and a base-material resistivity in the neighborhood of 10 ohm-centimeters. A 1000-volt device would require a base thickness 8 mils or more and a base-material resistivity of 50 ohm-centimeters or more. Because the wider base width increases on-state voltage, power dissipation, and thermal resistance, and decreases surge rating and switching speeds, this choice is a fundamental consideration and must be optimized for each voltage class, as a compromise between voltage and other characteristics.

The voltage ratings of thyristors are given for both steady-state and transient operation and for forward- and (SCR only) reverse-blocking conditions. For SCR's, voltages are considered to be in the forward or positive direction when the anode is positive with respect to the cathode. Negative voltages for SCR's are referred to as reverse-blocking voltages. For triacs, voltages are considered to be positive when main terminal No. 2 is positive with respect to main terminal No. 1. Alternatively, this condition may be referred to as operation in the first quadrant.

When the voltage applied to a thyristor is in the polarity for which switching to the on state is possible, the voltage-blocking capability of the device is temperature-sensitive. The maximum junction temperature for thyristors is usually between 100°C and 150°C. The selection of the maximum operating temperature represents a compromise which assures that a sufficient number of devices provide the required blocking-voltage capability (for which a low junction temperature

is desirable) and which allows the highest possible current rating for the thyristors (for which a high junction temperature is desirable). Increases in junction temperature above this maximum value result in a greater reliability stress and adversely affect the switching characteristics of thyristors.

**OFF-State Voltage**—The **repetitive peak off-state** voltage is the maximum value of off-state voltage that the thyristor should be required to block under the stated conditions of temperature and gate-to-cathode resistance. If this voltage is exceeded, the thyristor may switch to the on state. The circuit designer should insure that the rating is not exceeded to assure proper operation of the thyristor.

The effect of increased temperature is accentuated in thyristors because of the regenerative action upon which the operation of these devices is dependent. Thermally generated currents tend to be multiplied. If this blocking current crosses the gate-to-cathode junction, its effect on the thyristor is similar to that of the gate current and thus tends to reduce the breakover voltage $V_{BO}$. For this reason, off-state voltage ratings are specified at the maximum rated junction temperature.

A gate-to-cathode shunting resistance can be used to provide a path for the blocking current that bypasses the gate-to-cathode junction. The use of this shunt resistance improves the off-state blocking capability, but reduces the gate sensitivity. Off-state voltage ratings, therefore, are usually specified with the gate open to represent worst-case conditions.

Under relaxed conditions of temperature or gate impedance, or when the blocking capability of the thyristor exceeds the specified rating, it may be found that a thyristor can block voltages far in excess of its repetitive peak off-state voltage rating $V_{DROM}$. Because the application of an excessive voltage to a thyristor may produce irreversible effects, an absolute upper limit should be imposed on the amount of voltage that may be applied to the main terminals of the device. This voltage rating is referred to as the **nonrepetive peak off-state voltage** $V_{DSOM}$. It should be noted that the peak off-state voltage has a single rating irrespective of the voltage grade of the thyristor. This rating is a function of the construction of the thyristor and of the surface properties of the pellet. The $V_{DM}$ rating should not be exceeded under either continuous or transient conditions.

Fig. 229 shows a simple, inexpensive test circuit that may be used to valuate the off-state voltage capabilities of thyristors.

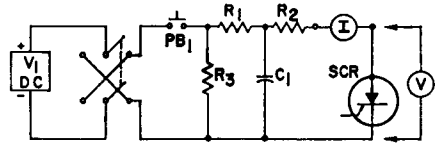



Figure 229. Test circuit used to determine dc forward- and reverse-voltage-blocking capabilities and leakage current of thyristors.

(The circuit may also be used for reverse-blocking and leakage tests of thyristors.) Resistor $R_1$ and capacitor $C_1$ are included in the test circuit to limit the rate of rise of applied voltage to the thyristor under test. Resistor $R_2$ limits the discharge of capacitor $C_1$ through the thyristor in the event that the thyristor is turned

on during the test. Resistor $R_3$ provides a discharge path for capacitor $C_1$.

**Reverse Voltages** (For SCR's) —Reverse-voltage ratings are given for SCR's to provide operating guidance in the third quadrant, or reverse-blocking mode.

The **repetitive peak reverse voltage** $V_{RSOM}$ is the maximum allowable value of reverse voltage, including all repetitive transient voltages, that may be applied to the SCR. Because reverse power dissipation is small at this voltage, the rise in junction temperature because of this reverse dissipation is very slight and is accounted for in the rating of the SCR.

The **nonrepetitive peak reverse voltage** $V_{RSOM}$ is the maximum allowable value of any nonrepetitive transient reverse voltage which may be applied to the SCR. These nonrepetitive transient voltages are allowed to exceed the steady-state ratings, even though the instantaneous power dissipation can be significant. While the transient voltage is applied, the junction temperature may increase, but removal of the transient voltage in a specified time allows the junction temperature to return to its steady-state operating temperature before a thermal runaway occurs.

The test circuit shown in Fig. 229 may be used for reverse voltage tests of an SCR.

## Maximum Junction Temperature

The maximum junction temperature is the second most important consideration in thyristor design. Several factors must be considered in determination of a maximum junction temperature rating, as indicated in Fig. 230.

At the upper end of the temperature scale, the maximum allowable junction temperature is restricted by material limits defined not so much by the silicon, but by peripheral materials such as the solders used on the pellet, solders used in lead attachments, encapsulating resins, and plastic package materials, where used. Temperature excursions into this area cause material and structural damage.

The next lower range of limiting temperature is determined by

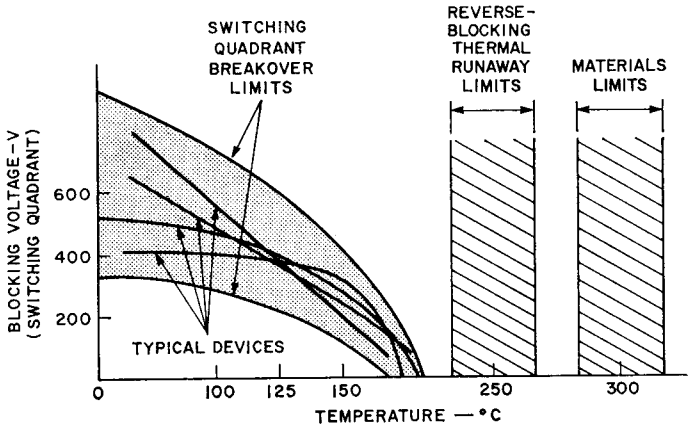



Figure 230. Basic design chart used to select maximum junction-temperature rating for a thyristor.

thermal runaway of the reverse-blocking junction in SCR's. This runaway is the familiar mechanism in which reverse-blocking losses generated in the junction increase with junction temperature at a faster rate than the dissipation capability of the heat sink. Fig. 231 illustrates thermal-runaway conditions for two common situations.

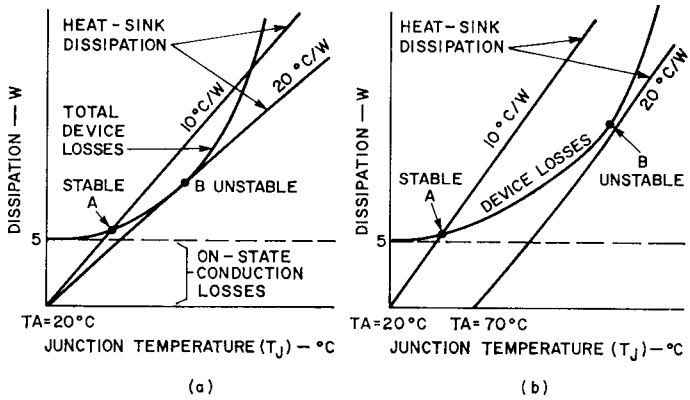



Figure 231. Two thermal situations for which the junction temperature may increase at a rate greater than the dissipation-capability of the heat sink: (a) effect of a decrease in the size of the heat sink; (b) effect of an increase in the ambient temperature.

The curves indicate that some level of power dissipation is associated with conduction in the on state and that the blocking power dissipation of a device increases with junction temperature. The thermal properties of the heat sink are assumed to be such that the junction temperature increases with dissipation at a rate of 10°C per watt. A device operating in this condition operates stably as a thermal system at point A with some temperature rise. If the size of the heat sink is decreased or the air flow is constricted in some way, the rate of increase in junction temperature rises. A new slope is then required to represent the characteristic of the heat sink. This situation is indicated by the 20°C-per-watt slope shown in Fig. 231(a). For this condition, the point of operation of the device-and-heat-sink combination shifts from A to B. This point is an unstable one for the thermal system. Thermal runaway occurs, and the temperature of the device increases uncontrollably into the range in which material degradation occurs.

Fig. 230(b) illustrates another situation in which the thermal system initially operates stably at point A. If the ambient temperature is increased, however, the same heat sink operates along a new dissipation curve. The operpoint of the thermal system shifts to point B and beyond. Again, because of the rapid increase in device temperature, limiting temperatures for material degradation are reached.

The lowest range of thyristor temperature limits, as shown in Fig. 230, is defined by the temperature sensitivity of the blocking capability in a switching

quadrant. In practice, various devices within a design family exhibit a range of temperature sensitivity. At temperatures greater than 200°C, very few thyristors retain their blocking capability. The temperature sensitivity is improved greatly by decreasing gate sensitivity or by emitter shorting.

It should be noted that turn-on induced by temperature is not in itself damaging, because it results from the excess generation of hole-electron pairs that forward-bias the cathode emitter in much the same manner as a gate signal. Device problems from over-temperature turn-on arise when the "gating" signal is too low in magnitude for the rate of rise of load current (di/dt) required. Temperature-induced turn-on is usually an application problem, therefore, because control of the load is lost.

For a given range of blocking capability, a junction-temperature rating is selected to provide adequate yield of devices that have the required voltage ratings. Typically, for thyristors, junction-temperature ratings have been selected in the range from 100° to 125°C.

It is also necessary to assure the adequacy of other parameters of the device which are temperature-sensitive, such as commutation capability in triacs, turn-off time in fast-switching SCR's, and static dv/dt ratings for all thyristors.

## On-State Current Ratings

Thyristor current ratings define maximum values for normal or repetitive currents and for surge or nonrepetitive currents. These maximum ratings are determined on the basis of the maximum junction-temperature rating, the junction-to-case thermal resistance, the internal power dissipation that results from the current flow through the thyristor, and the ambient temperature. The effect of these factors in the determination of current ratings is illustrated by the following example.

Fig. 232 shows curves of the maximum average forward power dissipation for the RCA-2N3873

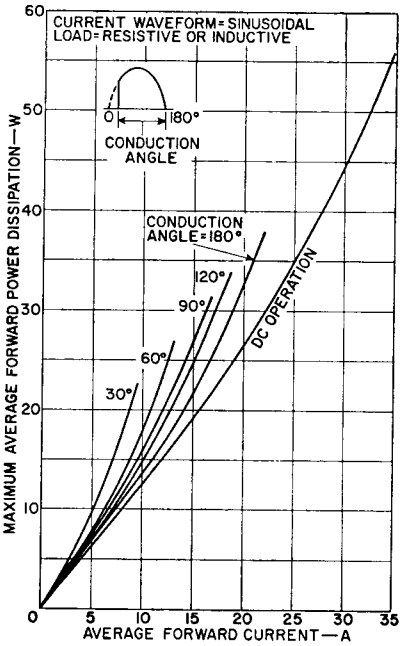



Figure 232. Power-dissipation rating chart for the RCA-2N3873 SCR.

SCR as a function of average forward current for dc operation and for various conduction angles. For the 2N3873, the junction-to-case thermal resistance $\theta_{J\text{-}C}$ is 0.92°C per watt and the maximum operating junction temperature $T_J$ is 100°C. If the maximum case tem-

perature $T_{C(max)}$ is assumed to be 65°C, the maximum average forward power dissipation can be determined as follows:

$$P_{AV(max)} = \frac{T_{J(max)} - T_{C(max)}}{\theta_{J-C}} \quad (243)$$

$$= \frac{(100 - 65)\ °C}{0.92\ °C/watt}$$

$$= 38 \text{ watts}$$

The maximum average forward current rating for the specified conditions can then be determined from the rating curves shown in Fig. 232. For example, if a conduction angle of 180 degrees is assumed, the average forward current rating for a maximum dissipation of 38 watts is found to be 22 amperes.

These calculations assume that the temperature is uniform throughout the pellet and the case. The junction temperature, however, increases and decreases under conditions of transient loading or periodic currents, depending upon the instantaneous power dissipated within the thyristor. The current rating must take these variations into account.

The on-state current ratings for a thyristor indicate the maximum values of average, rms, and peak (surge) current that should be allowed to flow through the main terminals of the device, under stated conditions, when the thyristor is in the on state. For heat-sink-mounted thyristors, these maximum ratings are based on the case temperature; for lead-mounted thyristors, the ratings are based on the ambient temperature.

**Steady-State Ratings**—The example used to show the effect of various factors on maximum current ratings pointed out that these ratings are determined on the basis of the internal power dissipation, the junction-to-case thermal resistance, and the difference between the maximum operating junction temperature and the maximum case temperature. Because the maximum operating junction temperature is fixed, the maximum on-state current ratings may be given by curves that relate current to case temperature. The maximum allowable current approaches zero as the case temperature approaches the maximum operating junction temperature because this current is directly proportional to the ratio of the difference between case and junction temperatures to the junction-to-case thermal resistance.

The basic ratings for thyristors define the boundary conditions for a basic current/case-temperature rating chart, such as that shown in Fig. 233. The various limits for a specific design rating are indicated curves 1 through 5. The first item is the choice of junction-temperature rating, as has been described, to insure an adequate yield of devices with useful voltage ratings. This first choice sets the upper boundary on the rating chart.

There is a practical level of ambient temperature, represented by curve 2, which must be allowed for in the design of equipment. Also, in the design of equipment, there must be some differential between the ambient and the case temperature. These considerations set the two lower boundaries, curves 2 and 3. The leads attached to the top of the thyristor, that is, to the cathode of the SCR or main terminal No. 1 of the triac, are not thermally

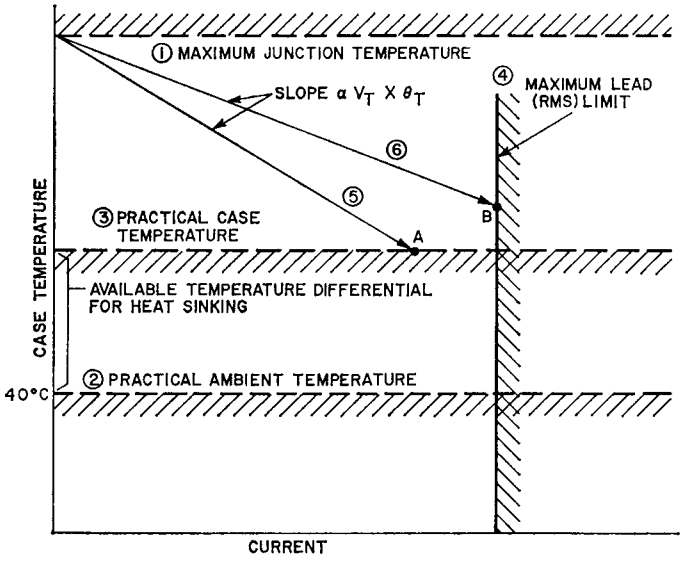



Figure 233. Basic current/case-temperature rating chart for thyristors.

connected to the heat sink of the thermal system that maintains the case temperature. The thermal limit of these leads sets an upper rms current limit for the package, as shown by curve 4.

The case temperature of the thyristor is derated from point 1 as current is increased to maintain a constant junction temperature. The slope of the derating curve 5 is proportional to the on-state dissipation and the junction-to-case thermal resistance. Both of these values are inversely proportional to pellet area. The device defined by derating curve 5 should be terminated in its current rating at point A for practical heat-sink arrangements and ambient conditions.

A device defined by curve 6 may have a larger chip area, a lower-thermal-resistance package, or a narrower base width in comparison to a device defined by curve 5. This particular device is terminated at the maximum rms current capability of the lead materials.

The **maximum average on-state current rating** is usually specified for a half-sine-wave current at a particular frequency. Fig. 234 shows curves of the maximum allowable average on-state current $I_{T(AV)}$ for the RCA-2N3873 SCR as a function of case temperature. Because peak and rms currents may be high for small conduction angles, the curves in Fig. 234 also show maximum allowable average currents as a function of conduction angle. The maximum operating junction temperature for the 2N3873 is 100°C. The rating curves indicate, for a given case temperature, the maximum average on-state current for which the average temperature of the pellet will not exceed the maximum allowable value. The rating curves may be used for only resistive or inductive loads. When ca-

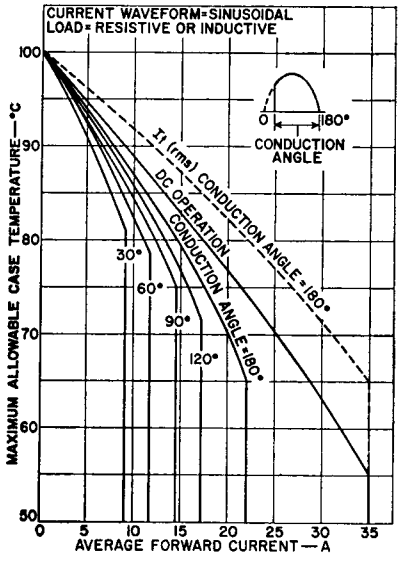



Figure 234. Current rating chart for the RCA-2N3873 SCR.

pacitive loads are used, the currents produced by the charge or discharge of the capacitor through the thyristor may be excessively high, and a resistance should be used in series with the capacitor to limit the current to the rating of the thyristor.

The ratio of rms to average values for a sinusoidal current waveform through an SCR is 1.57. The maximum average on-state current rating $I_{T(AV)}$, therefore, can be readily converted to the **maximum rms on-state current rating** $I_{T(RMS)}$. For example, as may be determined from Fig. 234, the maximum average on-state current for the 2N3873 is 22 amperes for a conduction angle of 180 degrees and a maximum case temperature of 65°C. For these same conditions, the rms current rating may be determined as follows:

$$I_{T(RMS)} = I_{T(AV)} \times 1.57$$
$$= 22 \text{ amperes} \times 1.57$$
$$= 35 \text{ amperes}$$

The dashed-line curve in Fig. 234 shows the rms current rating for the 2N3873 as a function of case temperature for a conduction angle of 180 degrees.

The on-state current rating for a triac is given only in rms values because these devices normally conduct alternating current. Fig. 235 shows an rms on-state current rating curve for a typical triac as a function of case temperature. As with the SCR, the triac curve is derated to zero current when the case temperature rises to the maximum operating junction temperature. Triac current ratings are given for full-wave conduction under resistive or inductive loads. Precautions should be taken to limit the peak current to tolerable levels when capacitive loads are used.

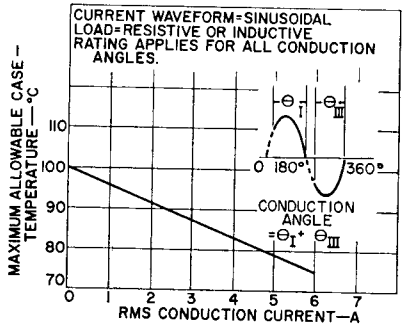



Figure 235. Current-rating curve for a typical RCA triac.

**Surge Ratings**—the **surge on-state current rating** $I_{TSM}$ indicates the maximum peak value of a short-duration current pulse that should be allowed to flow through a thyristor during one

on-state cycle, under stated conditions. This rating is applicable for any rated load condition. During normal operation, the junction temperature of a thyristor may rise to the maximum allowable value; if the surge occurs at this time, the maximum limit is exceeded. For this reason, a thyristor is not rated to block off-state voltage immediately following the occurrence of a current surge. Sufficient time must be allowed to permit the junction temperature to return to the normal operating value before gate control is restored to the thyristor. Fig. 236 shows a surge-current rating curve for the 2N3873 SCR. This

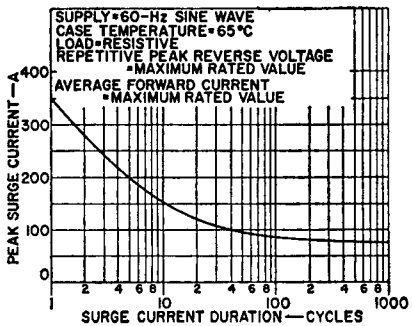



Figure 236. Surge-current rating curve for the RCA-2N3873 SCR.

curve shows peak values of half-sine-wave forward (on-state) current as a function of overload duration measured in cycles of the 60-Hz current. Fig. 237 shows surge-current rating curves for a typical triac. For triacs, the rating curve shows peak values for a full-sine-wave current as a function of the number of cycles of overload duration. Multicycle surge curves are the basis for the selection of circuit breakers and fuses that are used to prevent damage to the thyristor in the event of accidental short-circuit of the device. The number of surges permitted over the life of the thyristor should be limited to prevent device degradation.

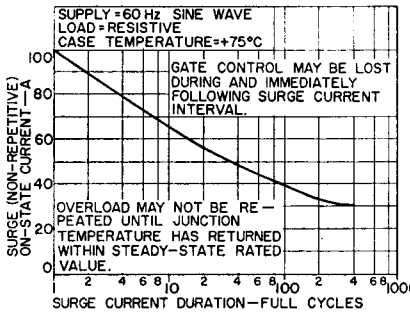



Figure 237. Surge-current rating chart for a typical triac.

## Holding and Latching Currents

After a thyristor has been switched to the on-state condition, a certain minimum value of anode current is required to maintain the thyristor in this low-impedance state. If the anode current is reduced below this critical holding-current value, the thyristor cannot maintain regeneration and reverts to the off or high-impedance state. Because the holding current ($I_H$) is sensitive to changes in temperature (increases as tempreature decreases), this rating is specified at room temperature with the gate open.

The latching-current specification of a thyristor denotes a value of anode current, slightly higher than the holding current, which is the minimum amount required to sustain conduction immediately after the thyristor is switched from the off state to the on state and the gate signal is removed. Once the latching current ($I_L$) is reached, the thyristor remains in the on, or low-impedance, state until its anode current is decreased below the holding-current value. The latch-

ing-current rating is an important consideration when a thyristor is to be used with an inductive load because the inductance limits the rate of rise of the anode current. Precautions should be taken to insure that, under pulse-gating conditions, the gate signal is present until the anode current rises to the latching value so that complete turn-on of the thyristor is assured.

Fig. 238 shows the latching current $I_L$ as a function of the gate triggering current $I_{GT}$ for each operating mode of a thyristor. In the $I^+$, $III^-$, and $III^+$ modes, the latching-current value

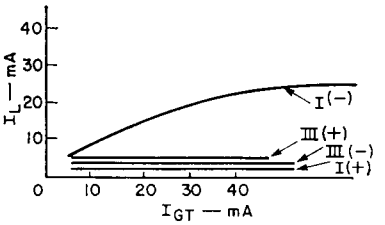



Figure 238. Latching current as a function of gate current for each triggering mode of a triac.

remains relatively constant with changes in the gate current. In the $I^-$ mode, however, larger values of latching current are required as the amount of gate current driven into the triac is increased. "Starved" gating in the $I^-$ mode, therefore, can be a problem because of the dependence of the latching-current value upon the amount of gate current. This dependence is one of the reasons that a triac should be overdriven whenever possible.

Fig. 239 shows a simple test circuit that may be used to determine the holding and latching currents of a triac. For the holding-current test, the triac is gated on by closing the toggle switches $S_1$ and $S_2$, depressing the pushbutton switches $PB_1$ and $PB_2$, and adjusting potentiometer $R_L$ to a low value. The pushbutton switch $PB_1$ is then opened, and the value of potentiometer $R_L$ is increased until the triac turns off. The current reading indicated on the meter $M_1$ just prior to complete turn-off is the holding-current value.

For the latching-current test, toggle switches $S_1$ and $S_2$ are closed to select the triggering mode, potentiometer $R_{GT}$ is adjusted for the desired value of gate current [usually the $I_{GT}$ (max) value for the device being tested], and the pushbutton switches $PB_1$ and $PB_2$ are held closed as potentiometer $R_L$ is adjusted to some value for which the triac is maintained in the off state. The value of $R_L$ is then gradually decreased in small increments until the gate signal

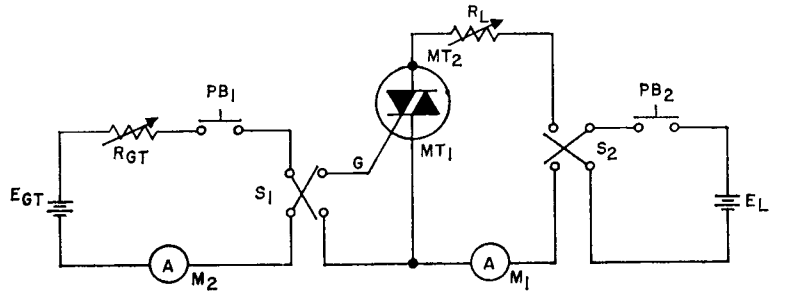



Figure 239. Test circuit used to determine holding and latching currents of thryistors.

injected at each $R_L$ setting by depression of $PB_1$ turns on the triac. The reading on meter $M_1$ indicates the latching-current value. The pushbutton switch $PB_1$ should be alternately opened and closed for this value of $R_L$ to assure that the triac was originally off.

## Critical Rate of Rise of On-State Current (di/dt)

When a thyristor is turned on by application of a gate trigger pulse, conduction does not instantly occur throughout the entire pellet. The initial flow of current is concentrated in very small areas near the gate contact. A short interval of time is required for the current to spread sufficiently so that the entire pellet is in conduction. If the rate at which the load current increases is high in comparison to the rate at which current spreads laterally across the pellet, considerable energy will be concentrated in the turned-on areas, and localized high-temperature regions (hot spots) may develop. These hot spots may adversely affect other characteristics of the thyristor or, in extreme cases, may cause permanent damage to the pellet. For these reasons, thyristor manufacturers usually specify a limiting value to define the critical rate of rise of on-state current for their products.

The waveshape for testing the di/dt capability of the RCA 2N3873 is shown in Fig. 240. The critical rate of rise of on-state current is dependent upon the size of the cathode area that begins to conduct initially, and the size of this area is increased for larger values of gate trigger current. For this reason, the di/dt rating is specified for a specific value of gate trigger current.

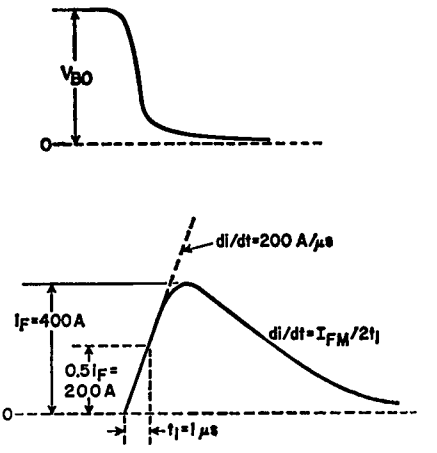



Figure 240. Voltage and current waveforms used to determine di/dt rating of the RCA-2N3873 SCR.

## Critical Rate of Rise of Off-State Voltage (dv/dt)

An important parameter for thyristors is the "critical rate of rise of off-state voltage." A source voltage can be suddenly applied to an SCR or a triac which is in the off state through either closure of an ac line switch or transient voltages as a result of an ac line disturbance. If the fast rate of rise of the transient voltage (dv/dt) exceeds the device rating, the thyristor may switch from the off state to the conducting state in the absence of a gate signal. If the thyristor is controlling alternating voltage, "false" (non-gated) turn-on resulting from a transient imposed voltage is limited to no more than half the applied voltage because turn-off occurs during the zero current crossing. However, if the source voltage suddenly applied to the off thy-

ristor is a dc voltage, the device may switch to the on state and turn-off could then be achieved only by circuit interruptions. The switching from the off state is caused by the internal capacitance of the thyristor. A steep-rising voltage impressed across the terminals of a thyristor causes a capacitance-charging current to flow through the device. This charging current ($i = Cdv/dt$) is a function of the rate of rise of applied off-state voltage. If the rate of rise of voltage exceeds a critical value, the capacitance-charging current exceeds the gate trigger current and causes device turn-on. Operation at elevated junction temperatures reduces the thyristor ability to support a steep-rising (high-dv/dt) voltage because less gate current is required for turn-on. The effect of temperature on the critical rate of rise of off-state voltage is shown in Fig. 241.

The use of the shorted emitter construction in RCA thyristors has resulted in a substantial increase in the dv/dt capability of these devices by providing a shunt path around the gate-to-cathode junction. Typical units can withstand rates of voltage rise up to 200 volts per microsecond under worst-case conditions. The dv/dt capability of a thyristor decreases as the temperature rises and is increased by the addition of an external resistance from gate to reference terminal. The dv/dt rating, therefore, is given for the maximum junction temperature with the gate open, i.e., for worst-case conditions.

Fig. 242(a) shows a simple test circuit that may be used to determine the dv/dt capability of a

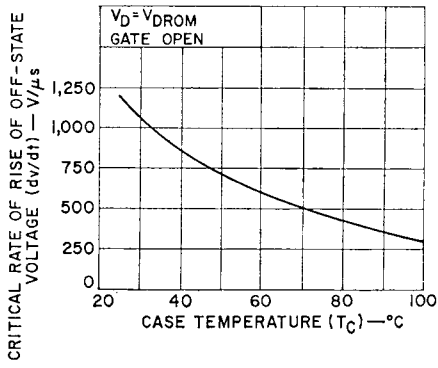
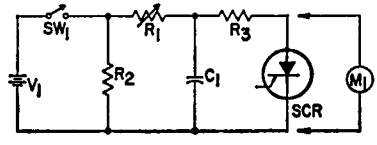


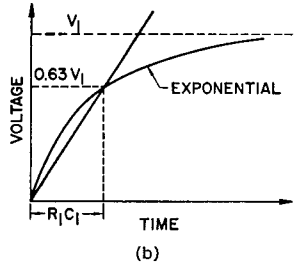

Figure 241. Critical rate of rise of off-state voltage as a function of case temperature.

thyristor. The curves in Fig. 242(b) define the critical values for linear and exponential rates of increase in reapplied forward off-state voltage for an SCR. The critical value for the exponential rate of rise of forward




$V_1$ = anode supply voltage (variable)
$SW_1$ = mercury-wetted relay
$R_1$ = noninductive variable resistor
$R_2$ = discharge resistor
$R_3$ = current-limiting resistor
$M_1$ = oscilloscope

(a)




(b)

Figure 242. (a) Test circuit and (b) waveforms used to determine dv/dt capability of a thyristor.

voltage is the rating given in the manufacturer's test specifications.

This rating is determined from the following equation:

$$\frac{dv}{dt} = \frac{V_{DROM} \text{ rating}}{RC \text{ time constant}} \times 0.632 \quad (244)$$

The dv/dt specification allows a circuit designer to design an RC time-constant network that can be used to limit the rate of rise of a transient voltage below the critical value of the thyristor.

Voltage transients which occur in electrical systems as a result of disturbance on the ac line caused by various sources such as energizing transformers, load switching, solenoid closure, contactors, and the like may generate voltages which are above the ratings of thyristors and result in spike voltages that exceed the critical rate of rise of off-state voltage. Thyristors, in general, switch from the off state to the on state whenever the breakover voltage of the device is exceeded, and energy is then transferred to the load. Good practice in the use of thyristors exposed to a heavy transient environment is to provide some form of transient suppression.

For applications in which low-energy, long-duration transients may be encountered, it is advisable to use thyristors that have voltage ratings greater than the highest voltage transient expected in the system to provide protection against destructive transients. The use of voltage clipping cells is also effective. In either case, analysis of the circuit application will reveal the extent to which suppression should be employed. In an SCR application in which there is a possibility of exceeding the reverse-blocking voltage rating, it is advisable to add a clip cell or to use an SCR with a higher reverse-blocking voltage rating to minimize power dissipation in the reverse mode. Because triacs generally switch to a low conducting state, if the di/dt buildup of the principal current flow after turn-on is within device ratings it is safe to assume that reliable operation will be achieved under the specified conditions.

The use of an RC snubber is most effective in reducing the effects of the high-energy short-duration transients more frequently encountered in thyristor applications. When an RC snubber is added at the thyristor terminals, the rate of rise of voltage at the terminals is a function of the load impedance and the RC values used in the network. In some applications, "false" (non-gated) turn-on for even a portion of the applied voltage cannot be tolerated, and circuit response to voltage transients must be determined. An effective means of generating fast-rising transients and observing the circuit response to such transients is shown in Fig. 243. This circuit makes use of the "splash" effects of a mercury-wetted relay to transfer a capacitor charge to the input terminals of a control circuit. This approach permits generation of a transient

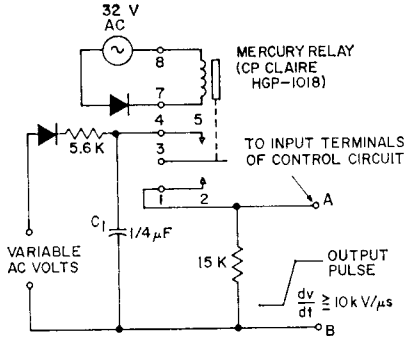



Figure 243. Circuit used to generate fast-rising transients.

of known magnitude whose rate of rise of voltage can easily be displayed on an oscilloscope. For a given load condition, the values in the RC snubber network can be adjusted so that the transient voltage at the device terminals is suppressed to a tolerable level. This approach affords the circuit designer with meaningful information as to how a control circuit will respond in a heavy transient environment. The circuit is capable of generating transient voltages in excess of 10 kilovolts per microsecond, which exceeds industrial generated transients. The response of a 100-millihenry solenoid control circuit exposed to a fast-rising transient is shown in Fig. 244.

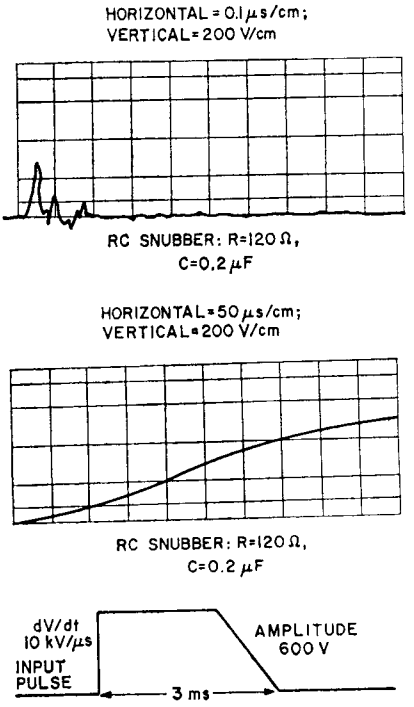



Figure 244. Waveforms showing response of a 100-millihenry solenoid control circuit to a fast-rising transient.

## GATE CHARACTERISTICS

SCR's and triacs are specifically designed to be triggered by a signal applied to the gate terminal. The manufacturer's specifications indicate the magnitudes of gate current and voltage required to turn on these devices. Gate characteristics, however, vary from

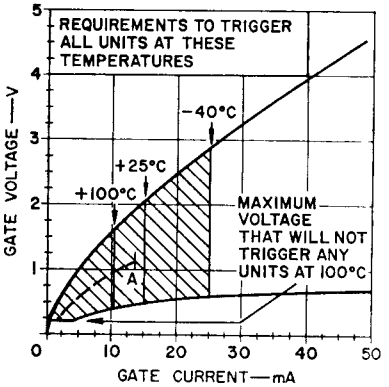



Figure 245. Gate-characteristic curves for a typical RCA SCR.

device to device even among devices within the same family. For this reason, manufacturer's specifications on gating characteristics provide a range of values in the form of characteristic diagrams. A diagram such as that shown in Fig. 245 is given to define the limits of gate currents and voltages that may be used to trigger any given device of a specific family. The boundary lines of maximum and minimum gate impedance on this characteristic diagram represent the loci of all possible triggering points for thyristors in this family. The curve OA represents the gate characteristic of a specific device that is triggered within the shaded area.

## Effect of Gate Signal on Breakover Voltage

The breakover voltage of a thyristor can be varied, or controlled, by injection of a signal at the gate, as indicated by the family of curves shown in Fig. 246. Although this family of curves is shown in the first quadrant typical of SCR operation, a similar set of curves can also be drawn for the third quadrant to represent triac operation.

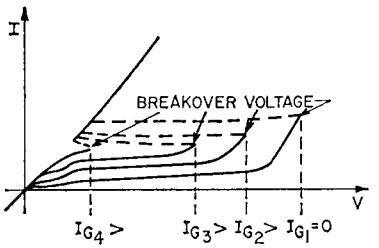



Figure 246. Curves showing breakover characteristics of a thyristor for different values of gate current.

When the gate current $I_G$ is zero, the applied voltage must reach the breakover voltage of the thyristor before switching occurs. As the value of gate current is increased, however, the ability of a thyristor to support applied voltage is reduced and there is a certain value of gate current at which the behavior of the thyristor closely resembles that of a rectifier. Because thyristor turn-on, that results from exceeding the breakover voltage, can produce instantaneous power dissipation during the switching transition, an irreversible condition may exist unless the magnitude and rate of rise of principal current is restricted to tolerable levels. For normal operation, therefore, thyristors are operated at applied voltages lower than the breakover voltage, and are made to switch to the on state by gate signals of sufficient amplitude to assure complete turn-on independent of the applied voltage. Once the thyristor is triggered to the on state, the principal-current flow is independent of gate voltage or gate current, and the device remains in the on state until the principal-current flow is reduced to a value below the holding current required to sustain regeneration.

Because of its complex structure, a triac can be triggered by either a positive or negative signal regardless of the polarity of the voltage across the main terminals of the device. The direction of the principal current influences the gate trigger current; the current required to trigger the triac, therefore, differs for each operating mode. The operating modes in which the principal current is in the same direction as the gate current require less gate trigger current than the operating modes in which the principal current is in opposition to the gate current. Fig. 247 shows the directions of the gate current and the

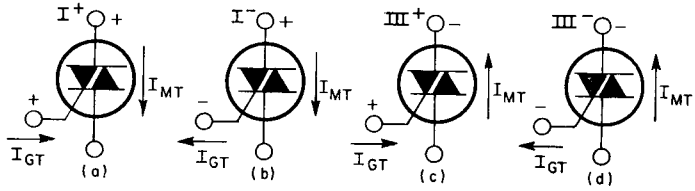



Figure 247. Four gating conditions of a triac.

principal current for each triggering mode.

## Trigger Level

The magnitude of gate current and voltage required to trigger a thyristor varies inversely with junction temperature. As the junction temperature increases, the level of gate signal required to trigger the thyristor becomes smaller. Worst-case triggering conditions occur, therefore, at the minimum operating junction temperature.

The maximum value of gate voltage below the level required to trigger any unit of a specific thyristor family is also an important gate characteristic. At high operating temperatures, the level of gate voltage required to trigger a thyristor approaches the minimum value, and undesirable noise signals may inadvertently trigger the device. The maximum nontriggering gate voltage at the maximum operating junction temperature of the device, therefore, is a measure of the noise-rejection level of a thyristor.

The gate voltage and current required to switch a thyristor to its low-impedance state at maximum rated forward anode current can be determined from the circuit shown in Fig. 248. The value of resistor $R_2$ is chosen so that maximum anode current, as specified in the manufacturer's current rating, flows when the device latches into its low-impedance state. The value of resistor $R_1$ is gradually decreased until the device under test is switched from its high-impedance state to its low-impedance state. The values of gate current and gate voltage immediately prior to switching are the gate voltage and current required to trigger the thyristor.

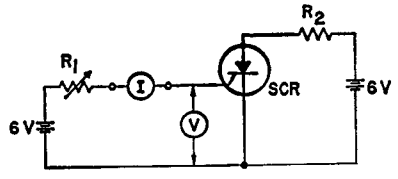



Figure 248. Test circuit used to determine gate-trigger-pulse requirements of thyristors.

The **gate nontrigger voltage** $V_{gnt}$ is the maximum dc gate voltage that may be applied between gate and cathode of the thyristor for which the device can maintain its rated blocking voltage. This voltage is usually specified at the rated operating temperature (100°C) of the thyristor. Noise signals in the gate circuit should be maintained below this level to prevent unwanted triggering of the thyristor.

## Pulse Triggering

The gate current specified in published data for thyristors is the dc gate trigger current required to switch an SCR or triac into its low-impedance state. For practical purposes, this dc value can be considered equivalent to a pulse current that has a minimum pulse width of 50 microseconds. For gate-current pulse widths smaller than 50 microseconds, the pulse-current curves associated with a particular device should be used to assure turn-on.

When pulse triggering of a thyristor is required, it is always advantageous to provide a gate-current pulse that has a magnitude exceeding the dc value required to trigger the device. The use of large trigger currents reduces variations in turn-on time, increases di/dt capability, minimizes

the effect of temperature variation on triggering characteristics, and makes possible very short switching times. When a thyristor is initially triggered into conduction, the current is confined to a small area which is usually the more sensitive part of the cathode. If the anode-current magnitude is great, the localized instantaneous power dissipation may result in irreversible damage unless the rate of rise of principal current is restricted to tolerable levels to allow time for current spreading over a larger area. When a much larger gate signal is applied, a greater part of the cathode is turned on initially; as a result, turn-on time is reduced, and the thyristor can support a much larger peak anode inrush current.

In the past, the maximum value of gate signal that could be used to trigger a thyristor was severely restricted by minimum dc triggering requirements and limitations on maximum gate power. The coaxial gate structure and the "shorted-emitter" construction techniques now used in RCA thyristors, however, has greatly extended the range of limiting gate characteristics. As a result, the gate-dissipation ratings of RCA thyristors are compatible with the power-handling capabilities of other elements of these devices. Advantage can be taken of the higher peak-power capability of the gate to improve dynamic performance, increase di/dt capability, minimize interpulse jitter, and reduce switching losses. This higher peak-power capability also allows greater interchangeability of thyristors in high-performance applications.

The "shorted-emitter" technique makes use of the resistance path within the gate layer which is in direct contact with the cathode electrode of the thyristor. When gate current is first initiated, most of the current bypasses the gate-to-cathode junction and flows from the resistive gate layer to the cathode contact. When the IR drop in this gate layer exceeds the threshold voltage of the gate-to-cathode junction, the current across this junction increases until the thyristor is triggered.

When an SCR is triggered by a gate signal just sufficient to turn on the device, the entire junction area does not start to conduct instantaneously. Instead, as pointed out in the discussion on **Critical Rate of Rise of On-State Current,** the device current is confined to a small area, which is usually the most sensitive part of the cathode. The remaining cathode area turns on as the anode current increases. When a much larger signal is applied to the gate, a greater part of the cathode is turned on initially and the time to complete the turn-on process is reduced. The peak amplitude of gate-trigger currents must be large, therefore, when thyristors have to be turned on completely in a short period of time. Under such conditions, the peak gate power is high, and pulse triggering is required to keep the average gate dissipation within the values given in the manufacturer's specifications. New gate ratings, therefore, are required for this type of application.

The forward gate characteristics for thyristors, shown in Fig. 249, indicate the maximum allowable pulse widths for various peak values of gate input power. The pulse width is determined by the relationship that exists between gate power input and the increase in the temperature of the thyristor pellet that results from the ap-

plication of gate power. The curves shown in Fig. 249(a) are for RCA SCR's that have relatively small current ratings (2N4101, 2N4102, and 40379 types), and the curves shown in Fig. 249(b) are for RCA SCR's that have larger current ratings (2N3670, 2N3873, and 2N3899 types). Because the higher-current thyristors have larger pellets, they also have greater thermal capacities

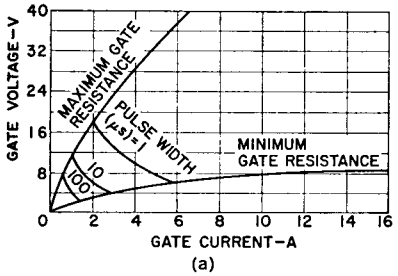



Figure 249. Forward-gate characteristics for pulse triggering of RCA SCR's: (a) low-current types, (b) high-current types.

than the smaller-current devices. Wider gate trigger pulses can therefore be used on these devices for the same peak value of gate input power.

Because of the resistive nature of the "shorted-emitter" construction, similar volt-ampere curves can be constructed for reverse gate voltages and currents, with maximum allowable pulse widths for various peak-power values, as shown in Fig. 250. These curves

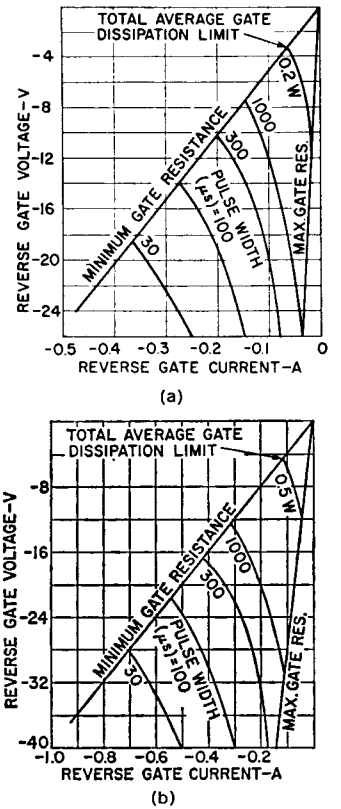



Figure 250. Reverse gate characteristics of RCA SCR's: (a) low-current types, (b) high-current types.

indicate that reverse dissipations do not exceed the maximum allowable power dissipation for the device.

## Trigger-Circuit Requirements

The basic gate trigger circuit for a thyristor can be represented by a voltage source and a series resistance, as shown in Fig. 251. The series resistance should include both the external circuit resistance and the internal generator resistance. With this type of equivalent circuit, the conventional load-line approach to gate trigger-circuit design can be used.

With pulse triggering, it is assumed initially that the turn-on time required to trigger all thyristors of the same type is known, and that the maximum allowable gate trigger-pulse widths for specific gate-power inputs are to be determined.

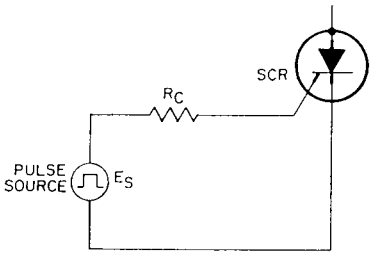



Figure 251. Equivalent diagram of the basic gate-trigger circuit for a thyristor.

The magnitude of gate-trigger current required to turn on all SCR's of a given type can be determined from the turn-on characteristics shown in Fig. 252. The spread or band of turn-on characteristics for the same gate current results from the variation of gate-trigger characteristics among devices of the same family. Because of the greater over-drive factor involved, the same gate current applied to a device obviously turns on a low-gate-current device in much less time than that required to turn on a higher-gate-current device. For example, a gate-trigger current of 100 milliamperes overdrives an SCR that requires a trigger current of only 2 milliamperes by a factor of 50 and causes the device to turn on very quickly, while an SCR that requires 10 milliamperes of trigger current is overdriven by a factor of 10 and is turned on more slowly. As the gate current increases, the band of turn-on characteristics becomes narrower, and an increase in gate current does not effectively decrease the turn-on time.

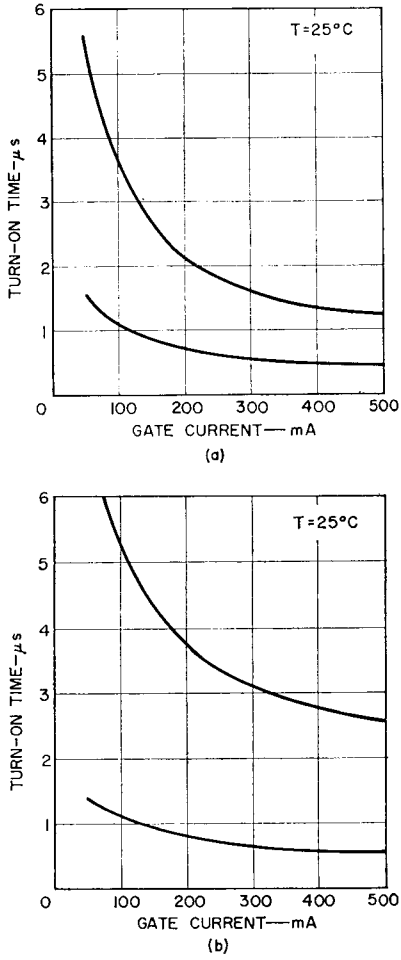



Figure 252. Turn-on time distribution among RCA SCR's: (a) low-current types, (b) high-current types.

The turn-on characteristics shown in Fig. 252 indicate that a gate-trigger current of 1 ampere is required to assure that all devices of this type will turn on in 2.5 microseconds (the 2.5-microsecond ordinate level intersects the upper curve at 500 milliamperes).

In addition, the width of the gate-trigger pulse should be at least 2.5 microseconds to ensure that the SCR remains on after it is triggered. Actually, the minimum requirement is that the pulse width be wide enough for the SCR anode current to achieve the latching value. Conservative design, however, requires the pulse width to be at least equal to the turn-on time. For inductive loads, the turn-on time is larger than indicated in the characteristics curves because of the slow rise of current through the inductance.

A straight load line can then be plotted on the pulse triggering characteristics, as shown in Fig. 253. The two points that determine the position of this line are the source voltage (20 volts) and a point slightly above the intersection of the required gate current (500 milliamperes) and the curve of maximum gate resistance. The load line should lie below the pulse-width curve required to trigger all SCR's (in this example, the 2.5-microsecond curve). The maximum allowable pulse width is obtained by estimation of the pulse-width curve tangent to the load

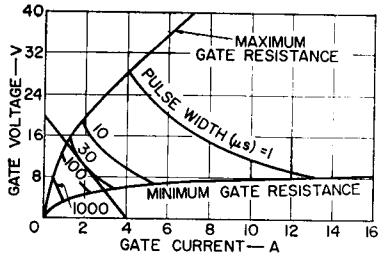



Figure 253. Forward gate characteristics of typical RCA SCR's showing load line for a source of 20 volts and a required gate current of 1 ampere.

line. In this example, the pulse width is estimated to be 30 microseconds (the pulse-width curves are logarithmically spaced). The load line intersects the abscissa at the 4-ampere point. The maximum circuit resistance, therefore, is 5 ohms. The peak gate power is the product of gate voltage and gate current at the point of tangency of the pulse-width curve, and is approximately 20 watts (10 volts x 2 amperes).

When gate pulses are used to trigger SCR's, the maximum allowable operating frequency f is dependent upon the average power rating of the gate $P_{g(avg)}$ and can be determined from the following equation:

$$f = P_{g(avg)}/P_{g(pk)} \times PW_g \quad (245)$$

where $P_{g(pk)}$ is the peak gate power and $PW_g$ is the gate pulse width.

If it is assumed that only half the total average gate-dissipation rating, or 0.25 watt, is used to trigger the device, this value is used in the frequency calculation. For example, if this value is 0.25 watt for the SCR selected, then the maximum allowable operating frequency is determined as follows:

$$f = \frac{0.25\ W}{20W \times 2.5 \times 10^{-6}\,\text{second}}$$

$$= 5000\ Hz$$

If there is no reverse gate power dissipation, the maximum allowable frequency can be 10,000 Hz. If the maximum allowable pulse width is 30 microseconds, the maximum allowable operating frequency is proportionately reduced to 416 Hz.

The trigger-circuit design is usually fixed by the requirements for reliable triggering, and reverse gate dissipation is considered after the values of source voltage and circuit resistance have been determined. Reverse gate

power dissipation results from reverse gate-bias conditions or circuit reaction caused by some switching function. As in the case of the forward gate characteristics, a load-line approach can also be used to determine the reverse gate characteristics. The maximum anticipated value of reverse gate potential is used as the source voltage, and the external circuit resistance is used to determine the slope of the load line. The load line on the reverse gate characteristics shown in Fig. 254 represents a reverse gate-source

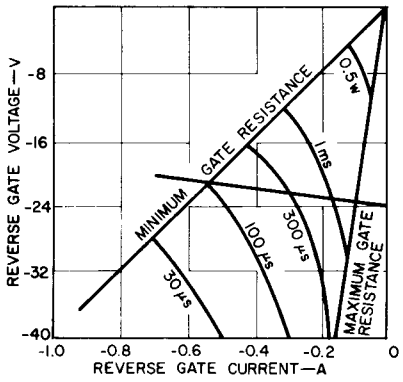



Figure 254. Reverse gate characteristics for typical RCA SCR's showing load line for a reverse gate-source voltage of 24 volts and an external circuit resistance of 5 ohms.

voltage of 24 volts and an external-circuit resistance of 5 ohms. From the relationship that exists among pulse width, average gate power, peak gate power, and frequency, a maximum pulse width can be calculated for the actual operating frequency. For a reverse gate dissipation of 0.25 watt, peak gate power of 10 watts, and a frequency of 5000 Hz, the maximum allowable pulse width PW is calculated as follows:

$$PW = \frac{0.25\ W}{5000\ Hz \times 10W} \quad (246)$$

$$= 5\ \text{microseconds}$$

This reverse gate-pulse width should be less than the maximum allowable pulse width, as determined by the curve that lies just below the load line on Fig. 254. In this example, the maximum allowable pulse width for reverse dissipation is 100 microseconds.

The total average dissipation caused by gate-trigger pulses is the sum of the average forward and reverse dissipations. This total dissipation should correspond to the average gate power dissipation shown in the published data for the selected SCR. If the average gate dissipation exceeds the maximum published value, as the result of high forward gate-trigger pulses and transient or steady-state reverse gate biasing, the maximum allowable forward-conduction-current rating of the device must be reduced to compensate for the increased rise of junction temperature caused by the increased gate power dissipation.

The trigger-circuit design considerations described for RCA SCR's also apply to RCA triacs. Although both types of devices are triggered in the same manner, the triac can be triggered by either positive or negative gate-trigger pulses independent of the polarity of the voltage between the main terminals.

## SWITCHING CHARACTERISTICS

The ratings of thyristors are based primarily upon the amount of heat generated within the device pellet and the ability of the device package to transfer the internal heat to the external case. For high-frequency applications in which the peak-to-average current ratio is high, or for high-perfor-

mance applications that require large peak values but narrow current pulses, the energy lost during the turn-on process may be the main cause of heat generation within the thyristor. The switching properties of the device must be known, therefore, to determine power dissipation which may limit the device performance.

## Turn-on Time

When a thyristor is triggered by a gate signal, the turn-on time of the device consists of two stages, a delay time $t_d$ and a rise time $t_r$, as shown in Fig. 255. The total turn-on time $t_{gt}$ is defined as the time interval between the initiation of the gate

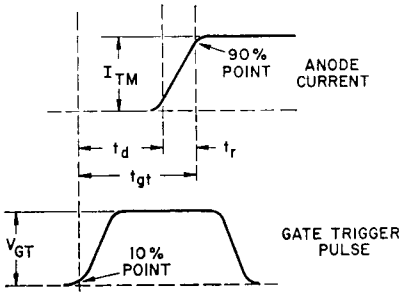



Figure 255. Gate current and voltage turn-on waveforms for a thyristor.

signal and the time when the resulting current through the thyristor reaches 90 per cent of its maximum value with a resistive load. The **delay time** $t_d$ is defined as the time interval between the 10-per-cent point of the leading edge of the gate-trigger voltage and the 10-per-cent point of the resulting current with a resistive load. The **rise time** $t_r$ is the time interval required for the principal current to rise from 10 to 90 per cent of its maximum value. The **total turn-on time,** therefore, is the sum of both the delay and rise times of the thyristor.

Although the turn-on time is affected to some extent by the peak off-state voltage and the peak on-state current level, it is influenced primarily by the magnitude of the gate-trigger current pulse. Fig. 256 shows the variation in turn-on time with gate-trigger current for the RCA-2N3873 SCR.

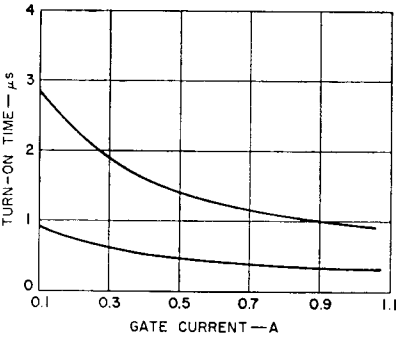



Figure 256. Turn-on time characteristics for the RCA-2N3873 SCR.

When larger currents are available from the gate-trigger pulses, the delay-time portion of the turn-on period is reduced, and the over-all turn-on time is decreased. When it is desirable to reduce the variation in turn-on time among devices of the same type, higher gate-drive signals should be used.

Fig. 257 shows a simple test circuit used to determine turn-on times of thyristors. The value of resistor $R_1$ is chosen so that the rated value of current flows through the thyristor. Turn-on time is specified by the thyristor manufacturer at the rated blocking voltage.

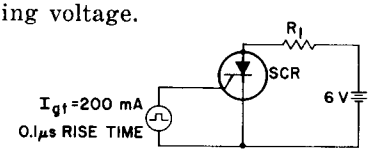



Figure 257. Test circuit used to determine turn-on time of thyristors.

When a thyristor is turned on by a gate-current pulse, current does not start to flow throughout the entire junction instantaneously; instead, the current is confined initially to a small area adjacent to the gate. The voltage drop across the thyristor at this time is large because the current density in the small area that is turned on is high. As the conduction area increases, the current density is reduced, and the voltage drop across the thyristor becomes smaller. Eventually, the boundaries of the high-current-density region propagate across the entire junction area. The time required for completion of this action is considerably longer than the specified turn-on time. For resistive loads, the turn-on time can be defined as the time interval between the 10-per-cent point at the beginning of the gate voltage and the instant at which the applied blocking voltage decreases to 10 per cent of its original value.

For thyristors operated at low blocking voltages, the 10-per-cent value is insignificant from the standpoint of device dissipation. For thyristors operated at blocking voltages in the order of hundreds of volts, however, 10 per cent is sufficiently high in magnitude to represent an appreciable amount of device dissipation. Moreover, the typical turn-on time, as defined for certain gate drives, may be in the order of 2 to 3 microseconds, while the time required for conduction to spread over the entire junction area may be in the order of 20 microseconds. During this spreading time, the dynamic voltage drop is high, and the current density can produce localized hot spots in the pellet area in conduction. In order to guarantee reliable operation and to provide guidance for equipment designers in applications having short conduction periods, published data for RCA thyristors give the voltage drop at a given instantaneous forward current and at a specified time after turn-on from an OFF-state condition. The wave-shapes for the initial ON-state voltage for the RCA-2N3873 SCR are shown in Fig. 258. This initial voltage, together with the time required for reduction of the dynamic forward voltage drop during the spreading time, is an indication of the current-switching capability of the thyristor.

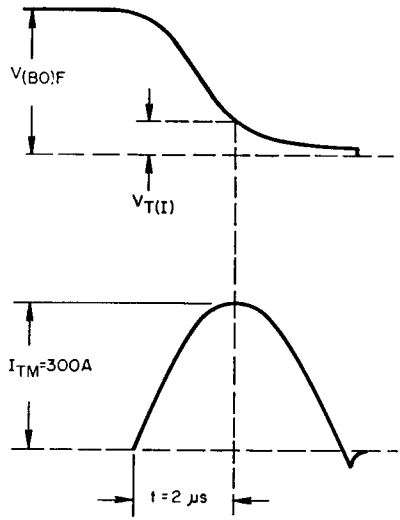



Figure 258. Initial ON-state voltage and current waveforms for the 2N3873 SCR.

When the entire junction area of a thyristor is not in conduction, the current through that fraction of the pellet area in conduction may result in large instantaneous power losses. These turn-on switching losses are pro-

portional to the current and the voltage from cathode to anode of the device, together with the repetition rate of the gate-trigger pulses. The instantaneous power dissipated in a thyristor under such conditions is shown in Fig. 259. The curves shown in this figure indicate that the peak power dissipation occurs in the short interval immediately after the device starts to conduct, usually in the first microsecond. During this time interval, the peak junction temperature may exceed the maximum operating temperature given in the manufacturer's data; in this case, the thyristor should not be required to block voltages immediately after the conduction interval. If the thyristor must block voltages immediately following the conduction interval, the junction-temperature rating must not be exceeded, and sufficient time must elapse to allow the junction temperature to decrease to the operating temperature before blocking voltage is re-applied to the device.

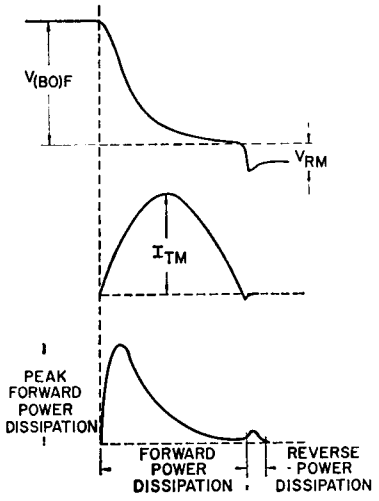



Figure 259. Instantaneous power dissipation in a thyristor during turn-on.

The transient temperature rise may have a major effect on the turn-off time of a thyristor. As a result, when transient effects have to be considered, turn-off time measurements should be made under pulsed conditions.

## Turn-off Time (for SCR's)

Turn-off time of a thyristor is associated only with SCR's. In triacs, a reverse voltage cannot be used to provide circuit-commutated turn-off voltage because a reverse voltage applied to one half of the triac structure would be a forward voltage for the other half.

When the forward current of an SCR is reduced to zero at the end of the conduction period, the application of forward voltage between the anode and the cathode terminals must be delayed for a definite length of time if the device is expected to block the reapplied forward voltage. This required minimum amount of time is referred to as the **turn-off time** of the SCR. In most practical applications, the forward current is removed from the SCR by the reversal of current flow in the circuit with a gradual, controlled rate of change. The decreasing forward current passes through zero and becomes negative before the SCR ceases to conduct and blocks the reverse voltage impressed on the device by the circuit. The **turn-off time** $t_q$ is measured from the time at which the decreasing forward current $I_T$ passes through zero to the point at which the reverse voltage blocked by the SCR passes through zero and becomes positive, as shown in Fig. 260.

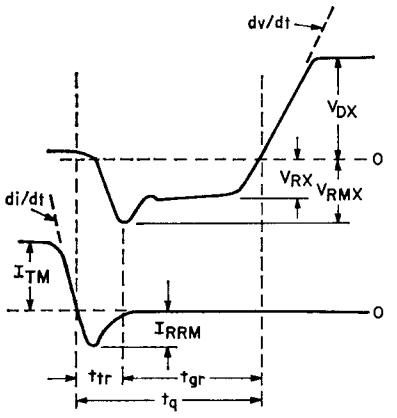



Figure 260. Circuit-commutated turn-off voltage and current waveforms for an SCR.

After forward conduction, the reverse current in the circuit will continue to flow through the SCR until a depletion layer has developed across the reverse-blocking junction. The reverse current reaches a peak value ($I_{RRM}$) and then starts to decay to zero, as shown in Fig. 261. Before the reverse current starts to decay, the rate of change of this current ($-di/dt$) is controlled by the circuit, and a positive voltage is maintained across the terminals by the stored charges in the SCR. During decay, the rate of change of the reverse current is mainly controlled by the SCR, and reverse voltage increases across the cathode-to-anode terminals until the full reverse-blocking voltage impressed on the SCR by the circuit is achieved. This time interval, which is designated $t_{rr}$ in Fig. 261, is referred to as the **reverse-recovery time.** The definition of the reverse-recovery time for an SCR is the same as for rectifier diodes (refer to section on **Silicon Rectifiers**).

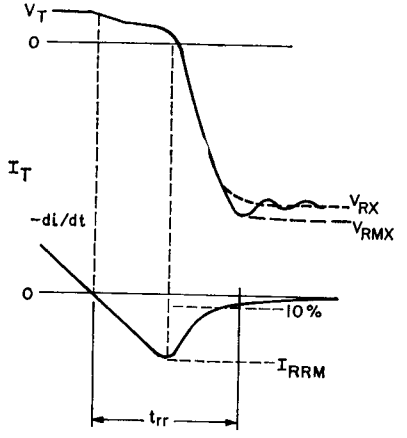



Figure 261. SCR voltage and current waveforms during the recovery of the reverse-blocking junction.

The decay of the reverse current from the peak value in most SCR's is fast and produces an under-damped oscillation superimposed on the reverse blocking voltage. The amplitude and frequency of this oscillation depend mainly on the "snap-off" characteristics of the reverse recovery current, the reverse-blocking junction capacitance of the SCR, the wire or circuit inductance, and stray circuit capacitance.

The turn-off time presented to the SCR by the circuit is referred to as the **circuit turn-off time.** For successful operation of the SCR in the circuit, the circuit turn-off time under all operating conditions should be greater than the device turn-off time. The SCR turn-off time is measured by reducing the circuit turn-off time to a minimum value at which the reapplied forward voltage is still blocked.

**Dependence of SCR Turn-off Time on Operating Conditions**— For a given SCR, the turn-off time varies significantly with different waveforms and temperatures presented to the device. For reliable operation, it is essential

during circuit design that the variations in turn-off time as a function of operating conditions be taken into account. With the proper anticipation of the operating conditions and SCR turn-off time, the circuit turn-off time may be selected without excessive margin that would result in poor utilization efficiency of the SCR.

**Temperature:** Among all the parameters, temperature has the greatest effect on the turn-off time. The turn-off time of the SCR increases with increasing junction temperature. Fig. 262 shows turn-off time as a function

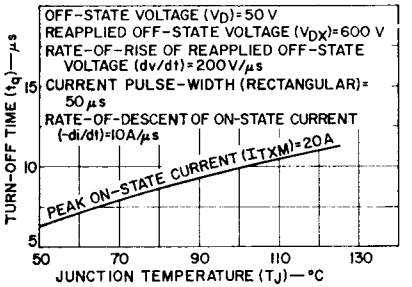



Figure 262. Typical variation of turn-off time with junction temperature (rectangular pulse).

of junction temperature for a typical RCA SCR. Turn-off time is normally measured at elevated temperatures.

**On-state current:** An increase in on-state current causes a corresponding increase in the turn-off time. The effect of the on-state current on turn-off time is somewhat greater at higher temperatures, as shown in Fig. 263. Turn-off time is also affected by the

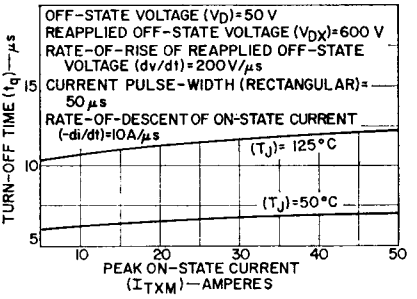



Figure 263. Typical variation of turn-off time with on-state current.

rate of change of on-state current prior to turn-off. In other words, the SCR seems to "remember" the on-state current history for several microseconds prior to turn-off. For example, the turn-off time measured for on-state current waveforms such as (a) and (b) in Fig. 264 may be shorter than that measured for the waveform (c). The prior-to-turn-off history of the on-state current (c) in Fig. 264 is more severe than that of the currents (a) and (b). The sensitivity of the SCR's to on-state current history varies from device to device.

**Magnitude and rate of rise of reapplied forward voltage:** If the rate of rise (dv/dt) of the reapplied forward blocking voltage is held constant, the turn-off

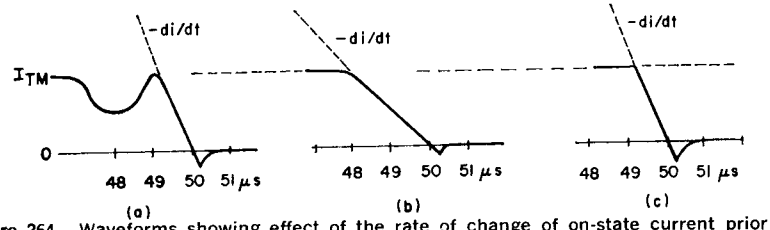



Figure 264. Waveforms showing effect of the rate of change of on-state current prior to turnoff on turnoff time.

time increases with increasing forward blocking voltage. The rate of increase of turn-off time as a function of reapplied forward-blocking voltage becomes greater at higher dv/dt values, as shown in Fig. 265. The SCR turn-off time is affected to a greater degree by variation in the dv/dt of the reapplied forward blocking voltage. The effect of increased dv/dt on the device turn-off time may be off-set by the application of negative bias to the gate of the SCR.

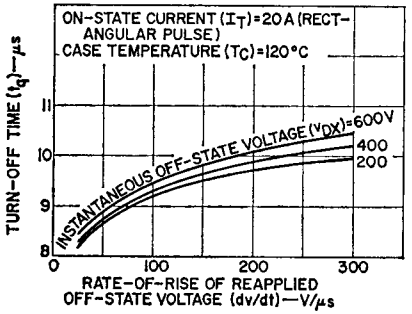



Figure 265. Typical variation of turn-off time with rate of rise of reapplied off-state voltage (rectangular pulse).

**Negative gate bias:** Normally, during the turn-off time measurement, the gate of the SCR is connected to zero voltages which is applied through a specified gate resistor. If a negative bias is applied to the gate, the turn-off time of the SCR may be reduced (as shown in Fig. 266), and the dv/dt capability improved. The effect of negative bias on turn-off time is more pronounced at higher junction temperatures. At higher bias voltages, the effectiveness of the negative bias slowly diminishes.

Fig. 266 shows that the dv/dt capability of the SCR can be improved greatly by the application of a few volts of negative bias

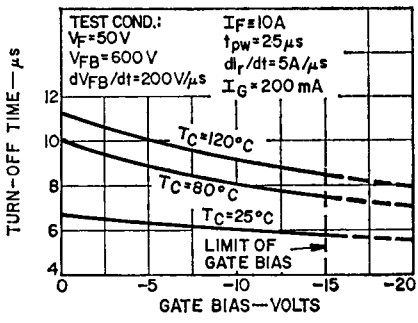



Figure 266. Typical variation of turn-off time as a function of negative gate bias.

to the gate during the rise of the reapplied forward blocking voltage. The negative voltage reverse-biases the gate-cathode junction and provides a path for the displacement current that charges the forward-blocking junction capacitance. Without the negative bias, the displacement current flows through the shorted-emitter resistor (and through the gate-cathode junction). If the magnitude of this current is too high (excessive amount of dv/dt stress), the forward bias on the gate-cathode junction will turn on the SCR. In a reliable design, the amplitude and duration of the negative bias should be selected so that it does not exceed the gate-to-cathode reverse breakdown voltage and the maximum allowable gate power dissipation.

**Reverse blocking voltage:** The effect of the reverse voltage impressed on the anode-to-cathode terminals on the turn-off time becomes minor above a certain voltage level. Normally, the turn-off time decreases with increasing reverse voltage.

**Concentrated turn-on losses:** If the SCR is stressed beyond its capability during turn on, some hot spots may be generated in the

device, as explained in the section on **Critical Rate of Rise of On-State Circuit.** If the rate at which conduction spreads in the SCR is slow compared to the rate at which the load current rises, an excessive amount of energy is dissipated in a small volume of the pellet. The temperature in this small volume then increases to unusually high magnitudes. This localized high temperature ("hot spot") increases the turn-off time of the device. The interaction of the turn-on capability of the device with the turn-off capability may be indicated by the appropriate selection of test conditions. A narrow pulse width with high peak current at rated voltage maximizes the turn-on stresses on the device. The narrow pulse width also prevents the hot spots from cooling down before the blocking voltage is reapplied, as shown in Fig. 267.

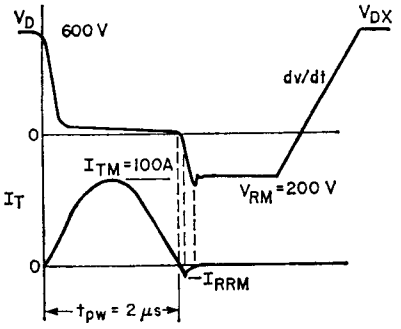



Figure 267. Waveforms used to test the turn-off time of an SCR with severe turn-off stresses.

For applications in which turn-on stresses are not encountered, the turn-off time of the device may be measured with a wide square wave current pulse that turns on the SCR from a low supply voltage. The low voltage will minimize the turn-on dissipation, and the wide current pulse allows time for any "hot spots" to cool. The waveforms shown in Fig. 268 illustrate this condition.

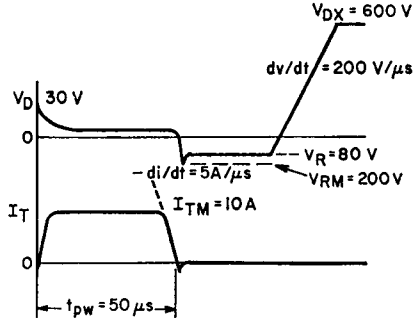
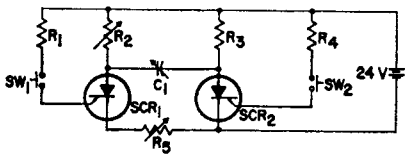



Figure 268. Waveforms used to test the turn-off time for an SCR for wide square-wave applications and applications in which turn-on stresses are negligible.

**Turn-off Time Test Circuit**— Because the turn-off time of an SCR depends upon a number of circuit parameters, the manufacturer's turn-off time specification is meaningful only if these critical parameters are listed and the test circuit used for the measurement is indicated.

Fig. 269 shows a simple test circuit used to measure turn-off time. The circuit subjects the SCR to current and voltage waveforms similar to those encountered in most typical applications. In the circuit diagram, $SCR_1$ is the device under test. Initially, both SCR's are in the off-state; push-button switch $SW_1$ is momentarily closed to start the test. This action turns on $SCR_1$ and load current flows through this SCR and resistor $R_2$. Capacitor $C_1$ charges through resistor $R_3$ to the voltage developed across $R_2$. If the second push-button switch $SW_2$ is then closed, $SCR_2$ is turned on. $SCR_1$ is then reverse-biased by the voltage across capacitor $C_1$. The discharge of this capacitor causes




$R_1$, $R_4$ = 100 ohms
$R_2$ = variable resistor, 0.7 to 50 ohms
$R_3$ = 5000 ohms
$R_5$ = variable resistor, 0.1 to 1 ohm
$C_1$ = variable capacitor, 0.1 to 1 μF, 150 V
$SCR_1$ = SCR under test
$SCR_2$ = RCA-40378

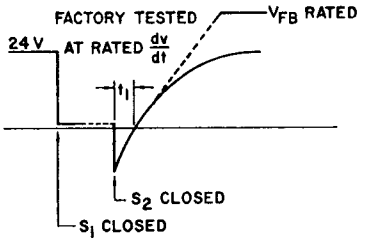



Figure 269. Test circuit and voltage waveforms used to determine turn-off times of thyristors.

a short pulse of reverse current to flow through $SCR_1$ until this device recovers its reverse-blocking capability. At some time $t_1$, the anode-to-cathode voltage of $SCR_1$ passes through zero and starts to build up in a forward direction at a rate dependent upon the time constant of $C_1$ and $R_2$. The peak value of the reverse current during the recovery period can be controlled by adjustment of potentiometer $R_5$. If the turn-off time of $SCR_1$ is less than the time $t_1$, the device will turn off. The turn-off interval $t_1$ can be measured by observation of the anode-to-cathode voltage of $SCR_1$ with a high-speed oscilloscope. A typical waveform is shown in Fig. 269.

## Commutating Capability (of Triacs)

As explained in the preceding section, turn-off times are not associated with triacs because of the physical structure of the device. The ability of a triac to commutate a fixed value of current under specified conditions, however, is an important circuit design consideration. This ability is characterized as the **critical rate of rise of commutation voltage**, or more simply as the **commutating dv/dt capability** of the device.

In ac power-control applications, a triac must switch from the conducting state to the blocking state at each zero-current point, or twice each cycle, of the applied ac power. This action is called commutation. If the triac fails to block the circuit voltage (turn off) following the zero-current point, this action is not damaging to the triac, but control of the load power is lost. Commutation for resistive loading presents no special problems because the voltage and current are essentially in phase. For inductive loading, however, the current lags the voltage so that, following the zero-current point, an applied voltage opposite to the current and equal to the peak of the ac line voltage occurs across the thyristor. The maximum rate of rise of this voltage which can be blocked without the triac reverting to the on state is termed the critical rate of rise of commutation voltage, or the commutating dv/dt capability, of the triac.

SCR's do not experience commutation limitations because turn-on is not possible for the polarity of voltage opposite to current flow.

The commutating dv/dt is a major operating characteristic used to describe the performance capability of a triac. The charac-

teristic can be more easily understood if the triac pellet, shown in Fig. 270, is considered to be divided into two halves. One half conducts current in one direction, the other half conducts in the opposite direction. The main

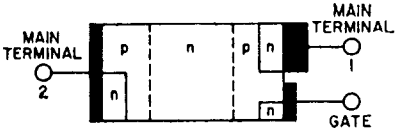



Figure 270. Junction diagram for a triac pellet.

blocking junctions and a lightly doped n-type base region in which charge can be stored are common to both halves of the triac pellet. (The base region is the section shown between the dotted lines in Fig. 106.)

Charge is stored in the base when current is conducted in either direction. The amount of charge stored at the end of each half-cycle of conduction depends on the commutating di/dt, i.e., the rate of decrease of load current as commutation is approached. The junction capacitance of the triac at commutation is a function of the remaining charge at that time. The greater the di/dt, the more remaining charge, and the greater the junction capacitance. When the voltage changes direction, the remaining charge diffuses into the opposite half of the triac structure. The rate of rise of this voltage (commutating dv/dt) in conjunction with the junction capacitance results in a current flow which, if large enough, can cause the triac to revert to the conducting state in the absence of a gate signal.

The commutating dv/dt capability is specified in volts per microsecond for the following conditions:

1. the maximum rated on-state current [$I_{T(RMS)}$];
2. the maximum case temperature for the rated value of on-state current;
3. the maximum rated off-state voltage ($V_{DROM}$);
4. the maximum commutating di/dt (where $di/dt = I_{pk} \sin \omega t$ and $\omega = 2\pi f$).

Commutating dv/dt can be understood as a function of a practical circuit and related to the above conditions by consideration of a simple triac circuit that consists of an inductive load such as a motor. This circuit is shown in Fig. 271, together with waveforms for the circuit. If a gate signal that allows continuous conduction is applied to the triac, the load current ($I_L$) lags the line voltage ($V_{LINE}$) by approximately 90 degrees for the inductive load. During the triac on-state conduction, the voltage across the triac ($V_T$) is in phase with the line current $I_L$ and typically will have an amplitude of $\pm$ 1.5 volts.

When the gate signal is removed, the triac begins to commutate off near the end of the half-cycle, when the load current drops to a value below the holding current ($I_H$) of the triac. At the instant the triac commutates off, the voltage ($V_T$) reverses direction and climbs to the peak of the line voltage ($V_{Line}$), its rate of rise (commutating dv/dt) and overshoot being a characteristic of the circuit components. After the triac successfully commutates off, the voltage ($V_T$) is in phase with the line voltage. If the commutating dv/dt of the circuit is greater than the commutating dv/dt capability of the triac, the

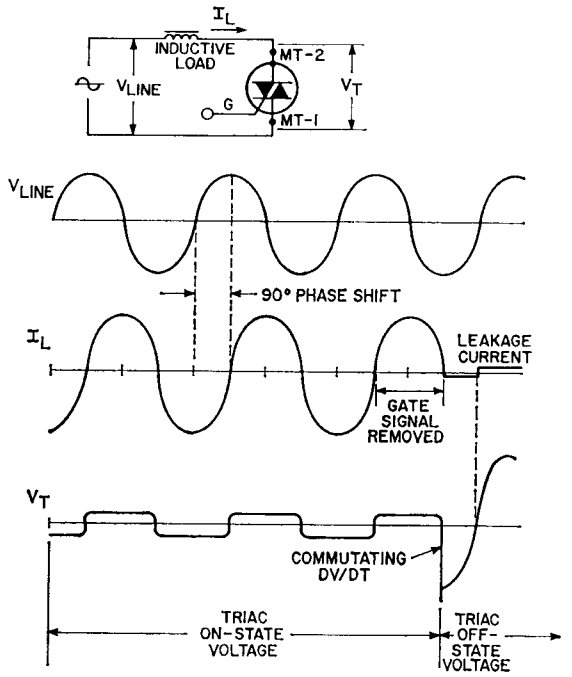



Figure 271. Waveforms for simple triac circuit with inductive load.

triac does not turn off, but reverts to the on-state. With no gate signal applied, the triac again attempts to turn off in the next half-cycle (opposite polarity). If it succeeds, it remains off; if its capability is again exceeded, it remains on until the circuit power is interrupted.

In Fig. 272, the last half-cycle of triac conduction current ($I_L$) has been superimposed on the commutating dv/dt capability. This diagram indicates that the circuit capability exceeds the device capability; therefore, the triac will not commutate off. An additional "snubber network", consisting of a resistor and a capacitor, placed across the main terminals of the triac will reduce the commutating dv/dt of the circuit to within the capability of the device, and thus allow the triac to commutate off. Typical values for the "snubber network" are 0.1 microfarad and 100 ohms; however, these values will vary depending upon the load and rating of the triac. The resistor should be as large as possible to minimize overvoltage ringing and to protect against high capacitor-discharge currents through the triac.

Fig. 273 indicates how the commutating dv/dt capability of a triac depends on current and frequency. A particular triac has a specific commutating dv/dt capability at the rated 60-Hz on-state current. If this 60-Hz on-state current is reduced (dashed-line), then its associated commutating dv/dt capability is increased. It should be noted that although the

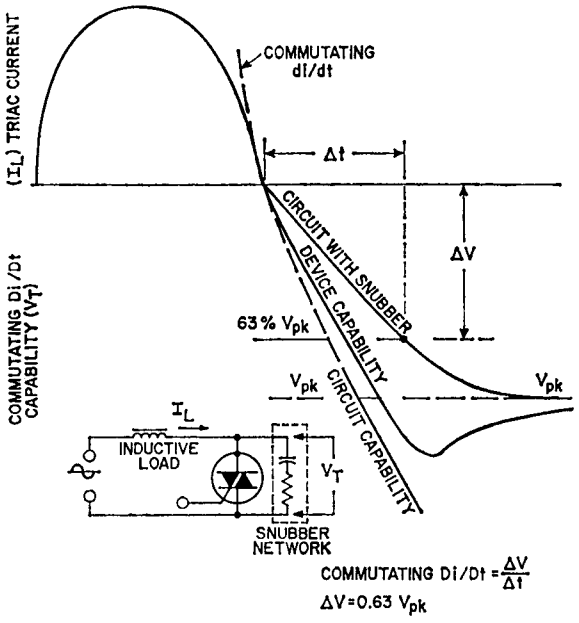



Figure 272. Use of snubber network to reduce commutating dv/dt of circuit.

sine-wave current is decreased in magnitude, the commutating di/dt is also decreased. For a 400-Hz on-state current of the same magnitude, it is evident that the commutating di/dt is much greater than at 60 Hz and, therefore, the commutating dv/dt capability is greatly reduced. These relationships indicate that a triac capable of 400-Hz operation must have an extremely high commutating capability. RCA offers a complete line of triacs rated for 400-Hz operation. Applications of such devices are described in the section on **Thyristor AC Power Controls**.

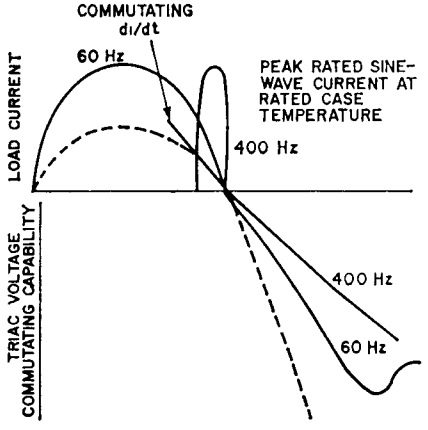



Figure 273. Dependence of triac commutating capability on current and frequency.

It should be evident that 400 Hz is not an upper limit on frequency capability for triacs; 400 Hz is a characterization point simply because it is a standard operating frequency. Figs. 273 and 274 indicate how the frequency capability of a typical RCA 400-Hz triac can be increased. Fig. 274 shows that reduction of load current increases frequency capability. Maximum rated junction temperature and minimum rated commutating dv/dt are held constant for this

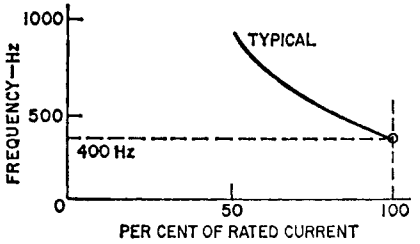



Figure 274. Frequency capability of a 400-Hz triac as a function of load current.

test of capability. Fig. 275 shows the effects of junction temperature on frequency capability. For this test, rated current and minimum rated dv/dt are held constant. Therefore, if a typical 400-Hz triac is used at less than its maximum rated junction temperature and less than its rated current, its frequency capability is greater enhanced.

It is apparent from the preceding discussions that the relationship between commutating dv/dt capability and the commutating di/dt is an important circuit-design consideration. Fig. 276 shows curves of this relationship for RCA 25-ampere and 40-ampere triacs. As a frame of reference, Table XIX shows the di/dt values for sine-wave currents of 25 and 40 amperes rms at frequency of 50, 60, and 400 Hz.

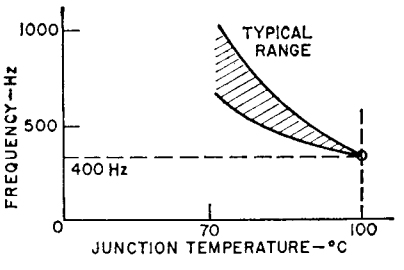



Figure 275. Frequency capability of a 400-Hz triac as a function of junction temperature.

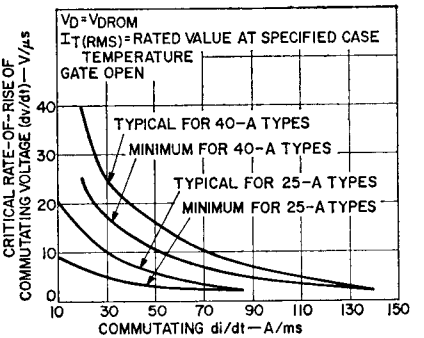



Figure 276. Commutating voltage as a function of commutating current.

One other factor that greatly affects commutating capability is temperature. All commutating characteristic data is specified for maximum operating case temperature at maximum rated steady-state current. If the operating case temperature is below the rated value, the commutating capability is increased.

Table XIX—Commutating di/dt for Various Currents and Frequencies

| Commutating di/dt A/ms | Sine Wave Amperes RMS | Operating Frequency HERTZ |
|---|---|---|
| 141 | 40 | 400 |
| 88 | 25 | 400 |
| 21.1 | 40 | 60 |
| 13.2 | 25 | 60 |
| 17.6 | 40 | 50 |
| 11.0 | 25 | 50 |

# Power Hybrid Circuits

A hybrid circuit is essentially a functional combination of solid-state devices and associated passive elements (resistors and capacitors) within a single package. A simple power hybrid circuit may contain several power transistors and resistors connected to form a high-current switch. A more complex power hybrid circuit, such as a high-power voltage regulator, may consist of a number of power transistors and resistors, one or more diodes and capacitors, and an integrated circuit, all mounted in one compact package.

The primary advantage of power hybrid circuits to system designers is cost effectiveness. In addition, the substitution of one functional package for a number of discrete elements usually results in smaller size, lighter weight, and improved ruggedness, reliability, and performance.

Power hybrid circuits are used in a wide variety of applications. In high-speed printers, high-gain Darlington amplifiers are used for hammer-driver and stepper-motor applications. Voltage regulators containing foldback protection and crow-bar driver load-fault protection are used in computer power supplies. Linear power amplifiers are used in such diverse areas as audio systems, servo controls, linear motor controls, sonar, and ultrasonics. High-power arrays are used as power switches, inverters, bridge circuits, motor controls, and shunt regulators.

## FABRICATION

Fabrication of power hybrid circuits usually involves the use of thick-film chip and wire techniques and the use of solid-state power devices in subassembly (building-block) form.

A typical thick-film circuit is shown in Fig. 277. The fabrica-

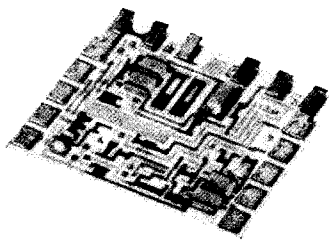


Figure 277. A typical hybrid circuit that contains active-device and capacitor chips and trimmed thick-film resistors.

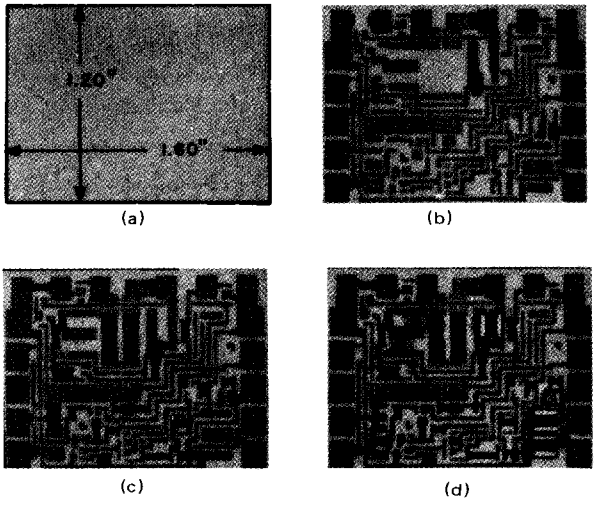



Figure 278. Steps in fabrication of a thick-film circuit: (a) blank substrate; (b) conductor pattern; (c) resistors printed on circuit; (d) semiconductor and capacitor chips added.

tion sequence starts with a blank ceramic substrate, as shown in Fig. 278(a). In the sequence of steps shown in (b), (c), and (d), conductor patterns are printed and fired, resistor inks are printed and fired, and all semiconductor and capacitor chips are then added. Each resistor is trimmed to the specified value by means of an abrasive sandblast prior to chip mounting. The 22 chips in the circuit (seven capacitors, six diodes, and nine transistors) are mounted directly on the printed circuit with an electrically and thermally conductive epoxy. Electrical connections to the tops of the semiconductor chips are provided by 0.002-inch-diameter aluminum wires that are ultrasonically bonded to the contact points.

Chips or pellets used in hybrid circuits may be mounted directly on the circuit substrate, as described above, or may first be assembled into "building blocks" for ease of handling. A "building block" is a complete functional device fabricated on the smallest possible substrate compatible with chip size, terminal pins, and thermal considerations.

Power-resistor building blocks may also be fabricated by use of thick-film substrate technology. The building blocks desired for a particular circuit are high-temperature soldered to a common baseplate, and interconnections are then added to provide the required circuit functions. Final sealing of the circuit can be accomplished by either plastic encapsulation or hermetic sealing techniques. Fig. 279 illustrates the steps in the construction of a typical power-device building block.

These two basic fabrication methods are often combined in the construction of specific circuits. A typical power hybrid circuit is described in the following paragraphs.

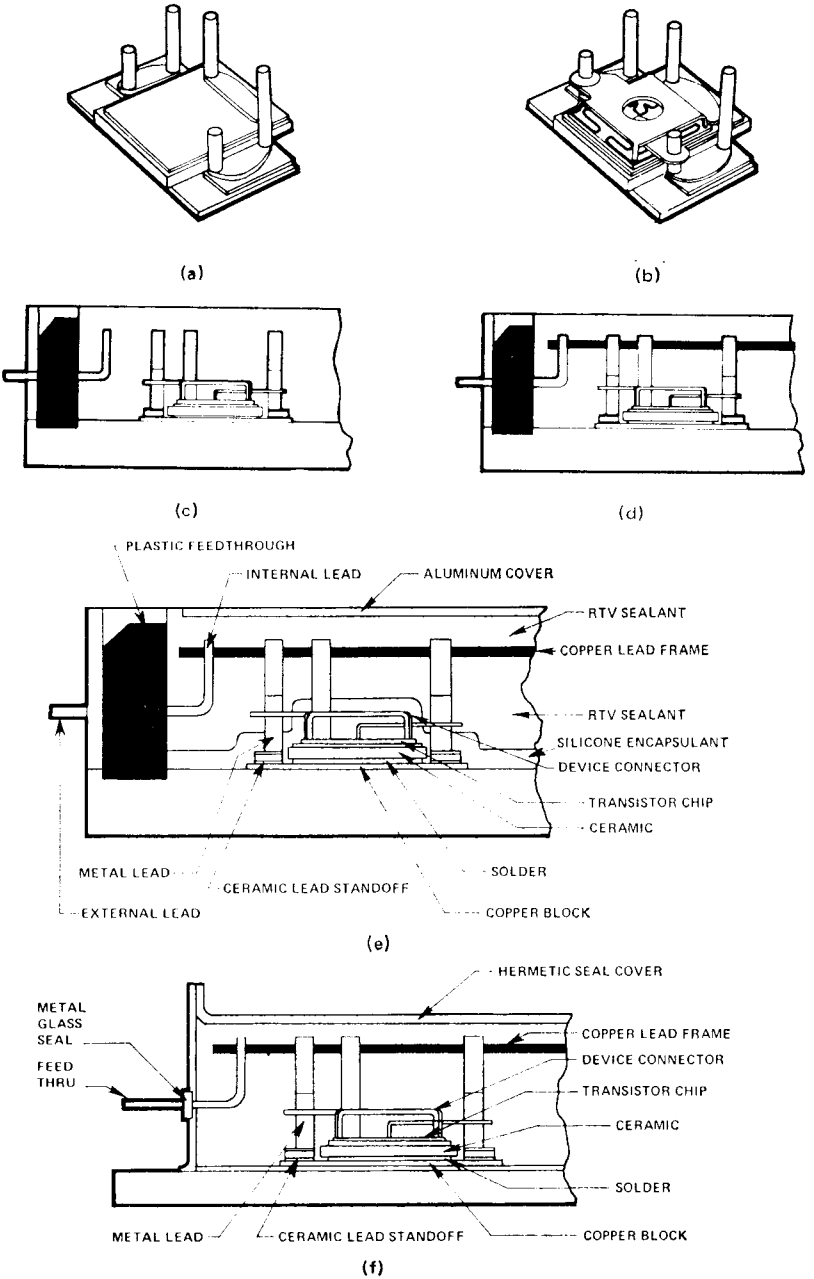



Figure 279. Steps in the assembly of power-device building blocks: (a) basic building-block structure; (b) transistor chip added to basic block; (c) building block mounted on baseplate in package; (d) building blocks interconnected; (e) nonhermetic package; (f) hermetic package.

## THE HC1000 LINEAR POWER AMPLIFIER

The HC1000 hybrid power amplifier consists of two sections. The input-and-driver section, which includes a load-line-limiting protection circuit, drives an output section that consists of two power transistors operated in a quasi-complementary-symmetry class B mode. A diode across each output transistor protects the amplifier from inductive voltage spikes generated by transformer-coupled and motor loads.

The construction of the HC1000 uses a combination of thick-film and building-block techniques. The driver section is the thick-film circuit described previously and shown in Fig. 277. After fabrication, the driver circuit is tested functionally for distortion, output power, sensitivity, load-line limiting, and balance. Units that pass these tests are then ready for assembly to the output section.

The output section, shown in Fig. 280, consists of two output power-transistor chips mounted on copper heat-spreader blocks, and two rectifier chips. The output devices are electrically isolated from the metal baseplate by the output ceramic substrate. A dual-purpose lead frame makes connections to the output-section devices and acts as an interconnect to the driver substrate.

At final assembly, a tested driver substrate is bonded to the baseplate of a pretested output with a thermally conductive adhesive, and the external leads are clamped into position on their respective metallized pads. After electrical testing of the complete amplifier circuit, all connections are soldered in place by the reflow process. Fig. 281 shows the completed assembly ready for encapsulation and shearing of the leads.

Prior to plastic encapsulation, the amplifier is coated with a resilient resin that serves as a buffer for stresses induced by differences in thermal-expansion coefficients of the various materials used. The plastic material used for the molded encapsulation forms a very tenacious ad-

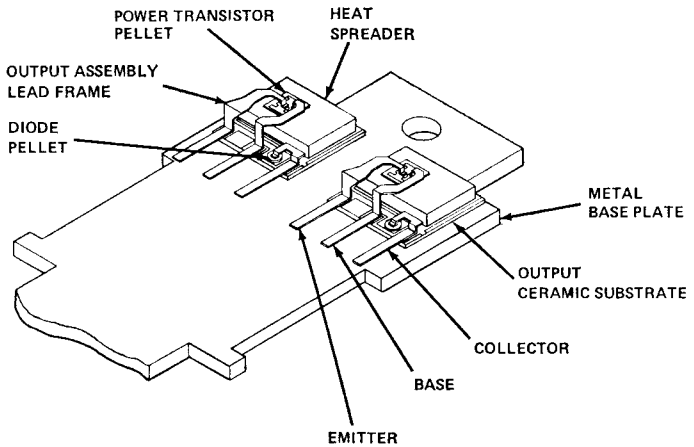



Figure 280. HC1000 output transistors and diodes mounted on baseplate.

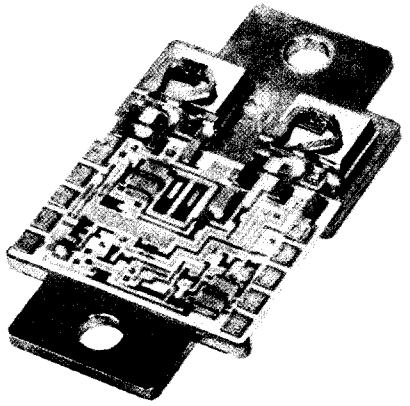


Figure 281. Complete HC1000 assembly before encapsulation.

hesive bond to the baseplate and leads, and virtually eliminates the possibility of moisture paths along the plastic-metal interface. Structures of this type have been subjected to several thousand electrically induced thermal cycles at maximum power dissipation without failure.

A photograph of the completed HC1000 hybrid amplifier is shown in Fig. 282. Because the baseplate is electrically isolated from the circuit, it can be bolted directly to a heat sink.

The HC1000 is designed for use in audio-amplifier systems, servo amplifiers, deflection amplifiers, voltage regulators, and driven inverters.

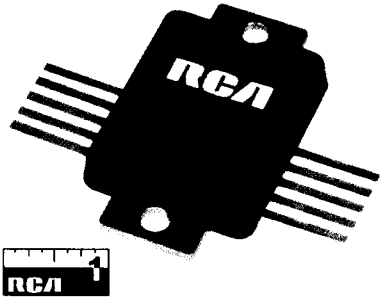



Figure 282. Photo of complete HC1000 assembly.

## OPERATING CONSIDERATIONS

Power hybrid circuits are thermally, mechanically, and electrically rugged devices. They can be designed to provide current capabilities of several hundred amperes and power-handling capabilities approaching a kilowatt. Protection circuits can be built in as part of the circuitry, and provision can be made for use of external circuit elements to extend the device capability and applications flexibility.

Because of the size and shape of the HC1000 package, flat-back-type heat sinks are recommended to simplify assembly. A silicone grease or heat-sink compound, such as Dow Corning 340 or equivalent, should be used in mounting the unit to the heat sink.

For power hybrid circuits that use tinned ribbon leads, such as the HC1000 (shown in Fig. 282), the simplest form of connection is to solder wires to the leads and insulate each connection with a sleeve. If the leads require shaping, the bending practice illustrated in Fig. 283 is recommended.

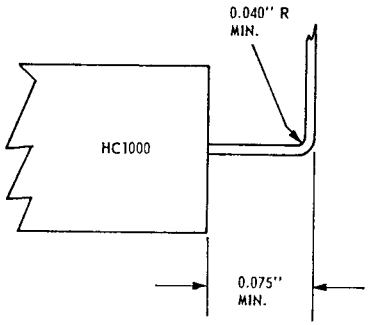



Figure 283. Recommended lead-bending specifications for power hybrid circuits that use ribbon leads.

# DC Power Supplies

A dc power supply converts the power from an ac line to direct current and a steady voltage of a desired value. The ac input power is first rectified to provide a pulsating dc, and then filtered to produce a smooth voltage. Finally, the voltage may be regulated to maintain a constant output level despite fluctuations in the power-line voltage or circuit loading. The rectification, filtering, and regulation steps in a dc power supply are illustrated in Fig. 284.

A dc power supply need not include all three of the elements shown in Fig. 284. Electroplating supplies and battery chargers require only rectification of the ac, and broadcast receivers and phonograph amplifiers need only the rectifier and filter steps. However circuits such as oscillators, high-gain ampliers, and low-voltage logic, which have exacting frequency, stability, or output requirements, can be critically affected by variations in dc supply voltages. Therefore, some type of regulation is frequently required to prevent significant changes in the output of a dc power supply as a result of line-voltage fluctuations or variations in circuit loading.

## RECTIFICATION

The most suitable type of rectifier circuit for a particular application depends on the dc voltage and current requirements, the amount of rectifier "ripple" (undesired fluctuation in the dc output caused by an ac component) that can be tolerated in the circuit, and the type of ac power available. Figs. 285 through 291 show seven basic rectifier configurations.

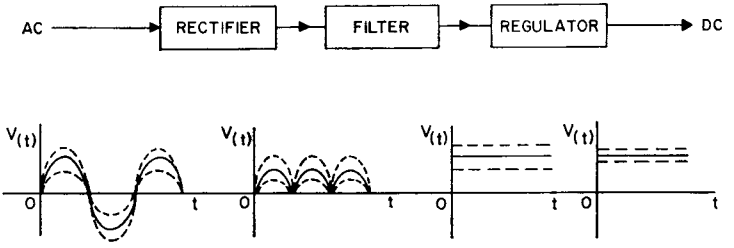



Figure 284. Block diagram of a regulated dc power supply. The waveforms show the effects of rectification, filtering, and regulation. (Dashed lines indicate voltage fluctuations as a result of input variations)

(Filters used to smooth the rectifier output are not shown for each circuit, but are discussed later.) These illustrations also include the output-voltage waveforms for the various circuits and the current waveforms for each individual rectifier in the circuits. Ideally, the voltage waveform should be as flat as possible (i.e., approaching almost pure dc). A flat curve indicates a peak-to-average voltage ratio of one.

The single-phase half-wave circuit shown in Fig. 285 delivers only one phase of current for each cycle of ac input voltage. As shown by the current waveform, the

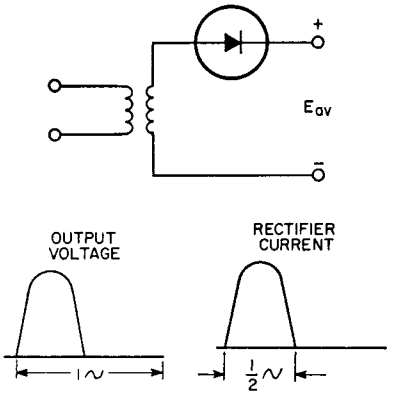



Figure 285. Single-phase half-wave circuit.

single rectifier conducts the entire current flow. This type of circuit contains a very high percentage of output ripple.

Fig. 286 shows a single-phase full-wave circuit that operates from a center-tapped high-voltage transformer winding. This circuit has a lower peak-to-average voltage ratio than the circuit of Fig. 295 and about 65 per cent less ripple. Only 50 per cent of the total current flows through each rectifier. This type of circuit is

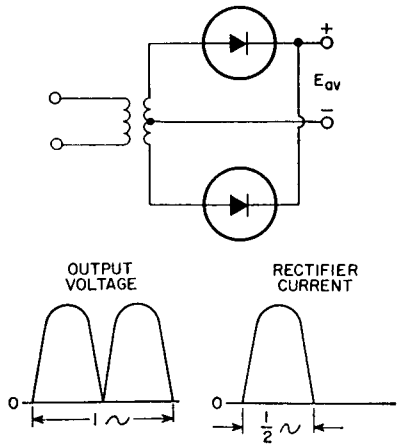



Figure 286. Single-phase full-wave circuit with center-tapped power transformer.

widely used in television receivers and large audio amplifiers.

The single-phase full-wave bridge circuit shown in Fig. 287 uses four rectifiers, and does not require the use of a transformer center-tap. It can be used to supply twice as much output voltage

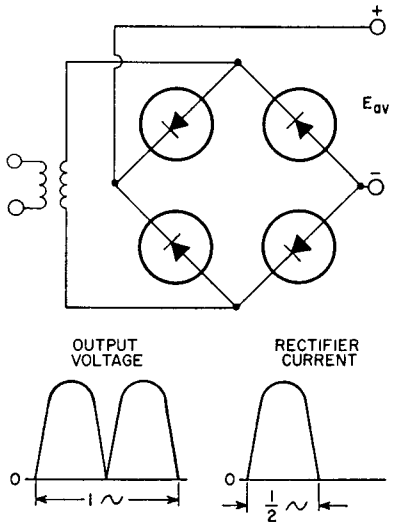



Figure 287. Single-phase full-wave circuit without center-tapped power transformer (i.e., bridge-rectifier circuit).

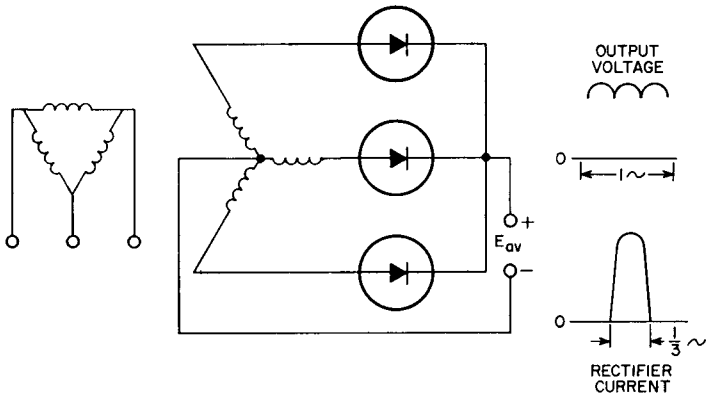



Figure 288. Three-phase "Y" half-wave circuit.

as the circuit of Fig. 286 for the same transformer voltage, or to expose the individual rectifiers to only half as much peak reverse voltage for the same output voltage. Only 50 per cent of the total current flows through each rectifier. This type of circuit is popular in amateur transmitter use.

The three-phase circuits shown in Figs. 288 through 291 are usually found in heavy industrial equipment such as high-power transmitters. The three-phase Y half-wave circuit shown in Fig. 288 uses three rectifiers. This circuit has considerably less ripple than the circuits discussed above. In addition, only one-third of the total output current flows through each rectifier.

Fig. 289 shows a three-phase full-wave bridge circuit which uses six rectifiers. This circuit delivers twice as much voltage output as the circuit of Fig. 288 for the same transformer conditions. In addition, this circuit, as well as those shown in Figs. 290 and 291, has an extremely small percentage of ripple.

In the six-phase "star" circuit shown in Fig. 290, which also uses six rectifiers, the least amount of the total output current (one-sixth) flows through each output

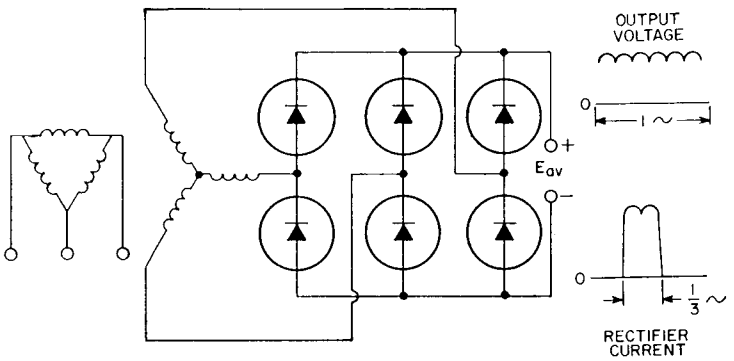



Figure 289. Three-phase "Y" full-wave circuit.

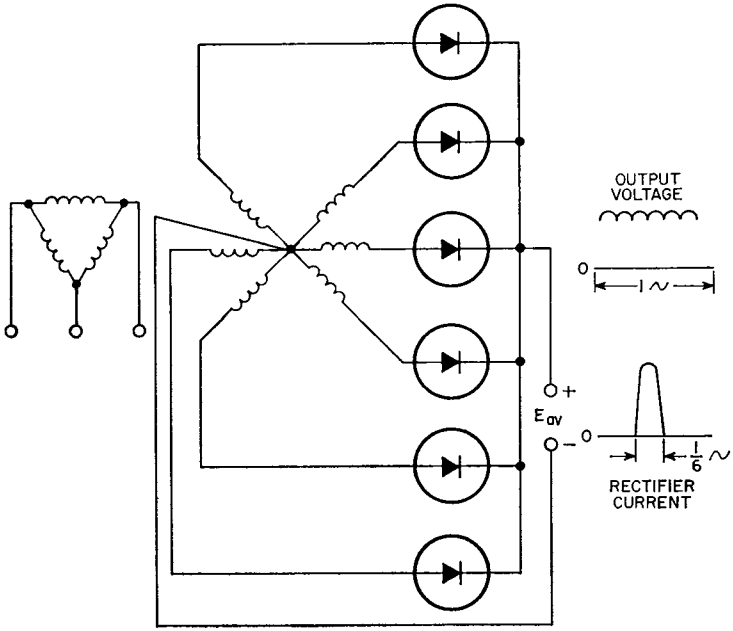



Figure 290. Six-phase "star" circuit.

rectifier. The three-phase double-Y and interphase transformer circuit shown in Fig. 291 uses six half-wave rectifiers in parallel. This arrangement delivers six current pulses per cycle and twice as

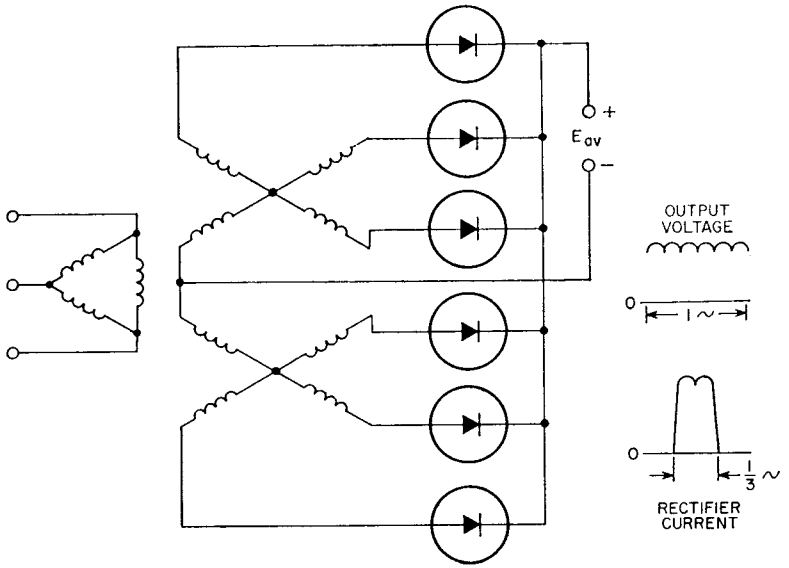



Figure 291. Three-phase "double-Y" and interphase-transformer circuit.

much output current as the circuit shown in Fig. 288.

Table XX lists voltage and current ratios for the circuits shown in Figs. 285 through 291 for resistive or inductive loads. These ratios apply for sinusoidal ac input voltages. It is generally recommended that inductive loads rather than resistive loads be used for filtering of rectifier current, except for the circuit of Fig. 285. Current ratios given for inductive loads apply only when a filter choke is used between the output of the rectifier and any capacitor in the filter circuit. Values shown do not take into consideration voltage drops which occur in the power transformer, the silicon rectifiers, or the filter components under load conditions. When a particular rectifier type has been selected for use in a specific circuit, Table XX can be used to determine the parameters and characteristics of the circuit.

In Table XX, all ratios are shown as functions of either the average output voltage $E_{av}$ or the average dc output current $I_{av}$, both of which are expressed as unity for each circuit. In practical applications, the magnitudes of these average values will, of course, vary for the different circuit configurations.

## FILTERING

Filter circuits are generally used to smooth out the ac ripple in

Table XX—Voltage and Current Ratios for Rectifier Circuits Shown in Figs. 285 Through 291. Fig. 285 Uses a Resistive Load, and Figs. 286 Through 291 an Inductive Load

| CIRCUIT RATIOS | Fig. 285 | Fig. 286 | Fig. 287 | Fig. 288 | Fig. 289 | Fig. 290 | Fig. 291 |
|---|---|---|---|---|---|---|---|
| Output Voltage: | | | | | | | |
| Average | $E_{av}$ | $E_{av}$ | $E_{av}$ | $E_{av}$ | $E_{av}$ | $E_{av}$ | $E_{av}$ |
| Peak (x $E_{av}$) | 3.14 | 1.57 | 1.57 | 1.21 | 1.05 | 1.05 | 1.05 |
| RMS (x $E_{av}$) | 1.57 | 1.11 | 1.11 | 1.02 | 1.00 | 1.00 | 1.00 |
| Ripple (%) | 121 | 48 | 48 | 18.3 | 4.3 | 4.3 | 4.3 |
| Input Voltage (RMS): | | | | | | | |
| Phase (x $E_{av}$) | 2.22 | 1.11* | 1.11 | 0.855• | 0.428• | 0.74• | 0.855• |
| Line-to-Line (x $E_{av}$) | 2.22 | 2.22 | 1.11 | 1.48 | 0.74 | 1.48† | 1.71‡ |
| Average Output (Load) Current | $I_{av}$ | $I_{av}$ | $I_{av}$ | $I_{av}$ | $I_{av}$ | $I_{av}$ | $I_{av}$ |
| RECTIFIER CELL RATIOS | | | | | | | |
| Forward Current: | | | | | | | |
| Average (x $I_{av}$) | 1.00 | 0.5 | 0.5 | 0.333 | 0.333 | 0.167 | 0.167 |
| RMS (x $I_{av}$): | | | | | | | |
| resistive load | 1.57 | 0.785 | 0.785 | 0.587 | 0.579 | 0.409 | 0.293 |
| inductive load | — | 0.707 | 0.707 | 0.578 | 0.578 | 0.408 | 0.289 |
| Peak (x $I_{av}$): | | | | | | | |
| resistive load | 3.14 | 1.57 | 1.57 | 1.21 | 1.05 | 1.05 | 0.525 |
| inductive load | — | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.500 |
| Ratio peak to average: | | | | | | | |
| resistive load | 3.14 | 3.14 | 3.14 | 3.63 | 3.15 | 6.30 | 3.15 |
| inductive load | — | 2.00 | 2.00 | 3.00 | 3.00 | 6.00 | 3.00 |
| Peak Reverse Voltage: | | | | | | | |
| x $E_{av}$ | 3.14 | 3.14 | 1.57 | 2.09 | 1.05 | 2.42 | 2.09 |
| x $E_{rms}$ | 1.41 | 2.82 | 1.41 | 2.45 | 2.45 | 2.83 | 2.45 |

* to center tap • to neutral † maximum value ‡ maximum value, no load

the output of a rectifier circuit. Filters consist of two basic types, inductive "choke" input and capacitive input. Combinations and variations of these types are often used; some typical filter circuits are shown in Fig. 292.

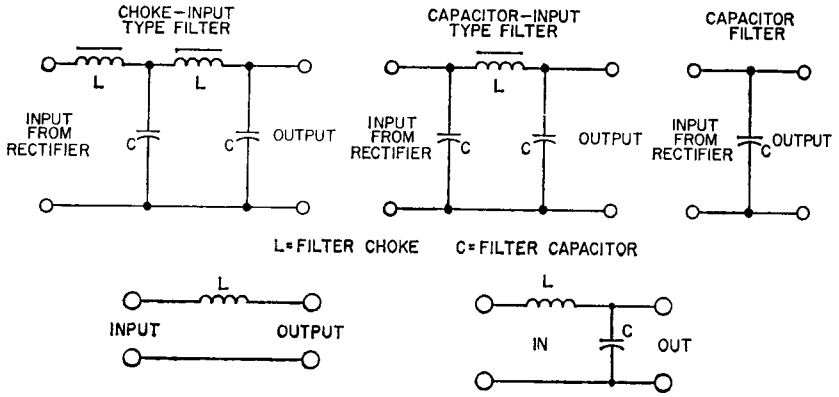



Figure 292. Typical filter circuits.

The simplest of these filtering circuits is the capacitive input. This type of filtering is most often used in low-current circuits in which a fairly large amount of ripple can be tolerated. Such circuits are usually single-phase, half-wave or full-wave. In this type of filter, the capacitor charges up to approximately the peak of the input voltage on each half-cycle that a rectifier conducts. The current into the load is then supplied from the capacitor rather than from the power supply until the point in the next half-cycle when the input voltage again equals the voltage across the capacitor. A rectifier circuit that uses a smoothing capacitor and the voltages involved are shown in Fig. 293.

Higher average dc output voltages and currents can be obtained from this type of circuit by the use of larger capacitors. A larger capacitor also tends to reduce the ripple. However, care must be taken that the capacitor is not so large that excessive peak and rms currents cause overheating of the rectifier. The effects of capacitive loading on a rectifier circuit are discussed in detail in the section on **Capacitive-Load Circuits.**

The next simplest filter is the inductive input filter. This filter

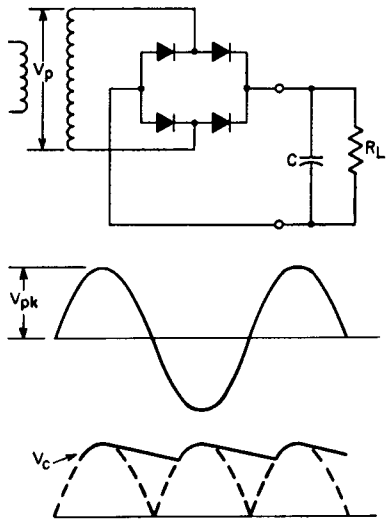



Figure 293. Bridge-rectifier circuit with capacitor input filter.

performs the same function as a capacitive input filter in that it smooths the load current by storing energy during one part of the cycle and releasing it to the load during another part of the cycle. However, the inductor acts in a different way by extending the time during which current is drawn from a rectifier. When a smoothing inductor is used in series with a full-wave rectifier circuit, the conduction period of each rectifier may be extended so that conduction does not stop in one rectifier until the other rectifier starts conducting. As a result of this spreading action, any increase in inductance to reduce ripple results in a decrease in the average output voltage and current.

The smoothing capabilities of capacitors and inductors can be combined as shown in the other filters of Fig. 292 to take advantage of the best feature of each. Filters which provide maximum output and minimum ripple and use reasonably small components can thus be designed.

## CAPACITIVE-LOAD CIRCUITS

When rectifiers are used in circuits with capacitive loads, the rectifier current waveforms may deviate considerably from their true sinusoidal shape. This deviation is most evident for the peak-to-average-current ratio, which is somewhat higher than that for a resistive load. For this reason, capacitive-rating calculations are generally more complicated and time-consuming than those for resistive-load rectifier circuits. However, the simplified rating system described below allows the designer to calculate the characteristics of capacitive-load rectifier circuits quickly and accurately.

Fig. 294 shows typical half-wave and voltage-doubling rectifier circuits that use capacitive loads. In such circuits, the low forward voltage drop of the silicon rectifiers may result in a very high surge of current when the capacitive load is first energized.

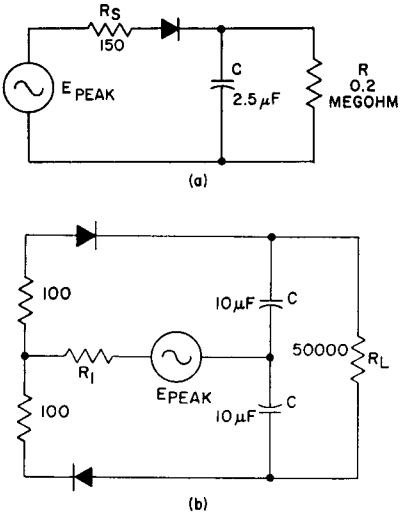



Figure 294. Typical rectifier circuits using capacitive loads: (a) half-wave rectifier circuit; (b) voltage doubler.

Although the generator or source impedance may be high enough to protect the rectifier, additional resistance must be added in some cases. The sum of this resistance plus the source resistance is referred to as the total limiting resistance $R_s$. The magnitude of $R_s$ required for protection of the rectifier may be calculated from surge rating charts such as those shown in Figs. 295 and 296. Each point of these curves defines a surge rating by indicating the maximum time for which the device can safely carry a specific value of rms current.

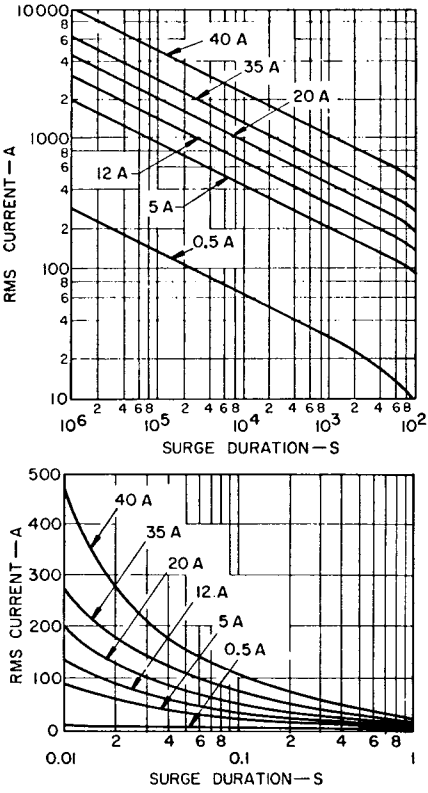



Figure 295. Universal surge rating charts for RCA rectifiers.

With a capacitive load, maximum surge current occurs if the circuit is switched on when the input voltage is near its peak value. When the time constant $R_sC$ of the surge loop is much smaller than the period of the input voltage, the peak current $I_{peak}$ is equal to the peak voltage $E_{peak}$ divided by the limiting resistance $R_s$, and the resulting surge approximates an exponentially decaying current with the time constant $R_sC$.

Surge-current ratings for rectifiers are often given in terms of the rms value of the surge current and the time duration t of the surge. For rating purposes, the surge duration t is defined by the time constant $R_sC$. The rms surge current $I_{rms}$ is then approximated by the following equations:

$$I_{rms} = 0.7\ (E_{peak}C/R_sC)$$

$$= 0.7\ (E_{peak}C/t) \qquad (247)$$

and

$$I_{rms}t = 0.7\ E_{peak}\ C \qquad (248)$$

where $E_{peak}$ and C are the values specified by the circuit design. This equation may then be plotted on the surge-rating chart, which has axes labeled $I_{rms}$ and t. Because $R_sC$ is equal to t, any given value of $R_s$ defines a specific time t, and hence a specific point on the plot of the equation for $I_{rms}t$. However, $R_s$ must be large enough to make this point fall below the rating curve for the rectifier used.

The following example illustrates the use of this simplified procedure for the half-wave rectifier circuit shown in Fig. 294(a), which has a frequency f of 60 Hz and a peak input voltage $E_{peak}$ of 4950 volts. The values shown for $E_{peak}$ and C are substituted in the equation for $I_{rms}t$ as follows:

$$I_{rms}t = 0.7\ (4950)\ (2.5 \times 10^{-6})$$

$$= 0.0086$$

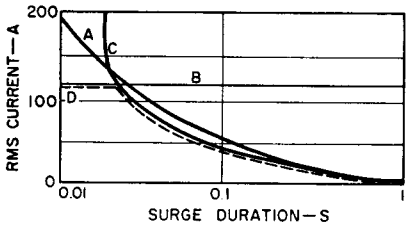



Figure 296. Typical coordination chart for determination of fusing requirements: Curve A—surge rating for 20-ampere rectifier; Curve B—expected surge current in half-wave circuit; Curve C—opening characteristics of protective device; Curve D—resulting surge current in modified circuit.

When this value is plotted on the surge-rating chart of Fig. 297, the resulting line intersects the rectifier rating curve at $3.3 \times 10^{-4}$ second. The minimum limiting resistance which affords adequate surge protection is then calculated as follows:

$$R_s C \geq 3.3 \times 10^{-4}$$

$$R_s \geq \frac{3.3 \times 10^{-4}}{2.5 \times 10^{-6}} = 132 \text{ ohms}$$

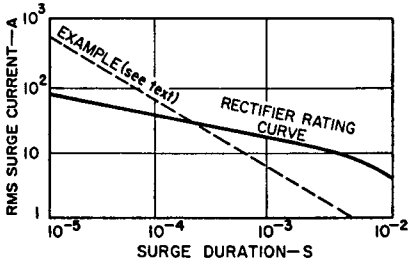



Figure 297. Surge rating chart for stack rectifier CR210.

Therefore the value of 150 ohms shown for $R_S$ in Fig. 294(a) provides adequate surge-current protection for the rectifier.

The design of rectifier circuits having capacitive loads often requires the determination of rectifier current waveforms in terms of average, rms, and peak currents. These waveforms are needed for calculation of circuit parameters, selection of components, and matching of circuit parameters with rectifier ratings. Although actual calculation of rectifier current is a rather lengthy process, the current-relationship charts shown in Figs. 298 and 299 can be used to determine peak or rms current if the average current is known, or vice versa.

The ratios of peak-to-average current ($I_{peak}/I_{av}$) and rms-to-average current ($I_{rms}/I_{av}$) are shown in Fig. 297 as functions of

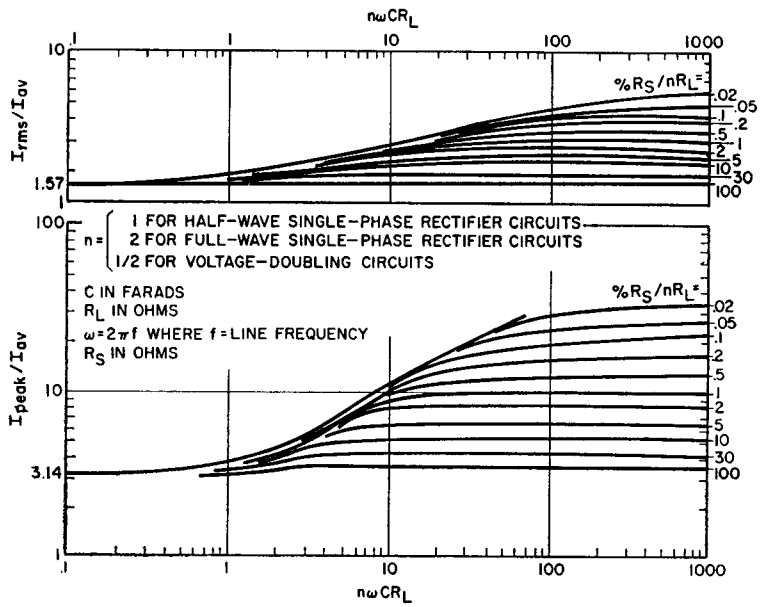



Figure 298. Relationship of peak, average, and rms rectifier currents in capacitor-input circuits.

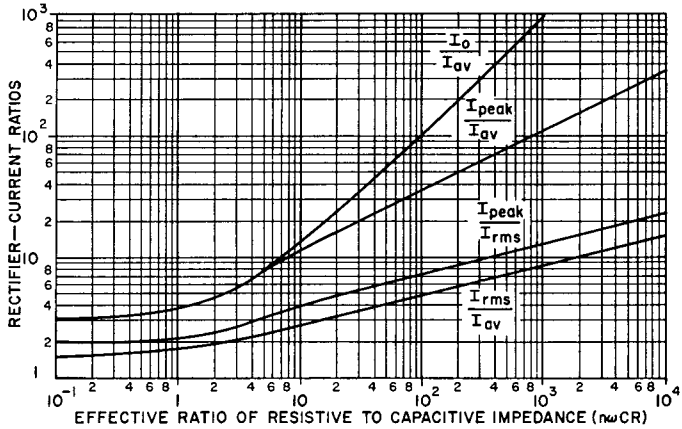



Figure 299. Forward-current ratios for rectifiers in capacitor-input circuits in which $R_s$ is much less than 1/C.

the circuit constants $n\omega CR_L$ and $R_s/nR_L$. The quantity $\omega CR_L$ is the ratio of resistive-to-capacitive reactance in the load, and the quantity $R_s/R_L$ is the ratio of the limiting resistance to the load resistance. The factor n, referred to as the "charge factor," is simply a multiplier which allows the chart to be used for various circuit configurations. The value of n is equal to unity for half-wave circuits, to 0.5 for doubler circuits, and to 2 for full-wave circuits. (These values actually represent the relative quantity of charge delivered to the capacitor on each cycle.)

In many silicon rectifier circuits, $R_s$ may be neglected when compared with the magnitude of $R_L$. In such circuits, the calculation of rectifier currents is simplified by use of Fig. 299, which gives current ratios under the limitation that $R_s/R_L$ approaches zero. Even if this condition is not fully satisfied, the use of Fig. 299 merely indicates a higher peak and higher rms current than will actually flow in the circuit, i.e., the rectifiers will operate more conservatively than calculated. As a result, this simplified solution can be used whenever a rough approximation or a quick check is needed on whether a particular rectifier will fit a specific application. When more exact information is needed, the chart of Fig. 298 should be used.

Average output voltage $E_{av}$ is another important quantity in capacitor-input rectifier circuits because it can be used to determine average output current $I_{av}$. The relationships between input and output voltages for half-wave, voltage-doubler, and full-wave circuits are shown in Figs. 300, 301, and 302, respectively. Fig. 303 shows curves of output ripple voltage (as a percentage of $E_{av}$) for all three types of circuits.

The following example illustrates the use of these curves in rectifier-current calculations. Both exact and approximate solutions are given. For the half-wave circuit of Fig. 294(a), the resistive-to-capacitive reactance $\omega CR_L$ is given by

$$\omega CR_L = 2\pi \times 60 \times 2.5 \times 10^{-6} \times 200{,}000 = 189$$

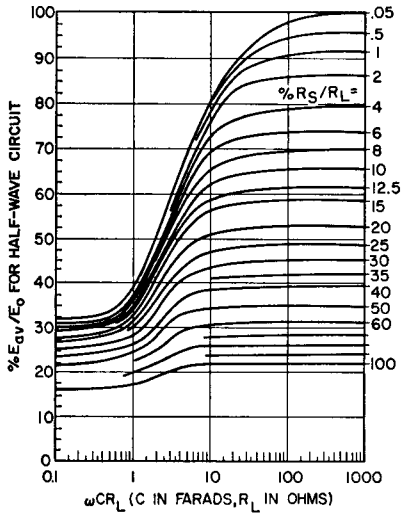



Figure 300. Relationship of applied ac peak voltage to dc output voltage in half-wave capacitor-input circuit.

For an exact solution using Fig. 298, the ratio of $R_S$ to $R_L$ is first calculated as follows:

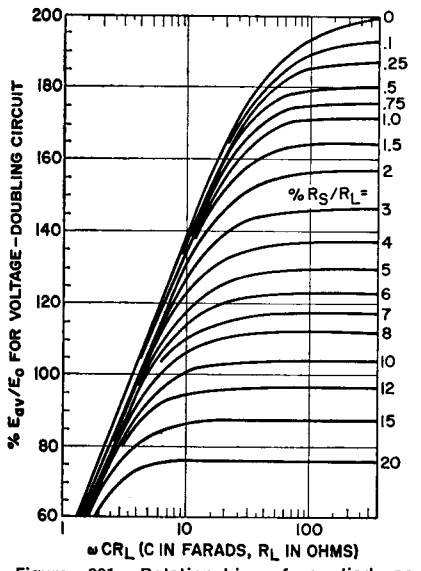



Figure 301. Relationship of applied ac peak voltage to dc output voltage in capacitor-input voltage-doubler circuit.

$$\frac{R_s}{R_L} = \frac{150}{200{,}000} = 0.075$$

The values for $\omega CR_L$ and $R_s/R_L$ are then plotted in Fig. 300 to determine the average output voltage $E_{av}$ and the average output current $I_{av}$ as follows:

$E_{av}/E_{peak} = 98$ per cent
$E_{av} = 0.98 \times 4950 = 4850$ volts
$I_{av} = E_{av}/R_L$
$I_{av} = 4850$ volts/200,000 ohms
$= 24.2$ milliamperes

This value of $I_{av}$ is then substituted in the ratio of $I_{rms}/I_{av}$ obtained from Fig. 298, and the exact value of rms current $I_{rms}$ in the rectifier is determined as follows:

$I_{rms}/I_{av} = 4.4$
$I_{rms} = 4.4 \times 24.2$
$= 107$ milliamperes.

For a simplified solution using Fig. 299, it is assumed that the average output current $I_{av}$ is approximately equal to the peak input voltage $E_{peak}$ divided by the load resistance $R_L$, as follows:

$I_{av} = E_{peak}/R_L$
$I_{av} = 4950/200{,}000$
$= 24.7$ milliamperes

This value of $I_{av}$ is then substituted in the ratio of $I_{rms}/I_{av}$ obtained from Fig. 299, and the approximate rms current is determined, as follows:

$I_{rms}/I_{av} = 5.7$
$I_{rms} = 5.7 \times 24.7$
$= 141$ milliamperes

Current-versus-temperature ratings for rectifiers are usually given in terms of average current

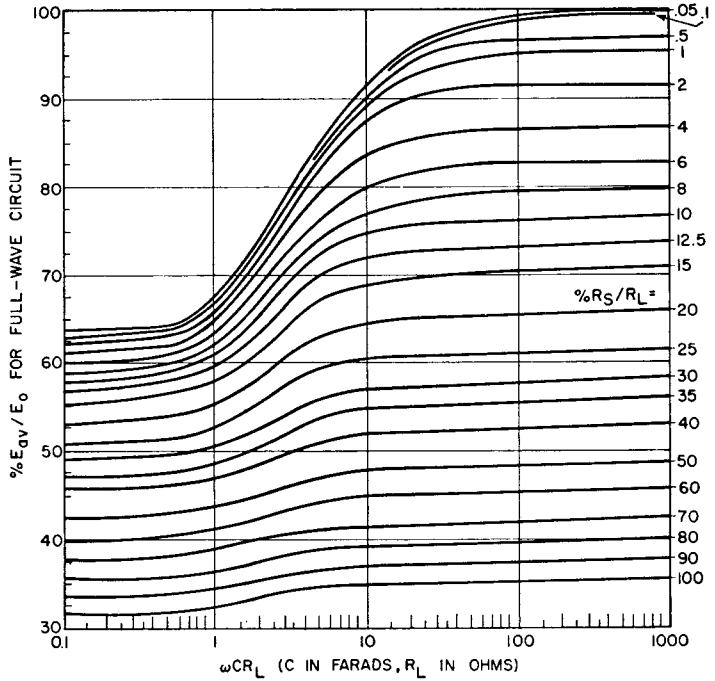



Figure 302. Relationship of applied ac peak voltage to dc output voltage in full-wave capacitor-input circuit.

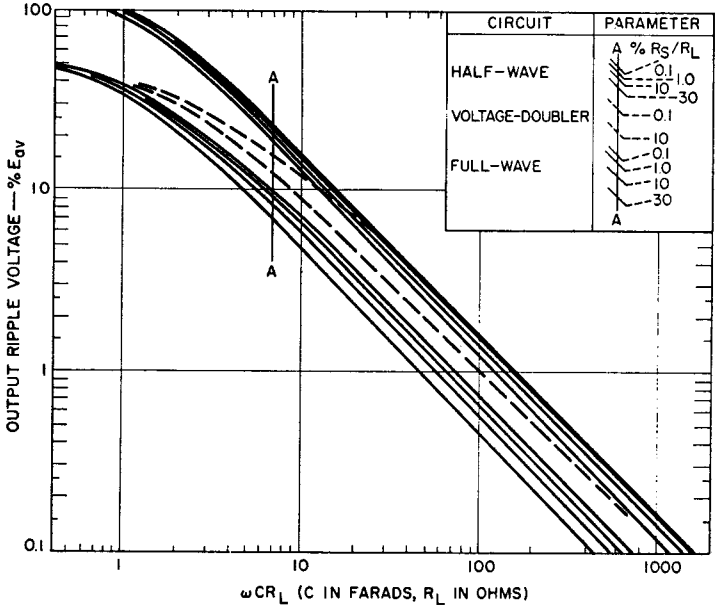



Figure 303. RMS ripple voltage in capacitor-input circuits.

for a resistive load with 60-Hz sinusoidal input voltage. When the ratio of peak-to-average current becomes higher (as with capacitive loads), however, junction heating effects become more and more dependent on rms current rather than average current. Therefore, capacitive-load ratings should be obtained from a curve of rms current as a function of temperature. Because the ratio of rms-to-average current for the rated service is 1.57 (as shown by $I_{rms}/I_{av}$ at low $\omega$CR on Figs. 298 and 299), the current axis of the average-current rating curves for a sinusoidal source and resistive load can be multiplied by 1.57 to convert the curves to rms rating curves. Fig. 304 shows an example of this conversion for RCA stack-rectifier rating curves.

## REGULATION

The regulation of a dc power supply is usually accomplished by some type of feedback circuit that senses any change in the dc output and develops a control signal to cancel this change. As a result, the output is maintained essentially constant. The nature of the control exercised by the feedback circuit (regulator) is determined by the type of circuit arrangement (series or shunt) and the mode of operation of the variable-resistance **pass element,** which is a transistor or an SCR. In a transistor regulator, the output voltage from the dc power supply is compared with a reference voltage, and the difference signal is amplified and fed back to the base of a pass transistor. In response to the feedback signal, the conduction of the pass transistor is varied, either linearly or as a switch, to regulate the output voltage. When the pass transistor can be operated at any point between cutoff and saturation, the regulator circuit is referred to as a **linear voltage regulator.** When the pass transistor operates only

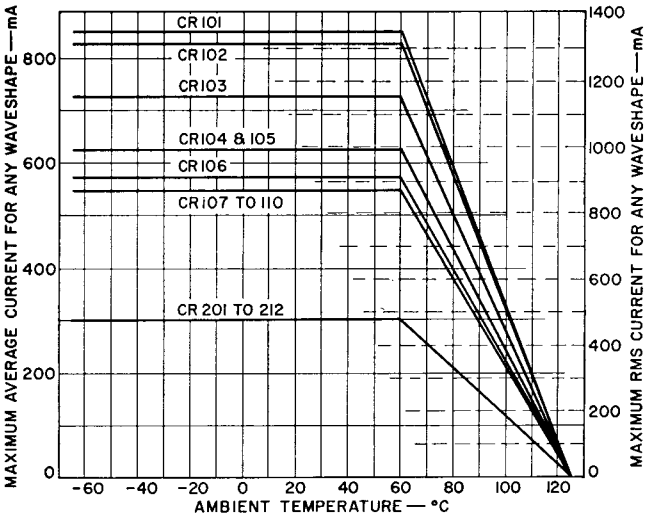



Figure 304. Current-temperature ratings for silicon stack rectifiers.

at cutoff or at saturation, the circuit is referred to as a **switching regulator.** All SCR regulators are by nature of SCR operation switching regulators.

All linear voltage regulators can be classified as either **series** or **shunt** types, as determined by the arrangement of the pass element with respect to the load. In a series regulator, as the name implies, the pass transistor is connected in series with the load. Regulation is accomplished by variation of the current through the series pass transistor in response to a change in the line voltage or circuit loading. In this way, the voltage drop across the pass transistor is varied and that delivered to the load circuit is maintained essentially constant. In the shunt regulator, the pass transistor is connected in parallel with the load circuit, and a voltage-dropping resistor is connected in series with this parallel network. If the load current tends to fluctuate, the current through the pass transistor is increased or decreased as required to maintain an essentially constant current through the dropping resistor.

## Series Regulators

Series-regulated power supplies may be either voltage-regulating types, voltage-regulating current-limiting types, current-regulating types, or voltage-regulating current-regulating types. Fig. 305 shows the response characteristics for each type of series-regulated power supply.

Linear series regulators provide an excellent means for prevention of large variations in power-supply load current or output voltage. Fast response time provided by the linear control circuit makes possible close control of the output voltage. However, because the series pass transistor is equivalent to a variable resistance in series with the load, the transistor must dissipate a large amount of power at low output voltages. Another disadvantage of the series regulator is that the total fault current passes through the regulating transistor if the load becomes short-circuited. As a result, overload and short-circuit protection in the form of current-limiting or drive-reduction networks that operate rapidly must be used to protect the transistor.

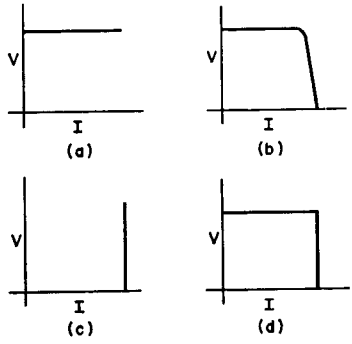



Figure 305. Typical response characteristics for series-regulated power supplies: (a) voltage-regulating types; (b) voltage-regulating current-limiting types; (c) current-regulating types: (d) voltage-regulating current-regulating types.

**Basic Circuit Configurations**—Fig. 306 shows a basic configuration for a linear series regulator which is representative of the type used in **voltage-regulating power supplies.** In this type of regulator, the series pass transistor is usually operated as an emitter-follower, and the control (error) signal used to initiate the regulating action is applied to the base. The base control is developed by a dc amplifier. This amplifier,

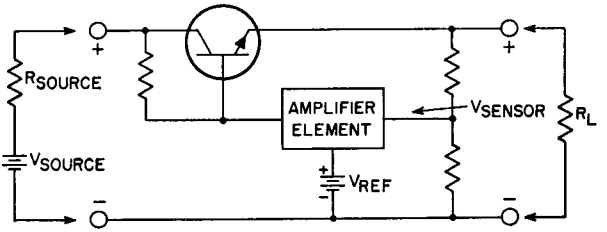



Figure 306. Basic series voltage regulator.

which is included in the feedback loop from the load circuit to the pass transistor, senses any change in the output voltage by comparison of this voltage with a known reference voltage. If an error exists, the error voltage is amplified and applied to the base of the pass transistor. The conduction of the pass transistor is then increased or decreased in response to the error signal input as required to maintain the output voltage at the desired value.

Voltage-regulating power supplies are required to maintain a constant output voltage, independent of the load current, as shown in Fig. 305(a). The supply, therefore, usually has a very low output impedance. For this reason, voltage-regulating supplies must often be made current-limiting to protect the regulator from very high current drawn at the output terminal, such as may be caused by a short circuit. In **voltage-regulating current-limiting power supplies,** the load current is prevented from rising above some predetermined design value by reduction of the power-supply output voltage when this current limit is reached, as shown in Fig. 305(b).

Fig. 307 shows the basic configuration for a linear regulator circuit used in **current-regulating power supplies.** This regulator senses the voltage across a resistor in series with the load, rather than the voltage across the load circuit as in the linear voltage regulator. Because the voltage across the series resistor is directly proportional to the load current, a detected error signal can be used to cancel any tendency for a change in load current from the desired value. Ideally, the linear current regulator has an infinite output impedance and out-

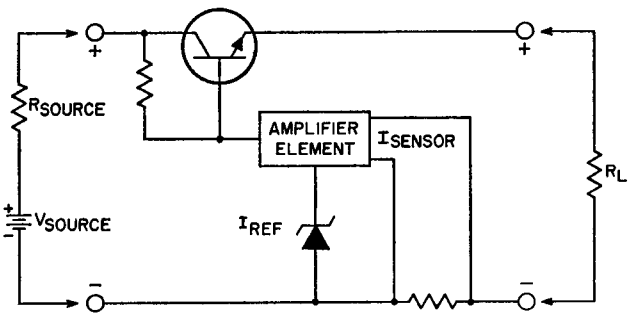



Figure 307. Basic series regulator modified for current sensing.

put characteristics as shown in Fig. 305(c).

The regulator circuit used with **voltage-regulating current-regulating power supplies** is essentially a combination of the other types of linear regulators. As shown in Fig. 305(d), the output response characteristics of this type of regulated supply exhibit a crossover point at which the supply switches from voltage regulation to current regulation.

Fig. 308 shows a block diagram of a voltage-regulating current-regulating power supply. The input ac power is rectified and filtered and is then applied to the regulating circuit. When preregulators are used, as is normally the case, switching types are preferred. The efficiency of the switching regulator is extremely high, and a fast response time to load or line variations is not required at this point in the circuit. (The operation and characteristics of switching regulators are discussed later in the section on **Switching Regulators.**)

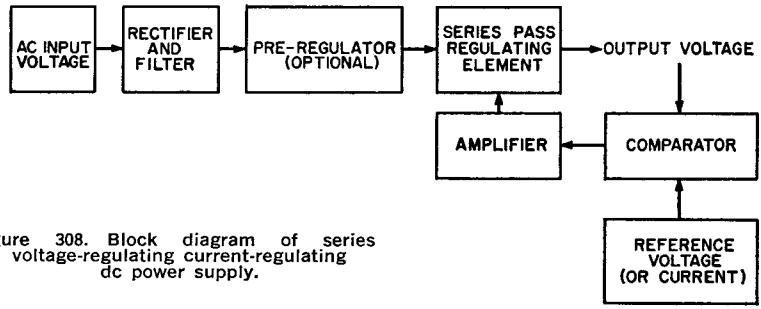



Figure 308. Block diagram of series voltage-regulating current-regulating dc power supply.

The output from the preregulator is transferred to the series pass element which provides the fast response time for the entire regulating circuit. At this point in the circuit, a sample of the output voltage is compared with a reference voltage and the resulting error signal, which is proportional to the difference between these voltages, is amplified and delivered to the base of the pass transistor to correct the output voltage.

In this type of system, the resulting output voltage is highly dependent upon the accuracy of the reference supply. Such a voltage source may be a temperature-compensated zener diode in series with a very constant source of current so that the diode incremental resistance has no effect on the output voltage. The sensitivity of the regulator is an inverse function of the gain of the drive amplifier. The smaller the variation to be sensed, the higher the required gain of the amplifier. A higher gain, however, results in less stability.

**Performance Parameters**—Most voltage-regulated power supplies are required to provide voltage regulation for wide variation in load current. It is important, therefore, to specify the output impedance of the supply, $\Delta V_{out}/\Delta I_{out}$, over a large band of frequencies. This parameter indicates the ability of the power supply to maintain a constant output voltage during rapid changes in load. The output impedance of a typical voltage-regulated supply is normally less than 0.1 ohm at all

frequencies below 2 kHz. Above this frequency, the impedance increases and may be as much as several ohms.

A power supply must continue to supply a constant voltage (or current) regardless of variations in line voltage. An index of its ability to maintain a constant output voltage or current during input variation is called the **line regulation** of the supply, which is defined as 100 $(V_o'/V_o)$, or as the change in output voltage $\Delta V_o$, for a specified change in input voltage, expressed in per cent. Typical values of line regulation are less than 0.01 per cent.

Another important power-supply parameter is **load regulation,** which specifies the amount that the regulated output quantity (voltage or current) changes for a given change in the unregulated quantity. Load regulation is mainly a function of the stability of the reference source and the gain of the feedback network.

A power-supply parameter referred to as **recovery time** denotes the time required for the regulated quantity (voltage or current) to return to the specified limits when a step change in load is applied, as shown in Fig. 309. Recovery time is a function of the frequency response of the feedback network of the power supply. For voltage-regulated supplies, the "roll-off" of the feedback network increases the output impedance at high frequencies, and the impedance becomes inductive. As a result, the high-frequency harmonics of the step change in the load current induce a spike of voltage at the output.

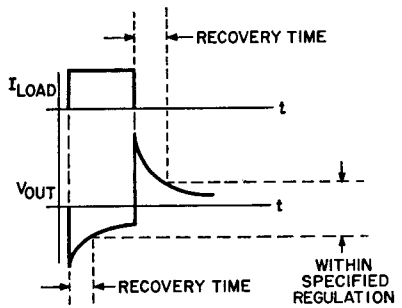



Figure 309. Typical recovery-time characteristics for regulated dc power supplies.

The amount of change in the output voltage of the regulated power supply from an initial value over a specified period of time is referred to as **drift.** This parameter is measured after an initial warm-up period with a constant input voltage and load applied and the ambient temperature held constant.

**Transistor Requirements**—In linear series regulators, the transistor parameters that affect circuit design and performance are collector dissipation, maximum collector current $I_C(max)$, leakage current ($I_{CER}$ in most cases), current gains $h_{FE}$ and $h_{fe}$, collector-to-emitter saturation voltage $V_{CE}(sat)$, collector-to-emitter breakdown voltage $V_{CEO}(sus)$, and second breakdown.

The collector-dissipation rating limits the amount of power which the series transistor can safely dissipate when the power supply is short-circuited. The maximum collector current $I_C(max)$ limits the total current which the regulator can handle. A low value of leakage current is required to maintain the stability of the circuit and, possibly, to prevent thermal runaway. This requirement makes silicon transistors especially suitable for use as the regulator pass element because leakage current is generally much lower in silicon transistors than

in germanium types. The current-gain parameters $h_{FE}$ and $h_{fe}$ determine the amount of drive current needed at various collector current levels. The ac forward-current transfer ratio $h_{fe}$ also determines the output impedance of the supply. A high $h_{fe}$ results in a low output impedance. The saturation voltage $V_{CE}(sat)$ is one factor that determines the required input voltage to the regulator for a specified output voltage and current. The collector-to-emitter breakdown voltage $V_{CEO}(sus)$ limits the maximum output voltage of the power supply. Second-breakdown considerations in circuit applications of transistors were discussed previously in the section on **Second Breakdown.**

**Current-Limiting Techniques**—One of the problems encountered in the design of series transistor voltage regulators is protection of the series control element from excess dissipation because of current overloads and short circuits.

In some series voltage-regulator circuits, overloading results in permanent damage to the series control transistor. For example, when the output terminals of the regulator circuit shown in Fig. 310 are shorted, the full input voltage and available current are applied to the series control transistor. This power usually is many times greater than the dissipation ratings of the series transistor.

A series fuse is sometimes used in an attempt to protect the series transistor from this excessive dissipation. A series fuse cannot usually provide the necessary protection under all overload conditions, however, because the thermal time constant of the fuse

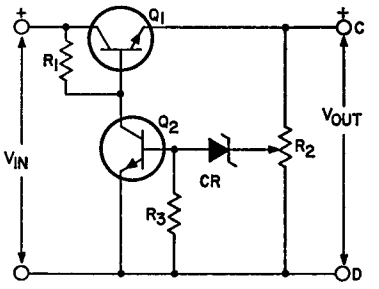



Figure 310. Series voltage regulator without current limiting.

is normally much greater than that of the transistor.

Protection for all overload conditions may be accomplished by use of a circuit which limits the current to a safe value, as determined from the dissipation rating of the series regulator transistor. An effective current-limiting circuit must respond fast enough to protect the series transistor and yet permit the circuit to return to normal regulator operation as soon as the overload condition is removed. It is desirable to achieve current-overload protection with minimum degradation of regulator performance.

One method of achieving limiting is to use a resistor in series with the regulator transistor. The large resistance normally required, however, dissipates a large amount of power and degrades the regulator performance.

The current-limiting section (dashed line) of the regulator circuit shown in Fig. 311(a) is designed to appear as a large series resistance during current overload and as a negligible resistance during normal operating conditions. The value of resistance $R_5$ is designed so that, during normal regulator operation, transistor $Q_4$ operates in the saturated condition. For the overload condi-

tion, $R_4$ is adjusted so that the maximum allowable value of overload current through this resistor produces a voltage drop large enough to cause silicon rectifier $CR_1$ to conduct. Conduction of $CR_1$ reduces the bias to $Q_4$, so that the transistor appears as an increasing series resistance in the regulator circuit.

Under short-circuit conditions, the entire value of input voltage $V_{in}$ appears across $Q_4$ simultaneously with the limiting value of current. Transistor $Q_4$ must be capable of withstanding the resulting dissipation. When the current limit is reached, the junction temperature of $Q_4$ rises to a value considerably above the ambient temperature. This increase in junction temperature causes the value of short-circuit current to rise slightly because of the inherent variation of the base-to-emitter voltage $V_{BE}$ with temperature in transistors. This effect is minimized by mounting silicon rectifier $CR_1$ and transistor $Q_4$ on a common heat sink so that their respective junction temperatures may reach the same value (the values of their respective $V_{BE}$ and forward-voltage-drop temperature coefficients are comparable).

Performance characteristics for the transistor series voltage regulator of Fig. 311(a) are shown in Fig. 311(b).

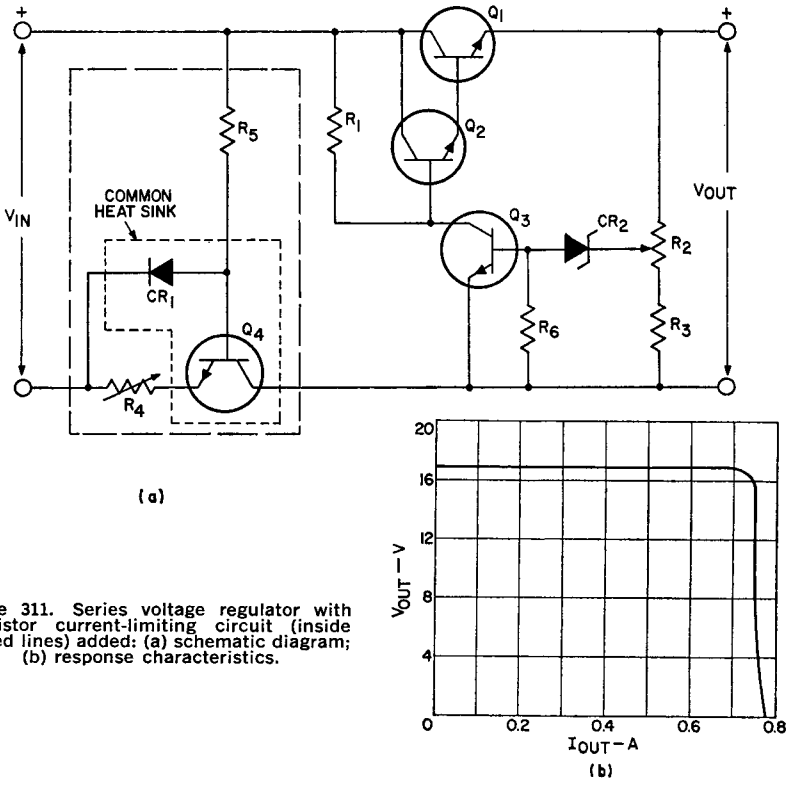
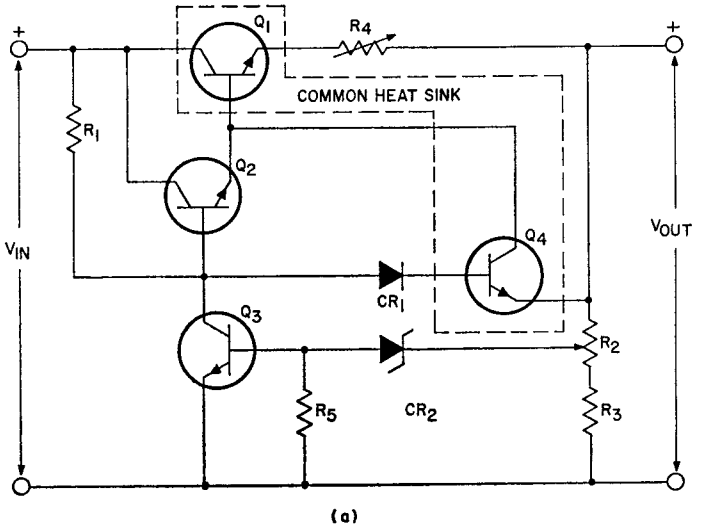


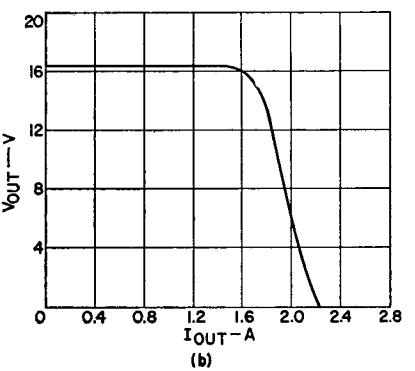

Figure 311. Series voltage regulator with transistor current-limiting circuit (inside dashed lines) added: (a) schematic diagram; (b) response characteristics.

Although the series-regulator circuit shown in Fig. 311(a) provides adjustable current limiting with simple circuitry and minimum power loss during normal operation, it has the disadvantage of requiring a second series transistor capable of withstanding short-circuit output current and total input voltage simultaneously.

In many high-current high-voltage regulator circuits, it is necessary to use parallel or series connections of pass transistors so that the voltage, current, and power ratings of the series control element are not exceeded. The method shown in Fig. 311(a) may not be practical in this application because of the additional series transistor required. The circuit shown in Fig. 312(a) eliminates the need for an additional series transistor by use of the series regulator transistor as the current-limiting element. This method is very effective when a Darlington connection is used for the series control transistor. A desirable feature of this circuit in high-current regulators is that it functions well even when the value of resistor $R_4$ is reduced to zero.




(a)



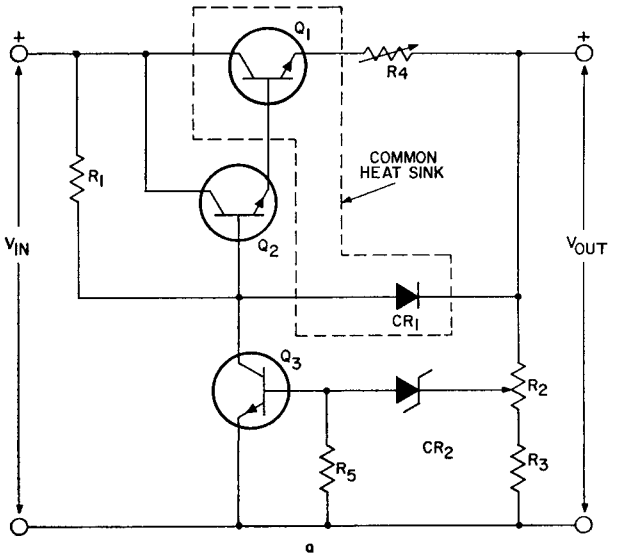

(b)

Figure 312. Series voltage regulator using pass transistor as part of current-limiting circuit: (a) schematic diagram; (b) response characteristics (for $R_4 = 0$).

In the circuit shown in Fig. 312(a), current limiting is achieved by the combined action of the components shown inside the dashed lines. The voltage developed across $R_4$ and the base-to-emitter voltages of $Q_1$ and $Q_2$ are proportional to the circuit output current. During current overload, these voltages add up to a value great enough to cause $CR_1$ and $Q_4$ to conduct. As $CR_1$ and $Q_4$ begin to conduct, $Q_4$ shunts a portion of the bias available to the series regulator transistor. This action, in turn, increases the series resistance of $Q_1$. The value of current in the circuit, under current-limiting conditions, is adjusted by varying the value of resistance $R_4$.

Higher current ranges may be obtained by increasing the number of rectifiers represented by $CR_1$. Temperature drift is minimized by mounting transistors $Q_1$ and $Q_4$ on a common heat sink. Performance characteristics for this circuit (for $R_4 = 0$) are shown in Fig. 312(b).

The circuit shown in Fig. 313(a) is a variation of that



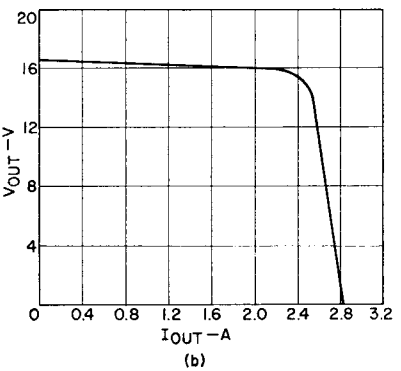




Figure 313. Series voltage regulator which uses additional transistor-diode network and series pass transistor to accomplish current-limiting function: (a) schematic diagram; (b) response characteristics.

shown in Fig. 312(a). Current limiting is adjusted by varying $R_4$ and by changing the number of silicon rectifiers represented by $CR_1$. Temperature drift is minimized by mounting the series control transistor $Q_1$ and silicon rectifier $CR_1$ on a common heat sink. Performance characteristics for this circuit are shown in Fig. 313(b). The circuits shown in Figs. 312 and 313 are both applicable to high-current high-voltage regulators because additional series power transistors are not required.

Fig. 314(a) shows another current-limiting circuit in which the regulator series control transistor is used as the current-limiting element. The series element must be capable of withstanding input voltage and short-circuit current simultaneously. The value of short-circuit current is selected by adjusting the value of resistor $R_4$. Performance characteristics of this circuit are shown in Fig. 314(b). The circuit functions equally well with resistor $R_4$ located in the positive output lead.

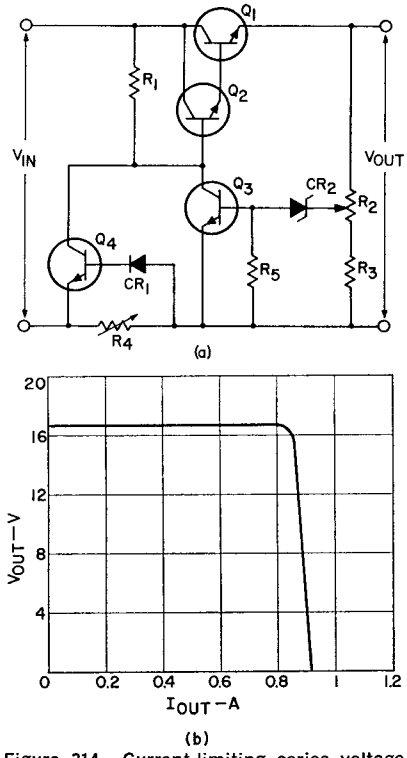



Figure 314. Current-limiting series voltage regulator in which series pass transistor must be capable of withstanding input voltage and short-circuit current simultaneously: (a) schematic diagram; (b) response characteristics.

**Design of Basic Current-Limited Regulated Supply**—Fig. 315 shows the circuit diagram of a voltage-regulated current-limited power supply. The function of the differential amplifier is to maintain the output voltage equal to the voltage at the top of $R_{ADJ}$, which is at point $V_f$. Because the input impedance of the differential amplifier is high, a negligible amount of current flows into its terminals. Therefore, essentially all of the current supplied by the reference voltage supply $V_{REF}$ flows through the resistive voltage divider consisting of $R_{CAL}$ and $R_{ADJ}$. Under quiescent conditions, the ratio of the reference voltage to the output voltage $V_o$ is given by

$$V_{REF}/V_o = R_{CAL}/R_{ADJ} \qquad (249)$$

As long as the reference voltage remains constant, a constant current $I_{REF}$ flows through $R_{CAL}$. Essentially, this same current (minus a negligible amount that flows into the differential amplifier) flows through $R_{ADJ}$. The output voltage can be expressed by the following product:

$$V_o = I_o R_{ADJ} \qquad (250)$$

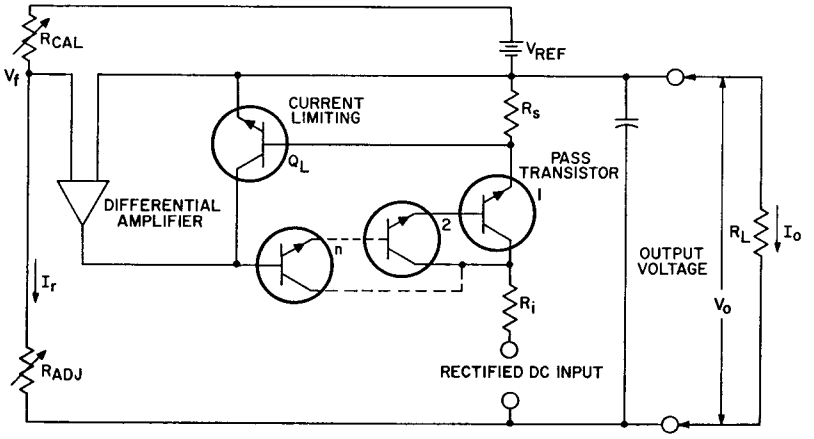



Figure 315. Voltage-regulated current-limited power supply.

If a current of 0.01 ampere flows through $R_{ADJ}$, the output voltage is then adjusted at the rate of 1/0.01 or 100 ohms per volt. The resistance $R_{ADJ}$ may be located at a remote point from the supply to make the supply remotely programmable.

Eq. (249) implies that the stability of the dc output voltage of the power supply is a direct function of the stability of the reference voltage. Stability of the ac feedback system of the power supply is maintained by the addition of a large value of capacitance in parallel with the output voltage of the power supply, as shown in Fig. 316, which illustrates the feedback mechanism of a voltage-regulated power supply. There are three points in the circuit at which phase shifts may occur: the pass transistor, the driver amplifier, and the differential amplifier. If the sum of these phase shifts is approximately 180 degrees, and if there is a point in the circuit at which the gain is greater than unity, the entire system will become unstable. The addition of a capacitor at the output decreases the gain at higher frequencies

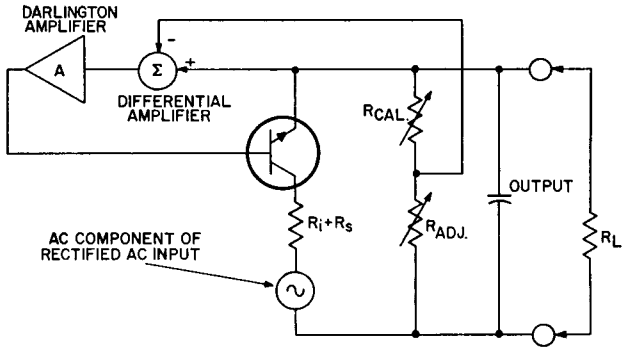



Figure 316. AC feedback circuit for a voltage-regulated power supply.

when the total phase shift (including that created by the capacitor itself) is 180 degrees or more. In addition, because of the inverse relationship of capacitance reactance to frequency, this capacitor decreases the total output impedance at higher frequencies.

Transistor $Q_L$ and resistor $R_S$ form the current-limiting configuration for this power supply. When the output current $I_o$ exceeds the current-limiting value of load current $I_o(max)$, the corresponding voltage drop across $R_S$ becomes large enough to forward-bias transistor $Q_L$. Transistor $Q_L$ then diverts all drive current greater than that needed to supply $I_o(max)$ from the differential amplifier to reference ground. At the same time, the impedance of the pass transistor increases to maintain the output current at a value essentially equal to $I_C(max)$, while the differential amplifier remains temporarily in saturation.

**Design equations and procedure:** The following is a step-by-step procedure for the design of a practical voltage-regulated current-limited power supply such as that shown in Fig. 315:

1. The desired input and output conditions for the power supply and the expected deviations from these values are determined. In this determination, the following parameters must be considered:

**Input Conditions**
Input voltage $V_i$
Possible positive change in input voltage caused by line variation, $\Delta V_i$
Possible negative change in input voltage caused by line variation and ripple, $\Delta V_n$
Input voltage source impedance $R_i$
Maximum case temperature $T_C(max)$

**Output Conditions**
Maximum output voltage $V_o(max)$
Maximum output current $I_o(max)$

2. A value is selected for the output voltage of the reference supply that satisfies the following condition:

$$V_o/20 < V_{REF} < V_o/10 \quad (251)$$

3. One of the following equations for a programmable power supply is used to calculate the value for $R_{ADJ}$. (Typically, a programmable power supply will have an $R_{ADJ}$ equal to 100 ohms per volt or 1000 ohms per volt.)

$$R_{ADJ} = 100\ V_o, \text{ or } R_{ADJ} = 1000\ V_o \quad (252)$$

4. The resistance value for $R_{CAL}$ is calculated by use of the following equation, which is obtained by rearrangement of terms in Eq. (249):

$$R_{CAL} = (V_{REF} \times R_{ADJ})/V_o \quad (253)$$

5(a). Initially, the transistor should satisfy the following requirements:

$$V_{CEO}(sus) \geq V_i + \Delta V_i \quad (254)$$
$$I_C(max) \geq I_o(max)$$

where $V_i$ is the input supply voltage, $\Delta V_i$ is the maximum variation in line voltage, $I_C$ is the

maximum rating for collector current, and $I_o(max)$ is the maximum output current.

**(b)** The maximum available output current from the differential amplifier $I_d$ is then determined.

**(c)** The total current gain $A_I$ of the cascaded driver stages and the pass unit is given by

$$A_I = h_{FE}(min)_1 \times h_{FE}(min)_2 \times \ldots \times h_{FE}(min)_n \quad (255)$$

Therefore, the appropriate number of driver stages is selected so that the total current gain will satisfy the following requirement:

$$A_I \geq I_o(max)/I_d \quad (256)$$

In addition, each driver transistor in the cascaded configuration should have a collector-to-emitter sustaining voltage $V_{CEO}(sus)$ equal to or greater than that of the pass transistor $Q_1$.

**(d)** The value of resistor $R_S$ is determined from the following relationship:

$$R_S = V_{BE}/I_o(max) \quad (257)$$

where $V_{BE}$ is the base-to-emitter voltage of transistor $Q_L$.

**(e)** The current-limiting transistor $Q_L$ is selected. The maximum collector-to-emitter voltage rating of this transistor $V_{CEO}$ should be greater than that of the differential amplifier.

**(f)** To assure that the maximum output voltage $V_o$ is obtained under maximum output-current conditions, the following quantities must be defined:

$$V_{BE1} + V_{BE2} + \ldots + V_{BE_{n-1}} + V_{CE}(sat)_n = V_x \quad (258)$$

$$V_{CE}(sat)_1 = V_y \quad (259)$$

Under the prescribed conditions, one of the two quantities $V_x$ or $V_y$, whichever is larger, appears across the pass transistor $Q_1$. (Whether $V_x$ or $V_y$ is larger is determined by the saturation characteristics of the transistor types used for the driver and pass transistors.) This larger value (the quantity $V_x$ or $V_y$) must be less than the effective input voltage minus the output voltage, i.e.,

$$V_x \text{ or } V_y < [V_i - \Delta V_n - I_{max}(R_i + R_s)] - V_o \quad (260)$$

where the bracketed term in the inequality defines the effective input voltage.

**(g)** The selection of the proper transistor for use as the pass element is based partly on the maximum power that can be dissipated by the device. This maximum power value is calculated according to the following procedure:

First, the output voltage from the regulator is determined from the following equation:

$$V_o = V_i - I_o(R_s + R_i) - V_{pass} + \Delta V_i \quad (261)$$

This equation is rewritten to obtain an expression for the voltage dropped across the pass transistor, as follows:

$$V_{pass} = V_i + \Delta V_i - V_o - I_o(R_i + R_S) \quad (262)$$

The power dissipated by the pass transistor is given by

$$P_{pass} = I_o V_{pass} \quad (263)$$

Substitution of Eq. (262) into Eq. (263) results in the following expression for the power dissipation in the pass transistor:

$$P_{pass} = I_o (V_i+\Delta V_i) - I_o V_o - I_o^2 (R_i+R_S) \quad (264)$$

To determine at what value of output current $I_o$ the power dissipation is maximum, the partial derivative of power with respect to current is set equal to zero, as follows:

$$dP_{pass}/dI = (V_i+\Delta V_i) - V_o - 2I_o (R_i+R_S) = 0 \quad (265)$$

The maximum power in the circuit occurs, therefore, when

$$I_o = \frac{V_o + \Delta V_i - V_o}{2(R_i+R_S)} \quad (266)$$

Eq. (266) shows that the absolute maximum power is dissipated by the pass transistor when the output voltage $V_o$ is zero. For this condition, the equation for the output current $I_o$ becomes

$$I_o = \frac{V_i + \Delta V_i}{2(R_i+R_S)} \quad (267)$$

Therefore, the maximum power dissipated by the pass transistor occurs when the output voltage is zero (i.e., when the output terminals are short-circuited). If the power supply is current-limited to a value of current less than $(V_i + \Delta V_i)/2(R_i + R_S)$, then the maximum power dissipated by the pass transistor can be expressed as follows:

$$P_{pass(max)} = I_o (V_i+\Delta V_i) - I_o^2 (R_i+R_S) \quad (268)$$

where $I_o$ is the maximum value of current allowed by the current-limiting circuit.

Consequently, if a transistor is to be specified for safe operation as a pass element, this maximum power dissipation $P_{pass}$ must be less than or equal to the maximum power rating of the transistor at the maximum case temperature.

**(h)** In addition, the maximum safe operating region for the transistor in this supply (a rectangle in which the upper right-hand corner is $V_i$, $I_o$) should be within the maximum safe operating region of this device.

**Sample design:** The following example illustrates the use of the basic design procedure and equations in the design of a practical voltage-regulating current-limiting power supply. It is assumed that the rectified ac power source is the output from a bridge rectifier that has a 200-microfarad filter capacitance, isolated from the line by a 1:1 transformer.

1. The desired operating conditions are selected as follows:

**Input Conditions**

Input voltage $V_i = 150$ V
Possible positive change in input voltage $\Delta V_i = 20$ V
Possible negative change in input voltage $\Delta V_n = 40$ V
Input source impedance $R_i = 10$ ohms
Maximum case temperature $T_C(max) = 75°C$

**Output Conditions**

Output voltage $V_o = 100$ V
Maximum output current $I_o(max) = 0.3$ A

2. A value of 8.2 volts is selected for the reference voltage on the basis of the following requirement:

$$V_o/20 < V_{REF} < V_o/10$$

3. From Eq. (252), the value of $R_{ADJ}$ is calculated as follows:

$$R_{ADJ} = 1000\ V_o = 1000\ (120)$$
$$= 120{,}000 \text{ ohms}$$

4. From Eq. (253), the value of $R_{CAL}$ is calculated to be

$$R_{CAL} = (V_{REF}/V_o)\ R_{ADJ}$$
$$= (8.2 \times 100 \times 10^3)/150$$
$$= 5500 \text{ ohms}$$

**5(a).** The RCA-2N5240 transistor is selected for the pass transistor because it satisfies the following conditions:

$$V_{CEO}(sus) \geq V_i \times \Delta V_i$$
$$225 \geq 150 + 20 = 170$$
$$I_C(max) \geq I_o(max)$$
$$5A \geq 0.3\ A$$

**(b)** A differential amplifier that has an available drive $I_d$ equal to 0.4 milliampere is used.

**(c)** From Eq. (255), the total current gain of the driver stage is determined as follows:

$$A_I \geq I_o(max)/I_d = 300\ mA/0.4\ mA$$
$$= 750$$

**(d)** The 2N3440 is a suitable drive transistor because it fulfills the following requirement:

$$h_{FE}(min)_1 \times h_{FE}(min)_2 > A_I$$
$$20 \times 40 = 800 > 750$$

**(e)** The 2N2102 transistor is suitable as the current-limiting device because it meets the following requirement:

$$V_{CEO} \text{ (of } Q_L) > V_o(max)$$
$$\text{(of differential amplifier)}$$
$$65\ V > 8.2\ V$$

**(f)** The quantities $V_x$ and $V_y$ are determined as follows:

$$V_{BE1} = V_{BE} \text{ (of 2N5239}$$
$$\text{at } I_o = 0.3\ A) = 1\ V$$
$$V_{CE}(sat)_2 = V_{CE}(sat) \text{ (of 2N3440}$$
$$\text{at } I_C = 15\ mA) = 0.5\ V$$
$$V_{BE1} + V_{CE}(sat)_2 = 0.64 + 0.75$$
$$= 1.39\ V = V_x$$
$$V_{CE}(sat)_1 \text{ (for 2N5239)} = 2.5V$$
$$= V_y$$
$$V_y > V_x$$

Therefore, $V_z = 2.5$ V ($V_z$ is used to represent the larger of the two quantities $V_x$ and $V_y$). The quantity $V_z$ must meet the following requirement:

$$V_z < (V_i - \Delta V_n) - V_o$$
$$- I_o(max)\ (R_i + R_S)$$
$$2.5 < 170 - 40 - 120 - 0.3(10+2)$$
$$2.5 < 6.1$$

**(g)** The maximum power point is then determined as follows:

$$I_o(max) < (V_i + \Delta V_i)/2(R_i + R_S)$$
$$0.3 < (150+20)/2(10+2)$$
$$= 170/24 = 7.1$$

Therefore, $P_{pass}(max)$ is determined from Eq. (268) as follows:

$$P_{pass}(max) = I_o\ (V_i + \Delta V_i)$$
$$- I_o^2\ (R_i + R_S)$$
$$= (0.3)(170)$$
$$- (0.3)^2\ (10+2)$$
$$= 51 - 1.08 \simeq 50 \text{ watts}$$

The maximum power rating of the 2N5240 is 70 watts at 75°C; therefore, the power-handling capability of this transistor makes it suitable for use as the pass element.

**(h)** The maximum operating region for the 2N5240 in this supply is within the maximum safe operating region of this device, as shown in Fig. 317. Fig. 318 shows the complete circuit diagram for the voltage-regulated current-limited power supply; Fig. 319 shows the schematic for the reference-voltage power supply.

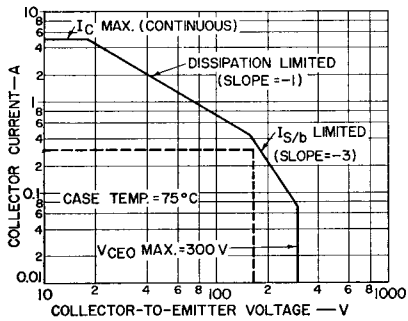



Figure 317. Safe-area curve for the RCA 2N5240 transistor. Dashed-line rectangle indicates that transistor operates within ratings in the 100-volt current-limiting (0-3 ampere) series-regulated power supply.

**Foldback Current Limiting**— Foldback current limiting is a form of protection against excessive current. If the load impedance is reduced to a value that would draw more than the predetermined maximum current, the foldback circuit reduces output voltage and thus reduces the current. Further reduction of load impedance causes further decrease of output voltage and current; therefore a regulated power supply that includes a foldback current-limiting circuit has the voltage-current characteristic shown in Fig. 320. The foldback process is reversible; if the load impedance is increased while the circuit is in the limiting mode, the output voltage and current increase. When the current reaches the threshold level,

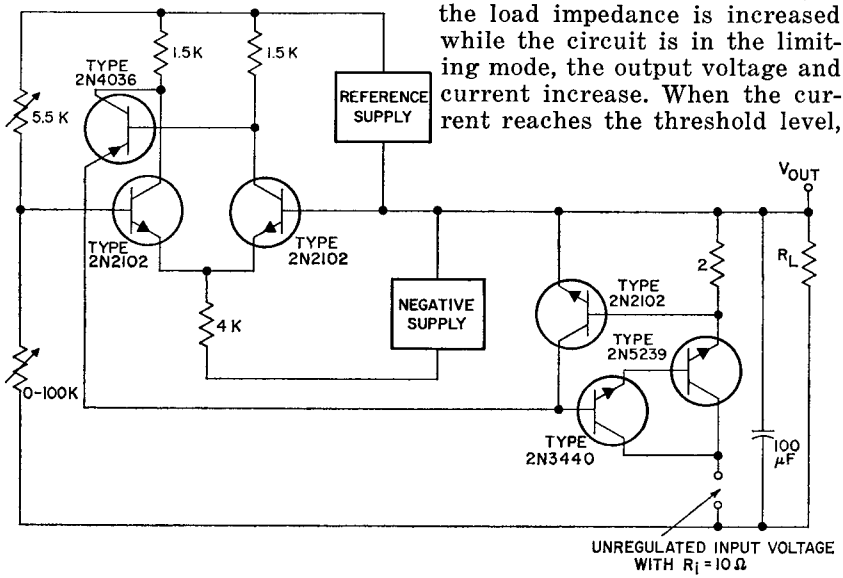



Figure 318. Schematic diagram of 100-volt series-regulated dc power supply in which output current is limited to a maximum value of 0.3 ampere.

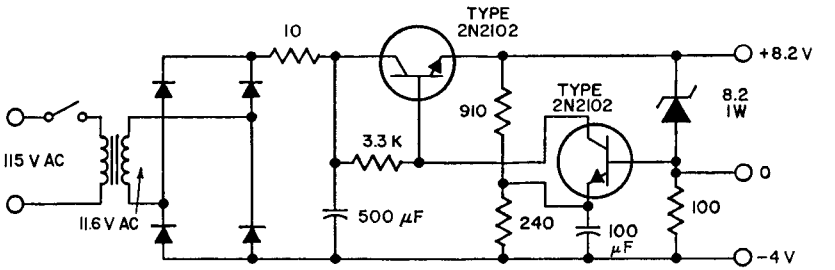



Figure 319. Schematic diagram of reference voltage supply for regulated power supply shown in Fig. 328.

the regulator is re-activated, and the power supply returns to normal operation.

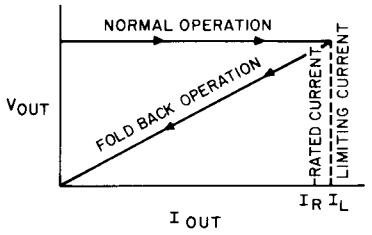



Figure 320. Output characteristic of a regulated power supply with foldback current-limiting protection for pass transistor.

A foldback current-limiting circuit is shown in Fig. 321. At low output current, transistor $Q_5$ is cut off; the value of resistor $R_5$ is selected so that $Q_5$ has zero bias when the output current reaches its rated value, $I_R$. When the load current $I_{OUT}$ reaches the limiting value, $I_X$, $Q_5$ begins to conduct; current flows through resistor $R_2$, transistor $Q_4$ turns on, and the base-to-emitter voltage of transistor $Q_3$ is reduced. Therefore, the base-to-emitter voltage of transistor $Q_2$ decreases, and the output voltage of the power supply decreases.

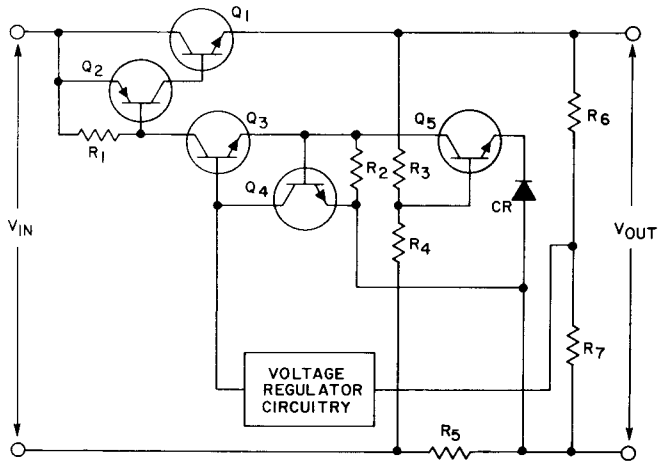



Figure 321. Foldback-current-limiting circuitry in a series voltage regulator.

This decrease in the output voltage $V_{OUT}$ reduces the output current, so that $Q_5$ continues to conduct at the same emitter current. If the load impedance is reduced further, $Q_5$ is driven even harder, and the output voltage and current decrease even further.

The performance of this foldback circuit can be analyzed with the aid of the loop equation for the base-emitter circuit of $Q_5$. The general loop equation may be expressed in the following form:

$$I_{OUT} R_5 = V_D + V_{BE} + V_{R4} \quad (268)$$

where $V_D$ is the voltage drop across the diode CR, and $V_{BE}$ is the base-emitter circuit of $Q_5$.

At the rated current $I_R$, it is desirable that $V_{BE} = 0$. For this condition, Eq. (268) can be rewritten as follows:

$$I_R R_5 = V_D + V_{R4} \quad (269)$$

At the limiting current $I_L$, just before foldback is initiated, the loop equation becomes

$$I_L R_5 = V_D + V_{BE} + V_{R4} \quad (270)$$

At the short-circuit current, $I_{SC}$, $V_{OUT} = 0$; therefore, $V_{R4} = 0$. The loop equation is then expressed by the following relationship.

$$I_{SC} R_5 = V_D + V_{BE} \quad (271)$$

By use of the appropriate loop-equation forms, the ratio of the limiting current to the rated current can be expressed as follows:

$$\frac{I_L}{I_R} = \frac{V_D + V_{BE} + V_{R4}}{V_D + V_{R4}} \quad (272)$$

and the ratio of the short-circuit current to the rated current is given by

$$\frac{I_{SC}}{I_R} = \frac{V_D + V_{BE}}{V_D + V_{R4}} \quad (273)$$

The following relationship defines the necessary conditions for the limiting current to be close to rated current:

$$V_D + V_{BE} + V_{R4} \approx V_D + V_{R4} \quad (274)$$

Therefore,

$$(V_D + V_{R4}) \gg V_{BE} \quad (275)$$

If $V_D$ is the drop across a single diode (0.7 volt for a silicon diode) and if $(V_D + V_{R4})$ is 3 volts as a compromise, then $V_{R4} = 2.3$ volts. For the 2N3055 transistor $Q_5$, $V_{BE}$ is approximately 0.7 volt. The ratio of the limiting current to the rated current is given by

$$\frac{I_L}{I_R} = \frac{0.7 + 2.3 + 0.7}{0.7 + 2.3} = 1.23$$

and the ratio of the short-circuit current to the rated current is given by

$$\frac{I_{SC}}{I_R} = \frac{0.7 + 0.7}{0.7 + 2.3} = 0.47$$

Improved foldback-circuit performance can be achieved by use of a differential amplifier instead

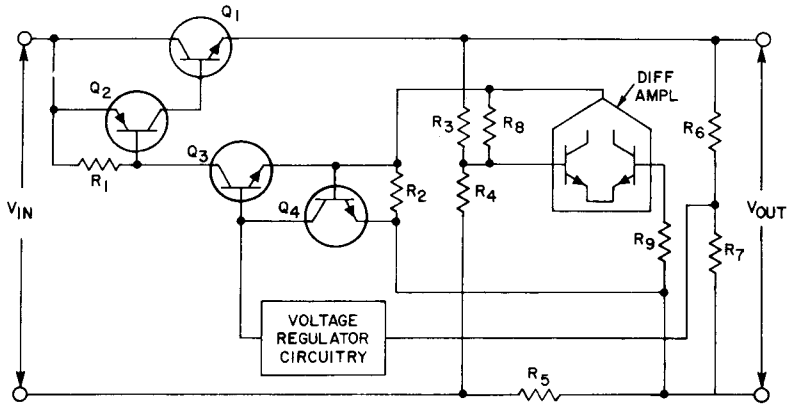



Figure 322. Foldback-current-limiting circuit with a differential amplifier for greater sensitivity.

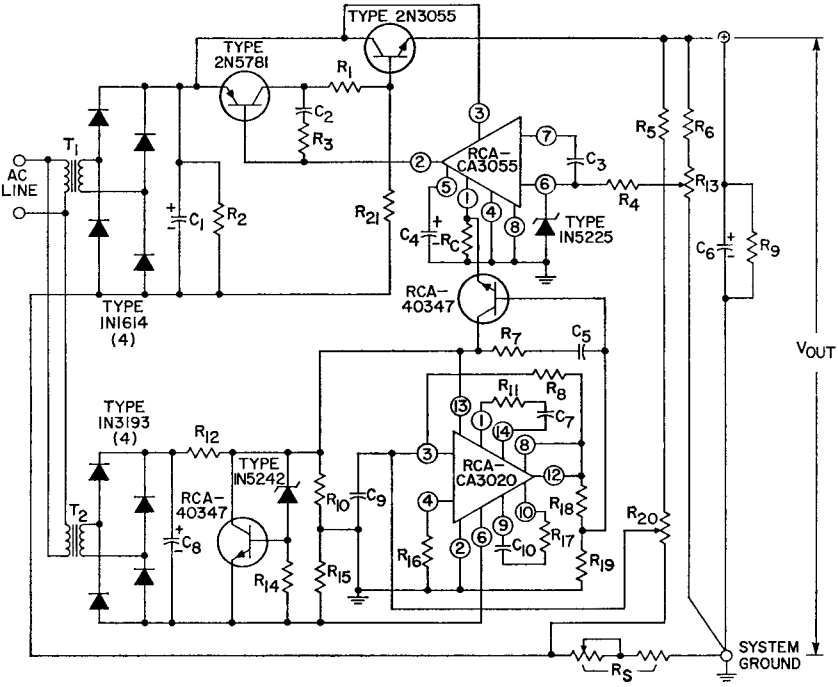
of single-ended amplifier $Q_5$. With the improved circuit, illustrated in Fig. 322, the following current-ratio values are obtained:

$$\frac{I_L}{I_R} = 1.05 \quad \text{and} \quad \frac{I_{SC}}{I_R} = 0.042$$

**Foldback-Limited Regulated Supply**—Fig. 323 shows a series regulated power supply with foldback current limiting. This supply can deliver currents of up to 3 amperes at 20 volts. The circuit uses integrated circuits for the regulation and protection functions; the voltage regulator is an RCA-CA3055 and the foldback limiter uses an RCA-CA3030 operational amplifier as a linear differential amplifier.

**Foldback-Limited Supply That Uses a Hybrid-Circuit Regulator**—The RCA line of power hybrid circuits includes a series voltage regulator, shown in Fig. 324, designed for use as the regulating element in foldback-current-limited regulated dc power supplies. The hybrid-circuit regulator includes an RCA-CA3085A monolithic-integrated-circuit voltage-regulator chip for voltage regulation, stability, and temperature compensation. This integrated circuit supplies a regulated signal to a two-stage high-current booster circuit that consists of a p-n-p driver chip $Q_2$ and an n-p-n hometaxial-base transistor chip $Q_4$ (RCA-2N3055 type) used as the series pass transistor. This two-stage output circuit makes possible a load-current capability of 4 amperes without the use of external booster devices. With the use of two external booster transistors, the hybrid circuit can provide regulation at load currents up to 12 amperes. For load currents greater than 12 amperes, increased current-handling capability can be achieved by use of the regulator circuit as a Darlington driver.

The internal circuitry of the hybrid regulator also includes a foldback-current-limiting circuit, a crowbar trigger circuit, and three ballast resistors. The ballast resistors are provided to assure current sharing between the internal pass transistor and two




$T_1$ = Signal Transformer Co., Part No. 24-4 or equivalent

$T_2$ = Signal Transformer Co., Part No. 12.8-0.25 or equivalent

$C_1$ = 5900 μF, 75 V. Sprague Type 36D592F075BC or equivalent

$C_2$ = 0.005 μF, ceramic disc, Sprague TGD50 or equivalent

$C_3$, $C_7$, $C_{10}$ = 50pF, ceramic disc, Sprague 30GA-Q50 or equivalent

$C_4$ = 2μF, 25 V, electrolytic, Sprague 500D G025BA7 or equivalent

$C_5$ = 0.01 μF, ceramic disc, Sprague TG510 or equivalent

$C_6$ = 500 μF, 50 V, Cornell-Dubilier No. BR500-50 or equivalent

$C_8$ = 250 μF, 25 V, Cornell-Dubilier BR 250-25 or equivalent

$C_9$ = 0.47 μF, film type, Sprague Type 220P or equivalent

$R_1$ = 5 ohms, 1 watt, IRC type BWH or equivalent

$R_2$ = 1000 ohms, 5 watts, Ohmite type 200-5 ¼ or equivalent

$R_3$ = 1200 ohms, ½ watt, carbon, IRC Type RC ½ or equivalent

$R_4$ = 100 ohms, ½ watt, carbon, IRC Type RC ½ or equivalent

$R_5$ = 430 ohms, 2 watts, wire wound, IRC Type BWH or equivalent

$R_6$ = 9100 ohms, 2 watts, wire wound, IRC Type BWH or equivalent

$R_7$ = 470 ohms, ½ watt, carbon, IRC type ½ or equivalent

$R_8$ = 5100 ohms, ½ watt, carbon, IRC type RC ½ or equivalent

$R_9$, $R_{14}$ = 1000 ohms, 2 watts, wire wound, IRC type BWH or equivalent

$R_{10, 15}$ = 250 ohms, 2 watts, 1% wire wound, IRC type AS-2 or equivalent

$R_{11}$, $R_{17}$ = 1000 ohms, ½ watt, carbon, IRC type RC ½ or equivalent

$R_{12}$ = 82 ohms, 2 watts, IRC type BWH or equivalent

$R_{13}$ = 1000 ohms, potentiometer, Clarostat Series U39 or equivalent

$R_{16}$ = 1200 ohms, 2 watts, wire wound, IRC type BWH or equivalent

$R_{38}$ = 510 ohms, ½ watt, carbon, IRC type RC ½ or equivalent

$R_{19}$ = 10,000 ohms, ½ watt, carbon, IRC type RC ½ or equivalent

$R_{20}$ = 300 ohms, potentiometer, Clarostat Series U39 or equivalent

$R_{21}$ = 510 ohms, 3 watts, wire wound, Ohmite type 200-3 or equivalent

$R_C$ = 240 ohms, 1%, wire wound, IRC type AS-2 or equivalent

$R_S$ = (See text for fixed portion); 1 ohm, 25 watts, Ohmite type H or equivalent

**Miscellaneous**

(1 Req'd) — Heat Sink, Delta Division Wakefield Engineering NC-423 or equivalent

(3 Req'd) — Heat Sink, Thermalloy #2207 PR-10 or equivalent

(1 Req'd) — 8-pin socket Cinch #8-1CS or equivalent

(1 Req'd) — 14-pin DIL socket, T.I., #ICO14ST-7528 or equivalent

(2 Req'd) — TO-5 socket ELCO #05-3304 or equivalent

Vector Board #838AWE-1 or equivalent

Vector Receptacle R644 or equivalent

Chassis—As required

Cabinet—As required

Dow Corning DC340 filled grease

Figure 323. Schematic diagram of a dc power supply that uses integrated circuits in the voltage regulator and foldback-current-limiting circuitry.

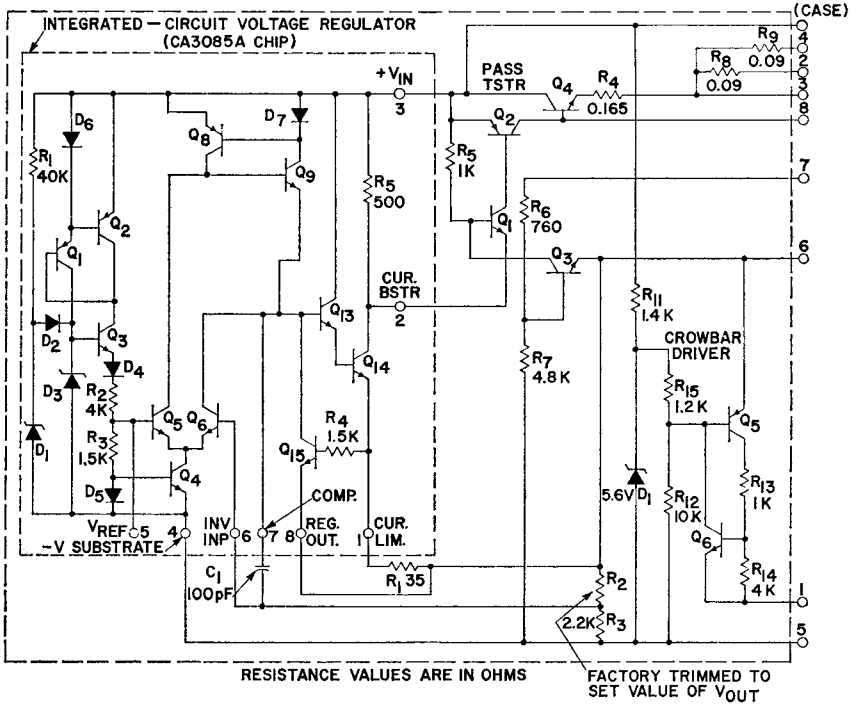



Figure 324. RCA high-current hybrid-circuit series voltage regulator.

external pass transistors when regulation is required at current levels between 4 and 12 amperes.

Because the CA3085A chip is rated for a maximum supply voltage of 40 volts and a feedback voltage to the inverting input (terminal 6) of 1.8 volts, the output-voltage capability of the hybrid-circuit regulator is limited to a range from 2 to 32 volts. Standard-design regulator circuits that provide a regulated output of 5, 8, or 12 volts are available. For each type, the output voltage is regulated to within ± 1 per cent for typical line-voltage, load-current, and temperature variations.

The values of the voltage-divider resistors $R_2$ and $R_3$ establish the level of the output voltage. The junction of these resistors is directly coupled to the inverting input of the CA3085A voltage-regulator chip. The values of the resistors are selected to divide the output voltage so that, for the rated output, the voltage applied to the CA3085A inverting input is approximately 1.6 volts. Any change in output voltage produces a corresponding change at the inverting input of the CA3085A voltage-regulator circuit, and this circuit then develops an output to cancel the change in output voltage. For example, an increase in load resistance causes the output voltage to rise. The resultant increase in the voltage at the in-

verting input of the CA3085A circuit is applied to the base of transistor $Q_6$ on the CA3085A chip, and the collector current of this transistor increases. The base current of transistor $Q_{13}$ then decreases because it is derived from a constant-current source. This action, in turn, causes the base and collector currents of transistor $Q_{14}$ to decrease so that the drive for the p-n-p driver transistor $Q_2$ and the n-p-n pass transistor $Q_4$ is reduced. The load current then decreases to return the output voltage to its original value.

Fig. 325 illustrates the use of the hybrid-circuit series voltage regulator in a foldback-current-limited regulated dc power supply. This supply provides an output of 5 volts regulated to within ± 1 percent. The two external 2N3055 transistors $Q_{B1}$ and $Q_{B2}$ are used as booster pass transistors to increase the load-current capability of the basic regulator circuit to 10 amperes.

The values of the three ballast resistors ($R_4$, $R_8$, and $R_9$) in the hybrid circuit are selected so that (1) the voltage drop across each resistor at the rated current is approximately 450 millivolts and (2) the current through each external pass transistor is 1.5 times that through the internal pass transistor. These resistors, therefore, force the required current sharing between the internal and external pass transistors.

The foldback-current-limiting circuit consists of transistors $Q_1$ and $Q_3$ and resistors $R_1$, $R_5$, $R_6$, and $R_7$ in the hybrid circuit (shown in Fig. 324). Transistor $Q_3$ senses the voltages across the branches of the bridge circuit formed by resistors $R_6$ and $R_7$ on

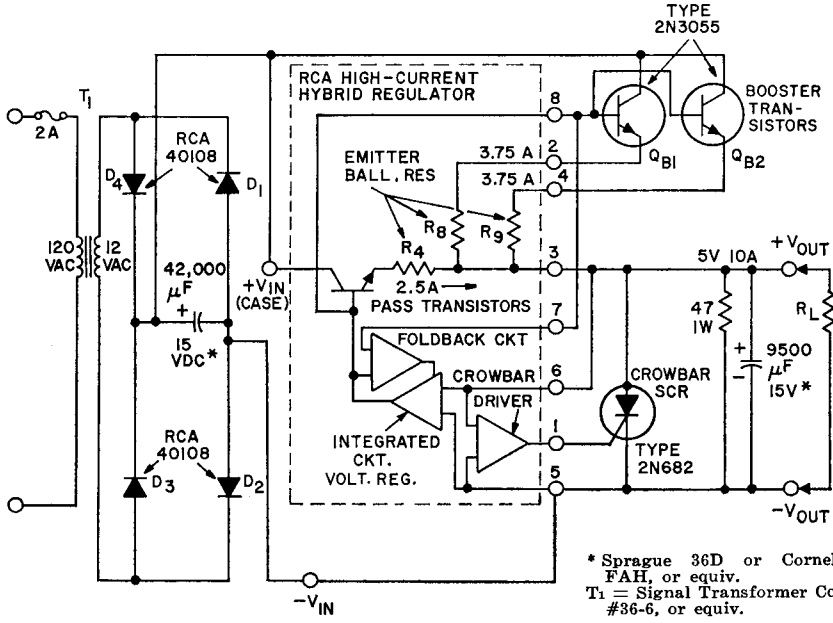



Figure 325. Foldback-current-limited regulated dc power supply that uses an RCA high-current power hybrid circuit as the series voltage regulator.

one side and the base-emitter junction of the series pass transistor $Q_4$, resistor $R_4$, and the load resistance on the other side. Because the base-to-emitter voltage of transistor $Q_4$ increases with the collector current, inclusion of the base-emitter junction of this transistor in the bridge increases circuit sensitivity.

During normal operation, transistor $Q_1$ is operated in the saturation region, and transistor $Q_3$ is cut off. When an overload occurs, transistor $Q_3$ is driven into conduction, and the collector current of this transistor flows through resistor $R_5$ to decrease the base-to-emitter voltage of transistor $Q_1$. This effect reduces the base drive to the p-n-p driver transistor $Q_2$ and subsequently the voltage drop across the load resistor. The voltage-feedback condition reaches a stable point on the load-resistance characteristic because the loop gain is less than unity. Transistor $Q_1$ effectively protects transistors $Q_2$ and $Q_4$ and the main pass transistor $Q_{14}$ on the CA3055 integrated-circuit chip (shown in Fig. 324), but it does not protect transistor $Q_{13}$. Protection of this latter transistor is provided by transistor $Q_{15}$, which turns on when the current through resistor $R_1$ reaches 20 milliamperes.

The crowbar trigger circuit in the hybrid circuit provides a trigger input to the external 2N682 crowbar SCR in response to an overvoltage that ranges from 105 to 125 per cent of the rated output value. This overvoltage may result from short-duration transient currents generated by either the load, the supply, or a pass transistor that becomes short-circuited. Resistor $R_{11}$ and zener diode $D_1$ provide a stable reference voltage that is reduced in value by the voltage divider formed by resistors $R_{12}$ and $R_{15}$. The voltage at the junction of these resistors is compared to the output voltage by transistor $Q_5$. When an overvoltage occurs, transistor $Q_5$ turns on and provides the base drive to turn on the transistor $Q_6$. This action is regenerative, and the collector currents of transistors $Q_5$ and $Q_6$ are limited only by resistor $R_{13}$, which limits the base current of transistor $Q_6$. Resistor $R_{14}$ provides a leakage path for the collector-base junction of transistor $Q_6$. The output of the trigger circuit is connected to the gate of the SCR. The SCR is triggered on by the gate current supplied by this circuit to provide a low-impedance path to shunt excessive currents generated by the overvoltage condition away from the load circuit.

## Shunt Regulators

Although shunt regulators are not as efficient as series regulators for most applications, they have the advantage of greater simplicity. The shunt regulator includes a shunt element and a reference-voltage element. The output voltage remains constant because the shunt-element current changes as the load current or input voltage changes. This current change is reflected in a change of voltage across the resistance $R_1$ in series with the load. A typical shunt regulator is shown in Fig. 326.

The shunt element contains one or more transistors connected in the common-emitter configuration in parallel with the load, as shown in Fig. 327.

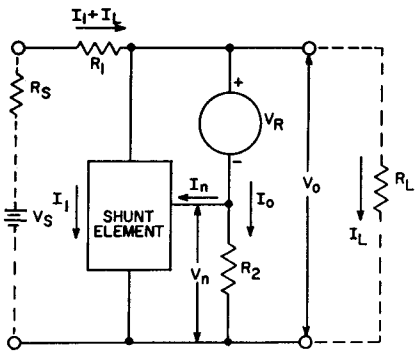



Figure 326. Basic configuration for a typical shunt regulator.

**Design Procedure and Equations**—The following step-by-step procedure is recommended for the design of transistor shunt-type voltage regulators:

1. The desired input requirements, load conditions, and output-voltage requirements are defined in terms of the following parameters:

Input voltage $V_S$

Input-voltage variation $\Delta V_S$

Source resistance $R_S$

Output load resistance $R_{Lo}$

Output voltage $V_o$

Output-voltage variation $\Delta V_o$

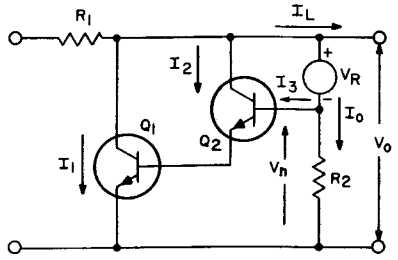



Figure 327. Shunt regulator circuit using two transistors as the shunt pass element.

The terms $V_S$ and $R_{Lo}$ are design-center values; $\Delta V_S$ and $\Delta R_L$ are maximum deviations from these values.

2. The transistor type selected must operate within ratings for the following values of $V_1(max)$, $I_1(max)$, and maximum dissipation $P_1(max)$ across the shunt element when both line and load regulation are required:

$$I_1(max) = I_L(max) = V_o(R_{L_o} - \Delta R_L) \quad (276)$$

$$V_1(max) = V_o \text{ (under forward-bias conditions)} \quad (277)$$

$$P_1(max) = I_1(max)\, V_1(max) \quad (278)$$

3. A value for resistance $R_2$ to provide a current $I_o$ greater than the minimum value required to supply the reference voltage (i.e., to break down a voltage-reference diode, for example) is determined. The following equation may be used as a guide:

$$R_2 = n/I_o \quad (279)$$

where n is the number of stages in the shunt element.

4. The output resistance $R_o$ of the regulator is given by

$$R_o = (2\Delta V_o/V_o)/R_{L_o} \quad (280)$$

If a value of series resistance is assumed for the reference $R_f$, this equation can be solved for $h_{fc}$. It can be shown that

$$R_o = \frac{R_f + \dfrac{h_{ie} R_2}{h_{ie} + R_2}}{1 + h_{fe} \dfrac{R_2}{R_2 + h_{ie}}} \quad (281)$$

where $h_{fe} = h_{fe1} h_{fe2} \ldots h_{fen}$, and $h_{fen}$ and $h_{ie}$ are the ac current transfer ratio and the input impedance, respectively, of the $Q_n$ stage.

5. The values of $I_n$ and $V_n$ for the shunt element are determined as follows:

$$I_n = I_1/(h_{FE1} h_{FE2} \ldots h_{FEn-1}) \quad (282)$$

where $h_{FEn-1}$ is the dc current gain of the $Q_{n-1}$ stage of the shunt element measured at a collector current of $I_{n-1}$.

$$V_n = V_1 + V_2 + \ldots + V_{n-1} \quad (283)$$

where $V_n$ is the base-to-emitter voltage of the $Q_{n-1}$ stage at a collector current of $I_{n-1}$.

6. A voltage reference source is selected which has a resistance less than the value that had been assumed for $R_f$ (or $h_{fe}$ is recomputed using a new value of $R_f$), a voltage $V_R = V_o - V_n$, a maximum current greater than $I_o + I_n$, and a maximum dissipation rating greater than $V_R (I_o + I_n)$.

7. The value of series resistance R, including both source resistance $R_S$ and external resistance $R_1$, is determined. The value of R depends on the value of the input voltage $V_S$ and its variation $\Delta V_S$; R may be expressed in terms of these quantities as follows:

$$V_S + \Delta V_S = V_o + R\,[I_L(\max) + I_1(\max)] \quad (284)$$

$$V_S - \Delta V_S = V_o + R\,I_1(\max) \quad (285)$$

For the usual case, $I_1(\max)$ is equal to $I_L(\max)$.

**Sample Design**—The following example illustrates the procedure used in the design of a practical shunt type of voltage regulator such as that shown in Fig. 328. When the step-by-step procedure refers to components by reference designations, Fig. 227 indicates the component being considered.

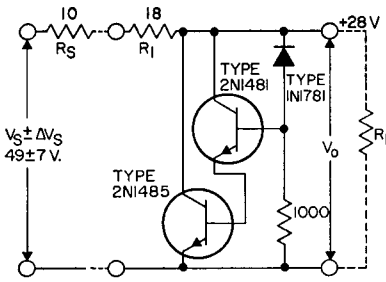



Figure 328. Schematic diagram of 28-volt shunt regulator circuit.

1. The first step in the design of a regulated power supply is to establish the circuit operating conditions and requirements. For the example chosen, the following parameter values are assumed:

Source resistance $R_S = 10$ ohms
Output load resistance $R_{Lo} = 110$ ohms
Possible variation in output load resistance $\Delta R_L = \pm 55$ ohms
Input voltage $V_S = 49$ volts
Possible change in input voltage $\Delta V_S = \pm 7$ volts
Output voltage $V_o = 28$ volts
Possible change in output voltage $\Delta V_o = \pm 0.0125$ volt
Maximum transistor case temperature $T_C(\max) = 55°C$

2. A transistor for use as the shunt pass element is then selected on the basis of the maximum current, voltage, and power dissipation that the pass element will be subjected to in the power-supply circuit. These maximum values are determined as follows:

$$I_1(max) = I_L(max) = V_o/(R_{L_o} - \Delta R_L) = 0.5 \text{ ampere}$$

$$V_1(max) = V_o = 28 \text{ volts}$$
under forward-bias conditions

$$P_1(max) = V_1(max)\, I_1(max) = 28 \times 0.5 = 14 \text{ watts}$$

The data on the RCA-2N1485 indicate that this transistor can operate within ratings for these circuit conditions and, therefore, is suitable for use as the pass element.

**3.** A voltage-reference diode to supply the voltage $V_R$ is selected. If an output current $I_o$ of 2 milliamperes is required to supply the reference voltage $V_R$ (i.e., to break down the voltage-reference diode) and if two stages are used for the shunt pass element, the value of the resistance $R_2$ in series with the voltage-reference diode is calculated as follows:

$$R_2 = 2/(2\times10^{-3}) = 1000 \text{ ohms}$$

**4.** The output resistance of the regulator is calculated as follows:

$$R_o = (2\Delta V_o/V_o)/R_{L_o} = (0.025/28)/110 = 0.10 \text{ ohm}$$

**5.** If the series resistance $R_f$ is assumed to be 5 ohms and the $Q_n$ stage is assumed to have a typical input impedance $h_{ie}$ of 50 ohms, the ac current transfer ratio $h_{fe}$ of the pass element is determined from the following calculation:

$$R_o = \frac{R_f + \dfrac{R_2\, h_{ie}}{h_{ie} + R_2}}{1 + h_{fe}\dfrac{R_2}{h_{ie} + R_2}}$$

$$0.10 = \frac{5 + \dfrac{1000 \times 50}{50 + 1000}}{1 + h_{fe}\dfrac{1000}{50 + 1000}}$$

$$h_{fe} = 52.9/0.095 = 560$$

Consequently, two stages are required for the shunt element, with a product $h_{fe1} \times h_{fe2} = 560$. The 2N1485 selected in step (2) for the first stage $Q_1$ has the following design-center values:

$I_C = I_1 = V_o/R_{L_o} = 250$ mA
$f_{fe1} = 56$
$h_{FE} = 50$
$V_{BE} = 0.8$ volt

For the second stage $Q_2$, therefore, the following values are required:

$h_{fe2} = 560/56 = 10$
$I_C = I_1/h_{FE1} = 250 \text{ mA}/50 = 5$ mA

An RCA-2N1481 transistor meets these requirements. The following design-center values can be obtained from published data for the 2N1481 for a collector current of 5 milliamperes:

$h_{fe2} = 20$ $\quad h_{ie} = 50$
$h_{FE2} = 25$ $\quad V_{BE} = 0.7$ V

The value of $h_{ie} = 50$ is determined from the slope of the typical base-characteristics curve ($V_{BE}$ vs. $I_B$) at the 5-milliampere collector-current operating point, where $I_B = I_C/h_{FE2} = 5/25 = 0.2$ milliampere. If actual measurements indicate a different value from that assumed above, the new value is used and $h_{fe}$ is recomputed.

6. The current $I_2$ and voltage $V_2$ are calculated as follows:

$$I_2 = I_1/(h_{FE1}\, h_{FE2}) = 250/(50\times25) = 0.20 \text{ mA}$$

As listed in step (5), the base-to-emitter voltage of the 2N1485 for the design-center collector current of 250 milliamperes is 0.8 volt. For the 2N1481, the base-to-emitter voltage for the design-center collector current of 5 milliamperes is 0.7 volt. Therefore, $V_2$ is given by

$$V_2 = 0.8 + 0.7 = 1.5 \text{ volts}$$

7. A 1N1781 silicon voltage-reference diode is selected on the basis of the following design conditions:

$R_f = 5$ ohms
$V_R = V_o - V_2 = 28 - 1.5 = 26.5$ volts
$I(\text{max}) = I_o + I_2 = 2 + 0.2 = 2.2$ mA
$P_1 = 26.5 \times 2.25 = 60$ milliwatts

8. The series resistance R, which includes both the source resistance $R_S$ and the external resistance $R_1$, is determined for the condition $I_1(\text{max}) = I_L(\text{max})$ as follows:

$$V_S - \Delta V_S = V_o + [R\ I_1(\text{max})]$$
$$49 - 7 = 28 + (0.5\ R)$$
$$R = 28 \text{ ohms}$$

8. The external resistance $R_1$, therefore, becomes

$$R_1 = R - R_S = 28 - 10 = 18 \text{ ohms}$$

The circuit diagram of the shunt voltage regulator that results from this step-by-step procedure is shown in Fig. 328.

## Switching Regulator

Fig. 329 shows the basic configuration for a switching type of

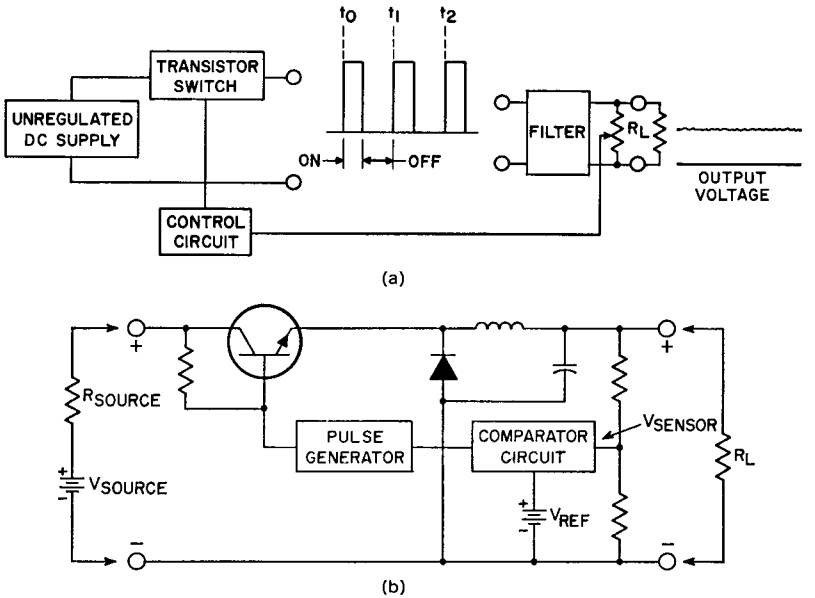



Figure 329. Basic configuration of switching type of transistor voltage regulator: (a) block diagram; (b) schematic diagram.

transistor voltage regulator. In this circuit, the pass transistor is connected in series with the load, and regulation of the output voltage is accomplished by on-off switching of the pass transistor through a feedback circuit. The feedback circuit samples the output voltage and compares it to a reference voltage. The difference (error signal) between the two voltages is used to control the on-off duty cycle of the pass transistor. If the output voltage tends to decrease below the reference voltage, the duration of the on-time pulse increases. The pass transistor then conducts for a longer period of time so that the output voltage increases to the desired level. If the output voltage tends to rise above the reference voltage, the duration of the on-time pulse decreases. The shorter conduction period of the pass transistor then results in a compensating decrease in output voltage. Some type of filter is required between the pass transistor and the load to obtain a smooth dc output. A commonly used filter consists of an LC network and a commutating diode.

The major advantage of the switching regulator over the linear regulator is the higher efficiency that results from the mode of operation of the series pass transistor. In this mode of operation, the transistor is operated in its two most efficient stages, either at cutoff or at saturation. As a result, dissipation is considerably less than when the transistor is operated in the linear region. The response time of the switching regulator, however, is usually slower than that of the linear regulator, but can be improved by operation of this circuit at higher frequencies.

Design considerations and some applications of switching regulators are discussed in the following paragraphs. In addition, other uses of switching regulators are discussed in the section on **Ballast Circuits for Mercury-Arc Lamps.**

**Filter Considerations**—A fundamental part of every switching regulator is the filter. Fig. 330 shows the various types of filters that can be used. Selection of the optimum filter for a power supply is based on the load requirements of the particular circuit and consideration of the basic disadvantages of the various types of filters.

A capacitive filter, shown in Fig. 330(a), has two primary disadvantages: (1) because large peak currents exist, R must be made large enough to limit peak transistor current to a safe value; and (2) the resistance in this circuit introduces loss.

An inductive filter, shown in Fig. 330(b), has three disadvan-

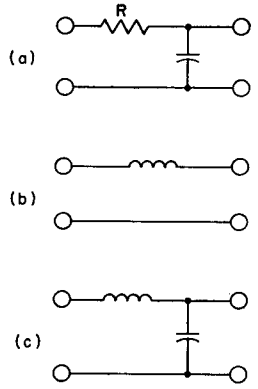



Figure 330. Typical filter circuits for use between pass element and load in a switching regulator: (a) capacitive filter; (b) inductive filter; (c) inductive-capacitive filter.

tages: (1) The inductance may produce a destructive voltage spike when the transistor turns off. This problem, however, can be solved effectively by the addition of a commutating diode, as shown in Fig. 331. This diode

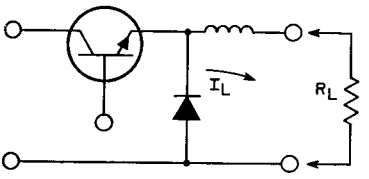



Figure 331. Use of inductance and commutating diode as filter network between pass transistor and load in switching voltage regulator.

commutates the current flowing through the inductor $I_L$ when the transistor switches off. (2) An abrupt change in the load resistance $R_L$ produces an abrupt change in output voltage because the current through the load $I_L$ cannot change instantaneously. (3) A third disadvantage of the inductive filter becomes evident during light loads. The energy stored in an inductor is given by

$$E = \tfrac{1}{2} LI^2 \qquad (286)$$

As a result, the capability of the inductor to store energy varies with the square of the load current. Under light load conditions, the inductor must be much larger to provide a relatively constant current flow when the transistor is off than is required for a heavy load.

Most of the problems associated with either a capacitive filter or an inductive filter can be solved by use of a combination of the two as shown in Fig. 330(c). Because the energy stored in an inductor varies directly as current squared, whereas the energy output at constant voltage varies directly with current, it is not usually practical to design the inductor for continuous current at low current outputs. The addition of a capacitor eliminates the need for a continuous flow of current through the inductor. With the addition of a commutating diode, this filter has the following advantages.

(1) No "lossy" elements are required.

(2) The inductive element need not be oversized for light loads because the capacitance maintains the proper output voltage $V_{out}$ if the inductive current becomes discontinuous.

(3) High peak currents through the transistor are eliminated by the use of the inductive element.

In summary, the switching-regulator filter can take on various forms depending upon the load requirements. However, if a wide range of voltage and current is required, an LC filter is used in combination with a commutating diode.

A practical rule of thumb is to design the inductor to be large enough to dominate the performance during maximum-load conditions. The filter capacitor is chosen to be large enough to dominate performance at mid-range current values and the full range of output voltages.

A primary advantage of the transistor switching regulator is that the switching frequency can be made considerably higher than the line frequency. As a result, the filter can be made relatively small and light in weight.

The means by which the switching regulator removes the line-frequency ripple component is illustrated in Fig. 332. The on time increases under the valley points of the unregulated supply and decreases under the peaks. The net result is to remove the 60-Hz component of ripple and introduce only ripple at the switching frequency which is relatively high frequency and easily filtered out.

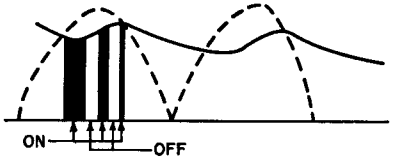



Figure 332. Effect of high-frequency switching of the switching regulator on power-supply ripple component.

**Transistor Parameters**—The transistor parameters affecting the performance of a switching regulator are the current gain $h_{FE}$, the collector-to-emitter saturation voltage $V_{CE(sat)}$, the leakage current $I_{CER}$, forward-bias second-breakdown voltage, and switching times. The forward-current transfer ratio $h_{FE}$ determines the amount of drive current needed. The collector-to-emitter saturation voltage $V_{CE}(sat)$ is important because it determines part of the power loss in the circuit and the dissipation of the transistor during the ON period. The amount of leakage current is important because the transistor essentially conducts this amount of current during the OFF period and thus increases dissipation. If this leakage current is large enough, the transistor can enter into a condition of thermal runaway. Silicon transistors, with their inherently lower leakage-current value, do not often exhibit this problem. Collector-breakdown voltage should be higher than the supply voltages encountered or the maximum voltage that is to be switched by the transistor if several units are connected in series.

The transistor safe-area rating determines the maximum power that can be handled by the transistor and by the supply. This parameter and its implications are explained in detail in the section on **Safe-Area Ratings.** It should be noted, however, that the peak power dissipated by the transistor is also a function of the switching time of the commutating diode. This fact can be demonstrated by examination of the circuit operation. It is assumed that the transistor is off and the commutating diode is conducting. When the transistor turns on, the diode requires some finite time to turn off; therefore, a power pulse is generated during this interval. When the transistor turns off, the inductive load maintains current through the transistor or load while the collector-to-emitter voltage $V_{CE}$ rises to the value of the input voltage $V_{in}$. This condition alone results in a power pulse which is increased by the additional pulse created because the diode does not conduct immediately when the voltage across it reverses. As a result, both turn-on and turn-off transients should be investigated carefully when the peak power requirement of the transistor is determined.

Switching time is also a factor in determination of the maximum power that the transistor is

capable of dissipating. When the transistor turns on, the current flows into the load and into the output capacitor through the inductor. Energy is stored in the inductor and the capacitor so that when the switch is open (i.e., the transistor is cut off) this energy is available to supply the load. During the on time, the current through the inductor is a linear ramp. The rate of increase of current di/dt is determined by L and the voltage across it ($V_{in} - V_{out}$), as follows:

$$dI/dt = (1/L)\,(V_{in} - V_{out}) \qquad (287)$$

The peak current is, therefore, given by

$$I_p = [(V_{in} - V_{out})/L]\,t_{on} \qquad (288)$$

The switching times, $t_r$ (rise time) and $t_f$ (fall time), are of prime consideration in selection of a transistor to be used as the switch. For good regulation over a wide range of input voltage and output current, the duty cycle must be variable from at least 10 to 90 per cent (i.e., the pulse width could be a minimum of one-tenth of the period 1/10f). For low switching losses, the rise and fall times should each be less than 10 per cent of the minimum pulse width. These requirements are summarized as follows:

$$t_r \leq 1/100f;\ t_f \leq 1/100f \qquad (289)$$

where f is the frequency of the pulse generator.

**Switching Arrangement**—The transistor switching arrangement usually takes on one of two forms as illustrated in Fig. 333. If isolated supplies appear in the drive circuits of $Q_1$ and $Q_2$, performance of the two circuits is basically the same. However, if no isolated supplies are used, then the circuit of Fig. 333(b) has the disadvantage that the $V_{CE}$ of $Q_2$ cannot be reduced below the $V_{BE}$ of $Q_2$. This condition results because the base of $Q_2$ cannot be tied to a point more positive than the plus voltage of the power supply.

The circuit of Fig. 333(a) can avoid this problem if the collector of the driver unit is connected to the positive side of the supply. The disadvantage is that current in the driver does not flow through the load; the power associated with this current, therefore, is lost.

The circuit of Fig. 333(b) is usually preferred when the power that results from a high $V_{CE}$(sat) can be tolerated.

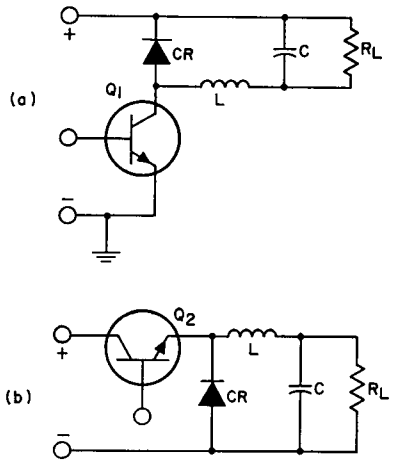



Figure 333. Basic transistor switching arrangements: (a) filter elements and load impedance in collector circuit of switching transistor; (b) filter elements and load impedance in emitter circuit of switching transistor.

**Design of a Practical Switching-Regulator Power Supply**—Power supplies that use switching regulation usually are smaller and lighter and operate more efficiently than conventional supplies. These improvements result from elimination of the need for a 60-Hz power transformer and heat sinks for the transistors.

A complete switching-regulator power supply is described in detail in the sections below. This supply produces 250 watts at 5 volts with an efficiency of 70 per cent. It uses two switching transistors in a push-pull arrangement with variable pulse width; the switching rate is 200 kHz. The complete supply weighs only 10 pounds and occupies only 470 cubic inches.

**Power-supply elements:** The switching-regulator power sup-

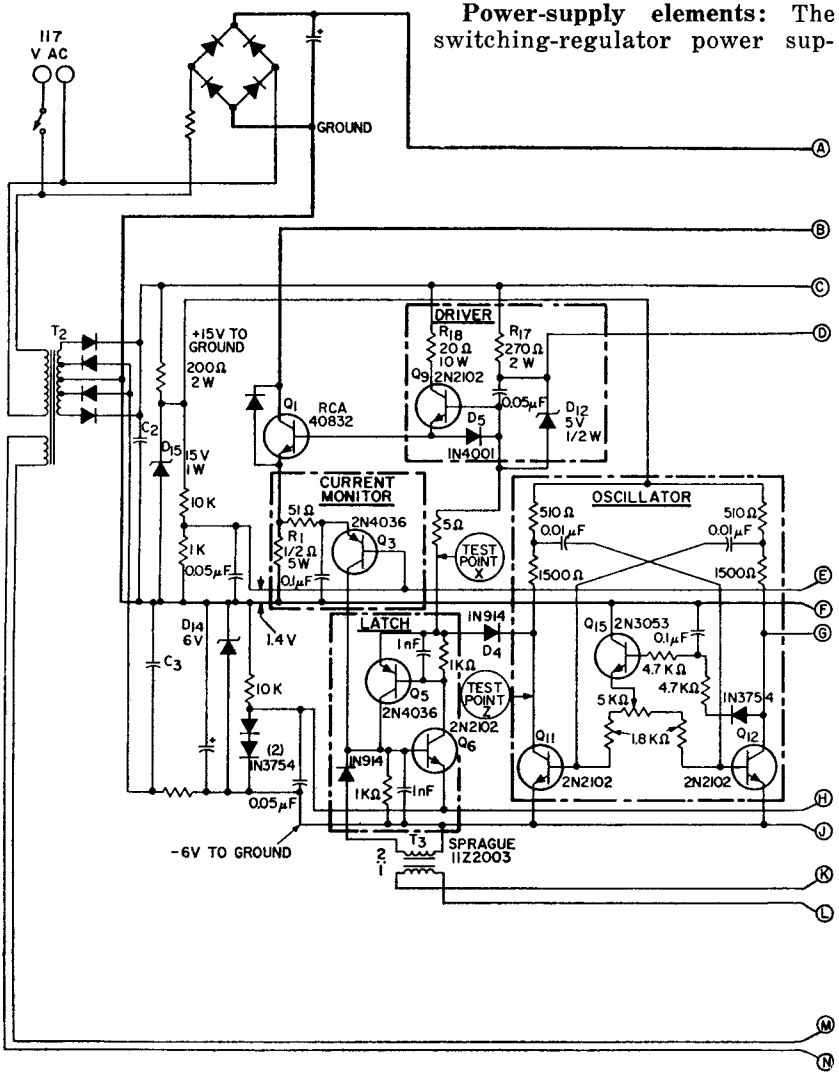



Figure 334. Diagram of switching-regulator power supply (continued on page 293).

ply includes the six major elements shown in the schematic diagram of Fig. 334: (1) the main power supply, (2) the power-switching transistors, (3) the isolation transformer, (4) the modulator circuits, (5) the power rectifiers, and (6) the filter. The important parameters of these elements are discussed below.

The **main power supply** provides the power that ultimately becomes the output power. It rectifies and filters the line voltage without use of a 60-Hz transformer. For a switching-regulator type of power supply, the main supply may be designed for high ripple without increased regulator losses (such as would

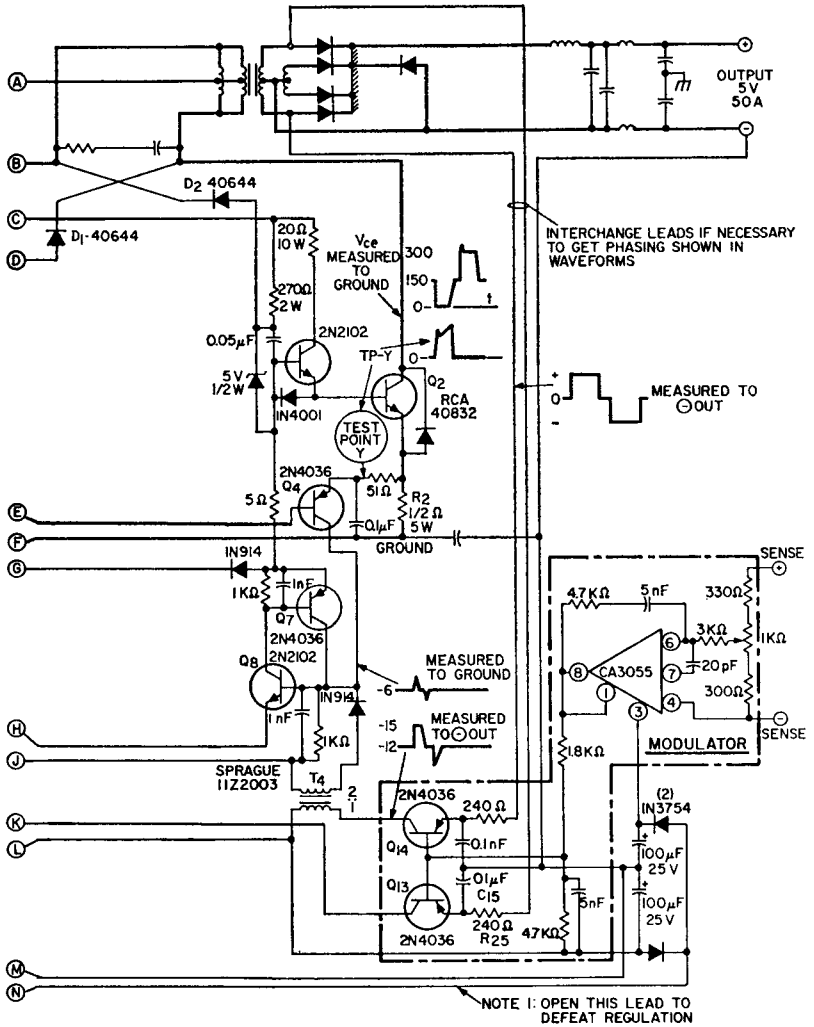



Figure 344. Diagram of switching-regulator power supply (continued from page 292).

occur in a conventional series regulator). Therefore, smaller capacitors and lower-cost rectifiers can be used. Some resistance must be added in series with the power line to prevent damage to the rectifiers during turn-on. The voltage delivered by the main power supply varies with line-voltage and load variations. The peak output voltage of the main supply at the maximum line conditions (with transients) determines both the collector-voltage rating required for the power-switching transistors and the turns ratio of the isolation transformer.

The performance capabilities of the power supply are determined by **the switching transistors,** because they are the parts least able to withstand overloads such as those caused by load faults or misuse. Therefore, the switching transistors must have the following characteristics (listed in order of importance):

(1) High forward-bias second-breakdown capability.

(2) Ability to withstand the required collector voltage while in the cut-off condition.

(3) Short rise and fall times ($t_r$ and $t_f$) for low power dissipation in the transistors and thus high efficiency of the power supply.

(4) Reasonably low $V_{CE}(sat)$ for low dissipation and economical transistor heat sinks.

(5) Stable leakage current ($I_{CEV}$). The magnitude of the leakage is not important (even 20 milliamperes at 500 volts contributes less than 5 watts to the average dissipation per transistor), but it should be stable.

The **isolation transformer** is a ferrite-core transformer that operates at 20 kHz. Its design involves three basic problems: core material selection, windings to keep peak flux below saturation, and compensation for unbalanced direct current.

If a core has too much loss, it will overheat. If it has too many turns, the flux density will be below saturation, but the copper losses will be greater than necessary. The number of turns is kept low to avoid unnecessary copper losses, but must be great enough to keep the peak flux in the core below saturation.

The core will saturate if its cross section is too small, if there are not enough turns in the primary winding, or if the primary direct current is unbalanced. Core saturation causes the power-switching transistors to draw excessive currents that can increase collector dissipation to destructive levels. To prevent these high currents, the power supply includes a monitor circuit that cuts off the base drive to the switching transistors when emitter current reaches the maximum safe value.

Fig. 335 shows the emitter-current waveform of a power-switching transistor, monitored at point Y in Fig. 334, for different numbers of primary turns. If the emitter current is excessive, the circuit reduces the duty cycle to protect the power-switching transistor. Fig. 336 shows the waveforms for unbalanced dc drive. These unbalanced currents result from unequal duty cycles, caused by oscillator unbalance or by unbalance or faults in the modulator. Because such unbalances occur in normal operation, the protective circuits must be included in the design.

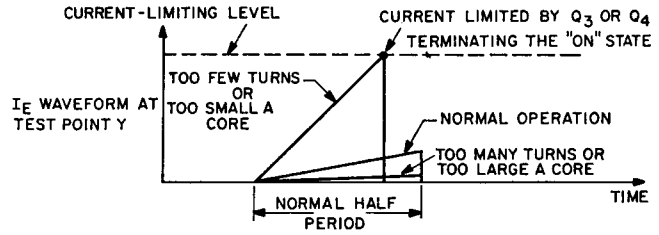



Figure 335. Waveform of emitter current in power-switching transistor showing effects of core size and number of primary turns, with regulation defeated (see note on Fig. 334).

The **modulator circuit (oscillator, drivers, modulators, and latches)**, which is indicated in the circuit diagram, delivers the base drive to the power-switching transistors. The forward drive must be sufficient to keep the transistors saturated under all conditions, and must have a short rise time to provide fast transistor turn-on and low dissipation. The reverse drive must have short rise time and a magnitude equal to or greater than the forward base drive.

The oscillator frequency should be stable to minimize rectifier losses, and should be greater than 20 kHz to eliminate sound. All of the circuits should be insensitive to component-value variations, component drift, and random or stray interference.

Most of the losses in the power supply occur in the **power rectifiers**. Because of the many variables (and unknowns) involved, the rectifier losses should be determined by measurement of circuit efficiency or heat-sink temperature. A total rectifier loss of 45 per cent of the rated output of the regulator is to be expected.

The use of ac power to generate dc outputs that are free of ac signals requires a good **filter**. Moreover, in a power supply that delivers high current, the filter components must be of high quality: the inductor must have high Q, and the capacitor must have both low resistance and low inductance.

The inductance value used is a compromise between the need for a high value to limit peak currents and thus permit good transistor utilization, and the need for a low value to permit fast response to sudden current de-

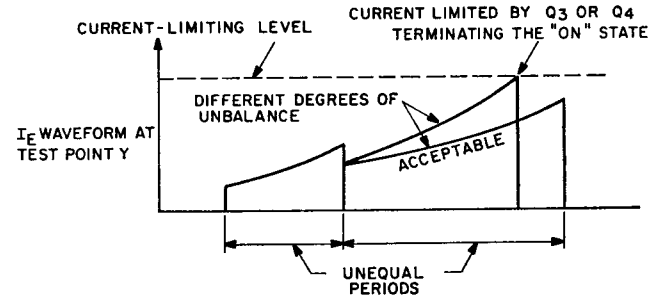



Figure 336. Waveform of emitter current in power-switching transistor showing effect of unbalanced direct current, with regulation defeated and load current of 25 amperes.

mands. The minimum value of inductance is determined by the peak collector current allowed, as follows:

$$L_{min} = \frac{t_{off}(max)\ E_{out}}{n_T\ I_c(peak) - I_{load}} \quad (290)$$

where $n_T$ is the turns ratio of the isolation transformer.

The filter capacitors for this application must be selected for 20-kHz operation. Ceramic and paper types are recommended, but tantalum or high-quality aluminum electrolytics can be used for large values of capacitance. The capacitance must be sufficient to prevent the output voltage from decreasing excessively when the load is suddenly increased and the inductor supplies less than the load current. The minimum capacitance is given by

$$C_{min} = \frac{I_{load}[t_{dis} + 2t_{off}(max)]}{2(\Delta V)\ allowed} \quad (291)$$

where

$$t_{dis} = \frac{L\ I_{load}}{\frac{V_{CC}(min)}{n_T} - V_o - 1.0} \quad (292)$$

and $t_{off}(max)$ is 12.5 microseconds for this design.

**Power-supply performance:** The power supply shown in Fig. 334 can deliver a load current of 50 amperes at 5 volts. All of the pulse-width modulation circuits, drivers, and latches are duplicated for each power-switching transistor. This duplication uses more than the minimum number of components, but it provides wide design margins and reliable operation.

Voltage regulation and overload regulation are accomplished by reducing the duty cycle of the power-switching transistors. The duty cycle is reduced by triggering the latches on, either from pulse transformers $T_3$ and $T_4$ to regulate the output voltage, or from transistors $Q_3$ and $Q_4$ to prevent excessive emitter currents in the power-switching transistors. The excessive currents could be caused by overloads at the output or by transformer core saturation resulting from unbalanced duty cycles.

Input-to-output isolation is maintained through the main isolation transformer( $T_1$), the 60-Hz transformer ($T_2$), and the pulse transformers ($T_3$ and $T_4$). This circuit isolation is indicated in Fig. 334.

This power supply is capable of operating into any load impedance, including short circuits, without damage. It can operate at duty cycles from less than 10 per cent to 100 per cent. With a duty cycle of 100 per cent, the supply operates as a straight inverter at the full capacity of the transistors, transformers, and rectifiers.

The base drive for the power-switching transistors is direct-coupled, and is supplied by an unregulated low-voltage power supply that operates from a 60-Hz transformer. Direct coupling of the base drive provides positive control over transistor bias. The reverse base drive is supplied by the two-transistor latch circuits $Q_5$ and $Q_6$ or $Q_7$ and $Q_8$, or by the oscillator transistors ($Q_{11}$ and $Q_{12}$) if the duty cycle is 100 per cent. The reverse base voltage is obtained from a 6-volt regulated supply.

The frequency is controlled by the astable transistor oscillator that operates from 15-volt and −6-volt regulated sources. A potentiometer for equalization of the duty cycle is shown, but is not normally required. Transistor $Q_{15}$ insures that the oscillator does not "hang up."

Common-mode conduction is reduced by cross-coupled diodes $D_1$ and $D_2$. These diodes conduct when the $V_{CE}$ of the power-switching transistor is less than 5 volts (breakdown of the zener diode), and provent conduction of the opposite power-switching transistor. These diodes are of critical importance because the storage time of the power-switching transistors is several microseconds at light load conditions ($I_{B_1} > 0.5$ amperes and $I_C < 0.5$ amperes).

A major consideration in the design of this power supply is the protection of the switching transistors and the load circuit from damage caused by transients or faults in the modulator. The faults most likely to occur are lock-up in the oscillator, transient turn-on of the latching transistors caused by dv/dt at point X in Fig. 334, and magnetic pickup in the pulse transformers. The circuit is designed so that any of these faults will cause the power-switching transistors to turn off; this design protects the transistors and keeps the output voltage low. The overcurrent protection circuit is made independent of the proper functioning of the output regulator or its associated circuits, and is dc-coupled to minimize the possibility of failure. Finally, if the low-voltage supplies fail, the output voltage merely falls to zero without any harmful surges.

**Step-Down Switching Regulator**—A transistor switching regulator can be used as a dc step-down transformer. This circuit is a very efficient means of obtaining a low dc voltage directly from a high-voltage ac line without the need for a step-down transformer. Fig. 337 shows a typical step-down transistor switching regulator. This regulator utilizes the dc voltage obtained from a rectified 117-volt line to provide a constant 60-volt supply. Fig. 338 shows the performance characteristics for this circuit.

## Phase-Controlled Regulated Power Supply

In a different type of pulse-width-modulated switching regulator, the pass element is switched at the line frequency and the conduction angle is varied to obtain the desired pulse width. This type of control is generally used with SCR's because turn-on of an SCR is simple and turn-off is accomplished automatically when the line voltage reverses.

Fig. 339 shows the circuit configuration for a regulated dc power supply that uses an SCR as a series pass element. This type of circuit is designed to provide approximately 125 volts, regulated to ± 3 per cent for both line and load. Ripple is less than 0.5 per cent rms.

The power supply is basically a half-wave phase-controlled rectifier. The 5-microfarad capacitor between the cathode and gate of the SCR charges up during half of each cycle and is discharged by the firing of the SCR. The firing angle of the SCR is advanced or retarded by the charg-

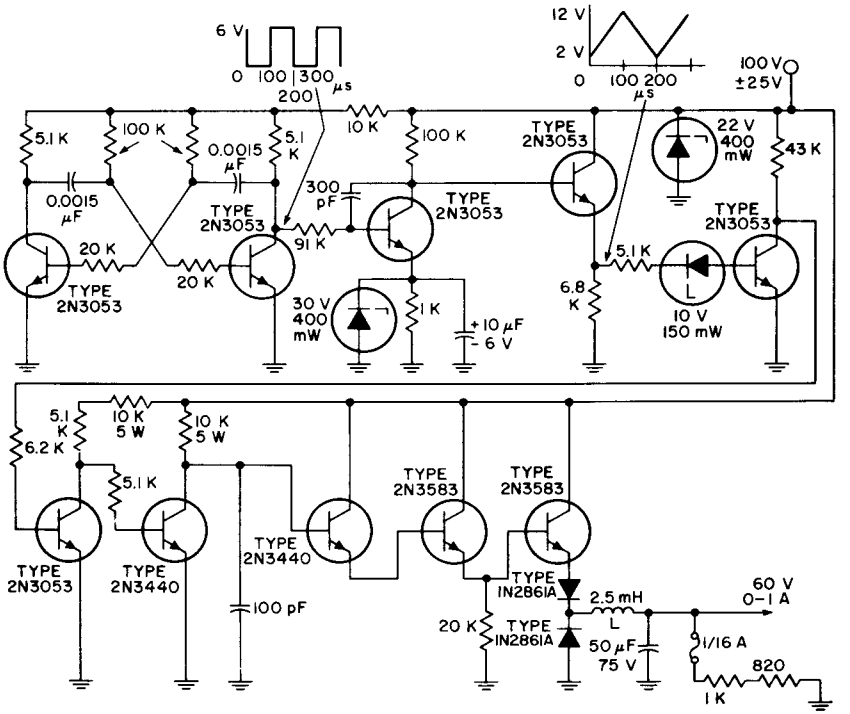
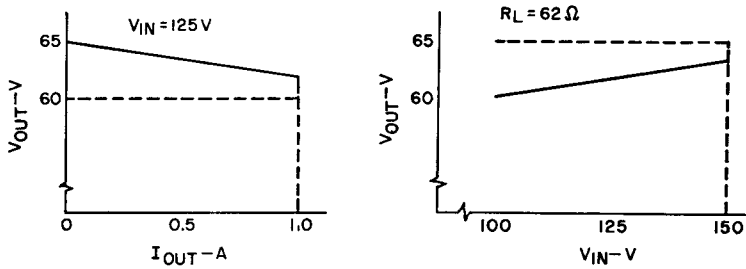



Figure 337. Typical step-down transistor switching regulator.

ing current flowing into the capacitor. Some of the current which would normally charge this capacitor is shunted by the collector of the RCA-40424 control transistor. As the current in the control transistor increases, current is shunted around the capacitor, through the ballast lamp, so that the capacitor charging time is increased. As a result, the firing angle of the SCR is retarded, and a lower output voltage results.




Figurfe 338. Performance characteristics of step-down switching regulator shown in Fig. 337.

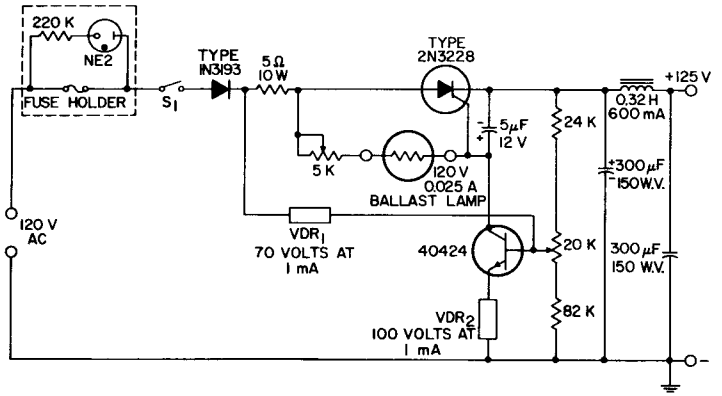



Figure 339. SCR regulated power supply.

The controlling voltage on the control transistor is derived from both the dc output and the line voltage in such a manner as to provide load and line regulation respectively. The voltage-dependent resistor ($VDR_2$) in the base circuit of the control transistor decreases resistance for an increase in line voltage and thus increases base current (and collector current) as line voltage is increased. In addition, the ballast lamp exhibits an increase in resistance with increasing line voltage, and thus tends to retard the firing angle of the SCR. Changes in dc output voltage that result from variations in load current are fed back to the base of the control transistor by a voltage divider at the input to the filter in the proper polarity to adjust collector current in a direction to compensate for changes in dc output voltage.

# Power Conversion

IN many applications, the optimum value of voltage is not available from the primary power source. In such instances, dc-to-dc converters or dc-to-ac inverters may be used, with or without regulation, to provide the optimum voltage for a given circuit design.

An **inverter** is a power-conversion device used to transform dc power to ac power. If the ac output is rectified and filtered to provide dc again, the over-all circuit is referred to as a **converter.** The purpose of the converter is then to change the magnitude of the available dc voltage.

Power-conversion circuits, both inverters and converters, consist basically of some type of "chopper," which is used to develop a wave shape that is acceptable to a transformer, and the transformer. The design of the transformer is an important consideration because this component determines the size and frequency of the converter (or inverter), influences the amount of regulation required after the conversion or inversion is completed, and provides the transformation ratio necessary to assure that the desired value of output voltage is delivered to the load circuit. The chopping or switching function in the inverter circuit is usually performed by high-speed transistors or SCR's connected in series with the primary winding of the output transformer.

Inverters may be used to drive any equipment which requires an ac supply, such as motors, ac radios, television receivers, or fluorescent lighting. In addition, an inverter can be used to drive electromechanical transducers in ultrasonic equipment, such as ultrasonic cleaners and sonar detection devices. Similarly, converters may be used to provide the operating voltages for equipment that requires a dc supply.

Transistor and SCR inverters can be made very light in weight and small in size. They are also highly efficient circuits and, unlike their mechanical counterparts, have no moving components.

## TRANSISTOR INVERTERS AND CONVERTERS

Several types of transistor circuits may be used to convert a steady-state dc voltage into either an ac voltage (inversion) or another dc voltage (conversion). The simplest converter circuit is the

blocking-oscillator, or ringing-choke, power converter which consists of one transistor and one transformer. More complex circuits use two transistors and one or two transformers.

## Basic Circuit Configurations

Fig. 340 shows the basic circuit configuration for a **ringing-choke dc-to-dc converter.** In this converter, a blocking oscillator (chopper circuit) is transformer-coupled to a half-wave rectifier

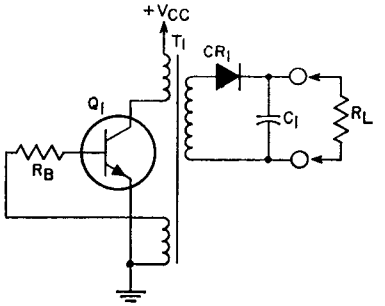



Figure 340. Basic circuit configuration for a ringing-choke dc-to-dc converter.

type of output circuit. The rectifier converts the pulsating oscillator output into a fixed-value dc output voltage.

When the oscillator transistor $Q_1$ conducts (as a result of either a forward bias or external drive), energy is transferred to the collector inductance presented by the primary winding of transformer $T_1$. The voltage induced across the transformer feedback winding connected to the transistor base through resistor $R_B$ increases the conduction of $Q_1$ until the transistor is driven into saturation. The rectifier diode $CR_1$ in series with the secondary winding of transformer $T_1$ is oriented so that no power is delivered to the load circuit during this portion of the oscillator cycle.

With transistor $Q_1$ in saturation, the collector current through the primary inductance of transformer $T_1$ rises linearly with time ($-di/dt = E/L$) until the base drive supplied by the transformer feedback winding can no longer maintain $Q_1$ in saturation. As the current through $Q_1$ decreases from the saturation level, the voltage induced into the feedback winding decreases, and transistor $Q_1$ is rapidly driven beyond cutoff. The energy stored in the collector inductance (primary of transformer $T_1$) is released by the collapsing magnetic field and coupled by the secondary winding of transformer $T_1$, through rectifier diode $CR_1$, to the load resistance $R_L$ and filter capacitor $C_1$. The filter capacitor stores the energy it receives from the collector inductance. When no current is supplied to the load circuit from the oscillator (i.e., during conduction of transistor $Q_1$), capacitor $C_1$ supplies current to the load resistance $R_L$ to maintain the output voltage at a relatively constant value. The switching action of rectifier diode $CR_1$ prevents any decrease of the energy stored by capacitor $C_1$ because of the negative pulse coupled from the oscillator during the periods that transistor $Q_1$ conducts.

The operating efficiency of the ringing-choke inverter is low, and the circuit, therefore, is used primarily in low-power applications. In addition, because power is delivered to the output circuit for only a small fraction of the oscillator cycle (i.e., when $Q_1$ is not conducting), the circuit has a relatively high ripple factor

which substantially increases output filtering requirements. This converter, however, provides definite advantages to the system designer in terms of design simplicity and compactness.

The **push-pull switching inverter** is probably the most widely used type of power-conversion circuit. For inverter applications, the circuit provides a square-wave ac output. When the inverter is used to provide dc-to-dc conversion, the square-wave voltage is usually applied to a full-wave bridge rectifier and filter. Fig. 341 shows the basic configuration for a push-pull switching converter. The single saturable transformer controls circuit switching and provides the desired voltage transformation for the square-wave output delivered to the bridge rectifier. The rectifier and filter convert the square-wave voltage into a smooth, fixed-amplitude dc output voltage.

When the voltage $V_{CC}$ is applied to the converter circuit, current tends to flow through both switching transistors $Q_1$ and $Q_2$. It is very unlikely, however, that a perfect balance can be achieved between corresponding active and passive components of the two transistor sections; therefore, the initial flow of current through one of the transistors is slightly larger than that through the other transistor. If transistor $Q_1$ is assumed to conduct more heavily initially, the rise in current through its collector inductance causes a voltage to be induced in the feedback windings of transformer $T_1$ which supply the base drive to transistors $Q_1$ and $Q_2$. The base-drive voltages are in the proper polarity to increase the current through $Q_1$ and to decrease the current through $Q_2$. As a result of regenerative action, the conduction of $Q_1$ is rapidly increased, and $Q_2$ is quickly driven to cutoff.

The increased current through $Q_1$ causes the core of the collector inductance to saturate. The inductance no longer impedes the rise in current, and the transistor current increases sharply into the saturation region. For this condition, the magnetic field about the collector inductance is constant, and no voltage is induced in the feedback windings of transformer $T_1$. With the cutoff base voltage removed, current is allowed to flow

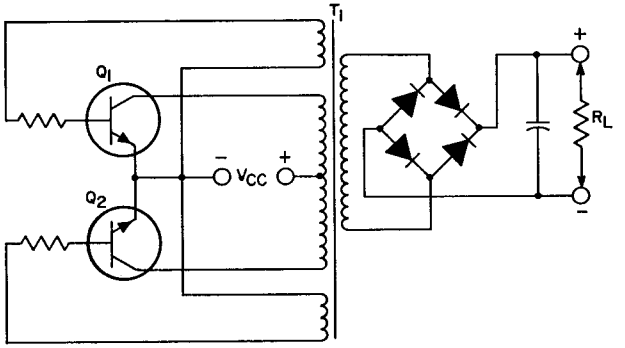



Figure 341. Basic circuit configuration of a single-transformer push-pull switching converter.

through transistor $Q_2$. The increase in current through the collector inductance of this transistor causes voltages to be induced in the feedback windings in the polarity that increases the current through $Q_2$ and decreases the current through $Q_1$. This effect is aided by the collapsing magnetic field about the collector inductance of $Q_1$ that results from the decrease in current through this transistor. The feedback voltages produced by this collapsing field quickly drive $Q_1$ beyond cutoff and further increase the conduction of $Q_2$ until the core of the collector inductance for this transistor saturates to initiate a new cycle of operation. The square wave of voltage produced by the switching action of transistors $Q_1$ and $Q_2$ is coupled by transformer $T_1$ to the bridge rectifier and filter, which develop a smooth, constant-amplitude dc voltage across the load resistance $R_L$. The small ripple produced by the square wave greatly simplifies filter requirements.

Push-pull transformer-coupled converters with full-wave rectification provide power to the load continuously and are, therefore, well suited for low-impedance, high-power applications. Although not as economical as the ringing-choke design, the push-pull configuration provides higher efficiency and improved regulation.

Fig. 342 shows a **four-transistor, single-transformer bridge configuration** that is often used in inverter or converter applications. In this type of circuit, the primary winding of the output transformer is simpler and the breakdown-voltage requirements of the transistors are reduced to one-half those of the transistors in the push-pull converter shown in Fig. 341.

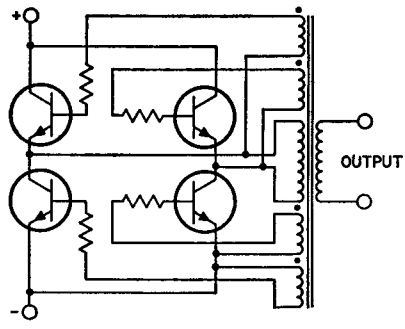



Figure 342. Basic circuit configuration of a four-transistor, single-transformer bridge inverter.

Fig. 343 shows the schematic diagram for a **two-transistor, two-transformer converter.** In this circuit, a small saturable transformer provides the base drive for the switching transistors, and a nonsaturable output transformer provides the coupling and desired voltage transformation of the output delivered to the load circuit. With the exception that it uses a separate saturable transformer, rather than feedback windings on the output transformer, to provide base drive for the transistors, this converter is very similar in its operation to the basic push-pull converter shown in Fig. 341. The saturable-transformer technique may also be applied in the design of a bridge converter, as shown in Fig. 344.

## Transformer Considerations

The selection of the proper core material in the design of a transformer to be used in an inverter depends on the power-handling requirements, operating frequency, and operating temperature of the

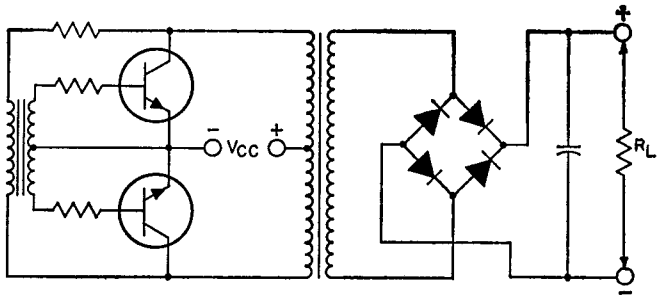



Figure 343. Basic circuit configuration of a two-transformer push-pull switching converter.

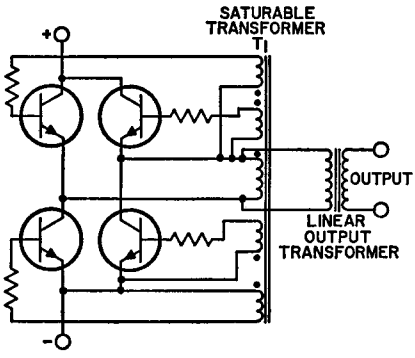



Figure 344. Basic configuration of a four-transistor bridge inverter that uses a saturable output transformer.

inverter. For high-frequency applications, the ferrite core is superior to the iron type in both performance and economy. Even at low frequencies, ferrite cores may be more economical because the iron type must be made in thin laminations or in the form of a tape-wound toroid.

Power loss in ferrite is approximately a linear function of frequency up to 40 kHz. Above this frequency, eddy-current losses decrease the efficiency of most ferrites. Laminated iron cores are normally restricted to frequencies below 10 kHz.

The operating temperature of the transformer is an important consideration in the choice of the particular ferrite core. For many ferrite cores, the Curie temperature is low. The manufacturer's data on ferrite material indicate the maximum operating temperature which, together with the variation in flux density as a function of temperature and the desired flux density (B), must be considered to select the proper core.

Another important consideration is the efficiency of the transformer. The transformer efficiency desired can be used to obtain an approximation of allowable magnetic power, $P_M$, dissipated by the transformer. When $P_M$ and the core loss factor are known, the maximum volume of core material which can be used is estimated. The core loss factor at the operating frequency is obtained from the manufacturer's data.

The remaining design considerations follow the conventional rules of transformer design. The size of the wire must be large enough to assure that copper losses are low. The selection is made on the basis of a 50-per-cent duty cycle. If the wire size is too small, copper losses will be appreciable and cause an increase in core temperatures. In high-power, high-frequency inverters,

a large number of turns in the primary should be avoided to minimize copper losses and maintain a low value of leakage inductance. Moreover, because of the relatively small size of the core and the large size of wire that must be used, a large number of turns may be physically impossible. Good balance and close coupling between primaries is normally achieved by the use of bifilar windings.

## Additional Considerations

In addition to the previous considerations for ringing-choke-type and transformer-coupled types of dc-to-dc converters, there are other factors which may require consideration in practical designs, such as starting-bias methods, the use of voltage-multiplication techniques, and maximum operating temperature. Excellent starting under heavy load conditions may be obtained by the use of a transistor-type switch which will provide a large starting bias and then be cut off by the buildup of the output voltage. It is also possible to obtain satisfactory starting by the use of a fixed bias resistance, provided the value of this resistance is high enough so that it does not materially affect normal switching.

For dc output voltages higher than those given in the particular design procedure, a voltage-multiplier-type rectifier circuit may be used to avoid use of larger transformer step-up ratios. Although the use of a voltage-multiplier circuit results in a reduction in over-all efficiency, this condition is more acceptable than one which results in higher copper losses, magnetic-coupling problems, and higher core losses that may result from the use of higher transformer step-up ratios.

The transistor requirements given later in Tables XXI through XXIV are for operation at a case or flange temperature of 55°C. To convert case or flange temperature to ambient temperature, it is necessary to know the thermal resistance between the transistor and free air. This resistance is a function of the contact resistance between the transistor case or flange and the chassis; the thermal resistance of any insulating washer used; the size, thickness, and material of the chassis; and the method used to cool the chassis (for example, forced-air cooling, water cooling, or simple convection cooling).

To assure reliable operation at any permissible ambient temperature, care must be taken that the collector-junction temperature of the transistor is not greater than that specified by the manufacturer. The average temperature of the junction $T_{J(av)}$ is equal to the ambient temperature plus the product of the average power dissipated in the transistor and the thermal resistance between junction and case plus the case-to-air thermal resistance as indicated by the following equation:

$$T_{J(AV)} = T_A + P_{AV}\,\theta_{J-C} + \theta_{C-A} \quad (293)$$

The average junction temperature calculated by use of Eq. (293) is equivalent to the effective case temperature $T_{C(eff)}$ usually given on transistor safe-area-rating charts. The effects of switching on the instantaneous temperature must be evaluated by use of standard safe-area techniques, as described in the section on **Safe-Area Ratings.**

## Design of Practical Transistor Inverters

The design of practical inverter (or converter) circuits involves, essentially, selection of the proper transistors and design of the transformers to be used. The particular requirements for the transistors and transformers to be used are specified by the individual circuit design. The following paragraphs discuss the design of three basic inverter circuits: the simple one-transistor, one-transformer (ringing-choke) type and two push-pull switching converters (a two-transistor, one-transformer type and a two-transistor, two-transformer type). The operation of each circuit is described, design equations are derived, and a sample design is presented.

**One-Transistor, One-Transformer Converter**—Fig. 345(a) shows the basic configuration for a practical circuit of a ringing-choke converter, which is basically a one-transistor, one-transformer circuit. Fig. 345(b) shows the waveforms obtained during an operating cycle.

During the "on" or conduction period of the transistor ($t_{on}$), energy is drawn from the battery and stored in the inductance of the transformer. When the transistor switches off, this energy is delivered to the load. At the start of $t_{on}$, the transistor is driven into saturation, and a substantially constant voltage, waveform A in Fig. 345(b), is impressed across the primary by the battery. This primary voltage produces a linearly increasing current in the collector-primary circuit, waveform B. This increasing current induces substantially constant voltages in the base windings, shown by waveform C, and in the secondary winding.

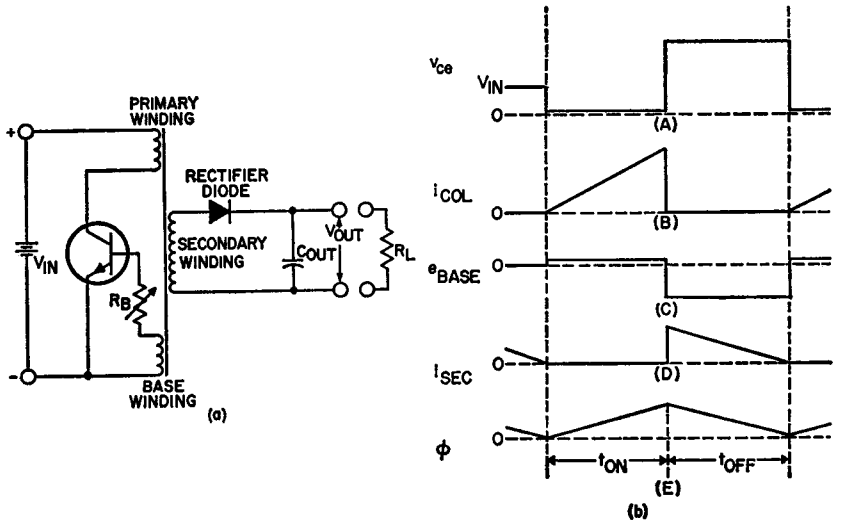



Figure 345. Ringing-choke converter circuit: (a) Schematic diagram; (b) Typical operating waveforms in a ringing-choke converter—(A) primary voltage; (B) primary current; (C) base-to-emitter voltage; (D) secondary current; (E) magnetic flux in transformer core.

The resulting base current is substantially constant and has a maximum value determined by the base-winding voltage, the external base resistance $R_B$, and the input conductance of the transistor. Because the polarity of the secondary voltage does not permit the rectifier diode to conduct, the secondary is open-circuited. Therefore, during the conduction period of the transistor $t_{on}$, the load is supplied only by energy stored in the output capacitor $C_{out}$.

The collector-primary current increases until it reaches a maximum value $\hat{I}_p$ which is determined by the maximum base current and base voltage supplied to the transistor. At this instant, the transistor starts to move out of its saturated condition with the result that the collector-primary current and the voltage across the transformer windings rapidly decrease, and "switch-off" occurs.

After the transistor has switched off, the circuit starts to "ring", i.e., the energy stored in the transformer inductance starts to discharge into the stray capacitance of the circuit, with the result that the voltages across the primary, base, and secondary windings reverse polarity. These reverse voltages rapidly increase until the voltage across the secondary winding exceeds the voltage across the output capacitor. At this instant the diode rectifier starts to conduct and to transfer the energy stored in the inductance of the transformer to the output capacitor and load. Because the output capacitor tends to hold the secondary voltage substantially constant, the secondary current decreases at a substantially constant rate, as shown by waveform D in Fig. 345(b). When this current reaches zero the transistor switches on again, and the cycle of operation repeats.

**Design equations:** In the following derivation of the equations used to design the one-transistor, one-transformer inverter, all references to transformer windings, circuit components, voltages, and currents are for the circuit shown in Fig. 345(a). All references to time periods and waveforms apply to those shown in Fig. 345(b).

The average value of the current in the primary winding of the transformers $I_{P(AV)}$ may be expressed as follows:

$$I_{P(AV)} = I_{out} (V_{out}/\eta V_{in}) \quad (294)$$

where $I_{out}$ is the load current in amperes, $V_{out}$ is the dc output voltage in volts, $\eta$ is the circuit-efficiency factor as given in Table XXI, and $V_{in}$ is the dc input voltage in volts.

The peak value $\hat{I}_p$ of the triangular waveform of primary current, waveform B in Fig. 345(b), is expressed by the following equation:

$$\hat{I}_P = 2I_{P(AV)} (T/t_{on}) = (2T/\eta t_{on}) (V_{out}/V_{in}) I_{out} \quad (295)$$

where $T = t_{on} + t_{off}$, and is in seconds.

If the saturation resistance of the transistor ($R_{sat}$) and the resistance of the primary winding ($R_p$) are sufficiently small, essentially the full dc input voltage $V_{in}$ is impressed across the primary during $t_{on}$, and the required primary inductance $L_P$ in henries may be determined from the following relation:

$$L_P (dI_P/dt) = V_{in} \quad (296)$$

For the triangular current waveform, waveform B in Fig. 345(b), the following equation relates the rate of change of primary current ($dI_P/dt$) to peak current ($\hat{I}_P$) and the transistor conduction period $t_{on}$:

$$dI_P/dt = \hat{I}_P/t_{on} \qquad (297)$$

When this relationship [Eq. (297)] is used, Eq. (296) may be rewritten in the following form:

$$L_P = (V_{in}/\hat{I}_P)\, t_{on} \qquad (298)$$

With $\hat{I}_p$ defined as indicated in Eq. (295), the primary inductance $L_p$ may then be expressed as follows:

$$L_P = \frac{\eta\, V_{in}^2\, t_{on}^2}{2\, V_{out}\, I_{out}\, T} = \left(\frac{\eta\, V_{in}^2}{2\, P_{out}}\right)\left(\frac{t_{on}^2}{T^2}\right)\left(\frac{1}{f}\right) \qquad (299)$$

where f is the operating frequency in hertz, and is equal to 1/T.

Before Eq. (299) can be evaluated, the ratio $t_{on}/T$ must be determined. From Eq. (298), the conduction period of the transistor $t_{on}$ is given by

$$t_{on} = L_P\hat{I}_P/V_{in} \qquad (300)$$

An equation for the off period ($t_{off}$) can be derived from the following relationship:

$$V'_{out} = L_P\, (dI_s'/dt) \qquad (301)$$

where $I_s'$ is the secondary current referred to the primary [$I_s' = I_s\,(N_s/N_P)$], $V'_{out}$ is the secondary voltage referred to the primary [$V'_{out} = V_{out}\,(N_o/N_s)$], $N_P$ is the number of turns on the primary, and $N_s$ is the number of turns on the secondary.

The derivative $dI_s'/dt$ defines the slope of the secondary current (referred to the primary) as a function of time (i.e., $dI_s'/dt = \hat{I}_P/t_{off}$). The off period, therefore, may be determined as follows:

$$t_{off} = L_P\, (\hat{I}_P/V'_{out}) \qquad (302)$$

The total period of oscillation ($T = t_{on} + t_{off}$), as determined from the summation of Eqs. (300) and (302), is given by the following equation:

$$T = L_P\hat{I}_P\left(\frac{1}{V_{in}} + \frac{1}{V'_{out}}\right) \qquad (303)$$

The ratio of on time to total period of oscillation, therefore, can be expressed as follows:

$$\frac{t_{on}}{T} = \frac{L_P\hat{I}_P\left(\frac{1}{V_{in}}\right)}{L_P\,\hat{I}_P\left(\frac{1}{V_{in}} + \frac{1}{V'_{out}}\right)} = \frac{V'_{out}}{V_{in} + V'_{out}} \qquad (304)$$

This relationship is used in Eq. (299) to obtain the following expression for the required primary inductance:

$$L_P = \left(\frac{\eta\, V_{in}^2}{2\, P_{out}\, f}\right)\left[\frac{(N_P/N_S)^2\, V_{out}^2}{V_{in}^2 + (N_P/N_S)^2 + 2\, V_{in}\, V_{out}\, (N_P/N_S)}\right] \qquad (305)$$

The required primary inductance $L_P$ can also be expressed in terms of the maximum permissible flux-density swing in the transformer core, which is given by

$$\frac{\Delta\phi}{A} = \hat{I}_P\left(\frac{4\pi\, N_P}{10}\right)\left[\frac{1}{l_i/\mu_i + l_a/\mu_a}\right] \qquad (306)$$

where A is the cross-sectional area of the core in square centimeters, $l_i$ is the length of the magnetic path in centimeters, $l_a$ is the length of the air gap in centimeters, $\mu_i$ is the permeability of the core material, and $\mu_a$ is the permeability of air ($\mu_a = 1$).

The maximum permissible flux-density swing may also be expressed as follows:

$$\frac{\Delta\phi}{A} = \frac{\phi_{(max)}}{A} - \frac{\phi_{(res)}}{A} \qquad (307)$$

where $\phi_{(max)}$ is the saturation flux density for the core material and $\phi_{(res)}$ is the residual flux density in the core. The primary inductance can be defined in terms of $\Delta\phi$ by means of the following relation:

$$\Delta\phi = (L_P \hat{I}_P / N_P) \times 10^8 \qquad (308)$$

or

$$L_P = (N_P \Delta\phi \times 10^{-8})/\hat{I}_P \qquad (309)$$

Eqs. (307), (308), and (309) can be combined to provide the following expression for the required primary inductance:

$$L_P = \left(\frac{4\pi N_P^2}{10^9}\right)\left[\frac{1}{(l_i/\mu_i)\,A + (l_a/\mu_a)}\right] \qquad (310)$$

The length of air gap $l_a$ should be adjusted to assure operation of the core near but not in the saturation region (i.e., at a maximum flux density slightly less than the saturation value for the core material used). The value of the flux density can be checked by means of Eq. (308).

The induced voltage in the base windings must provide a base-to-emitter voltage sufficiently large to supply the required peak primary current for any transistor of the type to be used in the circuit. The primary voltage $V_P$ at the end of the conduction time $t_{on}$ is expressed by the following equation:

$$V_P = V_{in} - \hat{I}_P (R_{sat} + R_P + R_{supply}) \qquad (311)$$

The required number of turns for the base winding is given by

$$N_B = \frac{N_P V_B}{V_{in} - \hat{I}_P (R_{sat} + R_P + R_{supply})} \qquad (312)$$

$V_B$ is chosen to be twice the necessary $V_{BE(max)}$ for the transistor used. One-half of the voltage $V_B$ is dropped across $R_B$. This voltage division helps compensate for variations in $V_{BE}$ from one transistor to another and at different temperatures. Eq. (312) may be rewritten to express $N_B$ in terms of $V_{BE}$ as follows:

$$N_B = \frac{2N_P V_{BE(max)}}{V_{in} - \hat{I}_P (R_{sat} + R_P + R_{supply})} \qquad (313)$$

The required number of turns for the secondary $N_S$ is determined from the relationship of input voltage $V_{in}$, output voltage $V_{out}$, number of turns on the primary $N_P$, and maximum allowable collector-to-base voltage $V_{CB(max)}$ for the transistor (the value during $t_{off}$), as follows:

$$V_{CB(max)} = V_{in} + (N_P/N_S)\,V_{out} \qquad (314)$$

or

$$N_S = \frac{N_P V_{out}}{V_{CB(max)} - V_{in}} \qquad (315)$$

The external base resistance $R_B$ is necessary to compensate for individual differences in base-to-emitter voltage $V_{BE}$ of the transistors used. The required value

for this resistance may be determined from the following relationship:

$$R_B = V_{BE(max)}/I_B \quad (316)$$

where $V_{BE(max)}$ is the maximum allowable value of $V_{BE}$ for the transistor type used and $I_B$ is the typical value of base current required to provide the peak primary current $\hat{I}_P$ at a base-to-emitter voltage $V_{BE}$. Transient losses that occur during switch-off because of hole-storage effects can be minimized if the value of $R_B$ is maintained as small as possible.

The peak secondary current $\hat{I}_s$ and the peak secondary voltage $V_s$ can be expressed by the following equations:

$$\hat{I}_S = \hat{I}_P (N_P/N_S) \quad (317)$$

$$V_S = (V_{CB(max)} - V_{in}) (N_S/N_P) \quad (318)$$

The value of the output capacitor $C_{out}$ should be such that the time constant $C_{out} R_L$ is at least 10 times larger than $t_{on}$ to assure that the output voltage $V_{out}$ remains essentially constant. The capacitance value required for this condition is determined as follows:

$$R_L = V_{out}^2/P_{out}$$
$$C_{out}R_L = 10\ t_{on}$$
$$C_{out} = (10\ t_{on}\ P_{out})/V_{out}^2 \quad (319)$$

The optimum ratio of primary-to-secondary winding space $K^2$ for the transformer is determined from the following expression:

$$K = \frac{1}{\sqrt{(N_S/N_P)\ (V_{in}/V_{out})+1}} \quad (320)$$

**Sample design:** The application of the relationships derived in the preceding section to the design of a practical ringing-choke inverter can best be illustrated by use of a numerical example. In this example, a converter is designed to convert 12 volts dc to 150 volts dc with a maximum power output of 1 watt.

The first step in the design of any converter is the selection of the transistor(s) to be used. For the ringing-choke circuit, the list of typical design parameters given in Table XXI provides an excellent basis for this selection. For specified values of dc output voltage $V_{out}$, output power $P_{out}$, and dc input voltage $V_{in}$, this table indicates the maximum allowable transistor saturation resistance $R_{sat}$ and the minimum transistor rating requirements for collector-to-base breakdown voltage $V_{CB(sat)}$, peak collector current $I_{C(pk)}$, and maximum allowable power dissipation $P_T$. The transistor selected must satisfy these basic requirements.

For $V_{out}$ = 150 volts, $P_{out}$ = 1 watt, and $V_{in}$ = 12 volts, Table XXI indicates that the transistor used in the converter should have a saturation resistance $R_{sat}$ that does not exceed 10 ohms, a collector-to-base breakdown voltage $V_{CB(max)}$ of at least 35 volts, a peak collector current rating $I_{C(pk)}$ of at least 400 milliamperes, and a dissipation rating $P_T$ at 55°C of at least 80 milliwatts. The collector-to-emitter voltage of the transistor must also be considered. During the transistor off period, the worst-case value of this voltage is equal to the product of the base-winding-to-primary-winding turns ratio ($N_B/N_P$) and the maximum collector-to-base voltage [$V_{CB(max)}$]. If this product exceeds the base-to-emitter breakdown voltage of the transistor, a

Table XXI—Typical Design Parameters For Ringing-Choke-Type DC-To-DC Converters That Have Output Ratings Up To 50 Watts.

| APPLICATION REQUIREMENTS | | | TRANSISTOR REQUIREMENTS | | | | TRANSFORMER-CORE PARAMETERS | | CIRCUIT EFFICIENCY |
|---|---|---|---|---|---|---|---|---|---|
| $P_{out}$ (W) | Max. $V_{out}$ (V) | DC $V_{in}$ (V) | Max. $R_{sat}$ (Ω) | Min. $V_{CR}$(max) (V) | Min. $I_C$(pk) (A) | Min. $P_D$* (W) | Area A ($cm^2$) | Length $l_i$ (cm) | Factor $\eta$ |
| | 250 | 6-10 | 5 | 25 | 0.5 | 0.1 | | | |
| 1 | 500 | 10-15 | 10 | 35 | 0.4 | 0.08 | 0.5-1.5 | 2.5-10 | 0.75 |
| | 750 | 15-20 | 20 | 45 | 0.3 | 0.07 | | | |
| | 250 | 6-12 | 1 | 30 | 3 | 1.5 | 0.5-5 | 2.5-12 | 0.75 |
| 5 | 500 | 12-20 | 2 | 45 | 2 | 1 | | | |
| | 750 | 20-28 | 8 | 60 | 1 | 0.5 | 0.5-5 | 2.5-12 | 0.7 |
| | 300 | 6-12 | 0.8 | 30 | 6 | 3 | | | |
| 10 | 500 | 12-18 | 1 | 45 | 4 | 2 | 1-7.5 | 2.5-15 | 0.7 |
| | 750 | 18-28 | 1.2 | 60 | 2 | 1 | | | |
| | 400 | 10-18 | 0.5 | 45 | 10 | 10 | | | |
| 25 | 600 | 18-26 | 0.8 | 60 | 6 | 5 | 1-10 | 5-15 | 0.65 |
| | 750 | 26-36 | 1 | 80 | 3 | 2 | | | |
| 50 | 500 | 12-24 | 0.5 | 60 | 15 | 20 | 2-15 | 7.5-20 | 0.6 |
| | 750 | 24-36 | 0.5 | 80 | 8 | 7.5 | | | |

*Case or Flange Temperature = 55°C.

diode must be inserted in series with the base.

A suitable transistor for use in the 12-to-150-volt converter is the RCA-2N3053. This transistor has a peak-collector-current rating of 700 milliamperes, a maximum collector-to-base voltage rating of 60 volts, a maximum saturation resistance of 9.5 ohms, and a maximum dissipation rating at 25°C of 5 watts. [The maximum saturation resistance is not given in the published data on the transistor, but is readily determined from the maximum rating for the collector-to-emitter saturation voltage $V_{CE(sat)}$ = 1.4 volts at a collector current $I_C$ of 150 milliamperes.]

Table XXI also provides data on the required cross-sectional area A and length of magnetic path $l_i$ for the transformer core and on the expected circuit efficiency factor $\eta$. The data in Table XXII provide a basis for selection of the core material for a suitable operating frequency. For the specified converter operating condition (i.e., $V_{out}$ = 150 volts, $V_{in}$ = 12 volts, and $P_{out}$ = 1 watt), the data in these tables indicate that a suitable transformer would use a ferrite core that has a cross-sectional area A of 1.3 square centimeters, such as Ferroxcube Part No. 9F520 (E type core, type 3C material) or equivalent. A suitable operating frequency is 8 kHz, and the expected circuit efficiency is 75 per cent.

After a suitable transistor has been selected and the parameters of the magnetic core have been determined, the following step-by-step procedure is used to complete the design of the transformer and

Table XXII—Optimum Core Materials For Different Operating Frequencies.

| Transformer Material | Operating Frequency (kHz) |
|---|---|
| Ferrite | 1-20 |
| Silicon Iron (Grain-Oriented) | 0.1-1 |
| Silicon Steel | 0.1-1 |

to determine the constants for the output rectifier and filter:

1. The transformer secondary-to-primary turns ratio $N_s/N_P$ is determined from Eq. (315) on the basis of the desired dc output voltage $V_{out}$, the available dc input voltage $V_{in}$, and the transistor collector-to-base breakdown voltage rating $V_{CB(max)}$, as follows:

$$\frac{N_S}{N_P} = \frac{V_{out}}{V_{CB(ma)} - V_{in}} = \frac{150}{60 - 12} = 3.1$$

[The use of the 60-volt $V_{CB(max)}$ rating of the RCA-2N3053 rather than the 35-volt value obtained as the minimum $V_{CB(max)}$ rating from Table XXI reduces the required step-up ratio.]

2. The value determined for $N_s/N_P$, and Eq. (305) are used in the following calculation of the required primary inductance:

$$L_P = \frac{(0.7)(12)^2}{(2)(1)(8 \times 10^3)} \left[\frac{(150/3.1)^2}{(12)^2 + (150/3.1)^2 + 2(12) + (150/3.1)}\right] = 4 \text{ mH}$$

3. The required number of primary turns $N_P$ is calculated, for an air-gap length $l_a$ of 0.01 cm, from Eq. (310) as follows:

$$N_p = \sqrt{\frac{4 \times 10^{-3}}{4\pi}} \times \sqrt{\left(\frac{8.1}{1.3 \times 10^3} + \frac{0.01}{1.3}\right) 10^9} = 63 \text{ turns}$$

4. From Eq. (304), the ratio of conduction time $t_{on}$ to the total period of oscillation T is calculated as follows:

$$\frac{t_{on}}{T} = \frac{V_{out}\,(N_p/N_s)}{V_{in} + V_{out}\,(N_p/N_s)} = \frac{150/3.1}{12 + 150/3.1} = 0.78$$

5. From Eq. (295), the peak primary current is calculated to be

$$\hat{I}_p = \left(\frac{2\,P_{out}}{\eta\,V_{in}}\right)\left(\frac{T}{t_{on}}\right) = \frac{2\,(1)}{(0.7)(12)(0.78)} = 0.3 \text{ ampere}$$

6. The maximum flux density in the transformer core is determined on the basis of the relation expressed by Eq. (308) from the following calculation:

$$B_M = \frac{\Delta\phi}{A} = \frac{L_p\,I_p}{N_p\,A} \times 10^8 = \frac{(4 \times 10^{-3})(0.3)}{63(1.3)} = 600 \text{ gauss}$$

This calculation shows that the Ferroxcube 9F520 core is acceptable as the core material because the calculated value of flux density does not exceed the saturation value of this core, which is 4600 gauss. (Table XXIII shows typical values of magnetic parameters for commercial transformer-core materials.)

7. From the published data on the RCA-2N3053 transistor, the typical values of $V_{BE}$ and $I_B$ required for a peak primary current

Table XXIII—Typical Values Of Magnetic Parameters For Commercial Transformer-Core Materials.

| Material | Maximum Permeability ($\mu_m$) | Maximum Flux Density ($B_m$)—Gauss |
|---|---|---|
| Ferrite | 1000-4000 | 2000—5000 |
| Silicon Iron (Grain-Oriented) | 8500 | 10,000—15,000 |
| Silicon Steel ("Hi-Mu" Type) | 30,000 | 15,000—20,000 |
| Nickel-Iron Alloy | 70,000 | 15,000—20,000 |

$I_p$ of 0.3 ampere are, respectively, 1.2 volts and 4.3 milliamperes. The maximum values for these parameters are 2.3 volts (this value includes the drop across the diode which must be used in series with the base, as shown below) and 11 milliamperes, respectively.

If $R_p + R_{supply}$ is assumed to be 2 ohms, which is generally the case, the required number of turns for the base winding is calculated from Eq. (313), as follows:

$$N_B = \frac{2 N_P V_{BE(max)}}{V_{in} - \hat{I}_P (R_{sat} + R_P + R_{supply})}$$

$$= \frac{(63)(2.3)(2)}{10 - (0.3)(12)} = 34 \text{ turns}$$

8. The following calculation is used to determine the base-to-emitter voltage of the transistor:

$$V_{BE} = (N_B/N_p) V_{CB(max)}$$
$$= (34/63)\, 60 = 32.5 \text{ volts}$$

Because this value exceeds the base-to-emitter breakdown-voltage rating of the 2N3053, a diode must be used in series with the base.

9. From Eq. (316), the required value of the external base resistance is calculated to be

$$R_B = V_{BE(max)}/I_B$$
$$= (2.3)/(11 \times 10^{-3}) = 210 \text{ ohms}$$

10. From Eq. (315), the required number of turns for the secondary winding is calculated to be

$$N_S = N_p (V_{out})/[V_{CB(max)} - V_{in}]$$
$$= 63(150)/(60-12) = 200 \text{ turns}$$

11. The calculation of the peak secondary current and the peak secondary voltage, from Eqs. (317) and (318), respectively, yields the following results:

$$\hat{I}_s = \hat{I}_p (N_p/N_s)$$
$$= 300/3.1 = 0.097 \text{ ampere}$$

$$\hat{V}_s = [V_{CB(max)} - V_{in}] (N_s/N_p)$$
$$= (60-12)(3.1) = 150 \text{ volts}$$

12. The values obtained for $\hat{I}_s$ and $V_s$ dictate that the diode selected for use in the rectifier circuit must be capable of handling a peak current of 0.097 ampere and must have a peak-inverse-voltage rating of more than 150 volts. The RCA-1N1763A silicon recti-

fier has a maximum peak-inverse rating of 400 volts and a maximum dc-forward-current rating of 3.5 amperes and, therefore, fulfills these requirements.

13. The value of an output capacitor $C_{out}$ capable of storing the energy required by the load and delivering this energy to the load during the conduction period $t_{on}$ at a substantially constant voltage can be calculated from the following relationship [Eq. (319)]:

$$C_{out} \geq (10\ t_{on}\ P_{out})/V_{out}^2$$

From Step 5, the ratio of "ON" time to total period of oscillation was calculated as $t_{on}/T = 0.78$. The period of oscillation T, however, is the reciprocal of the operating frequency (f = 8 kHz). The value of $t_{on}$, therefore, can be determined as follows:

$$t_{on} = (0.78/8) \times 10^{-3} - 97 \times 10^{-6}$$

From the initial conditions, the output power $P_{out}$ is 1 watt and the output voltage $V_{out}$ is 150 volts. The value of the output capacitor then is calculated to be

$$C_{out} \geq \frac{10\ (97 \times 10^{-6})(1)}{22.5 \times 10^3} = 0.042\ \mu F$$

A RETMA standard value of 0.047 microfarad would be suitable.

15. From the peak currents in the primary, secondary, and base windings (the peak base current $\hat{I}_B$ is the maximum $I_B$ required to produce the necessary value of $\hat{I}_p$), the wire sizes based on the conservative rating of 700 circular mils per ampere are as follows:

Primary = No. 26 for $\hat{I}_P$ of 0.300 ampere
Secondary = No. 32 for $\hat{I}_S$ of 0.097 ampere
Base = No. 36 for $\hat{I}_B$ of 0.011 ampere

16. From Eq. (320), the optimum ratio of primary-to-secondary winding space K is calculated as follows:

$$K = \frac{1}{\sqrt{(N_S/N_P)(V_{in}/V_{out}) + 1}}$$

$$= \frac{1}{\sqrt{(3.1)(12/150) + 1}} = 0.9$$

The best coupling is achieved when the winding order with respect to the core is primary, base, and secondary.

The design described could be improved by use of a transformer core having a cross-sectional area greater than 1.3 square centimeters. Because such a core would permit the use of fewer turns on the primary, base, and secondary windings, and a larger window area or winding space, it would be possible to use wires of larger sizes, and thus to achieve a substantial reduction in copper loss with only a slight increase in core losses because of the larger core.

**Two-Transistor, One-Transformer Converter**—Fig. 346 shows a push-pull, transformer-coupled, dc-to-dc converter that uses one transformer and two transistors. Fig. 347 shows the waveforms obtained from this circuit during one complete operating cycle.

During a complete cycle, the flux density in the transformer core varies between the saturation value in one direction and the saturation value in the opposite direction, as shown by waveform A in Fig. 347. At the start of the

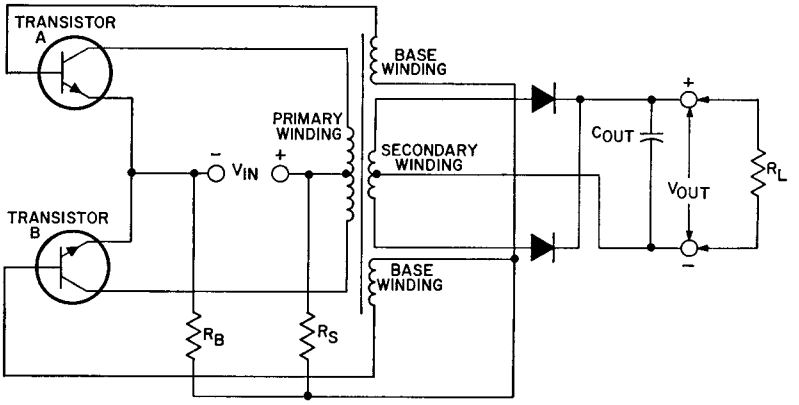



Figure 346. Two-transistor, one-transformer push-pull switching converter.

conduction period for one transistor, the flux density in the core is at either its maximum negative value ($-B_{sat}$) or its maximum positive value ($+B_{sat}$).

For example, transistor A switches "ON" at $-B_{sat}$. During conduction of transistor A, the flux density changes from its initial level of $-B_{sat}$ and becomes positive as energy is simultaneously stored in the inductance of the transformer and supplied to the load by the battery. When the flux density reaches $+B_{sat}$, transistor A is switched off and transistor B is switched on. The transformer assures that energy is supplied to the load at a constant rate during the entire period that transistor A conducts. This energy-transformation cycle is repeated when transistor B conducts.

Initially, sufficient bias is applied to saturate transistor A. As a result, a substantially constant voltage, waveform B in Fig. 347, is impressed across the upper half of the primary winding by the dc source $V_{in}$. This bias voltage can be a temporary bias, a small fixed bias, or even a small forward bias

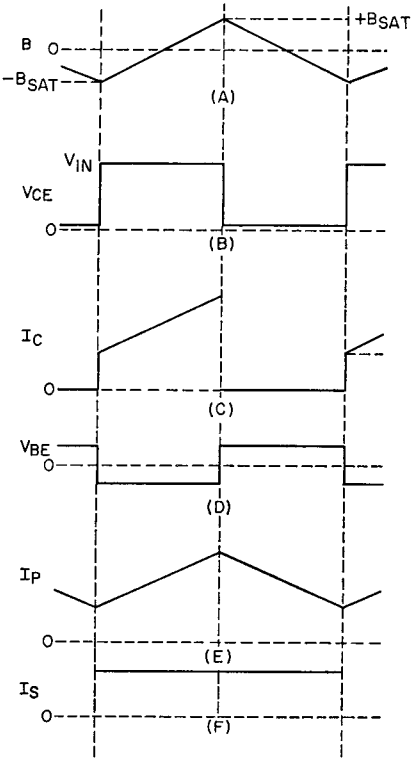



Figure 347. Typical operating waveforms for a two-transistor, one-transformer switching converter: (A) flux density in transformer core; (B) collector voltage of one transistor; (C) collector current of one transistor; (D) base voltage of one transistor; (E) primary current; (F) secondary current.

developed across the bias winding as a result of leakage and saturation current flowing in the transformer primary. The constant primary voltage causes a dc component and a linearly increasing component of current, waveform C in Fig. 347, to flow through transistor A. As in the ringing-choke converter, the linearly increasing primary current induces substantially constant voltages, waveform D in Fig. 347, in the base winding and secondary winding. The induced voltage in the base winding limits the maximum value of the base current and, therefore, of the collector current.

In the push-pull transformer-coupled converter, the transition to switch-off is initiated when the transformer begins to saturate. As long as the transistor is not saturated, the product of the transformer inductance and the time rate of change of the collector current remains constant. When the transformer core saturates, however, the inductance decreases rapidly toward zero, with the result that the time rate of change of the collector current increases towards infinity. When the collector current reaches its maximum value, transistor A moves out of saturation and the winding voltages decrease and then reverse and thereby cause transistor A to switch off. The reversal of the winding voltages switches transistor B on, and the switching operation is repeated.

**Design equations:** If it is assumed that the full dc supply voltage is impressed across one-half of the primary winding, the current flowing in the collector circuit of the conducting transistor may be determined by means of the following equation:

$$L_P\,(dI_P/dt) = V_{in} \qquad (321)$$

where $L_P$ is the inductance of one-half the primary winding in henries, $dI_P/dt$ is the rate of change of primary current in amperes per second, $I_P$ is the magnitude of the change in collector current during one conduction interval, and $V_{in}$ is the dc supply voltage in volts.

For the triangular current waveform, waveform C in Fig. 347, the instantaneous rate of change of current can be approximated as follows:

$$dI_P/dt = 2\hat{I}_P/0.5T = 4I_P/T = 4f\hat{I}_P \qquad (322)$$

where T is the total period of oscillation in seconds and f is the operating frequency in hertz ($f = 1/T$).

Eqs. (321) and (322) are combined and terms are re-arranged to obtain the following expression for the peak value of the current in the collector of the conducting transistor:

$$\hat{I}_P = V_{in}/4f\,L_P \qquad (323)$$

The average value of collector current in the conducting transistor is given by:

$$I_{AV} = P_{out}/\eta\,V_{in} \qquad (324)$$

The maximum primary current then can be expressed by the following equation:

$$I_{P(max)} = \hat{I}_P + I_{AV} = (V_{in}/4f\,L_P) + (P_{out}/\eta\,V_{in}) \qquad (325)$$

The required inductance for one-half of the primary is de-

termined from the following relationship:

$$L_P = (N_P \Delta\phi / \hat{I}_P) \times 10^{-8} \quad (326)$$

In the push-pull transformer-coupled converter, however, the swing on the B-H saturation curve is symmetrical about the origin. The residual flux density $\theta_{res}/A$, therefore, is zero, and the maximum permissible magnetic swing is determined as follows:

$$\frac{\Delta\phi}{A} = \frac{\phi_{max}}{A} - \frac{\phi_{res}}{A} = \frac{\phi_{max}}{A} = B_{max} \quad (327)$$

The required inductance for one-half of the primary may, therefore, be expressed in terms of the maximum flux density $B_{max}$, as shown in the following equation:

$$L_P = (N_P B_{max} A \times 10^{-8} / \hat{I}_P) \quad (328)$$

where $B_{max}$ is expressed in gauss.

If Eq. (323) is substituted into Eq. (328), the following result can be obtained:

$$V_{in} = 4 N_P f B_{max} A \times 10^{-8} \quad (329)$$

Eq. (225) may be rewritten to obtain the following expression for the required number of turns in the primary:

$$N_P = (V_{in} \times 10^8) / 4f B_{max} A \quad (330)$$

Because no air gap is required, the required inductance for one-half the primary is given by

$$L_P = (4\pi N_P^2 / 10^9) (\mu_i A / l_i) \quad (331)$$

To determine the required number of turns for each section of the base winding, it is necessary to know the maximum base-to-emitter voltage $V_{BE}(max)$ at which the transistors provide the peak primary current $\hat{I}_p$. This voltage may be obtained from the published data for the transistor type used or from the transistor manufacturer. The number of turns for each half of the base winding is then expressed as follows:

$$N_B = N_P [2 V_{BE(max)} / V_{in}] \quad (332)$$

The required number of turns for each half of the secondary winding is expressed as follows:

$$N_S = (N_P / V_{in}) (V_{out} + R_{out} I_{out}) \quad (333)$$

where $R_{out}$ is the resistance of the secondary and reflected primary resistance. Because $R_{out}$ is usually very small in transformer-coupled converters, it can be neglected in the initial calculations.

The value of the external base resistance $R_B$ is found in the same manner as for the ringing-choke converter. If extreme flexibility of operation is desired, a separate external resistor may be used in each base circuit.

In the push-pull transformer-coupled circuit, the maximum allowable dc input voltage $V_{in}$ is limited by the collector-to-base-voltage rating for the transistor type used. The maximum permissible supply voltage $V_{in}$ is given by

$$V_{in(max)} \leq [V_{CB(max)} - V_{BE(wind)}]/2 \quad (334)$$

where $V_{CB}(max)$ is the maximum collector-to-base-voltage rating for the transistor type used and $V_{BE}(wind)$ is the induced voltage in one-half of the base winding. Eq. (334) is based on the assumption that the leakage inductance in the transformer is zero. In practice, $V_{in}$ should not be more

than about 90 per cent of the value given in Eq. (334).

The peak secondary current is approximately equal to the dc load current, i.e.,

$$\hat{I}_S = I_{RL} = P_{out}/V_{out} \qquad (335)$$

and the peak secondary voltage is given by

$$V_S = V_{in}\,(N_S/N_P) \qquad (336)$$

For good filtering of the output voltage, the value of $C_{out}$ should be chosen so that the output time constant is at least 10 times the period of the oscillator; i.e.,

$$R_L\,C_{out} = 10T = 10/f$$

or

$$C_{out} = 10/R_L\,f \qquad (337)$$

where the load resistance $R_L$ is determined as follows:

$$R_L = (\hat{V}_S/\hat{I}_S) \qquad (338)$$

The starting resistor $R_s$ is chosen so that a voltage of 0.6 volt appears at the center tap of the feedback winding when the supply voltage is applied, i.e.,

$$R_S = \left(\frac{V_{in} - 0.6}{0.6}\right) R_B \qquad (339)$$

Slightly higher starting voltages may be required for operation at low temperatures.

**Sample design:** The design data shown in Table XXIV can be used as the starting point in the design of a practical single-transformer type of push-pull converter. For specified operating conditions (application requirements), the data in the table provide the criteria used in the selection of the converter transistors, give the parameters of the transformer core, and indicate the expected circuit efficiency. When these design data are used as the starting point, a practical single-transformer push-pull converter can be designed by proper application of the design relationships derived in the preceding section, as shown by the following numerical example.

In the example, it is assumed that the converter circuit to be designed is required to develop a dc output voltage $V_{out}$ of 110 volts and a dc power output $P_{out}$ of 100 watts from a dc input voltage of 13.6 volts. For these application requirements, the data in Table XXIV indicate that the transistor selected for use in the converter should have a maximum saturation resistance $R_{sat}$ less than 0.5 ohm, a collector-to-base breakdown voltage $V_{CB}$(max) greater than 45 volts, a minimum collector-current rating $I_{c(pk)}$ of at least 15 amperes, and a minimum dissipation rating of 15 watts at a temperature of 55°C. The published data on the RCA-40251 transistor indicate that it would be suitable for use in the converter circuit. This transistor has a maximum saturation resistance $R_{sat}$ of 0.2 ohm [i.e., $V_{CE(sat)} = 1.5$ volts at a collector current of 8 amperes], a collector-to-base breakdown voltage $V_{CB}$(max) of 50 volts, a peak collector current rating $I_{C(pk)}$ of 15 amperes, and a dissipation rating of 117 watts at a case temperature of 25°C.

If a suitable operating frequency for the converter is assumed to be 3.5 kHz, Table XXIV indicates a ferrite-core type of transformer should be used. From Table XXIV it is determined that the core should have a cross-sectional area A of 3 square centimeters and a magnetic path length $l_i$ of 25 centimeters. On the basis

Table XXIV—Design Data for Push-Pull, Transformer-Coupled DC-to-DC Converters.

| APPLICATION REQUIREMENTS | | | TRANSISTOR REQUIREMENTS | | | | TRANSFORMER-CORE PARAMETERS | | CIRCUIT EFFICIENCY FACTOR |
|---|---|---|---|---|---|---|---|---|---|
| $P_{out}$ (W) | Max. $V_{out}$ (V) | DC $V_{in}$ (V) | Max. $R_{sat}$ (Ω) | Min. $V_{CB(max)}$ (V) | Min. $I_{C(pk)}$ (A) | Min. $P_D$* (W) | Area A ($cm^2$) | Length $l_t$ (cm) | $\eta$ |
| 2 | 250 | 6-12 | 2 | 30 | 0.5 | 0.1 | 0.5-4 | 2.5-10 | 0.85 |
| | 500 | 12-20 | 4 | 45 | 0.4 | 0.075 | | | |
| 10 | 400 | 12-18 | 1.5 | 45 | 2 | 1 | 0.5-5 | 2.5-10 | 0.85 |
| | 600 | 18-28 | 3 | 60 | 1 | 0.5 | | | |
| 25 | 400 | 12-18 | 1 | 45 | 5 | 3 | 1.5 | 5-15 | 0.85 |
| | 600 | 18-28 | 2 | 60 | 3 | 1.5 | | | |
| | 250 | 8-18 | 0.5 | 45 | 12 | 10 | | | |
| 50 | 500 | 18-28 | 0.8 | 60 | 8 | 5 | 2-7.5 | 7.5-20 | 0.85 |
| | 800 | 28-38 | 1 | 80 | 5 | 2 | | | |
| | 400 | 12-18 | 0.5 | 45 | 18 | 15 | | | |
| 100 | 600 | 18-28 | 0.5 | 60 | 10 | 10 | 3-12 | 10-25 | 0.85 |
| | 800 | 28-38 | 0.5 | 80 | 7.5 | 5 | | | |
| | 400 | 12-24 | 0.2 | 60 | 20 | 25 | | | |
| 200 | 600 | 24-36 | 0.2 | 80 | 15 | 15 | 5-15 | 15-35 | 0.8 |
| | 800 | 36-48 | 0.5 | 100 | 10 | 10 | 7.5-20 | 20-40 | 0.75 |

*Case or Flange Temperature = 55°C.

of these data, a core such as the Allen-Bradley T3000 H106B (or equivalent) should be used. This core is a toroid of Allen-Bradley RO-3 ferrite material (or equivalent); it has a flux density $B_{sat}$ of 4000 gauss and a permeability $\mu_i$ of 3500. (Because the transformer core must be operated in the saturation region, an air gap $l_a$ is not normally required.) The expected circuit efficiency $\eta$, given in Table XXIV, is 85 per cent.

After the transistor type to be used, the magnetic-core parameters, and the expected circuit efficiency have been determined, the following step-by-step procedure results in the design of a practical converter circuit.

1. From Eq. (334), the maximum permissible value of supply voltage $V_{in}$ is calculated as

$$V_{in} \leq 0.9\left[\frac{V_{CB(max)} - V_{BE(max)}}{2}\right]$$

$$= \frac{0.9\,(50-3.5)}{2} = 21 \text{ volts}$$

The desired supply voltage of 13.6 volts is substantially less than this maximum permissible value.

2. From Eq. (330), the required number of turns for each half of the primary is calculated as

$$N_p = (V_{in}\,10^8)/4fAB_{(sat)}$$

$$= \frac{(13.6)(10^8)}{4(3.5)(10^3)(3)(4)(10^3)} = 8 \text{ turns}$$

The total number of primary turns, therefore, is given by

$$2\,N_p = 2(8) = 16 \text{ turns}$$

3. The required inductance for one half of the primary is determined from Eq. (322) as follows:

$$L_p = (4\pi\,N_p^2/10^9)\,(\mu_i A/l_i)$$

$$= \left[\frac{4\pi\,(64)}{10^9}\right]\left[\frac{(3500)(3)}{25}\right]$$

$$= 0.335 \text{ mH}$$

4. From Eq. (325), the maximum primary current is determined as follows:

$$I_p = (V_{in}/4fL_p) + (P_o/\eta V_{in})$$
$$= \frac{13.6}{(4)(3.5)(10^3)(0.335)(10^{-3})}$$
$$+ \frac{100}{(13.6)(485)} = 11.5 \text{ amperes}$$

5. The published data on the RCA-40251 give the maximum values of base-to-emitter voltage $V_{BE}$ and base current $I_B$ as follows:

$$V_{BE(max)} = 2.5 \text{ volts}$$
$$I_{B(max)} = 0.8 \text{ ampere}$$

6. From Eq. (332), the required number of turns for each section of the base winding then may be determined from the following calculation:

$$N_B = N_p \left[\frac{2\, V_{BE(max)}}{V_{in}}\right]$$
$$= \frac{8(2)(2.5)}{13.6} = 3 \text{ turns}$$

7. The required value for the external base resistance $R_B$ is equal to $V_{BE}(max)/I_B$, where $V_{BE}(max)$ and $I_B$ are the values required to obtain a collector current $I_C$ (primary current $I_p$) of 11.5 amperes for the 40251 transistor. The resistance $R_B$, therefore, is calculated as follows:

$$R_B = 2.5/0.8 = 3 \text{ ohms}$$

The 3-turn-per-side drive winding (as calculated in step 6) produces a reverse bias voltage of 5 volts across the "off" transistor [i.e., $E = (N_B/N_p)\, E_{in} = (3/8)(13.6) = 5$ volts]. The 2.4-volt drop across the 3-ohm resistor increases this reverse bias voltage to 7.4 volts, which exceeds the $V_{EBO}$ rating (7.0 volts) of the 40251 transistor. A special selection of this transistor, which has a minimum emitter-to-base breakdown voltage of 8.0 volts, must be used.

8. From Eq. (333), the required number of turns for each half of a center-tapped secondary winding or the total number of turns for an untapped secondary winding is given by

$$N_s = N_p\, (V_{out}/V_{in}) = 8\,(110)/13.6$$
$$= 66 \text{ turns}$$

9. From Eqs. (335) and (336), the peak secondary current $\hat{I}_s$ and the peak secondary voltage $V_s$ are calculated as follows:

$$\hat{I}_s = I_{R_L} = P_{out}/V_{out} = 100/110$$
$$= 0.9 \text{ ampere}$$

$$\hat{V}_s = V_{in}\, (N_s/N_p) = (13.6)(66)/(8)$$
$$= 110 \text{ volts}$$

The load resistance $R_L$ is then calculated to be

$$R_L = 110/0.9 = 125 \text{ ohms}$$

10. On the basis of the values calculated for the peak secondary current and voltage, the RCA-1N1202A diffused-junction silicon rectifier is determined to be suitable for the rectifier diodes. For a single-phase, full-wave rectifier circuit that does not use a center-tapped secondary winding, four of these rectifiers are required.

From Eq. (337), a suitable value for the output capacitor $C_{out}$ is calculated as follows:

$$C_{out} = 10/R_L\, f$$
$$= \frac{10}{(125)(3.5)(10^3)} = 2.3\,\mu F$$

11. The value of the starting resistor $R_s$ is determined from Eq. (338) to be

$$R_s = \frac{(13.6 - 0.6)(3)}{0.6} = 65 \text{ ohms}$$

Fig. 348 shows the circuit schematic and Fig. 349 shows the performance curves for the 13.6-to-110-volt converter.

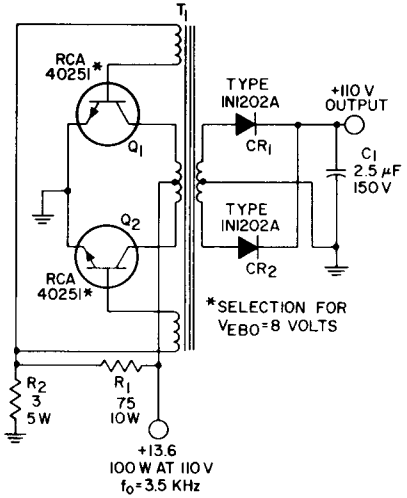



Figure 348. Schematic diagram of 13.6-to-110-volt transformer-coupled push-pull converter.

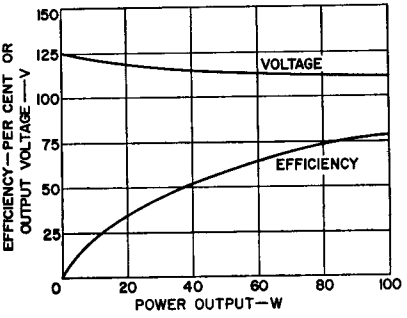



Figure 349. Output voltage and efficiency as a function of power output for the converter circuit shown in Fig. 348.

**Two-Transistor, Two-Transformer Inverters**—There are three basic disadvantages associated with the two-transistor, one-transformer inverter. First, the peak collector current is independent of the load. This current, therefore, depends on the available base voltage, the gain of the transistor, and the input characteristic of the transistor. Second, because of the dependence of the peak current on transistor characteristics, the circuit performance depends on the particular transistor used because there is a wide spread in transistor characteristics. Third, the transformer, which is relatively large, must use expensive square-loop material and must have a high value of flux density at saturation. These disadvantages can be overcome by the use of two transformers in various circuit arrangements, such as that shown in Fig. 350.

In this type of circuit, a saturable base-drive transformer $T_1$ controls the inverter switching operation at base-circuit power levels. The linearly operating output transformer transfers the output power to the load. Because the output transformer $T_2$ is not allowed to saturate, the peak collector current of each transistor is determined principally by the

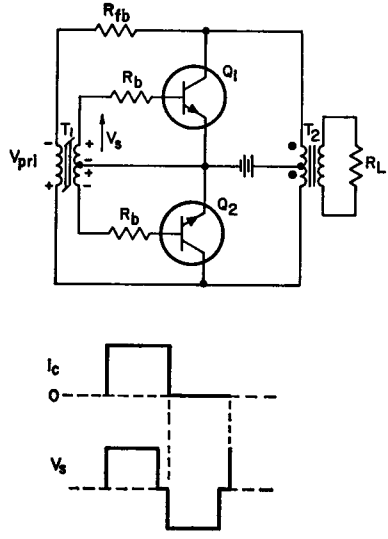



Figure 350. Two-transistor, two-transformer push-pull switching inverter.

value of the load impedance. This feature provides high circuit efficiency. The operation of the inverter circuit is described as follows:

It is assumed that, because of a small unbalance in the circuit, one of the transistors, $Q_1$ for example, initially conducts more heavily than the other. The resulting increase in the voltage across the primary of output transformer $T_2$ is applied to the primary of base-drive transformer $T_1$ in series with the feedback resistor $R_{fb}$. The secondary windings of transformer $T_1$ are arranged so that transistor $Q_1$ is driven to saturation. As transformer $T_1$ saturates, the rapidly increasing primary current causes a greater voltage drop across feedback resistor $R_{fb}$. This increased voltage reduces the voltage applied to the primary of transformer $T_1$; thus, the drive input and ultimately the collector current of transistor $Q_1$ are decreased.

In the circuit arrangement shown in Fig. 350, the base is driven hard compared to the expected peak collector current (forced beta of ten, for example). If the storage time of the transistor used is much longer than one-tenth of the total period of oscillation T, the transistors begin to have an appreciable effect on the frequency of operation. In Fig. 350, the storage time could conceivably be quite long because there is no turn-off bias (the drive voltage only decreases to zero) for $Q_1$ until the collector current of $Q_1$ begins to decrease.

Two methods of overcoming this problem by decreasing the storage time are shown in Fig. 351. In Fig. 351(a), a capacitor is placed in parallel with each base resistor $R_B$. When $V_s$ is positive, the capacitor charges with the polarity shown. When $V_s$ decreases to zero, this capacitor provides turn-off current for the transistor. In Fig. 351(b), a feedback winding from the output transformer is placed in series with each base. The base-to-emitter voltage $V_{BE}$ is then expressed as follows:

$$V_{BE} = V_s - V_{rb} - V_T \qquad (340)$$

If $V_s$ decreases to zero and the collector current does not begin to decrease, then the base-to-emitter voltage is expressed simply by

$$V_{BE} = V_{rb} - V_T \qquad (341)$$

A turn-off bias is thus provided to decrease the collector current.

The energy stored in the output transformer by its magnetizing current is sufficient to assure a smooth changeover from one transistor to the other. The release of this stored energy allows the inverter-circuit switching to be accomplished without any possibility of a "hang-up" in the crossover region during the short period when neither transistor is conducting.

The operation of the high-speed converter is relatively insensitive to small system variations that may cause slight overloading of the circuit. Under such conditions, the base power decreases; however, this loss is so small that it does not noticeably affect circuit performance. At the same time, the amount of energy stored in the output transformer also increases. Although this increase results in a greater transient dis-

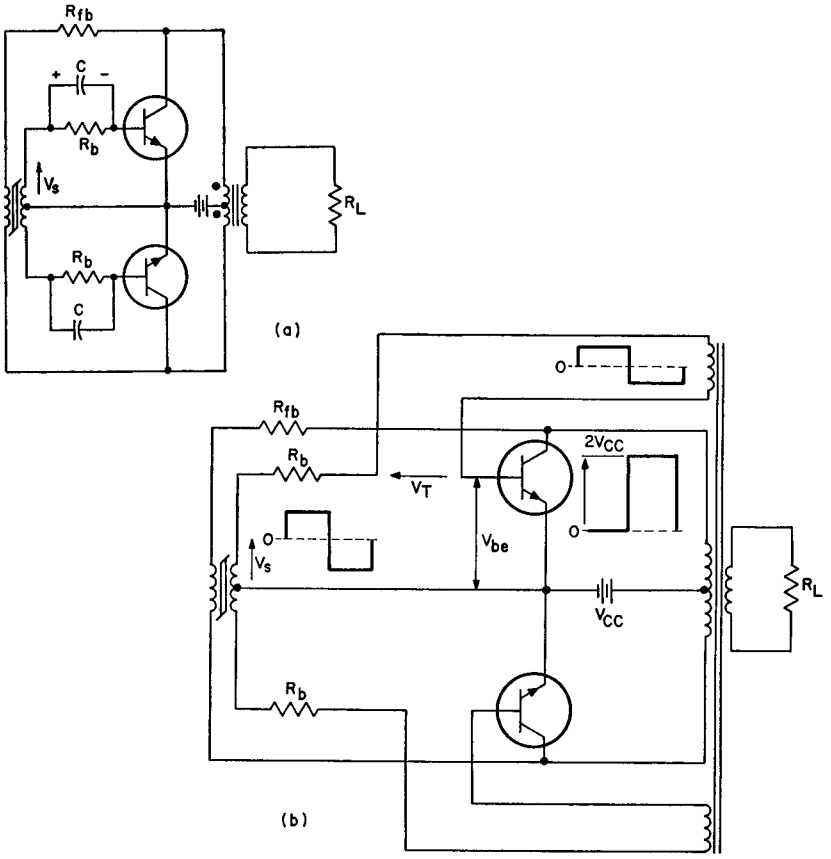



Figure 351. Two-transistor, two-transformer push-pull switching inverters in which transistor-storage times are reduced: (a) Capacitor in parallel with each base resistor assures sharp turn-off of associated transistor; (b) Feedback winding from output transformer in series with base of each transistor assures sharp cutoff characteristics.

sipation, the inverter switching is still effected smoothly.

A practical design of the high-speed converter should include some means of initially biasing the transistors into conduction to assure that the circuit will always start. Such starting circuits, as described later, can be added readily to the converter, and are much more reliable than one which depends on circuit imbalance to shock the converter into oscillation.

**Design Equations and procedure:** The design of a high-speed two-transistor, two-transformer converter is based on the available supply voltage, the required output voltage and power, and the range of ambient temperature over which the converter is required to operate. Moreover, the converter specifications usually provide additional preliminary design information such as size and weight limitations, operating frequency, and

stability requirements for the operating frequency.

The first step in the design of a practical converter is the selection of the transistors to be used in the circuit. After suitable transistors have been chosen on the basis of the pre-established criteria, a value for the maximum case temperature $T_C$ is determined, and the transistor parameters given in the manufacturer's data for this value are then used in the following step-by-step procedure to design the converter circuit:

1. The power input to the output transformer $T_2$ ($P'_{out}$) is computed as follows:

$$P'_{out} = P_{out}/\eta_2 \qquad (342)$$

where $P_{out}$ is the required output power of the inverter circuit and $\eta_2$ is the transformer efficiency (a transformer efficiency of 90 to 95 per cent is usually assumed).

2. An estimate of the transistor collector current for a square wave ($I'_C$) can then be obtained from the ratio of $P'_{out}$ to the supply voltage $V_S$, as follows:

$$I'_C = P'_{out}/V_s \qquad (343)$$

3. From the manufacturer's data, the transistor saturation voltage $V_{CE}(sat)$ that corresponds to the collector current $I'_C$ and case temperature $T_C$ is determined. A second estimate of the transistor collector current is then computed as follows:

$$I_C'' = P'_{out}/[V_s - V_{CE(sat)}] \qquad (344)$$

4. The manufacturer's data on the transistor is then consulted to determine the base-to-emitter voltage $V_{BE}$ required for the collector-to-emitter saturation voltage $V_{CE}(sat)$, as given in step 3, at the collector current $I'_C$ and the case temperature $T_C$. The common-emitter forward-transfer ratio $h_{FE}$ at this collector current and case temperature is also obtained from the manufacturer's data. A forced value of the ratio $h'_{FE}$ that is low enough to insure saturation (usually, $h'_{FE}$ is about one-half of $h_{FE}$) is then used, together with the value determined for $V_{BE}$, to estimate the base-circuit input power $P_{in}$, as follows:

$$P_{in} = V_{BE}\left(\frac{I_C''}{h'_{FE}}\right) + \left(\frac{I_C''}{h_{FE}}\right)^2 R_B \qquad (345)$$

The base stabilizing resistance $R_B$ is small and is usually chosen so that the voltage drop across it is about one-half of $V_{BE}$.

5. The input power to the base-drive transformer $T_1$ can be approximated on the basis of the base-circuit input power $P_{in}$ and the transformer efficiency $\eta_1$, as follows:

$$P'_{in} = P_{in}/\eta_1 \qquad (346)$$

6. The collector current can then be approximated on the basis of the total power developed in the converter circuit:

$$I_C = \frac{P_{out}' + P_{in}'}{V_S - V_{CE(sat)}} \qquad (347)$$

If the collector current given by Eq. (347) is significantly higher than that given by Eq. (344), steps 4, 5, and 6 should be repeated with this higher value of collector current substituted for $I_C''$. The collector current also includes the magnetizing current $I_m$ of the output transformer $T_2$. In steps 2 through 6 it is assumed that $I_m$ is less than 10 per cent of $I_C$.

7. The turns ratio of the output transformer $T_2$ is computed on the basis of the specified load impedance $Z_L$ and the reflected impedance $Z_L'$ determined as follows:

$$Z_L' = (V_s - V_{sat})/I_C \quad (348)$$

Thus, the turns ratio $N_2$ for $T_2$ is determined from the following equation:

$$N_2^2 = Z_L/Z_L' \quad (349)$$

8. The value of the feedback resistor $R_{fb}$ is usually chosen so that approximately one-half of the available voltage is dropped across this resistor. Thus, the primary voltage $V_{pri1}$ is equal to $(V_S - V_{sat})$ and the total primary current $I_{pri1}$ is determined as follows:

$$I_{pri1} + I_{m1} = (P_{in}'/V_{pri1}) + I_{m1} \quad (350)$$

9. The turns ratio for transformer $T_1$ is given by

$$N_1 = (V_{BE} + I_B R_B)/V_{pri1} \quad (351)$$

where the base current $I_B$ is equal to $I_C/h'_{FE}$.

10. The minimum number of turns in the primary winding of the output transformer $T_2$ is given by

$$N_p = (V_{pri} \times 10^8)/4fAB \quad (352)$$

where $V_{pri}$ is the primary voltage in volts, f is the operating frequency in hertz, A is the area of the transformer core in square centimeters, and B is the flux density in gauss. Eq. (352) is also used to determine the number of turns required in the primary winding of the base-drive transformer $T_1$ to produce the proper frequency.

11. The magnetizing current in the primary of $T_1$ is determined from the following equation:

$$I_{m1} = (H_s l_i)/(1.26 N_p) \quad (353)$$

where $N_p$ is the number of turns in the primary winding, $l_i$ is the magnetic-path length in centimeters, and $H_s$ is the value of the magnetizing field strength in oersteds at the value of B used in Eq. (352).

This value of $I_{m1}$ must be added to $I_{pri1}$ when the value of $R_{fb}$ is determined (step 8). Eq. (353) is also used in the design of the output transformer $T_2$ to assure that $I_{m2}$ is small compared to $I_C$. If $I_{m2}$ is not small enough, the minimum number of turns as given by Eq. (352) should be increased.

**Special transistor requirements:** The type of transistor selected for use in a high-speed converter circuit is dictated by the following conditions:

1. In a high-speed converter, the peak value of the collector-to-emitter voltage of each transistor is equal to twice the supply voltage plus the amplitude of the voltage spikes generated by transient elements. Therefore, the collector-to-emitter breakdown voltage $V_{CEO}$ of the transistors should be slightly greater than twice the supply voltage (usually an additional 20 per cent is sufficient).

2. The transistors must be capable of handling the currents necessary to produce the required output power at the given supply voltage, and their saturation voltage at these currents must be low enough so that the high efficiency desired can be obtained.

3. The junction-to-case thermal resistance of the transistors $\theta_{J\text{-}C}$ must be low enough so that the manufacturer's maximum ratings, for the given ambient temperature

and the available heat sink and cooling apparatus, are not exceeded.

The maximum collector current, the dissipation, and the heat-sink thermal resistance of the transistors can be approximated on the basis of these limiting conditions as follows:

The maximum collector current $I_C$ is approximately given by

$$I_C = P_{out}\,\eta/[V_S - V_{CE(sat)}] \quad (354)$$

where $V_S$ is the supply voltage, $V_{CE(sat)}$ is the transistor collector-to-emitter saturation voltage (for a specific $I_C$), $P_{out}$ is the required power output, and $\eta$ is the desired efficiency of the output transformer (usually 90 to 95 per cent).

The transistor dissipation can be approximated as follows (because the base dissipation is very small, it is neglected in this approximation):

$$P_D = (T_1/T)\,(V_{CE(sat)}\,I_C + 2I_{CEX}\,V_S) + [(t_{on}+t_f)/T]\,(V_S I_C/3) \quad (355)$$

where $V_S$ is the supply voltage, $V_{CE(sat)}$ is the transistor saturation voltage (for a specific $I_C$); $I_C$ is the collector current, as given by Eq. (354); $I_{CEX}$ is the collector current with the base reverse-biased (for $V_{CE} = 2V_s$); $t_{on}$ is the transistor "turn-on" time [at $I_C$ given by Eq. (354) and $h'_{FE}$ given in step 4 of the general procedure]; $t_f$ is the transistor "fall" time (at $I_C$ given by Eq. (354) and $h'_{FE}$ given in the general procedure); T is the period reciprocal of the operating frequency; and $T_1 = \frac{1}{2}\,[T-(t_{on} + t_f)]$.

Eq. (355) is used as a guide for the first stages of design; the exact dissipation is determined experimentally. The transistor saturated-switching characteristics must be fast enough to prevent the transient dissipation from becoming excessive.

The required heat-sink thermal resistance may be approximated by the following equation:

$$\Theta_{C-A} = (\Delta T/P_D) - \Theta_{J-C} \quad (356)$$

where $\Delta T$ is the permissible junction temperature rise ($\Delta T = T_{J(max)} - T_A$); $P_D$ is the transistor dissipation; and $\theta_{C-A}$ is the case-to-air thermal resistance, including mounting, interface, any insulation material, and heat sink.

The estimate of the required heat-sink thermal resistance, together with the manufacturer's maximum rating curve or safe operating region, completes the determination of transistor requirements.

**Second-breakdown considerations:** A high-speed, high-power inverter requires transistors that have high power-handling capabilities and very fast saturated-switching speeds. Reverse-bias second breakdown (which is discussed in an earlier section of this Handbook) is a factor that must also be considered in the design of these circuits.

Reverse-bias second breakdown can be analyzed as follows: During the turn-off time $t_{off}$, the transistor is subjected to high energy as a result of energy stored in the output-transformer leakage inductance. This leakage inductance can be made small by careful winding of the transformer to obtain close coupling. An approximation of the value of leakage inductance can be obtained by measuring the inductance of one-half the primary with the other half of the primary short-circuited. As is shown in the sample design at

the end of this section, the leakage-inductance value and the peak collector current can be used to provide an analysis of reverse-bias second breakdown.

**Feedback resistance:** The value of feedback resistance $R_{fb}$ is computed as the resistance required to produce the difference in voltage that should exist between the collector-to-collector voltage of the two transistors and the voltage applied to the primary of transformer $T_1$ at a given primary current $I_{pri1}$. The optimum value of the feedback resistor is then determined experimentally. A decrease in the value of $R_{fb}$ increases the loss that results from the circuit resistance and that in the transformer core because the magnetizing current increases. The voltage across the primary of the transformer then increases and, as may be inferred from Eq. (352), the operating frequency increases. An increase in the value of $R_{fb}$ causes a greater voltage drop across this resistance, and less voltage is then available to the primary of transformer $T_1$; therefore, the frequency decreases. Thus, $R_{fb}$ can be used to control frequency over a limited range only.

**Starting circuits:** The circuits shown in Fig. 350 and 351 will not necessarily begin to oscillate, especially under a heavy load. As a result, a starting bias must be applied so that the circuit has a loop gain greater than unity and is always capable of initiating oscillation. This bias arrangement can be such that it is connected only during starting, or can be connected permanently within the circuit. Two practical starting circuits are described in the following paragraphs.

Fig. 352 shows an inverter that uses a resistive voltage-divider network to supply the necessary starting bias. The value of resistor $R_1$ can be determined by use of Eq. (339). With this circuit,

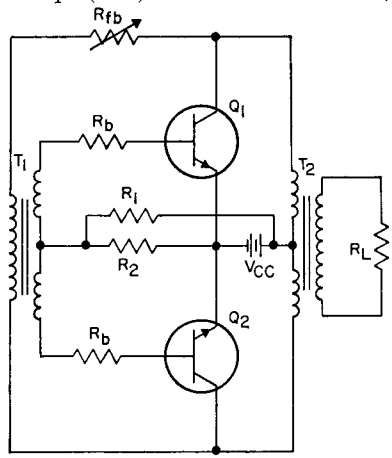



Figure 352. Two-transistor, two-transformer push-pull inverter that uses a resistive voltage-divider network to provide starting bias.

a compromise of reliable starting and tolerable bleeder current must be reached.

Fig. 353 shows a diode starting circuit in which the bases of the

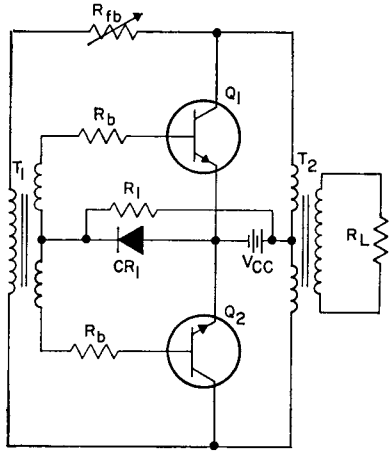



Figure 353. Two transistor, two-transformer push-pull inverter that uses a diode starting circuit.

two inverter transistors are supplied by a resistance $R_1$, which is determined as follows:

$$R_1 = V_{CC}/2I_B \qquad (357)$$

As the inverter begins to oscillate, the base current is conducted through the base-emitter diode and through the forward direction of the starting diode. Usually, additional drive is needed to compensate for the diode voltage drop. Low-voltage silicon diodes capable of carrying the base current continuously are normally used.

**Sample design:** The following paragraphs explain the use of the design procedure given in the preceding section to design a practical high-speed, two-transformer, push-pull converter. The operating requirements upon which the design of the converter is based are as follows:

dc power output $P_{out} = 250$ watts
dc supply voltage $V_s = +28$ volts
operating frequency $f = 50$ kHz
load resistance $R_L = 25$ ohms
ambient temperature $T_A = 25°C$

The design of the converter circuit is performed in four basic parts, as follows:

1. Selection of Transistors. The first step in the selection of the transistors for the high-speed converter is to compute the power input to the output transformer, $P_{out}'$; a transformer efficiency of 95 per cent is assumed. Thus, from Eq. (342),

$$P'_{out} = 250/0.95 = 262.5 \text{ watts}$$

Eq. (343) is then used to make the initial estimate of the transistor collector current necessary to produce the required power output:

$$I_C' = 262.5/28 = 9.4 \text{ amperes}$$

The transistors used in the inverter circuit must have a collector-to-emitter breakdown voltage $V_{CEO}$ equal to at least twice the supply voltage plus an additional 20 per cent to allow for voltage spikes. The value of $V_{CEO}$ is thus given by

$$V_{CEO} \geq 2\,(29)(1.20) = 67 \text{ volts}$$

The RCA-2N3265 power transistors selected for the converter circuit have a $V_{CEO}(sus)$ of 90 volts, and a collector-to-emitter saturation voltage $V_{CE}(sat)$ of 0.75 volt (given in the manufacturer's data for a collector current $I_C$ of 15 amperes), which is low enough to insure that the desired high operating efficiency can be obtained. The switching times for the 2N3265 transistor are as follows:

Fall time $t_f = 500$ nanoseconds
(at $I_C = 15$ amperes)

On time $t_{on} = 500$ nanoseconds
(at $I_C = 15$ amperes)

These switching times are short in comparison to the 20-microsecond period at the 50-kHz operating frequency.

It is now possible to recompute the transistor collector current to obtain a more accurate approximation of the maximum value of this parameter in the converter circuit. Eq. (344) is used to obtain the following result:

$$I_C'' = 262.5/(28-0.75) = 9.62 \text{ amperes}$$

The data given for the 2N3265 transistor are used to determine

the $h_{FE}$ ratio and the base-to-emitter voltage $V_{BE}$ of the transistor at this level of collector current. The $h_{FE}$ ratio is found to be 60 (5th percentile) at a collector current of 10 amperes, which is close enough to the value calculated for $I_C''$. The forced value for this ratio, $h'_{FE}$, is chosen to be 20, which is small enough to assure that the transistor will saturate. The base-to-emitter saturation voltage $V_{BE}(sat)$ at the collector current of 10 amperes is found to be 1.3 volts. The values for the base current and base input resistance can then be computed as follows:

$$I_B = I_C''/h_{FE}' = 9.62/20$$
$$= 0.481 \text{ ampere}$$
$$R_{in} = V_{BE(sat)}/I_B = 1.3/0.481$$
$$= 2.7 \text{ ohms}$$

The total base-circuit input resistance, $R'_{in}$, is the sum of the quantity $R_{in}$ and the transistor bias resistor $R_B$. The value of $R_B$ is chosen to be 1 ohm. Thus, $R'_{in}$ is equal to 3.7 ohms. The base-circuit input voltage $V'_{in}$ can be readily calculated either as the product of $R'_{in}$ and $I_B$ or as follows:

$$V_{in}' = V_{BE(sat)} + I_B R_B = 1.3 + 0.481$$
$$= 1.781 \text{ volts}$$

In the design of a high-speed inverter circuit, the value of the feedback resistor is usually chosen so that the available voltage is divided equally across this resistor and the primary of the base-drive transformer. The voltage across the primary $V_{pri}$ is determined, therefore, as follows:

$$V_{pri} = (0.5)(2)\,[V_S - V_{CE(sat)}]$$
$$= (0.5)(2)(28 - 0.75)$$
$$= 27.25 \text{ volts}$$

The base-circuit input power $P_{in}$ is determined from Eq. (241) or from the product of $V'_{in}$ and $I_B$, as follows:

$$P_{in} = (1.781)(0.581) = 0.86 \text{ watt}$$

If a transformer efficiency of 95 per cent is assumed, the power input to the base-drive transformer is given by

$$P_{in}' = 0.86/0.95 = 0.902 \text{ watt}$$

The primary current is then determined as follows:

$$I_{pri} = 0.902/27.25 = 0.0332 \text{ ampere}$$

The value of the bias resistor $R_1$ (a resistive voltage-divider starting circuit is used) required to produce a starting current of 0.481 ampere is determined as follows:

$$R_1 = \frac{V_{CC} - 0.6}{0.6}$$
$$= \frac{28 - 0.6}{0.6} = 45.7 \text{ ohms}$$

2. Output Transformer Calculations. It is then possible to calculate the transistor collector current on the basis of total power in the inverter circuit, $P'_{out} + P'_{in}$. The value obtained is given by

$$I_C = (262. + 0.902)/27.25$$
$$= 9.65 \text{ amperes}$$

The impedance reflected into the primary of the output transformer $R'_L$, is computed on the basis of this value of collector current as follows:

$$R_L' = 27.25/9.65 = 2.84 \text{ ohms}$$

The ratio of the specified circuit load impedance ($R_L = 25$ ohms) and this reflected impedance defines the transformer turns ratio $N_2$ as follows:

$$N_2^2 = R_L/R_L' = 25/2.84 = 8.85$$
$$N_2 = 2.98$$

On the basis of a transformer efficiency of 95 per cent, the power magnetically dissipated in the output transformer is given by

$$P_M = P_{out}\,(1.00-.95) = 12.5 \text{ watts}$$

For an operating frequency f of 50 kHz, the Allen-Bradley type WO-3 ferrite core material, or equivalent, is acceptable. From the manufacturer's data sheet for this ferrite, the maximum usable core temperature is 125°C. For linear operation at this temperature, the flux density $B_M$ should be 1000 gauss.

The core loss factor $\rho$ as determined from the manufacturer's specification data for $B_M = 1000$ gauss and $f = 50$ kHz is given by

$$\rho = 3.2\mu \text{ W/cm}^3 \text{ Hz}$$

Thus, at 50 kHz the frequency-dependent core loss, $\rho'$, is calculated as follows:

$$\rho' = (3.2\mu \text{ W/cm}^3 \text{ Hz})\,(50\times10^3 \text{ Hz})$$
$$= 0.160 \text{ W/cm}^3$$

The maximum permissible volume of the core for a transformer efficiency of 95 per cent, therefore, is given by

$$\text{Vol} = 12.5 \text{ W}/(0.16 \text{ W/cm}^3) = 78 \text{ cm}^3$$

For a pair of "C" cores, Allen-Bradley Type No. U2625C133A, or equivalent, which have a cross-sectional area A of 2.04 square centimeters and a length $l_i$ of 16.4 centimeters, the volume is only 40 cubic centimeters. As a result, the core loss is less than 7 watts instead of 12.5 watts, and the transformer efficiency is greater than the assumed value of 95 per cent.

The manufacturer's specifications do not include information for estimation of the temperature rise in the core. If the transformer overheats, a new one which uses a core that has a lower loss factor must be designed.

When the two C cores mentioned above are used, the number of turns in the transformer primary can be calculated by use of Eq. (352) as follows:

$$N_p = \frac{27.25\times10^8}{(4)(5)(10^4)(2.04\times10^3)}$$
$$= 6.55 \text{ turns}$$

If $N_p = 6$ turns, then $N_s = 6(2.98) = 18$ turns.

From the manufacturer's data sheet, it is determined that for linear operation the value of $H = 0.189$ oersted results in a magnetizing current $I_m$ given by

$$I_m = (16.4)(0.189)/(1.26)(6.55)$$
$$= 0.376 \text{ ampere}$$

This value is less than 10 per cent of $I_C$.

The transformer wire size should be large enough to prevent excess copper losses, and the primary should be bifilar wound. The transformer should be constructed with a minimum amount of tape applied to the core to reduce the core temperature rise.

3. Base-Drive Transformer Calculations. The Allen-Bradley type RO-3 rectangular-loop ferrite core material, or equivalent, is suitable for use in the base-drive transformer. The flux density $B_m$ of the drive transformer should be 3000 gauss and the saturation field strength $H_s$ should be 1 oersted.

The core-loss factor for a flux density of 3000 gauss and an operating frequency of 50 kHz is given

by the manufacturer's data as

$$\rho = 63\ \mu W/cm^3\ Hz$$

The core loss at 50 kHz is then calculated as follows:

$$\rho' = 63 \times 50 \times 10^3 = 3.15\ W/cm^3$$

If a transformer efficiency of 95 per cent is assumed, the magnetically dissipated power in the drive transformer is given by

$$P_m = P_{in}\ (1.00-0.95) = (0.86)(0.05) = 0.043\ watt$$

The maximum volume is then calculated as follows:

$$Vol = P_m/\rho' = 0.043/3.15 = 0.0136\ cm^3$$

An Allen-Bradley Type No. TO620H101A core, or equivalent, is chosen. This core has an area A of 0.119 square centimeter, a length of 3.9 centimeters, and a volume of 0.465 cubic centimeter. This volume is about thirty times the maximum allowable volume for 95-per-cent efficiency. The magnetic losses, therefore, are about 1.3 watts, and the transformer efficiency is low.

The number of turns in the primary is determined from Eq. (352) as follows:

$$N_p = \frac{(27.25)(10^8)}{(4)(5\times10^4)(0.119)(3000)} = 37.5\ turns$$

The turns ratio $N_1$ for the base-drive transformer $T_1$ is determined as follows:

$$R_{in}' = 3.7\ ohms$$
$$R_{pri} = 27.25/0.033 = 830\ ohms$$
$$N_1^2 = 830/3.7 = 224$$
$$N_1 = 15$$

Therefore, the number of secondary turns is given by

$$N_s = 37.5/15 = 2.5$$

The magnetizing current is determined from Eq. (353) as follows:

$$I_{m1} = (3.9\times1)/(1.26\times37) = 0.084\ ampere$$

The total primary current is then

$$I_{pri} = 0.033 + 0.084 = 0.117\ ampere$$

The feedback resistance $R_{fb}$ is then calculated as follows:

$$R_{fb} = 27.25/0.117 = 235\ ohms$$

4. Thermal-Resistance Calculations. From Eq. (355), the average transistor dissipation is given by

$$P_D = [(20-1)/(2\times20)][(0.75\times9.65) + 2\ (0.020\times28)] + (1/20)\ [(28\times9.65)/3] = 8.65\ watts$$

where $I_{CEX}$, as taken from the manufacturer's data, is equal to 1 to 20 milliamperes.

For a junction temperature of 125°C, the maximum temperature rise is given by

$$T = 125°C - 25°C = 100°C$$

The total junction-to-air thermal resistance, including heat sink, mounting, and junction-to-case thermal resistance, is determined as follows:

$$\Theta_{J-A} = 100/8.65 = 11.6°C/W$$

For the 2N3265, the junction-to-case thermal resistance $\theta_{J-C}$ is given as 1°C/W. The mounting thermal resistance is about 0.25° C/W. Thus, the heat-sink-to-air thermal resistance $\theta_{HS-A}$ is 11.6 − 1.25 = 10.35°C/W.

**Experimental results:** The leakage inductance of the output transformer, as measured on a Q meter, is about 0.5 microhenry. The peak collector current is calculated to be about 10 amperes, and the reverse base-to-emitter bias voltage is about −2 volts. The 2N3265 transistor has an assured capability to withstand second breakdown at currents in excess of 10 amperes for a collector inductance of 90 microhenries and a reverse bias of 6 volts. The published data on the 2N3265 indicate that a reduction in bias voltage or in collector inductance allows the transistor to handle larger amounts of reverse-bias energy. The operating conditions for the output transformer are well within the safe area. Both transformers should be constructed with a minimum of tape to provide as much surface area as possible to ensure a low core temperature.

Fig. 354 shows the schematic diagram for the completed circuit. The values for the feedback resistance and the bias-starting resistance were arrived at experimentally with the calculated values used as a beginning.

Fig. 355 shows the output characteristics of the converter as a function of the load. The output characteristics were measured at the load at the output terminals of the rectifier bridge. Thus, the efficiency shown represents the total circuit efficiency. The range of values indicated on the efficiency curve (i.e., 82 to 88 per cent) takes into account the transistor dissipation, transformer losses, rectifier-bridge losses, and all other circuit IR losses.

Fig. 356 shows the experimental transistor load line for a load resistance of 25.6 ohms and a supply voltage of 28 volts. The area enclosed by the load line shows that high dissipation occurs during switching. This area is decreased somewhat when loads having a small capacitive reactance are used.

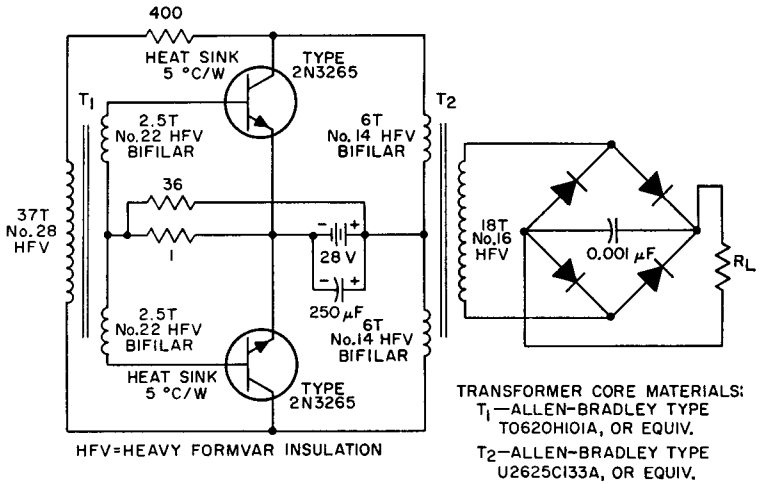



Figure 354. Schematic diagram of 250-watt, 50-kHz push-pull dc-to-dc converter.

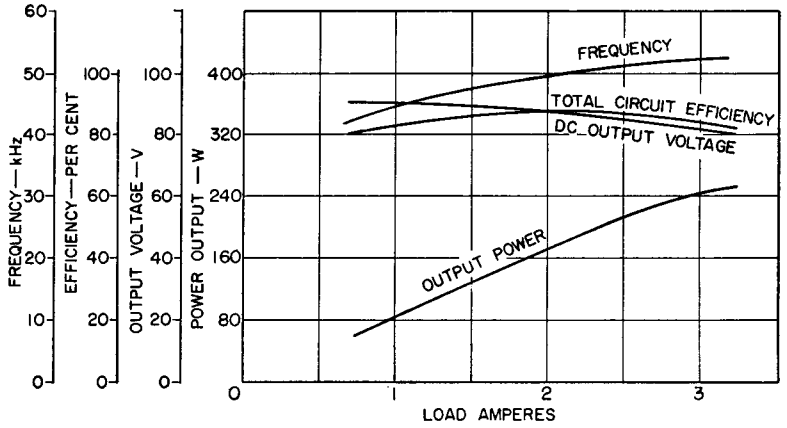



Figure 355. Output characteristic (i.e., frequency, efficiency, voltage, and power) of the 250-watt converter as a function of the load.

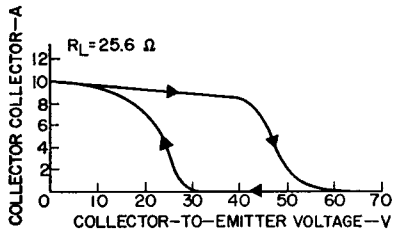



Figure 356. Experimental load line for the 2N3265 transistor using a load impedance of 25.6 ohms and a supply voltage of 28 volts.

Fig. 357 shows the collector current and voltage waveforms. The collector-current waveform exhibits the transformer saturation current. The collector-voltage waveform exhibits the voltage spikes resulting from the transformer leakage inductance. Fig. 358 shows the collector current on an expanded time scale to illustrate the current rise and fall times.

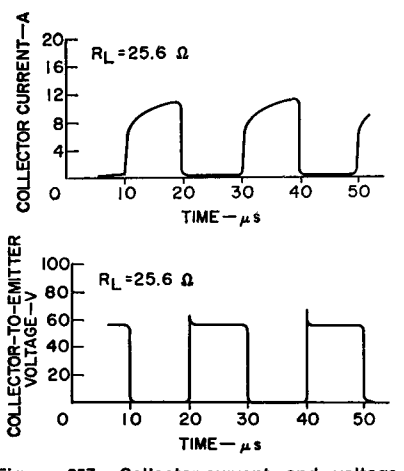



Figure 357. Collector-current and voltage waveforms for the 2N3265 transistors used in the 250-watt converter.

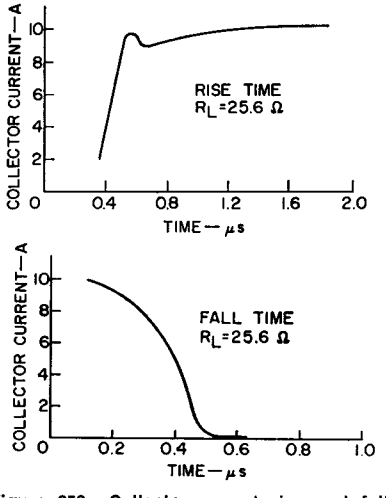
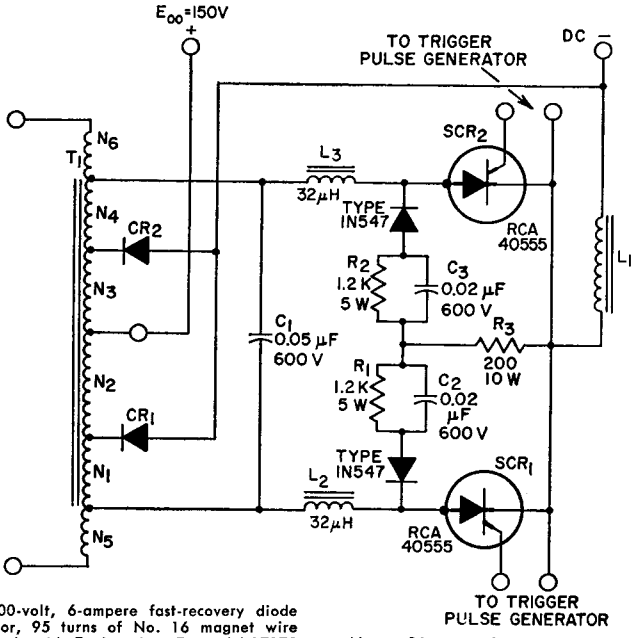



Figure 358. Collector-current rise and fall times.

## SCR INVERTER

Fig. 359 shows a typical high-frequency SCR switching inverter; Fig. 360 shows the waveshapes across the SCR and the output of the transformer. For resistive loads, this inverter is capable of delivering 500 watts of output power at an operating frequency of 8 kHz, and is provided with regulation from a no-load condition to full load. With proper output derating, this circuit can also accommodate inductive and capacitive loads. Under a capacitive load the power dissipation of the SCR's is increased; under an inductive load the turn-off time is decreased.

The inverter can be operated at any optional frequency up to 10 kHz provided that a suitable output transformer is used and the timing capacitors are changed in




$D_1$, $D_2$ = 600-volt, 6-ampere fast-recovery diode
$L_1$ = Inductor, 95 turns of No. 16 magnet wire wound on Arnold Engineering Type A4-17172 (or equiv.) core
$T_1$ = Output transformer: $N_1$ = $N_4$ = 9 turns of No. 18 magnet wire, two strands; $N_2$ = $N_3$ = 36 turns of No. 18 magnet wire; $N_5$ = $N_6$ = 21 turns of No. 18 magnet wire; core, two sets of Siemens Type B66251-A0000-R026 (or equiv.) with 4-mil air gap

Figure 359. High-frequency (10-kHz) SCR push-pull switching inverter.

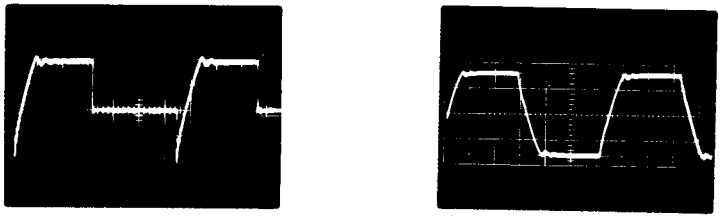


Figure 360. Typical operating waveforms for SCR inverter shown in Fig. 359.

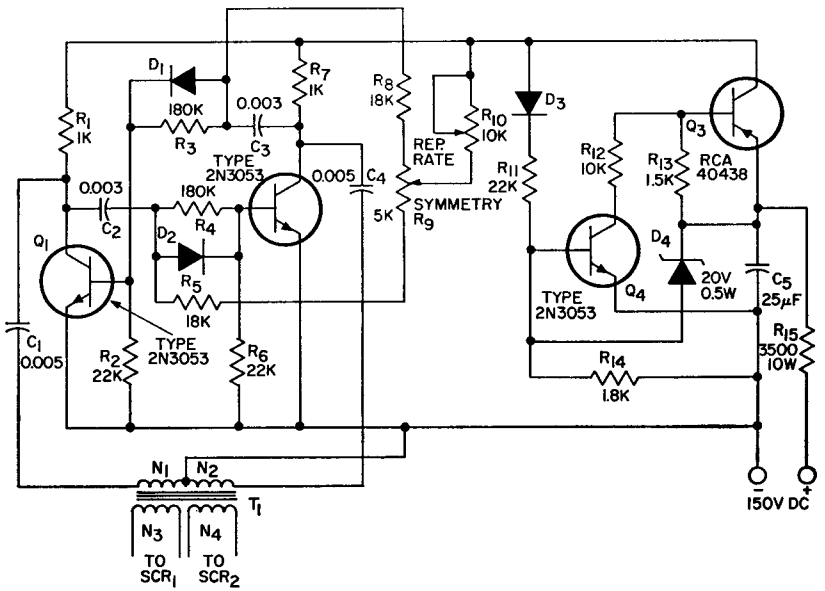
the gate-trigger-pulse generator. A change in operating frequency, however, does not require any change in the commutating components $C_1$ and $L_1$. The operation of the SCR inverter is very similar to that of the two-transistor push-pull inverter except that external gate-trigger signals are required to initiate the SCR switching action.

## Circuit Operation

Fig. 359 shows the two thyristors $SCR_1$ and $SCR_2$ connected to the output transformer $T_1$. These thyristors are alternately triggered into conduction by the gate-trigger-pulse generator shown in Fig. 361 to produce an alternating current in the primary of the power transformer.

The thyristors are commutated by capacitor $C_1$, which is connected between the anodes of $SCR_1$ and $SCR_2$. The flow of current through the circuit can be traced more easily if it is assumed that initially $SCR_1$ is conducting and $SCR_2$ is cut off and that the common cathode connection of the SCR's is the reference point. For this condition, the voltage at the anode of $SCR_2$ is twice the voltage of the dc power supply, i.e., $2\,E_{oo}$. The load current flows from the dc power supply through one-half the primary winding of transformer $T_1$, inductor $L_2$, $SCR_1$, and inductor $L_1$. When the firing current is applied to the gate of $SCR_2$, this SCR turns on and conducts.

During the on period of $SCR_2$, the capacitor $C_1$ begins to discharge through $L_3$, $SCR_2$, $SCR_1$,




$D_4$ = Zener diode, 20-volt, ½-watt
$T_1$ = Pulse transformer; center-tapped primary: $N_1 = N_2 = 150$ turns of No. 36 wire; split secondaries: $N_3 = N_4 = 100$ turns of No. 36 wire; core material: Indiana General Type No. CF902-05, or equiv.

Figure 361. Gate-trigger pulse generator for SCR inverter shown in Fig. 359.

and $L_2$. Inductors $L_2$ and $L_3$ function to limit the rate of rise of the discharge current di/dt so that the associated stresses are maintained within the capability of the device during the turn-on of the SCR. The effect of this control is to decrease the turn-on dissipation, which becomes a significant portion of the total device dissipation at high repetition rates.

The discharge current through $SCR_1$ flows in a reverse direction, and after the carriers are swept out (and recombined) the $SCR_1$ switch opens (i.e., $SCR_1$ switches to the off state). At this time, the voltage across the capacitor $C_1$, which is approximately equal to $-2\ E_{co}$, appears across $SCR_1$ as reverse voltage. This voltage remains long enough to allow the device to recover for forward blocking. Simultaneously during this interval, the conducting $SCR_2$ establishes another discharge path for capacitor $C_1$ through transformer $T_1$ and inductors $L_1$ and $L_3$. The role of inductor $L_1$ is to control the rate of discharge of the capacitor to allow sufficient time for turn-off.

After capacitor $C_1$ is discharged from $-2\ E_{co}$ to zero, it starts to charge in the opposite direction to $+2\ E_{co}$. When $C_1$ is charged to $+2\ E_{co}$, because of the phase shift between voltage and current the flux at that time in the inductor $L_1$ is a maximum. This reactive energy stored in the inductor is normally transferred to the capacitor and causes an "overvoltage" or "overcharge", which in this particular case is undesirable. Voltages on the capacitor higher than $2\ E_{co}$ produce a negative voltage at the anode of $SCR_2$ with respect to the negative terminal of the dc power supply. This condition is prevented by use of a clamping diode $CR_2$ connected to an extra tap on the transformer oriented close to the anode of $SCR_2$. As a result, the amount of "overcharge" of the capacitor is considerably reduced. The energy stored in inductor $L_1$ causes current to flow through diode $CR_2$, the $N_4$ transformer winding inductor $L_3$, and $SCR_2$. Transformer windings $N_4$ and $N_3$ act as an autotransformer through which the energy stored in the inductor is fed back to the power supply.

When the firing current is applied to the gate of $SCR_1$, this device conducts and the process described above is repeated.

Each time the SCR's turn off to interrupt the reverse recovery current, a certain amount of energy remains in the inductor. This energy is transferred to the device capacitance, which is relatively small, and thus a high-voltage transient is generated. This high-voltage transient may exceed the rating of the device, produce undesirable stresses, and increase the switching dissipation. A transient-suppressor network consisting of two 1N547 diodes, resistors $R_1$, $R_2$, and $R_3$, and capacitors $C_2$ and $C_3$ prevents this transient voltage from exceeding the maximum rating of the SCR's.

### Gate-Trigger-Pulse Generator

The gate-trigger-pulse generator, as shown in Fig. 361, is a conventional astable (free-running) multivibrator, combined with a threshold-sensitive switch consisting of transistors $Q_3$ and $Q_4$ which turns the generator on and off. The square-wave output of the generator is differentiated and fed to the gates of $SCR_1$ and $SCR_2$ through the $N_3$ and $N_4$ windings of pulse transformer $T_1$. The threshold-sensitive switch holds the generator off until the re-

quired dc level is achieved in the power supply. This minimum level is necessary to maintain a nominal repetition rate and to supply sufficient current to trigger both SCR's. As dc power is applied through resistor $R_{15}$ to charge capacitor $C_5$, the gradually increasing voltage at the emitter of transistor $Q_3$ eventually rises to a value above the zener voltage of the zener diode $D_4$ connected between the emitter of transistor $Q_3$ and the base of transistor $Q_4$. So long as this voltage is not exceeded, the base current of transistor $Q_4$ is zero. Because transistor $Q_4$ is cut off, transistor $Q_3$ also remains cut off. As the voltage of the power supply increases and exceeds the zener voltage of $D_4$, the zener diode conducts current to the base of transistor $Q_4$ and causes the transistor to conduct. The collector current of $Q_4$ then flows into the base of $Q_3$ and causes this transistor to conduct. The collector current of $Q_3$ is then applied to the astable multivibrator. A polarity-sensitive positive feedback loop consisting of diode $D_3$ and resistor $R_{11}$ provides regenerative feedback to transistors $Q_4$ and $Q_3$ when the zener diode $D_4$ is conducting. In the event that the power-supply voltage decreases and current ceases to flow through the zener diode, this feedback network maintains transistor $Q_3$ in saturation until the voltage in the circuit drops to a few volts.

The collector current through transistors $Q_1$ and $Q_2$ does not maintain perfect balance as the base currents of transistors $Q_1$ and $Q_2$ increase. Any slight unbalance in collector current is amplified through the positive feedback loops. As a result, one transistor is cut off and the other is turned on at the extreme limit of unbalance. If transistor $Q_1$ is assumed turned on, the base of transistor $Q_2$ is driven negative by capacitor $C_2$, which is connected to the collector of $Q_1$. The negative bias on the base of $Q_2$ drives the transistor into the cutoff state. Capacitor $C_3$ connected to the base of $Q_1$ is then charged through the load resistor $R_7$ of transistor $Q_2$, and the base drive on transistor $Q_1$ increases until the capacitor is fully charged. Capacitor $C_2$, with its negatively charged plate connected to the base of transistor $Q_2$ through a resistor divider consisting of $R_4$ and $R_6$, is discharged through resistor $R_5$. Resistor $R_5$ is connected to a potentiometer $R_9$ which controls the waveshape symmetry and another potentiometer $R_{10}$ which is connected to the positive supply voltage and serves as the repetition-rate control.

When the negative bias decreases to zero and the base of $Q_3$ become positive, transistor $Q_2$ turns on and causes $Q_1$ to turn off. The capacitor $C_4$ which was charged through load resistor $R_7$ starts to discharge through the $N_2$ primary windings of the pulse transformer $T_1$ after $Q_2$ is turned on. This discharge current is fed to the gate of the SCR, in the appropriate direction to fire the device. During the alternate half-cycle of multivibrator operation, capacitor $C_1$ discharges through the $N_1$ primary windings of the pulse transformer to trigger $SCR_1$.

## Applications

Some of the applications of the SCR inverter are as follows:

1. DC-to-dc converter. Conversion can be accomplished by the use of small, light-weight, low-cost transformers, inductors, and capacitors. This circuit is suitable for use in computer power supplies, telephone equipment, radio transmitters, battery chargers, and similar equipment.

2. High-frequency fluorescent-lighting supply. Because of the high frequency of the inverter circuit, the size and weight of the inductive ballast is considerably reduced; in addition, half of the inductive components can be replaced with low-cost capacitors to maintain a unity power factor in the circuit. The over-all system efficiency can also be improved; for example, the 20- to 26-per-cent power dissipation as a result of the low-efficiency ballast at 60 Hz can be reduced to a few per cent by use of high-frequency, high-efficiency inductors at moderate cost. This decrease in power dissipation in a large industrial building can mean less burden on the air-conditioning system.

# Thyristor AC Line-Voltage Controls

THE use of thyristors is becoming increasingly important for power-control applications ranging from low voltages to more than 1000 volts at current levels from less than half an ampere to more than 1000 amperes. When power control involves conversion of ac voltages and/or currents to dc and control of their magnitude, SCR's are used because of their inherent rectifying properties. SCR's are also used in dc switching applications, such as pulse modulators and inverters, because the currents in the switching device are unidirectional. A triac provides symmetrical bidirectional electrical characteristics. Triacs have been developed specifically for control of ac power.

## TYPES OF THYRISTOR TURN-ON CONTROLS

Thyristors are excellent devices for use in the control of ac power. In general, thyristors initially assume a blocking, or high-impedance state, and remain in that state until triggered to the on or low-impedance state. Once triggered, the thyristor remains on until the current is reduced to zero. The thyristor then returns to its blocking state. Because the current decreases to zero during every half-cycle in an ac supply, turn-off is guaranteed every half-cycle. All that is necessary for ac power control, therefore, is a trigger circuit to control thyristor turn-on so that whole or partial cycles may be switched to the load.

### Phase Control

In many power-control applications of thyristors, partial cycles of the applied ac voltage are switched to the load. Because the power delivered to the load is controlled by variation of the phase angle at which the thyristor switching initiates current flow, this type of operation is usually referred to as **phase control.** The electrical angle of the applied ac voltage waveform at which thyristor current is initiated is termed the **firing angle** ($\theta_F$). The **conduction angle** ($\theta_C$) is the number of electrical degrees of the applied ac voltage waveform during which the thyristor is

in conduction. The conduction angle is equal to $180° - \theta_F$ for a half-wave circuit and $2(180° - \theta_F)$ for a full-wave circuit. The voltage waveforms across the thyristor and the load for each type of circuit are illustrated in Fig. 362.

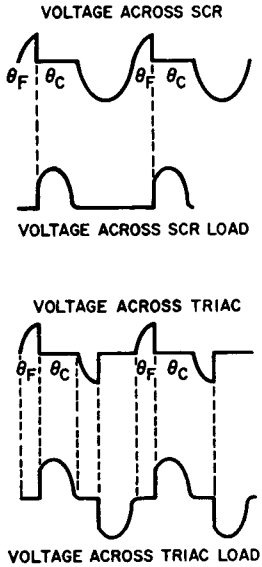



Figure 362. Voltage waveforms showing conduction angle for half-wave operation (SCR) and full-wave operation (triac) of thyristor phase-control circuits.

Phase control of thyristor-and-diode combinations may be employed to provide many different ac and dc output waveforms to a load circuit. Some basic combinations, together with the corresponding voltage waveforms at the load for two complete cycles of operation, are shown in Fig. 363. In general, triac circuits are more economical for full-wave power control than are circuits that use two SCR's. For partial range control when the load is not sensitive to a nonsymmetrical waveform, such as resistive loads, a control circuit that uses a diode and an SCR is acceptable.

In the design of thyristor power-control circuits of the types shown in Fig. 363, it is often necessary to determine the specific values of peak, average, and rms current that flow through the thyristors. For conventional rectifiers, these values are readily determined by use of the current ratios shown in Table V, given in the section on **Silicon Rectifiers.** For thyristors, however, the calculations are more difficult because the current ratios become functions of the conduction angle and the firing angle of the device.

The curves in Figs. 364, 365, and 366 show several current ratios as functions of conduction or firing angles for three basic SCR circuits. These curves can be used in a number of ways to calculate desired current values. For example, they can be used to determine the peak or rms current in an SCR when a certain average current is to be delivered to a load during a specific part of the conduction period. It is also possible to work backwards and determine the necessary period of conduction if, for example, a specified peak-to-average current ratio must be maintained in a particular application. Another use of the curves in Figs. 364, 365, and 366 is in the calculation of the rms current at various conduction angles when it is necessary to determine the power delivered to a load, or power losses in transformers, motors, leads, or bus bars. Although the curves are presented in terms of device current, they are equally useful for the calculation of load current and voltage ratios.

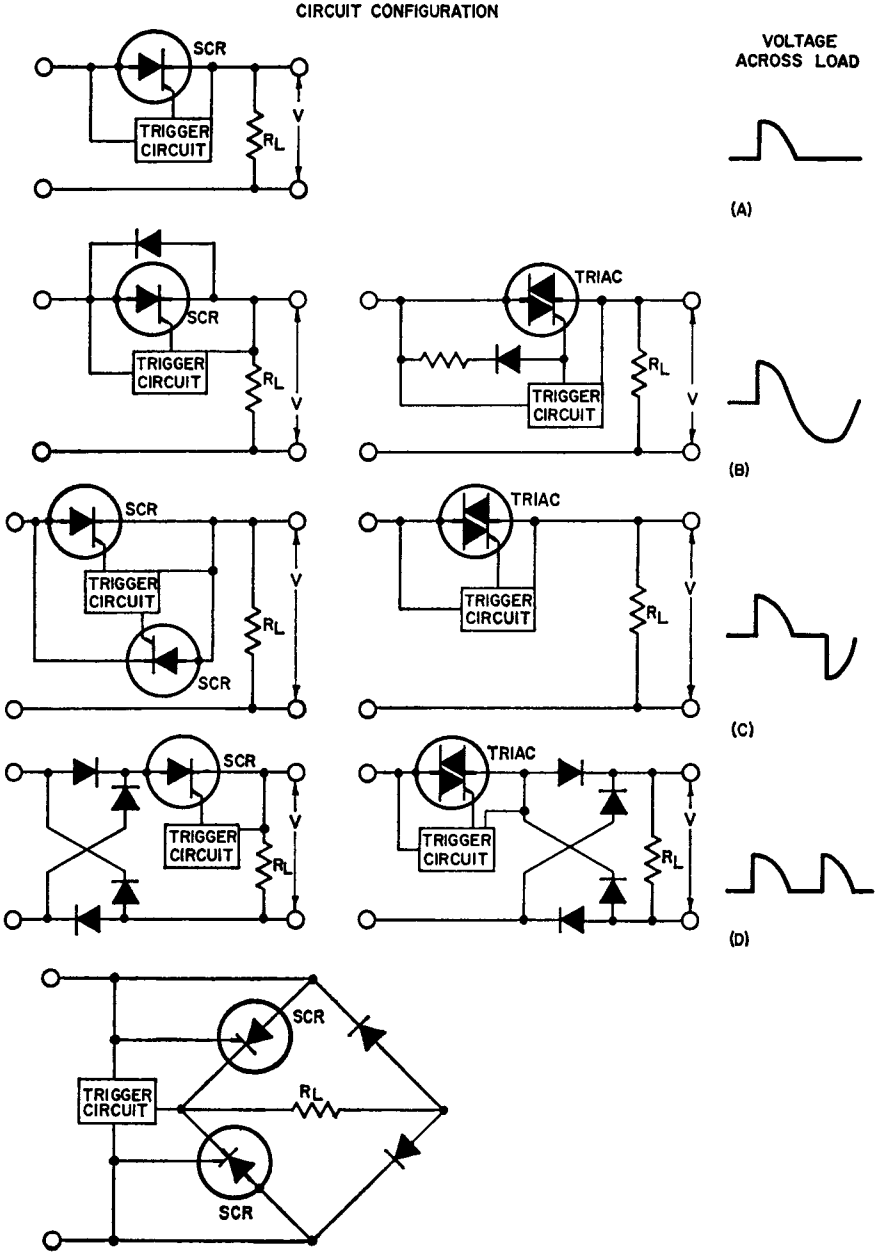



Figure 363. Basic circuit configuration for thyristor power controls and voltage waveform across the load for two complete cycles of operation.

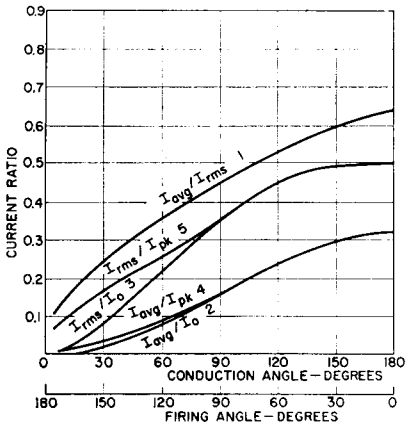



Figure 364. Ratio of SCR current as a function of conduction and firing angles for single-phase half-wave conduction into a resistive load.

The curves provide ratios that relate average current $I_{avg}$, rms current $I_{rms}$, peak current $I_{pk}$, and reference current $I_o$. The reference current is a circuit constant and is equal to the peak source voltage $V_{pk}$ divided by the load resistance $R_L$. The term $I_{pk}$ refers to the peak current that flows through the SCR during its period of forward conduction. $I_o$ is the maximum possible peak-current value during the peak of the sine wave. For conduction angles greater than 90 degrees, $I_{pk}$ is equal to $I_o$; for conduction angles smaller than 90 degrees, $I_{pk}$ is smaller than $I_o$.

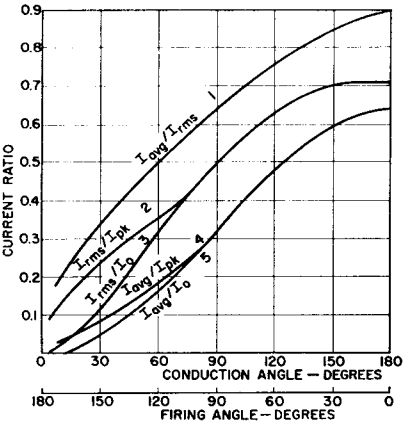



Figure 365. Ratio of thyristor current as a function of conduction and firing angles for single-phase full-wave conduction into a resistive load.

The general procedure for the use of the curves is as follows:

(1) Identify the unknown or desired parameter.

(2) Determine the values of the parameters fixed by the circuit specifications.

(3) Use the appropriate curve to find the unknown quantity as a function of two of the fixed parameters.

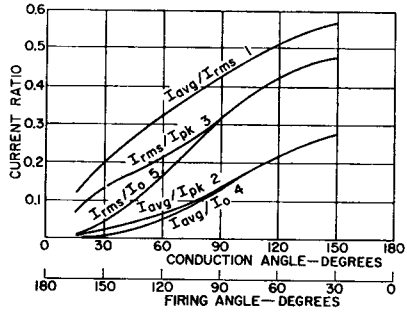
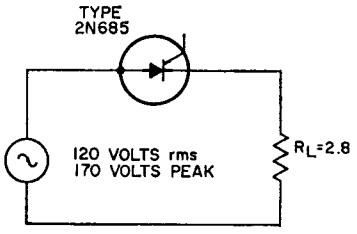



Figure 366. Ratio of SCR current as a function of conduction and firing angles for three-phase half-wave circuit having a resistive load.

**Example No. 1**—In the single-phase half-wave circuit shown in Fig. 367, a 2N685 SCR is used to control power from a sinusoidal 120-volt-rms (170-volt-peak) ac source into a 2.8-ohm load. This application requires a load current that can be varied from 2 to 25 amperes rms. It is necessary to determine the range of conduction angles required to obtain this range of load current.

First, the reference current $I_o$ is calculated as follows:




$I=0 (0° \leq \theta \leq \theta_f)$

$I=I_o \sin \theta \ (\theta_f \leq \theta \leq 180°)$

$I_{avg}=1/2\pi \int_{\theta_f}^{180°} I d\theta$

$I_{rms}=\left[1/2\pi \int_{\theta_f}^{180°} I^2 d\theta\right]^{1/2}$

$I_{pk}=I_o \ (0 \geq \theta_f \geq 90°)$

$I_{pk}=I_o \sin \theta_f \ (90° \leq \theta_f \leq 180°)$

Figure 367. Single-phase half-wave circuit that operates into a resistive load and the respective equations for SCR current.

$$I_o = \frac{V_{pk}}{R_L} = \frac{170}{2.8} = 61 \text{ amperes}$$

The ratio of $I_{rms}$ to $I_o$ for the minimum and maximum load-current values is then calculated as follows:

$$(I_{rms}/I_o)_{min} = 2/61 = 0.033$$

$$(I_{rms}/I_o)_{max} = 25/61 = 0.41$$

These current-ratio values are then applied to curve 3 of Fig. 364 to determine the corresponding conduction angles:

$$\theta_{C(min)} = 15 \text{ degrees}$$

$$\theta_{C(max)} = 106 \text{ degrees}$$

**Example No. 2**—In the single-phase full-wave bridge circuit (two legs controlled) shown in Fig. 368, a constant average load current of 7 amperes is to be maintained while the load resistance varies from 0.2 to 4 ohms. In this case, it is necessary to determine the variation required in the conduction angle. The average current through the SCR is one-half the load current, or 3.5 amperes. The applicable current ratios for this circuit are shown in Fig. 364 (the individual device currents are half-wave although the load current is full-wave).

Again, the first quantity to be calculated is the reference current $I_o$. Because the reference current varies with the load resistance, the maximum and minimum values are determined as follows:

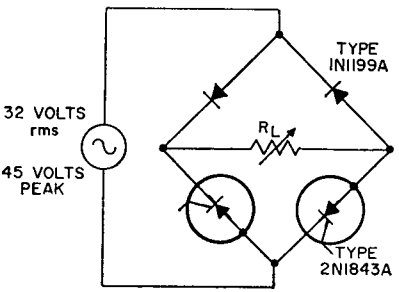
$$I_{o(max)} = V_{pk}/R_{L(min)} = 45/0.2 = 225 \text{ amperes}$$

$$I_{o(min)} = V_{pk}/R_{L(max)} = 45/4 = 11.2 \text{ amperes}$$

The corresponding ratios of $I_{avg}$ to $I_o$ are then calculated, as follows:

$$(I_{avg}/I_o)_{min} = 3.5/225 = 0.015$$

$$(I_{avg}/I_o)_{max} = 3.5/11.2 = 0.312$$




$I=0 (0° \leq \theta \leq \theta_f)$

$I=I_o \sin \theta \ (\theta_f \leq \theta \leq 180°)$

$I_{avg}=1/\pi \int_{\theta_f}^{180°} I d\theta$

$I_{rms}=\left[1/\pi \int_{\theta_f}^{180°} I^2 d\theta\right]^{1/2}$

$I_{pk}=I_o \ (0° \leq \theta_f \leq 90°)$

$I_{pk}=I_o \sin \theta_f \ (90° \leq \theta_f \leq 180°)$

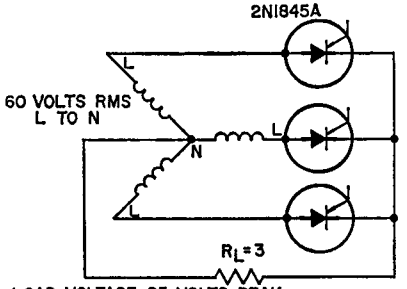
Figure 368. Single-phase full-wave bridge circuit that operates into a resistive load and the respective equations for SCR current.

Finally, these ratios are applied to curve 2 of Fig. 364 to determine the desired conduction values:

$$\Theta_{C(min)} = 25 \text{ degrees}$$

$$\Theta_{C(max)} = 165 \text{ degrees}$$

**Example No. 3**—In the three-phase half-wave circuit shown in Fig. 369, the firing angle is varied continuously from 30 to 155 degrees. In this case, it is necessary to determine the resultant variation in the attainable load power.




LOAD VOLTAGE=85 VOLTS PEAK
DEVICE VOLTAGE=85 VOLTS PEAK FORWARD
DEVICE VOLTAGE=149 VOLTS PEAK REVERSE

$I = I_o \sin\theta \quad (30° \le \theta \le 180°)$

$I_{avg} = 1/2\pi \int_{\theta_f}^{\theta_f+120} I\,d\theta \quad (30° \le \theta_f \le 60°)$

$I_{avg} = 1/2\pi \int_{\theta_f}^{180°} I\,d\theta \quad (60° \le \theta_f \le 180°)$

$I_{rms} = \left[1/2\pi \int_{\theta_f}^{\theta_f+120°} I^2\,d\theta\right]^{1/2} \quad (30° \le \theta_f \le 60°)$

$I_{rms} = \left[1/2\pi \int_{\theta_f}^{180°} I^2\,d\theta\right]^{1/2} \quad (60° \le \theta_f \le 180°)$

$I_{pk} = I_o \quad (30° \le \theta_f \le 90°)$

$I_{pk} = I_o \sin\theta_f \quad (90° \le \theta_f \le 180°)$

Figure 369. Three-phase half-wave circuit that uses a resistive load and the respective equations for SCR current.

Reference current for this circuit is determined as follows:

$$I_o = V_{pk}/R_L = 85/3 = 28 \text{ amperes}$$

Rectifier current ratios are determined from Fig. 366 for the extremes of the firing range, as follows:

$$\Theta_F = 30 \text{ degrees}; \; I_{rms}/I_o = 0.49$$

$$\Theta_F = 155 \text{ degrees}; \; I_{rms}/I_o = 0.06$$

These ratios, together with the reference current, are then used to determine the range of rms current in the rectifiers, as follows:

$$I_{rms(max)} = 0.49 \times 28 = 13.7 \text{ amperes}$$

$$I_{rms(min)} = 0.06 \times 28 = 1.7 \text{ amperes}$$

In this circuit, the rms current in the load is equal to the rms rectifier current multiplied by the square root of three; as a result, the desired power range of the load is as follows:

$$P_{max} = [I_{rms(max)} \sqrt{3}]^2 R_L = 1700 \text{ watts}$$

$$P_{min} = [I_{rms(min)} \sqrt{3}]^2 R_L = 26 \text{ watts}$$

Phase-control techniques can be used very effectively and efficiently to control ac input power in lamp-dimming, motor-speed-control, low-power electric-heating, and many other similar types of applications. Phase-control systems generate radio-frequency noise as a result of the random thyristor switching and often must include suppression circuits to minimize radio-frequency interference (RFI) in other electrical systems. In higher-power applications, the RFI is of such magnitude that the suppression circuits become excessively bulky and expensive.

## Zero-Voltage Switching

Power to an ac load may be controlled by switching of complete half-cycles or integral numbers of whole cycles of the ac power to the load. This type of control is usually referred to as **integral-cycle** or **zero-voltage-switching control.** Fig. 370 shows the relationship between line and

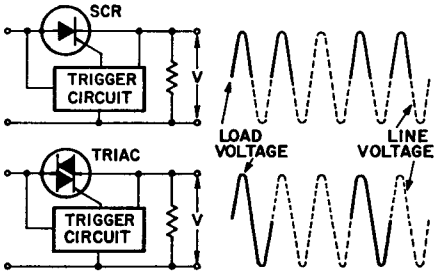



Figure 370. Integral-cycle thyristor power control circuits.

load voltages for both SCR (half-wave) and triac (full-wave) power-control circuits that employ this control technique. With this type of control, the RFI associated with phase-control circuits is substantially decreased, or even eliminated, because thyristor switching occurs at or near the 0- or 180-degree (zero-voltage) points on the ac line voltage.

In zero-voltage-switching controls, only two levels of input power are delivered to the load. The load receives the full amount of power for a period of time and zero power for a period of time; the average power delivered to the load, therefore, depends upon the ratio of the power-on interval to the power-off interval.

In solid-state power-control systems that employ zero-voltage-switching techniques, two modes of operation are possible. The controlled variable (for example, temperature in a heat-control system) may be sensed and used to turn the power on or off. Because the power-control element is a solid-state device and, therefore, is free of wearout mechanisms, the differential in the control variable that causes the switching can be made very small, and accurate control is achieved. Fig. 371 shows the response characteristics for a heating system that uses this type of control.

In control systems that have large time constants, such as a home heating system, on-off controls of the type described above may produce relatively large overshoots and undershoots. In this type of system, better regulation may be achieved by use of a control method referred to as proportional integral-cycle control.

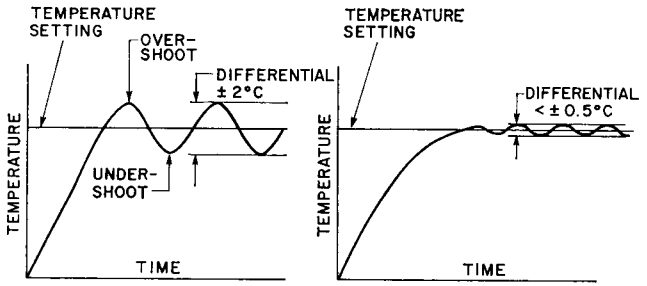



Figure 371. Transfer characteritics of (a) on-off and (b) proportional control systems.

The power excursions that result with on-off controls are substantially reduced, as shown in Fig. 371(b), by use of integral-cycle proportional control with synchronous switching. With integral-cycle proportional control, a time base is selected, and the "on" time of the thyristor is varied within the time base. The ratio of on-to-off times of the thyristor during this interval depends upon the amount of power to the load required to maintain a predetermined average level for the system. As this level is approached (as determined by a sensing element), less power is delivered to the load (i.e., the duty cycle is reduced). This type of control is usually selected for heating systems.

Fig. 372 shows the on-off-ratio of the triac. Within the time period, the on-time varies by an integral number of cycles from full on to a single cycle of input voltage.

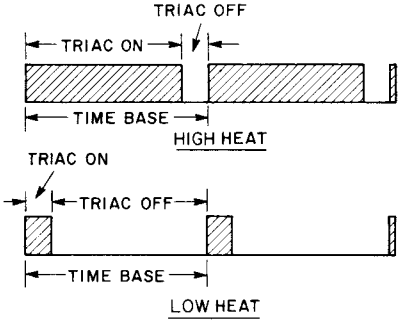



Figure 372. Triac duty cycle.

One method of achieving integral-cycle proportional control is to use a fixed-frequency sawtooth generator signal which is summed with a dc control signal. The sawtooth generator establishes the period or time base of the system. The dc control signal is obtained from the output of the temperature-sensing network. The principle is illustrated in Fig. 373. As the sawtooth voltage increases, a level is reached which turns on power to the heating elements. As the temperature at the sensor changes, the dc level shifts accordingly and changes the length of time that the power is applied to the heating elements within the established time.

When the demand for heat is high, the dc control signal is high and high power is supplied con-

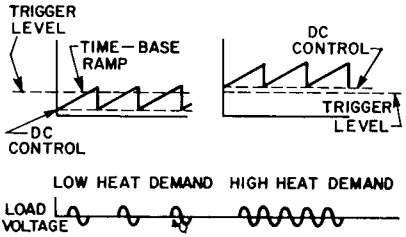



Figure 373. Proportional-controller waveshapes.

tinuously to the heating elements. When the demand for heat is completely satisfied, the dc control signal is low and low power is supplied to the heating elements. Usually a system using this principle operates continuously somewhere between full on and full off to satisfy the demand for heat.

## TRIGGERING TECHNIQUES

When thyristors are triggered, the primary requirement to assure sustained forward conduction is that the gate current is of sufficient magnitude to meet all requirements specified in the published data on the thyristor. These triggering requirements are usually stated in terms of dc voltage and current. Because it is often desirable to pulse-fire thyristors, it is also necessary to consider the duration of firing

pulse required. A trigger pulse that has an amplitude just equivalent to the dc requirements must be applied for a relatively long period of time (approximately 30 microseconds) to assure that the gate signal is provided during the full turn-on period of the thyristor. As the amplitude of the gate-triggering signal is increased, the turn-on time of the thyristor is decreased, and the width of the gate pulse may be reduced. When highly inductive loads are used, the inductance controls the current-rise portion of the turn-on time. For this type of load, the gate pulse must be long enough to assure that the principal current rises to a value greater than the latching-current level of the device. The latching current of RCA thyristors is usually twice the holding current.

The application usually determines the degree of sophistication of the circuit used to trigger a given thyristor. Triggering circuits can be as numerous and as varied as the applications in which they are used; this text discusses the basic types only.

## Basic Triggering Configurations

Many applications require that a thyristor be switched full "on" in a manner similar to the operation of a relay. The simplest method of accomplishing this type of triggering is illustrated by the circuits shown in Fig. 374.

The resistance $R_G$ maintains the gate current within the rating of the thyristor gate and the associated switch. After firing, the thyristor switches to its low-impedance state; depending on the forward-current magnitude, the voltage drop across the thyristor can be as high as a few volts. It cannot be assumed that if the resistor were removed from the gate circuit, the gate switch would carry only enough current to trigger the device and then decrease to zero. Because the gate has a low impedance, it carries a large percentage of the forward current. The gate resistor $R_G$ assures that the gate current will decrease to a negligible value after the thyristor is fired.

When an SCR is used with an ac supply, a diode may be required to keep the reverse polarity across the SCR from being impressed across the gate circuit. The allowable reverse dissipation is limited to that shown in the published data on the SCR. Fig. 374(b) illustrates the use of the SCR-and-diode combination.

When an ac resistive trigger network is used, only a certain

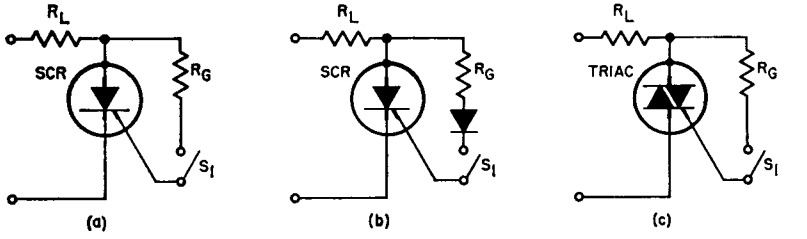



Figure 374. Simple thyristor triggering methods: (a) resistance triggering of SCR; (b) resistance-diode triggering of SCR; (c) resistance triggering of triac.

degree of phase-angle control can be accomplished. The degree of control varies from 90- to 180-degree conduction when an SCR is used and from 180- to 360-degree conduction when a triac is used. This degree of control is illustrated in Fig. 375. With maximum resistance in either circuit, the thyristor is off. As the resistance is reduced, a point is reached at which sufficient gate trigger current is provided at the peak of the voltage wave to trigger the thyristor. The thyristor initially turns on with a conduction angle $\theta_C$ of 90 degrees. A further reduction in resistance increases the conduction angle from 90 degrees toward 180 degrees for an SCR and from 90 degrees and 270 degrees to zero and 180 degrees, respectively, for a triac.

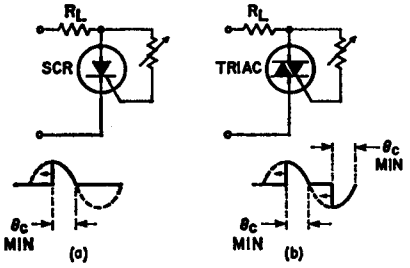



Figure 375. Degree of control over conduction angles when ac resistive network is used to trigger SCR's and triacs.

The easiest method to obtain a firing-angle delay greater than 90 degrees for half-wave operation is to use a resistance-capacitance triggering network, which is shown in its simplest form in Fig. 376(a). The polarity of the sine wave which reverse-biases the SCR charges the capacitor in the reverse direction to the peak of the line voltage through a diode. On the next half-cycle, the

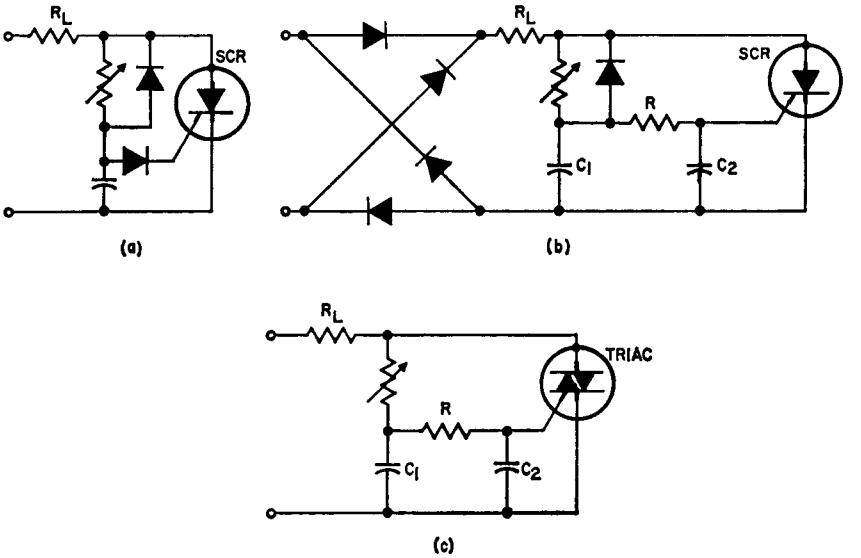



Figure 376. RC triggering networks used for phase-control triggering of thyristors: (a) series RC (single-time-constant) network and diode (for SCR); (b) RC lattice (double-time-constant) network and diode (for SCR); (c) RC lattice (double-time-constant) network without diode (for triac).

capacitor charges through the potentiometer to the relatively small positive voltage required to trigger the SCR. Controls of this type can have conduction angles from 0 to 180 degrees.

Resistance and resistance-capacitance trigger circuits have one great disadvantage: the gate voltage rises slowly to the triggering level. Because of variations in gate characteristics among thyristors (15 to 1 in gate-trigger current, 2 to 1 in gate-trigger voltage, and 4 to 1 with temperature), a given control-potentiometer resistance setting may yield a different conduction angle for different thyristors or temperature conditions. The performance of the circuit is improved somewhat by use of a double RC section as shown in Figs. 376(b) and 376(c), or by use of a negative voltage across the capacitor, as shown in Fig. 376(a). These techniques increase the rate of rise of gate voltage in the vicinity of the triggering potential and, therefore, minimize the effects of gate differences on the conduction angle.

## Triggering Devices

A variety of thyristor triggering devices are available to overcome the disadvantages noted for simple resistance or resistance-capacitance triggering circuits. These triggering devices have a smaller range of characteristics and are not so temperature-sensitive. Basically, a thyristor triggering device exhibits a negative resistance after a critical voltage is reached, so that the gate-current requirement of the thyristor can be obtained as a pulse from the discharge of the phase-shift capacitor. Because the gate pulse need be only microseconds in duration, the gate-pulse energy and the size of the triggering components are relatively small. Triggering circuits of this type employ elements such as neon bulbs, trigger diodes, unijunction transistors, and two-transistor switches.

**Basic Operation**—The most elementary form of triggering-device circuit is shown in Fig. 377. The voltage-current characteristic for the trigger device used

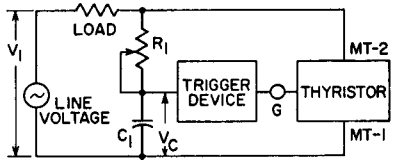



Figure 377. Thyristor power control in which switching is controlled by basic triggering-device circuit.

in this circuit is shown in Fig. 378.

When the variable resistor $R_1$ is adjusted for maximum resistance in the circuit, the circuit operates so that the series combination of $R_1$ and $C_1$ produces a voltage $V_C$, which is reduced in magnitude and shifted in phase with respect to the line voltage $V_1$. When the peak value of the

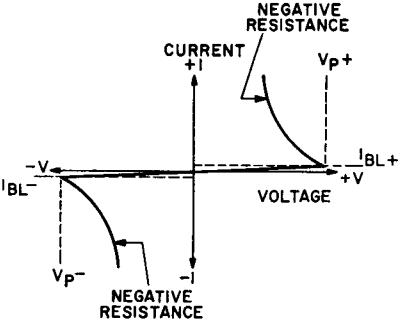



Figure 378. Voltage-current characteristic for triggering device shown in Fig. 377.

voltage $V_C$ is less than the triggering-device breakdown voltage $V_P$, the trigger device does not conduct; the thyristor receives no gate current and remains in the blocking state. The blocking action of the thyristor prevents the line voltage from appearing across the load. This condition is shown in Fig. 379(a).

As the value of $R_1$ is decreased, the voltage $V_C$ increases in magnitude and changes in phase with respect to the line voltage $V_1$. This change continues as long as the value of $R_1$ is decreased, and eventually a point is reached at which $V_C$ exceeds $V_P$. At this point, the trigger device instantaneously switches to a negative-resistance characteristic. This action causes a sudden discharge of capacitor $C_1$, which provides a pulse of gate current to the thyristor. The thyristor is then triggered into conduction and remains in the on state for the rest of that particular half-cycle of line voltage.

The magnitude and duration of the gate-current pulse are determined by the interaction of the capacitor $C_1$, the triggering-device characteristics, and the impedance of the thyristor gate. This interaction can be represented by the curves shown in Fig. 380.

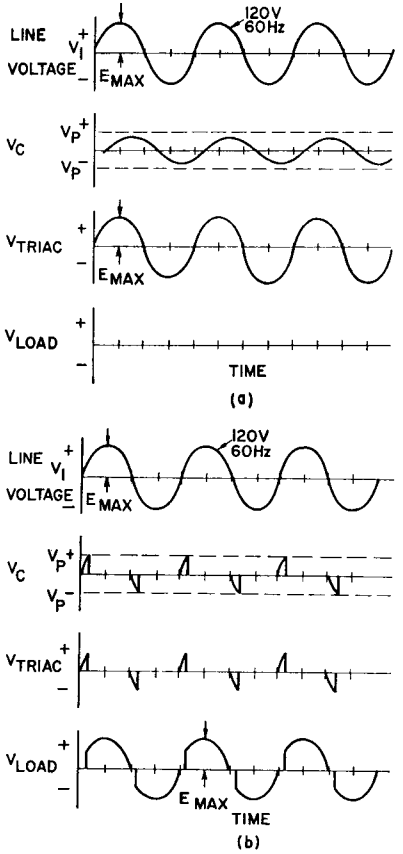



Figure 379. Voltage waveforms for circuit shown in Fig. 378: (a) thyristor in blocking state; (b) when thyristor is triggered at controlled intervals by triggering-device circuit.

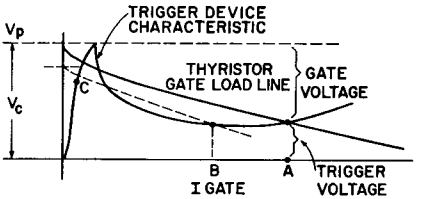



Figure 380. Load line for the circuit shown in Fig. 377.

The capacitor, which is charged to the voltage $V_C$, discharges through the negative-resistance slope of the trigger device, and the gate current rises to some magnitude A at which the total voltage drops in the circuit are equal to the voltage source $V_C$. The capacitor voltage immediately begins to decrease from its initial value $V_C$, at a rate determined by the current level reached and the size of the capacitor. As it does, the gate current decreases.

The load-line representation at some later instant is shown by the dashed line in Fig. 380, and gate current at that instant is defined by point B. As this process continues, the circuit enters the unstable negative-slope region of the triggering device characteristic and quickly reverts to a stable point, approximately indicated by C. Fig. 381 shows the typical shape of the gate-current pulse that is produced. (More specific magnitudes are shown in later diagrams for particular

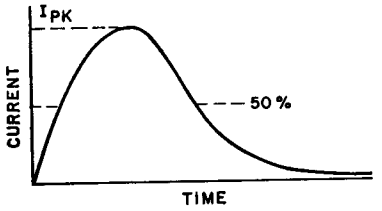



Figure 381. Typical gate-current waveform for circuit shown in Fig. 377.

triggering devices.) The delay in reaching the peak gate current is a function of the speed at which the triggering device is switched from its high-impedance to its low-impedance state. This delay in effect indicates that there is a dynamic or time-dependent characteristic of the trigger device, which traces out a shape somewhat different from the static characteristic shown in Fig. 380.

The magnitude and duration of the gate pulse produced by the triggering device and the capacitor must be adequate to fire the thyristor. A curve of turn-on time as a function of gate-pulse magnitude, provided in the published data on the thyristor, defines the minimum requirements.

Because the thyristor is triggered to the on state by the gate pulse, and the voltage source for the triggering circuit is taken from across the thyristor, the triggering circuit cannot go through another charge-discharge cycle after the first firing pulse. The capacitor discharges from point C through the potentiometer and the thyristor for the remainder of the ac line-voltage cycle, and the triggering process repeats on the next ac line-voltage half-cycle.

The maximum voltage applied to the load is limited by the breakover voltage of the trigger diode because the line voltage must rise to that value before the thyristor gate can be energized. This condition is illustrated by the voltage waveforms shown in Fig. 379(b).

Several types of devices commonly used to trigger RCA thyristors are discussed in the following paragraphs:

**Neon Bulbs**—Neon bulbs can be used as triggering devices for RCA thyristors. The breakover voltages for these devices range from 50 volts to 100 volts, with typical values of 80 volts. Tighter breakover voltage spreads can be obtained by manufacturer's selections. A typical current pulse resulting from a 0.1-microfarad capacitor discharging through a neon bulb and a thyristor gate is illustrated in Fig. 382.

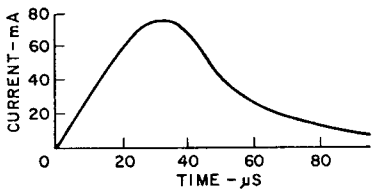



Figure 382. Typical current pulse that results from the discharge of a 0.1-microfarad capacitor through a neon bulb and a thyristor gate.

Fig. 383 illustrates the use of a neon bulb as a triggering device. The bilateral characteristic of the neon bulb allows it to trigger both SCR's and triacs.

Use of a neon bulb as a trigger device does have disadvantages. For example, when this type of device is used as a trigger on a 120-volt-rms ac line, an rms voltage loss as great as 10 per cent can occur at the load. The losses are caused by the relatively high breakdown voltage of the neon bulb. The neon bulb is also sensitive to radiation in that the breakdown point changes. When precise control is required, it may be necessary to shield the bulb or to obtain bulbs specially treated to minimize the effects of radiation. A major advantage of neon triggers is that relatively reliable and long-lived triggers can be obtained for a low price.

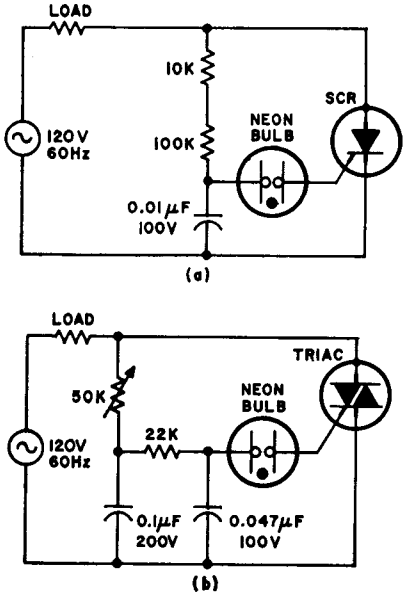



Figure 383. Circuits showing application of neon bulb as thyristor triggering device: (a) SCR power control circuit; (b) triac power control circuit.

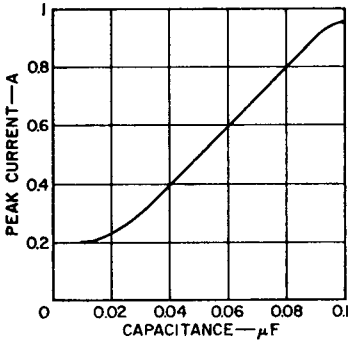
**Trigger Diodes**—A trigger diode is the solid-state replacement for a neon bulb in phase-control triggering circuits. These diodes offer the advantages of reduced requirements for peak-voltage firing, higher pulse-current capability, and longer life. The solid-state diodes have breakdown voltages in the range of 27 to 37 volts and are designed specifically for triggering bidirectional thyristors (triacs).

The trigger diodes, often referred to as **diacs**, are three-layer symmetrical avalanche devices which break over in the negative-resistance region whenever a particular voltage, termed the breakover voltage, is exceeded in either voltage polarity. In these devices, a maximum limit is usually imposed on the symmetry between positive and negative breakover voltages (voltage symmetry). The voltage-current characteristic of a bidirectional trigger diac is essentially the same as the triggering-device characteristic shown earlier in Fig. 378.

The slight current offset in the characteristic before the voltage breakover point is leakage current $I_{BO}$, which is usually in the order of 50 microamperes.

The magnitude and duration of the gate current pulse are determined by the value of the phase-shift capacitance, the change in voltage across and the dynamic impedance of the trigger diac, and the thyristor gate impedance. The interaction of all circuit impedances and the phase-shift capacitance can best be represented by the curve of peak current as a function of

the capacitance shown in Fig. 384.



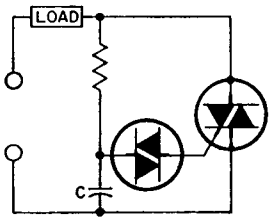




Figure 384. Peak pulse trigger currents as a function of the phase-shift capacitance.

**Unijunction Transistor**—The unijunction transistor is a three-terminal two-layer device formed by an emitter and a base. One lead is connected to the emitter and the other two leads are connected to the base. Between the two base connections there is an "interbase resistance." The basic operation of the unijunction transistor was described and application of this device in a pulse-triggering circuit was explained previously in the section on **Materials, Junctions, and Devices.**

The disadvantage of the unijunction device is that it is unilateral with regard to current flow and requires a dc voltage. These requirements indicate that diodes must be used to assure that no reverse voltage appears across the device when it is used in an ac circuit. The output pulses are positive-going and can be used to trigger SCR's directly. For triacs or inverse parallel SCR's, transformer or capacitive coupling is required, as shown in Fig. 385.

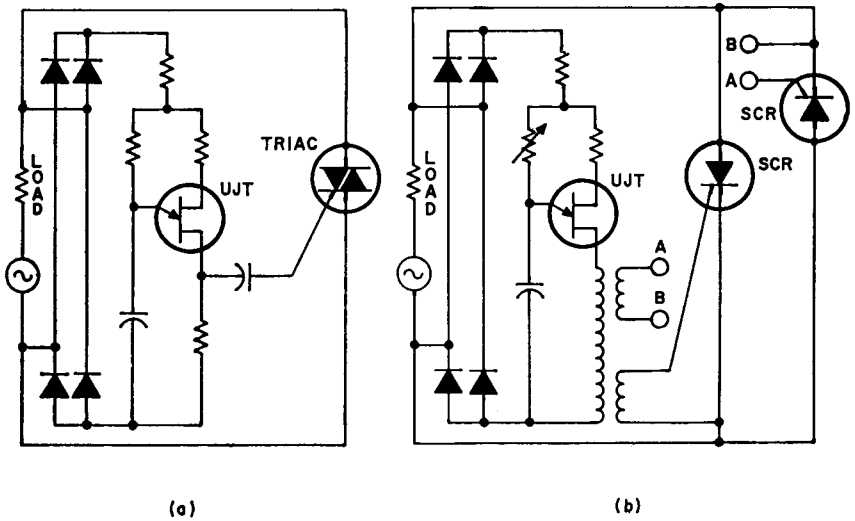



Figure 385. Circuit showing application of unijunction-transistor circuit for pulse triggering of triacs and inverse parallel SCR's: (a) triggering pulse is capacitively coupled to gate of triac; (b) triggering pulse is transformer-coupled to gate of SCR.

**Two-Transistor Trigger Circuit**—A two-transistor trigger circuit that has characteristics similar to those of a trigger diode is shown in Fig. 386. The regenerative action of this type of circuit when either transistor begins

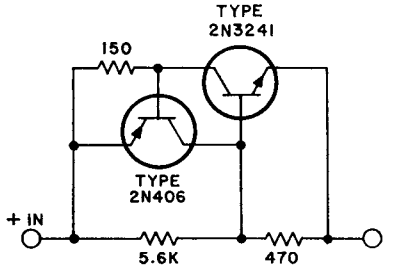



Figure 386. Two-transistor switch. Characteristics of this circuit are similar to those of a trigger diode.

to conduct causes switching comparable to avalanching in a trigger diode. Proper biasing of this circuit yields triggering voltages of 15 volts or less. The circuit shown in Fig. 387 can deliver trigger currents as high as 1 ampere and is more than capable of triggering all RCA thyristors.

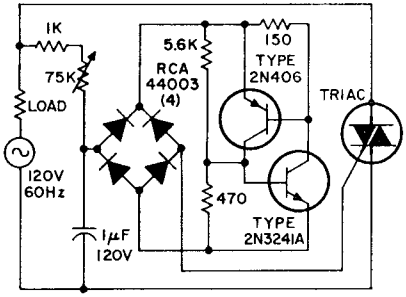



Figure 387. Circuit showing application of two-transistor switch as thyristor triggering device.

Fig. 388 shows an SCR circuit that uses the two-transistor regenerative-trigger network with a motor load. The phase-shift characteristics are still retained to provide conduction angles less than 90 degrees through the RC network of $R_1$, $R_2$, and $C_1$. Resistor $R_3$ provides turn-on current to the base of $Q_1$ when the voltage across $C_1$ becomes large enough during the positive half-cycle. The base current in $Q_1$ turns on this transistor. Transistor $Q_1$ then supplies base current to $Q_2$. When $Q_2$ turns on, it supplies more base current to $Q_1$. This regenerative action leads to the rapid saturation of transistors $Q_1$ and $Q_2$. Capacitor $C_1$ discharges through the saturated transistors into the gate of the SCR. When the SCR fires, the remaining portion of the positive half-cycle of ac power is applied to the load. Speed control is accomplished by adjustment of potentiometer $R_1$. For the component values shown on the schematic diagram in Fig. 388, the threshold voltage for firing the circuit is approximately 8 volts, and the maximum conduction angle is approximately 170 degrees. Table XXV shows values for operation of the circuit with various RCA SCR's.

An advantage of the two-transistor trigger circuit is its low threshold triggering voltage. For all practical purposes, a full 180-degree conduction angle can be obtained when an SCR is used. When two SCR's are to be triggered, a transformer must be used to couple a gate signal of the proper polarity to the SCR with the proper anode-to-cathode polarity. A triac, however, can be triggered in either direction with positive-polarity gate signals. The only requirement is that isolation be maintained between the dc and ac current.

**Application Guide for Triggering Devices**—Table XXVI provides a quick reference to the

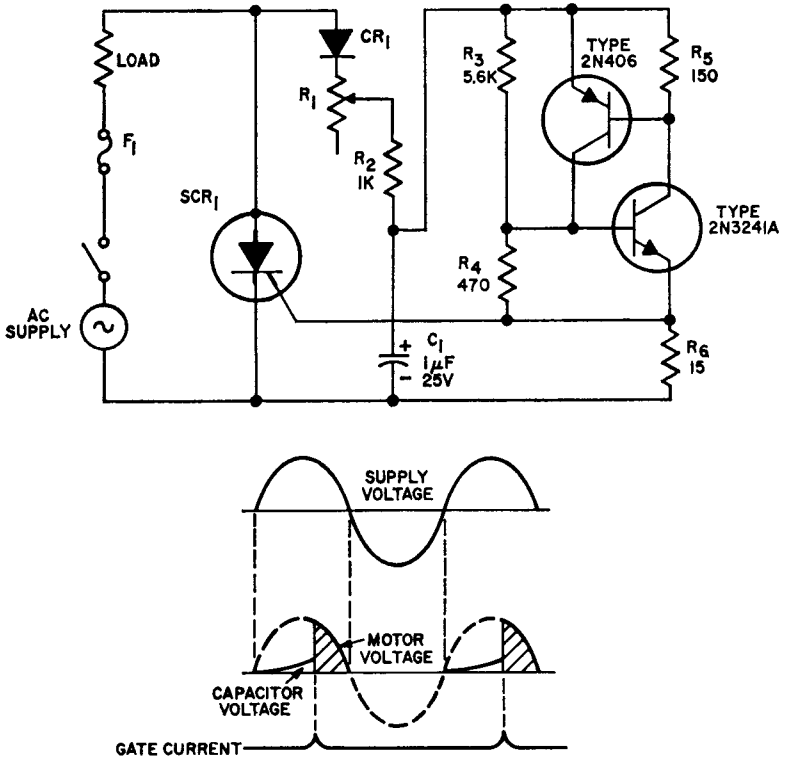



Figure 388. Half-wave SCR motor control circuit, without regulation.

Table XXV—Components For Circuit Shown in Fig. 388

| AC SUPPLY | AC CURRENT | $F_1$ | $CR_1$ | $R_1$ | $SCR_1$ |
|---|---|---|---|---|---|
| 120V | 1A | 3AG, 1.5A, Quick Act | RCA-44004 | 75K, ½W | RCA-2N3528 |
| 120V | 3A | 3AB, 3A | RCA-44004 | 75K, ½W | RCA-2N3228 |
| 120V | 7A | 3AB, 7A | RCA-44004 | 75K, ½W | RCA-2N3669 |
| 120V | 25A | 3AB, 25A | RCA-44004 | 75K, ½W | RCA-2N3897 |
| 240V | 1A | 3AG, 1.5A, Quick Act | RCA-44005 | 150K, ½W | RCA-2N3529 |
| 240V | 3A | 3AB, 3A | RCA-44005 | 150K, ½W | RCA-2N3525 |
| 240V | 7A | 3AB, 7A | RCA-44005 | 150K, ½W | RCA-2N3670 |
| 240V | 25A | 3AB, 25A | RCA-44005 | 150K, ½W | RCA-2N3898 |

prevalent types of applications for various triggering devices.

## Isolated Trigger Circuits

In many applications, phase control of a thyristor by use of an isolated, low-voltage control element is desirable. A number of techniques are available for use in this type of control.

**Step-Down Transformers** — A suitable step-down transformer can be used as shown in Fig. 389 to provide isolation between the load and the control element in a

Table XXVI—Triggering-Device Families

| | LOW COST | MEDIUM COST | HIGH COST |
|---|---|---|---|
| **Function** | Manual or Simple On-Off Power Control | Automatically Controlled or Regulated power | Power-Output Stage in Large Electronic or Electro-Mechanical System |
| **Typical Applications** | Light Dimmers<br>Tool Speed Controls<br>Appliance Speed Controls<br>Gas Ignition<br>Photoelectric Controls<br>Static On-Off Switches | Regulated Power Supplies<br>Temperature Controls<br>Commercial DC Motor Drives<br>Flashers<br>Time Delays<br>Static On-Off Power Relays | Bulk Power Conversion for Metal Refining and Electrochemical Processes<br>Large Industrial Motor Drives<br>Variable-Frequency Drives<br>Pulse Modulators<br>Precise Process-Temperature Controls<br>Logic-Arrays Power Output<br>eg.—Vending Machines<br>—Signs and Scoreboards<br>—Computer Printer Driver |
| **Common Characteristics of the Application** | Frequently Bidirectional<br>Small physical size an asset<br>Low performance demands | Frugal but not poor<br>Both dc and ac loads<br>Thyristor fired is higher cost<br>Technically oriented users and applications<br>Electrical feedback or sensor input in addition to manual control<br>Long firing pulses often required | Firing circuit small percentage of system cost<br>Rigid and extensive performance requirements<br>Firing circuit often merged into other system circuits<br>Custom engineered<br>Primarily electrical inputs from regulators or sensors<br>Often built up from standard logic and waveshaping circuits |
| **Trigger Devices in Approximate Order of Preference or Use** | 1. Diac<br>2. Neon bulb<br>3. Four-layer diode<br>4. Unijunction<br>5. Two-transistor regenerative circuit<br>NOTE: For on-off control, a switch contact or single transistor may form the firing circuit | 1. Unijunction transistor<br>2. Transistors<br>3. Integrated Circuits<br>4. Magnetic amplifier<br>NOTE: Firing circuit often includes several diodes, a zener, pulse transformers, control power transformers, and numerous passive components | 1. Transistors<br>2. Integrated Circuits |

thyristor power control circuit. The transformer provides impedance transformation and reflects the transformed value of potentiometer $R_1$ into the secondary circuit. As the potentiometer is adjusted from a maximum to a minimum value, the voltage at point A in Fig. 389 varies from a minimum to a maximum value. The trigger circuit may be either a modified diac-resistance-capacitance network or a unijunction type of circuit. The advantage of the circuit shown in Fig. 389 is that it provides low-voltage potentials and isolation for remote controls in a conductive environment.

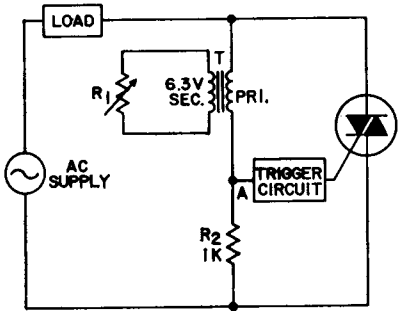



Figure 389. Triac power-control circuit in which a step-down transformer is used to provide isolation of the control potentiometer.

**Pulse Transformers**—When a trigger circuit provides a fast-rising current pulse for gated turn on, a pulse transformer provides a simple form of circuit isolation between the line-voltage system and the control system. Fig. 390 illustrates this application of a pulse transformer.

extended relay contact life. A circuit using this technique is illustrated in Fig. 391.

**Photocell Trigger Circuits**—Optical coupling between trigger circuits and load circuits provides complete dc isolation. In addition, this technique permits variable power control as well as on-off switching control functions. The optical control element of this system is a light source of the filament type, gas-discharge type, or the light-emitting-diode (LED) type. The sensing element is a photosensitive resistor, which forms part of the gate circuit.

Photoresistors are available in a wide range of dark-to-light resistance ratios and resistance values. Integral photocouplers which use any one of the three types of light sources described above are also available. Variable pow-

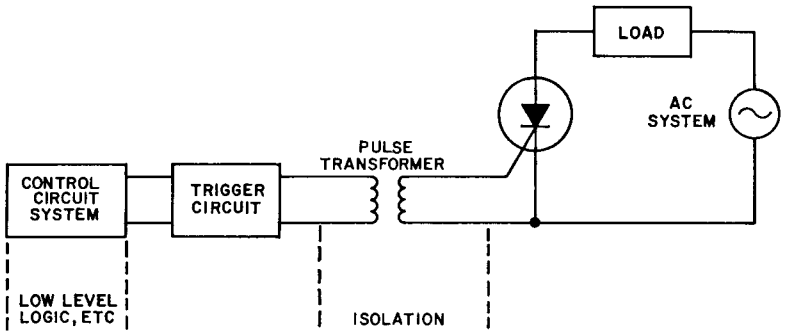



Figure 390. SCR power-control circuit that uses a pulse transformer to provide isolation between the line voltage and the control circuit.

**Reed Relays**—For applications in which an on-off switching function is required and isolation is necessary between the control circuit and load circuits, a current reed relay can be used to provide an effective method of control. The contacts of the reed relay carry only a small current, usually less than 100 milliamperes to assure an

er control can be achieved by changing the resistance value of the photoresistor by use of a low-power variable-intensity light source. The range of control depends on the type of light source and photoresistor. A photoresistor can be used by itself to sense ambient light intensity for a switching application or for modulating

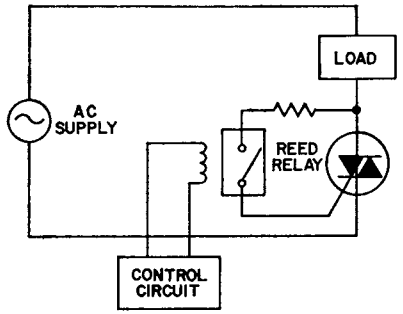



Figure 391. Triac power-control circuit that uses a reed relay to provide isolation between the control circuit and the load.

switch is a monolithic integrated circuit used primarily as a trigger circuit for the control of thyristors. The multistage circuit employs a diode limiter, a threshold detector, a differential amplifier, and a Darlington output driver to provide the basic switching action. The dc supply voltage for these stages is supplied by an internal zener-diode-regulated power supply that has sufficient current capability to drive external circuit elements, such as transistors and other integrated circuits. This

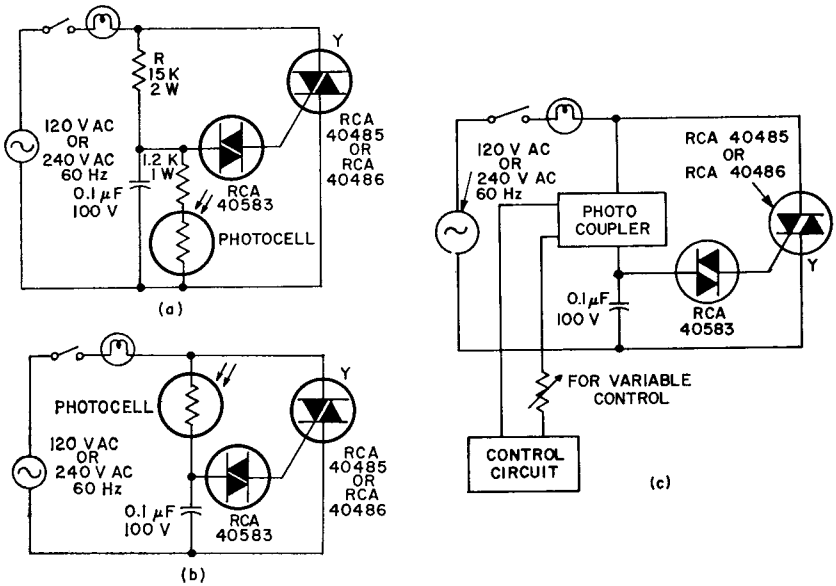



Figure 392. Triac power-control circuits that use optical coupling to provide complete dc isolation between the control circuit and the load: (a) light-controlled turn-off circuit with photocell sensor; (b) light-control turn-on circuit with photocell sensor; (c) light-controlled turn-off circuit with photocoupler sensor.

power as a function of ambient light intensity. Fig. 392 illustrates circuit applications using photo trigger circuits.

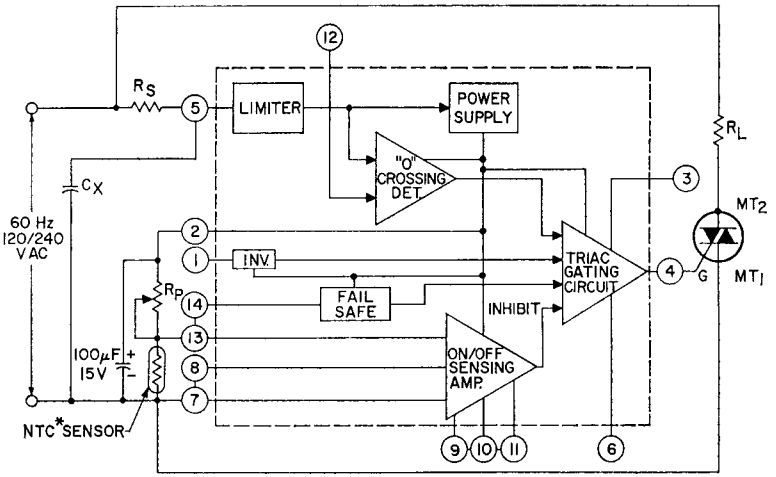
## Integrated-Circuit Zero-Voltage Switch

The RCA-CA3059 zero-voltage built-in power supply provides unique solutions to many application problems. An important feature of the CA3059 is that the trigger pulses developed by this circuit can be applied directly to the gate of a silicon controlled rectifier (SCR) or a triac. A built-in fail-safe circuit inhibits the application of these pulses to the

thyristor gate circuit in the event that the external sensor for the integrated-circuit switch should be inadvertently opened or shorted.

**Circuit Operation**—Fig. 393 shows a functional block diagram of the CA3059 integrated-circuit zero-voltage switch. Any triac that is driven directly from the output terminal of this circuit should be characterized for operation in the I(+) or III(+) triggering modes, i.e., with positive gate current (current flows into the gate for both polarities of the applied ac voltage).

The limiter stage of the CA3059 clips the incoming ac line voltage to approximately plus and minus 8 volts. This signal is then applied to the zero-voltage-crossing detector, which generates an output pulse during each passage of the line voltage through zero. The limiter output is also applied to a rectifying diode and an external capacitor that comprise the dc power supply. The power supply provides approximately 6 volts as the $V_{CC}$ supply to the other stages of the CA3059. The on/off sensing amplifier is basically a differential comparator. The triac gating cir-




* NTC=NEGATIVE TEMPERATURE COEFFICIENT

| AC Input Voltage (Volts) 50/60 or 400 Hz | Series Resistor $R_S$ (kΩ) | Power Rating of $R_S$ (Watts) |
|---|---|---|
| 24 | 2 | 0.5 |
| 120 | 10 | 2 |
| 208/230 | 20 | 4 |
| 277 | 25 | 5 |

Figure 393. Functional block diagram of the integrated-circuit zero-voltage switch.

cuit contains a driver for direct triac triggering. The gating circuit is enabled when all the inputs are at a high voltage, i.e., the line voltage must be approximately zero volts, the sensing-amplifier output must be "high," the external voltage to terminal 1 must be a logical "1," and the output of the fail-safe circuit must be "high."

Fig. 394 shows the circuit diagram of the CA3059. The zero-voltage threshold detector consists of diodes $D_3$, $D_4$, $D_5$, and $D_6$, and transistor $Q_1$. The differential amplifier consists of transistor pairs $Q_2$-$Q_4$ and $Q_3$-$Q_5$. Transistors $Q_1$, $Q_6$, $Q_7$, $Q_8$, and $Q_9$ comprise the triac gating circuit and driver stage. Diode $D_{12}$, zener diode $D_{15}$, and transistor $Q_{10}$ constitute the fail-safe circuit. The power supply consists of diodes $D_7$ and $D_{13}$, and an external resistor and capacitor connected to terminals 5 and 2, respectively, and to ground through pin 7. If the transistor pair $Q_2$-$Q_4$ and transistor $Q_1$ are turned off, an output appears at terminal 4. Transistor $Q_1$ is in the off state if the incoming line voltage is less than approximately the voltage drops across three silicon diodes (2.1 volts) for either the positive or negative excursion of the line voltage. Transistor pair $Q_2$-$Q_4$ is off if the voltage across the sensor, connected from terminals 13 to 7, exceeds the reference voltage from 9 to 7. If either of these

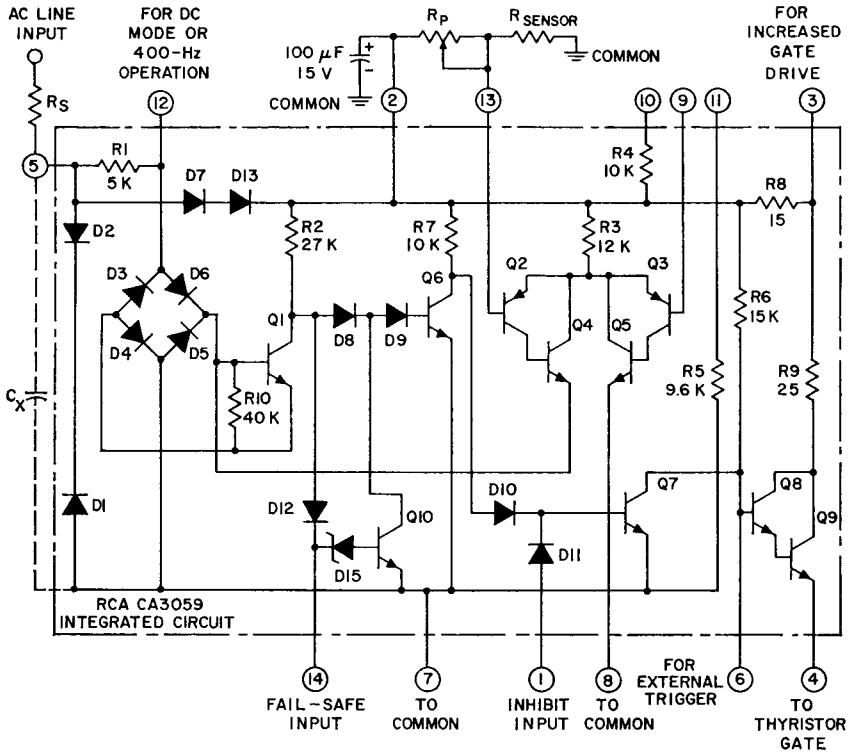



Figure 394. Circuit diagram for the CA3059 zero-voltage switch.

conditions is not satisfied, pulses are not supplied to terminal 4. Fail-safe operation requires that terminal 13 be connected to 14. The addition of hysteresis and elimination of half-cycling can be obtained by a resistive voltage divider connected from 13 to 8 and from 8 to 7.

Fig. 395 shows the position and width of the pulses supplied to the gate of a thyristor with respect to the incoming ac line voltage. The CA3059 can supply sufficient gate voltage and current to trigger most RCA thyristors at ambient temperatures of 25°C. However, under worst-case conditions (i.e., at ambient-temperature extremes and maximum triggering requirements), selection of the higher-current thyristors may be necessary for particular applications.

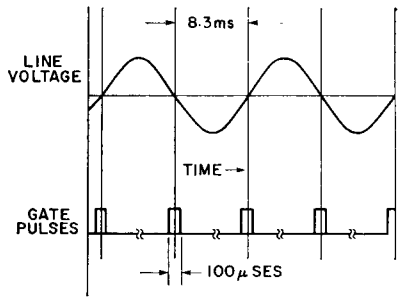



Figure 395. Timing relationship between the output pulses of the CA3059 and the ac line voltage.

**Effect of CA3059 on Thyristor Load Characteristics**—The CA3059 is designed primarily to gate a thyristor that switches a resistive load. Because the output pulse supplied by the CA3059 is of short duration, the latching current of the triac becomes a significant factor in determining whether other types of loads can be switched. (The latching-current value determines whether the triac will remain in conduction after the gate pulse is removed.) Provisions are included in the CA3059 to accommodate inductive loads and low-power loads. For example, for loads that are less than approximately 4 amperes rms or that are slightly inductive, it is possible to retard the output pulse with respect to the zero-voltage crossing by insertion of the capacitor $C_x$ from terminal 5 to terminal 7 as shown in Fig. 393. The insertion of capacitor $C_x$ permits switching of triac loads that have a slight inductive component and that are greater than approximately 200 watts (for operation from an ac line voltage of 120 volts rms). However, for loads less than 200 watts (for example, 70 watts), it is recommended that the user employ the RCA-40526 sensitive-gate triac with the CA3059 because of the low latching-current requirement of this triac.

For loads that have a low power factor, such as a solenoid valve, the user may operate the CA3059 in the dc mode. In this mode, terminal 12 is connected to terminal 7, and the zero-crossing detector is inhibited. Whether a "high" or "low" voltage is produced at terminal 4 is then dependent only upon the state of the differential comparator within the CA3059 integrated circuit, and not upon the zero crossing of the incoming line voltage. Of course, in this mode of operation, the CA3059 no longer operates as a zero-voltage switch. However, for many applications that involve the switching of low-current inductive loads, the amount of RFI generated can frequently be tolerated.

**Fail-Safe Feature**—As shown in Figs. 393 and 394, when terminal 13 is connected to terminal 14, the fail-safe circuit of the CA3059

is operable. If the sensor should then be accidentally opened or shorted, power is removed from the load (i.e., the triac is turned off). The internal fail-safe circuit functions properly, however, only when the ratio of the sensor impedance at 25°C, if a thermistor is the sensor, to the impedance of the potentiameter $R_p$ is less than 4 to 1. It is readily apparent that, if the potentiometer is adjusted for 1000 ohms and the sensor is 100,000 ohms, the zener diode $D_{15}$ (shown in Fig. 394) would conduct because virtually all the dc power-supply voltage (from terminal 2 to terminal 7) would appear across the sensor. The CA3059 would then detect this condition as an open sensor.

For ratios greater than 4 to 1, for example 100 to 1, the circuit shown in Fig. 396 may be employed to provide fail-safe operation. In this circuit, transistor $Q_1$ and diode $D_1$ are components external to the CA3059. Transistor $Q_1$ detects the sensor current which maintains this transistor in saturation so that terminal 1 is effectively shorted to terminal 7 through the collector-to-emitter junction of the transistor. Transistor $Q_1$ provides sufficient current gain to permit operation with a sensor impedance greater than 1 megohm. If the sensor becomes open-circuited, transistor $Q_1$ turns off, and current then flows into terminal 1, the inhibit terminal of the CA3059, and results in the removal of power to the load. For the shorted-sensor condition, the external diode $D_1$ conducts and causes triac $Y_1$ to turn off. Diode $D_2$ compensates for variations in the base-to-emitter voltage of transistor $Q_1$ with temperature. Terminals 13 and 14 on the CA3059 should not be connected when the external fail-safe circuit shown in this illustration is employed.

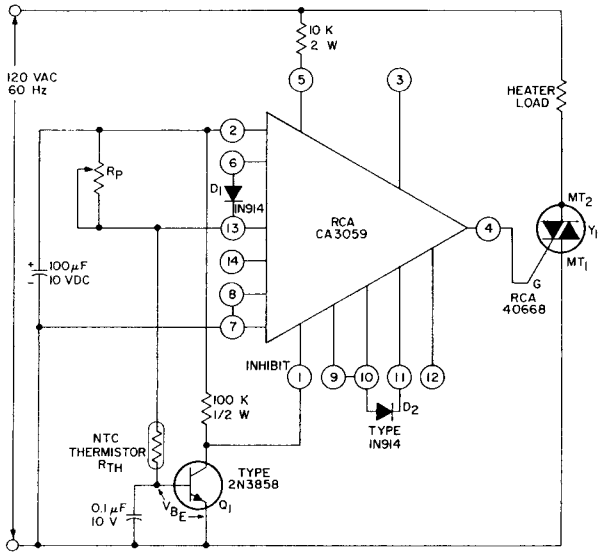



Figure 396. CA3059 on-off controller that uses an external fail-safe circuit.

**Half-Cycling and Hysteresis Characteristics**—The method by which the CA3059 senses the zero crossing of the ac power results in a half-cycling phenomenon at the control point. Fig. 397 illustrates this phenomenon. The CA3059 senses the zero-voltage crossing every half-cycle, and an output, for example pulse No. 4, is produced to indicate the zero crossing. During the remaining 8.3 milliseconds, however, the differential amplifier in the CA3059 may change state and inhibit any further output pulses. The uncertainty region of the differential amplifier, therefore, prevents pulse No. 5 from triggering the triac during the negative excursion of the ac line voltage.

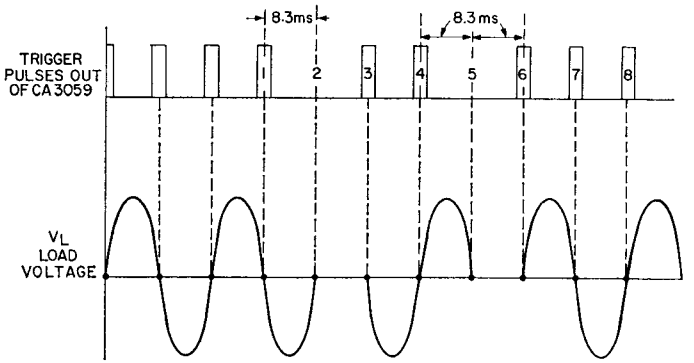



Figure 397. Half-cycling phenomenon in the CA3059.

Several solutions exist for elimination of the half-cycling phenomenon. If the user can tolerate some hysteresis in the control, then positive feedback can be added around the differential amplifier.

## THREE-PHASE TRIAC CONTROLS

The growing demand for solid-state switching in heating controls and other industrial applications has resulted in the increasing use of three-phase triac power-control circuits. The following paragraphs describe the use of triacs to control the application of ac power to both three-phase resistive and inductive loads. In the circuits described, the RCA-CA3059 integrated-circuit zero-voltage switch is used as an interface control from the low-power logic circuitry to the high-power load. The requirements of the three-phase triac controls are as follows:

1. The load should be connected in either a three-wire delta or wye configuration. Four-wire wye loads may be handled as three independent single-phase systems.
2. Only one logic command signal is available for the control circuits. This signal must be electrically isolated from the three-phase power system.
3. Three separate triac gating signals are required.
4. With resistive loads, the zero-voltage-switching technique should be used to minimize any RFI/EMI that may be generated.

The electrical isolation of the command signal required in the control circuits is achieved by photo-optic techniques. Other techniques, such as pulse transformer,

magnetoresistor, or reed relay, may also be used with some circuit modifications.

## Three-Phase Resistive Loads

Fig. 398 illustrates the basic phase relationships of a balanced three-phase resistive load, such as may be used in heater applications, in which the application of load power is controlled by zero-voltage switching. The following conditions are inherent in this type of application:

1. The phases are 120 degrees apart; consequently, all three phases cannot be switched on simultaneously at zero voltage.

2. A single phase of a three-wire system cannot be turned on.

3. Two phases must be turned on for initial starting of the system. These two phases form a single-phase circuit which is out of phase with both of its component phases. The single-phase circuit leads one phase by 30 degrees and lags the other phase by 30 degrees.

These conditions indicate that in order to maintain a system in which no appreciable RFI/EMI is generated by the switching action, from initial starting through the

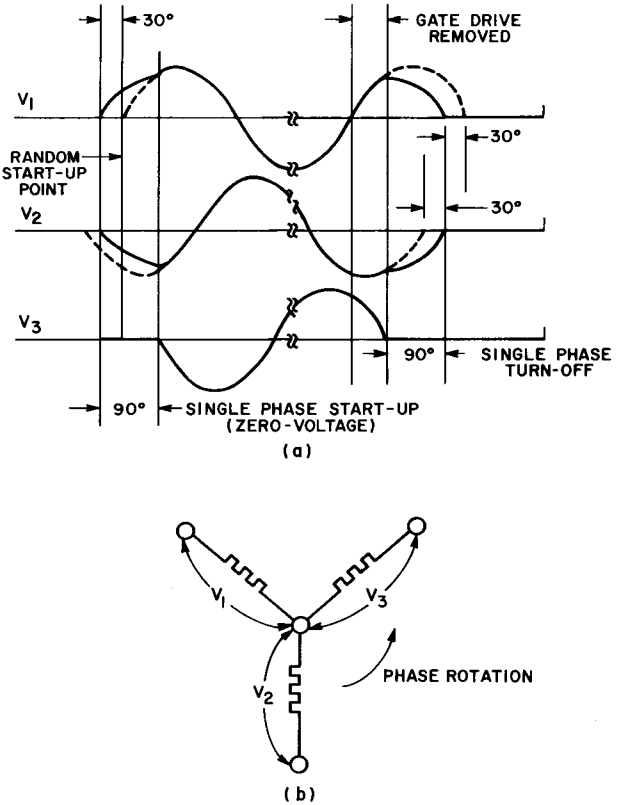



Figure 398. Voltage phase relationship for a three-phase resistive load when the application of load power is controlled by zero-voltage switching. (a) voltage waveforms; (b) load-circuit orientation of voltages.

steady-state operating condition, the system must first be turned on, by zero-voltage switching, as a single-phase circuit and then must revert to synchronous three-phase operation. Fig. 399 shows a simplified circuit configuration of a three-phase heater control that employs zero-voltage synchronous switching in the steady-state operating condition, with random starting. In this system, the logic command to turn on the system is given when heat is required, and the command to turn off the system is given when heat is not required. Time proportioning heat control is also possible through the use of logic commands.

The three photocoupled inputs to the three CA3059 circuits change state simultaneously in response to a "logic command". The CA3059 circuits then provide a positive pulse, approximately 100 microseconds in duration, only at a zero-voltage crossing relative to their particular phase. A balanced three-phase sensing circuit is set up with the three CA3059 circuits each connected to a particular phase on their common side (pin 7) and referenced at their high side (pin 5), through the current-limiting resistors $R_4$, $R_5$, and $R_6$, to an established artificial neutral point. This artificial point is necessary because the neutral on the load side is not accessible. Because only one triac is pulsed on at a time, the diodes ($D_1$, $D_2$, and $D_3$) are necessary to trigger the opposite-polarity triac and, in this way, to assure initial latching-on of the system. The three resistors ($R_1$, $R_2$, and $R_3$) are used for current limiting of the gate drive.

In critical applications that require suppression of all generated RFI/EMI, the circuit shown in Fig. 400 may be used. In addition to synchronous operating conditions, this circuit also incorporates a zero-voltage starting circuit. The start-up condition is zero-voltage synchronized to a single-phase, 2-wire, line-to-line circuit, comprised of phases A and B. The logic command engages the single-phase "start-up" CA3059 and three-phase photo-isolators $OCI_3$, $OCI_4$, $OCI_5$ through photo-isolators OCI and $OCI_2$. The single-phase CA-3059 which is synchronized to phase A and B starts the system at zero voltage. As soon as start-up is accomplished, the three photo-isolators $OCI_3$, $OCI_4$, and $OCI_5$ take control, and three-phase synchronization begins. When the "logic command" is turned off, all control is ended, and the triacs automatically turn off when the sine-wave current decreases to zero. Once the first phase turns off, the other two will turn off simultaneously, 90° later, as a single-phase line-to-line circuit, as is apparent from Fig. 398.

## Three-Phase Inductive Load

For inductive loads, zero-voltage turn on generally is not required because the inductive current cannot increase instantaneously; therefore, RFI/EMI generated is usually negligible. Also, because of the lagging nature of the inductive current, the triacs cannot be pulse fired at zero voltage. There are several ways in which the CA3059 may be interfaced to a triac for inductive-load applications. The most direct approach is to use the CA3059 in the dc mode, i.e., to provide a continuous dc output instead of pulses at points of zero-voltage crossing. This mode of operation is accom-

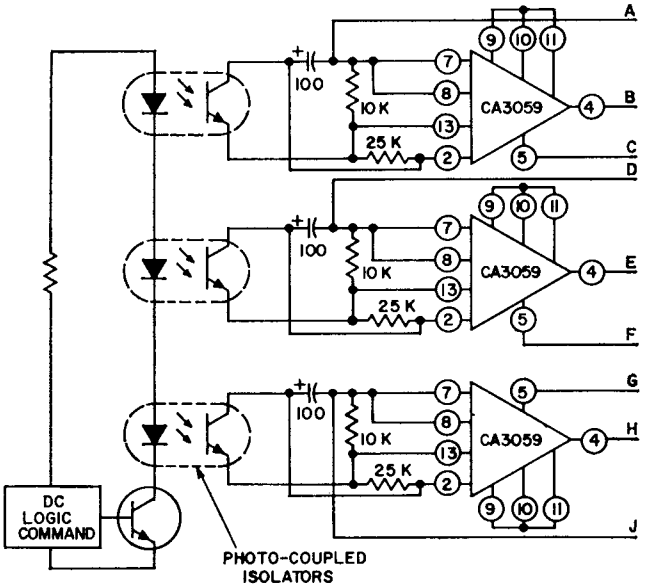



Figure 399. Simplified diagram of a three-phase heater control that employs zero-voltage synchronous switching in the steady-state operating condition (continued on page 367).

plished by connection of terminal 12 to terminal 7, as shown in Fig. 401. The output of the CA3059 should also be limited to approximately 5 milliamperes in the dc mode; and the use of a triac from the RCA 40530 family is recommended for this application. Terminal 3 is connected to terminal 2 to limit the steady-state power dissipation within the CA3059. For most three-phase inductive load applications, the current handling capability of the 40530 triac (2.5 amperes) is not sufficient. Therefore, the 40530 is used as a trigger triac to turn on any other currently available power triac that may be used. The trigger triac is used only to provide trigger pulses to the gate of the power triac (one per half cycle); the power dissipation in this device, therefore, will be minimal.

Simplified circuits using pulse transformers and reed relays will also work quite satisfactorily in this type of application. The RC networks across the three power triacs are used for suppression of the commutating dv/dt when the circuit operates into inductive loads. A detailed explanation of commutating dv/dt is provided in the basic discussion of **Thyristors** in an earlier section of this Handbook.

## HEATING CONTROLS

Thyristors may be used in most heating applications to provide reliable and economic control of the input power supplied to the heating element. Triacs are usually employed in such applications because full-wave control (a 100-per

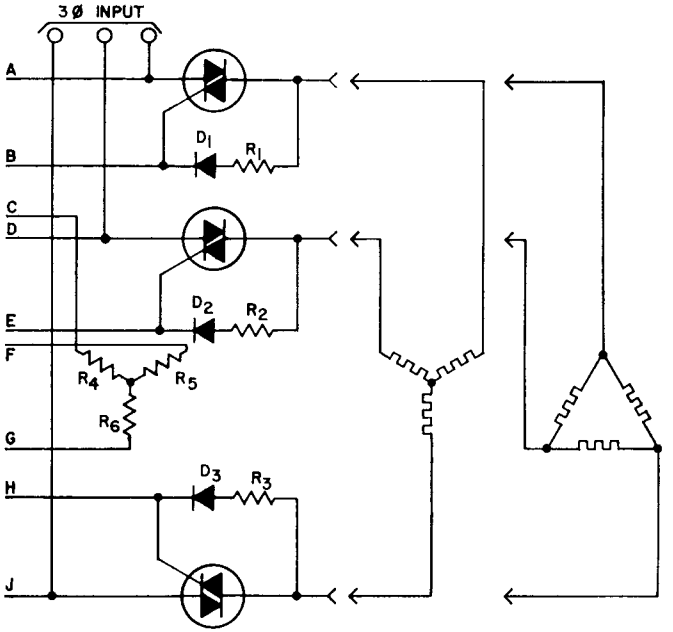



Figure 399. Simplified diagram of a three-phase heater control that employs zero-voltage synchronous switching in the steady-state operating condition (continued from page 366).

cent control range) can be achieved with a single device. Although two back-to-back SCR's may also be used to provide full-wave control, triacs are generally preferred because a single triac is less expensive than two SCR's that operate at the same current performance level. The application of power to the heating element may be controlled by a single SCR for applications in which half-wave control is adequate. An SCR and diode combination is used when only limited control is required and an always-on system is desired.

## General Design Considerations

The temperature of the heat acceptor can be controlled to an accuracy of ±20°C by use of relatively simple open-loop systems, i.e., systems in which no feedback is received from the heat acceptor. A typical example is an oven-heating system in which a variable source of power supplied to the oven may be increased or decreased as desired, but no provisions are made for control of temperature variations that may occur at any power-input level. Such temperature variations may result from changes in the input line voltage or in circuit loading.

Sophisticated closed-loop systems that use low-cost thyristors may be designed for use in applications that require precise temperature control. In such systems, a sensor is used to monitor the temperature of the heat acceptor, and an error (feedback) signal is

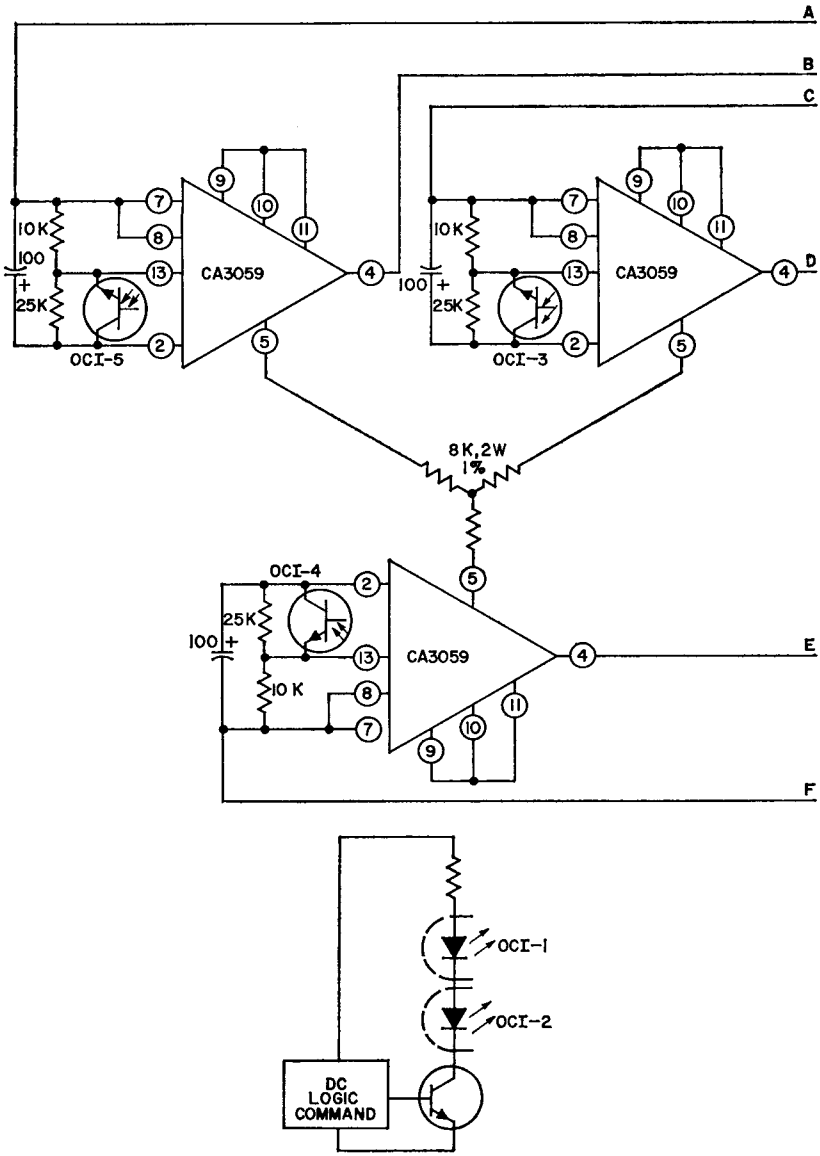



Figure 400. Three-phase power control that employs zero-voltage synchronous switching both for steady-state operation and for starting (continued on page 369).

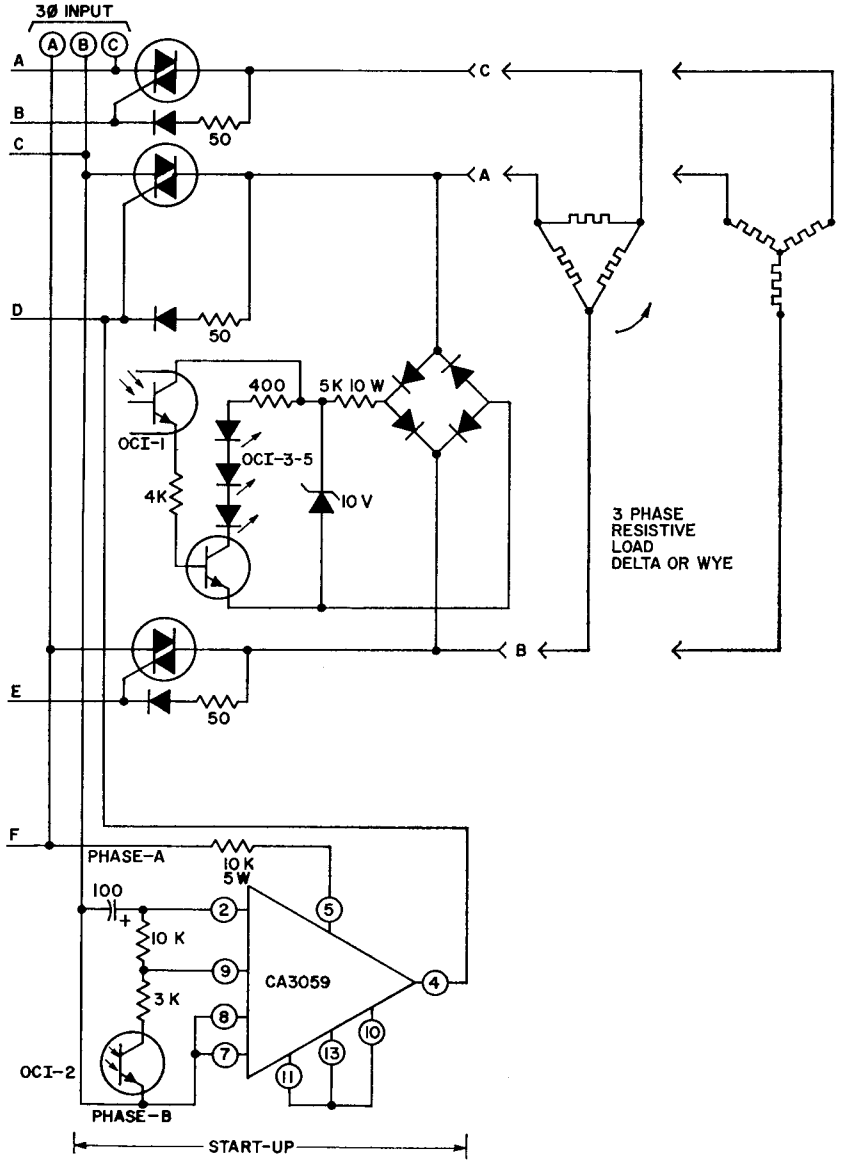



Figure 400. Three-phase power control that employs zero-voltage synchronous switching both for steady-state operation and for starting (continued from page 368).

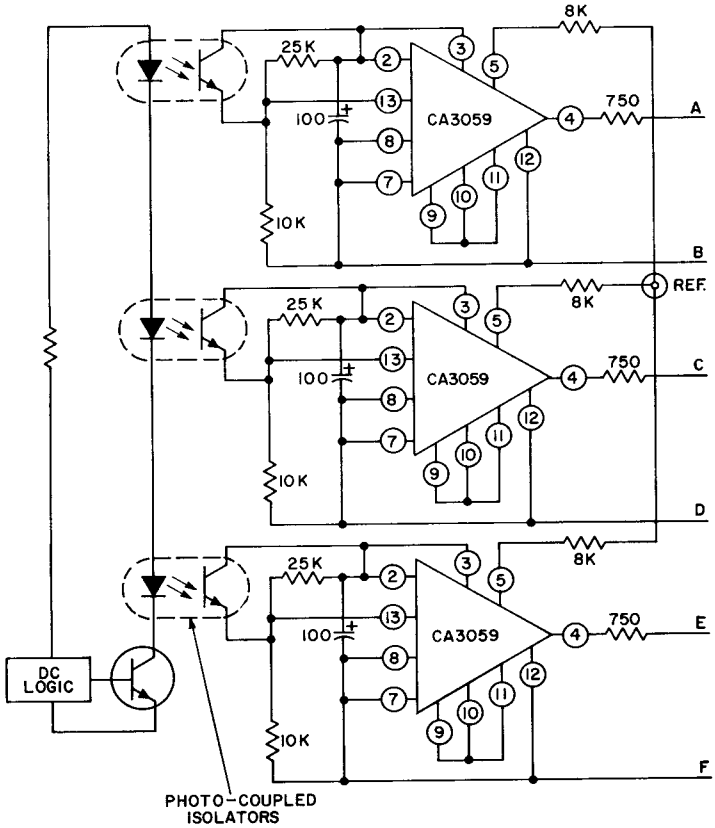



Figure 401. Three-phase triac control circuit for an inductive load, i.e., three-phase motor (continued on page 371).

developed to indicate whether the amount of power delivered to the heating element should be increased or decreased. The system operates in response to the feedback signal to compensate for the effects of any change in line voltage or circuit loading; as a result, variations in the temperature of the heat acceptor can be limited to very small values. With a closed-loop system, a control accuracy within $\pm 0.5°C$ is readily achieved.

An important practical consideration for heating-control systems is that when power is added to or removed from the ac line, the increments of change in power level should not exceed 5 kilowatts. This limitation is necessary to assure that transient effects produced by the power switching do not interfere with other electronic equipment on the same ac line.

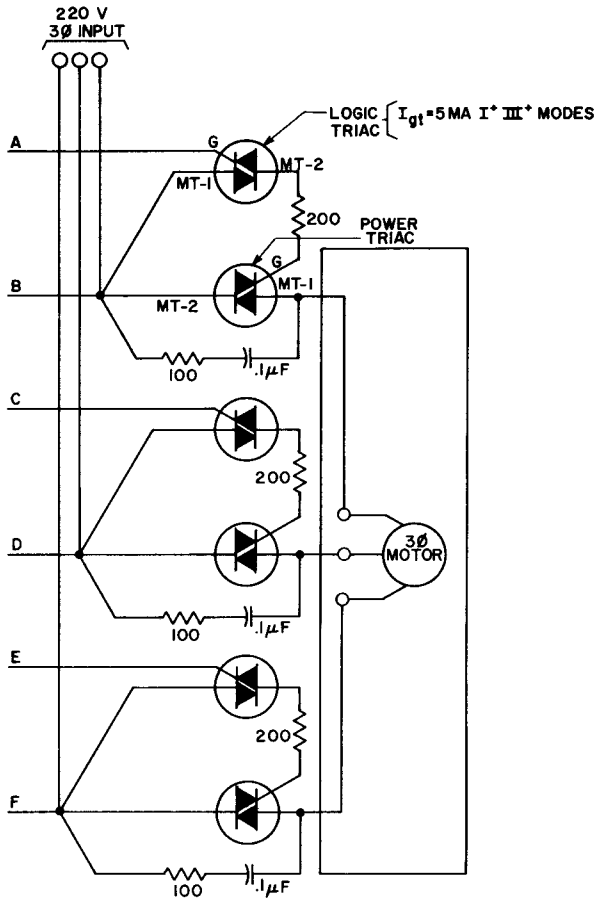



Figure 401. Three-phase triac control circuit for an inductive load, i.e., three-phase motor (continued from page 370).

These increments of power level can be easily controlled by use of digital circuitry and low-level control circuits.

In heating systems that use air as a heat-transfer medium, modulation of the air flow may be necessary to match the amount of heat energy being produced. With this technique, a minimum amount of heat can be provided at a low speed to reduce the temperature difference between the heated area and the stagnant air.

In many applications, isolation of the sensor and associated circuitry from the triac and ac line are desirable. As explained previously in the section on **Isolated Trigger Circuits,** either optical or magnetic techniques may be employed to provide this isolation. In polyphase control circuits this type of isolation is essential.

Typical isolation circuits and techniques used in such applications are described in the section on **Three-Phase Triac Controls.**

## Comparison of Heating-Control Techniques

Manual adjustment of the output of electrical heaters can be provided over an infinite range by use of the phase-control techniques described in the discussion on **Types of Thyristor Turn-on- Circuits.** Automatic temperature regulation may be incorporated into phase-control systems by the addition of a firing circuit that adjusts the conduction angle of the thyristor in response to feedback from a temperature sensor, such as a thermistor.

Fig. 402 shows a simple full-wave heat control circuit that employs the phase-control technique. This circuit illustrates a rather crude, but usable approach to heat control that may be employed for hot plates, solder pots, and other noncritical appliances. The circuit is the basic two-time-constant phase control discussed earlier, in the section on **Triggering Techniques.** The range of the positive-temperature coefficient (PTC) sensor must be sufficient to provide the desired control. Care must be taken to assure that the sensor is not self-heated as a result of the current that flows through it. The triac $Q_1$ provides full-wave control so that the amount of dc component on the ac line is very small.

If the sensor is replaced by a short circuit, an open-loop control is obtained. (The value of resistor $R_2$ must be increased.) Open-loop control can be advantageous for applications in which manual heat control is desired. A typical application is entry-way heating. A constant temperature is not required, but the temperature must be varied to maintain the temperature of the entry area at a level higher than the inside temperature.

The major disadvantage of phase controls for use in heating systems is that random switching of the high power levels usually involved in such applications generate large amounts of radio-frequency interference (RFI) in nearby electrical equipment and

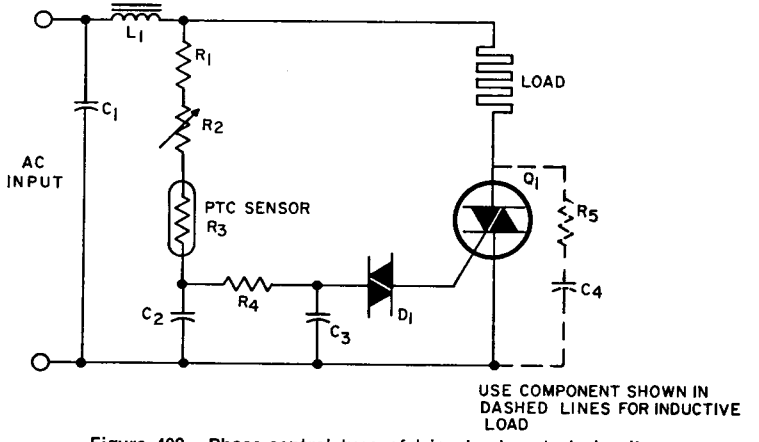



Figure 402. Phase-control type of triac heat-control circuit.

particularly in standard AM broadcast-band receivers. In the phase-control circuit shown in Fig. 402, the capacitor $C_1$ and the inductor $L_1$ form a filter network that is used to limit the RFI produced by the circuit. The value (and size) of inductor $L_1$ must be increased as the load is increased. At higher power levels, this requirement often causes packaging difficulties and substantially increases the over-all cost of the circuit. For this reason, zero-voltage switching is used for most heater controls.

The thermal-response time of a heater is generally much longer than the period of the ac line frequency. Modulation of the heater output, therefore, may be accomplished by application of full power to the load for a short time and complete removal of the power for a period of time. If the control circuit is made sufficiently sensitive to temperature changes, an extremely small temperature differential can be maintained between power-on and power-off states. Precise temperature control is then achieved by rapid and frequent switching from on to off. If this switching always occurs at or near the 0-degree or 180-degree (zero-voltage) points on the ac line voltage, the radio-frequency interference generated by thyristor switching can be almost completely eliminated.

## Zero-Voltage-Switched Triac Heater Controls

The RCA-CA3059 integrated-circuit zero-voltage switch is particularly suited for use in thyristor temperature-control applications. The integrated circuit may be employed as either an on-off type of controller or a proportional controller, depending upon the degree of temperature regulation required. The availability of numerous terminal connections to internal circuit points greatly increases the flexibility of the CA3059 and permits the circuit designer to exercise his creativity to employ the integrated switch in unique ways. A detailed description of the CA3059 and the block diagram (Fig. 393) and schematic (Fig. 394) of this integrated circuit are provided in the section on **Triggering Techniques** given earlier in this Handbook.

**On-off Control**—Fig. 403 shows a simple zero-voltage-switching control that uses a CA3059 and a triac. By selecting the triac for the current level involved, this simple circuit can control heater power at current and voltage levels from 2.5 amperes and 24 volts to 80 amperes and 600 volts. The components that must be changed for operation at different power levels are the series resistor $R_S$ and the triac. The chart shown in Fig. 403 illustrates the easy way in which circuits that use the CA3059 can be changed to accommodate various line voltages.

The circuit shown in Fig. 403 requires a negative-temperature-coefficient (NTC) thermistor for proper operation. If terminals 9 and 13 of the CA3059 are interchanged, positive-temperature-coefficient thermistors can be used, and advantage can then be taken of the shorted- and open-circuit fail-safe provisions of the CA3059.

The heater control operates in the off-on mode. Its temperature accuracy depends on the differential input sensitivity of the CA3059, or the thermistor used, and, to some extent, on the level of the temperature being con-

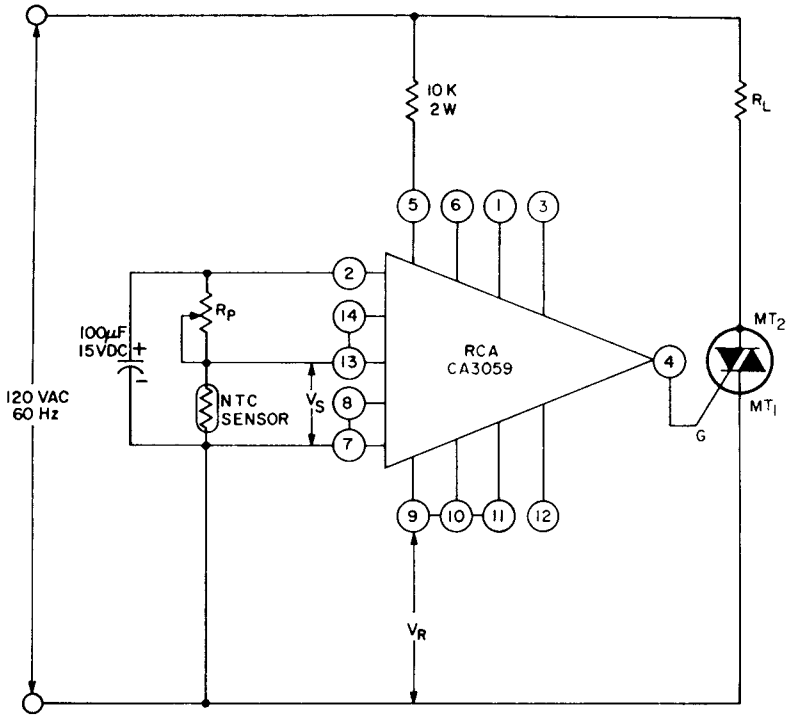



Figure 403. On-off temperature controller.

trolled. This behavior is controlled by the slope of the decrease in thermistor resistance with increasing temperature. These effects must be accommodated during the design of the circuits.

**Proportional Heating Controls** —When precise temperature control is required, closed-loop proportional control is recommended. Fig. 404 shows a typical example of a proportional type of heat control. The temperature accuracy of this circuit, which depends on the input sensitivity of the CA3059, the thermistor used, and the level of temperature being controlled, is in the order of ±0.2°C.

In proportional-control systems, either of two basic approaches may be used depending upon the performance desired. The simpler approach is to use a reference ramp that has an amplitude almost equal to the voltage swing that exists. A more sophisticated approach is to use a limited ramp that covers the set point.

Fig. 405 shows a basic circuit that may be used to perform the comparison between the reference ramp and a voltage divider formed by the sensor and calibration resistor. In the large-ramp approach, the ramp voltage varies from 0 volts to the voltage $V_S$. With this approach, the triac is not gated when the magnitude of the ramp voltage is less than the voltage $V_1$. This mode of operation produces an "on" period that decreases as the sensor temperature increases, as shown in Fig. 406.

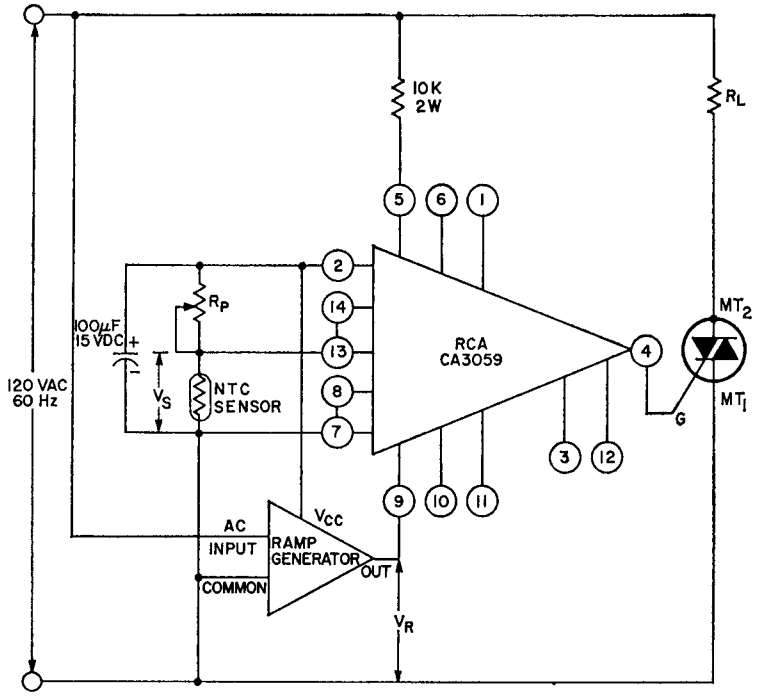



Figure 404. Proportional temperature controller.

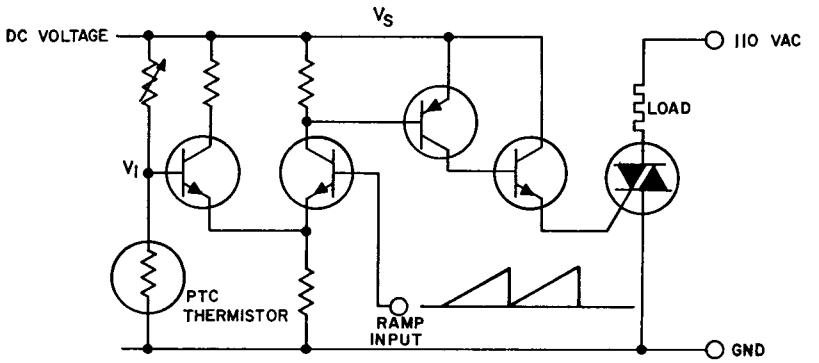



Figure 405. Basic circuit for comparison of ramp voltage and reference voltage in a proportional temperature controller.

Fig. 406 shows that, as the value of the reference voltage $V_1$ varies, the length of time the triac is on changes. The triac is never off all the time. The increments of power are based on the number of line voltage half-cycles that occur during the time-base period. A half-second time base allows 60 half-cycle periods of operation. The

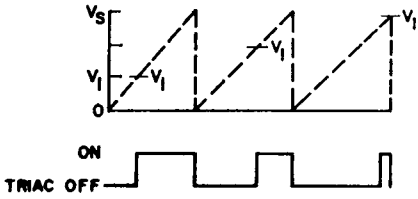



Figure 406. Waveform showing effect of the relative amplitude of the ramp voltage and the reference voltage on the conduction time of the triac.

heat can then be changed in increments of (1/60)100 or 1.6 per cent; if finer control is desired, the time base must be lengthened. The maximum amount of heating power that can be controlled is easily changed by selection of different triac and heater combinations. The increments of heat produced by this type of circuit are a function of the time base, the heater, and the differential-amplifier resolution.

In the more sophisticated limited-ramp approach, the ramp voltage traverses only small increments. In this case, the ramp voltage may vary from 0.4Vs to 0.6Vs. The heater receives full power until the voltage $V_1$ rises to 0.4Vs. At this point, the power supplied to the heater starts to decrease, and the heater is totally off when the voltage $V_1$ reaches 0.6Vs. This type of operation provides minimum warm-up time and the fastest response to loading. It is particularly useful in temperature-controlled processes. Fig. 407 compares the two approaches to proportional control with the same heaters and loads. Fig. 408 shows circuit diagrams of ramp generators for both approaches. The benefits of the sophisticated system may not be needed in all applications, and care must be exercised to assure the added cost is warranted. A complete heater control that uses the sophisticated type ramp generator is shown in Fig. 409.

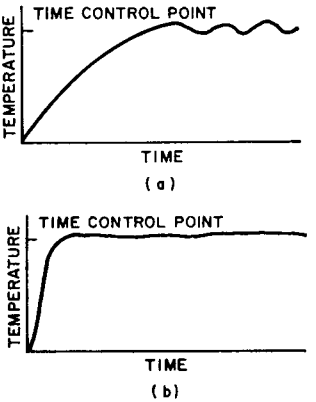



Figure 407. Response characteristics for (a) elementary and (b) sophisticated proportional heat-control systems.

**Heat Controls With Isolated Sensors**—In many industrial controls, isolation of the sensor from the thyristor and ac line is required. The main premise is to provide adequate operation and at the same time isolate the input from the power lines. Three schemes for achieving this isolation are shown in Figs. 410, 411, and 412.

Fig. 410 shows an isolation system that uses a low-voltage transformer and pulse transformer to provide sensor isolation and low-voltage operation of the CA3059. This circuit is a typical example

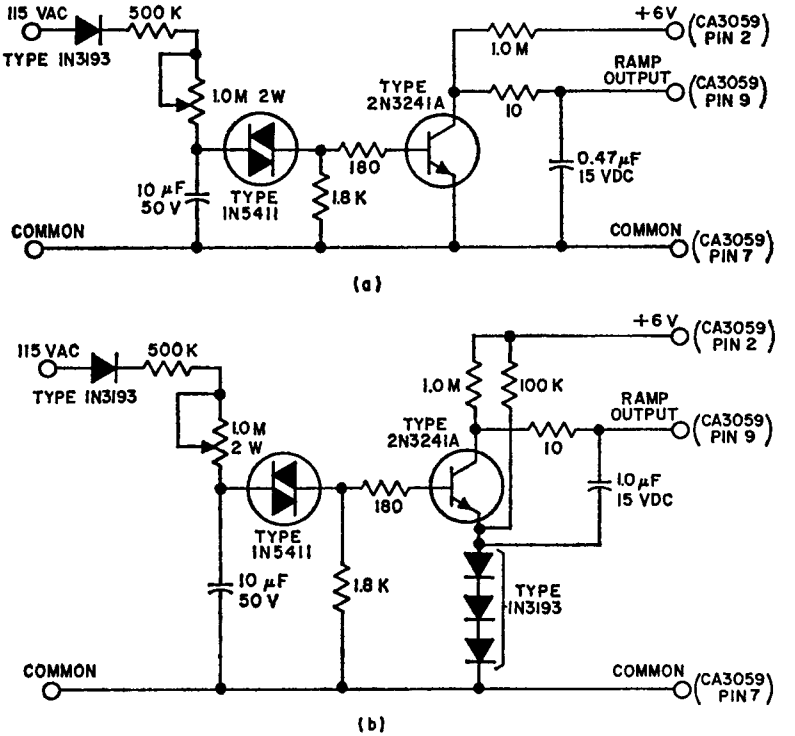



Figure 408. Ramp generators: (a) ramp amplitude from 0 to 6 volts; (b) ramp amplitude limited to 1.5 volts to 6 volts.

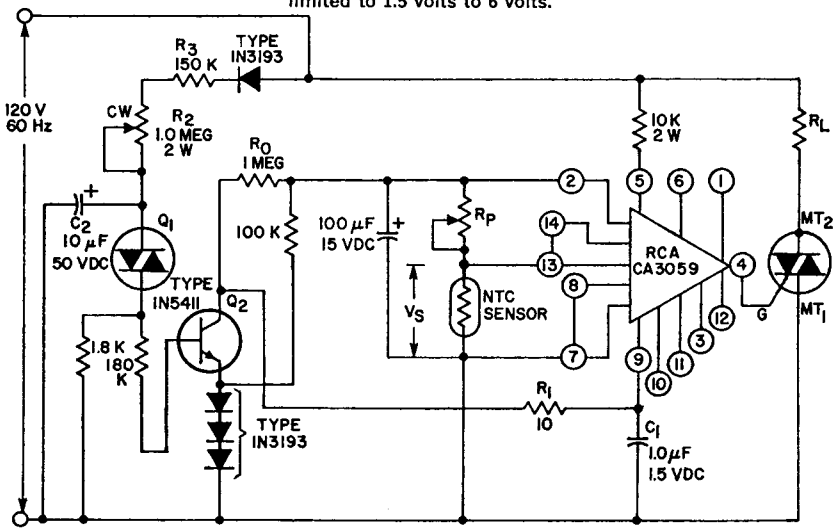



Figure 409. Proportional heat-control system.

of a heat-control circuit used in industrial applications.

Fig. 411 illustrates the use of an inverter stage for input isolation. The circuit shown in Fig. 411(a) uses the inverter isolation method to provide an on-off type of control in response to a dc level. This circuit can be used as a high-power switch in a numerical control system. In effect, it is simply a form of a solid-state ac relay and may be used for many other types of applications in addition to that of heating control.

with polyphase systems to provide isolation between phases. This isolation is accomplished most economically by use of multiple secondary windings. As the output of the inverter is coupled to the differential input of the CA3059, very little output current is required and power consumed by the inverter is basically its inherent losses.

Fig. 412 shows a system in which optical coupling is used to provide input isolation. Incandescent lamps, neon lamps, or light-emitting diodes may be used.

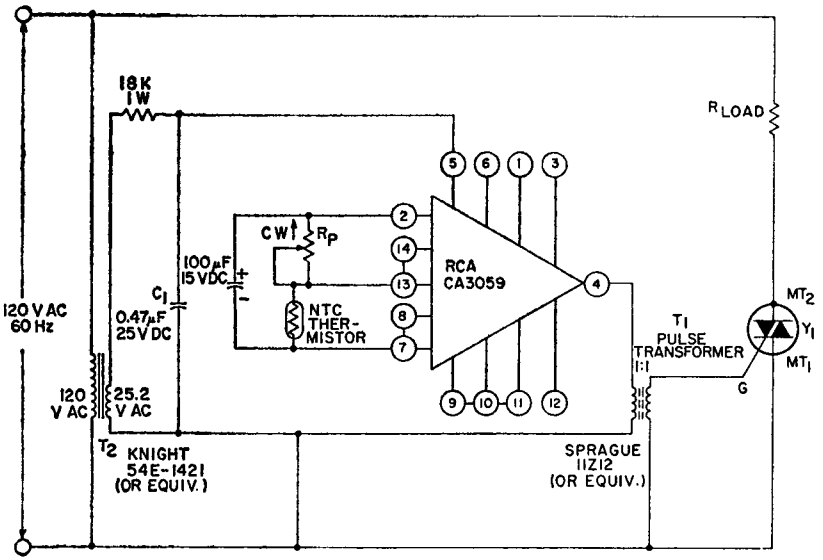



Figure 410. On-off controller that uses a step-down transformer to provide isolation of the sensor.

Fig. 411(b) shows the use of an inverter as an interface between a complex control system and the CA3059-triac combination. This type of circuit is usually employed for industrial control circuits in which the sensor is coupled to low-level logic systems. A control circuit of this type may also be used

Light-emitting-diode systems are preferred because they are rugged and have long life capability. The advantage of this system is that the input from the sensor to the differential amplifier can be proportional.

In proportional controls for polyphase systems, it is better for

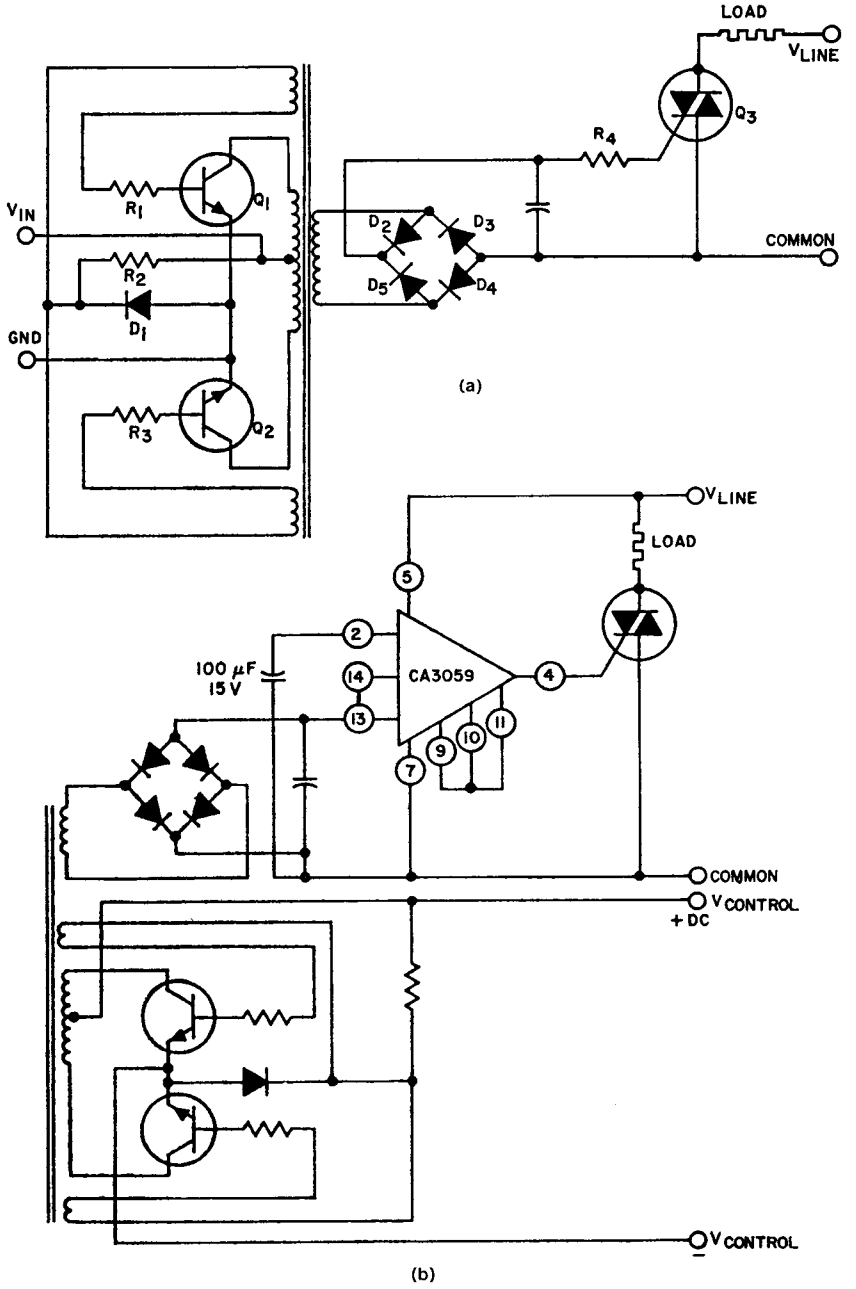



Figure 411. Proportional heat controls that use an inverter stage to provide input isolation.

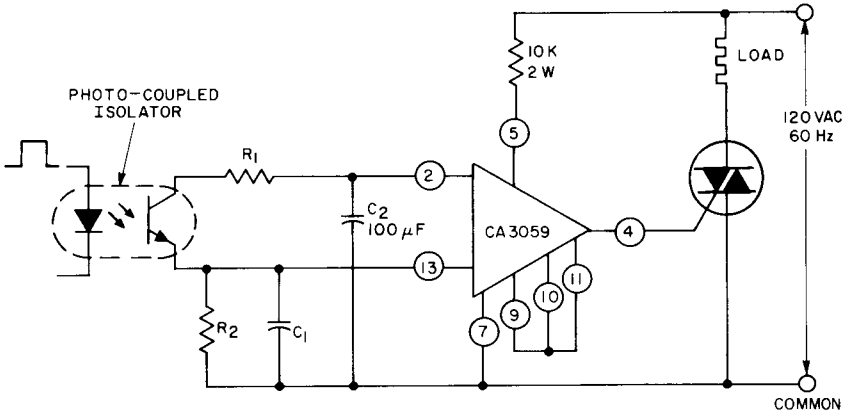



Figure 412. Proportional control system that uses optical coupling to provide input isolation.

the proportioning to be accomplished at the sensor and to supply a controlled signal to the optical elements. This method is preferred for operation with a magnetic isolation system. The alternate method is to use several ramp generators on each phase and hope the time bases are the same. The disadvantage of this latter method is that it permits chance and tolerances to cause poor operation.

In high-power systems (5 KW or higher), it is advisable to incorporate a stepped load system in addition to a proportionally controlled element. A block diagram of such a system is shown in Fig. 413.

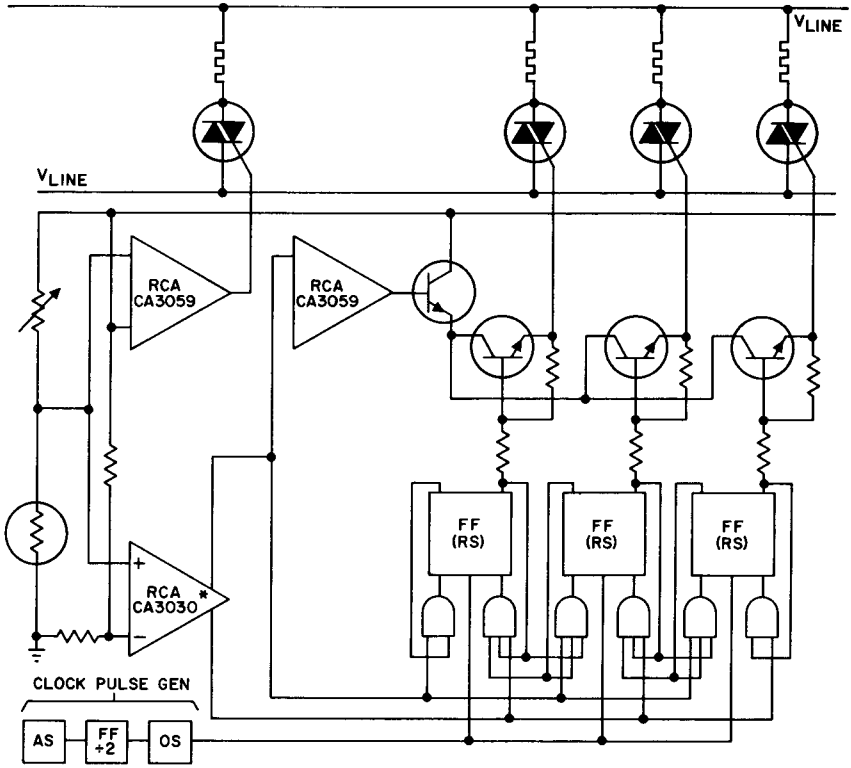
## INCANDESCENT LIGHTING CONTROLS

A popular application of thyristors, particularly of triacs, is in controls for incandescent lighting systems, such as lamp dimmers, traffic-signal lights, and warning flasher lights. In addition, triacs may be used to advantage in bulb preheat circuits that are employed to maintain the temperature of lamp filaments at a point just below that required for incandescence. Use of bulb preheat circuits makes it possible to avoid high initial current surges that result when full turn-on power is first applied to an incandescent lamp because of the low resistance of the cold filament.

### Surge-Current Considerations

An important consideration in the design of triac control circuits for incandescent lighting systems is the load and its effect on the requirements of the triac. Obviously, the triac must be capable of handling the steady-state load current. In addition, however, the triac should be capable of withstanding transient current surges that may result from bulb flashover or cold-filament inrush current.

**Flashover**—A short-duration, extremely high-current surge through the triac is initiated when a lamp filament ruptures. The rupture is most likely to occur as a result of a termination in bulb life; however it can be caused by




* For information on RCA3030 integrated-circuit operational amplifier the reader should refer to the *RCA Linear Integrated Circuits* Manual, Technical Series IC-42.

Figure 413. Non-isolated staged heat control.

a mechanical shock. The mechanism of flashover is initiated by the gap formed when rupturing occurs. The instantaneous value of line voltage across the break sets up an electric field that ionizes the gases in close proximity to the gap. The ionized gases, usually argon and nitrogen, provide an electrical conduction path across the gap, and the resulting current heats and ionizes more gases until an arc is formed across the filament lead-in wires. The arc is maintained as long as the regenerative heating and ionization continue. Finally, because of either increasing arc length or decreasing ac line voltage, or both, the electric field becomes too weak to sustain the arc, and the arc is extinguished.

Fig. 414 shows a flashover current pulse. Its magnitude and duration depend on many factors. The actual peak magnitude of the source voltage, the voltage phase at the instant of filament rupture, and the impedance of the lead wires and other circuitry (including RFI filters) all affect the duration and magnitude of the surge. Typical values can be given for the stress of flashover at a load center point. For bulbs of less than 75 watts, the duration of the surge can be typically less than 2 milliseconds. For bulbs of 100 to

150 watts, the duration of the surge can be typically less than 4 milliseconds. The magnitude of surge can vary considerably, with typical peak values ranging from 80 to 200 amperes when the flashover occurs near the maximum voltage point. If the flashover occurs at a zero-voltage crossing, the current surge may be reduced as a result of the dependence of the magnitude on the voltage phase at rupture.

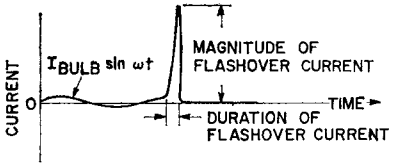



Figure 414. Flashover current at peak voltage point.

Because of the short duration of the flashover current, it is usually difficult to provide circuit fuse protection against flashover. Most incandescent bulbs are provided with a fuse built into one of the lead-in wires. This built-in fuse is not 100-per-cent effective against flashover and, therefore, cannot be depended upon to protect the triac.

**Inrush Current**—In tungsten-filament lamps, the cold filament resistance is approximately 1/18 to 1/12 of the hot filament resistance. The actual currents in a circuit under inrush and steady-state conditions do not vary in these ratios, however, because of the inductance and external limiting resistance of the circuitry, including the lead-in wires to the bulb. Furthermore, it is obvious that the highest inrush current will occur at the peak of the voltage sine wave in a lamp load circuit. If switching occurs at any other phase of the voltage sine wave, the peak current through the bulb is less than "worst case." Typically, the maximum inrush peak current can be ten times as great as the steady-state peak current, while the peak inrush current with zero-voltage switching can be approximately five times as great as the steady-state peak current, as shown in Fig. 415. Thus zero-voltage switching of a lamp effects a soft turn-on that reduces the initial peak of inrush current by half and greatly increases bulb life. This increase of bulb life by zero-voltage switching has been verified by test results; an increase in life of approximately ten times, with a 90 per cent confidence level, has been reported. Thus, maintenance costs are reduced and system reliability increased.

Fig. 415 shows how the current in a lamp circuit decreases to the steady-state value. The rate of decrease depends upon the thermal time constant of the tungsten filament. A 100-watt bulb typically

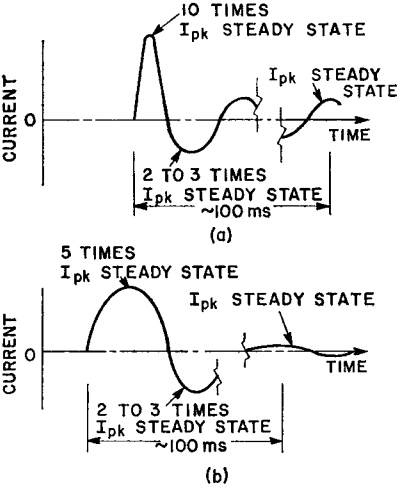



Figure 415. (a) Inrush current at peak voltage point and (b) inrush current at zero-voltage point.

might reach steady-state current within 100 milliseconds after turn-on, while a 1000-watt bulb typically requires 200 miliseconds to reach its steady-state current condition.

Flashover and inrush can occur in combination. Because a bulb is exposed to its most severe normal operating stress during inrush, the weakest spot of the filament often ruptures and causes a flashover at turn-on. Most often, switching and flashover occur at some point other than the peak voltage; therefore the resulting peak current is usually within the handling capability of the triac.

Fuses in incandescent-lamp circuits must not blow under the stress of inrush current, yet must blow under flashover current. For low-power bulbs, the flashover current is substantially greater than the peak inrush current, and fuse protection is simple. For example, a 100-watt bulb might have a typical flashover current of 100 to 200 amperes and a typical inrush current of 10 amperes. For large-wattage bulbs, however, fusing is difficult. For a 1000-watt bulb, the peak flashover current might still be between 100 and 200 amperes, while the peak inrush current is approximately 120 amperes. Fuses set to blow at 150 amperes peak flashover current of short duration may also blow under the long-duration, slightly-lower-amplitude stress of inrush. As a result, a fusing solution to the problem of triac protection would be marginally reliable. One solution is to use a 40-ampere triac (available in the RCA-2N5443 series), which has a single-cycle surge capability of 300 amperes, to control this 10-ampere load. Here again, system reliability would be improved, and maintenance costs reduced.

## Filament Preheat Circuit

Fig. 416 shows a filament preheat circuit that may be used to reduce the initial inrush current of an incandescent lamp. In this circuit, when transistor $Q_1$ is off, the logic interfacing triac $T_1$ is also off. Resistor $R_3$, which can be a fixed resistor of approximately 98 kilohms, is set so that triac $T_2$ is fired for only a small portion of the voltage cycle. This firing is accomplished by the standard double-time-constant gate circuitry of triac $T_2$. The low-conduction-phase

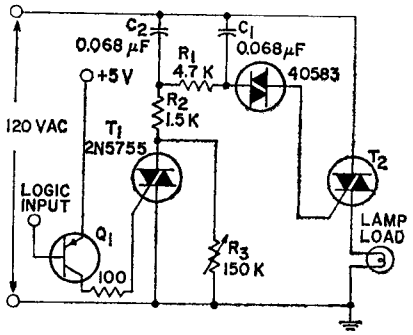



Figure 416. A circuit including a filament preheat arrangement.

firing of the bulb keeps the tungsten filament warm but not hot enough to radiate any readily visible light. When transistor $Q_1$ is turned on, triac $T_1$ is gated on and resistor $R_3$ is shorted, and the lamp load turns on.

The associated waveforms are shown in Fig. 417. For a 200-watt bulb in the circuit of Fig. 416, the first peak of current through the bulb is 7.5 amperes when the warm-up circuit is used and 25 amperes with cold-filament inrush.

## Lamp Dimmers

The light intensity of an incandescent lamp depends upon the

magnitude of the voltage impressed upon the lamp filament. Changes in the lamp voltage, therefore, vary the brightness of the lamp. When ac source voltages are used, a triac can be employed in series with an incandescent lamp to vary the voltage to the lamp by changing the conduction angle, i.e., the portion of each half-cycle of ac line voltage during which the triac conducts to provide power to the lamp filaments.

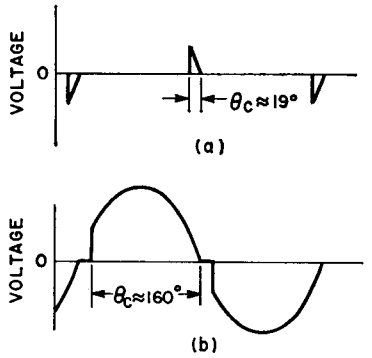



Figure 417. Waveforms for circuit in Fig. 416: (a) voltage on bulb when Q1 is off; (b) voltage on bulb when Q1 is on.

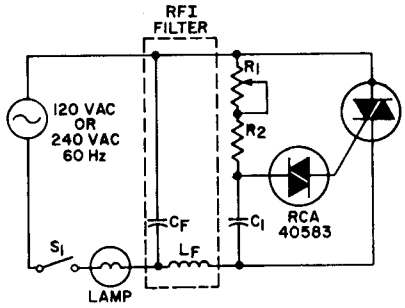
Phase-control switching of a triac is a very effective and widely used control technique for lamp-dimmer applications.

In 400-Hz (or higher-frequency) power systems, the thermal time constant associated with the lamp filament permits the intensity of an incandescent lamp to be varied (without noticeable flicker) by use of proportional integral-cycle triac controls to change the average power applied the lamp. This type of control may be desirable in high-power 400-Hz systems to reduce RFI components and inrush current surges.

**Single-Time-Constant Phase Control**—Fig. 418 shows the circuit diagram for a single-time-constant phase-control circuit. This circuit, which was described previously in the section on **Triggering Techniques**, represents the most economical type of lamp dimmer. However, it provides a high initial illumination level and exhibits hysteresis in the control-potentiometer setting at low light levels.

**Initial brightness:** The dimmer control turns on the incandescent lamp with an appreciable initial brilliance for the following reason. The first instantaneous discharge of $C_1$ reduces the voltage across it by some amount $\Delta V$. This voltage reduction is caused by the instantaneous discharge of $C_1$ through the trigger device and the gate circuit of the triac. This discharge causes triggering at earlier phase positions on succeeding half-cycles and, because the capacitor charges from a lower potential of opposite polarity, results in a "quick-turn-on" effect which produces fairly high levels of initial illumination. Continued rotation of the potentiometer shaft increases the voltage across the load until the points of maximum load voltage and lamp illumination are reached.

Rotation of the potentiometer shaft in the opposite direction increases the resistance value of $R_1$ and causes the phase position of the triac triggering to be increasingly delayed from the line-voltage crossover point. This delayed triggering reduces the effective load voltage gradually until a point is reached at which a small increment of additional resistance causes the triac to stop conducting. At this point, all voltage is removed from the load and the lamp is turned off. The value of resistance, $R_{TO}$, required to turn off the triac completely is greater than the value $R_{IC}$ required for the triac




| AC INPUT VOLTAGE | TRIGGER-CIRCUIT TIME-CONSTANT NETWORK | | | RFI FILTER | | RCA TRIAC TYPES |
|---|---|---|---|---|---|---|
| | $C_1$ | $R_1$ | $R_2$ | $L_F$ (typical) | $C_F$ (typical) | |
| 120 V 60 Hz | 0.1 μF 200 V | 100 K ½ W | 1 K ½ W | 100 μH | 0.1 μF 200 V | 40485 40509 40638 |
| 240 V 50/60 Hz | 0.05 μF 400 V | 200 K ½ W | 7.5 K 2 W | 100 μH | 0.1 μF 400 V | 40486 40510 40639 |

Figure 418. Single-time-constant phase-controlled lamp-dimmer circuit.

to conduct initially. Therefore, $R_{TO} > R_{IC}$ for a single-time-constant circuit configuration. This difference in potentiometer settings for turn-on and turn-off is commonly referred to as **hysteresis.** It should be noted that hysteresis and "quick-turn-on" are characteristics of the single-time-constant circuit.

**Hysteresis effect:** Fig. 419 shows the interaction between the RC network and the diac to produce the hysteresis effect. The capacitor voltage and the ac line voltage are shown as solid lines. As the resistance in the circuit is decreased from its maximum value, the capacitor voltage reaches a value which fires the diac. This point is designated A on the capacitor-voltage waveshape. When the diac fires, the capacitor discharges and triggers the triac at an initial conduction angle $\theta_1$. During the forming of the gate trigger pulse, the capacitor voltage drops suddenly. The charge on the capacitor is smaller than when the diac did not conduct. As a result of the different voltage conditions on the capacitor, the breakover voltage of the diac is reached earlier in the next half-cycle. This point is labeled B on the capacitor-voltage waveform. The conduction angle $\theta_2$ corresponding to point B is greater than $\theta_1$. All succeeding conduction angles are equal to $\theta_2$

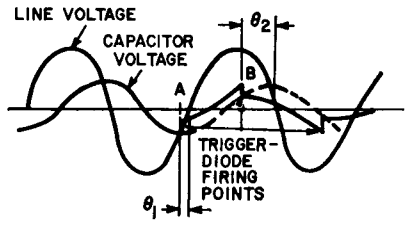
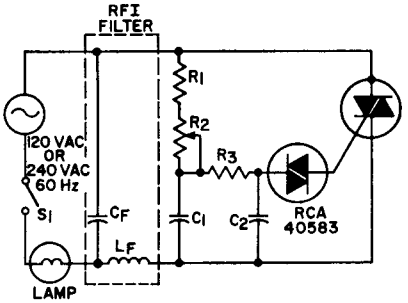



Figure 419. Waveforms showing interaction of control network and trigger diode.

in magnitude. When the circuit resistance is increased by a change in the potentiometer setting, the triac is still triggered, but at a smaller conduction angle. Eventually, the resistance in series with the capacitance becomes so great that the voltage on the capacitor does not reach the breakover voltage of the diac. The circuit then turns off and does not turn on until the circuit resistance is again reduced to allow the diac to be fired. The hysteresis effect makes the voltage load appear much greater than would normally be expected when the circuit is initially turned on.

**Trigger-device characteristics:** Other important factors that influence dimmer performance are the characteristics of the trigger device used in a given circuit configuration. In Fig. 378, the voltage-current characteristics for a triggering device were given, and several important parameters of the device were shown, including $V_P$, $I_{BL}$, and the negative-resistance characteristic $R_N$. Devices that have high values of $V_P$ and negative resistance produce large hysteresis and high values of $V_{IL}$. The value of $V_P$ is fundamental in determination of the value of $V_{IL}$, and the magnitude of the negative resistance $R_N$ is the primary factor in determination of the hysteresis.

**Double-Time-Constant Phase-Control**—The hysteresis effect can be substantially reduced by use of a double-time-constant lamp-dimmer control. The two basic forms of the double-time-constant circuit are shown in Figs. 420 and 421. Both these circuits produce less hysteresis and lower




| AC INPUT VOLTAGE | TRIGGER-CIRCUIT TIME-CONSTANT NETWORK | | | | | RFI FILTER | | RCA TYPES |
|---|---|---|---|---|---|---|---|---|
| | $C_1$ | $C_2$ | $R_1$ | $R_2$ | $R_3$ | $L_F$ (typical) | $C_F$ (typical) | |
| 120 V | 0.1 μF | 0.1 μF | 2.2 K | 100 K | 15 K | | 0.1 μF | 40485 40509 |
| 60 Hz | 200 V | 100 V | ½ W | ½ W | ½ W | 100 μH | 200 V | 40638 40668 |
| 240 V | 0.1 μF | 0.1 μF | 3.3 K | 250 K | 15 K | 200 μH | 0.1 μF | 40486 40510 |
| 50/60 Hz | 400 V | 100 V | ½ W | 1 W | ½ W | | 400 V | 40639 40669 |

Figure 420. Basic double-time constant lamp-dimmer circuit in which triac is cut off when potentiometer is set for maximum resistance.

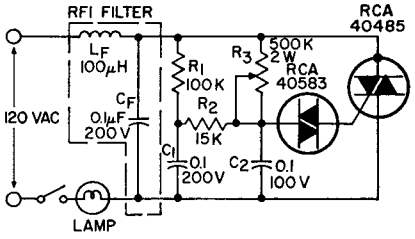



Figure 421. Double-time-constant lamp-dimmer circuit in which a small amount of triac current flows when potentiometer is set for maximum resistance.

initial load voltage $V_{IL}$ than the single-time-constant circuit because the triggering capacitor $C_2$ is recharged by capacitor $C_1$ after every trigger pulse. This recharge is shown in Fig. 422. The recharge from $C_1$ builds up the voltage on $C_2$ to a value slightly less than the breakdown voltage of the triggering device. This action results in relatively constant positive and negative reference voltages from which the capacitor $C_2$ is then subsequently charged to the next trigger-device breakover voltage. This voltage reduces the transient

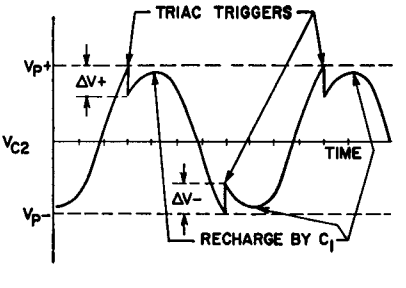



Figure 422. Voltage waveform across triggering capacitor $C_2$ in double-time-constant circuits shown in Figs. 420 and 421.

phase-back of triggering during the first few half-cycles of conduction and allows the lamp load to be initially energized with low effective voltage. The low effective voltage results in low levels of illumination. It is then necessary to rotate the potentiometer only a small amount for the triac to stop conducting and the lamp to be completely dark. Thus, the hysteresis effect is greatly reduced. A fundamental characteristic of the circuit of Fig. 422 is that the component values are chosen so that the triac is conducting a small amount of current when $R_3$ is at its maximum resistance setting. This current prevents the triac from going completely out of conduction and thus prevents hysteresis and quick-turn-on from occurring.

**Effect of series gate resistor:** In the lamp-dimmer circuit shown in Fig. 423, a resistor $R_3$ is connected in series with the triggering device to limit the magnitude of discharge current from capacitor $C_2$ and thus to reduce the instantaneous voltage drop across $C_2$. This technique results in a further reduction of hysteresis and quick-turn-on.

**Flash at turn off:** The performance characteristics of light dimmers which are influenced by the circuit configuration include hysteresis; the voltage across the load at initial turn-on, $V_{IL}$; the maximum voltage that can be developed

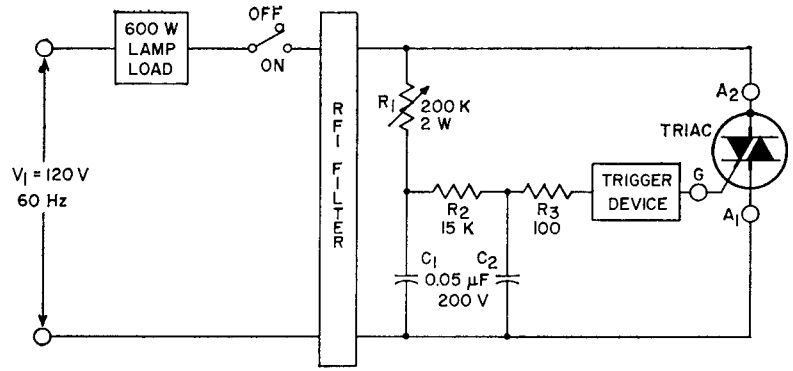



Figure 423. Double-time-constant circuit that uses a series gate resistor.

across the load, $V_L$(max); and a phenomenon called "flash at turn-off." All of these characteristics were considered previously except flash at turn-off. The flash at turn-off is produced in double-time-constant circuits when the potentiometer is adjusted to turn off the triac. This effect can also be achieved by a reduction in the magnitude of the line voltage applied to the circuit when the potentiometer is set at, or near, maximum resistance. A lower line voltage causes a shift in the phase of gate-triggering voltage to a point where the flashing condition occurs. The fundamental cause of this condition is a phase shift in the triggering voltage beyond the zero crossover point of the line voltage into the early portion of the next successive half-cycle, as shown in Fig. 424. The effective voltage on the load undergoes a transient change that lasts for a few cycles and results in the presence of an appreciable voltage across the load during the transient condition. This transient voltage causes the lamp filament to become brightly illuminated for a number of milliseconds and is manifested as a bright flash as the illumination of the lamp is being gradually reduced.

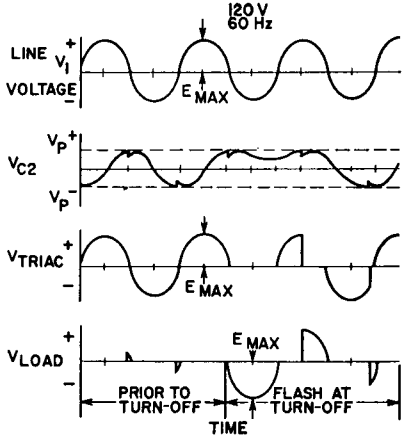



Figure 424. Waveforms during flash-at-turn-off condition in a double-time-constant phase-control circuit.

The double-time-constant circuit is capable of producing a phase shift of the triggering voltage which is greater than 90 degrees. If the circuit components are chosen to produce this result and, simultaneously, to maintain the voltage magnitude across the triggering capacitor above the trigger-device breakover voltage, then a flash at turn-off occurs. However, if the circuit components are

selected to limit the maximum amplitude of the triggering voltage at the 90-degree phase-delay condition below this critical value, there is no flash at turn-off. It should be noted that a trade-off exists between the flash-at-turn-off phenomenon and the hysteresis and $V_{IL}$ phenomena; i.e., elimination of the flash at turn-off produces slightly greater hysteresis and larger values of $V_{IL}$.

**Zero-Voltage-Switched Dimmer Circuit**—Fig. 425 shows a lamp-dimmer circuit in which the use of an RCA-CA3059 integrated-circuit zero-voltage switch in conjunction with a 400-Hz triac results in minimum RFI. (The CA3059 is described in the section on **Triggering Techniques**.)

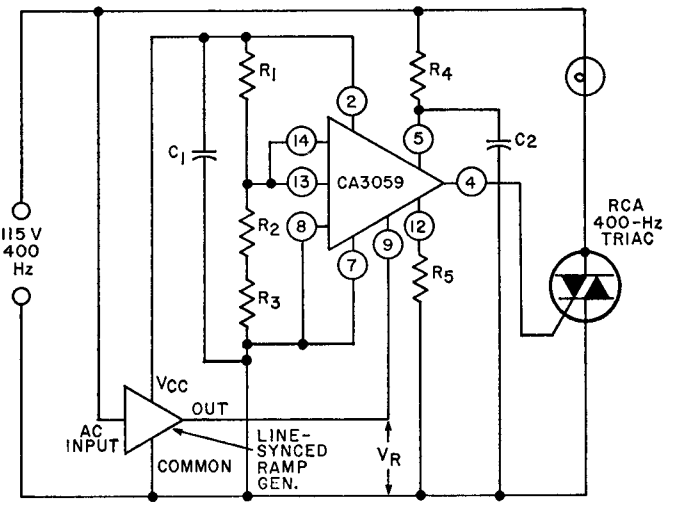



Figure 425. Circuit diagram for 400-Hz zero-voltage-switched lamp dimmer.

Lamp dimming is a simple triac application that demonstrates an advantage of 400-Hz power over 60 Hz. Fig. 426 shows a means of controlling power to the lamp by the zero-voltage-switching technique. Use of 400-Hz power makes possible the elimination of complete or half cycles within a period (typically 17.5 milliseconds) without noticeable flicker. Fourteen different levels of lamp intensity can be obtained in this manner. In the circuit shown in Fig. 426, a line-synced ramp is set up with the desired period and applied to terminal No. 9 of the differential amplifier within the CA3059. The other side of the differential amplifier (terminal No. 13) uses a variable reference level, set by the potentiometer $R_2$. A change of the potentiometer setting changes the lamp intensity.

In 400-Hz applications, it may be necessary to widen and shift the CA3059 output pulse (which is typically 12 microseconds wide and centered on zero voltage crossing) to assure that sufficient latching current is available. The resistor

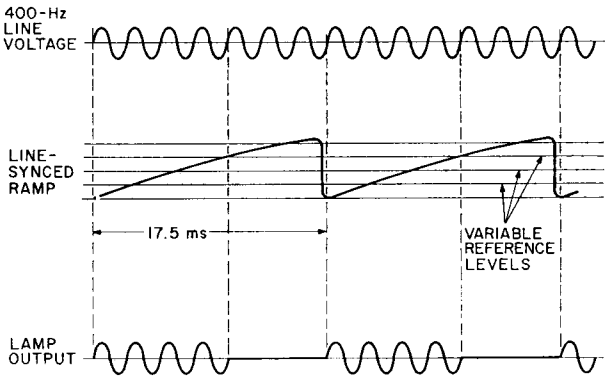



Figure 426. Waveforms for 400-Hz zero-voltage-switched lamp dimmer.

$R_5$ (terminal No. 12 to common) and the capacitor $C_2$ (terminal No. 5 to common) are used for this adjustment.

## Photocell-Operated On-Off Lamp Controls

Photocells can be used in conjunction with light-dimmer circuits to provide light-operated controls. These controls can be designed so that the lamp turns on or off as the ambient light level changes from dark to light, or vice versa. Two photocell control circuits are shown in Figs. 427 and 428.

The circuit shown in Fig. 427 causes the lamp load to turn on gradually as the light impinging on the photocell increases in intensity. As the ambient light intensity increases, the photocell resistance decreases to produce a higher effective voltage across the load.

The circuit shown in Fig. 428 causes the lamp load to turn off as the ambient light level in-

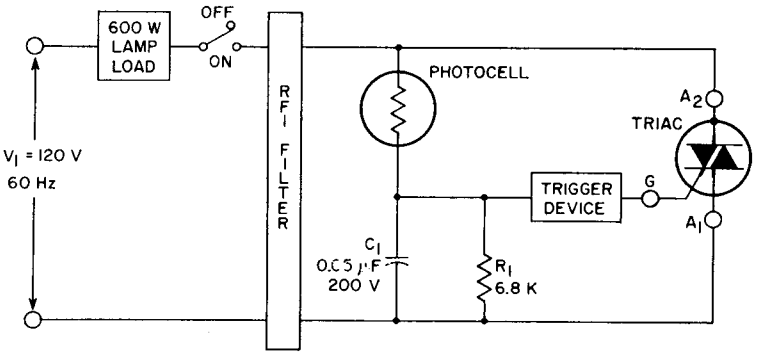



Figure 427. Photocell-operated on-off lamp control that energizes lamp load when photocell is illuminated.

creases. This behavior is caused by a decrease in photocell resistance as the ambient light intensity increases. The decreasing photocell resistance reduces the voltage across capacitor $C_1$ to values less than the breakover voltage of the triggering device and prevents the triac from being triggered. Circuits of this type are useful as outdoor lighting controls because they can be designed to turn on at dusk and turn off at dawn automatically.

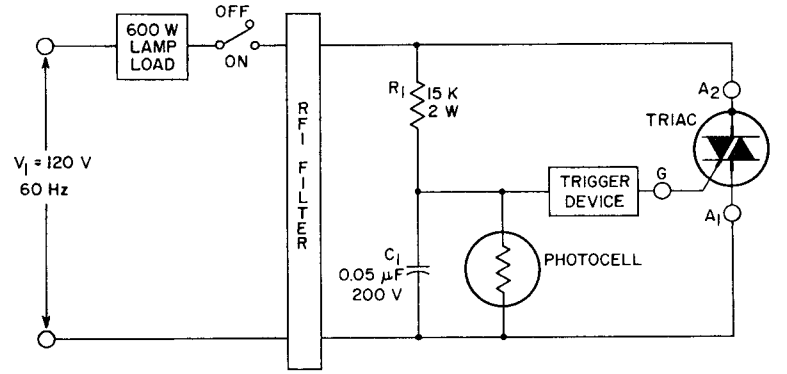



Figure 428. Photocell-operated lamp control circuit that energizes lamp load when photocell is not illuminated.

## Traffic-Signal-Lamp Controls

Triacs are ideally suited for use in the control of traffic signals. These devices can carry the high electrical power required for incandescent traffic-light bulbs and can be gated by the low-power signals from electronic control timers or monitoring computers.

The schematic diagram and operating voltage and current waveforms for two circuits that can be used in traffic-control applications are shown in Figs. 429 and 430 and Figs. 431 and 432. These circuits have the advantages of a common ground between logic and power circuitry, grounded bulbs, and isolation between the dc logic and the power circuitry afforded by use of the interfacing logic triacs.

In the positive-logic switching circuit shown in Fig. 429, logic triac $T_1$ is used to interface between the low-level logic and the load triac $T_2$. With triac $T_1$ gated on, capacitor $C_1$ is charged through resistor $R_1$ to the breakover voltage of the diac, at which point triac $T_2$ and the load are triggered on. As Fig. 430(d) shows, there is continuous gate power driving triac $T_2$ whenever triac $T_1$ is on, and thus the light is on hard.

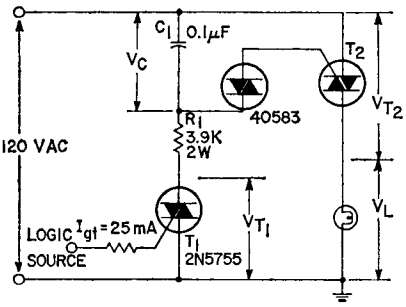



Figure 429. Positive-logic bulb-switching circuit.

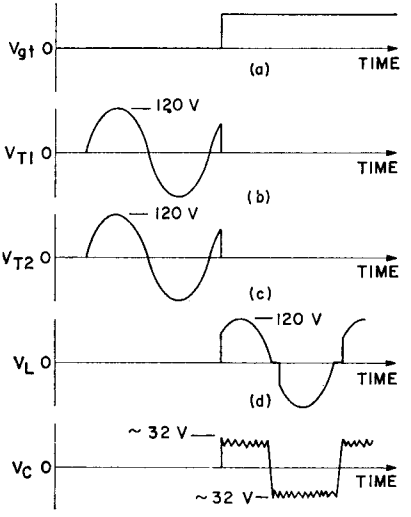



Figure 430. Waveforms for positive-logic switching.

A variation of this circuit with opposite (negative) logic is shown in Fig. 431. In this circuit, when triac $T_1$ is triggered on, triac $T_2$ and the lamp are off. When $T_1$ is off, capacitor $C_1$ can charge through resistors $R_1$ and $R_2$ to diac breakover, which discharges capacitor $C_1$ into the gate of triac $T_2$ and energizes the load. Little gate power is dissipated in this circuit because triac $T_2$ shorts across its gate circuitry when it is on.

Both of these circuits are shown with continuous gate drive into triac $T_1$. Logic power could be conserved by use of pulse drive, with no change of power-stage operation; however, the logic circuitry would be more complex.

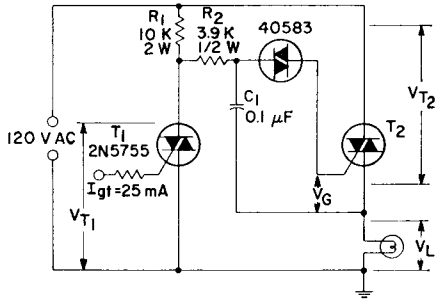



Figure 431. Negative-logic bulb-switching circuit.

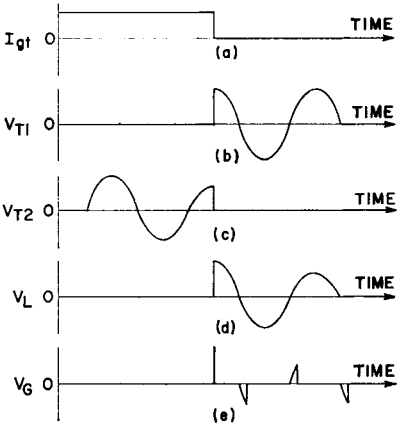



Figure 432. Waveforms for negative-logic switching.

## Traffic-Control Flasher

Thyristors can also be used to advantage in flasher-type traffic-control systems. In these applications, two lights are usually flashed on and off as a warning display. Fig. 433 shows a thyristor

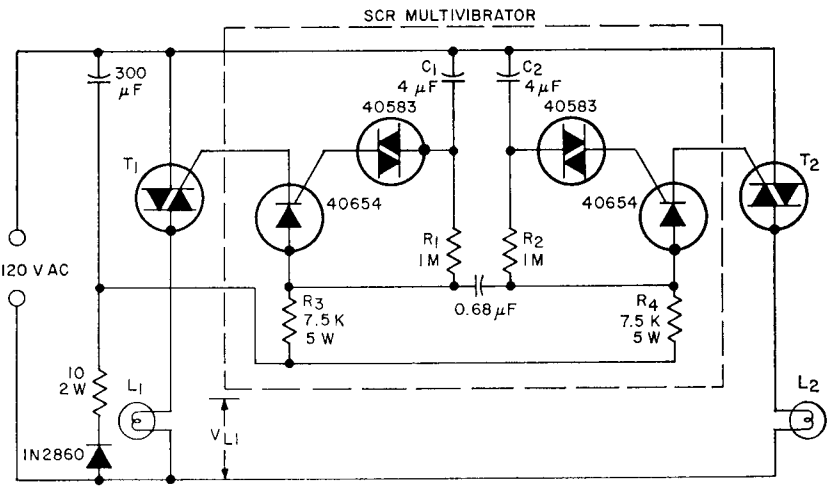



Figure 433. Thyristor flasher.

circuit that accomplishes this flashing function. As shown, a silicon-controlled-rectifier (SCR) multivibrator functions as the timer and flasher-triggering driver. The drive to the control triac is dc and is alternated between $T_1$ and $T_2$ according to the timing set in the multivibrator. A waveform for the component values shown is displayed in Fig. 434. The timing can be modified by selecting different values for any of the following components: $R_1$, $R_2$, $R_3$, $R_4$, $C_1$, $C_2$. The important features of this circuit are the simple, rugged dc power supply used and the use of SCR's as both timing and memory devices to trigger the triacs.

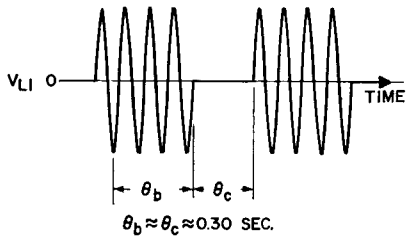



Figure 434. Timing of thyristor flasher.

## MOTOR CONTROLS

Thyristors have been widely accepted in power-control applications in industrial systems where high-performance requirements justify the economics of the application. The controls can be designed to provide good performance, maximum efficiency, and high reliability in compact packaging arrangements.

### Speed Controls for Universal Motors

Many fractional-horsepower motors are series-wound "universal" motors, so named because of their ability to operate directly from either ac or dc power sources. Fig. 435 is a schematic of this type of motor operated from an ac supply. In reality, the universal motor is a special form of series motor that has a laminated armature and field structure. Because most domestic applications today require 60-Hz power, universal motors are usually designed to

have optimum performance characteristics at this frequency. Most universal motors run faster at a given dc voltage than at the same 60-Hz ac voltage for which they are designed.

The field winding of a universal motor is in series with the armature and external circuit, as shown in Fig. 435. The current through the field winding produces a magnetic field which cuts across the armature conductors. The action of this field on the conductor of armature current subjects the individual conductors to a lateral thrust which results in armature rotation.

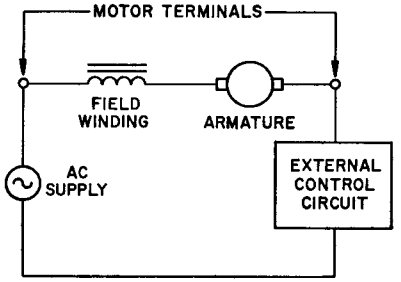



Figure 435. Schematic diagram for a series-wound universal motor.

The torque developed by a universal motor is a direct result of the magnitude of magnetic-field flux and armature current. The starting torque of a universal motor is high because the armature current at starting time is high. Similarly, at "stall" conditions, the armature current is high and results in a large torque. The stall torque of a series motor can be as high as 10 times the continuous rated torque.

High starting torque, adjustable speed characteristics, and small size are distinct advantages of a universal motor over a comparably rated single-phase induction motor. The speed can be adjusted by variation of the impressed voltage across the motor. Typical performance characteristic curves for a universal motor are shown in Fig. 436.

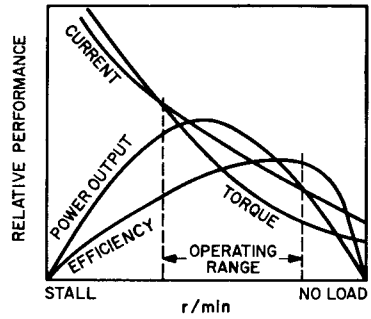



Figure 436. Typical performance curves for a universal motor.

One of the simplest and most efficient means of varying the impressed voltage and thus the speed of a universal motor is by control of the conduction angle of a thyristor (SCR or triac) placed in series with the load. Typical curves of the variation of motor speed with thyristor conduction angle for both half-wave and full-wave impressed motor voltages are shown in Fig. 437.

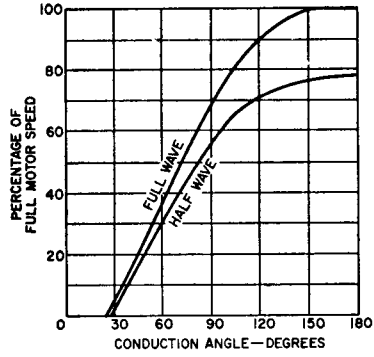



Figure 437. Typical performance curves for a universal motor with phase-angle control.

Fig. 438 shows a fundamental circuit of a direct-coupled SCR control with voltage feedback that is highly effective for speed control of universal motors. This circuit makes use of the counter emf induced in the rotating armature because of the residual magnetism in the motor on the half-cycle when the SCR is blocking.

This counter emf is a function of speed and, therefore, can be used as an indication of speed changes for mechanical-load variations. The gate-firing circuit is a resistance network consisting of $R_1$ and $R_2$. During the positive half-cycle of the source voltage, a fraction of the voltage is developed at the center-tap of the potentiometer and compared with the counter emf developed in the rotating armature of the motor. When the bias developed at the gate of the SCR from the potentiometer exceeds the counter emf of the motor, the SCR fires. AC power is then applied to the motor for the remaining portion of the positive half-cycle. Speed control is accomplished by adjustment of potentiometer $R_1$. If the SCR is fired early in the cycle, the motor operates at high speed because essentially the full-rated line voltage is applied to the motor. If the SCR is fired later in the cycle, the average value of voltage applied to the motor is reduced, and a corresponding reduction in motor speed occurs. On the negative half-cycle, the SCR blocks voltage to the motor. The voltage applied to the gate of the SCR is a sine wave because it is derived from the sine-wave line voltage. The minimum conduction angle occurs at the peak of the sine wave and is restricted to 90 degrees. Increasing conduction angles occur when the gate bias to the SCR is in-

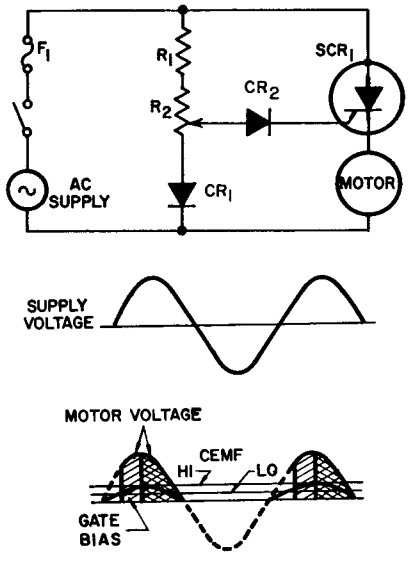



Figure 438. Half-wave SCR motor control circuit, with regulation.

Table XXVI—Components For Circuit Shown in Fig. 438

| AC SUPPLY | AC CURRENT | $F_1$ | $CR_1$, $CR_2$ | $R_1$ | $R_2$ | $SCR_1$ |
|---|---|---|---|---|---|---|
| 120V | 1A | 3AG, 1.5A, Quick Act | RCA-44004 | 2.7K, 4W | 1K, 2W | RCA-2N3528 |
| 120V | 3A | 3AB, 3A | RCA-44004 | 5.6K, 2W | 1K, 2W | RCA-2N3228 |
| 120V | 7A | 3AB, 7A | RCA-44004 | 5.6K, 2W | 500, 2W | RCA-2N3669 |
| 120V | 25A | 3AB, 25A | RCA-44004 | 2.7K, 4W | 500, 2W | RCA-2N3897 |
| 240V | 1A | 3AG, 1.5A, Quick Act | RCA-44005 | 10K, 5W | 1K, 2W | RCA-2N3529 |
| 240V | 3A | 3AB, 3A | RCA-44005 | 10K, 5W | 1K, 2W | RCA-2N3525 |
| 240V | 7A | 3AB, 7A | RCA-44005 | 5.6K, 7.5W | 500, 2W | RCA-2N3670 |
| 240V | 25A | 3AB, 25A | RCA-44005 | 5.6K, 7.5W | 500, 2W | RCA-2N3898 |

creased to allow firing at voltage values for which the phase delay with respect to the line voltage is less.

At no load and at the low-speed control setting, "skip-cycling" operation may occur, and motor speed may be erratic. Because no counter emf is induced in the armature when the motor is standing still, the SCR fires at low bias settings. The motor is then accelerated to a point at which the counter emf induced in the rotating armature exceeds the gate-firing bias of the SCR and prevents the SCR from firing. The SCR is not able to fire again until the speed of the motor is reduced (because of friction and winding losses) to a value for which the induced voltage in the rotating armature is less than the gate bias. At this time, the SCR fires again. Because the motor deceleration occurs over a number of cycles, there is no voltage applied to the motor (hence the term "skip cycling").

When a load is applied to the motor, the motor speed decreases and thus reduces the counter emf induced in the rotating armature. With a reduced counter emf, the SCR fires earlier in the cycle and provides increased motor torque to the load. Fig. 438 also shows variations of conduction angle with changes in counter emf. The counter emf appears as a constant voltage at the motor terminals when the SCR is blocking. Because the counter emf is essentially a characteristic of the motor, different potentiometer settings are required for comparable operating conditions for different motors. Component values for the circuit shown in Fig. 438 for use with the various RCA SCR's are listed in Table XXVI.

Fig. 439 shows a variation of the circuit in Fig. 388. The basic difference between the two circuits is that the circuit in Fig. 439 provides feedback for changing load conditions to minimize changes in motor speed. The feedback is provided by $R_7$, which is in series with the motor. A voltage proportional to the peak current through the motor is developed across the resistor. This voltage is stored on capacitor $C_2$ through diode $CR_2$, and is of a polarity that causes the bias on the resistance network of $R_3$ and $R_4$ to change in accordance with the load on the motor. With an increasing motor load, the speed tends to decrease. This decrease in motor speed causes more current to flow through the motor armature and field. When the current flowing through $R_7$ increases, the voltage stored on capacitor $C_2$ increases in the positive direction. This increase in capacitor voltage causes the transistors to conduct earlier in the cycle to fire the SCR, and to provide a greater portion of the power cycle to the motor. With a decreasing load, the motor current decreases, and the voltage stored across capacitor $C_2$ decreases. The transistors and SCR then conduct later in the cycle, and the resultant reduction in the average power supplied to the motor causes a reduced torque to the smaller load. Because motor current is a function of the motor itself, resistor $R_7$ has to be matched with the motor rating to provide optimum feedback for load compensation. Resistor $R_7$ may range from 0.1 ohm for larger-size universal motors to 1.0 ohm for smaller types. Circuit values for use with the various RCA SCR's are shown in Table XXVII.

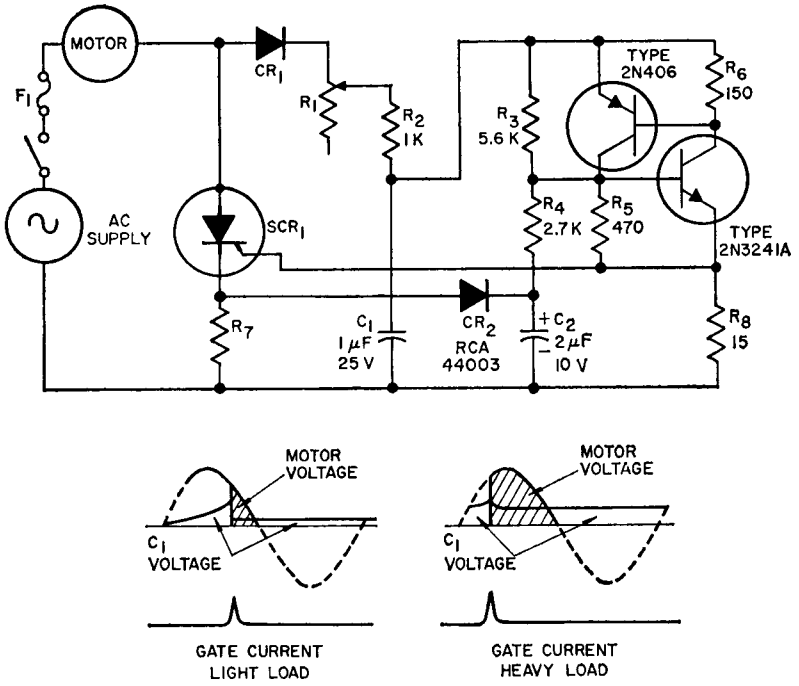



Figure 439. Half-wave SCR motor control circuit using two-transistor regenerative triggering, with regulation.

## Table XXVII—Components For Circuit Shown in Fig. 439

| AC SUPPLY | AC CURRENT | $F_1$ | $CR_1$ | $R_1$ | $SCR_1$ |
|---|---|---|---|---|---|
| 120V | 1A | 3AG, 1.5A, Quick Act | RCA-44004 | 75K, ½W | RCA-2N3528 |
| 120V | 3A | 3AB, 3A | RCA-44004 | 75K, ½W | RCA-2N3228 |
| 120V | 7A | 3AB, 7A | RCA-44004 | 75K, ½W | RCA-2N3669 |
| 120V | 25A | 3AB, 25A | RCA-44004 | 75K, ½W | RCA-2N3897 |
| 240V | 1A | 3AG, 1.5A, Quick Act | RCA-44005 | 150K, ½W | RCA-2N3529 |
| 240V | 3A | 3AB, 3A | RCA-44005 | 150K, ½W | RCA-2N3525 |
| 240V | 7A | 3AB, 7A | RCA-44005 | 150K, ½W | RCA-2N3670 |
| 240V | 25A | 3AB, 7A | RCA-44005 | 150K, ½W | RCA-2N3998 |

**Full-Wave Control**—This section discusses the application of thyristors (SCR's and triacs) to provide full-wave motor control. Fig. 440 shows three thyristor full-wave controls.

The speed of both universal and induction motors can be controlled by use of full-wave thyristor power control circuits. For fan motors, which have a particularly steep speed-torque curve, simple phase-control circuits can be used to provide suitably stable operating speeds. Other induction motors may require more complex feedback circuits to provide and maintain a variable speed. Cool-

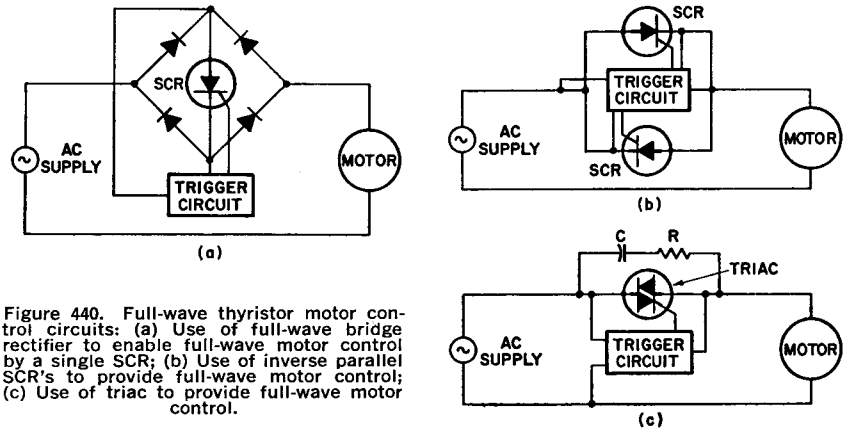



Figure 440. Full-wave thyristor motor control circuits: (a) Use of full-wave bridge rectifier to enable full-wave motor control by a single SCR; (b) Use of inverse parallel SCR's to provide full-wave motor control; (c) Use of triac to provide full-wave motor control.

ing provisions and bearings must be suitable for reduced-speed operation.

Figs. 418, 420, and 421 show full-wave thyristor circuits suitable for motor-speed controls.

An SCR full-wave circuit designed for applications requiring feedback for compensation of load changes is shown in Fig. 441. Operation is similar to that of the circuits discussed previously ex-

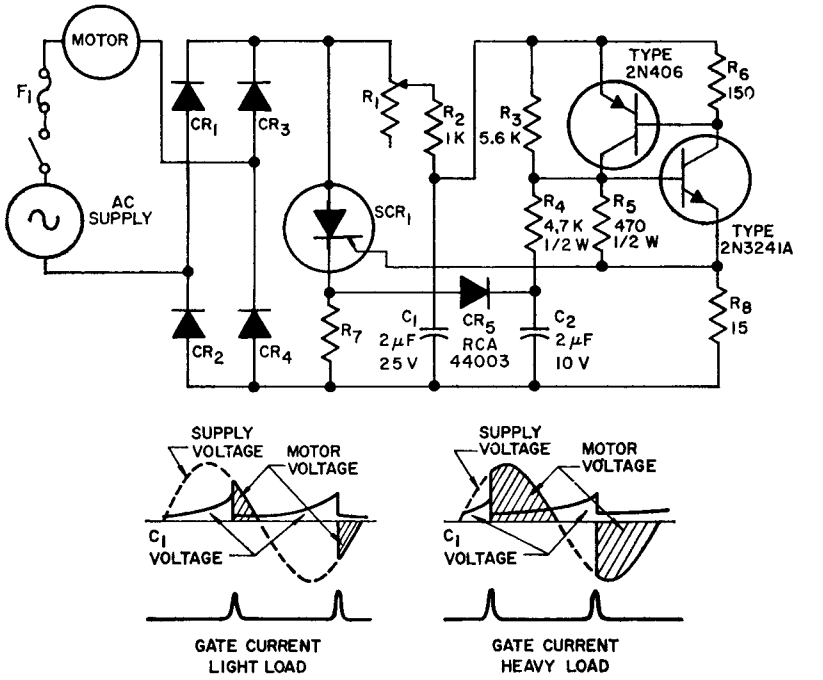



Figure 441. Full-wave SCR motor control circuit, with regulation.

Table XXVIII—Components For Circuit Shown in Fig. 441

| AC SUPPLY | AC CURRENT | $F_1$ | $CR_1$, $CR_2$, $CR_3$, $CR_4$ | $R_1$ | $SCR_1$ |
|---|---|---|---|---|---|
| 120V | 1A | 3AG, 1.5A, Quick Act | RCA-44003 | 50K, ½W | RCA-2N3528 |
| 120V | 3A | 3AB, 3A | RCA-40110 | 50K, ½W | RCA-2N3228 |
| 120V | 7A | 3AB, 7A | RCA-40110 | 50K, ½W | RCA-2N3669 |
| 240V | 1A | 3AG, 1.5A, Quick Act | RCA-44004 | 100K, ½W | RCA-2N3529 |
| 240V | 3A | 3AB, 3A | RCA-40112 | 100K, ½W | RCA-2N3525 |
| 240V | 7A | 3AB, 7A | RCA-40112 | 100K, ½W | RCA-2N3670 |

cept that this circuit has full-wave conduction with proportional control. Table XXVIII gives a component list for use of this circuit with various SCR's.

## Induction-Motor Reversing Controls

Triacs are finding increasing application in motor-reversing circuits. Motor-reversing systems that use electromechanical relays suffer from contact arcing, which results in short life and costly maintenance. Fig. 442 shows a triac-controlled motor-reversing circuit that uses a split-phase capacitor-run motor. When triac No. 1 is in the off state and triac No. 2 is in the on state, motor direction is controlled by triac No. 2. When triac No. 2 reverts to the off state and triac No. 1 turns on, the direction of motor rotation is reversed. Caution should be exercised in this type of circuit because, if triac No. 1 is turned on while triac No. 2 is on, a loop current that results from capacitor discharge will occur and may damage the triacs. A small resistance placed in the capacitor-loop current path limits the magnitude to tolerable levels and provides reliable operation.

Use of triacs in motor-reversing circuits is illustrated by electronic garage-door systems which use the principle of motor-reversing for garage-door direction control. The system contains a transmitter, a receiver, and an operator to provide remote con-

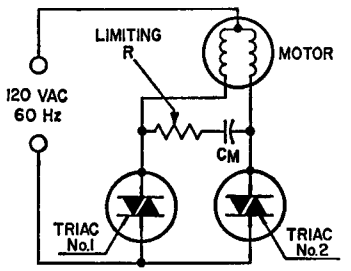



Figure 442. Motor-reversing circuit using two triacs.

trol for door opening and closing. The block diagram in Fig. 443 shows the functions required for a complete solid-state system.

When the garage door is closed, the gate drive to the DOWN triac is disabled as a result of the lower-limit closure, and the gate drive to the UP triac is inactive because of the conduction state of the flip-flop. A momentary keying of the transmitter causes the receiver to activate the time-delay monostable multivibrator, which changes the state of the flip-flop so that continuous gate drive is provided to the UP triac. The door continues to travel in the UP direction until the upper-limit closure

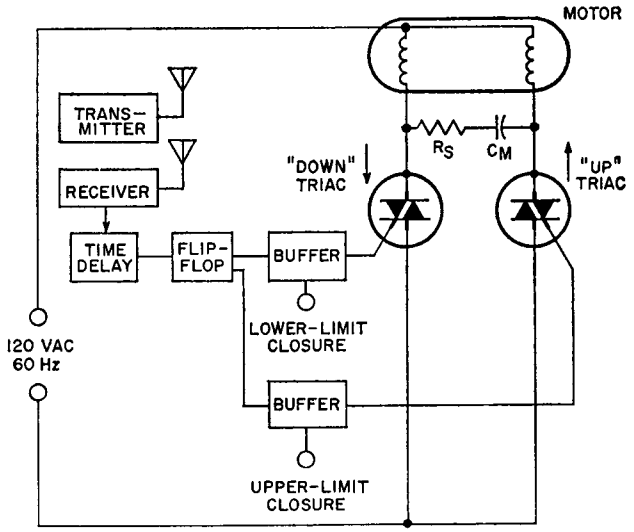



Figure 443. Block diagram of solid-state garage-door system.

switch disables the gate drive to the UP triac. A second keying of the transmitter provides the DOWN triac with gate drive and causes down travel until the DOWN triac is disabled by the lower-limit closure. The time during which the time-delay monostable multivibrator is active should override normal transmitter keying to eliminate erreoneous firing. Fig. 444 is a timing diagram that indicates the time the time-delay monostable multi-vibrator must be active in order to override transmitter keying, and Fig. 445 shows a circuit diagram which performs the required logic for motor direction. A significant feature of this system is that, during door travel, transmitter keying provides motor directions independent of the upper- and lower-limit closures. Additional features such as obstacle obstructions, manual control, or time-delay overhead garage lamps can be incorporated into the system very economically.

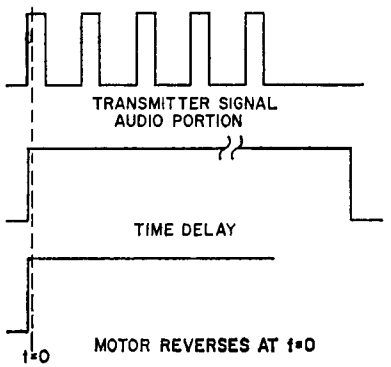



Figure 444. Timing diagram for reversing system.

## AC VOLTAGE REGULATOR

Thyristors are used in a basic ac-voltage regulating circuit that prevents ac rms or dc voltage from fluctuating more than ± 3 per cent

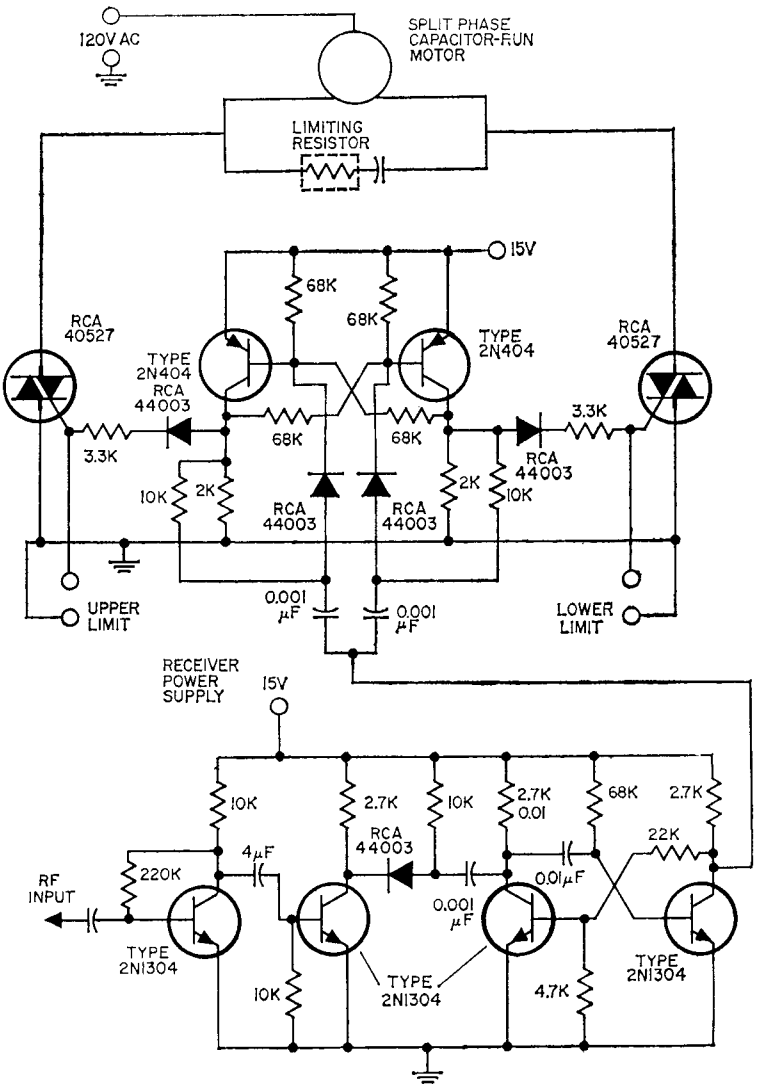



Figure 445. Garage-door control.

in spite of wide variations in input line voltage. Load voltage can also be held within ±3 per cent of a desired value despite variations in load impedance through the use of a voltage-feedback technique.

The voltage regulator can be used in photocopying machines, light dimmers, dc power supplies, and motor controllers (to maintain fixed speed under fixed load conditions).

## Circuit Operation

The basic configuration of the ac regulator is shown in Fig. 446. For simplicity, only a half-wave SCR configuration is shown; however, the explanation of circuit operation is easily extended to include a full-wave regulator that uses a triac.

The RCA-40583 diac used as the trigger device in Fig. 446, exhibits a high-impedance, low-leakage-current characteristic until the applied voltage reaches the breakover voltage $V_{BO}$, approximately 35 volts. Above this voltage, the device exhibits a negative resistance so that voltage decreases as current increases.

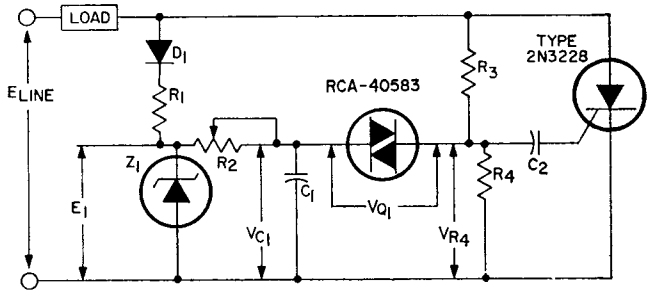



Figure 446. Basic ac regulator.

Capacitor $C_1$ in Fig. 446 is charged from a constant-voltage source established by zener diode $Z_1$. The capacitor is charged, therefore, at an exponential rate regardless of line-voltage fluctuations. A trigger pulse is delivered to the 2N3228 SCR when the voltage across capacitor $C_1$ is equal to the trigger voltage of the diac plus the instantaneous voltage drop developed across resistor $R_4$ during the positive half-cycle of line voltage. When the diac is turned on, the SCR is turned on for the remainder of the positive cycle of source voltage. Control of the conduction angle of the SCR regulates the rms voltage to the load.

Regulation is achieved by the following means: When line voltage increases, the voltage across $R_4$ increases, but the charging rate of $C_1$ remains the same; as a result, the voltage across $C_1$ must attain a larger value than required without line-voltage increase before the diac can be triggered. The net effect is that the pulse that triggers the SCR is delayed, and the rms voltage to the load is reduced. In a similar manner, as line voltage is reduced, the SCR turns on earlier in the cycle and increases the effective voltage across the load.

Fig. 447 shows the voltage waveforms exhibited by the ac regulator at both high and low line voltage. The charging voltage, $E_1$, for capacitor $C_1$ is equal to the zener voltage and remains constant up to the instant that the SCR is turned on. The capacitor voltage, $V_{C1}$, increases exponentially because the charging voltage $E_1$ is constant. The voltage across resistor $R_4$ conforms to the sinusoidal variations of the 60-Hz line voltage. At any given phase angle, the voltage across $R_4$ in-

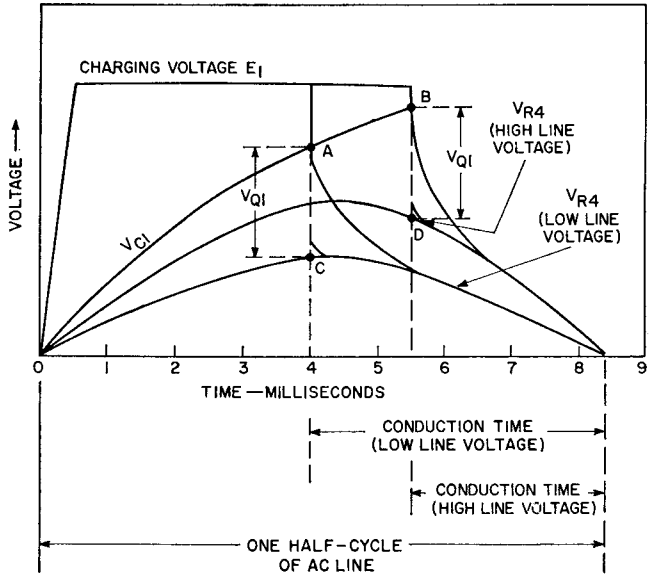



Figure 447. Voltage waveforms exhibited by the ac regulator in Fig. 446.

creases if line voltage increases and decreases if line voltage decreases.

The diac and SCR both trigger when the capacitor voltage, $V_{C1}$, equals the breakover voltage of the diac plus the instantaneous value of voltage developed across $R_4$ during the positive half-cycle of line voltage. This capacitor voltage is represented by points A and B for the low and high line-voltage conditions, respectively. The instantaneous voltages across resistor $R_4$ just before the SCR is triggered are represented by points C and D for the low and high line-voltage conditions, respectively. The voltage difference between points A and C and between points B and D is equal to the breakdown voltage of the diac.

Fig. 447 illustrates that the conduction time of the SCR is decreased as line voltage increases, and is increased when the line voltage decreases. By proper selection of the values of the voltage-divider-ratio resistors $R_3$ and $R_4$, it is possible to prevent the load voltage from varying more than 3 per cent with a 30-per-cent (approximate) change in line voltage.

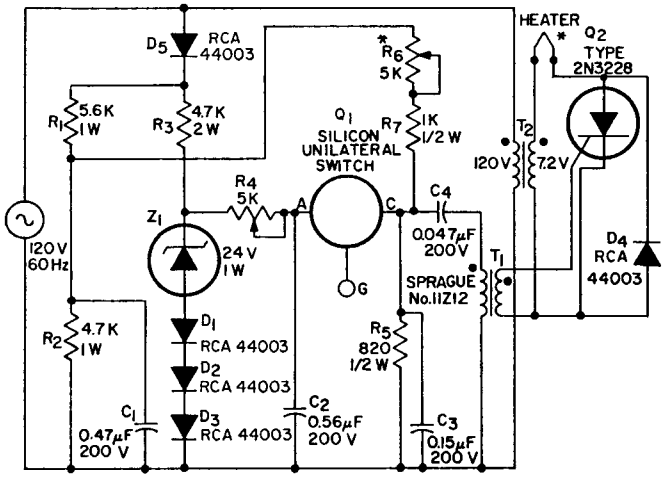
It should be mentioned that during measurements of load voltage careful consideration must be given to the measuring instruments. Many circuits produce a non-sinusoidal voltage across the load; the rms value of this voltage can be measured only with a true rms meter, such as a thermocouple meter. It is possible, however, that in certain applications the low input impedance of the thermocouple meter might load down the circuit being measured. In such cases, a high-input-impedance rms meter may be required.

## Heater Regulation

Fig. 448 shows a basic regulating technique for applications in

which it is desired to maintain constant voltage across a load such as a receiving-tube heater, the filament of an incandescent lamp, or possibly a space heater. It should be noted that this configuration is actually a half-wave regulator. However, the circuit of Fig. 448 differs from the circuit of Fig. 446, in which one half-cycle is blocked from the load and the other half-cycle is phase-controlled to provide regulation. In Fig. 448, essentially full voltage is applied to the load for one half-cycle by means of the 44003 silicon rectifier $D_4$; the other half-cycle is phase-controlled by the 2N3228 SCR to provide regulation.

regulation is provided by a silicon unilateral switch* and a control circuit, as follows: Capacitor $C_2$ is charged from a voltage source that is maintained constant by zener diode $Z_1$; the silicon rectifiers $D_1$, $D_2$, and $D_3$ compensate for the change in zener voltage with temperature. The voltage across $C_2$ increases until the sum of the breakover voltage of $Q_1$ and the instantaneous voltage across $R_5$ is exceeded. At this point, a positive pulse is coupled into the gate of the SCR by means of the pulse transformer $T_1$. The SCR then switches on for the remainder of the positive cycle of line voltage. Control of the conduction angle of




* IN THE CLOSED-LOOP REGULATOR, $R_6$ IS REPLACED BY A PHOTOCELL, AND A POTENTIOMETER IN SERIES WITH A 6-VOLT INCANDESCENT LAMP IS CONNECTED IN PARALLEL WITH THE HEATER TERMINALS
NOTE: ALL RESISTOR VALUES ARE IN OHMS

Figure 448. A circuit using a regulator to maintain voltage constant across a load.

The circuit in Fig. 448 is an open-loop regulator that features a high degree of safety; i.e., an open- or short-circuited component does not result in an excessive load voltage. Phase-controlled voltage the SCR varies the rms voltage to the heater.

* A silicon unilateral switch is a silicon, planar, monolithic integrated circuit that has thyristor electrical characteristics closely approximating those of an ideal four-layer diode. The device shown switches at approximately 8 volts.

As line voltage increases, the voltage across resistor $R_5$ also increases; because capacitor $C_2$ charges along the same exponential curve, however, the voltage across $C_2$ must attain a larger value before the SCR is turned on. The net effect is a delay in the trigger pulse and reduced rms voltage across the heater. In a similar manner, as line voltage is reduced, the SCR turns on earlier in the cycle and increases the effective voltage across the heater. By proper adjustment of potentiometer $R_6$ in conjunction with potentiometer $R_4$, it is possible to obtain excellent heater-voltage compensation over a range of line voltages. Fig. 449 shows the waveforms associated with the heater-regulator circuit.

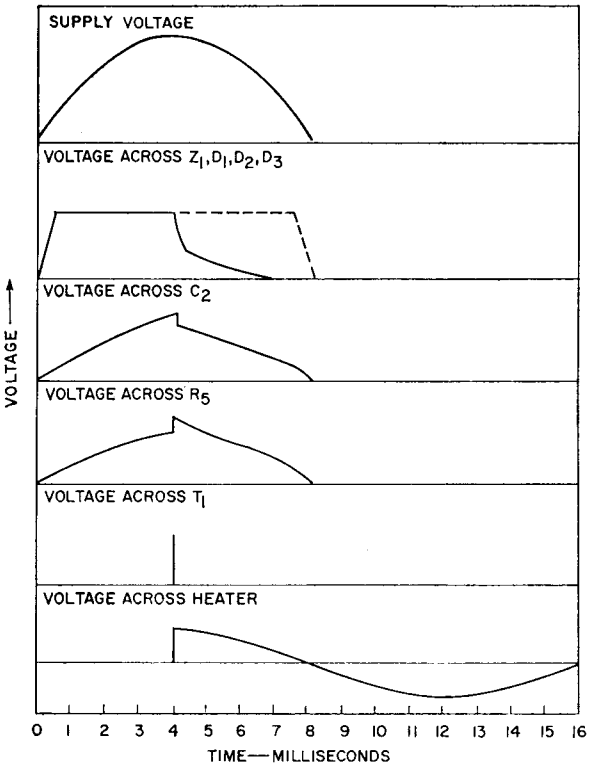



Figure 449. Voltage waveforms exhibited by the circuit of Fig. 448.

Curve A in Fig. 450 shows heater voltage as a function of line voltage for the open-loop regulator circuit shown in Fig. 448. Curve B in Fig. 450 shows a similar curve for a closed-loop regulator using a lamp-photocell module. The lamp, in series with a

limiting resistor, is connected across the heater terminals, and the photocell replaces potentiometer $R_6$. The lamp unit senses the phase-controlled true rms heater voltage. Changes in lamp brightness produced by heater-voltage variations change the photocell resistance in reverse proportion to the lamp voltage. The remainder of the circuit functions as previously described except that regulation is obtained not only through the monitoring of the instantaneous magnitude of line voltage, but also through the sensing of the true rms voltage across the heater. This characteristic identifies the circuit as an ac voltage regulator with closed-loop feedback control. The closed-loop regulator produces less error, is more resistant to the drift effects of components, and is easier to adjust than the open-loop regulator.

The lamp used in the closed-loop regulator is rated at 6 volts, but the series resistor limits the voltage to approximately 2 volts so that extremely long lamp life can be expected. An additional advantage at low voltage is that the light intensity varies linearly with the voltage across the lamp so that a small increase in voltage increases brightness markedly; near rated voltage the intensity does not vary linearly and the variation in brightness is not very apparent. A loss in sensitivity would result if the lamp were operated at its rated voltage.

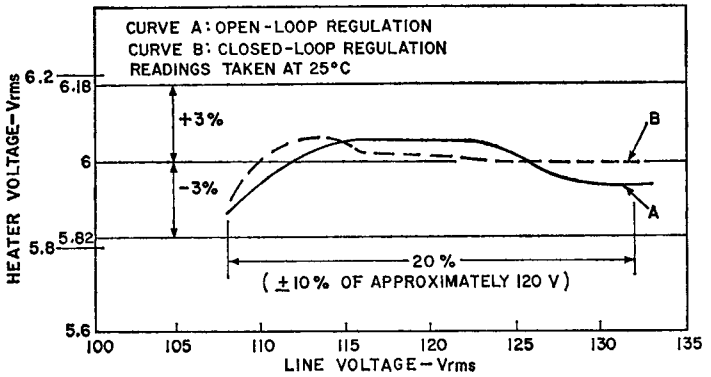
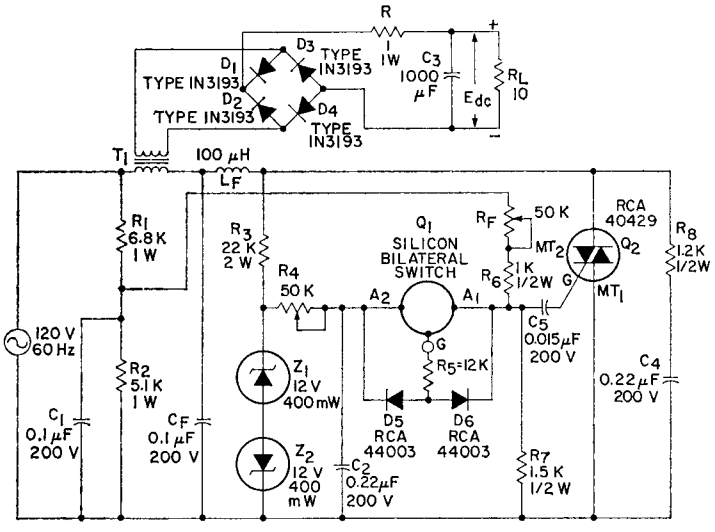



Figure 450. Heater voltage as a function of line voltage of the open- and closed-loop regulators.

The open-loop regulator can regulate 6 volts to within ±3 per cent within a temperature range from 10 to 40°C with an input-voltage swing of ±10 per cent. The closed-loop regulator can regulate 6 volts to within ±2 per cent within a temperature range from 0 to 60°C with an input-voltage swing of ±10 per cent.

## Line-Voltage-Regulated DC Supply

A simple but stable dc power supply that uses thyristors for line-voltage regulation is shown in Fig. 451. The power-supply section consists of the well-known full-wave bridge with RC filter. A




NOTE: ALL RESISTOR VALUES ARE IN OHMS

Figure 451. A voltage-regulated dc supply.

line-voltage transformer is employed to step down the supply voltage of 120 volts rms to approximately 12.5 volts rms. If a dc output voltage greater than 10 volts is desired, a transformer with a lower primary-to-secondary turns ratio should be employed.

The heart of the regulator shown in Fig. 451 is the phase-controlled triac on the primary side of the line transformer. Because the load presented to the triac is somewhat inductive, an RC network is used to assure proper commutation; $L_F$ and $C_F$ suppress rf interference. The circuit automatically compensates for wide variations in line voltage. Fig. 452 shows a curve of line voltage as a function of load voltage, $E_{dc}$, for a constant load of 10

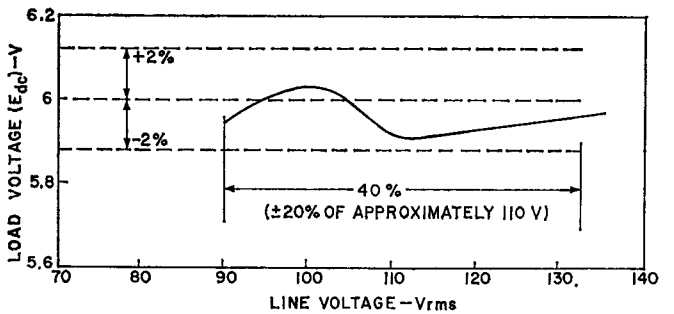



Figure 452. Load voltage as a function of line voltage for the circuit of Fig. 451; load resistance is constant at 10 ohms.

ohms. Fig. 453 shows the voltage waveforms associated with the circuit of Fig. 451.

If increased line, temperature, and load compensation is desired in the regulated dc supply of Fig. 452, a closed-loop type of control can be obtained by use of a photocell in place of potentiometer $R_F$ and connection of a lamp across the output terminals of the supply in such a way that the light from the lamp can impinge on the photocell surface.

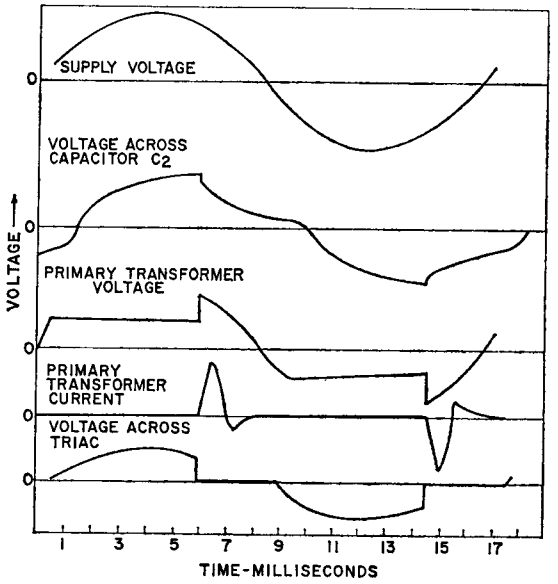



Figure 453. Voltage waveforms exhibited by the circuit of Fig. 451.

# Ballast Circuits for Mercury-Arc Lamps

RECENT advances in the voltage- and current-handling capabilities of power transistors have made possible the design of solid-state switching-regulator ballasts that offer significant advantages over conventional ballasting devices for high-pressure mercury-arc lighting systems. In addition to the usual transistor-circuit benefits of reduced weight and bulk, the solid-state ballasts provide unmatched power regulation for line-voltage fluctuations and exceptional versatility. The basic solid-state ballast circuit includes a built-in lamp-dimming feature that permits a single design to be used with lamps of various power ratings over a range of 50 to 150 per cent of the power rating specified for the ballast design. Moreover, transistor ballast circuits eliminate the annoying strobe effect associated with conventional ballasting devices and thereby make the long-life, efficient mercury-arc lamps suitable for use in studios and similar critical lighting areas.

## RELATIVE MERITS OF VARIOUS LIGHTING SYSTEMS

Table XXIX compares the characteristics and provides a brief cost analysis of incandescent, fluorescent, mercury-arc, Lucalox,* and sodium-lamp lighting systems. The over-all cost of each system is determined by three main factors: (1) power consumed during operation, (2) replacement and maintenance, and (3) initial installation. The cost of initial installation is almost insignificant when compared to the other cost items. In general, power-consumption costs are approximately seven times greater than the costs of initial installation. Replacement-and-maintenance costs, at present, represent two or three times the initial-installation costs, but are rising at a very rapid rate. Because of the higher efficiency and reduced maintenance requirements of gas-discharge (arc) lamps, lighting systems that use

* Trade name of the General Electric Company.

## Table XXIX—A Comparison of the Characteristics of Various Lighting Sources

| Type | Description | Ingredients | Light Quality | Percent Eff. | Life (hrs) | Warmup Time | Time Before Restart | 400W Bulb or Equivalent | | | Cents/ lumen-hr $\times 10^{-4}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Bulb Cost | Indoor Fixture Cost | Ballast Cost | |
| Incandescent | Filament (point light source) | Tungsten in Nitrogen | Good—much red, no blue (continuous spectrum) | 2.6 | 2,000 | None | None | $ 1.25 | $10.00 | — | 1.80 |
| Fluorescent | Low-pressure vapor with phosphor correction | Mercury | Good | 9.5 | 10,000 | Few Seconds | None | $16.00 | $25.00 | $ 20.00 | 0.56 |
| Mercury Arc (Color Corrected) | High-pressure vapor with phosphor correction (point source) | Mercury and Argon in Quartz burner | Slightly cold | 7.5 | 20,000 | 4 min. | 5 min. | $20.00 | $30.00 | $ 45.00 | 0.70 |
| Lucalox | High-pressure, high-temperature vapor (point source) | Sodium and Mercury in Alumina burner | Sunny, much yellow | 15.0 | 6,000 | 3.0 min. | 1 min. | $45.00 | $30.00 | $120.00 | 0.54 |
| Sodium Vapor | High-pressure vapor (point source) | Sodium, Neon | Yellow monochromatic | 15.0 | 6,000 | 18 min. | None | | | | |

NOTE: In the cost analysis, the maintenance factor proportional to life of bulb was not included. The electrical power cost was assumed to cost three cents per kilowatt hour, and the life of the ballast and fixture was estimated to be 60,000 hours.

these types of lamps have displaced those that use incandescent (tungsten-filament) lamps in most industrial and highway installations.

Fluorescent lighting systems are currently the most widely used of the various gas-discharge types. In view of the rapid rise in maintenance costs, however, the long-life (approximately 20,000 hours) mercury-arc bulbs have become increasingly attractive. The use of mercury-arc lighting systems is increasing at a rate that far exceeds that of fluorescent systems, and mercury-arc lamps are now being used in numerous applications for which fluorescent types were previously employed, as well as in many new applications in the home. In addition, greater expansion of the application of mercury-arc lamps is expected to result from new phosphors which will further develop the light characteristics of these devices.

Another important consideration in selection of a gas-discharge lighting system is whether the lamp is to be operated from an ac or a dc power source. Neither fluorescent nor Lucalox lamps are particularly well suited for dc operation. When fluorescent lamps are operated from dc voltages, the direct currents force the mercury atoms to one end of the arc tube with a resultant dimming of the other end. Moreover, the lamp efficiency for dc operation may be only 70 per cent of that for high-frequency ac operation, and the life of a dc-operated fluorescent is derated 20 per cent. The Lucalox arc tube cannot withstand the temperature differential between the electrodes that is characteristic of dc operation. This temperature differential results because the positive electrode is disproportionately heated by electron bombardment.

High-pressure mercury-arc lamps provide the same efficiency for either ac or dc operation. For dc operation, the mercury-arc lamp offers the advantage of no strobe effect. However, because only one arc-tube electrode is bombarded by electrons during dc operation, a slight decrease in tube life results from the overheating of this electrode. A redesign of the electrodes should alleviate this condition.

## CHARACTERISTICS OF MERCURY-ARC LAMPS

Fig. 454 shows the basic construction of a mercury-arc lamp. The arc tube is made of quartz to withstand the wide extremes and sharp gradients of temperature to which it is subjected. This quartz tube contains some argon in addition to the mercury which is evaporated and ionized to provide

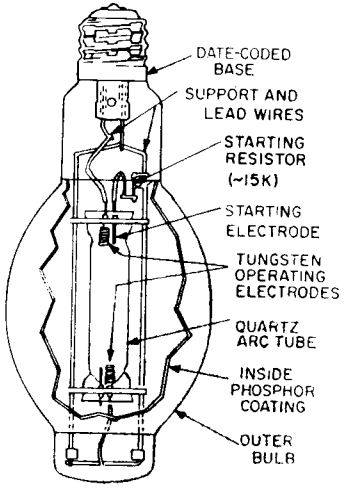



Figure 454. Cutaway view of a mercury-arc lamp.

the arc lighting. The argon is a starting aid and also prolongs the life of the lamp electrodes by retarding electron bombardment and evaporation of the electrodes.

A mercury-arc lamp is essentially a varying impedance which is driven from the ac line through an inductive ballast. Fig. 455 shows the voltage and current characteristics of the mercury-arc bulb during warmup. The argon in the arc tube ionizes when the voltage across the lamp electrodes rises to 200 volts (point 1 in Fig. 455); the voltage then decreases rapidly to 18 volts (point 2 in Fig. 455). The lag in current with respect to the bulb voltage, shown in Fig. 455(b), results because of the ballasting inductor in series with the lamp electrodes. Warm up of the mercury-arc bulb is completed in approximately 3 minutes. During this period, the mercury vaporizes, and a stable operating point is then attained (point 3 in Fig. 455). The inductive ballast is designed so that the slope of the change in voltage between points 2 and 3 results in a reduced warm-up time. If the mercury-arc bulb is turned off, the mercury cannot be re-ionized until approximately 5 minutes have elapsed, i.e., until the pressure and temperature in the arc tube have decreased sufficiently.

## CONVENTIONAL BALLASTING METHODS

For operation of the mercury-arc lamp in 120-volt line applications, a voltage step-up trans-

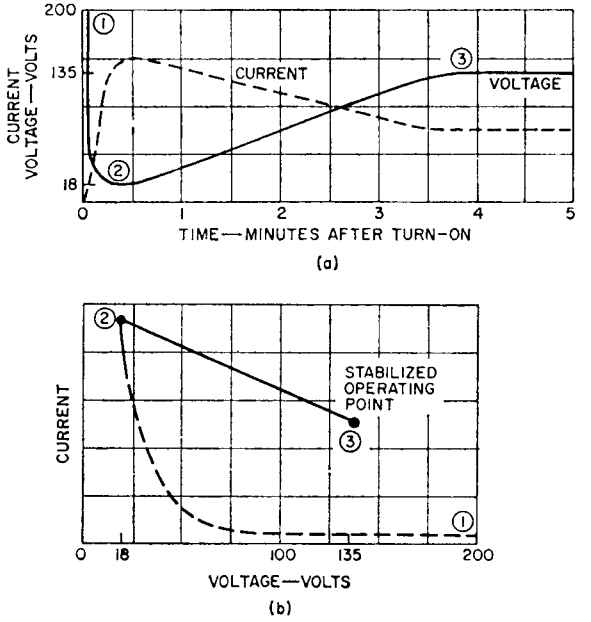



Figure 455. Warm-up characteristics of a typical (135-volt) mercury-arc lamp: (a) current and voltage as a function time; (b) current as a function of voltage.

former ballast must be used to develop the high starting potential (200 volts) and the required current-voltage slopes [shown in Fig. 455(b)]. This transformer ballast, however, must have a large leakage inductance to accommodate the varying bulb characteristics. For operation of the mercury-arc lamp from ac voltages of 220 volts or higher, ballasting may be provided by a simple series reactor. Fig. 456 shows the two ballasting arrangements. As shown in the circuit diagrams, a power-factor-correction capacitor (usually an oil type) should be used with each ballast circuit. The efficiency of these circuits ranges from 75 to 95 per cent.

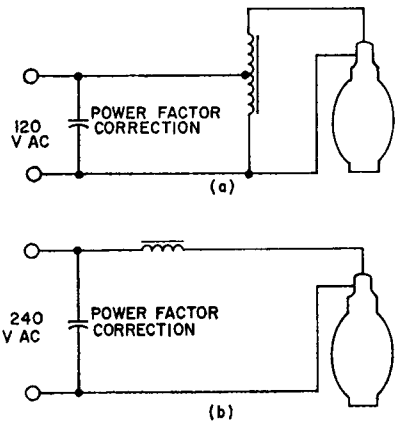



Figure 456. Conventional ballasts for 120- and 240-volt ac mercury-arc lamps.

A major disadvantage of conventional ballasting reactors is poor power regulation for line-voltage fluctuations. The power regulation can be improved, as shown in Fig. 457(a), by use of

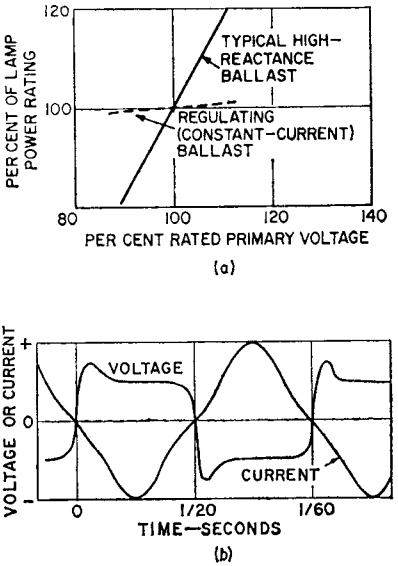



Figure 457. Characteristics of conventional mercury-arc-lamp ballasts: (a) regulation characteristics; (b) voltage and current as a function of time.

a saturating (constant-current) type of ballasting reactor. When this type of ballasting is employed, however, circuit efficiency is reduced, and a longer bulb warm-up period is required. Voltage and current waveshapes of conventional ballasts are shown in Fig. 457(b).

## SOLID-STATE BALLASTING CIRCUITS

The block diagram in Fig. 458 shows the basic requirements of an electronic type of ballasting circuit for mercury-arc lamps. This type of ballast may be operated from either an ac or dc voltage source; the rectifier bridge, or course, is not required for dc source voltages. AC input

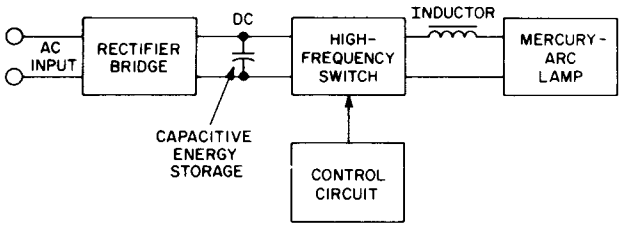



Figure 458. Block diagram of an electronic ballasting system for mercury-arc lamps.

voltages are first rectified, and the resultant dc voltage is then converted to the level required for application to the mercury-arc lamp by some type of inverter or converter (solid-state switch and associated control circuit).

Efficient conversion of a voltage from one level to another level requires the use of an inductive component. If a size advantage is to be realized from the use of an electronic ballasting circuit, the frequency of the solid-state switch must be high enough so that the converter inductor is significantly smaller than a conventional 60-Hz ballasting reactor. A small inductor, however, cannot maintain the arc in a mercury-arc bulb as the ac source voltage swings through zero. If no other storage element were included in the electronic ballasting circuit, the arc would be extinguished; the mercury-arc lamp must then be allowed to cool sufficiently before a new arc can be produced. The electronic ballast, therefore, includes a capacitor for additional energy storage when the circuit is operated from an ac voltage source.

Fig. 459 shows three prospective electronic ballasting circuits: a ringing-choke converter, a push-pull inverter, and a switching regulator. Table XXX summarizes the characteristics of each type. The important considerations in the selection of one circuit in preference to the other circuits are power-regulation capabilities, operating efficiency, small size, and requirements of the solid-state switching element.

The ringing-choke inverter offers the advantage of a dc output which is completely independent of the input voltage; its operating efficiency, however, is low in comparison to the other types of ballasting circuits. The push-pull inverter suffers from the fact that it provides an ac output with poor regulation. In addition, this circuit requires three magnetic components, which substantially add to the bulk of the ballast. The switching regulator is the most efficient and provides the best power regulation of the three types of electronic ballasting circuits. This ballasting circuit also imposes the least stringent requirements on the solid-state power-switching element, the most critical component of any electronic ballast. These factors make the switching regulator the most economical choice for an electronic ballasting circuit. An additional advantage of this circuit is that it requires only a single magnetic component; integrated-circuit construction tech-

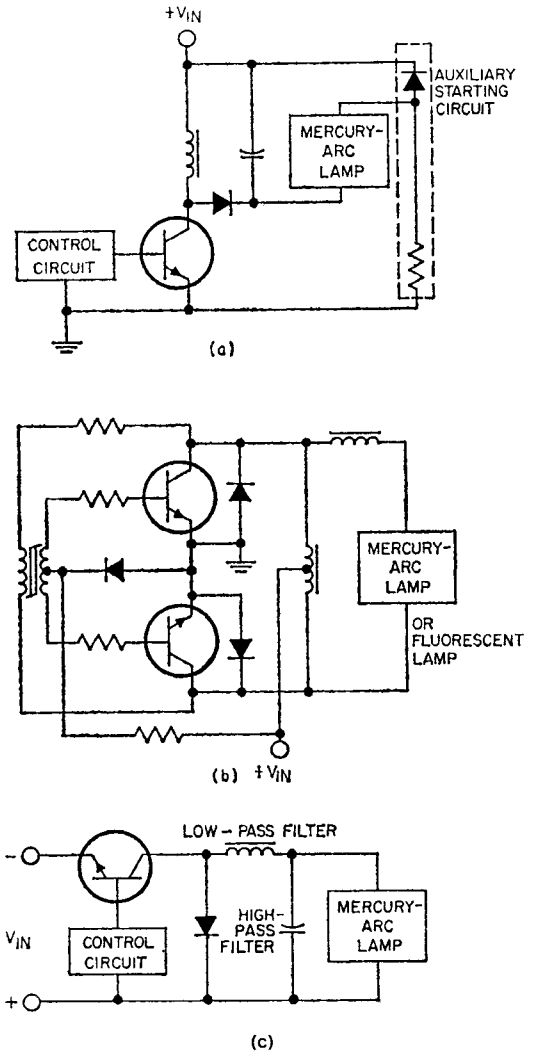



Figure 459. Three basic circuit configurations that may be used in electronic ballasting sytems: (a) ringing-choke converter; (b) push-pull inverter; (c) switching regulator.

niques, therefore can be readily applied to achieve the small sizes desired for ballasting elements. A disadvantage of the switching-regulator ballast is that the output voltage is always less than the input voltage, as indicated in Table XXX.

Table XXX—Characteristics of Various Electronic Ballasting Circuits

| CIRCUIT | $V_{IN}$–$V_{OUT}$ | DC or AC OUT | REMARKS | REGULATION | APPROX. EFF. | Switching Transistor No. of Devices | Switching Transistor $V_{CE}$ | Switching Transistor $I_C$(peak) |
|---|---|---|---|---|---|---|---|---|
| Ringing Choke | Independent | DC | Complex Circuit (Open-load protection) | Excellent | 70% | 1 | $V_{IN} + V_{OUT}$ | ~(4X) $I_{OUT}$ |
| Push-Pull | Independent | AC | Three magnetic elements | Limited | 80% | 2 | $2V_{IN}$ | (4X) $I_{OUT}$ |
| Switching Regulator * | $V_{IN} > V_{OUT}$ | DC | Simple Circuit - | Excellent | 90% | 1 | $V_{IN}$ | (2X) $I_{OUT}$ |

* The switching regulator offers the greatest efficiency and least stringent switching-transistor requirement.

## 120-Volt Switching-Regulator Ballast

For operation in 120-volt line applications, the basic switching-regulator circuit is modified, as shown in Fig. 460, so that the solid-state switching element (transistor $Q_1$) is operated in the positive feedback mode. The rectified 120-volt ac input appears as a dc voltage across the $V_{IN}$ terminals of the circuit. This voltage drives transistor $Q_1$ into saturation. The collector current of transistor $Q_1$ rises linearly through the primary ($L_1$) winding of transformer $T_1$ until the voltage drop across the current-sensing resistor $R_2$ increases above a predetermined threshold level. At this point, transistor $Q_3$ is turned on, and the collector current of this transistor, in turn, drives transistor $Q_2$ into conduction to create a virtual short between base and emitter of transistor $Q_1$. In this way, the drive input to transistor $Q_1$ is effectively removed. The inductive kick from the $L_1$ primary winding of transformer $T_1$ that results

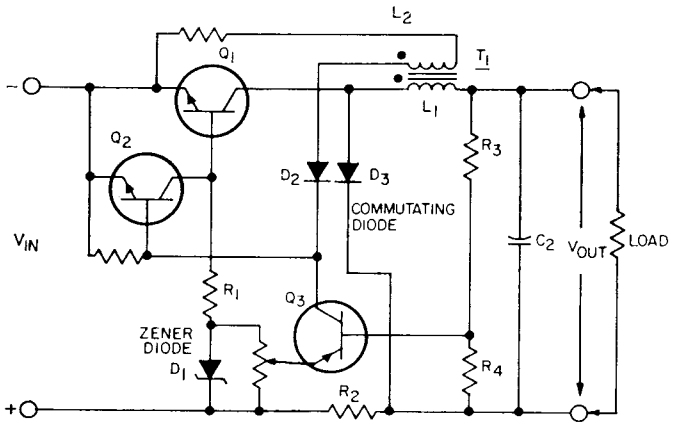



Figure 460. 120-volt switching regulator.

from the decrease in the collector current of transistor $Q_1$ is clamped by the commutating diode $D_3$ so that the current decays linearly through the winding. Positive feedback coupled from the secondary ($L_2$) winding of transformer $T_1$ holds switching transistor $Q_1$ in the "off" state until the current through the transformer primary winding decreases to zero. The cycle is then repeated. Fig. 461 shows the

During turn on, the voltage across the regulator inductor is essentially the algebraic difference between the input and output voltages (i.e., $E_L = V_{in} - V_{out}$). Because both of these voltages are constant, their difference results in a linearly increasing current through inductor $L_1$. The rate of change of the current (di/dt) is then the peak value to which the current rises divided by the turn-on period (i.e.,

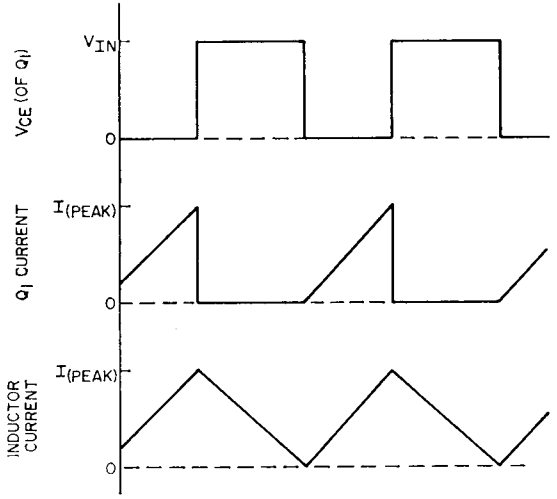



Figure 461. Typical voltage and current waveforms for the switching regulator shown in Fig. 460.

significant current and voltage waveshapes for the circuit. It is apparent from these waveshapes that switching losses occur only during turn-off.

The equations for the turn-on ($t_{on}$) and turn-off ($t_{off}$) times and the switching frequency (f) of the switching-regulator ballasting circuit can be derived from the following basic relationship for the voltage developed across an inductor:

$$E_L = L\frac{di}{dt} \qquad (358)$$

$di/dt = I_{peak}/t_{on}$). For those conditions, Eq. (358) may be rewritten in the following form:

$$V_{in} - V_{out} = L_1\frac{(I_{peak})}{t_{on}} \qquad (359)$$

If this equation is solved for $t_{on}$, the following result is obtained:

$$t_{on} = \frac{L_1\,(I_{peak})}{V_{in} - V_{out}} \qquad (360)$$

The equation for the turn-off time can be similarly derived. During this period, however, the voltage across inductor $L_1$ is essentially equal to the output voltage. The current decays linearly through the inductor so that the rate of change of current is constant over the turn-off period. When these conditions are imposed on Eq. (358), the following equation for the turn-off time can be derived:

$$t_{off} = \frac{L_1 (I_{peak})}{V_{out}} \qquad (361)$$

By use of Eqs. (360) and (361), the switching frequency of the switching-regulator ballast can be expressed in terms of the inductor $L_1$, the peak current, and the input and output voltages:

$$f = \frac{1}{t_{on} + t_{off}} = \frac{V_{out} + (V_{in} - V_{out})}{L_1 (I_{peak}) (V_{in})} \qquad (362)$$

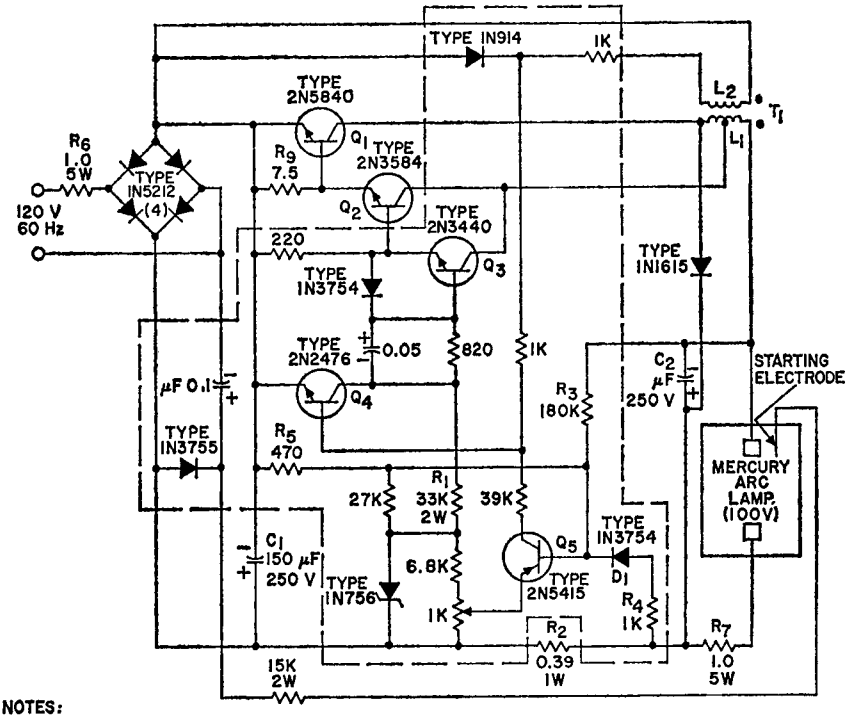
The peak current and associated output voltage of the switching-regulator circuit can be varied by adjustment of potentiometer $R_6$. For any given setting of the potentiometer, however, these quantities are constant and are independent of the input voltage. Another factor of interest, which is apparent from Eq. (362), is that a change in power level (i.e., in $V_{in}$ or $I_{peak}$) results in an inverse change in switching frequency.

Fig. 462 shows a practical 100-watt switching-regulator ballasting circuit designed for 120-volt line applications in which the output voltage and current are both sampled to reduce bulb warm-up time. This circuit has a voltage-current characteristic very similar to that shown in Fig. 455(b) for a conventional ballasting reactor.

The 120-volt ac input is rectified by a full-wave bridge rectifier. The dc output from the rectifier is developed across filter capacitor $C_1$. Because the input drive to the emitter-base circuit of the switching transistor is applied through a resistance network, the relatively high supply voltage can lead to serious $I^2R$ losses unless the drive current is maintained at a very small value. This condition is made possible by use of two transistors $Q_2$ and $Q_3$ in a Darlington configuration to provide the current gain necessary to increase the low value of drive current to the level required to saturate the switching transistor.

Because the switching regulator is a "down converter," has limited filtering, and operates from relatively low line voltages, a special low-voltage (100-volt rather than the more common 135-volt) mercury-arc lamp is used with the 100-watt, 120-volt ballasting circuit. The low-voltage arc tube contains slightly less mercury than the higher-voltage type. High starting potentials are obtained by use of a half-wave voltage doubler, wired to a separate starting electrode (with a current-limiting resistor).

Performance data of the 100-watt switching regulator are shown in Fig. 463. These data are shown as a function of the dc input voltage to filter capacitor $C_1$. The over-all efficiency of the circuit, including the rectifier




NOTES:

1. THIS CIRCUIT IS DESIGNED FOR OPERATION ABOVE 100 VRMS, WITH 250-V PEAK TRANSIENTS.
2. CONTROL CIRCUITRY, SHOWN WITHIN DOTTED LINE, MAY BE PRINTED CIRCUIT.
3. DIODE $D_1$ IS THERMALLY CONNECTED TO TRANSISTOR $Q_5$.

MERCURY ARC LAMP = 90-TO-100 VOLT TYPE WITH SEPARATE STARTING ELECTRODES.

$L_1$ = 120 TURNS OF No. 22 WIRE TAPPED 1 TURN FROM COLLECTOR

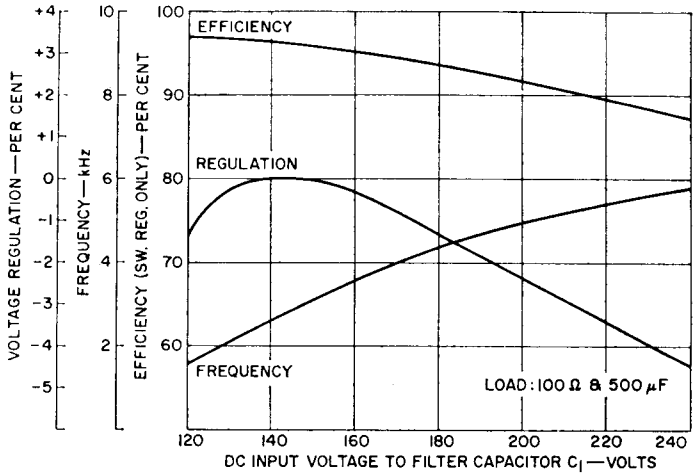
$L_2$ = 18 TURNS OF No. 34 WIRE

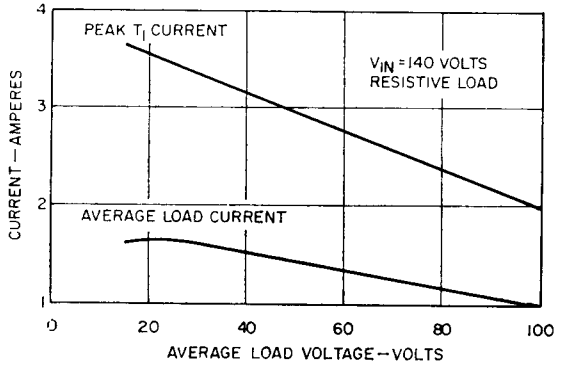
$T_1$ = ARNOLD AH 361 (OR EQUIV.) WITH 0.036" GAP. (TURNS RATIO, 6.7:1)

Figure 462. 100-watt, 120-volt ac switching-regulator ballasting circuit.

bridge and filter capacitor, is 87 per cent for a 120-volt ac input. The output is adjustable from 15 to 150 watts for operation of the circuit into a 100-ohm load impedance. The excellent regulation characteristics, shown in Fig. 463, are achieved in part, by the action of resistor $R_5$, which offsets a rise in output voltage with a corresponding rise in input voltage.

The 120-volt ballast circuit has a relatively small conduction angle, because of a necessarily large filter capacitor ($C_1$). The associated surge currents make the use of bulbs in excess of 200 watts impractical. The ballast has two 1-ohm surge-current-limiting resistors, $R_7$ and $R_{10}$. Resistor $R_{10}$ limits ac line transients; the resistor $R_7$ limits bulb current during ionization.




(b)

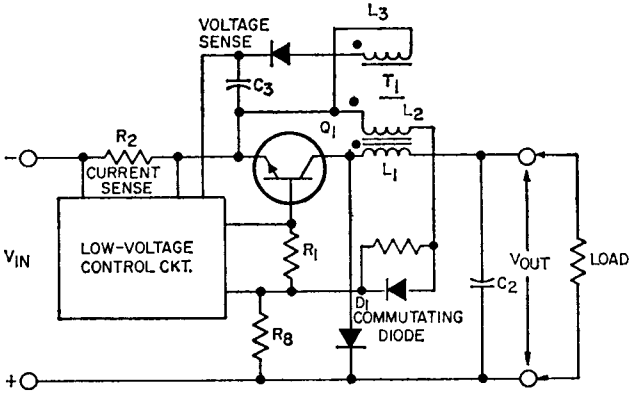
Figure 463. Performance characteristics of the 100-watt, 100-volt switching-regulator ballasting circuit: (a) voltage regulation, frequency reponse, and efficiency; (b) output characteristics.

## 200-to-300-Volt Switching-Regulator Ballasts

For industrial and highway lighting installations, 240-volt single-phase, 277-volt single-phase, and 208-volt three-phase ac power sources are readily available. For these voltages, a sufficient differential between the arc-tube voltage and input voltage exists to permit the transistor switching element to be driven from a secondary winding on the inductor of a low-pass filter. Relatively high drive currents can then be obtained without high power losses.

Fig. 464 shows the basic configuration for a switching regulator designed to operate from ac source voltages between 200 and 300 volts. Eqs. (358) through (362) and the waveshapes shown

in Fig. 461, given for the 120-volt switching-regulator ballasts, are also applicable to higher-voltage ballasts of the type shown in Fig. 464. A unique feature of the higher-voltage circuits is that only the high-current switching transistor $Q_1$ is required to have a breakdown-voltage capability sufficient to withstand the full value of the dc input voltage including transients applied across the $V_{IN}$ terminals. All the transistors in the control circuit are low-voltage, low-dissipation types. The design for the higher-voltage ballast also features built-in short-circuit protection.




Note: This circuit requires only one high-voltage switching element.

Figure 464. 200-to-300-volt ac switching regulator.

In the switching-regulator circuit shown in Fig. 464, the dc voltage applied to the $V_{IN}$ terminals drives a switching transistor ($Q_1$) that is slightly forward-biased by a small current (approximately 3 milliamperes) through a base-circuit resistor ($R_8$). Transistor $Q_1$ is immediately driven into saturation by the positive feedback from its collector circuit supplied by the $L_2$ secondary winding of transformer $T_1$. The $L_2$ secondary winding also supplies the drive power to the control circuit. The collector current of switching transistor $Q_1$ rises linearly through the $L_1$ primary winding of transformer $T_1$ until the voltage across the current-sensing resistor $R_2$ triggers the control circuit in shunt with the base-emitter junction of transistor $Q_1$. The transistor is then held cut off by the feedback voltage from the $L_2$ secondary winding of the transformer until the current through $L_1$ primary winding decays to zero. The inductive kickback that results from the decrease in current through $L_1$ is clamped by the commutating diode $D_1$ and, therefore, is the same as the output voltage on $C_2$. The $L_3$ winding of transformer $T_1$ then charges capacitor $C_3$ to a voltage proportional to the output voltage. During the next cycle, the control circuit samples a combination of the voltage across capacitor $C_3$ and the current through resistor $R_2$. In this way, an output characteristic similar to that of a

conventional ballast, shown in Fig. 455(b), is obtained.

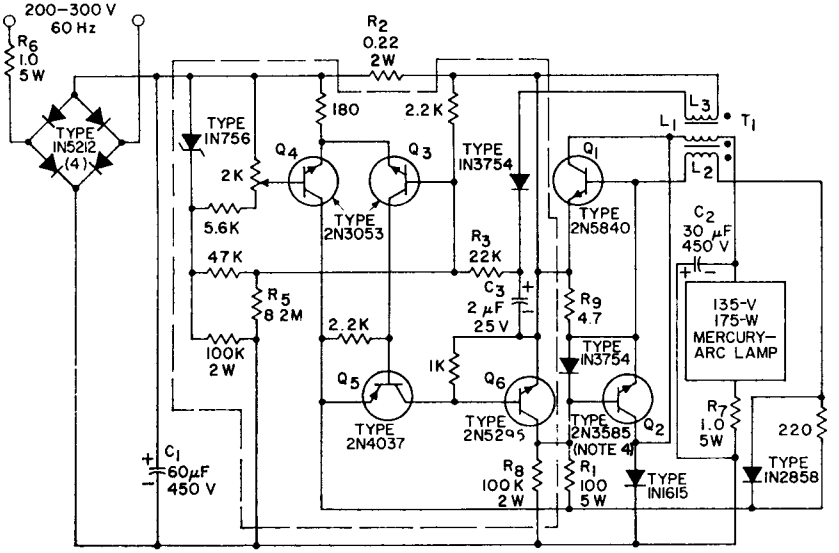
The schematic diagrams and performance data for two practical ballasting circuits, designed for use with 175-watt and 400-watt mercury-arc bulbs, that use the approach illustrated by the basic circuit configuration shown in Fig. 464 are shown in Figs. 465 and 466 and Figs. 467 and 468, respectively. Performance data are shown as a function of the dc input voltage to filter capacitor $C_1$. Excellent regulation is obtained for dc input voltages from 200 to 450 volts.

## DESIGN PROCEDURE

The design of solid-state switching-regulator ballasts for mercury-arc lamps involves three critical operations: (1) selection of the mercury-arc lamp and the peak starting current, (2) selection of the reactor element, and (3) selection of the switching transistor and other circuit components.

### Mercury-Arc Lamp and Peak Starting Current

The type of mercury-arc lamp used and the peak starting current that must be supplied to this lamp by the ballast circuit are dictated by the value of the ac source voltage, the amount of lamp power ($P_L$) required, and the warm-up time of the lamp. For operation from a 120-volt ac line at lamp power levels up to 200 watts, the special low-volt-




NOTES:

1. MAXIMUM TRANSIENT VOLTAGE = 450 V.
2. CONTROL CIRCUIT SHOWN WITHIN DOTTED LINE MAY BE PRINTED CIRCUIT.
3. TRANSISTORS $Q_3$ AND $Q_4$ ARE THERMALLY CONNECTED.
4. TRANSISTOR $Q_2$ IS SELECTED FOR A $V_{CER(SUS)}$ AT 200 OHMS GREATER THAN 500 VOLTS

$T_1$ = 2 x ARNOLD AH–108 (OR EQUIV.) WITH 0.054" AIR GAP 17:1.7:1 TURNS RATIO, 7:1

$L_1$ = 120 TURNS OF No. 22 WIRE

$L_2$ = 12 TURNS OF No. 32 WIRE

$L_3$ = 7 TURNS OF No. 32 WIRE

Figure 465. 175-watt, 200-to-300-volt switching-regulator ballasting circuit.

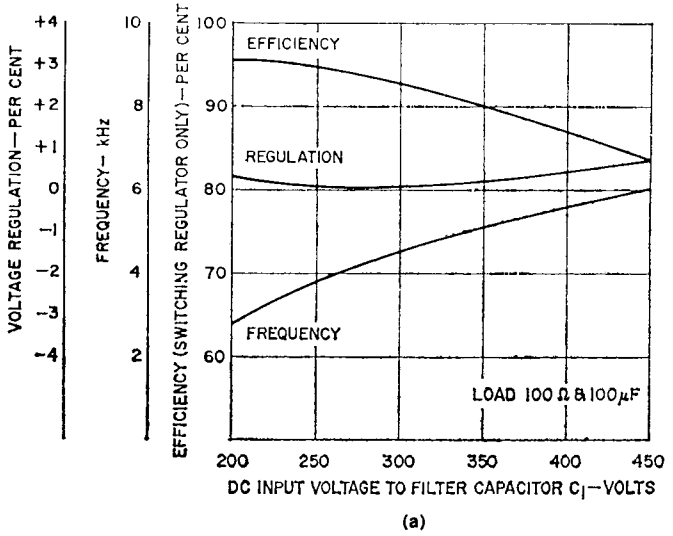
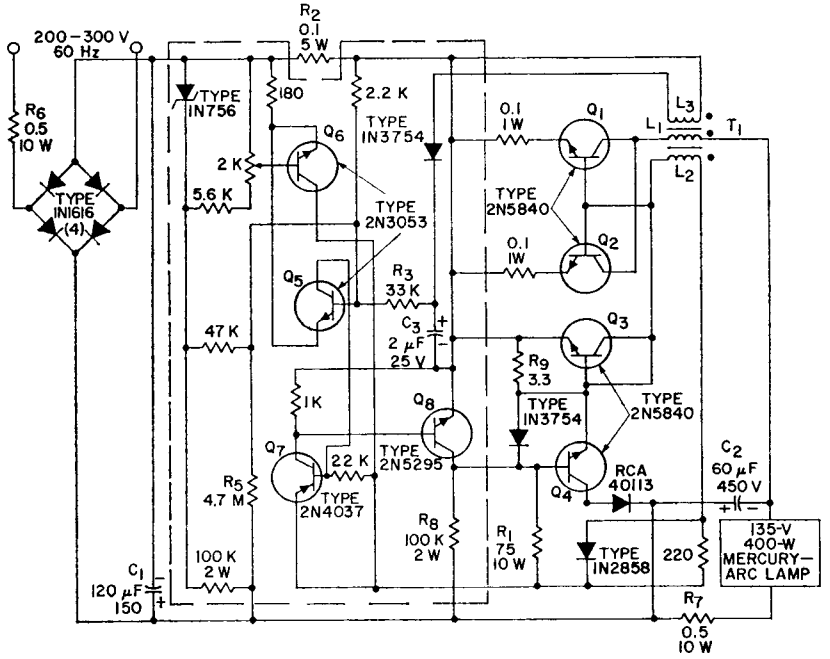



Figure 466. Performance characteristics of the 175-watt, 200-to-300-volt ballasting circuit: (a) voltage regulation, frequency response, and efficiency; (b) output characteristics.

age (90-to-100-volt) type of mercury-arc lamp should be used. The peak starting current is then determined from the following relationship:

$$I_{peak} = 4\left(\frac{P_L}{100V}\right) \qquad (363)$$

For operation from ac source voltages in the range of 200 to




NOTES:

1. MAXIMUM TRANSIENT VOLTAGE = 450 V.
2. CONTROL CIRCUIT SHOWN WITHIN DOTTED LINE MAY BE PRINTED CIRCUIT.
3. TRANSISTORS $Q_1$ THROUGH $Q_4$ ARE SELECTED TO HAVE A $V_{CER(SUS)}$ AT 20 OHMS GREATER THAN 500 V.
4. TRANSISTORS $Q_5$ AND $Q_6$ ARE THERMALLY CONNECTED.

$T_1$ = ARNOLD AH 223 (OR EQUIV.) WITH 0.125" AIR GAP 17:1. TURNS RATIO, 7:1

$L_1$ = 98 TURNS OF No. 18 WIRE.

$L_2$ = 10 TURNS OF No. 32 WIRE.

$L_3$ = 6 TURNS OF No. 32 WIRE.

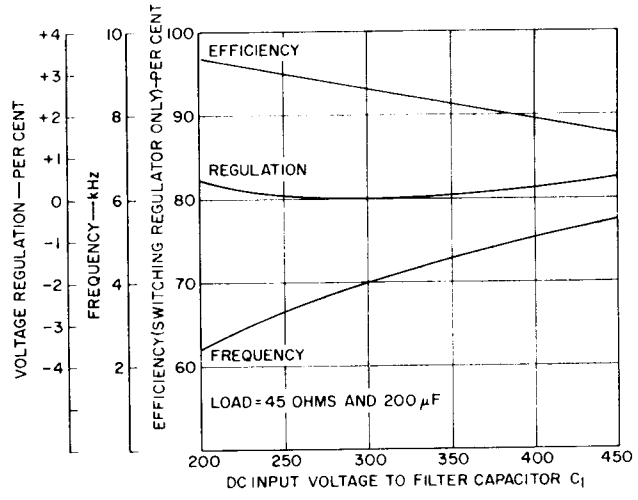
Figure 467. 400-watt, 200-to-300-volt switching-regulator ballasting circuit.

300 volts, the more conventional 135-volt type of mercury-arc lamp is used. The peak starting current, for a specified bulb power rating $P_L$, is then determined as follows:

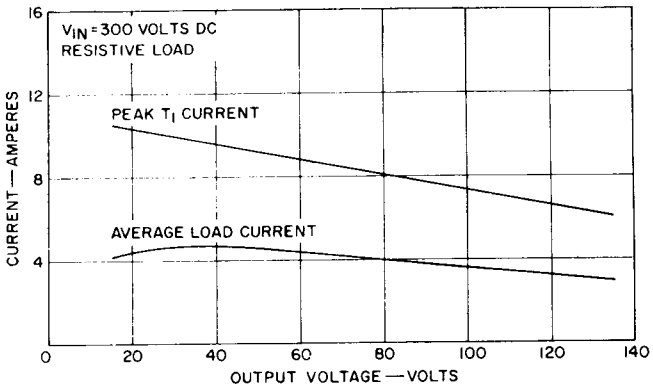
$$I_{peak} = 4\left(\frac{P_L}{135V}\right) \qquad (364)$$

## Switching-Regulator Reactor Element

The series inductor selected for the switching-regulator ballasting circuit should have a maximum core cross-sectional area and minimum air gap, consistent with the required inductance value, so that the minimum physical size is obtained. The circuit shown in Fig. 469 permits simple di/dt measurements that eliminate the need for repetitive calculations in determination of the required inductances. In this test circuit, the inductor is connected in series with a switching transistor and a dc voltage. The switching transistor is main-




(a)




(b)

Figure 468. Performance characteristics of the 400-watt, 200-to-300-volt switching-regulator ballasting circuit: (a) voltage regulation, frequency response, and efficiency; (b) output characteristics.

tained in the "on" state until the inductor saturates. The following equation then becomes the basis for the determination of the inductor parameters:

$$V_{in} = L_1 \left( \frac{I_{sat}}{t_{on}} \right) \qquad (365)$$

The desired flux density for the inductor is some fraction of that produced by the saturation current. The air gap, number of turns, and the core are selected as required to obtain the desired value. The turns ratio from the series inductor winding (primary) to the secondary windings

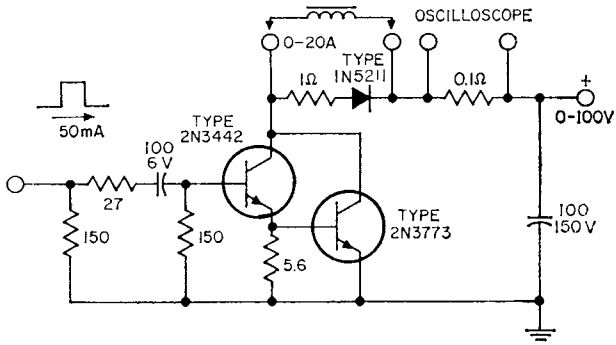



Figure 469. Inductor tester.

is as indicated in the circuit schematic (Fig. 462, 465, or 467) of the type of switching-regulator ballast being designed.

If an iron core is used for the inductor, the core laminations should be 4 mils thick (only a negligible increase in efficiency results from the use of thinner laminations). For stabilized operation and to avoid overheating of the inductor, the switching frequency of the ballasting circuit should be less than 5 kHz and the flux density in the inductor should be less than 6 kilogauss. For an inductor that uses a ferrite core, the flux density (determined for worst-case conditions) is usually 3 kilogauss, and the frequency is limited by only the transistor switching losses.

## Switching Transistor and Other Circuit Components

A switching transistor used as the switching element in a switching-regulator ballast must have a collector-to-emitter voltage-breakdown capability $V_{CER}(sus)$ high enough so that the device can withstand the total input dc voltage together with the maximum transient input voltage that may be developed in the circuit. In all the ballasting circuits described in this section, the transistor used as the high-current switching element is the type 2N5840. The specifications of this transistor are given in Table XXXI.

The Darlington transistor circuit in shunt with the emitter-base junction must drive the switching transistor well into the saturation region for the particular $I_{peak}$.

1. For the 120-volt ballast-circuit design,

$$I_B \text{ (max)} < 10 \text{ mA}$$

2. For the 200-to-300-volt ballast-circuit design,

$$I_B \text{ (max)} < 300 \text{ mA}$$

Approximately 20 per cent of the base drive to the switching transistor is diverted by resistor $R_9$ (in Figs. 462, 465, and 467) to achieve rapid turn-off of the transistor.

The power dissipated by the transistor selected for use as the switching element should not exceed 10 per cent of the power rating ($P_L$) of the mercury-arc

Table XXXI—Specifications for the Type 2N5840 Transistor

**TEST CONDITIONS AT 25°C ± 3°C**

| Parameter | $I_C$ | $R_{BE}$ | $V_{BE}$ | $V_{CE}$ | $I_B$ | L | Unit | Limit | |
|---|---|---|---|---|---|---|---|---|---|
| | A | OHMS | V | V | A | μH | | Min. | Max. |
| $V_{CER(sus)}$ | 0.2 | 50 | | | | | V | 375 | |
| $V_{CEO(sus)}$ | 0.2 | | | | | | V | 300 | |
| $I_B$ | 0.5 | | | 5.0 | | | A | 0.005 | |
| $V_{CE(sat)}$ | 3.0 | | | | 0.375 | | V | | 1.0 |
| $\theta_{J-C}$ | | | | | | | °C/W | | 1.75 |
| $I_S/_{b(1\ second)}$ | | | | 40 | | | A | 2.5 | |
| $E_S/_b$ | | 20 | -4 | | | 500 | A | 4.0 | |
| $t_{f(1)}$ | 3.0 | | | 200 | 0.375 | | μS | | 0.5 |

NOTES: 1. $I_{B1} = I_{B2} = 0.375$ A; ($h_{FE} = 8$)
2. The type 2N5840 transistor is an epitaxial-overlay switching transistor in a JEDEC TO-3 case.

bulb. The transistor power dissipation ($P_D$) is calculated for a hot, stabilized bulb ($I_{Cmax} = I_{STAB} = 2\ I_{avg}$) as follows:

$P_D$ = saturation loss + turn-off loss

$$= \frac{t_{on}}{t_{on} + t_{off}} \int_0^{I_{STAB}} i\ R_{(sat)}\ di + \frac{f_{(STAB)} V_{IN}\ t_f}{2} \left( \frac{I_{STAB}\ f}{2} \right) [t_{on}\ (I_{STAB})\ (R_{sat}) + V_{IN}\ t_f] \quad (366)$$

In Eq. 366, $R_{(sat)}$ is the saturation resistance of the switching transistor, and $t_f$ is its turn-off time for the particular circuit conditions. (It should be noted that the turn-off time is not directly related to the gain-bandwidth product $f_T$.)

The total base drive resistance of the switching-regulator ballasting circuits can be estimated on the basis of the current and voltage relationships for peak-current conditions.

1. For the 120-volt design, the voltage drop across the total of the resistors in the base drive circuit is the dc input voltage less the voltage (8.2 volts) across the 1N756 zener diode. The minimum value for the drive-circuit resistance $R_{IN}$, therefore, can be calculated by use of following equation:

$$R_{IN} = \frac{V_{IN\ (min)} - 8.2\ V}{I_{B(max)}} = \frac{100}{I_{B(max)}} \quad (367)$$

Eq. (367) indicates that the drive-circuit resistance for the 120-volt ballast design must be greater than 9000 ohms for a permissible $I_{B(max)}$ of 10 milliamperes.

2. For the 200-to-300-volt design, the total drive-circuit resistance is estimated as follows:

$$R_{in} = \frac{V_{in(min.)} - V_{out(min.)} \frac{L_2}{L_1} - 2\,V}{I_{B(max.)}} \qquad (368)$$

In this case, the drive-circuit resistance must be greater than 60 ohms for the 300 milliamperes of maximum permissible drive in the circuits presented.

The values of capacitors $C_1$ and $C_2$ and of resistors $R_2$, $R_7$, and $R_{10}$ are determined on the basis of the type of circuit being designed and the power rating of the mercury-arc lamp with which this circuit is to be used. When the lamp power rating ($P_L$) differs from that shown in the circuit diagrams of Figs. 465 and 467, the values of $C_1$, $C_2$, $1/R_2$, $1/R_7$, and $1/R_{10}$ should be increased or decreased in direct proportion to the change in the lamp power rating, i.e.,

$$\frac{C_1}{C'_1} = \frac{C_2}{C'_2} = \frac{R'_2}{R_2} = \frac{R'_7}{R_7} = \frac{R'_{10}}{R_{10}} = \frac{P_L}{P'_L} \qquad (369)$$

where the prime (') indicates the new circuit values.

The bridge-rectifier diodes and the commutation diode are selected on the basis of the maximum voltage and current requirements of the ballasting circuit.

The value of resistor $R_3$ is determined from the desired voltage-current slope of the ballast circuit.

$$VI_{(slope)} = -\frac{I_{bulb\ (hot)}}{V_{bulb\ (hot)}} \qquad (370)$$

An increase in the warm-up time for a given bulb and ballasting circuit arrangement can be achieved by the use of a larger resistor $R_3$ in both the 120-volt and 200-to-300-volt designs. This larger resistor would result in a smaller voltage-current slope, as shown in Fig. 470, and the collector current during starting ($I_{PEAK}$) would then be reduced.

The value of $R_5$ is selected to provide the best voltage regulation.

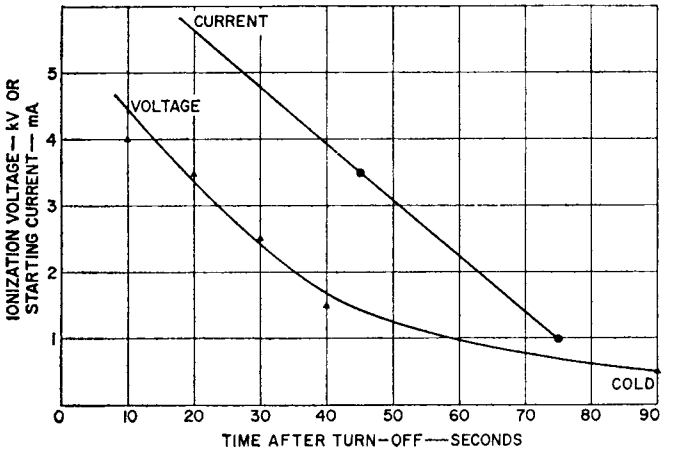



Figure 470. Typical warm-up characteristics for a 100-watt mercury-arc lamp.

## ADVANTAGES OF SOLID-STATE MERCURY-ARC-LAMP BALLASTING

The circuit configuration and design procedure for the solid-state ballasts present several noted advantages over conventional ballasts.

1. Because no strobe effect is associated with the solid-state ballasts, it is possible to use long-life, efficient mercury-arc lamps in studios and in similar critical lighting areas. In such applications, the low lighting cost and the advantage of more light with less heat are decisive factors in favor of mercury-arc lamps.

2. Solid-state ballasts provide unmatched power regulation for line-voltage fluctuations.

3. The new ballasts offer the physical advantages of reduced weight and bulk in comparison to conventional ballasts. For example, the weight of a 400-watt conventional ballast is approximately 13 pounds, while the weight of an equivalent solid-state ballast is only 2.4 pounds. It is anticipated that the weight and bulk of solid-state ballasts will be further reduced by the use of hybrid circuit techniques and ultrasonic operating frequencies.

4. A solid-state photocell control is required to switch only milliwatts of power to actuate a solid-state ballast, rather than the kilowatts that would be required for a conventional ballast.

5. The circuits permit adjustment of 70 to 150 per cent of rated bulb wattage. Outside this range, the negative-impedance characteristics of the bulb cause the arc to be extinguished. However, one basic ballast circuit may be used for bulbs of various power ratings.

6. The solid-state ballast supplies dc power to the bulb so that there are no RFI radiation problems.

In a comparison of solid-state and conventional ballasts, the initial cost factor must be considered. In regard to the initial cost, the simple magnetic ballast has a decided advantage. In general, however, initial cost is only 10 per cent of the total costs of the lighting system, and this advantage is clearly outweighed when a less efficient lighting means is displaced.

From the standpoint of reliability, proper design should result in solid-state ballasts that match the performance of conventional ballasts.

In the ballast circuit described in preceding paragraphs, only transistors were used. SCR's and triacs are also suited for use in ballasting circuits for arc-discharge lighting systems, particularly at high power levels. Significant future growth in mercury-arc lighting for both home and office should favor the transistor ballast at voltage and power levels below 120 volts and 100 watts.

# RF Power Amplifiers

POWER transistors are used in high-frequency amplifiers for military, industrial, and consumer applications. They are operated class A, B, or C, with frequency- or amplitude-modulation, single sideband or double sideband, in environments ranging from airborne to marine. The paragraphs below discuss the characteristics of rf power transistors and the circuit design techniques used in their applications.

## HIGH-FREQUENCY POWER TRANSISTORS

The increasing number of rf power transistors available today offers the circuit designer a wide selection from which to determine the optimum type for a particular application. The choice is based on factors such as maximum power output, maximum operating frequency, operating efficiency, power gain, reliability, and cost per watt of power generated. The ultimate choice of the transistors produced by any manufacturer, therefore, is dependent upon how well the devices perform in relation to these critical factors. RCA "overlay" silicon power transistors offer significant advantages for rf power applications at frequencies that extend well into the microwave region.

### Physical-Design Parameters

The exceptional high-frequency-power capabilities of the overlay power transistors result from the unique emitter construction used in these devices. At high current levels, the emitter current of a transistor is concentrated at the emitter-base edge. In overlay transistors, the size of the emitters is substantially reduced, and a large number (from 16 to several hundred) of separate emitter sites are connected in parallel. This method of construction results in the high emitter periphery-to-area ratios, and makes possible the high current-handling capabilities, low capacitances, and short transit times between emitter and collector, that are required for rf power transistors.

The overlay transistor takes its name from the emitter metallization, shown in Fig. 471, that lies

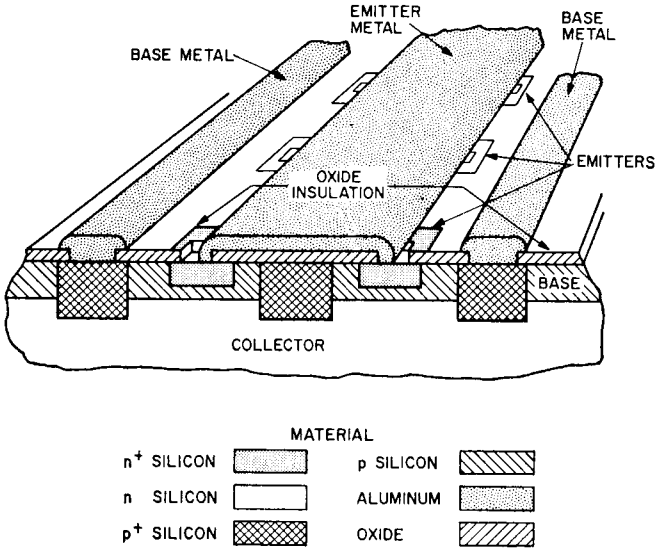



Figure 471. Top and cross-sectional view of a typical overlay transistor.

over the base instead of adjacent to it as in the interdigitated structure. The actual base and emitter areas beneath the metal pattern are insulated from one another by a silicon dioxide layer. The overlay arrangement provides a substantial increase in over-all emitter periphery without increasing the physical area of the device, and thus improves the power-frequency capability of the device.

In addition to the standard base and emitter diffusions, an added diffused region in the base serves as a conductor grid. This $p^+$ region offers three advantages: (1) it distributes base current uniformly over all the separate emitter sites, (2) it reduces distances between emitter and base, and (3) it reduces the base and contact resistances between the aluminum metallization and the silicon material.

For lower-power hf/vhf and small-signal uhf RCA transistors, an interdigitated structure is used. In this structure, as shown in Fig. 472, the emitters and bases are built like a set of interlocking combs. The sizes of the emitter and base areas are controlled by masking and diffusion. The oxide deposit, formed of silicon heated to a high temperature, masks the transistor against either an n- or p-type impurity. This oxide is removed by the usual photoetching techniques in areas where diffusion is required.

## Special Ratings Concepts

Unlike low-frequency high-power transistors, many rf devices can fail within the dissipation limits set by the classical junction-to-case thermal resistance during operation under conditions of high load VSWR, high collector supply voltage, or linear (Class A or AB) operation. Failure can be caused by hotspotting,

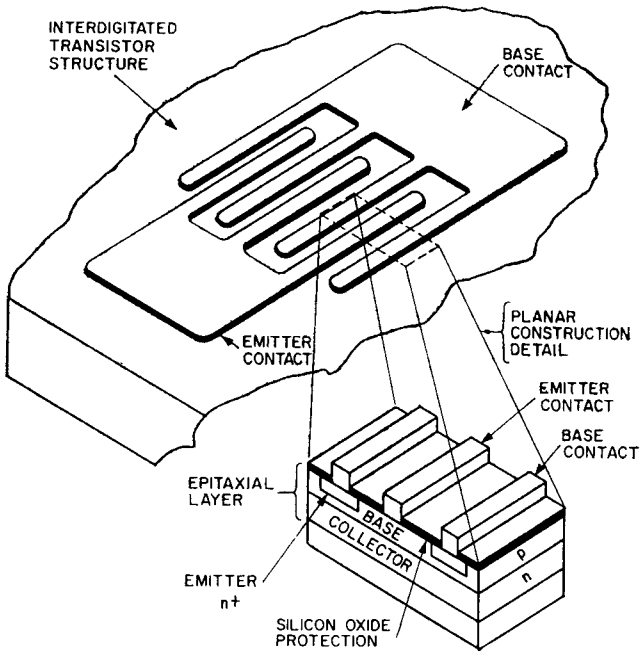



Figure 472. Top and cross-sectional view of a typical interdigitated transistor.

which results from local current concentration in the active areas of the device, and may appear as a long-term parameter degradation. Localized hotspotting can also lead to catastrophic thermal runaway.

The presence of hotspots can make virtually useless the present method of calculating junction temperature by measurements of average thermal resistance, case temperature, and power dissipation. However, by use of an infrared microscope, the spot temperature of a small portion of an rf transistor pellet can be determined accurately under actual or simulated device operating conditions. The resultant peak-temperature information is used to characterize the device thermally in terms of junction-to-case hotspot thermal resistance, $\theta_{JS\text{-}C}$.

The use of hotspot thermal resistance improves the accuracy of junction temperature and related reliability predictions, particularly for devices involved in linear or mismatch service.

**DC Safe Area**—The safe area determined by infrared techniques represents the locus of all current and voltage combinations within the maximum ratings of a device that produce a specified spot temperature (usually 200°C) at a fixed case temperature. The shape of this safe area is very similiar to the conventional safe area in that there are four regions, as shown in Fig. 473: constant current, constant power, derating power, and constant voltage.

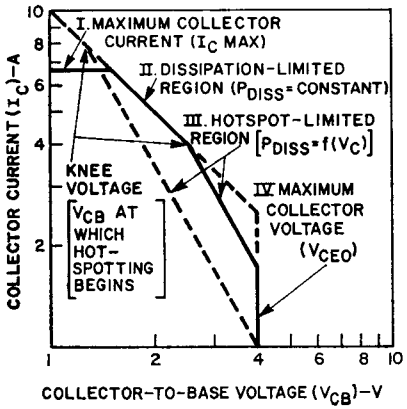



Figure 473. Safe-area curve for an rf power transistor determined by infrared techniques.

Regions I and IV, the constant-current and constant-voltage regions, respectively, are determined by the maximum collector current and $V_{CEO}$ ratings of the device. Region II is dissipation-limited; in the classical safe area curve, this region is determined by the following relationship:

$$P_{max} = \frac{T_J(max) - T_C}{\theta_{J-C}} \qquad (371)$$

where $T_C$ is the case temperature.

This relationship holds true for the infrared safe area; $P_{max}$ may be slightly lower because the reference temperature $T_{J(max)}$ is a peak value rather than an average value. The hotspot thermal resistance ($\phi_{JS-C}$) may be calculated from the infrared safe area by use of the following definition:

$$\theta_{JS-C} = \frac{T_{JS} - T_C}{P} \qquad (372)$$

where $T_{JS}$ is highest spot temperature [$T_{J(max)}$ for the safe area] and P is the dissipated power ( = I × V product in Region II).

The collector voltage at which regions II and III intersect, called the knee voltage $V_K$, indicates the collector voltage at which power constriction and resulting hotspot formation begins. For voltage levels above $V_K$, the allowable power decreases. Region III is very similar to the secondary breakdown region in the classical safe area curve except for magnitude. For many rf power transistors, the hotspot-limited region can be significantly lower than the second-breakdown locus. Generally $V_K$ decreases as the size of the device is increased.

Fig. 474 shows the temperature profiles of two transistors with identical junction geometries that operate at the same dc power

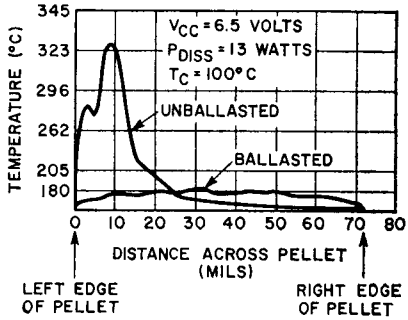



Figure 474. Thermal profiles of a ballasted and an unballasted power transistor during dc operation.

level. If devices are operated on the dissipation-limited line of their classical safe areas, the profiles show that the temperature of the unballasted device rises to values 130°C in excess of the 200°C rating. Temperatures of this magnitude, although not necessarily destructive, seriously reduce the lifetime of the device.

**Emitter Ballasting**—The profiles shown in Fig. 474 also demonstrate the effectiveness of emitter ballasting in the reduction of

power (current) constriction. In the ballasted device, a biasing resistor is introduced in series with each emitter or small groups of emitters. If one region draws too much current, it will be biased towards cutoff, allowing a redistribution of current to other areas of the device.

The amount of ballasting affects the knee voltage, $V_K$, as shown in Fig. 475. A point of diminishing returns is reached as $V_K$ approaches $V_{CEO}$.

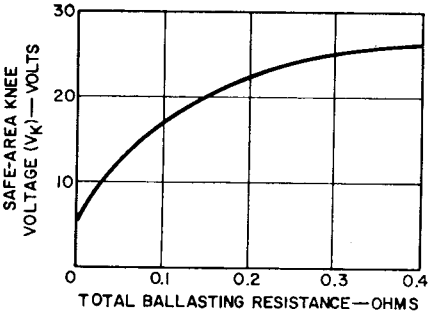



Figure 475. Safe-area voltage for an rf power transistor as a function of total ballasting resistance.

**RF Operation**—In normal class C rf operation, the hotspot thermal resistance is approximately equal to the classical average thermal resistance. If the proper collector loading (match) is maintained, $\theta_{JS-C}$ is independent of output power at values below the saturated- or slumping-power level, and is independent of collector supply voltage at values within + 30 per cent of the recommended operating level.

Power constriction in rf service normally occurs only for collector load VSWR's greater than 1.0. A transistor that has a mismatched load experiences temperatures far in excess of device ratings, as shown in Fig. 476(a)

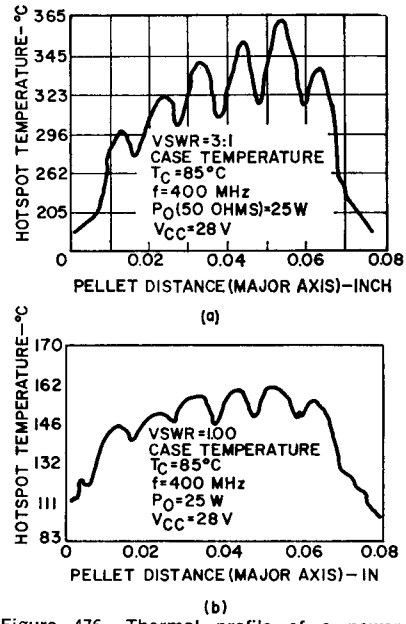



Figure 476. Thermal profile of a power transistor during rf operation: (a) under mismatched conditions; (b) under matched conditions.

for VSWR = 3.0. For comparison, the temperature profile for the matched condition is shown in Fig. 476(b).

Fig. 477 is a typical family of thermal-resistance curves that indicate the response of a device to various levels of VSWR and collector supply voltage. $\theta_{JS-C}$ responds to even slight increases

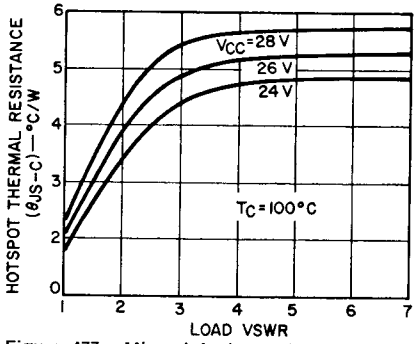



Figure 477. Mismatch-stress thermal characteristics for the 2N5071.

in VSWR above 1.0 and saturates at a VSWR in the range of 3 to 6. The saturated level increases with increasing supply voltage. Devices with high knee voltages tend to show smaller changes of $\theta_{JS-C}$ with VSWR and supply voltage. $\theta_{JS-C}$ under mismatch is independent of frequency and power level, and reaches its highest values at load angles that produce maximum collector current. Power level does, however, influence the temperature rise and probability of failure.

Device failure can also occur at a load angle that produces minimum collector current. Under this condition, collector voltage swing is near its maximum, and an avalanche breakdown can result. This mechanism is sensitive to frequency and power level, and becomes predominant at lower frequencies because of the decreasing rf-breakdown capability of the device.

Collector mismatch can be caused by the following conditions:

1. Antenna loading changes in mobile applications when the vehicle passes near a metallic structure.
2. Antenna damage.
3. Transmission-line failure because of line, connector, or switch defects.
4. Variable loading caused by nonlinear input characteristics of a following transistor (particularly broadband) or varactor stage.
5. Supply-voltage changes that reflect different load-line requirements in class C.
6. Tolerance variations on fixed-tuned or stripline circuits.
7. Matching network variations in broadband service.

**Case-Temperature Effects**—The thermal resistance of both silicon and beryllium oxide, two materials that are commonly used in rf power transistors, increases about 70 per cent as the temperature increases from 25 to 200°C. Other package materials such as steel, kovar, copper, or silver, exhibit only minor increases in thermal resistance (about 5 per cent). The over-all increase in $\theta_{JS-C}$ of a device depends on the relative amounts of these materials used in the thermal path of the device; typically the increase of $\theta_{JS-C}$ ranges from 5 per cent to 70 per cent. Fig. 478

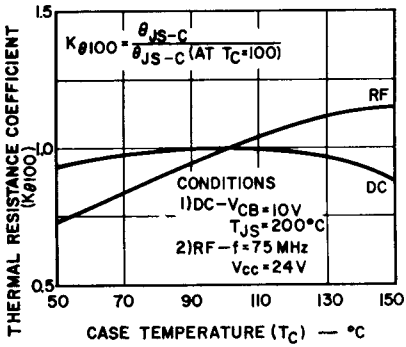



Figure 478. Thermal-resistance coefficient for the 2N5071.

shows the rf and dc thermal resistance coefficients for a typical rf transistor. For both cases, the coefficient is referenced to a 100°C case and is defined as follows:

$$K_{\theta 100} = \frac{\theta_{JS-C}}{\theta_{JS-C}\text{ at } T_C = 100°C} \quad (373)$$

The rf coefficient changes more than the dc coefficient, because of the power constriction that occurs in rf operation at elevated case temperature.

## Reliability Considerations

When the rf and thermal capabilities of a transistor have been established, the next step is to establish the reliability of the device for its actual application. The typical acceptable failure rate for transistors used in commercial equipment is 1 per cent per 1000 hours (100,000 hours MTBF); for transistors used in military and high-reliability equipment, it is 0.01 to 0.1 per cent per 1000 hours. Because it is not practical to test transistors under actual use conditions, dc or other stress tests are normally used to simulate rf stresses encountered in class B or class C circuits at the operating frequencies. Information derived from these tests is then used to predict the failure rate for the end-use equipment. The tests used to assure reliability include high-temperature storage tests, dc and rf operating life tests, dc stress step tests, burn-in, temperature cycling, relative humidity, and high-humidity reverse bias. The end-point measurement for these tests should include collector-to-emitter voltage $V_{CEO}$ and emitter-to-base voltage $V_{EBO}$ in addition to the common end-point collector-to-emitter current $I_{CEO}$, collector-to-base voltage $V_{CBO}$, collector-to-emitter saturation voltage $V_{CE}(sat)$, power output, and power gain.

One of the common failure modes in rf power transistors is degradation of the emitter-to-base junction. The high-temperature storage life test and the dc and rf operating life tests can accelerate this failure mode, and it can be detected by measurement of $V_{EBO}$.

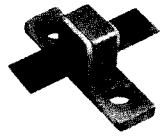
Plastic uhf power transistors are more sensitive to emitter-to-base-junction degradation than similar hermetic devices. The enhancement of this failure mode in plastic devices can be caused by moisture penetration into the very close geometries used in uhf power transistors. Thermal fatigue is also a problem that affects the reliability of uhf plastic power transistors because large thermal-expansion differences exist between the plastic encapsulant and the fine bonding wires (usually 1 mil) used in the devices.

## Packages

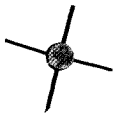
The package is an integral part of an rf power transistor. A suitable package for rf applications should have good thermal properties and low parasitic reactances. Package parasitic inductances and resistive losses have significant effects on such circuit performance characteristics as power gain, bandwidth, and stability. The most critical parastics are the emitter and base lead inductances. Table XXXII gives the inductances of some of the more important commercially available rf power-transistor packages. Photographs of the packages are shown in Fig. 479. The TO-60 and TO-39 packages were first used in devices such as the 2N3375 and the 2N3866. The base and emitter parastic inductance for both TO-60 and TO-39 packages is in the order of 3 nanohenries; this inductance represents a reactance of 7.5 ohms at 400 MHz. If the emitter is grounded internally in a TO-60 package (as in the RCA-2N5016), the emitter

**HF-21**
Hermetic
Ceramic-Metal
Coaxial Package,
Large
(JEDEC TO-201AA)

**HF-11**
Coaxial Package, Small
(JEDEC TO-215AA)



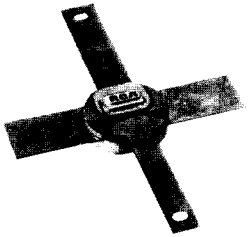
**HF-28**
Hermetic
Ceramic-Metal
Stripline Package
Grounded emitter
or base



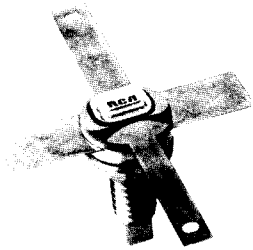
**HF-33**
Isolated
Electrodes

**HF-32**
Hermetic
Stripline Package


**JEDEC TO-39**

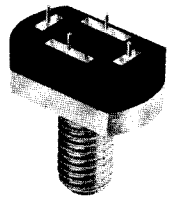


**HF-31**
Hermetic
Ceramic-Metal
Stripline Package
(Studless JEDEC TO-216AA)



**HF-19**
Hermetic Strip-Line Type
Ceramic-to-Metal Package
with Isolated Electrodes
(JEDEC TO-216AA)


**JEDEC TO-60**



**HF-12**
Molded Silicone-Plastic Case
(JEDEC TO-217AA)

Figure 479. Commercially available rf power transistor packages.

Table XXXII—Summary of RCA Transistor Packages

| PACKAGE | APPROXIMATE INDUCTANCE (nH) | UPPER FREQUENCY OF OPERATION (MHz) |
|---|---|---|
| TO-39 | 3 | 500 |
| TO-60 (isolated emitter) | 3 | 400 |
| TO-60 (internally grounded emitter) | 0.6 | 500 |
| UHF HERMETIC STRIPLINE | | |
| HF-19 (STUD) = JEDEC TO-216AA | 0.5 | 1000 |
| HF-31 = STUDLESS JEDEC TO-216AA | 0.5 | 1000 |
| HF-32 (FLANGED) | 0.5 | 1000 |
| MICROWAVE HERMETIC STRIPLINE | | |
| HF-28 (FLANGED) | 0.2 | 2500 |
| COAXIAL HERMETIC | | |
| HF-11 = JEDEC TO-215AA | 0.1 | 3000 |
| HF-21 = JEDEC TO-201AA | 0.2 | 2500 |

lead inductance is reduced to 0.6 nanohenry.

Hermetic low-inductance radial-lead packages are also available. The HF-19 package introduced by RCA for the 2N5919 utilizes ceramic-to-metal hermetic seals, has isolated electrodes, and has rf performance comparable to an rf plastic package. This package is also available in a studless version (HF-31) for miniaturized or low-power applications and in a grounded-emitter, flanged version (HF-32) for compact applications.

Low-parastic, hermetic packages are available for microwave applications. The HF-11, a medium-power, hermetic coaxial package first used for the RCA-2N5470, employs ceramic-to-metal construction and has parasitic inductances in the order of 0.1 nanohenry. A larger, higher-power version, the HF-21, uses the same constructional techniques and has parasitic inductances in the range of 0.2 nanohenry. The stripline equivalent of this package is the HF-28 and has approximately the same parasitic reactances as the HF-21.

Table XXXIII compares the performance of the TO-39 package, the HF-19 hermetric stripline package, and the HF-11 coaxial package with the same transistor chip. At a frequency of 1 GHz and an input power of 0.3 watt, the coaxial package performs significantly better than either the stripline or the TO-39 package. The coaxial package results in twice as much output power as the TO-39 package. In

Table XXXIII—Package Performances with Same Transistor Chip

| | f-GHz | $P_{in}$-W | $P_O$-W | P.G.-dB | $\eta c$(28V)-% |
|---|---|---|---|---|---|
| TO-39 | 1 | 0.3 | 1 | 5 | 35 |
| HF-19 | 1 | 0.3 | 1.5 | 7 | 45 |
| HF-11 | 1 | 0.3 | 2.2 | 8.6 | 50 |
| HF-11 | 2 | 0.3 | 1 | 5 | 35 |

addition, the coaxial-package transistor is capable of delivering an output of more than 1 watt with a gain of 5 dB at 2 GHz.

## DESIGN CONSIDERATIONS FOR RF POWER AMPLIFIERS

In the design of silicon-transistor rf power amplifiers for use in transmitting systems, several fundamental factors must be considered. As with any rf power amplifier, the class of operation has an important bearing on the power output, linearity, and operating efficiency. The matching characteristics of input and output terminations significantly affect power output and frequency stability and, therefore, are particularly important considerations in the design of transistor power amplifiers. The selection of the proper transistor for a given circuit application is also a major consideration, and the circuit designer must realize the significance of the various transistor parameters to make a valid evaluation of different types.

### Class of Operation

The class of operation of an rf amplifier is determined by the circuit performance required in the given applications. Class A power amplifiers are used when extremely good linearity is required. Although power gain in this class of service is considerably higher than that in class B or class C service, the operating efficiency of a class A power amplifier is usually only about 25 per cent. Moreover, the standby drain and thermal dissipation of a class A stage are high, and care must be exercised to assure thermal stability.

In applications that require good linearity, such as single-sideband transmitters, class B push-pull operation is usually employed because the transistor dissipation and standby drain are usually much smaller and operating efficiency is higher. Class B operation is characterized by a collector conduction angle of 180 degrees. This conduction is obtained by use of only a slight amount of forward bias in the transistor stage. In this class of service, care must be taken to avoid thermal runaway.

In a class C transistor stage, the collector conduction angle is less than 180 degrees. The gain of the class C stage is less than that of a class A or class B stage, but is entirely usable. In addition, in the class C stage, standby drain is virtually zero, and circuit efficiency is the highest of the three classes. Because of the high efficiency, low collector dissipation, and negligible standby drain, class C operation is the most commonly used mode in rf power-transistor applications.

For class C operation, the base-to-emitter junction of the transistor must be reverse-biased so that the collector quiescent current is zero during zero-signal conditions. Fig. 480 shows four methods that may be used to reverse-bias a transistor stage.

Fig. 480(a) shows the use of a dc supply to establish the reverse bias. This method, although effective, requires a separate supply, which may not be available or may be difficult to obtain in many applications. In addition, the bypass elements required for the separate supply increase the circuit complexity.

Figs. 480(b) and 480(c) show

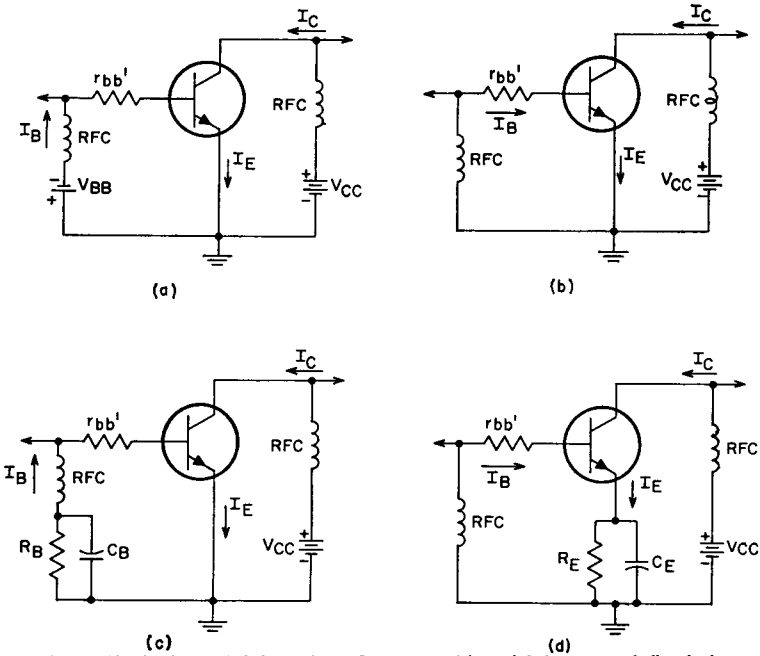



Figure 480. Methods for obtaining class C reverse bias: (a) by use of fixed dc supply $V_{BB}$; (b) by use of dc base current through the base spreading resistance $r_{bb'}$; (c) by use of dc base current through an external base resistance $R_B$; (d) by use of self bias developed across an emitter resistor $R_E$.

methods in which reverse bias is developed by the flow of dc base current through a resistance. In the case shown in Fig. 480(b), bias is developed across the base spreading resistance. The magnitude of this bias is small and uncontrollable because of the variation in $r_{bb'}$ among different transistors. A better approach, shown in Fig. 480(c), is to develop the bias across an external resistor $R_B$. Although the bias level is predictable and repeatable, the size of $R_B$ must be carefully chosen to avoid reduction of the collector-to-emitter breakdown voltage.

The best reverse-bias method is illustrated in Fig. 480(d). In this method, self-bias is developed across an emitter resistor $R_E$. Because no external base resistance is added, the collector-to-emitter breakdown voltage is not affected. An additional advantage of this approach is that stage current may be monitored by measurement of the voltage drop across $R_E$. This technique is very helpful in balancing the shared power in paralleled stages. The bias resistor $R_E$ must be bypassed to provide a very-low-impedance rf path to ground at the operating frequency to prevent degeneration of stage gain. In practice, emitter bypassing is difficult and frequently requires the use of a few capacitors in parallel to reduce the series inductance in the capacitor leads and body. Alternatively, the lead-inductance problem may be solved by formation of a self-resonant series circuit between the capacitor and its

leads at the operating frequency. This method is extremely effective, but may restrict stage bandwidth.

## Modulation (AM, FM, SSB)

Amplitude modulation of the collector supply of a transistor output stage does not result in full modulation. During down-modulation, a portion of the rf drive feeds through the transistor. Better modulation characteristics can be obtained by modulation of the supply to at least the last two stages in the transmitter chain. On the downward modulation swing, drive from the preceding modulated stages is reduced, and less feed-through power in the output results. Flattening of the rf output during up-modulation is reduced because of the increased drive from the modulated lower-level stages.

The modulated stages must be operated at half their normal voltage levels to avoid high collector-voltage swings that may exceed transistor collector-to-emitter breakdown ratings. RF stability of the modulated stages should be checked for the entire excursion of the modulating signal.

Amplitude modulation of transistor transmitters may also be obtained by modulation of the lower-level stages and operation of the higher-level stages in a linear mode. The lower efficiencies and higher heat dissipation of the linear stages override any advantages that are derived from the reduced audio-drive requirements; as a result, this approach is not economically practical.

Frequency modulation involves a shift of carrier frequency only. Carrier deviations are usually very small and present no problems in amplifier bandwidth. For example, maximum carrier deviations in the 50-MHz and 150-MHz mobile bands are only 5 kHz. Because there is no amplitude variation, class C rf transistor stages have no problems handling frequency modulation.

Single-sideband (SSB) modulation requires that all stages after the modulator operate in a linear mode to avoid intermodulation-distortion products near the carrier frequency. In many SSB applications, channel spacing is close, and excessive distortion results in adjacent-channel interference. Distortion is effectively reduced by class B operation of the rf stages, with close attention to biasing the transistor base-to-emitter junction in a near-linear region.

## Characterization of Large-Signal RF Power Transistors

The values of large-signal transistor parameters, such as the S and Y parameters, are different from those of small-signal transistors because (1) the values of transistor parameters change with power levels, and (2) the harmonic-frequency components that exist in a large-signal rf power amplifier must be considered in addition to the fundamental-frequency sinusoidal component in a small-signal amplifier. RF power-transistor characteristics are normally specified for a given circuit in a specific application.

The design of rf power-amplifier circuits involves the determination of dynamic input and load impedances. Before the input circuit is designed, the input

impedance at the emitter-to-base terminals of the packaged transistor must be known at the drive-power frequency. Before the output circuit is designed, the load impedance presented to the collector terminal must be known at the fundamental frequency. These dynamic impedances are difficult to calculate at microwave frequencies because transistor parameters such as $S_{11}$ and $S_{22}$ vary considerably under large-signal operation and also change with the power level. Small-signal equations that might serve as useful guides for transistor design cannot be applied rigorously to large-signal circuits. Because large-signal representation of rf power transistors has not yet been developed, transistor dynamic impedances are best determined experimentally with slotted-line or vector voltmeter measurement techniques.

The system used for determination of transistor impedances under operating conditions is shown in Fig. 481. This system consists of a well-padded power signal generator, a directional coupler (or reflectometer) for monitoring the input reflected power, an input triple-stub tuner, an input low-impedance line section, the transistor holder (or test jig), an output line section, a bias tee, an output triple-stub tuner, another directional coupler for monitoring the output waveform or frequency, and an output power meter. For a given frequency and input power level, the input and output tuners are adjusted for maximum power output and minimum input reflected power. Once the system has been properly tuned, the impedance across terminals 1-1 (with the transistor disconnected) is measured at the same frequency in a slotted-line set-up or with the vector voltmeter. The conjugate of this impedance is the dynamic input impedance of the transistor. Similarly, the impedance across terminals 2-2 (with the transistor disconnected) is the collector-load impedance presented to the transistor collector. Such measurements are performed at each frequency and power level. It should be noted that the circuit

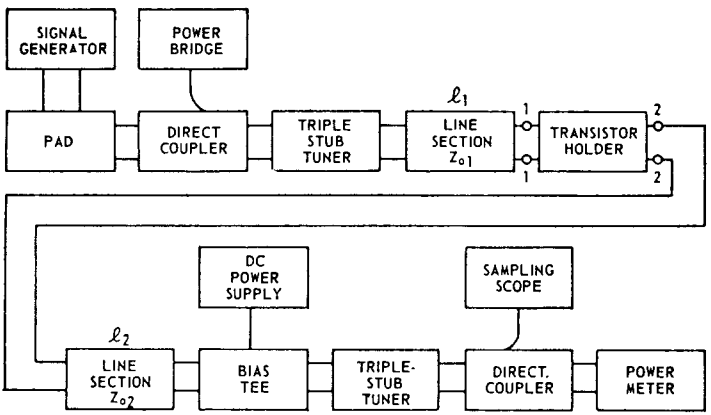



Figure 481. Set-up for measurement of rf transistor dynamic impedances.

arrangement of Fig. 481 is also useful for testing the performance of the transistor. Thus, power output, power gain, and efficiency are readily determined.

## Matching Requirements

A simplified high-frequency equivalent circuit of an "overlay" type of transistor is shown in Fig. 482. This circuit is similar to the hybrid-pi equivalent circuit

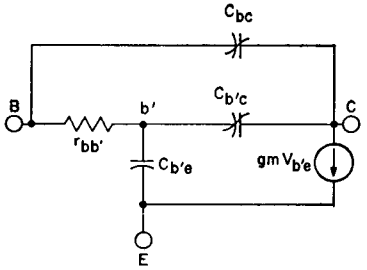



Figure 482. Simplified high-frequency equivalent circuit for an "overlay" transistor.

of a transistor except for the addition of the capacitance $C_{bc}$. This capacitance represents the high collector-to-base capacitance in the overlay transistor which is created by the large area of the collector-to-base junction together with the active area under the emitter. This capacitance and the capacitance $C_{b'c}$ vary nonlinearly with the collector-to-emitter voltage.

Maximum performance in a transistor rf amplifier can be obtained only if the base and collector terminals are properly terminated. The input network generally is required to match a 50-ohm source to the relatively low base-to-emitter impedance, which includes approximately 1 to 10 ohms of resistance and some series reactance. The output network must match a resistive component and the transistor output capacitance to a load impedance, which is generally about 50 ohms. In most applications, the output network also acts as a band-rejection filter to eliminate unwanted frequency components that may be included in the collector waveform. The filter presents a high impedance to these unwanted frequencies and also increases collector efficiency. The power output and collector-voltage swing determine the resistive component to be presented to the collector. The design and form of the output networks (resonant circuits for narrow-band operation or transmission lines for broad-band operation) are discussed in a later section.

## Multiple Connection of Power Transistors

Many applications require more rf power than a single transistor can supply. The parallel approach is the most widely used method for multiple connection of power transistors.

In parallel operation of transistors, steps must be taken to assure equal rf and thermal load sharing. In one approach, the transistors are connected directly in parallel. This approach, however, is not very practical from an economic standpoint because it requires the use of transistors that are exactly matched in efficiencies, power gains, terminal impedances, and thermal resistances. A more practical approach is to employ signal splitting in the input matching network. By use of adjustable components in each leg, adequate compensation can be made for variations in power gains and input impedances to assure equal load sharing between the transistors. For applications in which low supply voltages are used and high power outputs are desired, the

output impedance of the rf amplifier is very low. For this reason, it is beneficial, in the interest of paralleling efficiency, to split the collector loads. By use of separate collector coils, the power outputs may be combined at higher impedance levels at which the effect of any asymmetry introduced by lead inductances is insignificant and resistive losses are less. The use of separate collector coils also permits individual collector currents to be monitored.

## Transistor Selection

In selection of a transistor and circuit configuration for an rf power amplifier, the designer should be familiar with the following transistor and circuit characteristics:

(1) maximum transistor dissipation and derating,
(2) maximum collector current,
(3) maximum collector voltage,
(4) input and output impedance characteristics,
(5) high-frequency current-gain figure of merit ($f_T$),
(6) operational parameters such as efficiency, usable power output, power gain, and load-pulling capability.

Proper cooling must be provided to prevent destruction of the transistor because of overheating. Transistor dissipation and derating information reflect how well the heat generated within the transistor can be removed. This factor is determined by the junction-to-case thermal resistance of the transistor. A good rf power transistor is characterized by a low junction-to-case thermal resistance.

The current gain of an rf transistor varies approximately inversely with emitter current at high emitter-current levels. Peak collector current may be determined by the allowable amount of gain degradation at high frequencies. For applications in which amplitude modulation or low supply voltages are involved, peak current-handling capabilities are very important criteria to good performance.

The maximum collector voltage rating must be high enough so that junction breakdown does not occur under conditions of large collector voltage swing. The large voltage swing is produced under conditions of amplitude modulation or reactive loading because of load mismatch and circuit tuning operations.

Before the proper matching networks for an rf amplifier can be designed, transistor impedance (or admittance) characteristics at the expected operating conditions of the circuit must be known. It is important that the value and dependence of transistor impedances on collector current, supply voltage, and operating frequency be defined.

The term $f_T$ defines the frequency at which the current gain of a device is unity. This parameter is essential to the determination of the power-gain performance of an rf transistor at a particular frequency. Because $f_T$ is current-dependent, it normally decreases at very high emitter currents. Therefore, it should be determined at the operating current levels of the circuit. A high $f_T$ at high emitter or collector current levels characterizes a good rf transistor.

The operational parameters of an rf transistor can be considered to be those measured during the performance of a given circuit in which this type of transistor is

used. The information displayed by these parameters is of a direct and practical interest. Operating efficiencies can normally be expected to vary between 30 and 80 per cent. Whenever possible, a circuit should employ transistors that have operational parameters specified at or near the operating conditions of the circuit so that comparisons can be made.

In some rf power applications, such as mobile radio, the transistors must withstand adverse conditions because high SWR's are produced by faulty transmission cables or antennas. The ability of a transistor to survive these faults is sometimes referred to as load-pulling or mismatch capability, and depends on transistor breakdown characteristics as well as circuit design. The load-pulling effects that the transistor may be subjected to can be determined by replacement of the rf load with a shorted stub and movement of the short through a half wavelength at the operating frequency. Dissipation capabilities of a transistor subjected to load pulling must be higher than normal to handle the additional device dissipation created by the mismatch.

## Circuit Considerations

Frequency stability is an important consideration in the design of high-frequency transistor circuits. Most instabilities occur at frequencies well below the frequency of operation because of the increased gain at lower frequencies. With the gain increasing at 6 dB per octave, any parasitic low-frequency resonant loop can set the circuit into oscillation. Such parasitic oscillations can result in possible destruction of the transistor. These low-frequency loops can usually be traced to inadequate bypassing of power-supply leads, circuit component self-resonances, or rf choke resonances with circuit or transistor capacitances. Supply bypassing can be effected by use of two capacitors, one for the operating frequency and another for the lower frequencies. For amplifiers operated in the 25-to-70-MHz range, sintered-electrode tantalum capacitors can provide excellent bypassing at all frequencies of concern. At uhf and higher frequencies, these capacitors may be lossy and therefore not effective for bypassing. High-Q ceramic bypass capacitors are better suited for uhf use. RF chokes, when used, should be low-Q types and should be kept as small as possible to reduce circuit gain at lower frequencies. Chokes of the ferrite-bead variety have been used very successfully as base chokes. Collector rf chokes can be avoided by use of a coil in the matching network to apply dc to the collector.

Because of the variation of transistor parameters with changes in collector voltage and current, the stability of an rf transistor stage should be checked under all expected conditions of supply voltage, drive level, source mismatch, load mismatch, and, in the case of amplitude modulation, modulation swing.

Parametric oscillation is another form of instability that can occur in rf circuits that use power transistors. The transistor collector-to-base capacitance, as stated previously, is nonlinear and can cause oscillations that appear as low-level spurious frequencies not related to the carrier frequency.

Careful selection of components is necessary to obtain good performance in an rf transistor circuit. The components should be checked with an impedance bridge for parasitic impedances and self-resonances. When parasitic elements are encountered, their possible detrimental effects on circuit performance should be determined. This procedure helps the designer select coils and capacitances with low losses and high self-resonances (capacitors of the "bypass feedthrough" or "mica postage stamp" variety can have very high self-resonances). Resistors used in rf current paths should have low series inductance and shunt capacitance (generally, low-wattage carbon resistors are quite acceptable).

Circuit layout and construction are also important for good performance. Chassis should be of a high-conductivity material such as copper or aluminum. Copper is sometimes preferable because of its higher conductivity and the fact that components can be soldered directly to the chassis. Another chassis approach now becoming popular is the use of double-side laminated printed-circuit boards. The circuit, in this approach, may be arranged so that all the conductors are on one side of the board. The opposite-side foil is then employed as an additional shield. Whenever possible, the chassis should be designed on a single plane to reduce chassis inductance and to minimize unwanted ground currents.

It must be remembered that, at rf frequencies, any conductor has an inductive and resistive impedance that can be significant when compared to other circuit impedances in a transistor amplifier. It follows, therefore, that wiring should be as direct and short as possible. It is also helpful to connect all grounds in a small area to prevent chassis inductance from causing common-impedance gain degeneration in the emitter circuit. Busses or straps may be used, but it should be remembered that these items have some inductance and that the point at which a component is connected to a buss can affect the circuit.

Coils used in input and output matching networks should be oriented to prevent unwanted coupling. In some applications, such as high-gain stages, coil orientation alone is not enough to prevent instability or strange tuning characteristics, and additional shielding between base and collector circuits must be used.

In common-emitter circuits, stage gain is very dependent on the impedance in series with the emitter. Even very small amounts of inductive degeneration can drastically reduce circuit gain at high frequencies. Although emitter degeneration results in better stability, it should be kept as low as possible to provide good gain and to reduce tuning interaction and feedback between output and input circuits. The emitters of many rf power transistors are internally connected to the case so that the lowest possible emitter-lead inductance is achieved. This technique substantially reduces the problems encountered when the transistor is fastened directly to the chassis. If a transistor with a separate emitter lead is used, every attempt should be made to provide a low-inductance connection to the chassis, even to the point of connecting the chassis directly to the lead (or pin) as close to the transistor body as practicable. In extreme cases, emitter

tuning by series resonating of the emitter-lead inductance is employed.

Another important area of concern involves the removal of heat generated by the transistor. Adequate thermal-dissipation capabilities must be provided; in the case of lower-power devices, the chassis itself may be used. Finned heat sinks and other means of increasing radiator area are used with higher-power devices. Consideration must also be given to ambient variations and mismatch conditions during tuning operations or load pulling, when transistor dissipation can increase. Under such conditions, the thermal resistance of the transistor may be the limiting factor, and may dictate either a change to another device of lower thermal resistance or a parallel mode of operation using the existing transistor.

## MATCHING NETWORKS

Matching networks for rf amplifiers perform two important functions. First, they transform impedance levels as required by the active and fixed elements (e.g., transistor output to antenna impedance). Second, they provide frequency discrimination by virtue of the "quality factor" (Q) of the resonant circuit, transform harmonic energy into desired output-frequency energy, and prevent the presence of undesired frequency components in the output.

### Design Objectives of Matching Circuits

The design of matching circuits is based on the following requirements:

(1) desired or actual network output impedance specified by the series reactance $X_s$ or shunt conductance $G_p$ and shunt susceptance $B_p$;

(2) desired or actual network input impedance specified by $R_s$ and $X_s$ or $G_p$ and $B_p$;

(3) loaded circuit Q calculated with input and output terminations connected.

The usual approach is to use L, T, or twin-T matching pads or tuned-transformer networks. More sophisticated systems may use exponential lines and balun transformers.

**Input Matching**—In practically all power-transistor stages, the input circuit must provide a match between a source impedance that is high compared to the transistor input impedance and the transistor input. When several stages are used, both the input and output impedance of a driver stage are usually higher than those of the following stage.

In most good rf transistors, the real part of the input impedance is usually low, in the order of a few tenths of an ohm to several ohms. In a given transistor family, the resistive part of the common-emitter input impedance is always inversely proportional to the area of the transistor and. therefore, is inversely proportional to the power-output capability of the transistor, if equal emitter inductances are assumed.

The reactive part of the input impedance is a function of the transistor package inductance, as well as the input capacitance of the transistor itself. When the capacitive reactance is smaller than the inductive reactance, low-frequency feedback to the base may be excessive. It is not uncommon to use an inductive input for high-power large-area transistors

because the input reactance is a series combination of the package lead inductance and the input capacitance of the transistor itself. Thus, at low frequencies, the input is capacitive, and at higher frequencies, it becomes inductive. At some single frequency, it is entirely resistive.

**Output Matching** — Although maximum power gain is obtained under matched conditions, a mismatch may be required to meet other requirements. Under some conditions, a mismatch may be necessary to obtain the required selectivity. In power amplifiers, the load impedance presented to the collector, $R_L$, is not made equal to the output resistance of the transistor. Instead, the value of $R_L$ is dictated by the required power output and the peak dc collector voltage. The peak ac voltage is always less than the supply voltage because of the rf saturation voltage. The collector load resistance $R_L$ may be expressed as follows:

$$R_L = (V_{CC})^2/2P_o \quad (374)$$

Designs for tuned, untuned, narrow-band high-Q, and broad-band coupling networks are considered later under specific applications. In some cases, particularly mobile and aircraft transmitters, considerations for safe operation must include variations in the load, both in magnitude and phase. Safe-operation considerations may include protective circuits or actual test specifications imposed on the transistor to assure safe operation under the worst-load conditions.

## Network Design

The basic components to be considered in the design of matching networks are shown in Fig. 483. For the input matching network, the source is assumed to be a generator that has a 50-ohm impedance. For the output matching

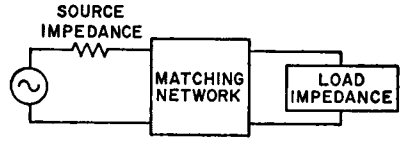



Figure 483. Basic components considered in the design of a matching network.

network, the source is the output of the transistor, which can be approximated as shown in Fig. 484.

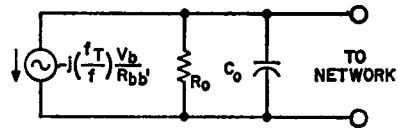



Figure 484. Equivalent circuit for the output of a transistor.

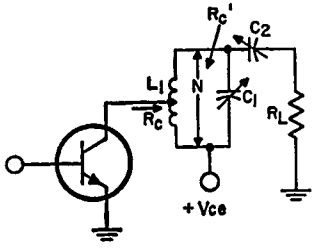
**Output-Circuit Design**—When the dc supply voltage and power output are specified, the circuit designer must determine the load for the collector circuit [$R_L = (V_{CE})^2/2P_o$]. Because an rf power amplifier is usually designed to amplify a specific frequency or band of frequencies, tuned circuits are normally used as coupling networks. The choice of the output tuned circuit must be made with due regard to proper load matching and good tuned-circuit efficiency.

As a result of the large dynamic voltage and current swings in a class C rf power amplifier, the collector current contains a large amount of harmonics. This effect is caused primarily by the nonlinearity in the transfer characteristics of the transistor. The

tuned coupling networks selected must offer a relatively high impedance to these harmonic currents and a low impedance to the fundamental current.

Class C rf power amplifiers are reverse-biased beyond collector-current cutoff; harmonic currents are generated in the collector which are comparable in amplitude to the fundamental component. However, if the impedance of the tuned circuit is sufficiently high at the harmonic frequencies, the amplitude of the harmonic currents is reduced and the contribution of these harmonic currents to the average current flowing in the collector is minimized. The collector power dissipation is therefore reduced, and the collector-circuit output efficiency is increased.

Figs. 485 and 486 illustrate the use of parallel tuned circuits to couple the load to the collector circuit. The collector electrode of the transistor is tapped down on the output coil. Capacitor $C_1$ provides tuning for the fundamental frequency, and capacitor $C_2$ provides load matching of $R_L$ to the tuned circuit. The transformed $R_L$ across the entire tuned circuit is stepped down to match the collector by the proper turns ratio of the coil $L_1$. If the value of the inductance $L_1$ is chosen properly and the portion of the output-coil inductance between the collector and ground is sufficiently high, the harmonic portion of the collector current in the tuned circuit is small. Therefore, the contribution of the harmonic current to the dc component of current in the circuit is minimized. The use of a tapped-down connection of the collector to the coil maintains the loaded Q of the circuit and minimizes variation in the bandwidth of the output circuit with changes in the output capacitance of the transistor.




FOR N:1 TURN RATIO

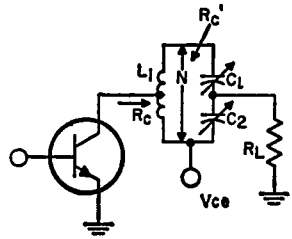
(1) $R_c = \frac{V_{ce}^2}{2P_O}$ (FOR CLASS C)

(2) $X_{L1} = \frac{R_c'}{Q_L} = \frac{N^2 R_c}{Q_L}$

(3) $X_{c2} = R_L\sqrt{\frac{N^2R_c}{R_L} - 1}$

(4) $X_{c1} = \frac{N^2R_c}{Q_L} \cdot \frac{1}{\left(1 - \frac{X_{C2}}{Q_L R_L}\right)}$

Figure 485. Tuned-circuit output coupling method and design equations in which output is transferred to load by a series coupling capacitor.




FOR N:1 TURN RATIO

(1) $R_c = \frac{V_{ce}^2}{2P_O}$ (FOR CLASS C)

(2) $X_{L1} = \frac{N^2R_c}{Q_L}$

(3) $X_{c1} = \frac{N^2R_cQ_L}{(Q_L^2+1)}\left[1 - \frac{R_L}{Q_L X_{C2}}\right]$

(4) $X_{c2} = \frac{R_L}{\sqrt{\frac{(Q_L^2+1)R_L}{N^2R_c} - 1}}$

Figure 486. Tuned-circuit output coupling method and design equations in which output to the load is obtained from a capacitive voltage divider.

Although the circuits shown in Figs. 485 and 486 provide coupling of the load to the collector circuit with good harmonic-current suppression, the tuned-circuit networks have a serious limitation at very high frequencies. Because of the poor coefficient of coupling in coils at very high frequencies, the tap position is usually established empirically so that proper collector loading is achieved. Fig. 487 shows several suitable output coupling networks that provide the required collector loading and also suppress the circulation of collector harmonic currents. These networks are not dependent upon coupling coefficient for load-impedance transformation.

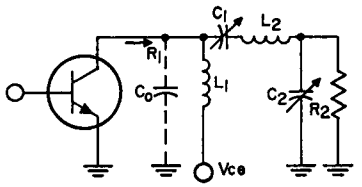
The collector output capacitance for the networks shown in Fig. 487 is included in the design equations. The collector output capacitance of a transistor varies considerably with the large dynamic swing of the collector-to-emitter voltage and is dependent upon both the collector supply voltage and the power output.

**Input-Circuit Design**—The input circuit of most transistors can be represented by a resistor $r_{bb}'$ in series with a capacitor $C_{in}$. The input network must tune out the capacitance $C_{in}$ and provide a purely resistive load to the collector of the driver stage. Fig. 488 shows several networks capable of coupling the base to the output of the driver stage and tuning out the input capacitance $C_{in}$. In the event that the transistor used has an inductive input, the reactance $X_{C_i}$ is made equal to zero, and the base inductance is included as part of inductor $L_1$ for networks such as that shown in Fig. 488(a) and is included as part of $L_2$ for networks of the type shown in Fig. 488(c). In Fig. 488(a), the input circuit is formed by the T network consisting of $C_1$, $C_2$, and $L_1$. If the value of the inductance $L_1$ is chosen so that its reactance is much greater than that of $C_{in}$, series tuning of the base-to-emitter circuit is obtained by $L_1$ and the parallel combination of $C_2$ and $(C_1 + C_o)$. Capacitors $C_1$ and $C_o$ provide the impedance matching of the resultant input resistance $r_{bb}'$ to the collector of the driving stage. Fig. 488(b) shows a T network in which the location of $L_1$ and $C_2$ is chosen so that the reactance of the capacitor is much greater than that of $C_{in}$; $C_2$ can then be used to step up $r_{bb}'$ to an appropriate value across $L_1$. The resultant parallel resistance across $L_1$ is transformed to the required collector load value by capacitors $C_1$ and $C_o$. Parallel resonance of the circuit is obtained by $L_1$ and the parallel combination $(C_1 + C_o)$ and $C_2$.

The circuits shown in Fig. 488(a) and 488(c) require the collector of the driving transistor to be shunt-fed by a high-impedance rf choke. Fig. 488(c) shows a coupling network that eliminates the need for a choke. In this circuit, the collector of the driving transistor is parallel tuned, and the base-to-emitter junction of the output transistor is series tuned. Fig. 489 shows several other forms of coupling networks that can be used in rf power-amplifier designs.

## The Impedance-Admittance Chart

One of the most useful tools for designing matching networks is the impedance-admittance chart. This chart can be described simply



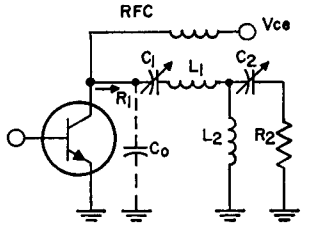

(a)

FOR $R_1 < R_2$

(1) $X_{C_1} = Q_L R_1$ (2) $X_{C_2} = \dfrac{R_2}{\sqrt{\dfrac{R_2(Q_L^2+1)}{R_1 Q_L^2}} - 1}$

(3) $X_{L_1} = \dfrac{Q_L R_1}{\left[\dfrac{Q_L R_1}{X_{co}} + 1\right]}$

(4) $X_{L_2} = Q_L R_1 \left[1 + \dfrac{R_2}{Q_L X_{C_2}}\right]$




(b)

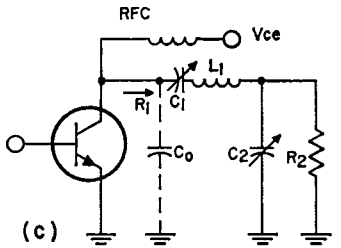
FOR $\dfrac{Q_L X_{co}}{\sqrt{R_1 R_2}} > 1$

(1) $X_{L_1} = \dfrac{Q_L X_{co}^2}{R_1}\left[1 - \dfrac{\sqrt{R_1 R_2}}{Q_2 X_{co}}\right]$

(2) $X_{L_2} = X_{co}\sqrt{R_2/R_1}$

(3) $X_{C_1} = \dfrac{Q_L X_{co}^2}{R_1}\left[1 - \dfrac{R_1}{Q_L X_{co}}\right]$

(4) $X_{C_2} = \dfrac{R_2}{Q_L}\left[\dfrac{Q_L X_{co}}{\sqrt{R_1 R_2}} - 1\right]$

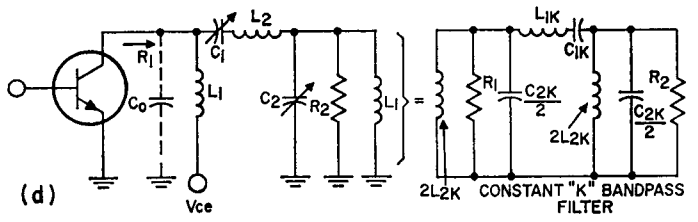



(c)

(1) $X_{C_1} = \dfrac{Q_L X_{co}^2}{R_1}\left[1 - \dfrac{R_1}{Q_L X_{co}}\right]$

(2) $X_{C_2} = \dfrac{R_2}{\sqrt{\dfrac{(Q_L^2+1)}{Q_L^2}\dfrac{R_1 R_2}{X_{co}^2}} - 1}$

(3) $X_{L_1} = \dfrac{Q_L X_{co}^2}{R_1}\left[1 + \dfrac{R_2}{Q_L X_{C_2}}\right]$



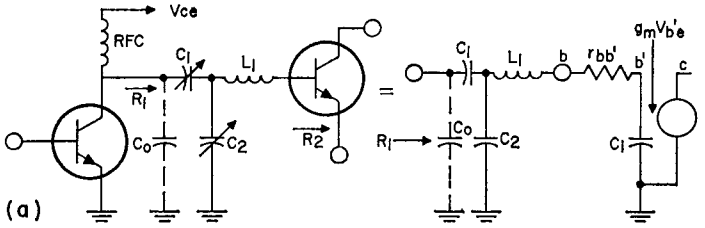

(d)

LET $C_{2K} = 2\ C_{OUT}$; $R_1 = R_2$; $f_1$ = LOW FREQ. CUTOFF; $f_2$ = HI-FREQ. CUTOFF

(1) $(f_2 - f_1) = \dfrac{1}{2\pi C_0 R_L}$ (2) $L_2 = L_{IK}\dfrac{R_1}{\pi(f_2 - f_1)}$ (3) $L_1 = 2L_{2K} = \dfrac{(f_2 - f_1)R_1}{2\pi f_1 f_2}$

(4) $C_1 = C_{IK} = \dfrac{f_2 - f_1}{4\pi f_1 f_2 R_1}$ (5) $C_2 = C_0 = \dfrac{C_{2K}}{2}$

Figure 487. Additional transistor output-coupling networks including transistor output capacitance.

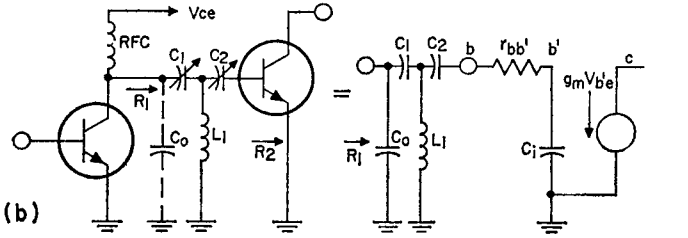



FOR $X_{L_1} >> X_{c_1}$; $R_1 > R_2 = r_{bb'}$

(1) $X_{L_1} = Q_L R_2 = Q_L r_{bb'}$

(2) $X_{c_1} = X_{co}\left[\sqrt{\frac{(Q_L^2+1)\, r_{bb'}}{R_1}} - 1\right]$

(3) $X_{c_2} = \frac{r_{bb'}(Q_L^2+1)}{Q_L} \cdot \frac{1}{\left[1-\sqrt{\frac{R_1 r_{bb'}(Q_L^2+1)}{X_{co}^2 Q_L^2}}\right]}$

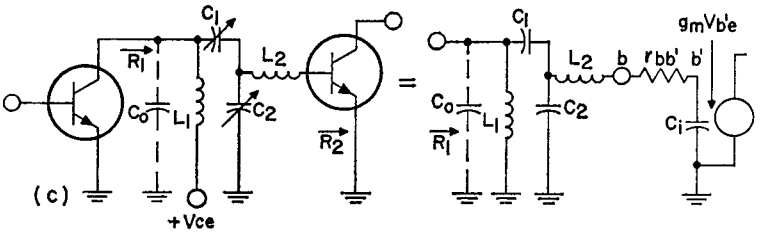



FOR $X_{c_2} >> X_{c_1}$; $R_1 > R_2 = r_{bb'}$

(1) $X_{L_1} = \frac{R_2(Q_L^2+1)}{Q_L} \cdot \frac{1}{\left[1+\sqrt{\frac{R_1 R_2}{X_{co}^2} \cdot \frac{(Q_L^2+1)}{Q_L^2}}\right]} = \frac{r_{bb'}(Q_L^2+1)}{Q_L} \cdot \frac{1}{\left[1+\sqrt{\frac{R_1 r_{bb'}}{X_{co}^2} \frac{(Q_L^2+1)}{Q_L^2}}\right]}$

(2) $X_{c_1} = X_{co}\left[\sqrt{\frac{r_{bb'}(Q_L^2+1)}{R_1}} - 1\right]$

(3) $X_{c_2} = Q_L R_2 = Q_L r_{bb'}$




FOR $R_1 > R_2$; $R_2 = r_{bb'}$; $X_{L_2} \gg X_{c_1}$

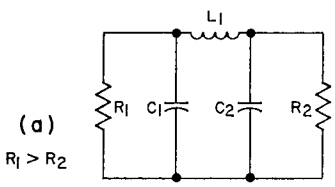
(1) $X_{L_1} = \frac{R_1}{Q_L}$

(2) $X_{L_2} = \frac{R_2}{Q_L} \cdot \frac{\left[\sqrt{\frac{R_1}{R_2}} - 1\right]}{\left[1 - \frac{R_1}{Q_L X_{co}}\right]}$

(3) $X_{c_1} = \frac{R_1}{Q_L} \cdot \frac{\left[1-\sqrt{\frac{R_2}{R_1}}\right]}{1 - \frac{R_1}{Q_L X_{co}}}$

(4) $X_{c_2} = \frac{R_1}{Q_L} \cdot \frac{\sqrt{\frac{R_2}{R_1}}}{\left[1 - \frac{R_1}{Q_L X_{co}}\right]}$

Figure 488. Transistor input-circuit coupling networks.



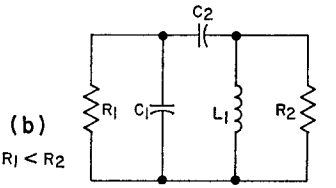

(a)

$R_1 > R_2$

(1) $X_{C_1} = \frac{R_1}{Q_L}$

(2) $X_{C_2} = \frac{R_2}{\sqrt{\frac{R_2}{R_1}(Q_L^2+1)-1}}$

(3) $X_{L_1} = \frac{Q_L R_1}{Q_L^2+1}\left[1+\frac{R_2}{Q_L X_{C_2}}\right]$



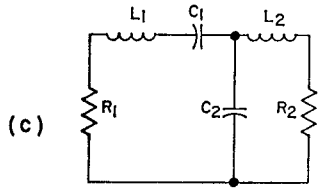

(b)

$R_1 < R_2$

(1) $X_{C_1} = \frac{R_1}{Q_L}$

(2) $X_{C_2} = \frac{Q_L R_1}{(Q_L^2+1)}\left[\frac{R_2}{Q_L X_{L_1}} - 1\right]$

(3) $X_{L_1} = \frac{R_2}{\sqrt{\frac{R_2}{R_1}(Q_L^2+1)-1}}$



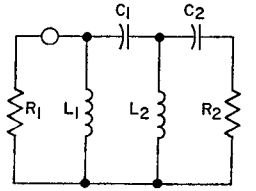

(c)

(1) $X_{L_1} = Q_L R_1$

(2) $X_{L_2} = \frac{R_2}{Q_L}\left[\sqrt{\frac{R_1(Q_L^2+1)}{R_2}} - 1\right]$

(3) $X_{C_1} = \frac{R_1(Q_L^2+1)}{Q_L}\left[1-\sqrt{\frac{R_2}{R_1(Q_L^2+1)}}\right]$

(4) $X_{C_2} = \frac{R_1}{Q_L}\sqrt{\frac{R_2(Q_L^2+1)}{R_1}}$




(d)

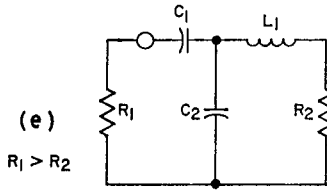
$R_1 > R_2$

(1) $X_{L_1} = \frac{Q_L R_1}{\left[1+\sqrt{\frac{R_1}{R_2}}\right]}$

(2) $X_{L_2} = \frac{Q_L R_2}{\left[1+\sqrt{\frac{R_2}{R_1}}\right]}$

(3) $X_{C_1} = Q_L R_1\sqrt{\frac{R_2}{R_1}}$

(4) $X_{C_2} = Q_L R_2$



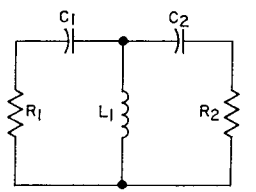

(e)

$R_1 > R_2$

(1) $X_{L_1} = Q_L R_2$

(2) $X_{C_1} = R_1\sqrt{\frac{R_2(Q_L^2+1)}{R_1} - 1}$

(3) $X_{C_2} = \frac{R_2(Q_L^2+1)}{Q_L} \cdot \frac{1}{\left[1-\frac{X_{C_1}}{Q_L R_1}\right]}$




(f)

(1) $X_{C_1} = R_1\sqrt{\frac{R_2(Q_L^2+1)}{R_1} - 1}$

(2) $X_{C_2} = Q_L R_2$

(3) $X_{L_1} = \frac{R_2(Q_L^2+1)}{Q_L} \cdot \frac{1}{\left[1+\frac{X_{C_1}}{Q_L R_1}\right]}$

Figure 489. Other suitable rf-amplifier coupling networks for maximum power transfer.

as the plane of reflection coefficient for admittances, and provides an easier and faster method of circuit analysis than that offered by rectangular admittance or impedance charts. The chart displays graphically all ladder-type matching networks, and shows the applicable tuning ranges for variable components. Lumped-component values for a given frequency may be determined directly from the chart in normalized values. The chart can be used for idealized equivalent circuits, as well as for circuits that employ transformers or tapped coils.

Fig. 490 shows the basic layout of the chart. Shunt elements in a ladder follow the admittance circles (shown dotted). Values of shunt elements correspond to values on the intersection arcs. Series elements follow the impedance circles; corresponding values are read from corresponding intersection arcs.

**Rules for Plotting Networks and Components**—When a single component L, C, or R is added to a known impedance, one of the following parameters does not change: resistance (R), reactance (X), conductance (G), or susceptance (B). (Non-ideal components must be divided into two separate ideal components; e.g., a lossy inductor into separate L and R

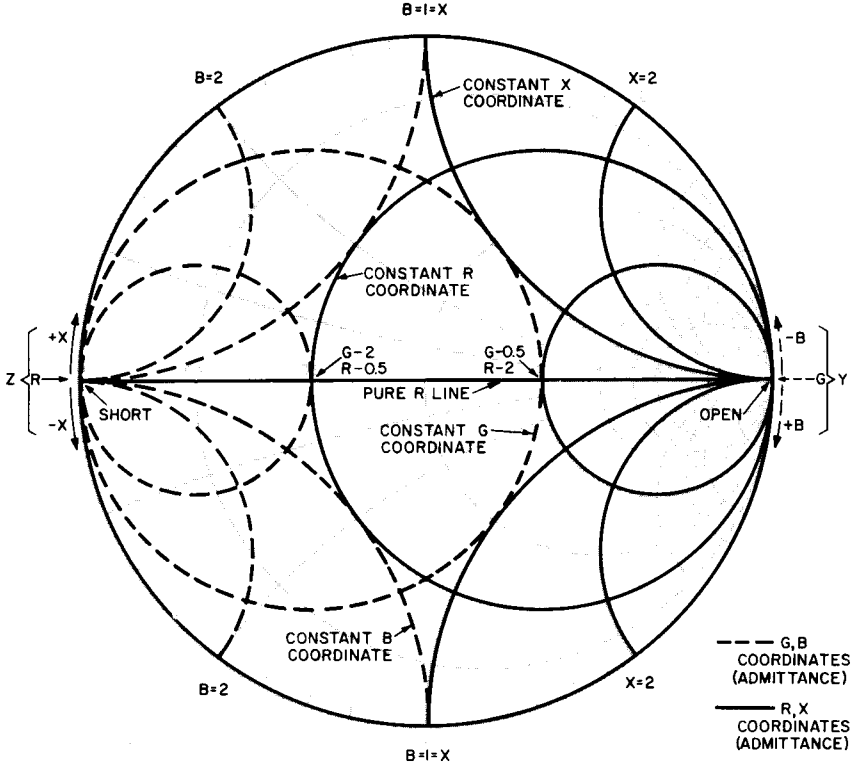



Figure 490. Impedance-admittance chart.

components.) Therefore, the component follows that constant-parameter curve. For example, an inductor added in series with the circuit does not change the series resistance curve. The procedure for each type of component is listed in Table XXXIV, which, together with Fig. 491, indicates the direction of travel along the curve and makes it unnecessary to determine the plus or minus sign

Table XXXIV—Procedure for Plotting Component Values on Impedance-Admittance Chart

| To Add | Use Chart | Follow a Curve of | Direction | Component Value |
|---|---|---|---|---|
| Series L | Z | Constant Series R | CW | $X_L = X_f - X_i$ |
| Series C | Z | Constant Series R | CCW | $X_C = X_f - X_i$ |
| Series R | Z | Constant X | toward open | $R_S = R_f - R_i$ |
| Shunt + L | Y | Constant Parallel R (G) | CCW | $B_L = B_f - B_i$ |
| Shunt + C | Y | Constant Parallel R (G) | CW | $B_C = B_f - B_i$ |
| Shunt + R | Y | Constant B | toward short | $1/R_p = G_f - G_i$ |

Calculate the change in X, B, G and R by disregarding the + and — signs of the points on the chart. However, be sure to measure the entire change in X, R, B, or G. For example, a series capacitor which changes $X_i = 0.4$ inductive (above pure R line) to $X_i = 0.3$ capacitive (below pure R line) has a value of 0.7.

**Note:** Shunt refers to components with one terminal grounded and in parallel with the rest of the network.

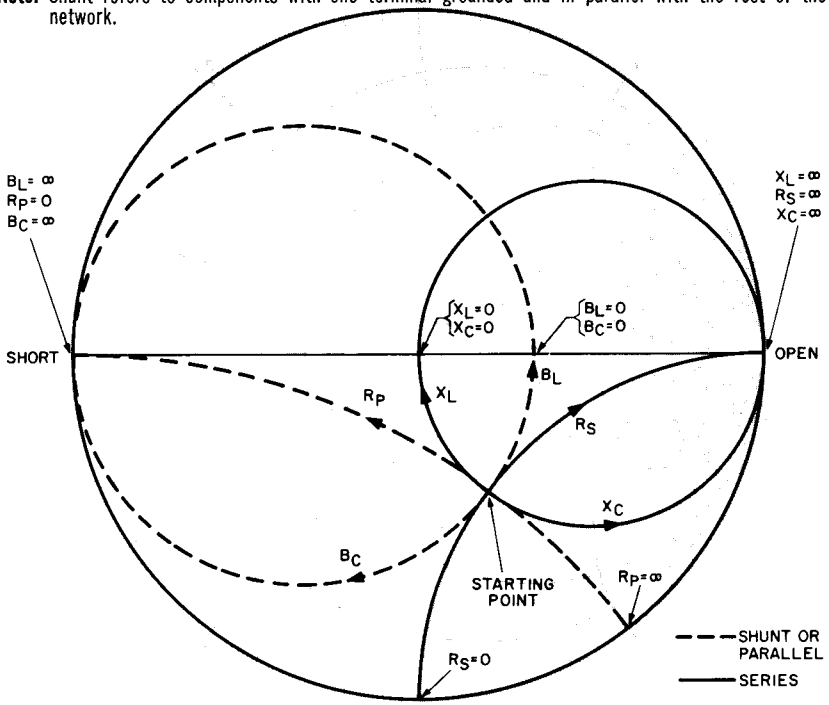



Figure 491. Method used to trace constant-parameter curves for matching-network components.

on the reactances and susceptances.

**Quality Factor, Q**—The operating Q must be specified, together with the input and output impedances, in the design of a matching network. The magnitude of the operating Q is a compromise between efficiency and harmonic rejection.

Unfortunately, the exact operating Q of a complex circuit cannot always be determined by calculations at a single frequency. When circuit design equations are used, this problem is circumvented by defining an operating Q which is easily calculated and which approximates the actual Q. The graphical technique uses the same type of approximation, but more simply and more visibly. The Q of each node of the circuit plot is determined by the constant-Q curves shown in Fig. 492. The node that has the highest Q dominates; this Q is then defined as the operating Q of the circuit.

**Normalized Values**—Impedance charts use normalized values. This graphical technique requires that normalized impedance and admittance values be consistent. The examples use $1\Omega = 1\upsilon = [50\Omega]$ (for impedances) $= [(1/50\upsilon]$ (for admittances). (Note: Brackets are used here, and in succeeding text and illustrations, to indicate the actual impedances or

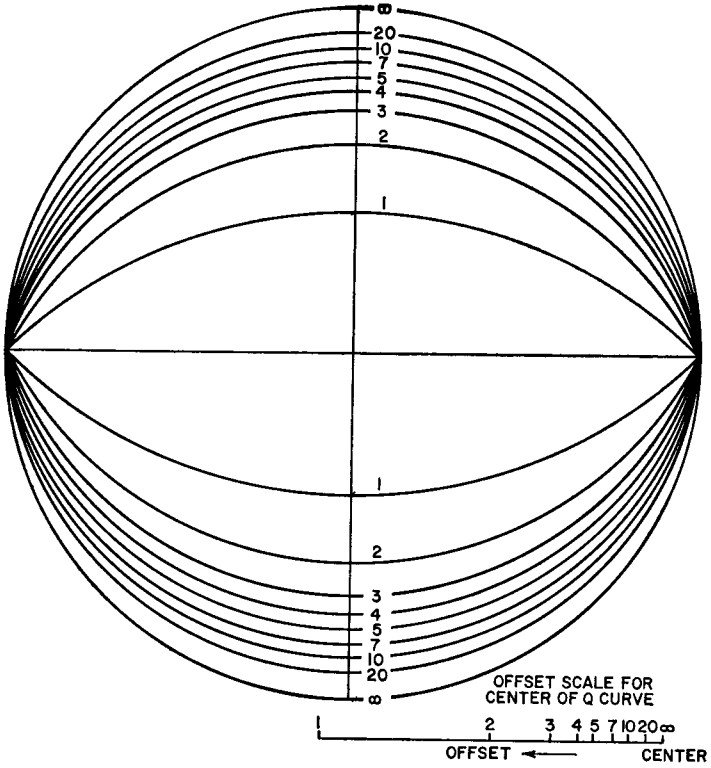



Figure 492. Chart of constant-Q curves.

admittances represented by the normalized values.) The ohm (Ω) and mho (℧) symbols are retained on the chart to distinguish between impedance and susceptance. A 50-ohm normalizing factor is used because this value represents a common rf-amplifier load impedance.

**Mapping Technique** — The matching network can be designed or analyzed by use of a network map, which is prepared by plotting each component (including input and output impedances) on the impedance chart. Dual impedance-admittance charts, such as that shown in Fig. 490, are available for this purpose, but the many curves required make these charts difficult to read. A more practical network map is prepared on tracing paper. The tracing paper is placed over either an impedance or an admittance chart to trace curves, read values, and compare impedances. Figs. 493 and 494 show simplified versions of a standard impedance chart and a standard admittance chart, respectively. The admittance coordinates have the same shape as the impedance coordinates except that they have been rotated 180 degrees around the chart. This statement can be easily verified by superposition of Fig. 493 on Fig. 494 with the short and open points of one chart aligned with the open and short points, respectively, of the other chart.

The first step in the preparation of a network map is to trace the perimeter of the impedance chart (line of pure R) and those stand-

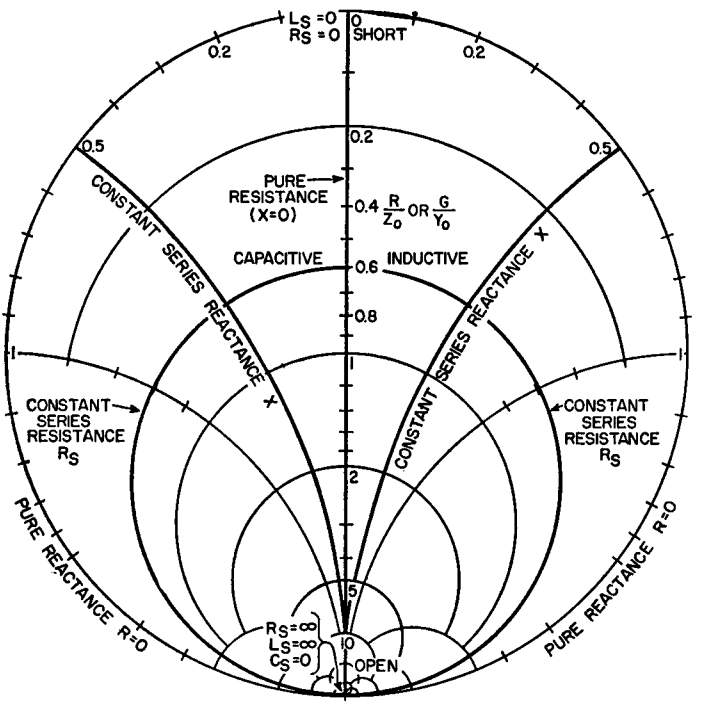



Figure 493. Impedance chart.

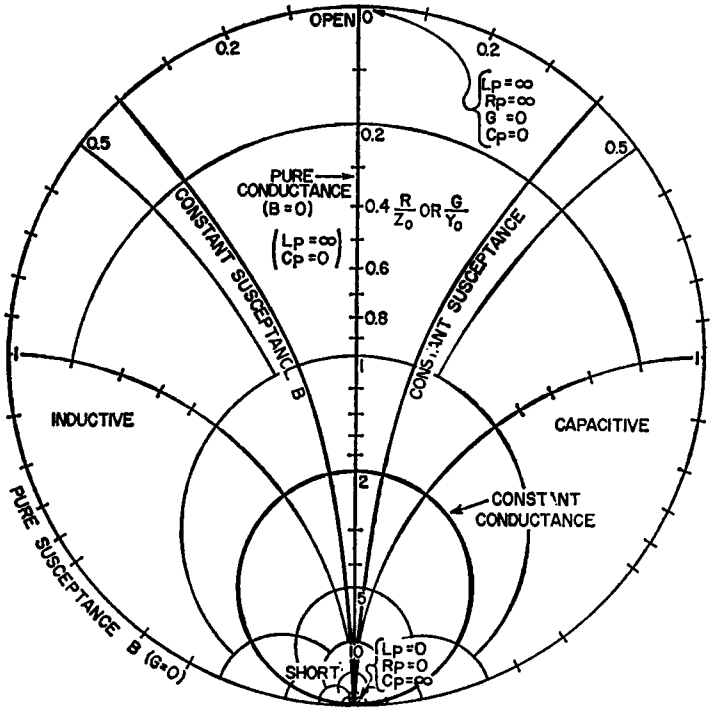



Figure 494. Admittance chart.

ard R, G, X, and B curves which are absolutely necessary. The "open" and the "short" points (or the pure R line) should also be marked to assist in accurate realignment of the tracing paper. The values of components may be determined with sufficient accuracy from curves that are traced from impedance or admittance charts placed under the tracing paper.

The following numerical examples illustrate the use of the mapping technique in the determination of the parameters of various types of matching networks.

**Determination of input impedance:** Fig. 495 shows a typical output matching network, together with the network map used to determine the required input impedance for this network. The reactance of each component in the matching network is known (from component values and operating frequency) and the input impedance is to be determined. The component reactances are plotted by use of curves traced from the Z, Y, and Q charts as follows: The output load impedance, 50 ohms normalized to 1 ohm, is located on the Z chart. Next, the series $C_2$ curve is plotted on a constant-R curve through 1 ohm, as indicated in Fig. 495. The series $C_2$ curve must change the reactance by its given value, 100 ohms normalized to 2 ohms, and the required normalized constant X = 2 ohms curve is traced from

the Z chart. The $C_2$ curve ends at point A, where B = 0.04 mho, and the shunt L curve begins at this point. The shunt inductor has a normalized susceptance of 1 mho. The curve of this inductor follows the constant $R_p$ admittance coordinate, passes through point A, and ends at $B_f = B_L - B_i = 1 - 0.4 = 0.6$ mho. (Table XXXIV summarizes this procedure. It should be remembered that, in this case, the curve crosses the pure R axis.) This point is labeled point B on the network map shown in Fig. 495. Similarly, the series $C_1$ curve is plotted along the constant $R_S$ coordinate, passes through point B, and ends at $X_f = X_c - X_i = 1.5 - 1.5 = 0$ ohms. The normalized value of input impedance is read on the Z chart as 0.5 ohm. The Q of the matching network is read from the Q curves at both points A and B. The Q is 2 at point A and 3 at point B. The higher value 3 is taken as most representative of the network Q.

**Determination of network components:** In the following examples, the graphical procedures used in the design of four different types of matching network are given. For the first type, a detailed explanation of the

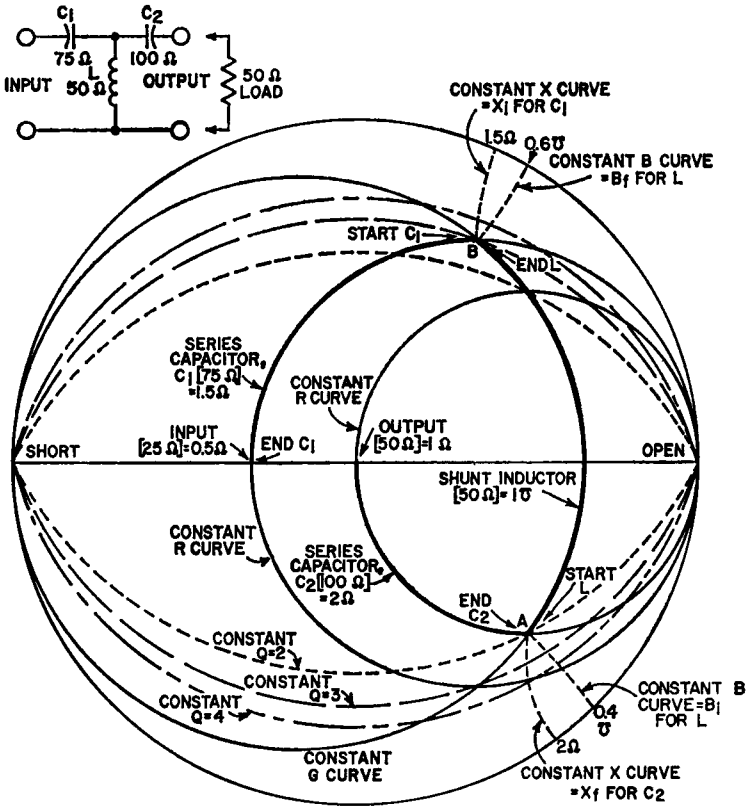



Figure 495. Typical output matching network and network map used to determine required input impedance for this network.

graphical procedure is given. For the other types, a tabular list of the steps required is considered sufficient because of the basic similarity of the graphical processes. The network maps for these examples show only the curves that are required to determine network parameters; all other curves are omitted for clarity. (In the examples, component curves are plotted as described in Table XXXIV.)

1. **Design of tapped-C network.** Fig. 496 shows the circuit configuration and the network map used to determine the component values for a tapped-C matching network that is required to transform 50 ohms to 20 ohms with a Q of 6. The procedures used to prepare the network map are as follows:

(a) The normalized input and output points (i.e., points 1 + j0Ω and point 0.4 + j0Ω are located on the impedance-chart coordinates.

(b) The Q = 6 curve is traced

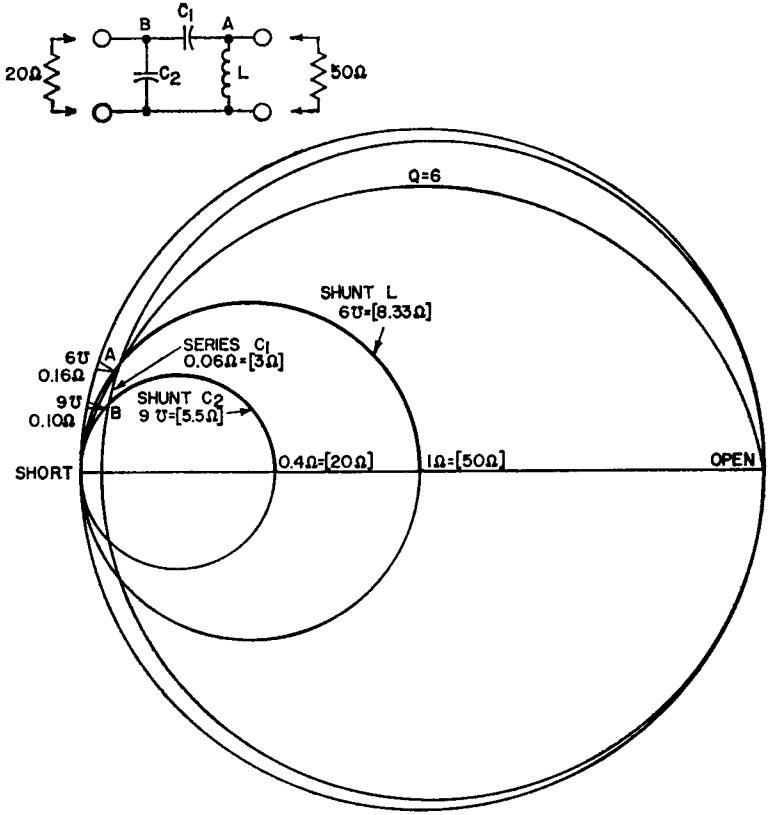



Figure 496. Circuit configuration and network map used to determine the component values for a tapped-C matching network.

from the Q chart, Fig. 492.

(c) The curve for the shunt L is traced along the constant $R_p = 1.0\upsilon$ curve (from the admittance chart) from the termination point $1 + j0\Omega$ to the intersection of this curve with the $Q = 6$ curve. This intersection is labeled A for further reference.

(d) A constant $R_S$ curve for the series $C_1$ is traced from the impedance chart. The starting point for this curve is point A.

(e) The curve for the shunt $C_2$ is then traced between the termination point $0.4 + j0\Omega$ and the intersection of this curve with that for the series $C_1$. The intersection of the $C_1$ and $C_2$ curves is labeled point B. (Although the intersection B is determined after the curve for the shunt $C_2$ is traced, this intersection is considered as the starting point for the shunt $C_2$ curve.)

(f) As a routine matter, the reactance X and the susceptance B values for the intersection points A and B are determined by means of series and parallel charts.

For point A, $X = 0.16$ ohm, and $B = 6$ mho.

For point B, $X = 0.10$ ohm, and $B = 9$ mho.

(g) As indicated in Table XXXIV, normalized reactance values for the shunt inductor L, the series capacitor $C_1$, and the shunt capacitor $C_2$ are determined by subtraction of the values at the starting point of the curves for these components from the values at the end point of these curves. The following values are obtained:

$$\Delta B_L = B_{(at\,A)} - B_{(at\,1+j0\Omega)} = 6\upsilon - 0 = 6 \text{ mhos}$$

$$\Delta X_{c1} = X_{(at\,B)} - X_{(at\,A)} = 0.10 - 0.16 = 0.06 \text{ ohm}$$

$$\Delta B_{C2} = B_{(at\,0.4+j0\Omega)} - B_{(at\,B)} = 0 - 9\upsilon = 9 \text{ mhos}$$

(h) The actual reactance values for L, C, and $C_2$ can then be determined as follows:

$$X_L = 50/\Delta B_L = 50/6 = 8.3 \text{ ohms}$$

$$X_{C_1} = 50(\Delta X_{C_1}) = 50(0.06) = 3 \text{ ohms}$$

$$X_{C_2} = 50/\Delta B_{C_2} = 50/9 = 5.5 \text{ ohms}$$

(i) The component values for the filter can then be calculated on the basis of the reactances and the operating frequency.

The detailed step-by-step procedure given above is summarized in Table XXXV. For one familiar with the basic graphical processes, this table provides sufficient information for the design of the filter network.

An intuitive analysis of the tapped-C network indicates that the shunt inductor L reduces the 50-ohm output impedance to the value represented by point A on the network map and that the shunt capacitor $C_2$ reduces the impedance of the 20-ohm input to a nearly equal value, as represented by point B. The series capacitance $C_1$ makes up the difference in the reactance of the two impedance points A and B and provides resonance. The values of both capacitors $C_1$ and $C_2$ must be changed together to maintain resonance when the input impedance is changed. The Q is deter-

Table XXXV—Procedure for Determining Component Values for Tapped-C Matching Network

| Step No. | Component Curves COMPONENT | INITIAL POINT | FINAL POINT AND HOW IT IS DETERMINED |
|---|---|---|---|
| 1 | Q = 6 | | |
| 2 | Shunt L | (1 + j0)Ω | A, determined by intersection of L and Q = 6 curves |
| 3 | Series $C_1$ | A | B, determined by intersection of $C_1$ and $C_2$ curves |
| 4 | Shunt $C_2$ | B | 0.4Ω + j0Ω = [20Ω] |

**Intersection Parameter Values from X and Y Charts**

| | INTERSECTION | REACTANCE X | SUSCEPTANCE B |
|---|---|---|---|
| 5 | A | 0.16Ω | 6υ |
| 6 | B | 0.10Ω | 9υ |

**Computing Component Values**

| | COMPONENT | INITIAL POINT | FINAL POINT | PARAMETER CHANGE (NORMALIZED VALUES) | COMPONENT VALUE |
|---|---|---|---|---|---|
| 7 | Shunt L | (1 + j0)Ω | A | ΔB = 6 − 0 = 6υ | $X_L = 50/\Delta B = 8.33\Omega$ |
| 8 | Series $C_1$ | A | B | ΔX = 0.1 − 0.16 = 0.06Ω | $X_{C1} = \Delta X(50) = 3\Omega$ |
| 9 | Shunt $C_2$ | B | (0.4 + j0)Ω | ΔB = 0 − 9 = 9υ | $X_{C2} = 50/\Delta B = 5.5\Omega$ |

mined by inductor L and the 50-ohm load impedance. At the input side, the transformation ratio is smaller, and the Q must be smaller.

2. **Design of pi network.** Fig. 497 shows the circuit configuration and the network map for a pi-type matching network required to transform 50 ohms to 20 ohms with a Q of 5. The network-mapping procedures used to determine the component values for the pi network are given in Table XXXVI.

In the pi matching network, the shunt C across the 50-ohm output reduces the output impedance to the value represented by point A on the network map. The shunt capacitor across the 20-ohm input reduces the input impedance to the nearly equal value represented by point B. The Q at the input is smaller because the change in impedance is less. The series inductor connects the input and output and cancels the reactances of the two capacitors. The impedance transformation is determined by the difference in the input and output Q.

3. **Design of lossy-L network.** Fig. 498 shows the circuit configuration and network map for a lossy-L matching network required to transform 50 ohms to 10 ohms with a Q of 5. Table XXXVII gives the graphical procedure used to determine the component values for this network.

In the lossy-L network, the series inductor increases the impedance of the 10-ohm input and determines the operating Q of the network. The series capacitor increases the impedance of the 50-ohm output, and the shunt capacitor tunes out the surplus reactance. In spite of the large impedance transformation (10 to 50 ohms), all component values have nearly equal impedances (56, 90, and 100 ohms). These relatively large values make the components quite practical, and are particularly advantageous for matching into the base of a transistor in which the impedance is only a few ohms.

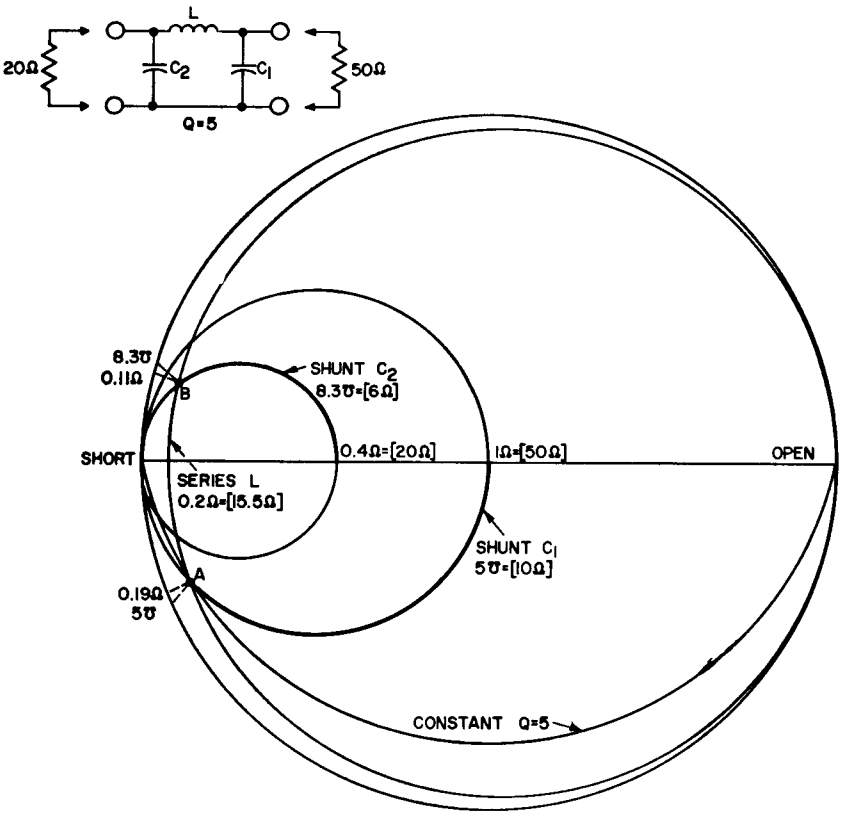



Figure 497. Circuit configuration and the network map used to determine the component values for a pi matching network.

## Table XXXVI—Procedure for Determining Component Values for Pi Matching Network

| Step No. | Component Curves<br>COMPONENT | INITIAL POINT | FINAL POINT AND HOW IT IS DETERMINED |
|---|---|---|---|
| 1 | Q = 5 | | |
| 2 | Shunt $C_1$ | (1 + j0)Ω | A, intersection of $C_1$ and Q = 5 |
| 3 | Series L | A | B, intersection of $C_1$ and $C_2$ (step 4) |
| 4 | Shunt $C_2$ | B | (0.4 + j0)Ω |
| | **Intersection Parameter Values** | | |
| | **INTERSECTION** | **SUSCEPTANCE B** | **REACTANCE X** |
| 5 | A | 5℧ | 0.19Ω |
| 6 | B | 8.3℧ | 0.11Ω |
| | **Compute Component Values** | | |

| | COMPONENT | INITIAL POINT | FINAL POINT | PARAMETER CHANGE (NORMALIZED VALUES) | COMPONENT VALUE |
|---|---|---|---|---|---|
| 7 | Shunt $C_1$ | (1 + j0)Ω | A | $\Delta B = 5 - 0 = 5℧$ | $X_{C_1} = 50/\Delta B = 10\Omega$ |
| 8 | Series L | A | B | $\Delta X = 0.11 + 0.19 = 0.3\Omega$ | $X_L = 50(\Delta X) = 15\Omega$ |
| 9 | Shunt $C_2$ | B | (0.4 + j0)Ω | $\Delta B = 0 - 8.33 = 8.33℧$ | $X_{C_2} = 50/\Delta B = 6\Omega$ |

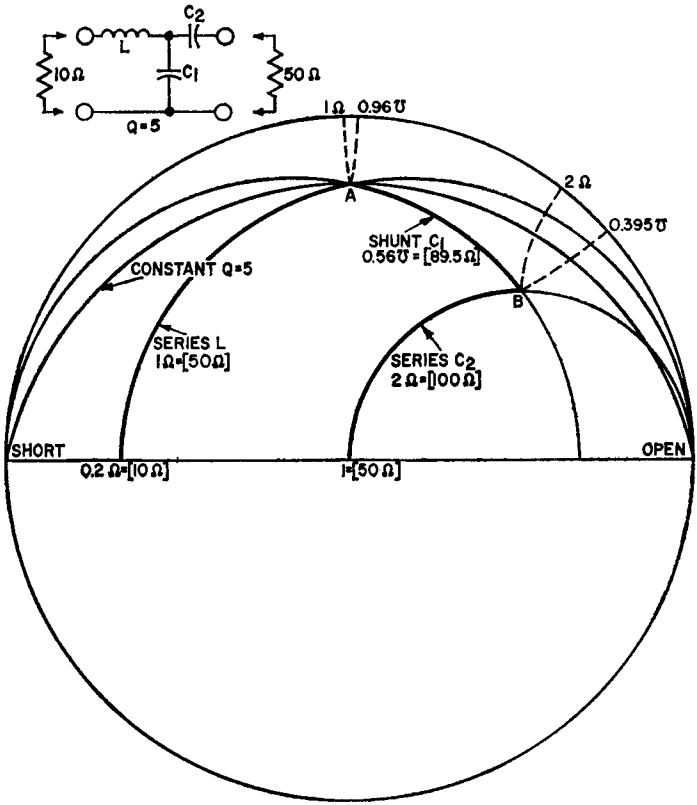



Figure 498. Circuit configuration and the network map used to determine the component values for a "lossy"-L matching network.

## Table XXXVII—Procedure for Determining Component Values for Lossy-L Matching Network

| Step No. | Component Curves COMPONENT | INITIAL POINT | FINAL POINT AND HOW IT IS DETERMINED |
|---|---|---|---|
| 1 | Q = 5 | | |
| 2 | Series L | 0.2Ω | A, intersection of L and Q = 5 curves |
| 3 | Shunt $C_1$ | A | B, intersection of $C_1$ and $C_2$ curves |
| 4 | Series $C_2$ | B | 1.0Ω |

**Determine Component Values**

| COMPONENT | INITIAL POINT | FINAL POINT | PARAMETER CHANGE (NORMALIZED VALUES) | COMPONENT VALUE |
|---|---|---|---|---|
| L | 0.2Ω | X = 1.0 | $\Delta X = 1.0 - 0 = 1.0\Omega$ | $X_L = 50\Omega$ |
| $C_1$ | A,B = 0.96 | B, B = 0.395 | $\Delta B = 0.395 - 0.96 = 0.56\upsilon$ | $X_{CL} = 89.5\Omega$ |
| $C_2$ | B,X = 2.0 | 1.0 | $\Delta X = 0 - 2.0 = 2.0\Omega$ | $X_{C2} = 100\Omega$ |

4. **Design of network containing four unspecified components.** Fig. 499 shows the circuit configuration and network map for a matching section required to transform a 50-ohm load impedance to 12.5 ohms for the collector load impedance of a transistor amplifier that has a Q of 5. The transistor collector has a parallel output capacitive reactance of 250 ohms. This network has four unspecified components ($L_1$, $L_2$, $C_1$, and $C_2$) and three required conditions. The values for only three of the components can be determined by the graphical techniques; the value of the fourth component must be arbitrarily assigned. The value of $L_1$ is normally selected so that this component is nearly resonant with $C_{in}$ at the operating frequency. In this example, however, the value selected for $L_1$ is small to demonstrate the flexibility in the choice.

The first step in the preparation of the map is to plot all known and assigned values. The three remaining components are then plotted and calculated as in examples 1, 2, and 3. The graphical procedures are outlined in Table XXXVIII.

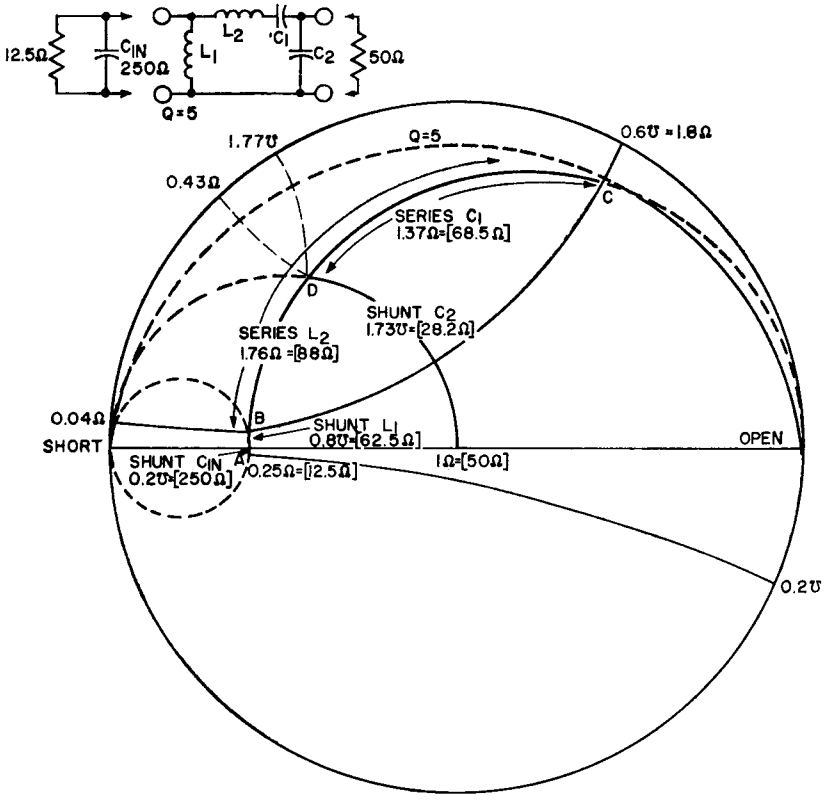



Figure 499. Circuit configuration and network map used to determine component values for matching network that includes four unspecified components.

Table XXXVIII—Procedure for Determining Component Values for Matching Network in Which Four Unspecified Components Are Used

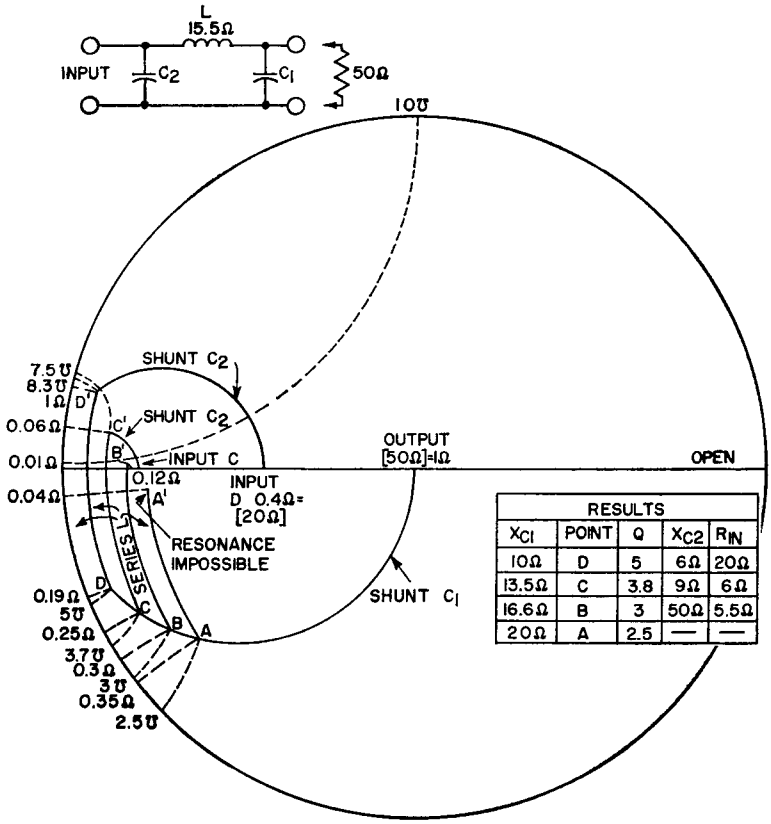
| Step No. | COMPONENT | INITIAL POINT | FINAL POINT AND HOW IT IS DETERMINED |
|---|---|---|---|
| 1 | Shunt $C_{in}$ | $(0.25 + j0)\Omega$ | A, determined by given value for $XC_{in}$ |
| 2 | Shunt $L_1$ | A | B, determined by assigned value for $X_{L1}$ |
| 3 | $Q = 5$ | — | — |
| 4 | Series $L_2$ | B | C, determined by intersection of $L_2$ and $Q = 5$ |
| 5 | Series $C_1$ | C | D, determined by intersection of $C_1$ and $C_2$ (step 6) |
| 6 | Shunt $C_2$ | D | $(1 + j0)\Omega$ |

| COMPONENT | INITIAL POINT | FINAL POINT | PARAMETER CHANGE (NORMALIZED VALUES) | REACTANCE VALUE |
|---|---|---|---|---|
| $C_{in}$ | $0.25\Omega$ | $A,B = 0.2\upsilon$ | $\Delta B = 0.20 = 0.2\upsilon$ | $X_{C1n} = 250\Omega$ |
| $L_1$ | $A,B = 0.2\upsilon$ | $B,B = 0.6\upsilon$ | $\Delta B = 0.6 + 0.2 = 0.8\upsilon$ | $X_{L1} = 62.5\Omega$ |
| $L_2$ | $B,X = 0.04\Omega$ | $C,X = 1.8\Omega$ | $\Delta X = 1.8 - 0.04 = 1.76\Omega$ | $X_{L2} = 88\Omega$ |
| $C_1$ | $C,X = 1.8\Omega$ | $D,X = 0.43\Omega$ | $\Delta X = 0.43 - 1.8 = 1.37\Omega$ | $X_{C1} = 68.5\Omega$ |
| $C_2$ | $D,B = 1.77\upsilon$ | $1.0\Omega$ | $\Delta B = 0 - 1.77 = 1.77\upsilon$ | $X_{C2} = 28.2\Omega$ |

This network provides the best separation of impedance transformation, resonance adjustment, and operating Q. The components $L_1$ and $C_{in}$ nearly resonate; however, perfect resonance is not required. The circuit tunes well even for large errors in $C_{in}$ or $L_1$. The capacitor $C_2$ reduces the 50-ohm output impedance to the series resistance required at the input. The capacitor $C_2$, therefore, is the principal loading adjustment for the amplifier. The components $L_2$ and $C_1$ form a series resonant circuit which compensates for the differences in input and output reactances. Inductor $L_1$ and the 12.5-ohm input determine the operating Q relatively independent of resonance. The Q, therefore, is rather tightly controlled. Capacitor $C_1$ compensates for the additional inductor $L_2$ needed to provide the proper Q but not needed to match the input to the output. Therefore, $C_1$ is the resonance adjustment, and $C_2$ is the loading adjustment.

**Effect of Component Changes in Graphical Design**—A particular advantage of graphical network design is that changes in component values can be easily evaluated. The pi network in example 2 (Fig. 497) was designed for an input impedance of 20 ohms, but it may be changed by means of variable components. Two components must be varied to (1) change the impedance, and (2) maintain resonance. For this pi network, $X_{C1}$ is increased in steps and L is kept constant, as shown in Fig. 500. Also shown is the $X_{C2}$ required to produce resonance, and the resulting Q and input impedance.

It should be noted that the first step increase in $X_{C1}$ (35%) changes the Q and $R_{in}$ greatly, but requires little change in $X_{C2}$. The second step increase in $X_{C1}$ (23%) changes the Q slightly, changes the input very little (8%), but requires a large change in $X_{C2}$. Any further change in $X_{C2}$ makes resonance impossible.




| RESULTS | | | | |
|---|---|---|---|---|
| $X_{C1}$ | POINT | Q | $X_{C2}$ | $R_{IN}$ |
| 10Ω | D | 5 | 6Ω | 20Ω |
| 13.5Ω | C | 3.8 | 9Ω | 6Ω |
| 16.6Ω | B | 3 | 50Ω | 5.5Ω |
| 20Ω | A | 2.5 | — | — |

Figure 500. Circuit configuration for pi matching network and network map showing tuning range for variable components used in this type of matching network.

## Transmission-Line Matching Techniques

The network-design techniques discussed in the previous sections apply largely to lumped-constant circuits operated in the vhf and uhf ranges. At uhf and higher frequencies it may be more desirable to use short sections of transmission line to provide the reactive elements needed in the previous discussions. The Smith Chart is generally used in these determinations also. There are many special-case conditions which the circuit designer can use without resorting to the general transmission-line equation or the graphical method of the Smith Chart. A few of the more useful expressions are presented in this section.

**Half-Wave Line Sections**—Sections of uniform transmission lines which are electrically an integral number of half-wavelengths ($\lambda/2$) in length are useful in transferring an impedance from one point to another, i.e., the terminating impedances on the line are equal, or $Z_S = Z_L$.

**Quarter-Wave Line Sections**—Sections of uniform transmission lines which are electrically a quarter of a wavelength ($\lambda/4$) in length have a number of interesting and useful properties. A quarter-wave line which is short-circuited at one end provides a very high impedance at the open end. This property can be used to provide high-resistance stub supports for rf structures as well as to provide rf-choke action for dc bias circuits.

The quarter-wave lines are also useful as an impedance transformer between real impedances. The characteristic impedance can be determined as follows:

$$Z_o = (R_S \times R_L)^{1/2} \quad (375)$$

where $R_S$ is the source or input impedance and $R_L$ is the load or output impedance.

If quarter-wave transformers are used to match a real impedance to an active device, as shown in Fig. 501, the reactive component of the complex impedance (the admittance) of the active device must be tuned out. For example, in the input circuit of a power-transistor amplifier circuit, the quarter-wave transformer matches the resistive component of the complex admittance of the device. An external capacitance $C_0$ or a stub provides the necessary susceptance needed to cancel the reactive component of the device. In the output portion of the circuit, a stub or a lumped element at the collector is used to bring the impedance to a real value and then to a quarter-wave line that goes to the actual load.

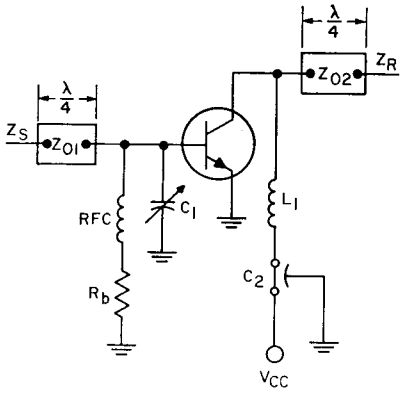


Figure 501. Quarter-wave transformers for rf power-transistor amplifiers.

Direct transformation between the transistor (complex impedance) and a given source or load (real resistance) is also possible. The characteristic impedance $Z_0$ and length $l$ of the transmission line required to provide direct transformation from a pure resistance $R_1$ to an impedance $Z_2 = R_2 + jX_2$ can be determined by use of the following equations:

$$Z_o = \sqrt{R_1 R_2} \times \sqrt{1 - \frac{X_2^2}{R_2 (R_1 - R_2)}} \quad (376)$$

$$\tan \beta l = Z_o \left( \frac{R_1 - R_2}{R_1 X_2} \right) \quad (377)$$

If the impedance $Z_2$ is a resistance (i.e., $X_2 = 0$), the expression for $Z_0$ reduces to the quarter-wave transformer equation, and $l = \lambda/4$.

**Eighth-Wave Line Sections**—Eighth-wave ($\lambda/8$) sections of uniform line have additional useful properties. If the eighth-wave section of line is terminated in a pure resistance, the input impedance will have a magnitude equal to the characteristic impedance $Z_0$ of the line. Conversely, an eighth-wave section of line

which is terminated in an impedance whose magnitude is equal to $Z_o$ must have a real input impedance. Therefore, for an eighth-wave line section, $Z_L$ is real if the following relationship is valid:

$$Z_o = |Z_s| = (R_S^2 + X_S^2)^{1/2} \quad (378)$$

Fig. 502 shows the use of eighth-wave transformers in a typical rf power-amplifier circuit.

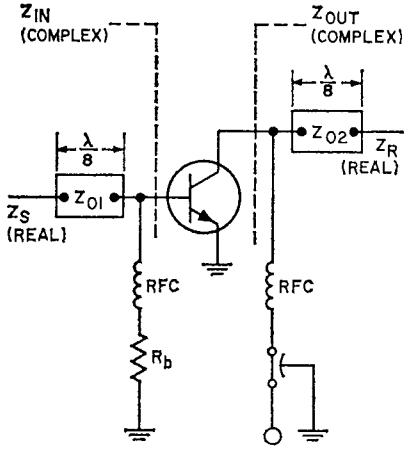



Figure 502. Eighth-wave transformers in a typical rf power-amplifier circuit.

The real impedance $Z_L$ can be determined from the Smith Chart or by use of the following equation:

$$Z_L = \frac{R}{1 - \frac{X}{|Z_S|}} \quad (379)$$

where R and X are the real and imaginary parts, respectively, of the complex impedance $Z_S$.

**Tapered Line Section**—Quarter-wave or eighth-wave line sections in which the impedance changes exponentially (or hyperbolically) have additional properties useful to the circuit designer. These line sections can be tapered directly to a desired real impedance rather than a predetermined impedance as was the case for a uniform line. In addition, because of the nature of the TEM mode of propagation in these tapered lines, substantial reductions in effective line lengths and increased transformation bandwidths are possible. The design of an amplifier circuit at a frequency of 2 GHz is described as an example. The dynamic input impedance $Z_{in}$ and collector load impedance $Z_{CL}$ of the transistor to be used are as follows:

$$Z_{in} = 7.5 + j\,8.0 \text{ ohms} \quad (380)$$

$$Z_{CL} = 6.5 + j\,33 \text{ ohms} \quad (381)$$

The amplifier uses an air-line input circuit and an air-line output circuit similar to that shown in Fig. 503.

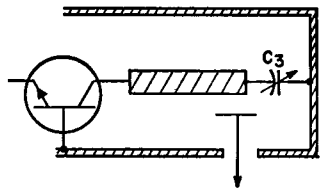



Figure 503. Capacitive-probe-coupled output cavity.

For the design of the output circuit, the optimum characteristic impedance $Z_o$ of the output line is calculated from Eq. (378) to be 31 ohms. The collector load impedance $Z_{CL}$ is normalized as follows:

$$Z_{CL}' = Z_{CL}/Z_o = 0.18 + j0.915 \quad (382)$$

Point $Z_{CL}'$ is then located on the Smith Chart shown in Fig. 504 and rotated about the constant-VSWR circle toward the load. The

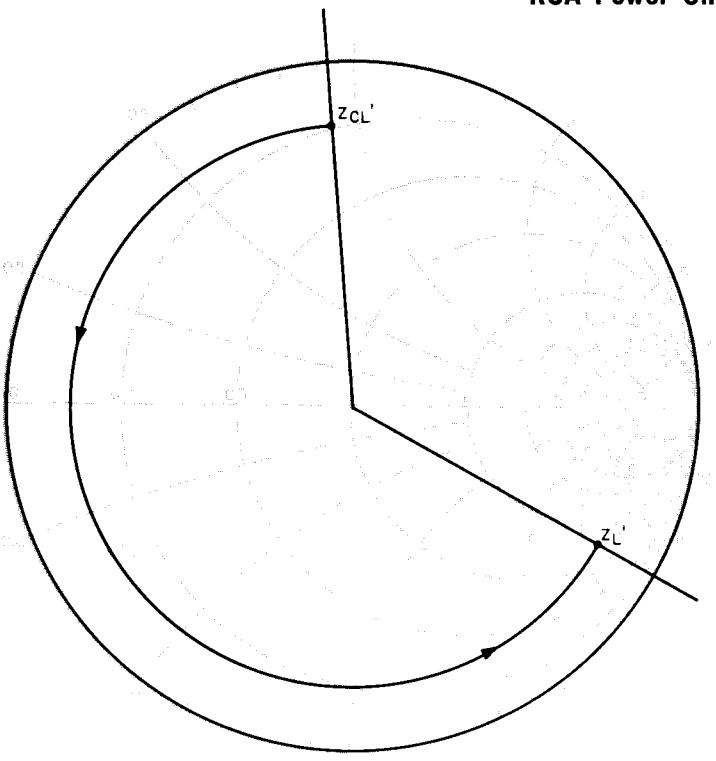



Figure 504. Smith-chart admittance plot for design of output transmission-line matching section.

intersection of the VSWR circle and the 1.39 constant-resistance circle is denoted as point $Z_L'$ (the load resistance is assumed to be 50 ohms and the normalized load resistance is, therefore, 1.39 ohms). At point $Z_L'$, the normalized impedance is given by

$$Z_L' = 1.39 - j\,3.3 \qquad (383)$$

The load impedance $Z_L$ is then equal to

$$Z_L = Z_o Z_L' = 36(1.39 - j\,3.3)$$
$$= 50 - j\,119 \text{ ohms} \qquad (384)$$

The line length required to transform the transistor collector load impedance from 0.5 ohm to a load impedance of 50 ohms is determined from Fig. 504 to be equal to 0.33λ (where λ is the wavelength in air). At 2 GHz, λ is equal to 5.9 inches, and the length of output line is calculated to be 1.95 inches. A capacitive reactance component with a value equal to 119 ohms is needed to complete the output circuit, as shown in Fig. 505(a).

For the design of the input circuit, a characteristic impedance $Z_o$ of 11 ohms is calculated from Eq. (378). The input impedance $Z_{in}$ is normalized as follows:

$$Z_{in}' = (Z_{in}/Z_o) = 0.68 + 0.725 \qquad (385)$$

Point $Z_{in}'$ is then located on the Smith Chart shown in Fig. 506

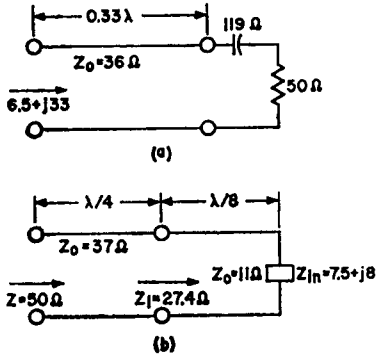



Figure 505. Input and output transmission lines with transmission-line matching sections added to provide required impedance transformation: (a) output line; (b) input line using series matching section; (c) input line using series matching section foreshortened by reactive elements.

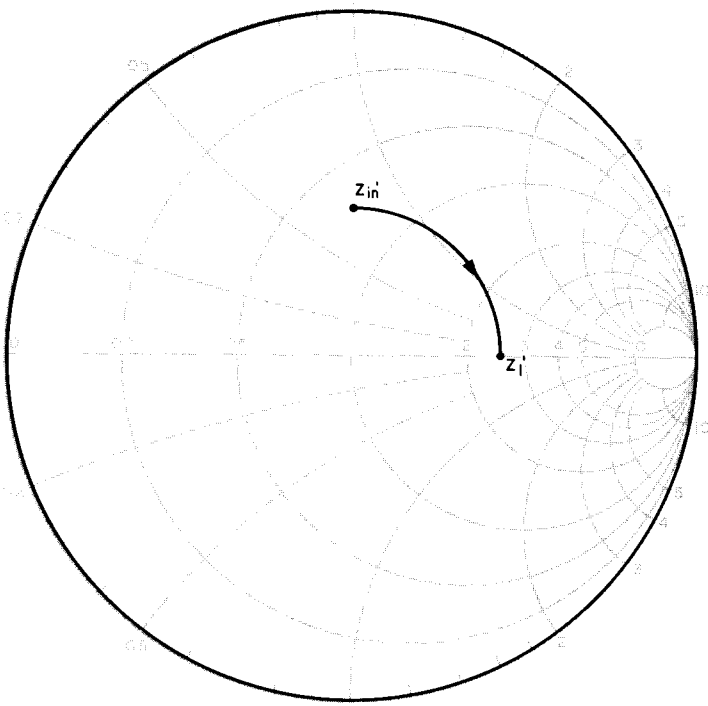



Figure 506. Smith-chart admittance plot used for design of input transmission-line matching sections.

and rotated about the constant-VSWR circle toward the generator to locate the intersection between the VSWR circle and the 4.55-ohm constant-resistance circle. (The driving-source impedance is assumed to be 50 ohms and the normalized source impedance is, therefore, 4.55 ohms.) However, Fig. 506 shows that such an intersection is not possible and that a more sophisticated input circuit is needed. One possible circuit employs another section of line. For minimum VSWR in the added section of line, the line length for the 11-ohm line must be $\lambda/8$ or 0.75 inch, as discussed previously. This point, denoted as $Z_1'$ in Fig. 506, is equal to 2.5 ohms; the impedance $Z_1$ is then 27.4 ohms. The characteristic impedance of the added line section required to transform a resistance of 27.4 ohms to a 50-ohm source is calculated from Eq. (375) to be 37 ohms. Such an input circuit is shown in Fig. 505(b). Another possible input circuit uses added reactive elements, as shown in Fig. 505(c), to foreshorten the additional line section.

The design of microstripline circuits is the same as that described for air-line circuits, except that the wavelength of the line must be modified by a factor of $1/\sqrt{\epsilon}$, where $\epsilon$ is the dielectric constant of the insulator material of the stripline.

## SINGLE SIDEBAND SYSTEMS

Single-sideband communication systems have many advantages over AM and FM systems. In areas where reliability of transmission as well as power conservation are of prime concern, SSB transmitters are usually employed. The main advantages of SSB operation include reduced power consumption for effective transmission, reduced channel width to permit more transmitters to be operated within a given frequency range, and improved signal-to-noise ratio.

In a conventional 100-per-cent modulated AM transmitter, two-thirds of the total power delivered by the power amplifier is at the carrier frequency, and contributes nothing to the transmission of intelligence. The remaining third of the total radiated power is distributed equally between the two sidebands. Because both sidebands are identical in intelligence content, the transmission of one sideband would be sufficient. In AM, therefore, only one-sixth of the total rf power is fully utilized. In an SSB system, no power is transmitted in the suppressed sideband, and power in the carrier is greatly reduced or eliminated; as a result, the dc power requirement is substantially reduced. In other words, for the same dc input power, the peak useful output power of an SSB transmitter, in which the carrier is completely suppressed, is theoretically six times that of a conventional AM transmitter.

Another advantage of SSB transmission is that elimination of one sideband reduces the channel width required for transmission to one-half that required for AM transmission. Theoretically, therefore, two SSB transmitters can be operated within a frequency spectrum that is normally required for one AM transmitter.

In a single-sideband system, the signal-to-noise power ratio is eight times as great as that of a fully modulated double-sideband system for the same peak power.

## Analysis of SSB Signal

A single-sideband signal is usually generated at low level and then amplified through a chain of linear amplifiers to the desired power. The two most commonly used methods of generating sideband signals are with the filter-type generator and the phasing-type generator.

It can be shown mathematically that a single-sideband signal is derived from an amplitude-modulated wave. If an rf carrier frequency is modulated by an audio frequency, the resulting AM wave can be expressed by the following equation:

$$e = E_o (1+m \cos 2\pi f_m t) \sin 2\pi f_c t$$

$$e = E_o \sin 2\pi f_c t + m E_o \cos 2\pi f_m t \sin 2\pi f_c t \quad (386)$$

in which $f_m$ is the audio modulating frequency, $f_c$ is the carrier frequency, and m is the per-cent-modulation factor. Expansion of the last term of this equation into functions of sum and difference angles by the usual trigonometric formulas results in the following expression:

$$e = E_o \sin 2\pi f_c t + (m E_o/2) \sin 2\pi (f_c+f_m) t + (m E_o/2) \sin 2\pi (f_c-f_m) t \quad (387)$$

This equation contains three components, each of which represents a wave. The first wave, represented by the term $E_o \sin 2\pi f_c t$, is called the carrier. It is present with or without modulation and maintains a constant average amplitude at a frequency $f_c$. The other two components of the equation represent waves that have equal amplitude, but frequencies above and below the carrier frequency by the amount of the modulating frequency. These components contain identical intelligence and are called sideband frequencies. The amplitude of the sideband frequencies depends on the degree of modulation (m). The higher the m factor, the greater the "talk power." Because only the sidebands transmit intelligence and because each sideband is a mirror image of the other, it is reasonable to assume that if the carrier and one sideband are eliminated, the remaining sideband is adequate for transmission of intelligence. This technique is applied in single-sideband transmission.

As mentioned previously, the elimination of one sideband reduces the bandwidth required by one half. This advantage is not fully realized unless the transmitter has the capability to amplify a signal linearly without introducing distortion products. Excessive distortion nullifies the advantage of reduced bandwidth in SSB transmission by generating unwanted frequencies which occupy segments of the spectrum that are allocated for other transmitters. The main objection to this distortion is not that it seriously affects intelligibility of the signal in the passband, but that it radiates rf energy on both sides of the passband and interferes with adjacent channels.

## Linearity Test

For an amplifier to be linear, a relationship must exist such that the output voltage is directly proportional to the input voltage for

all signal amplitudes. Because a single-frequency signal in a perfectly linear single-sideband system remains unchanged at all points in the signal path, the signal cannot be distinguished from a cw signal or from an unmodulated carrier of an AM transmitter. To measure the linearity of an amplifier, it is necessary to use a signal that varies in amplitude. In the method commonly used to measure nonlinear distortion, two sine-wave voltages of different frequencies are applied to the amplifier input simultaneously, and the sum, difference, and various combination frequencies that are produced by nonlinearities of the amplifier are observed. A frequency difference of 1 to 2 kHz is used widely for this purpose. A typical two-tone signal without distortion, as displayed on a spectrum analyzer, is shown in Fig. 507. The resultant signal envelope

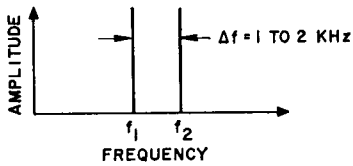



Figure 507. Frequency spectrum for a typical two-tone signal without distortion.

varies continuously between zero and maximum at an audio-frequency rate. When the signals are in phase, the peak of the two-frequency envelope is limited by the voltage and current ratings of the transistor to the same power rating as that for the single-frequency case. Because the amplitude of each two-tone frequency is equal to one-half the cw amplitude under peak power condition, the average power of one tone of a two-tone signal is one-fourth the single-frequency power. For two tones, conversely, the peak envelop power (PEP) rating of a single-sideband system is two times the average power rating.

## Intermodulation Distortion

Nonlinearities in an amplifier generate intermodulation distortion IMD. The important IMD products are those close to the desired output frequency, which occur within the pass band and cannot be filtered out by normal tuned circuits. If $f_1$ and $f_2$ are the two desired output signals, third-order IM products take the form $2f_1 - f_2$ and $2f_2 - f_1$. The matching third-order terms are $2f_1 + f_2$ and $2f_2 + f_1$, but these matching terms correspond to frequencies near the third harmonic output of the amplifier and are greatly attenuated by tuned circuits. It is important to note that only odd-order distortion products appear near the fundamental frequency. The frequency spectrum shown in Fig. 508 illustrates the frequency relationship of some distortion products to the test signals $f_1$ and $f_2$. All such products are either in the difference-frequency region or in the harmonic regions of the original frequencies. Tuned circuits or filters following the nonlinear elements can effectively remove all products generated by the

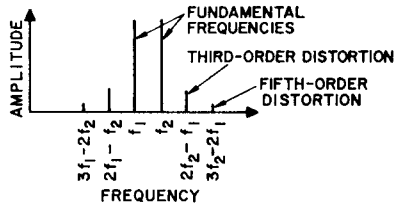



Figure 508. Frequency spectrum showing the frequency relationship of some distortion products to two test signals $f_1$ and $f_2$.

even-order components of curvature. Therefore, the second-order component that produces the second harmonic does not produce any distortion in a narrow-band SSB linear amplifier. This factor explains why class AB and class B rf amplifiers can be used as linear amplifiers in SSB equipment even through the collector-current pulses contain large amounts of second-harmonic current. In a wideband linear application, however, it is possible for harmonics of the operating frequency to occur within the pass band of the output circuit. Biasing the output transistor further into class AB can greatly reduce the undesired harmonics. Operation of two transistors in the push-pull configuration can also result in cancellation of even harmonics in the output.

The IMD ratio (in dB) is the ratio of the amplitude of one test frequency to the amplitude of the strongest distortion product. A signal-to-distortion specification of −30 dB means that no distortion product will exceed this value for a two-tone signal level up to the PEP rating of the amplifier. A typical presentation of IMD for an RCA-2N6093 transistor at various output-power levels is shown in Fig. 509.

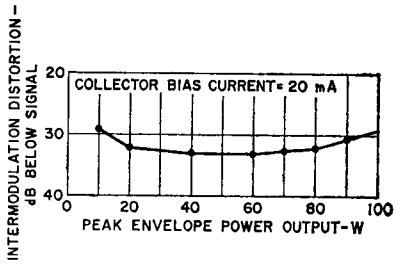



Figure 509. Typical intermodulation distortion in an RCA-2N6093 transistor at various output power levels.

## Transistor Requirements

Most high-frequency power transistors are designed for class C operation. Forward biasing of such devices for class AB operation places them in a region where second breakdown may occur. The susceptibility of a transistor to second breakdown is frequency-dependent. Experimental results indicate that the higher the frequency response of a transistor, the more severe the second-breakdown limitation becomes. For an rf power transistor, the second-breakdown energy level at high voltage (greater than 20 volts) becomes a small fraction of its rated maximum power dissipation. This behavior is one of the reasons that vacuum tubes have traditionally been used in single-sideband applications.

A power transistor designed especially for use as a linear amplifier is required to perform satisfactorily when forward-biased for class AB operation, as well as to exhibit the desired high-frequency response. The ability of the transistor to withstand second breakdown is improved by subdividing the emitter into many small sites and resistively ballasting the individual sites. An RCA-2N6093 transistor designed specifically for linear-amplifier service in SSB applications has an overlay structure with 540 parallel emitter sites, interconnected with metal fingers. Current-limiting resistors are placed in series with each emitter site between the metalizing and the emitter-to-base junction. The SSB RCA-2N6093 transistor has a high emitter-periphery-to-collector-area and a high emitter-periphery-to-emitter-area

ratio, and thereby combines good high-current performance with low capacitance.

Physically, second breakdown is a local thermal-runaway effect induced by severe current concentrations. The evidence of the random distribution of hot spots over the surface of the unit indicates that second breakdown may occur anywhere in the transistor. When a ballast resistor is used in each emitter site, current concentration is minimized. Fig. 510 is a schematic representation of the transistor showing the separate emitters with resistors in series with

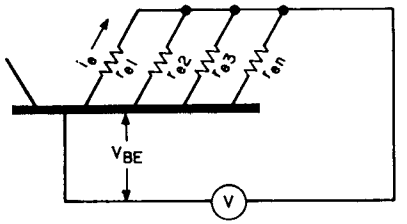



Figure 510. Schematic representation of the separate emitter sites in an "overlay" transistor with a resistor connected in series with each site.

each site. The voltage drop across each site is expressed by the following equation:

$$V = V_{BE} + i_e r_e \qquad (388)$$

Changes in $V_{BE}$ have an exponential effect on the emitter current $i_e$, as follows:

$$i_e = i_s [e^{(q/kT) V_{BE}} - 1]$$
$$= i_s [e^{(q/kT) (V - i_e r_e)} - 1] \qquad (389)$$

This equation indicates that when a constant voltage V is applied across the emitter-to-base junction and resistor network, an increase in $i_e$ at any one site causes a rise in the $i_e r_e$ voltage drop which, in turn, results in a decrease in the current to that site, i.e., the exponential term of the equation diminishes as the quantity $(V - i_e r_e)$ decreases. This condition effectively stabilizes that region. The addition of resistance to the emitters of the transistor has a degenerative effect on the device performance. However, if a large number of sites are connected in parallel, high-value individual resistors ($r_e$) can be sustained while a small total resistance ($R_t$) is still maintained at the input of the transistor, as indicated by the following relationship:

$$(1/R_t) = (1/r_{e1}) + (1/r_{e2}) + (1/r_{e3}) + \ldots + (1/r_{en}) \qquad (390)$$

A relatively large value of ballast resistance is desirable for prevention of second breakdown and for improvement of thermal stability and linearity of transfer characteristics. However, because ballast resistors are in series with the load, excessive ballasting can seriously degrade the rf performance of the transistor. Therefore, in a high-frequency power amplifier with low supply voltage, the impedance of the emitter resistance can become an appreciable portion of the reflected load presented to the collector and, as a result, can limit the power output. In determining the proper emitter-resistance value, a compromise must be made empirically so that sufficient second-breakdown protection is provided without serious effects on rf performance.

The adverse effect of high ballast resistance, besides reduced rf output power, is the increase in saturation voltage. Viewed externally, the total saturation voltage

also includes the voltage drop across the ballast resistance. This additional voltage makes the "soft" output characteristics of a transistor at high-current even softer. As a result, the available linear region through which the signal can swing is limited.

Examination of the relationship of intermodulation distortion to power output reveals that third-order distortion increases at both high and low output levels, as shown in Fig. 509. The inherent decrease in beta at high current, which causes variation in gain over a large portion of the collector dynamic characteristics, introduces additional distortion. The additional distortion is indicated by flattening of the peak of the sinusoidal swing.

The operation of a transistor near the saturation region has a pronounced effect on third-order distortion. All higher odd-order distortion products do not seem to be affected greatly by transistor operating conditions. The increase in distortion below 20 watts PEP can be attributed to lack of sufficient collector quiescent current. Nonlinearity caused by the voltage-current characteristic of the base-to-emitter junction affects distortion at low power levels. Third-order distortion is improved by use of a higher bias current, as shown in Fig. 511.

If collector bias current is set too high initially in an attempt to improve linearity at low power-output levels, the linear region of the collector characteristic is reduced. As a result, distortion because of saturation occurs much sooner. The controlling factor in determining the proper bias-current level is usually the maximum distortion that can be tolerated at a given power output. For a given transistor type, the bias point that yields the best compromise between linear performance and good collector efficiency must be determined experimentally. A

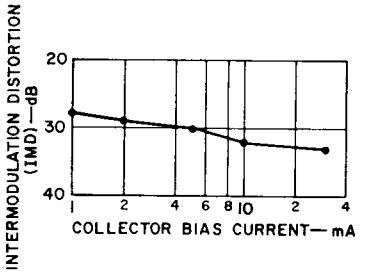



Figure 511. Intermodulation distortion as a function of collector bias current for the RCA-2N6093 SSB transistor.

collector bias current of from 2 to 20 milliamperes for the RCA-2N6093 SSB transistor is adequate to deliver 85 watts PEP. Fig. 512 shows a curve of power output as a function of supply voltage with distortion maintained at −30 dB.

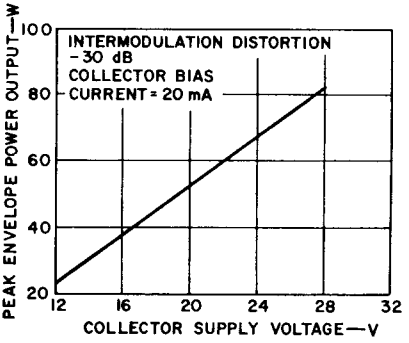



Figure 512. Peak envelope power as a function of collector supply voltage for the RCA-2N6093 SSB transistor.

## Bias Control

Operation of the transistor in a class AB amplifier to improve linearity requires the use of a

positive base voltage for an n-p-n silicon transistor. The magnitude of the positive voltage must be large enough to bias the transistor to a point slightly beyond the threshold of collector-current conduction. The class AB bias condition must be maintained over a wide temperature range to prevent an increase in idling current to the level at which the transistor can be destroyed as a result of thermal runaway and to minimize distortion that results from a shift in the quiescent point. Investigations of transistors that fail reveal that these devices exhibit a maximum $V_{BE}$ and then go into a negative-resistance region as shown in Fig. 513. The onset of

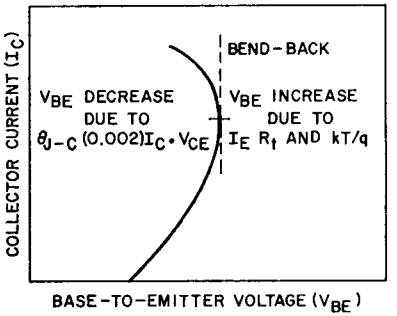



Figure 513. The bend-back phenomenon.

negative resistance, called bend-back, results in a runaway condition that ultimately destroys the transistor.

In most linear applications where the operating point of the device is biased with a voltage source, this $I_C$-$V_{BE}$ curve becomes an accurate means of predicting device stability. It is difficult to maintain a stable quiescent point of a transistor with low bend-back. Laboratory results indicate that a minimum bend-back current of 1 ampere at 22 volts is needed for a transistor to operate safely at 40-per-cent efficiency with approximately 50 watts of dissipation.

Bend-back occurs when the increase of $V_{BE}$ with collector current is just balanced by the decrease in $V_{BE}$ caused by junction-temperature rise. Therefore, at bend back,

$$kT/q + I_E R_t = \theta_{J-C}(0.002V/°C\ I_C\ V_{CE} \quad (391)$$

where

$kT/q = 0.032$ volt at 100°C
$R_t$ = total ballast resistance
$\theta_{jc}$ = junction-to-case thermal resistance
0.002V/°C = base-to-emitter junction temperature coefficient
$I_E$ = emitter current
$I_C$ = collector current
$V_{CE}$ = collector-to-emitter voltage

If $I_C = I_E$, Eq. (391) can be solved to find $I_E$ at bend-back:

$$I_E = \frac{-kT/q}{R_t - \theta_{J-C}(0.002V/°C)V_{CE}} \quad (392)$$

Thermal runaway can be attributed to the fact that the base-to-emitter junction of a transistor has a negative temperature coefficient. For example, the RCA-2N6093 transistor is forward-biased by 0.65 volt to produce a quiescent collector current of about 20 milliamperes at $V_{CC}$ = 28 volts. This operating point is shown as point A in Fig. 514. When rf drive is applied, the collector current increases to 3 amperes. If the efficiency is 40 per cent, the power dissipated in the transistor is given by

$$P_{diss.} = 28 \times 3\ (1 - 0.40)$$
$$= 50 \text{ watts}$$

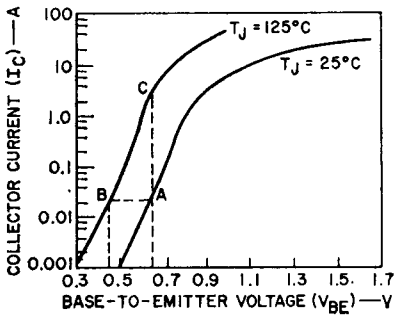



Figure 514. Collector current as a function of base-to-emitter voltage in the RCA-2N6093 for two values of junction temperature.

If the ambient temperature is 25°C, the case temperature is 50°C, and the thermal resistance is 1.5°C per watt, the junction temperature is given by

$$T_J = T_C + P_{diss.}\,\theta_{J-C}$$
$$= 50 + 50 \times 1.5 = 125°C. \quad (393)$$

The junction temperature is thus 100°C above ambient temperature. At this junction temperature, the $V_{BE}$ required to maintain a collector current of 20 milliamperes is only $0.65 - 100 \times 0.002 = 0.45$ volt, as shown at point B. If the bias voltage is fixed at 0.65 volt, however, and the drive is removed instantaneously, the quiescent current will no longer be

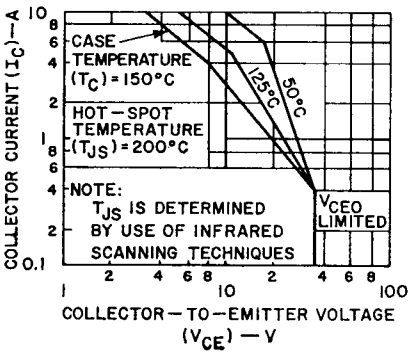



Figure 515. Safe area for dc operation of the RCA-2N6093.

20 milliamperes. Instead, the collector current will move to point C, where the operating point falls outside of the safe area of Fig. 515. Therefore, catastrophic failure will occur as a result of thermal runaway.

## Compensating Diode

To provide a bias voltage that varies with temperature in the same manner as $V_{BE}$ of the transistor, the 2N6093 incorporates a compensating diode as shown in Fig. 516(a). To insure fast thermal response time, this diode is mounted on the same beryllia disc as the transistor chip. The diode, forward-biased through $R_{Bias}$, serves as a temperature-sensing element. The voltage developed across the diode is amplified to provide a "stiff" bias-voltage source. Fig. 516(b) shows the block diagram of a temperature-compensated 30-MHz linear power amplifier that uses this transistor.

A bias-compensation circuit is included in the 30-MHz, 75-watt (PEP) amplifier shown in Fig. 517. The current amplifier uses Q1 and Q2 in a differential-amplifier arrangement so that the output voltage is independent of ambient-temperature variations. Q3 and Q4 provide the necessary current amplification. The bias current in rf transistor Q5 can be adjusted by varying $R_1$.

As shown in Fig. 518, with no rf signal the forward-biased transistor is statically stable up to a case temperature of 160°C. The dashed line in Fig. 518 shows that without temperature compensation the transistor tends to thermal runaway around 80°C. To further show the effectiveness of compensation, the third-order

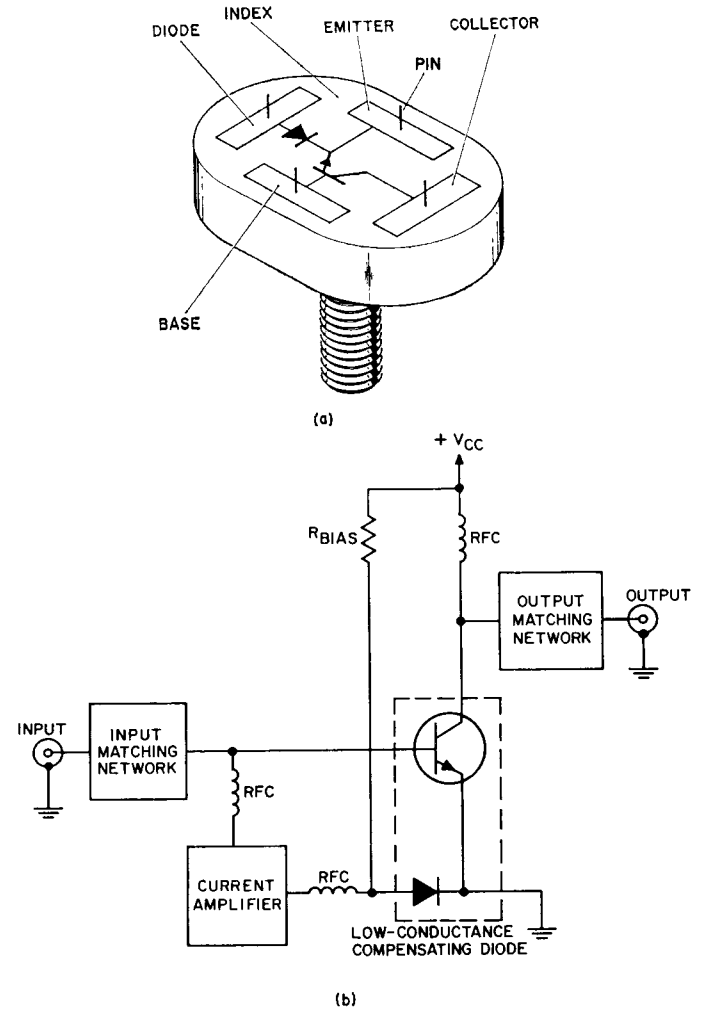



Figure 516. (a) Package outline of RCA 2N6093 showing internal compensating diode and (b) block diagram of temperature-compensated 30-MHz amplifier that uses this transistor.

distortion and output power are plotted as a function of case temperture in Fig. 519. The decrease in output power at high temperatures is caused by a drop in high-frequency gain and an increase in rf saturation voltage. The decrease in $h_{fe}$ produces a soft saturation knee that increases distortion.

## Typical Linear Amplifiers

The common-emitter configuration should be used for a linear power amplifier because of its stability and high power gain. Tuning is less critical, and the amplifier is less sensitive to variations in parameters among transistors. The class AB mode is used

to obtain low intermodulation distortion. Neither resistive loading nor neutralization is used to improve linearity because of the resulting drastic reduction in power gain; furthermore, neutralization is difficult for large signals because parameters such as output capacitance and output and input impedances vary nonlinearly over the limits of signal swing.

Fig. 517 shows a schematic diagram of a narrow-band, high-power, 30-MHz amplifier. The amplifier provides an output power of 75 watts PEP from a 28-volt power supply. The impedance of the base-to-emitter junction of the RCA-2N6093 SSB transistor in this circuit is transformed to 50 ohms to match the impedance of the drive source. The input circuit to the transistor can be represented as a resistance ($r_{bb'}$) in series with a capacitance $C_i$. The input network must tune out the capacitance $C_i$ and must present a pure resistive load to the driver. The input network is formed by the T-network consisting of capacitors $C_1$ and $C_2$ and inductor $L_1$. The value of $L_1$ is chosen so that the inductive reactance is much greater than the reactance of $C_i$. Series tuning of the base-to-emitter circuit is obtained by $L_1$ and the parallel combination of $C_2$ and $C_1$, together with the capacitance of the driver stage.

Inductor $L_2$ in the output circuit is selected to resonate with the transistor output capacitance. Capacitors $C_3$ and $C_4$ and inductor $L_3$ provide the proper impedance transformation from 50 ohms to 3.13 ohms at the resonant frequency. Base-bias voltage is obtained from the output of the compensating circuit. If the bias voltage is not temperature-compensated, both linearity and collecter efficiency can be affected.

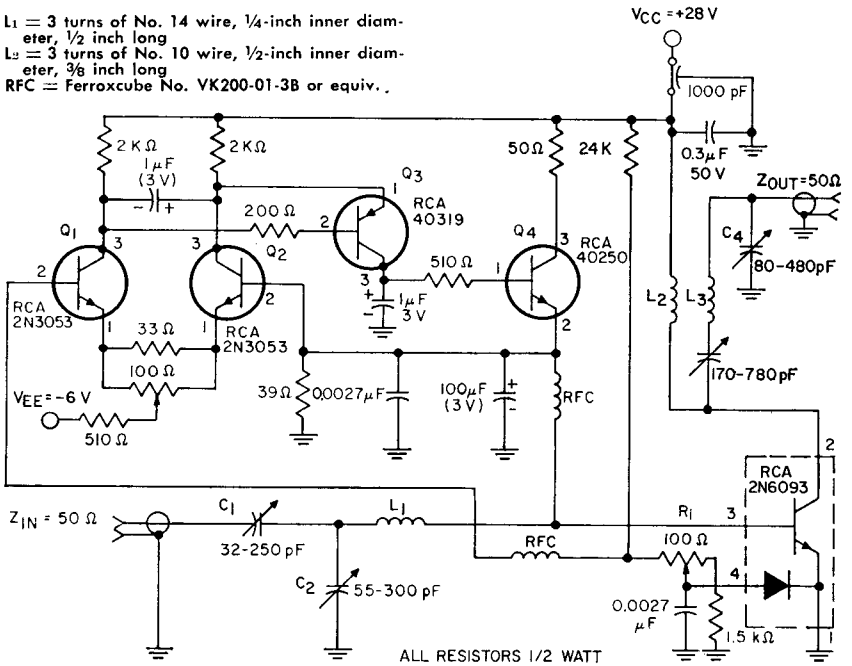



Figure 517. Use of the RCA-2N6093 in a 30-MHz, 75-watt (PEP) amplifier with temperature compensation.

When an rf signal is applied to the amplifier under high-power conditions, the rectifying property of the base-to-emitter junction charges any capacitance present in the base circuit of the transistor. This charge can alter the bias point and reduce the angle of conduction; the amplifier then operates more toward class C, and distortion and efficiency are both increased. Fig. 518 shows the effect of temperature compensation on the collector current of the 2N6093, and Fig. 519 shows the output and intermodulation-distortion characteristics of the 30-MHz amplifier as a function of temperature.

In low-power linear amplifiers, the use of temperature-compensating circuits is sometimes not necessary if the transistor output power is less than 50 per cent of its maximum cw power rating. The RCA-2N5070 transistor is useful in such applications. This transistor is specified for SSB applications without temperature compensation as follows:

Frequency = 30 MHz
$P_o$ (PEP) at 28 V = 25 W
Power Gain = 13 dB (min.)
Collector Efficiency = 40 % (min.)

Fig. 520 shows a 2-to-30-MHz wideband linear amplifier that uses other types of RCA rf power transistors. At 5 watts (PEP) output, IMD products are more than 40 dB below one tone of a two-tone signal. Power gain is greater than 40 dB.

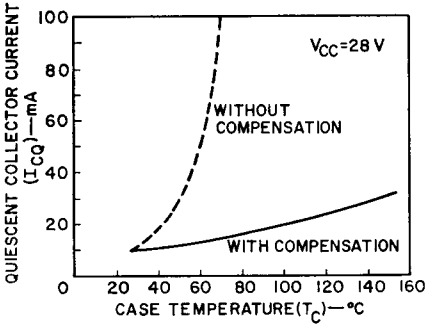



Figure 518. Quiescent collector current in the RCA-2N6093 as a function of case temperature with and without temperature compensation.

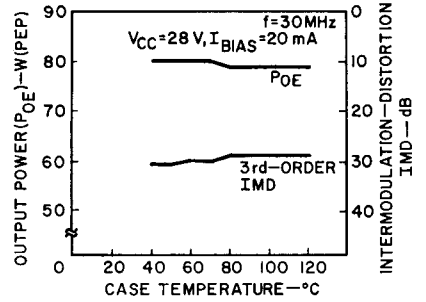



Figure 519. Output power and intermodulation distortion as a function of case temperature for the RCA-2N6093 amplifier shown in Fig. 517.

## Broadband Circuit Design

Before any circuit can be designed, the transistor input impedance and the collector load impedance over the required frequency band and at the desired levels of output power, IMD, case temperature, and collector supply voltage must be known or measured. The circuit designer must also know the transistor power gain over the same band. Curves of these characteristics for the RCA-2N6093 are shown in Figs. 521, 522, and 523. A broadband transistor should be selected for minimal impedance variation and low input Q across the frequency band. A transistor that has an $f_T$ well above the highest operating frequency, if available, can provide constant gain under broadband operation; such a transistor eliminates the need for additional gain-leveling circuitry. Because

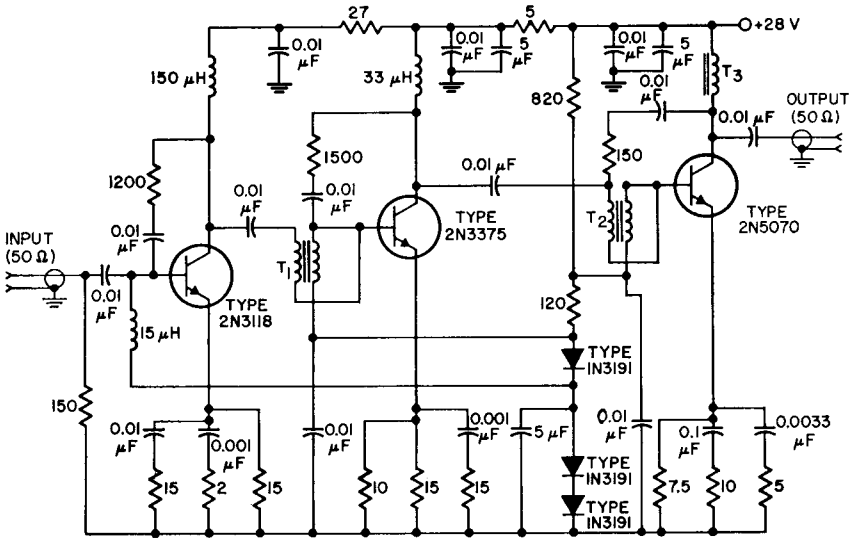



Figure 520. 2-to-30-MHz linear power amplifier.

circuit optimization becomes more difficult with high-power broadband operation, the need for thermal stability becomes more acute, and the necessity of diode compensation at high output powers becomes greater. To provide this stability, the transistor should have an internally mounted compensating diode.

The advantages which especially suit the 2N6093 for broadbanding are its low input Q and its internally mounted compensating diode. Its main disadvantage is a 15-dB gain decrease from 2 to 30 MHz that results from operation on a power-gain slope of 6 dB per octave.

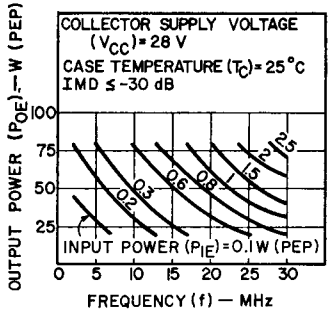



Figure 521. Typical output power as a function of frequency for the RCA-2N6093.

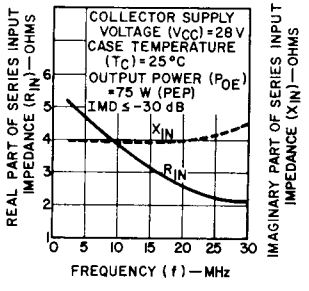



Figure 522. Typical large-signal series input impedance ($R_{in} + jX_{in}$) as a function of frequency for the RCA-2N6093.

After selection of the transistor and measurement of its broad-

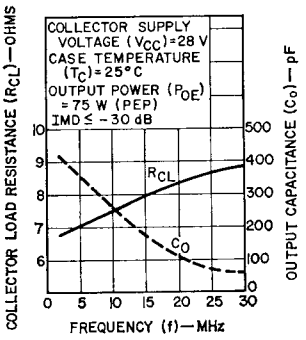



Figure 523. Typical large-signal parallel collector load resistance and parallel output capacitance as a function of frequency for the RCA-2N6093.

band parameters, the next step is to select the circuit approach. The most practical broadbanding method to provide an effective impedance transformation over four octaves (2 to 30 MHz) is a transmission-line-transformer/ferrite-core combination. The major disadvantage of a transmission line transformer is the limited number of impedance transformations available: 1:1, 4:1, 9:1, etc. The two fundamental configurations are the 1:1 reversing transformer and the 4:1 impedance transformer shown in Fig. 524.

**Ferrite Cores**—At low frequencies, a high primary reactance can be obtained with a few turns of transmission line on a high-permeability ferrite core. At high frequencies where length becomes critical, the permeability of the core decreases, thereby maintaining approximately the same levels of reactance with a short length of transmission line. Ferramic-Q core material is available in three high-frequency grades; a tabulation of their useful properties is given in Table XXXIX. Because the transformer performance is less dependent on core material at the higher-frequency end of its useful range, the poor intrinsic Q of Q-1 material above 20 MHz does not degrade the transformer operation at 30 MHz. Q-2 material, having lower permeability, requires more turns for operation at the lower frequencies.

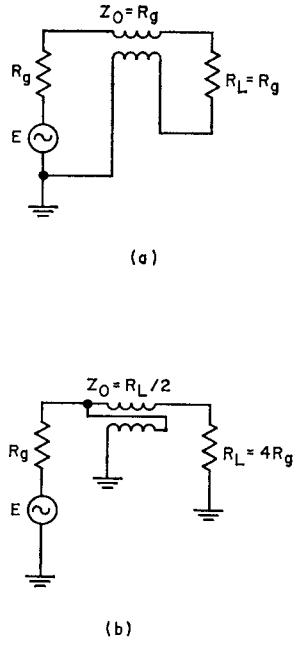



Figure 524. Transmission-line transformers: (a) 1:1 reversing/isolating transformer; (b) 4:1 impedance transformer.

**Hybrid Combiner/Dividers**—Hybrid combiner/dividers can be made by use of combinations of the 1:1 and 4:1 transformers on ferrite cores to provide high impedance-transformation ratios. As an example, Fig. 525 shows a

Table XXXIX—Permeability and Frequency Dependence of Ferramic-Q Materials

| Material | Permeability | Approximate Frequency at which core losses increase by a factor of 10 (MHz) |
|---|---|---|
| Q-1 | 125 | 10 |
| Q-2 | 40 | 90 |
| Q-3 | 16 | 225 |

180-degree-phase hybrid divider that matches a 50-ohm source to a 3.12-ohm push-pull configuration. Two 1:1 transformers are used to make the 4:1 transformation,

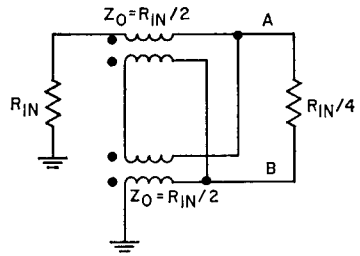



Figure 525. A 4:1 broadband transformation network that uses two 1:1 transformers to provide a balanced output.

rather than one 4:1 transformer, to provide the balanced output needed for a push-pull configuration. An equivalent transformation also can be made with one 1:1 transformer and one 4:1 transformer, as shown in Fig. 526.

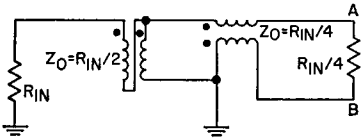



Figure 526. A 4:1 broadband transformation network that uses a 1:1 transformer and a 4:1 transformer to provide a balanced output.

Fig. 527 shows a 16:1 broadband transformation network for a push-pull configuration. The circuitry to the left of $V_2$ is the same as in Fig. 525; to the right of $V_2$ an extra transformer and dissipating resistor have been added. Points A and B are transistor base inputs, $R_2$ represents the resistive input to a conducting transistor, and $R_3$ is a resistor much larger than $R_2$ that is connected in shunt with each base-to-emitter junction. (Thus A-to-ground represents a conducting transistor, while B-to-ground represents a cut-off transistor, in Fig. 527.) $R_1$ dissipates any imbalances in power or phasing.

To find the input resistance to the network of Fig. 527, the network equations are written as follows:

$$\begin{aligned} I_1 &= I_2 = I_3 = I_4 & V_2 - V_4 &= V_4 - V_3 \\ I_5 &= I_6 = 2I_1 & V_1 &= 2(V_2 - V_3) \\ I_7 &= I_5 - I_8 & V_4 &= R_1 I_{10} \\ I_8 &= I_9 & V_2 &= I_7 R_2 \\ I_{11} &= I_9 + I_6 & V_3 &= R_3 I_{11} \\ I_{10} &= I_8 + I_9 & & \end{aligned} \qquad (394)$$

These equations yield $V_1/I_1$ as a function of $R_1$, $R_2$, and $R_3$:

$$R_{IN} = \frac{V_1}{I_1} = 16\left(\frac{R_1 R_2 + R_1 R_3 + R_2 R_3}{4R_1 + R_2 + R_3}\right) \qquad (395)$$

If $R_1 = 1/2R_2$ and $R_3 = 5R_2$, $R_{IN} = 16\ R_2$. Thus the 3.12-ohm-transistor resistance is transformed to 50 ohms.

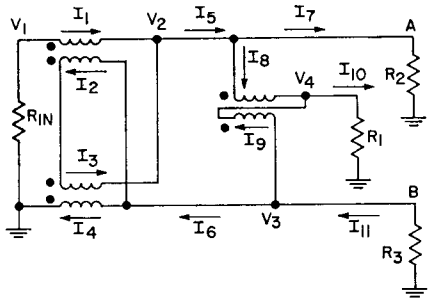



Figure 527. A 16:1 broadband transformation network with balanced output.

Because of symmetrical loading, the same hybrid configuration provides an 8:1 impedance transformation when used as a 180-degree-phase power combiner at the transistor collectors. This combiner operation of the network is shown in Fig. 528; the

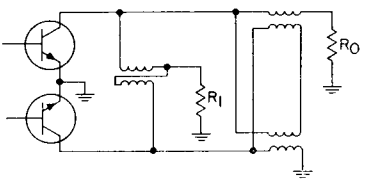



Figure 528. The network of Fig. 527 used as a 180-degree-phase power combiner.

output resistance is given by

$$R_{OUT} = \frac{V_{OUT}}{I_{OUT}} = 16\left(\frac{R_1R_2 + R_1R_3 + R_2R_3}{4R_1 + R_2 + R_3}\right) \quad (396)$$

If the collector load-line resistance is $R_L$, and if $R_1 = \frac{1}{2}R_L$ and $R_2 = R_3 = R_L$, then

$$R_{OUT} = 8R_L \quad (397)$$

Thus each collector is provided with a 6.25-ohm load line for $R_{OUT} = 50$ ohms. The inductance of the transmission line and its connectors is utilized to tune out both input and output negative reactances.

**2-to-30-MHz Circuit Design**—The push-pull configuration is used not only because the 180-degree-phase hybrids provide a high transformation ratio, but also because this configuration suppresses second harmonics and thus minimizes filter requirements at the output. If the output power level and the input and output impedance values at that power level are known, the circuit designer can use a combination of 180-degree-phase hybrids, hybrid resistance values, and additional transmission-line transformers to complete the proper transformation at the input and output. After the transformation closest to optimum match at the highest operating frequency has been selected, individual transformers are wound and measured over the desired frequency band. The HP 4815A vector impedance meter, RX Boonton Meter, or a similar instrument can be used for these measurements.

A 150-watt (PEP) linear amplifier for the 2-to-30-MHz frequencey range has been built with a pair of RCA-2N6093 transistors in push-pull, 180-degree-phase hybrid power combiner/dividers, and single-ended 4:1 transformers. The block diagram of this amplifier is shown in Fig. 529, and the circuit diagram is shown in Fig. 530.

Typical performance of this amplifier across the hf band is shown in Fig. 531. The power gain exhibits the same 6-dB-per-octave slope at mid-band low-frequency roll-off noted in the narrowband measurements (Fig.

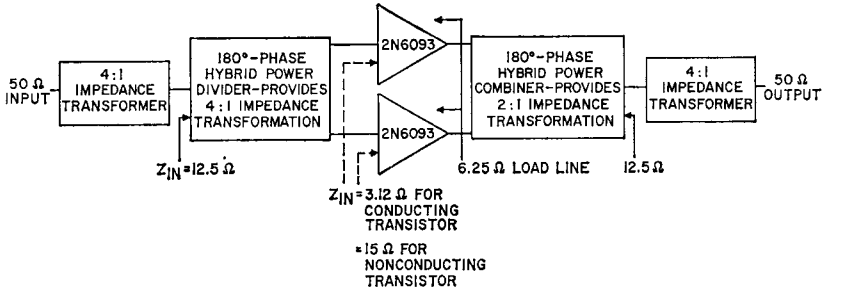
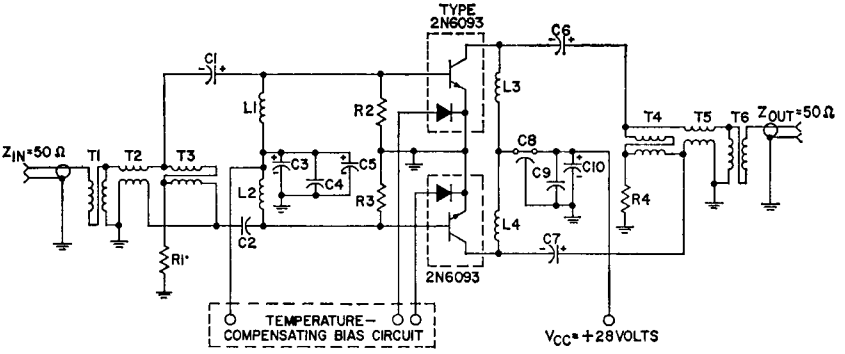



Figure 529. Block diagram of push-pull linear amplifier that provides 150 watts PEP




$C_1$, $C_2$ = 0.15 μF, electrolytic
$C_3$, $C_9$ = 0.04 μF, ceramic
$C_4$ = 0.0027 μF, ceramic
$C_5$ = 100 μF, electrolytic
$C_6$, $C_7$ = 0.1 μF, electrolytic
$C_8$ = 1000 pF, feedthrough, Allen-Bradley FA5C or equiv.
$R_1$ = 2.7 ohms, 0.5 watt in parallel with 3.3 ohms, 0.5 watt
$R_2$, $R_3$ = 30 ohms, 0.5 watt in parallel with 30 ohms, 0.5 watt
$R_4$ = 2.7 ohms, 1 watt
$L_1$, $L_2$ = 0.75 μH, RFC
$L_3$, $L_4$ = 7 turns of No. 20 wire wound on Indiana General Q1 CF-111 (or equiv.) ferrite core
$T_1$, $T_2$, $T_6$ = Two twisted pairs (9 turns per inch) of No. 26 wire paralleled 8 turns per inch; total length, 8 inches; five turns on Indiana General Q1 CF-111 (or equiv.) ferrite core
$T_3$, $T_4$, $T_5$ = Two copper strips (width = 85 mils, thickness = 8 mils) in parallel; total length, 3.5 inches; five turns on Indiana General Q1 CF-106 (or equiv.) ferrite core

Figure 530. Circuit diagram for 150-watt, 2-to-30-MHz push-pull linear amplifier.

521). Total gain variation is approximately 15 dB.

The intermodulation distortion exceeds −30 dB at frequencies below 6 MHz. The circuit is capable of −35 dB IMD over a good portion of the band if operated at the reduced output power of 100 to 110 watts PEP. If the same circuit components and transformation networks are utilized, the efficiency is somewhat reduced at

the reduced power level because the collector circuit is optimized for higher power.

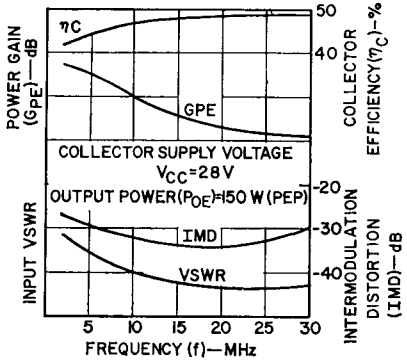



Figure 531. Typical performance of the broadband 150-watt (PEP) amplifier with two RCA-2N6093 transistors.

The efficiency of the amplifier is 40 to 50 per cent across the band. When operated at 150 watts PEP with a $V_{CC}$ of 28 volts, the amplifier becomes current limited at frequencies below 3 MHz. The increase in VSWR is related to the increase in the real part of the transistor input impedance (see Fig. 522).

Fig. 532 shows the performance of the 150-watt PEP amplifier as a function of case temperature at 30 MHz.

The main advantages of this type of circuit are its simplicity and compactness. The disadvantages are lack of gain leveling and low efficiency at lower frequencies because of increased VSWR.

Because the real value of the transistor input impedance increases with decreasing frequency, which affects both VSWR and IMD, a resistance-inductance series combination placed in parallel with the 50-ohm input or placed from base to base aids the transformation network in making a practical match at low frequencies. The impedance match is improved and some input power is absorbed at low frequencies; therefore, the VSWR improves and some gain leveling occurs. Other methods of gain leveling include collector-to-base feedback and loop feedback; for high-power circuits, the loop feedback system shown in Fig. 533 would be the most effective. In this system, input and output signals are compared, and gain differences are compensated by commensurate increases in input attenuation.

For higher powers, modules of push-pull pairs can be pyramided by the same hybrid-combining techniques.

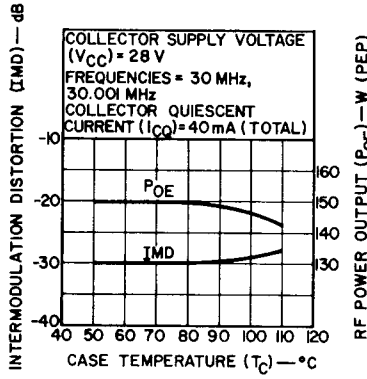



Figure 532. Performance of the 150-watt PEP amplifier as a function of case temperature at 30 MHz.

## RF AMPLIFIERS FOR MILITARY APPLICATIONS

The ruggedness, compactness, reliability, and efficiency of transistors make them especially useful in military environments. Air-

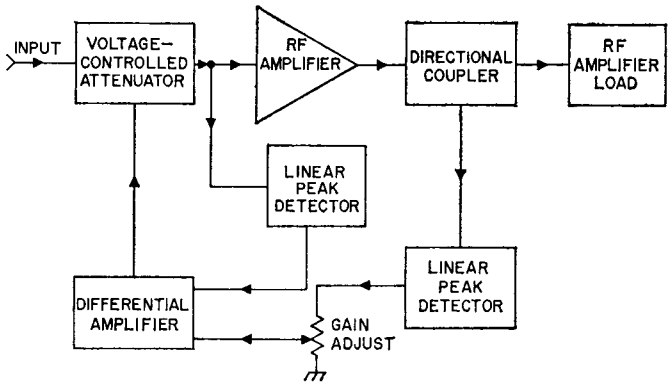



Figure 533. A loop feedback system for gain-leveling.

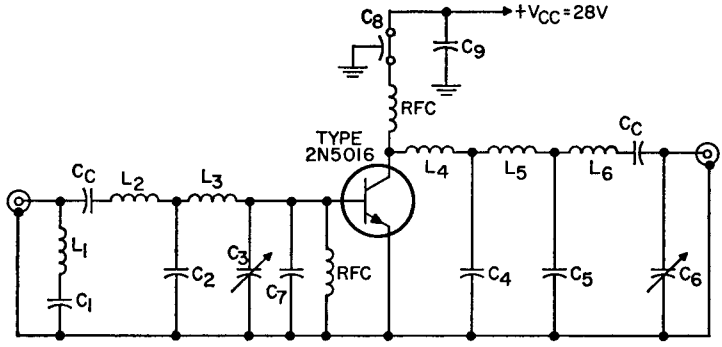
craft communication equipment, sonobuoy transmitters, and air-rescue beacons are typical military applications of rf power transistors.

## Military Aircraft Communications

The frequency range from 225 to 400 MHz is used in a large variety of relatively-high-power military communication systems. Equipments are usually amplitude-modulated and used for voice-communication purposes. The circuits discussed in this section are class B and class C amplifiers for use in driver or final output stages that provide power outputs in the range from 5 to 30 watts from a single transistor. Higher power can be obtained from combinations of transistors.

These amplifiers make extensive use of power combiners and broadband impedance matching. In an amplifier chain, all stages are designed to operate from a 50-ohm source into a 50-ohm load. For effective cascading, the input VSWR to any of the amplifiers in the chain must be as low as possible over the entire frequency band. Various techniques for broadbanding and for reducing input VSWR are discussed in following sections.

The lumped-constant circuit shown in Fig. 534 uses low-pass, LC ladder networks for impedance transformation. (The values given for the various components in the circuit diagram are measured at 400 MHz and include parasitic elements.) The output network transforms the 50-ohm load down to 20 ohms for the collector load. The dynamic output capacitance of the transistor provides the first shunt capacitor in the output network, and capacitor $C_C$ provides dc blocking. Similarly, the base input inductance of the 2N5016 transistor




$C_C$ = 2000 pF
$C_1$, $C_6$ = 7.5 pF
$C_2$ = 10 pF
$C_3$ = 1.5 to 30 pF (Johanson type or equiv.)
$C_4$ = 26.5 pF
$C_5$ = 17.5 pF
$C_7$ = 26.5 pF
$L_1$ = 4.5 nH
$L_2$ = 14 nH (includes inductance of input coupling capacitor $C_C$)
$L_3$ = 8.5 nH
$L_4$ = 5.6 nH
$L_5$ = 10 nH
$L_6$ = 19.5 nH (includes inductance of output coupling capacitor $C_C$)
**Note:**
All fixed components measured at 400 MHz

Figure 534. Lumped-constant 255-to-400-MHz power amplifier.

serves as the last series inductor of the input network, and capacitor $C_C$ is again used for dc blocking. The input match of the lumped-constant circuit is optimized at 400 MHz. The m-derived end section ($L_1$ and $C_1$) helps to provide the proper amount of mismatch at frequencies below 400 MHz to compensate for the gain characteristic of the transistor. With 6 watts of drive, this circuit provides approximately 17 watts of output power across the 225-to-400-MHz frequency band with a total output variation of 0.5 dB, as shown in Fig. 535.

Fig. 536 shows a schematic diagram of an amplifier that uses the RCA-2N5919. The circuit utilizes a lumped-element approach to broadband design. Typical amplifier performance is shown in Fig. 537. For a constant power output of 16 watts, response is fairly flat; the gain variation is within 1 dB across the band. Maximum input VSWR is 2:1.

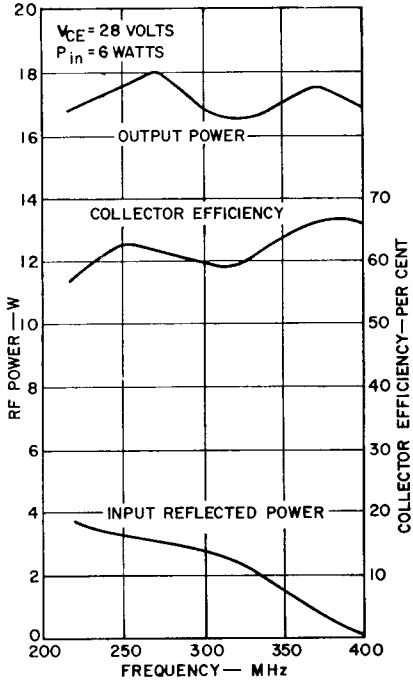
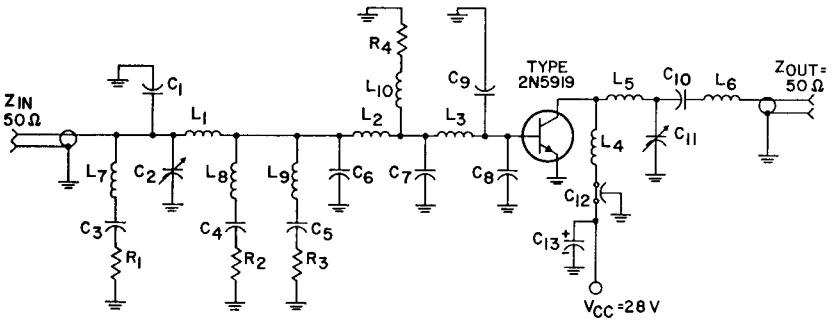



Figure 535. RF power output, input reflected power, and collector efficiency of the RCA-2N5016 transistor as functions of frequency.




$C_1$ = 10 pF silver mica
$C_2$ = 0.8-10 pF, Johanson 3957*
$C_3$ = 2.2 pF, Quality Components type 10% QC, "gimmick"*
$C_4$ = 1.0 pF, Quality Components type 10% QC, "gimmick"*
$C_5$ = 1.5 pF, Quality Components type 10% QC, "gimmick"*
$C_6$ = 36 pF, ATC-100*
$C_7$ = 51 pF, ATC-100*
$C_8$ = 47 pF, ATC-100*
$C_9$ = 68 pF, ATC-100*
$C_{10}$ = 12 pF, silver mica
$C_{11}$ = 0.8-20 pF, Johanson 4802*
$C_{12}$ = 1000 pF feedthrough type, Allen-Bradley FA5C*
$C_{13}$ = 1 μF electrolytic
$L_1$ = 1½ turns▲
$L_2$ = Copper strip ⅝ in. (15.875 mm) L; 5/32 in. (3.96 mm) W
$L_3$ = Transistor base lead, 3/6 in. (4.74 mm) L
$L_4$, $L_6$ = 3 turns▲
$L_5$ = 2 turns▲
$L_7$, $L_8$, $L_9$ = 0.18 μH RFC, Nytronics, P.#DD-0.18
$L_{10}$ = 0.1 μH RFC, Nytronics, P.#DD-0.10
$R_1$ = 100 Ω, 1 W, carbon
$R_2$, $R_3$ = 100 Ω, ½ W, carbon
$R_4$ = 5.1 Ω, ½ W, carbon
* Or equivalent
▲ All coils are 5/32 in. (3.96 mm) I. D., # 18 wire, 12 turns per inch.
Allen-Bradley Co., Milwaukee, Wis.
American Technical Ceramics, Huntington Station, N. Y. 11746
Johanson Mfg. Corp., Boonton, N. J. 07005
Nytronics, Inc., Berkeley Heights, N. J.

Figure 536. 16-watt broadband amplifier circuit using the RCA-2N5919.

Such flatness of response and low input VSWR are obtained by designing for the best possible match across the band and then dissipating some of the power at the low end of the band through dissipative RLC networks. The effectiveness of this technique can be evaluated by comparison of the gain and input VSWR curves in Fig. 537(a) with those in Fig. 537(b). The flatter the response, the smaller the dynamic range required in the output leveling system. Low input VSWR is necessary for effective cascading and protection of the driving stage in a cascade connection. The collector efficiency is not constant, but has a minimum value of about 63 per cent. The second harmonic of the 225-MHz signal is 12 dB

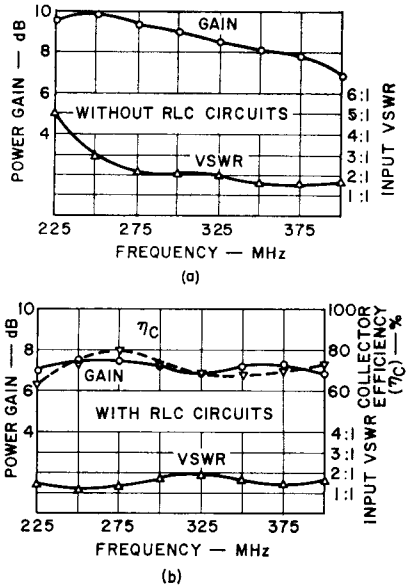
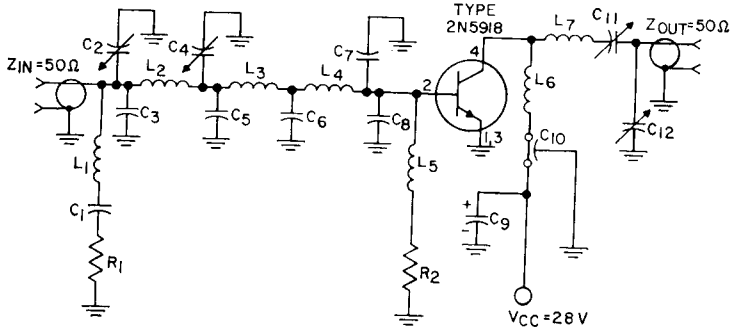



Figure 537. Typical performance of circuit of Fig. 536 from 225 to 400 MHz.

down and that of the 400-MHz signal is 30 dB down from the fundamental. Further reduction of the second harmonic of the 225-MHz signal is difficult to obtain because the amplifier bandwidth covers almost an octave.

In a cascade arrangement, a lower-power transistor, the 2N5918, is used to drive the 2N5919. The output circuit for the driver is modified to accommodate a higher collector load. The input circuit remains essentially the same as for the 2N5919. The 2N5918 amplifier schematic is shown in Fig. 538, and the performance of the two amplifiers connected in cascade is shown in Fig. 539. When the two stages are connected together, the broadband characteristics of the amplifiers minimize the number of adjustments required.

Fig. 540 shows the RCA-2N6105 high-power transistor in a 225-to-400-MHz broadband amplifier, and Fig. 541 shows the perform-




$C_1$ = 3 pF, ATC-100*
$C_2$ = 0.8-10 pF, Johanson 3957*
$C_3$ = 5 pF silver mica
$C_4$ = 2-18 pF, Amperex HTIOMA/218*
$C_5$ = 24 pF, silver mica
$C_6$ = 51 pF, ATC-100*
$C_7$ = 47 pF, ATC-100*
$C_8$ = 68 pF, ATC-100*
$C_9$ = 1 μF, electrolytic
$C_{10}$ = 1000 pF, feedthrough Allen-Bradley type, FA5C*
$C_{11}$ = 0.9-7 pF, Arco 400*
$C_{12}$ = 1.5-20 pF, Arco 402*
$L_1$ = 0.12μH RFC, Nytronics, P. No. DD-0.18*
$L_2$ = No. 18 wire, 0.64 in. long
$L_3$ = copper strip 5 mils thick, 150 mils wide, 670 mils long
$L_4$ = transistor base lead, 0.16 in. long
$L_5$ = 0.1μH RFC, Nytronics, P. No. DD-0.10*
$L_6$ = No. 18 wire, 1.08 in. long
$L_7$ = 2 turns, 5/32 in. I.D. No. 18 wire, 12 turns
$R_1$ = 100 ohms, ½ watt, carbon
$R_2$ = 5.1 ohms, ¼ watt, carbon
* Or equivalent

Figure 538. Driver amplifier using the RCA-2N5918.

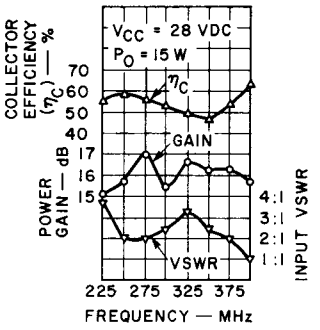
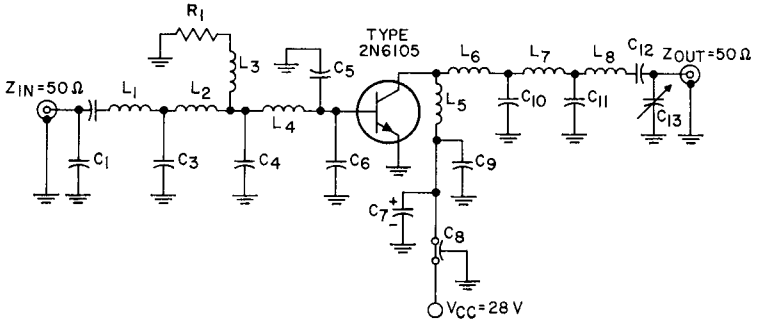



Figure 539. Performance characteristics of amplifiers shown in Figs. 536 and 538 connected in cascade.

ance of the amplifier. This circuit also utilizes lumped-circuit-element broadbanding. No special care is evident in this circuit to reduce the input VSWR. Two of these amplifiers can be combined by quadrature combiners, as shown in Fig. 542, to obtain higher output power. The input VSWR of the individual amplifier is not important in such a combination because of the high isolation characteristics of quadrature combiners; reflected power is dissipated in ports terminated with 50-ohm resistors. The performance of the 2N6105's combined by this method is shown in Fig. 543.

Another effective way to combine transistors is push-pull operation utilizing transmission-line techniques. The advantage




$C_1$ = 8.2 pF chip, Allen-Bradley*
$C_2$ = 18 pF silver mica
$C_3$ = 33 pF chip, Allen-Bradley*
$C_4$ = 47 pF chip, Allen-Bradley*
$C_5$ = 68 pF chip, ATC-100*
$C_6$ = 62 pF chip, ATC-100*
$C_7$, $C_8$ = 1000 pF, Feedthrough
$C_9$, $C_{12}$ = 1000 pF chip, Allen-Bradley*
$C_{10}$ = 22 pF chip, Allen-Bradley*
$C_{11}$ = 6.9 pF chip, Allen-Bradley*
$C_{13}$ = 0.8-10 pF variable air, Johanson No. 3957*

$L_1$ = 2 turns, 5/32-in. I.D. coil
$L_2$ = 17/32-in. long wire
$L_3$ = RFC, 0.1 μH, Nytronics*
$L_4$ = 5/32-in. long transistor base lead
$L_5$, $L_7$ = 13/16-in. long wire
$L_6$ = 9/16-in. long wire
$L_8$ = 7/8-in. long wire
$R_1$ = 5.0 Ω, ¼ W
All wire is No. 20 AWG

* Or equivalent

Figure 540. 225-to-400-MHz broadband amplifier using RCA 2N6105.

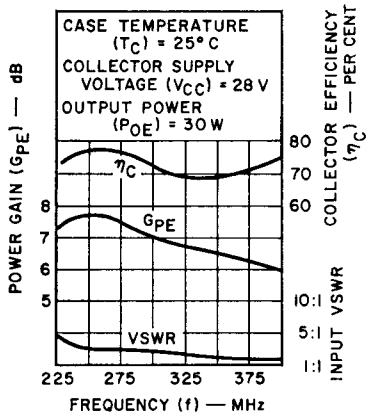



Figure 541. Typical performance of a 225-to-400-MHz broadband amplifier using RCA 2N6105 at $V_{CC}$ = 28V.

of low second harmonic in the push-pull configuration is especially important in the 225-to-400 MHz frequency band; because the second harmonic of the low frequency falls just outside the band, filtering it out presents considerable difficulty. The use of transmission lines results in a compact, relatively simple structure. The input VSWR in a push-pull amplifier is very high at the low end of the band, so this type of circuit is especially suitable for use with quadrature combiners. Fig. 544 shows details of an individual push-pull amplifier using two RCA-2N6105 transistors. Fig. 545 shows two of these amplifiers combined by quadrature combiners to make up a 100-watt broadband module. This approach results in a saving of four combiners, at least two of which are quadratures. The performance of this module for 100-watt constant output is shown in Fig. 546.

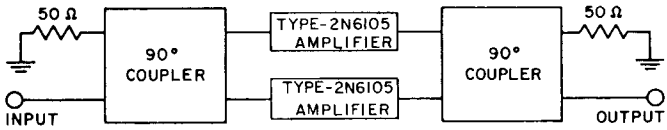



Figure 542. Two RCA-2N6105 amplifiers connected in parallel by use of quadrature couplers.

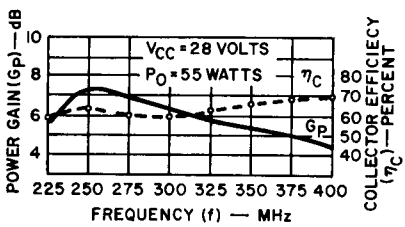
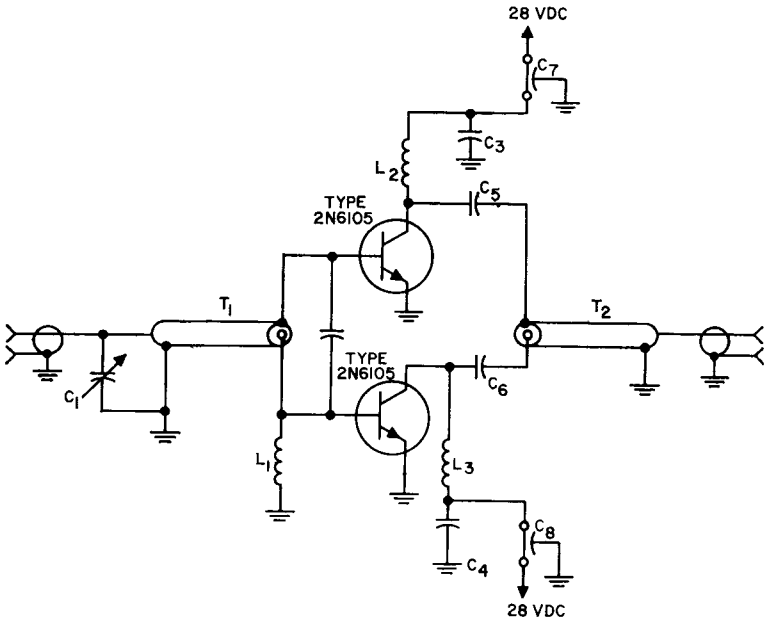



Figure 543. Power gain and efficiency as function of frequency for the broadband module shown in Fig. 542.

## Sonobuoy Transmitters

A sonobuoy is a floating submarine-detecting device that incorporates an underwater sound detector (hydrophone). The audio signals received are converted to a frequency-modulated rf signal which is transmitted to patrolling aircraft or surface vessels. The buoy is battery-operated and is designed to have a very limited active life.

Typical requirements for the rf-transmitter section of the sonobuoy are as follows:




$C_1$ = 0.8 to 10 pF, piston trimmer
$C_2$ = 56-pF chip, ATC-100 or equiv.
$C_3$, $C_4$, $C_5$, $C_6$ = 1000-pF chip, Allen-Bradley type or equiv.
$C_7$, $C_8$ = 1000 pF, feedthrough
$L_1$ = 0.18 μH, RFC, Nytronics type or equiv.
$L_2$, $L_3$ = ¾-inch-long No. 20 wire
$T_1$ = coaxial line, $Z_o$ = 25 ohms, 3¾ inches long
$T_2$ = coaxial line, $Z_o$ = 25 ohms, 4½ inches long

Figure 544. 225-to-400-MHz broadband push-pull amplifier using two 2N6105's.

Frequency = 165 MHz
Supply Voltage = 8 to 15 volts
CW Output = 0.25 to 1.5 watts
Over-all Efficiency = 50 per cent
Harmonic Output = 40 dB down from carrier

Fig. 547 shows a diagram of an experimental sonobuoy transmitter designed to produce a power output of 2 watts at 160 MHz. Only three stages, including the crystal-controlled oscillator section, are required. Efficiency is greater than 50 per cent (over-all) with a battery supply of 12 to 15 volts.

The 2N3866 or 2N4427 transistor can be used in a class A oscillator-quadrupler circuit which is capable of delivering 40 milliwatts of rf power at 80 MHz.

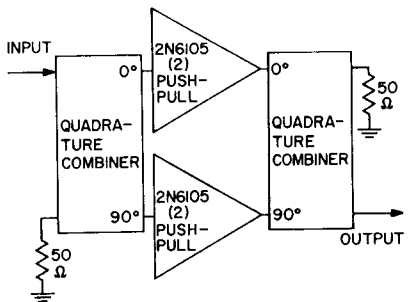



Figure 545. 100-watt 225-to-400-MHz broadband module.

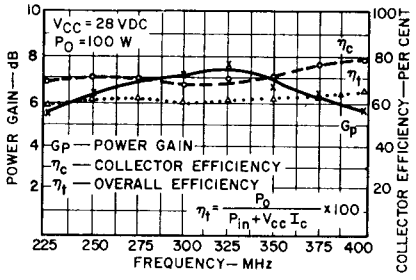
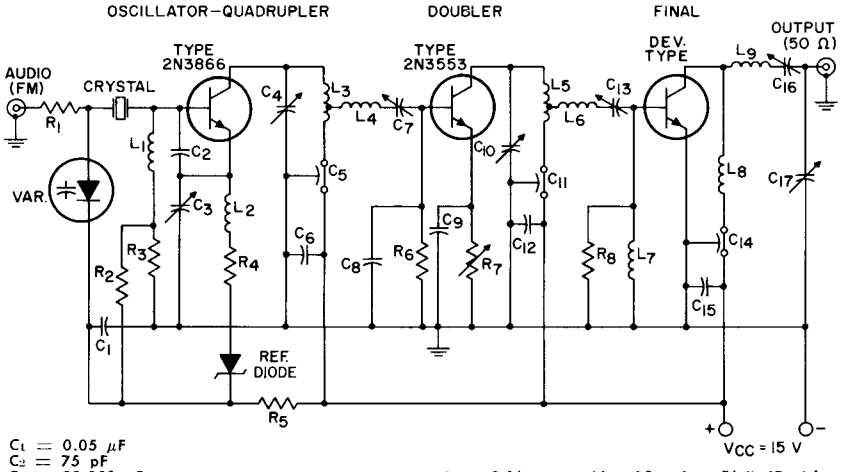



Figure 546. 100-watt, 225-to-400-MHz broadband module performance.

Narrow-band frequency modulation is accomplished by "pulling" of the crystal oscillator. The crystal is operated in its fundamental mode at 20 MHz. The oscillator is broadly tuned to 20 MHz in the emitter circuit and is sharply tuned to 80 MHz in the collector circuit. The supply voltage to the oscillator section is regulated at 12 volts by means of a zener diode. Spectrum-analyzer tests indicate that this stage is highly stable even though rather high operating levels are used.

The oscillator-quadrupler section is followed by a 2N3553 class C doubler stage. This stage delivers a power output of 250 milliwatts at 160 MHz from a 12- to 15-volt supply. The over-all output of the sonobuoy can be adjusted by varying the emitter resistance of this stage.

The final power output is developed by an RCA developmental transistor which operates

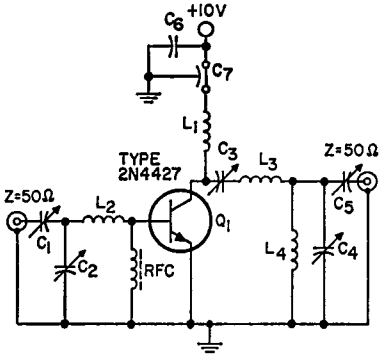



$C_1$ = 0.05 μF
$C_2$ = 75 pF
$C_3$ = 32-250 pF
$C_4$, $C_{10}$, $C_{16}$, $C_{17}$ = 3-35 pF
$C_5$ = 2200 pF, feedthru
$C_6$, $C_{12}$, $C_{15}$ = 0.01 μF
$C_7$ = 14-150 pF
$C_8$ = 50 pF
$C_9$ = 500 pF
$C_{11}$, $C_{14}$ = 1500 pF, feedthru
$C_{13}$ = 8-60 pF
$L_1$ = 22 μH
$L_2$ = 5.5 μH
$L_3$ = 3 turns No. 16 wire, ¼" ID x ½" long
$L_4$ = 5 turns No. 22 wire, 3/16" ID (close wound)
$L_5$ = 2-¼ turns No. 16 wire, ¼" ID x ⅜" long
$L_6$ = 2-¼ turns No. 18 wire, 7/32" ID (close wound)
$L_7$, $L_8$ = 1.0 μH
$L_9$ = 5 turns No. 16 wire, ¼" ID x ⅜" long
$R_1$ = 1000 ohms
$R_2$ = 1200 ohms
$R_3$ = 47 ohms
$R_4$ = 10 ohms
$R_5$ = 100 ohms
$R_6$ = 51 ohms
$R_7$ = potentiometer, 50 ohms
$R_8$ = 100 ohms
Var = varactor
Ref. Diode = 12-volt zener diode

Figure 547. 1.5-watt (rf power output) sonobuoy transmitter.

as a straight-through class C amplifier at 160 MHz. A pi network matches this output to the 50-ohm line. The spurious output (measured directly at the output port) is more than 35 dB down from the carrier. This suppression is achieved by means of series resonant trap circuits between stages and the use of the pi network in the output.

Many sonobuoy systems require power outputs in the range of only 0.25 to 0.5 watt, preferably with a supply voltage of 8 to 12 volts. The 2N4427 transistor is suitable for use as the doubler and also the final output device in such low-power applications. Fig. 548 shows a diagram of an output stage which uses the 2N4427 as a straight-through 175-MHz class C amplifier. This circuit can deliver output power of more than 500 milliwatts with a supply voltage of 10 volts and a drive power of 60 milliwatts.




$C_1$, $C_2$, $C_4$, $C_5$ = 7-to-100 pF Arco 423 or equiv.
$C_3$ = 14-to-150 pF, Arco 424, or equiv.
$C_6$ = 0.01 $\mu$F, 50 V
$C_7$ = 1000 pF, feedthru
$L_1$ = 0.75 $\mu$H
$L_2$ = 1 turn No. 18 wire, 5/32" ID
$L_3$ = 1½ turns No. 18 wire, ¼" ID
$L_4$ = 1¼ turns No. 18 wire, 3/16" ID
RFC = 450 ohms, ferrite

Figure 548. 0.5-watt 175-MHz sonobuoy rf power output stage.

For the lower power-output requirement at low supply voltages, the oscillator-quadrupler stage should use lower-power transistors such as the 2N1491 or 2N914. Only 10 to 15 milliwatts of fourth harmonic power is required in this case. The bias-network resistors ($R_2$ and $R_3$) should be adjusted for reliable oscillator starting conditions at the lower supply voltages.

Sonobuoy circuits, in general, must be reliable, simple, and low in cost. The three-stage transmitter circuit shown in Fig. 547 is intended to be representative of the general design techniques used in these systems. However, four-stage sonobuoy transmitter systems are also in common use at the present time. Typically, a four-stage arrangement consists of an oscillator-tripler stage, a second tripler stage, a buffer stage, and a final amplifier stage. Most present-day sonobuoy applications require cw power output between 0.25 and 1.5 watt.

## Air-Rescue Beacons

The air-rescue beacon is intended to aid rescue teams in locating airplane crew members forced down on land or at sea. The beacons are amplitude-modulated or continuous-tone line-of-sight transmitters. They are battery-operated and small enough to be included in survival gear.

Typical requirements for rescue beacons are as follows:

Frequency = 243 MHz (fixed)
Power Output = 300 milliwatts (carrier)
Efficiency = greater than 50 per cent

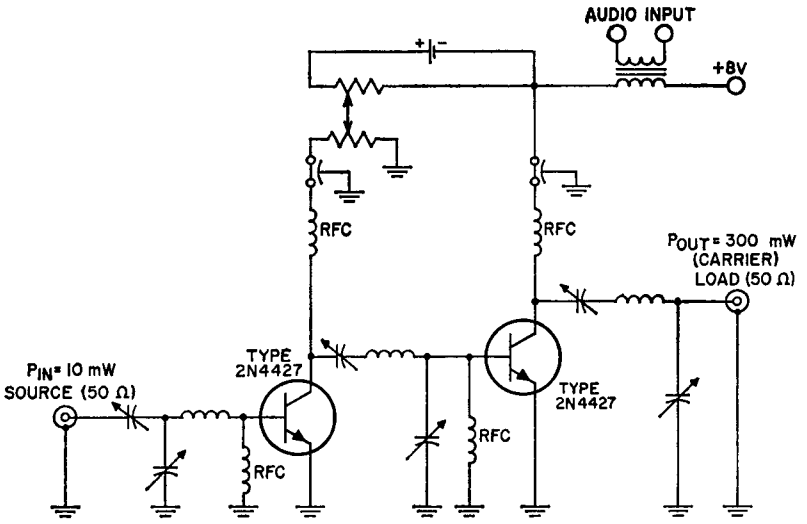



Figure 549. Driver and output stage for a 243-MHz beacon transmitter.

Supply Voltage = 6 to 12 volts
Modulation = AM, up to ±100 per cent

The 2N4427 transistor is especially suited for this service. A general circuit for the driver and output stages is shown in Fig. 549. Collector modulation, as well as some driver modulation, is used to achieve good downmodulation of the final amplifier. Conventional transformer-coupled modulation is used; however, a separate power supply and resistor network in the driver circuit are provided to adjust the modulation level of this stage independently of the output stage.

The rf-amplifier design is conventional; pi- and T-matching networks are used; simpler circuits (e.g., device-resonated tapped coils), however, could be used. The T-matching network at the driver input is used to match the amplifier to a 50-ohm source for test purposes. A 10-to-20-milliwatt input signal is needed to develop a 300-to-400-milliwatt carrier output level.

## Miniaturized Low-Power Oscillators

Low-power transistor oscillators are used as transmitters for telemetering or signal use in such devices as radiosondes, military fuses, beacons, and other remote sensing devices. Many of these units currently operate in the uhf range at output levels of about 0.25 to 1 watt. Battery supplies are normally used.

The 2N3866 and 2N4427 transistors are ideally suited for low-power oscillator service. Fig. 550 shows a simple microstripline circuit in which these transistors can provide power outputs of up to 1 watt in the frequency range of 400 to 600 MHz. The frequency of oscillation is primarily determined by capacitor C and the parasitic emitter-

lead inductance. The microstripline output circuit can be matched to a wide range of loads by use of taps along the line length.

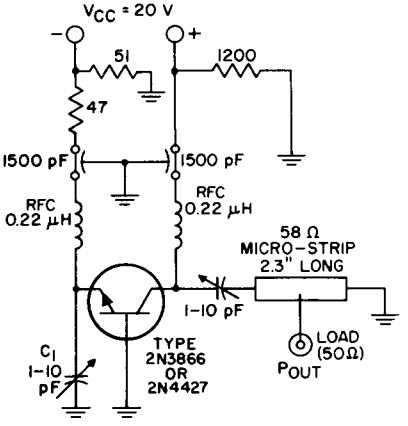



Figure 550. 1-watt, 500-MHz microstripline oscillator using the RCA 2N3866 or 2N4427 transistor.

Fig. 551 shows a very simple lumped-constant oscillator circuit for operation in the 700-to-1000-MHz frequency range. The parasitic emitter- and base-lead inductances are tuned directly with high-Q air dielectric capacitors, and no other external inductances are required for this frequency range. The collector is grounded directly to the ground plane for best dissipation of transistor heat. Capacitor $C_1$ primarily determines the oscillator frequency, and the output capacitors are used primarily for impedance matching. The 2N3866 is used for operation at supply voltages of 20 to 28 volts, and the 2N4427 is preferred for supply voltages of 15 to 20 volts. Power outputs in the order of 500 to 1000 milliwatts into a 50-ohm load can be developed by this simple circuit.

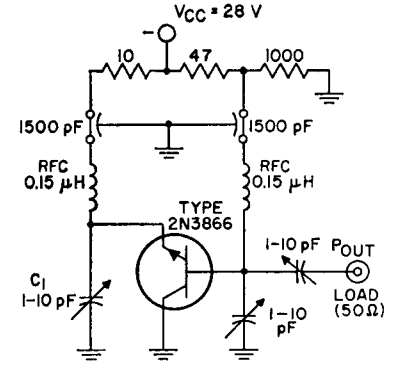



Figure 551. 0.5-watt, 1000-MHz lumped-constant oscillator using the RCA-2N3866 transistor.

## MOBILE AND MARINE RADIO

In the United States, three frequency bands have been assigned to two-way mobile radio communications by the Federal Communications Commission. These frequency bands are 25 to 50 MHz, 148 to 174 MHz, and 450 to 470 MHz. The low-frequency band for overseas mobile communications is 66 to 88 MHz.

Frequency modulation (FM) is used for mobile radio communications in the United States and most overseas countries. The modulation is achieved by phase-modulation of the oscillator frequencies (usually the 12th or 18th submultiple of the operating frequency). In vhf bands, the frequency deviation is ±5 kHz and channel spacing is 25 kHz. In uhf bands, at present, the modulation deviation is ±15 kHz and channel spacing is 50 kHz. In the United Kingdom, AM as well as FM is used in mobile communications.

The minimum mobile-transmitter power-output levels in the United States are 50 watts in the 50-MHz band, 30 watts in the 174-MHz band, and 15 watts in the

470-MHz band. Some of the transmitters used in the United States have power-output ratings well in excess of 100 watts. Overseas power requirements are more moderate and are often regulated by law; the most common power output level is 12 watts at the antenna.

## Transistor Requirements

The transistors in the rf power stages are the heart of every solid-state transmitter. In fact, the present power, load mismatch, and frequency capabilities constitute major design limitations for new mobile transmitters; the extension of the present limits are often achieved through transistor design tradeoffs. For instance, a high $f_T$ is necessary for gain optimization; however, a transistor with high $f_T$ is highly susceptible to mismatch in the load condition. Therefore a lower, flatter current-gain curve, which would allow good rf current gain at high currents, is a better choice. An indication of the frequency capability is the ratio of emitter periphery to base area, called the "design ratio;" a design ratio of 2 to 3 provides adequate gain and rugged performance.

## Package Considerations

The rf package design plays an important role in determining mobile transistor requirements. The minimization of emitter lead inductances, for instance, can decrease degeneration and significantly increase power gain; the elimination of basic parasitics can have a dramatic bandwidth widening effect. The thermal capability of the package determines the junction temperature at any given output power; a lower package $R_{\theta J\text{-}C}$ can increase high power performance (and mismatch capability) by allowing the transistor chip to operate at lower effective temperatures. Reliability can be greatly improved by both proper package-chip interfaces and hermetically isolating the chip from environmental changes.

Because of low device impedances, especially in 12 volt devices, all package losses must be eliminated; bond wires must be kept as small as possible.

## DC Operating Voltages

All-solid-state mobile transmitters can be divided into two basic types: transmitters that operate from 24-to-28-volt collector supply voltages, obtained from dc-to-dc converters, and transmitters that operate directly from the 12-volt electrical system of a vehicle.

Both types have advantages and disadvantages. The advantages of 24-to-28-volt operation include higher power gains per stage, good transient suppression, and fairly simple current and voltage limiting. The disadvantages are the additional cost of dc-to-dc converters and the somewhat higher power consumption and increased size of the radio. Direct operation from a 12-volt system permits savings in cost and size, as well as higher efficiency. Because 12-volt operation produces less gain per stage, however, additional rf stages are often needed. Transient suppression and voltage and current limiting are also somewhat more difficult. However, the savings in cost and size make 12-volt systems somewhat more de-

sirable. For this reason most mobile transmitters are designed for 12-volt operation.

Because of the two discrete voltage ranges used for mobile radios, the transistor must be designed specifically for either 24-to-28-volt operation or 12-volt operation. Devices designed for 24-to-28-volt operation have substantially higher collector-breakdown voltages. Devices designed for the 12-volt radio must have substantially higher current-handling characteristics.

## Matching Networks

The design of high-power, high-frequency transistor amplifiers presents unique problems. Low operating voltages and relatively high power levels result in impedances that become very small and circulating rf currents that become very large. For example, if an rf power output of 60 watts is required from an amplifier operating directly from a 12-volt supply, the collector load impedance to the final amplifier must be approximately 1 ohm. Under these conditions, the peak current can be as high as 20 amperes. At the same time, the series input impedance will be substantially below 1 ohm. Similar conditions often hold for powers as low as 10 watts. Because of the small magnitudes, all matching elements must have as high a Q as possible, line lengths must be minimized, and all stray inductances must be eliminated.

Microstrip circuit elements can provide high Q at low cost; Fig. 552 shows tunable and fixed-tuned microstrip circuits.

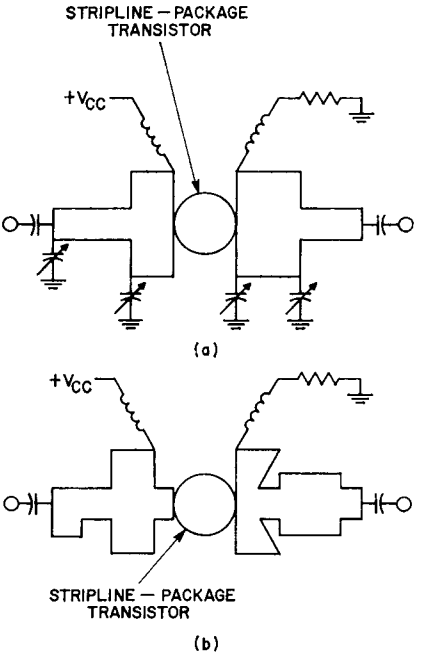



Figure 552. (a) Tunable and (b) fixed-frequency microstrip circuits.

## Instabilities in VHF/UHF Transistor Amplifiers

In vhf/uhf transistor power amplifiers, the most common instabilities occur at frequencies far below operating frequencies because the gain of the transistor increases at a rate of approximately 6 dB per octave as the frequency decreases. For example, a device that has a power gain of 5 dB at 174 MHz may have a gain of as much as 30 dB at 10 MHz. With such high gain, any kind of stray low-frequency resonant circuit can set the circuit into violent oscillation and even cause destruction of the transistor.

These low-frequency oscillations can be prevented by means of the following simple precautions, as indicated in Fig. 553:

(1) Because the base-emitter junction is highly capacitive at low frequencies, a resonant circuit can be easily formed with this capacitance and the choke RFC.

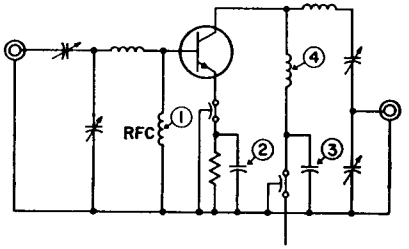



Figure 553. Circuit indicating areas where simple precautions prevent low-frequency oscillations.

This low-frequency resonant circuit can be avoided by the replacement of RFC with a low-Q, ferrite-type choke or even a wire-wound resistor.

In uhf circuits, a 0.22-microhenry choke in series with a 0.24-to-1.5-ohm resistor provides both a stable dc return and some base bias for better efficiency.

(2) The emitter bypassing should be effective not only at operating frequencies, but also at low frequencies; thus, two bypassing capacitors should be used. One of these capacitors should be effective at the operating frequency, and the other at low frequencies.

(3) DC-power wiring should have adequate bypassing both at operating and low frequencies to shunt out stray inductances in the wiring.

(4) Output-matching networks should make use of a coil as an integral part of matching for feeding dc to the collector. As a rule, the inductance of these coils is much smaller than that of self-resonant rf chokes, and thus the reactances are lower at low frequencies.

In higher-power uhf circuits, however, it is often desirable to series-tune the collector capacitance; this arrangement can result in better harmonic supression, and better transistor efficiency.

## Reliability

Mobile radio applications place severe requirements on transistor operation. At full rated input and output power, the device must consistently survive load mismatches from short to open circuit; this condition often occurs simultaneously with an increase in $V_{CE}$. The ability of a device to both survive such conditions and to avoid any permanent degradation can be realized only by uncompromising transistor design procedures coupled with the best available engineering judgment. Emitter ballasting, which forces transistor current to be equally shared through all active transistor emitter regions, is critical for ruggedness as are proper frequency and power tradeoffs. Because of the powers involved, thermal resistance must be minimized to avoid excessive junction temperatures. Long-term reliability can be enhanced by protection of critical junctions from shorting caused by migration, and by use of a hermetically sealed package to protect the sealed pellet.

The mobile transmitter manufacturer also contributes to reliability by proper heat-sink design and by building in voltage regulation and VSWR protection.

## 470-MHz Power Amplifier

Fig. 554 shows a 470-MHz amplifier chain that incorporates many of the techniques previously discussed. Operating from a 12.5-volt supply, the chain delivers 30 watts at 470-MHz, with an input of 150 milliwatts. The over-all efficiency for the chain is in excess of 47 per cent. The driver stages provide output powers of 0.8, 3, and 11 watts.

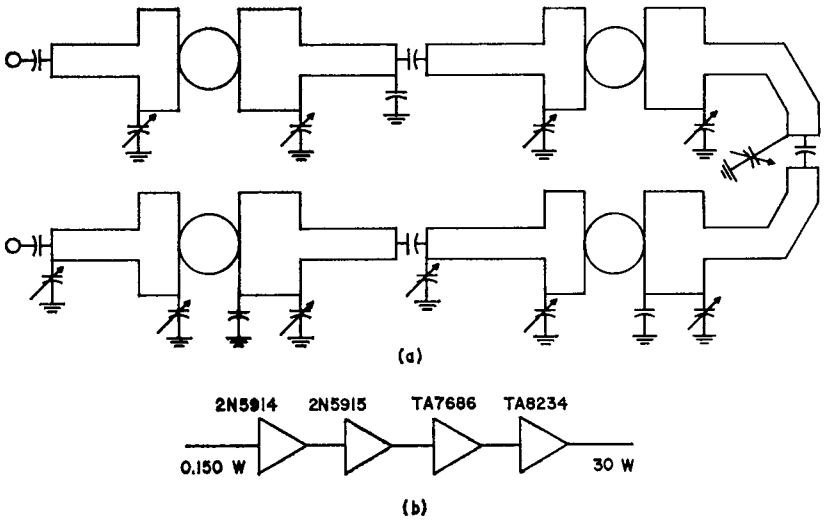



Figure 554. 30-watt, 470-MHz amplifier chain.

## 66-to-88-MHz Band

Transmitters that operate directly from vehicle batteries in the 66-to-88-MHz frequency band can use either AM or FM.

In an AM transmitter, the peak envelope power $P_p$ of a collector-modulated transistor is

$$P_p = P_c(1 + m)^2 \qquad (398)$$

where $P_c$ is the carrier power and m is the modulation index. If $m = 1$, the peak envelope power is four times the unmodulated carrier power. The peak rf voltage on the collector under 100-per-cent modulation is at least four times the supply voltage for the AM transistor. A suitable transistor for AM operation therefore must have high current handling capability for good upward modulation and high voltage ratings to prevent second breakdown.

The 2N5992 provides 7 watts of carrier power with a minimum power gain of 10 dB at 88 MHz. Modulation higher than 90 per cent can easily be achieved with the driver slightly modulated. Emitter-site-ballasting employed in this device not only improves the modulation capability, but

also enables the device to withstand high VSWR caused by antenna mismatch.

For FM applications, the 2N5993 can provide 18 watts of cw power with a power gain greater than 10 dB in the 66-to-88-MHz band. To insure device ruggedness, the 2N5992 and 2N5993 transistors are 100-percent tested under infinite VSWR through all phases at rated output power. For the AM type, the test is performed under full modulation. A typical circuit arrangement for testing output power, power gain, modulation index, and load-mismatch capability is shown in Fig. 555.

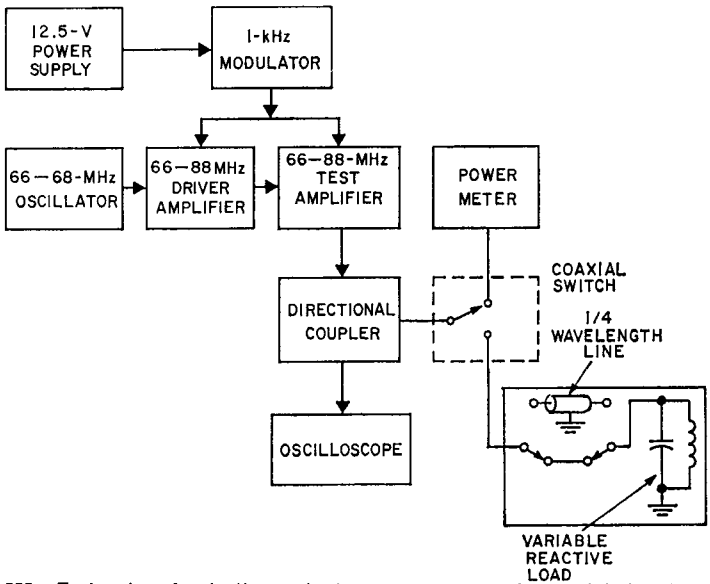



Figure 555. Test setup for testing output power, power gain, modulation index, and load-mismatch capability.

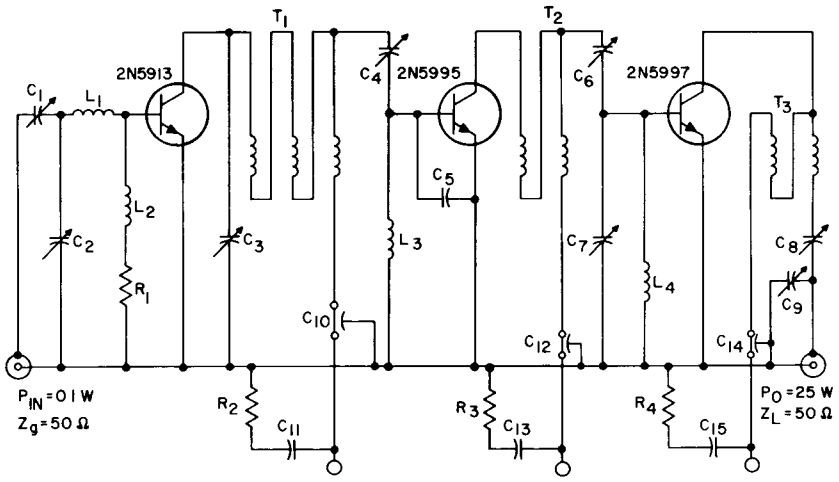
## 175-MHz Band

The circuit shown in Fig. 556 provides 25 watts for FM mobile transmission. It operates from a 12.5-volt supply and requires a drive power of 0.1 watt at 175 MHz. Transformers are used for interstage coupling, with matching networks that allow the circuit to be tuned for optimum performance.

Some of the design considerations discussed above for AM transmitters also apply to the design of FM circuits. In an FM transmitter, the transistor requirements are less stringent because the collector breakdown voltage needs to be only approximately twice the supply voltage, and the transistor can be driven very hard into saturation. In the absence of collector modulation, parasitic oscillations in a FM amplifier are also minimized.

## 156-MHz Marine Band

The increasing number of boating enthusiasts has generated a demand for low cost reliable communication equipment operating in the 156-MHz marine band. The amplifier chain shown in Fig. 557



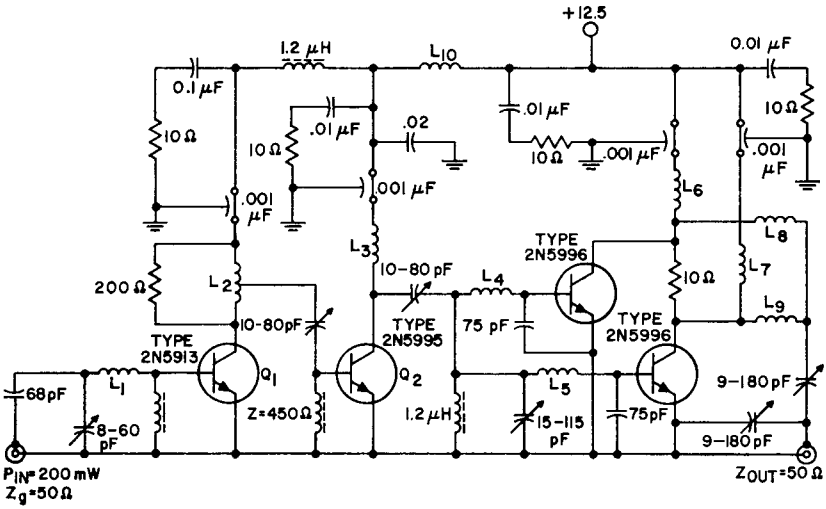

$C_1$, $C_2$ = 1.5 to 20 pF, Arco No. 402 or equiv.
$C_3$ = 10 pF, mica
$C_4$, $C_8$, $C_9$ = 14 to 150 pF, Arco No. 424 or equiv.
$C_5$ = 56 pF, mica
$C_6$, $C_7$ = 7 to 100 pF, Arco No. 423 or equiv.
$C_{10}$, $C_{12}$, $C_{14}$ = 1000 pF, feedthrough
$C_{11}$, $C_{13}$, $C_{15}$ = 0.01 μF, ceramic
$R_1$ = 33 ohms, ½ watt
$R_2$, $R_3$, $R_6$ = 10 ohms, ¼ watt
$L_1$ = 3 turns of No. 22 enamel wire, ¼-inch inner diameter, close wound
$L_2$, $L_3$, $L_4$ = No. 22 wire thread through Ferroxcube bead No. 56-590-65-4A or equiv.
$T_1$ = 3 twisted No. 22 wires, approximately 14 turns per inch, formed into a loop; ⅜-inch inner diameter; cross connected (i.e., end of one wire connected to beginning of the other)
$T_2$, $T_3$ = same as $T_1$, except only 2 twisted wires

Figure 556. 25-watt amplifier for mobile FM transmission.

incorporates the techniques discussed for calculating matching networks and avoiding low-frequency oscillations. The amplifier operates from a 12.5-volt collector supply and delivers 32 watts at 156 MHz with an input power of 200 milliwatts. The design is quite conservative; the output stage can withstand any load mismatch at its rated output power.

## COMMERCIAL AIRCRAFT RADIO

The aircraft radios discussed in this section are of the type used for communication between the pilot and the airport tower. The transmitter operates in an AM mode on specific channels between 118 and 136 MHz. Radios of this type are regulated by both the FCC (Federal Communications Commission) and the FAA (Federal Aeronautics Administration). The FCC assigns frequencies to airports and places some requirements on the transmitters, particularly as regards spurious radiation and interference. The FAA sets minimum requirements on radio performance which are based on the maximum authorized altitudes for the plane, whether paying passengers are carried, and on the authorization for instrument flying. The FAA gives a desirable TSO certification to radio equipment that satisfies their standards of airworthiness.




$L_1$ = 2 turns of No. 20 B.T., ¼-inch diameter, ⅛ inch long
$L_2$ = 5 turns of No. 20 B.T., ¼-inch diameter, ⅜ inch long, tap at 4½ turns from collector
$L_3$ = 5 turns of No. 20 enamel wire, 3/16-inch diameter, ¼ inch long
$L_4$, $L_5$ = 1 turn of No. 20 B.T., ⅛-inch diameter, ⅛ inch long
$L_6$, $L_7$ = 2 turns of No. 20 B.T., 3/16-inch diameter, ¼ inch long
$L_8$, $L_9$ = 2 turns of No. 18 B.T., ¼-inch diameter, 3/16 inch diameter
$L_{10}$ = 10 turns of No. 20 enamel wire, ¼-inch diameter, close wound

Figure 557. 32-watt, 156-MHz marine-band amplifier.

The FCC checks aircraft-radio transmitter designs for interference and other electrical characteristics (as it does all transmitters). Additional requirements are specified for radios intended for use by scheduled airlines by a corporation supported by the airlines themselves. The name of this corporation is ARINC (Aeronautical Radio, Inc., 2551 Riva Road, Annapolis, Maryland 21401).

All these specifications combine to generate radio-transmitter requirements for different types of aircraft, as indicated in Table XL.

## Desirable Features

Because multiple channel use is necessary, it is desirable that aircraft radios have all 360 channels. These channels are spaced every 50 kHz from 118 to 136 MHz, and are assigned to specific airports. Each must be crystal-controlled. Synthesizer techniques are used to reduce the number of crystals required.

Simple, foolproof operation is necessary because the pilot has little time to spare and little interest in adjustments to the radio equipment. The frequency settings are made by switches that provide a digital read-out. "Squelch," volume, and on/off controls are added.

Size is important because the instrument panel is crowded. On large aircraft, the transmitter is

Table XL—Four Popular Aircraft-Radio Transmitters
(Designs by Aircraft Type*)

| TYPICAL OWNER | NO. OF ENGINES IN AIRCRAFT | FAA & ARINC CLASS | VOLTAGE AVAILABLE | TRANSMITTER POWER (MIN.) | TYPICAL POWER RANGE | | TRANSMITTER FEATURES |
|---|---|---|---|---|---|---|---|
| Private Planes | 1 | I | 13 V | 1 W | >1.5 W | Type #1 | Low cost, few channels, may be portable |
| Owner/Pilot | 1 | I | 13 V | 4 W | >6.0 W | Type #2 | Panel mounted, 90 or 360 channels. |
| Private/Business | 2 | II | 28 V | 4 W | 6 to > 20 W | Type #2 | |
| Chartered & Cargo | 2-4 | III | 28 V | 16 W | >20 W | Type #3 | Remote Operation, 360 channels. |
| Scheduled Air Lines | 2-4 Jets | III & ARINC | 28 V | 25 W | 30 W | Type #4 | Maximum reliability |

* This chart is not complete or exact and is not intended to show actual requirements, but merely what is typical. Consult FAA for complete requirements.

operated by remote control by means of a set of switches on the panel. Weight and power drain are secondary considerations.

A primary consideration in all aircraft equipment is reliability. Spare radios are common in private aircraft, and are universal in aircraft equipped for instrument flying. The inherent reliability of transistorized equipment is a major advantage in aircraft radios.

## Design Problems

Amplitude modulation is an important design consideration for all transistor power amplifiers (as explained in the general section on AM). Amplitude-modulation requirements are set by the TSO at a minimum of 85 per cent, which corresponds to a PEP of 3 times the carrier power. Careful design is required to meet this specification because many factors tend to limit the PEP, including the decrease in transistor gain at high currents, transistor rf $V_{CE}(sat)$, and modulator losses.

The owners and pilots of aircraft require reliable, foolproof operation of their radio equipment. Unfortunately, they are not often technically trained and do not appreciate the importance of proper maintenance of the antenna and the transmission line. These vulnerable items directly affect the performance of the radio because optimum performance is achieved only when the transmission line VSWR is unity. With a mismatch (i.e., VSWR greater than 1), the power output may be low, and there may be spurious or distorted output. Even more important is the fact that antenna and transmission-line faults stress the transmitter output stage with high voltage-current products and/or high power dissipation. These characteristics can overstress and destroy a weak transistor. The likelihood and the drastic effects of a load mismatch make the transmitter output transistor a primary influence on equipment reliability and make mandatory the selection of a transistor rugged enough to withstand the possible stresses.

Aircraft radios must cover the entire frequency range from 118 to 136 MHz. The more expensive radios cover all 360 channels. This 18-MHz bandwidth is a major design challenge which may be met by use of either a narrow-band step-tuned transmitter or a broad-band transmitter.

A broad-band transceiver design is possible with transistors. In a power transistor, the input may be considered to consist of the base-lead inductance $L_b$ in series with $r_{bb}'$. If $L_b$ is minimized, the Q is reduced, and broad-band operation is possible. The output-circuit Q is less of a problem than that of the input. The Q is formed by $C_{ob}$ in parallel with the load impedance presented to the collector by the tuned circuit. Broad-band matching circuits between amplifier stages commonly use ferrite-core transformers of the transmission-line type (balun).

The use of broad-band amplifiers permits the largest portion of the transmitter to be remotely located without the need for expensive and complex servo tuning mechanisms. This feature is a great advantage in larger aircraft.

One problem encountered with a broad-band amplifier is reduction of harmonic output. Harmonics originate in the class C operating mode because of the nonlinear characteristics of transistors. These nonlinear characteristics, particularly the voltage sensitivity of $C_{bc}$, cause subharmonic-frequency generation as well as harmonic-frequency generation. The wide-band gain also increases the possibility of oscillation if any feedback exists. This condition is further intensified by the use of high-gain transistors or by excessive over-all gain.

The amount of harmonic output and transmitted interference permitted is rigidly specified by the FCC. A broad-band, band-pass filter should be added, therefore, after the transmitter.

## Power and Modulation

Because the only useful power in an AM transmitter is sideband power, it is reasonable to use this power as a reference in evaluation of the transmitter. When a single-tone sinusoidal modulating signal is used, the total sideband power $P_{SB}$ in a modulated wave is given by

$$P_{SB} = P_{AV}\left(\frac{m^2}{2 + m^2}\right) \qquad (399)$$

where $P_{AV}$ is the average power and m is the modulation index. This relationship is convenient to use because $P_{AV}$ is easy to measure and

$$P_{SB} = \frac{P_{AV}}{3} \qquad (400)$$

for 100-per-cent modulation.

The performance of an AM transmitter can also be expressed in terms of peak envelope power PEP. The peak envelope power is equal to 2.66 $P_{AV}$ in a 100-per-cent modulated wave. The value of PEP indicates the ultimate peak power-handling capabilities of the transistors being used.

It is unfortunate that carrier power is sometimes used as a reference in evaluation of the performance of AM transmitters, especially transistorized transmitters. Unlike the sideband power $P_{SB}$, the carrier power $P_C$ does not always have a definite relationship to $P_{AV}$ and PEP. When the carrier is used for a reference, "center shift" and

"upward modulation" must be considered. Use of these terms in conjunction with $P_C$ to define transmitter modulation only complicates the definition of per-cent amplitude modulation. For example, Fig. 558 shows an ampli-

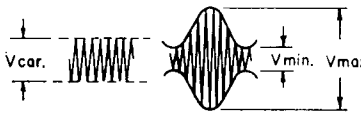



Figure 558. The amplitude modulated wave; $V_{car}$ is the amplitude of carrier before modulation.

tude-modulated wave. The amplitude modulation AM in per cent is defined as follows:

$$AM = \left(\frac{V_{max} - V_{min}}{V_{max} + V_{min}}\right) \times 100 \quad (401)$$

Use of this equation indicates that when $V_{min} = 0$, the wave is 100-per-cent modulated without reference to the carrier. The following expressions are based on carrier amplitude $V_{car}$ or carrier power $P_C$:

$$AM = \left(\frac{V_{max}}{V_{car}} - 1\right) \times 100 \quad (402)$$

$$P_{AV} = P_C\left(1 + \frac{m^2}{2}\right) \quad (403)$$

These expressions contain the tacit assumption that carrier level must not vary from the unmodulated state, which may not be the case. If the modulation is adjusted to 100 per cent by the use of Eq. (401) and $P_{AV}$ is measured, values can easily be computed for $P_{SB}$, PEP, and even $P_C$.

## Design Considerations

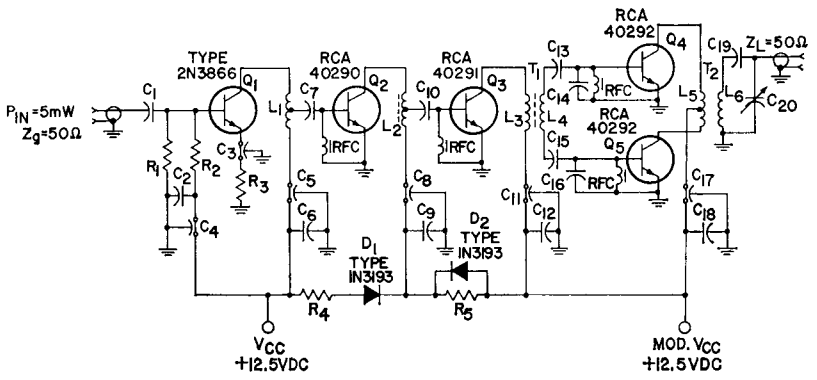
The need for wideband performance in aircraft transmitters precludes the use of sharply tuned circuits to reduce harmonic power in the output; instead, low-pass filters are used. Any configuration of active devices that reduces the harmonic content in the output helps to ease the requirements placed upon these filters. One such configuration is a push-pull amplifier, which inherently has low even harmonics in the output. The higher input impedance of a push-pull stage as compared to a single-ended parallel combination of two transistors is also advantageous for obtaining wider bandwidths because only one-half as much current is injected into the input of push-pull transistors as into parallel devices during one-half cycle.

The coupling circuits in the amplifier of Fig. 559 are basically double-tuned interstage circuits, as shown in Fig. 560. $R_1$ and $C_1$ represent the collector output resistance and the collector output capacitance of the driver transistor. $L_i$ and $R_i$ represent the input series inductance and the input series resistance of a transistor. (For simplicity, coil resistances are omitted.) Q values for the two circuits shown in Fig. 560 are expressed as follows:

$$Q_1 = \frac{R_1}{\omega L_1} \quad (404)$$

$$Q_2 = \frac{\omega (L_2 + L_i)}{R_i} \quad (405)$$

For large bandwidths, it is desirable that $Q_1$ be much larger than $Q_2$. $L_2$, $C_2$, and $L_i$ are series resonant at some frequency $f_o$ within the bandwidth; $L_1$ and $C_1$ can then be determined as follows:




$C_1 = 300$ pF, silver mica, Arco, or equiv.
$C_2 = 0.005$ μF, ceramic
$C_3$ $C_4$ $C_5$ $C_8$ $C_{11}$ $C_{17} = 1000$ pF, feedthrough
$C_6$ $C_9$ $C_{12}$ $C_{18} = 0.5$ μF, ceramic
$C_7 = 50$ pF, silver mica, Arco, or equiv
$C_{10}$ $C_{13}$ $C_{15} = 82$ pF, silver mica, Arco, or equiv.
$C_{14}$ $C_{16}$ $C_{19} = 150$ pF, silver mica, Arco, or equiv.
$C_{20} = 8$ to 60 pF, Arco #404, or equiv.
$R_1 = 470$ ohms, 0.5 W
$R_2 = 1500$ ohms, 0.5 W
$R_3 = 47$ ohms, 0.5 W
$R_4 = 15$ ohms, 0.5 W
$R_5 = 33$ ohms, 0.5 W

$L_1 = 7$ turns of No. 22 wire, 13/64" dia. 9/19" L. tap 1.5 T.
$L_2 = 5.5$ turns of No. 22 wire, 13/64" dia., closely wound, tap 2.0 turns
$L_3 = 6$ turns of No. 22 wire, 13/64" dia , interwind with L4
$L_4 = 4$ turns of No. 22 wire, 13/64" dia., interwind with L3
} Cambion IRN-9 core mat'l or equiv.
$L_5 = 5$ turns of No. 22 wire, 13/64" dia. C.T., interwind with L6
$L_6 = 5$ turns of No. 22 wire, 13/64" dia., interwind with L5
R.F.C. = 1 turn of No. 28 wire on ferrite bead, Ferroxcube #56-590-65/4B or equiv.

Figure 559. A 118-to-136-MHz 40-watt peak envelope power transistor amplifier.

$$L_1 C_1 = \frac{1}{(\omega_0)^2} \qquad (406)$$

In practice, the resonant frequency $f_0$ may not be exactly the center frequency of the passband, but may tend toward the high end of the bandwidth to compensate for degradation of the frequency response of the transistor itself. Normally, there is no problem obtaining relatively high values of $Q_1$ because transistors have large collector output resistance $R_1$. However, it is more difficult to obtain a low value of $Q_2$ in a transistor double-tuned interstage circuit because high-power transistors have low series input resistance $R_i$. The contribution of the induc-

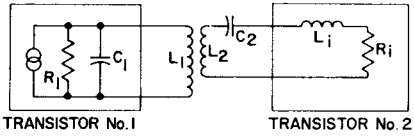



Figure 560. A double-tuned interstage.

tive series input reactance $L_i$ may be sufficient to raise the value of $Q_i$ to undesirable levels and thereby limit the obtainable bandwidth.

This problem can be solved by use of an L-section and its transforming properties. The inductive input impedance of a transistor may be represented by the solid lines of Fig. 561.

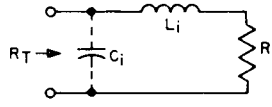



Figure 561. Transistor input as an L-section.

The definite Q value associated with this input impedance may be represented as $Q_i$. If a capacitor $C_i$ is added to the transistor input of Fig. 561, as shown by the dotted line, the resistance $R_i$ can be transformed up by the L-section to a new value $R_T$, as follows:

$$R_T = R_i (Q_i^2 + 1) \qquad (407)$$

The value of the capacitor $C_i$ is calculated as follows:

$$C_i = \frac{1}{\omega R_T}\sqrt{\frac{R_T}{R_i} - 1} = \frac{L_i}{\omega^2 L_i^2 + R_i^2} \qquad (408)$$

When an L-section is used in conjunction with a double-tuned interstage circuit, the value $Q'_2$ of the second circuit is given by

$$Q'_2 = \frac{\omega L_2}{R_T} \qquad (409)$$

This value is, of course, lower than that shown in Eq. (405). Consequently, an L-section can be used to match resistances of not-too-different magnitudes and at the same time maintain low values of Q. The value of $L_i$ in the circuit is given by

$$L_i = \frac{R_i}{\omega}\sqrt{\frac{R_T}{R_i} - 1} \qquad (410)$$

There are limits to the results that can be accomplished with this type of transformation. For some combination of $L_i$ and $R_i$, the required value of $C_i$ may be too large to be practically realizable. In addition, $R_T$ is a frequency-dependent parameter. For very low values of $Q_i$, the capacitor $C_i$ loses its effectiveness because $R_T$ becomes very nearly equal to $R_i$.

Double-tuned interstage coupling circuits are used throughout the amplifier shown in Fig. 559. When it is necessary to use a two-winding transformer, as in the case of $T_1$ and $T_2$, bifiliar windings are employed for tighter coupling. In other cases, autotransformers with their high coefficient of coupling are used quite successfully. Eq. (406) is used as the starting point for determination of the inductances in the primaries of the double-tuned interstages; the collector to base capacitance $C_{CB}$ of the transistor is substituted for $C_1$. Turn ratios are determined by the impedance levels to be transformed. The load resistance $R_L$ for each stage is determined as follows:

$$R_L = \frac{(V_{CC})^2}{2P_o} \qquad (411)$$

where $V_{CC}$ is the collector supply voltage and $P_o$ is the power output. The collector-to-emitter saturation voltage is omitted for simplicity.

A single 40292 transistor is capable of delivering 6 watts of output power with an input of 2 watts and a supply of 12.5 volts dc at 135 MHz. For these conditions, the load resistance $R_L$ is given by

$$R_L = \frac{(12.5)^2}{12} = 13 \text{ ohms}$$

This value of 13 ohms from one-half of the primary winding of $T_2$ is transformed to 50 ohms in the secondary winding. This impedance level allows the use of a 1:1 transformer, which is convenient for bifilar winding. For 40292 transistors, $R_i$ is approximately 6 ohms and $X_{Li}$ is about 3 ohms. An L-section is used in the inputs to the 40292 transistors in the push-pull amplifier. To maintain a low value of $Q_i$, the leads on the base-to-emitter capacitors ($C_{14}$ and $C_{16}$) are kept short, and the capacitors are placed as close to the base and the emitter as possible. The values of $C_{14}$ and $C_{16}$ of Fig. 559 are determined empirically. The effective capacitances may differ appreciably from the nominal value of 150 picofarads shown.

Drive power of about 3 to 3.5 watts is required for the push-pull amplifier. This power is provided by the 40291 driver transistor operating into a 24-ohm load

$$\left[ R_L = \frac{(V_{CE})^2}{2P_o} = (12.5)^2/6.5 \right]$$

Because the input resistance to the driver is sufficiently high (12 ohms), no L-section is used. The load resistance for the 40290 predriver transistor is selected to provide the required input to the driver of about 0.6 watt. The 100-milliwatt input required for the pre-driver stage is supplied by the 2N3866 class A input stage. Again, a double-tuned interstage circuit is used for coupling. The class A amplifier is biased to a quiescent current of 40 milliamperes for maximum gain, and has a load line of approximately 300 ohms, which is computed from

$$R_{\text{load line}} = \frac{V_{CC}}{I_C} \qquad (412)$$

An autotransformer is used to transform the 300-ohm load down to about 12 ohms at the predriver. The input of the 2N3866 stage is matched to the 50-ohm source. This stage has a gain of about 13 dB which increases the power from the 5-milliwatt input. The problem of subharmonic generation is solved by use of cores in the interstage transformers. Stable operation is obtained if the stages are kept 1.25 inches apart.

The final amplifier and the driver are modulated symmetrically about the carrier level. The predriver is modulated more in a positive direction as a result of the resistor-diode arrangement ($R_4$, $R_5$, $D_1$, and $D_2$).

Several precautions should be taken to avoid conditions which may lead to the destruction of transistors. For example, overmodulation should not be allowed to occur because excessive negative excursions of the collector voltage may forward-bias the collector-to-base junction to a destructive point. Also, when a transmitter is keyed off, a steady-state current flow in the order of 2 amperes is suddenly interrupted in the modulation transformer. The resulting transient voltages may easily exceed the transistor breakdown ratings. Use of a zener diode rated at twice the supply voltage in the collector circuit provides protection from this type of transient.

## Performance and Adjustment

The curves of Fig. 562 show typical values of average modu-

lated power $P_{AV}$ at an amplitude modulation of 95 per cent, and carrier power $P_C$, as measured by a bolometer-type power meter. The peak envelope power PEP is computed as follows:

$$PEP = P_{AV} + \frac{(1 + m^2)}{1 + \frac{m^2}{2}} \quad (413)$$

Output-power variation across the aircraft band is about 0.5 dB for both curves shown in Fig. 562. For this performance, the coil $L_1$ was stretched or compressed for maximum power output at 136 MHz and optimum bandwidth, and the trimmer $C_{20}$ was adjusted for the best combination of output flatness and efficiency. Efficiency is somewhat better at

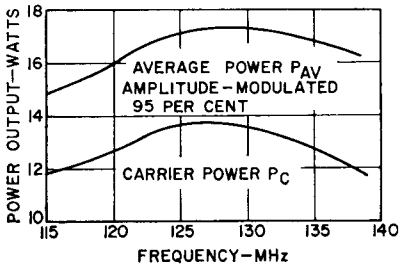



Figure 562. Typical output power as a function of frequency.

higher frequencies than at lower frequencies; harmonic rejection is better at lower frequencies, and may be as good as 20 dB. A spectrum analyzer is required for detection of subharmonics when the slugs in $L_2$ and $T_1$ are adjusted.

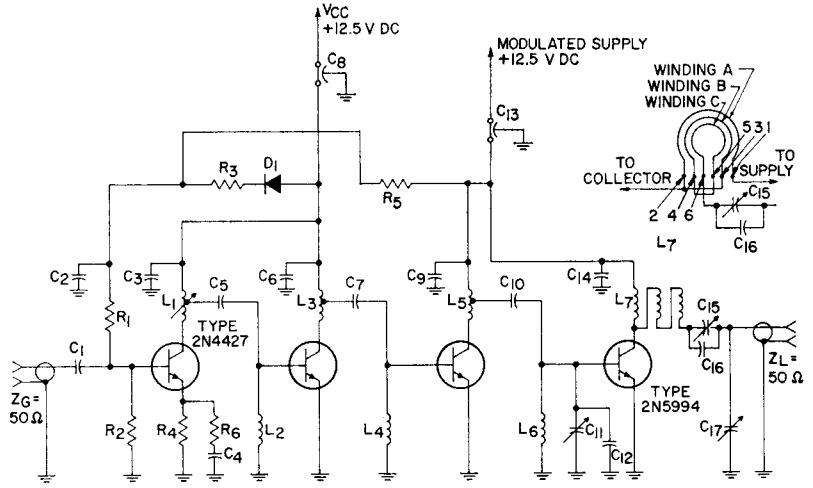
### 15-Watt Amplifier

The schematic diagram of a 15-watt single-ended broadband amplifier is shown in Fig. 563. The output transistor, a 2N5994, is completely tested for load mismatch capability at 118 MHz with a VSWR of infinity through all phases under full modulation. A minimum modulation of 85 per cent can be achieved in the 118-to-136-MHz band. Typical rf performance of this amplifier is shown in Fig. 564.

## COMMUNITY-ANTENNA TELEVISION

Community-antenna television (CATV) systems have experienced rapid growth in the last decade. These systems serve areas in which conventional antennas do not provide adequate television reception. The basic equipment consists of a "head-end" that picks up the signals, and a distribution system that delivers the signals to the subscriber's television receiver. The central antenna is erected at the most advantageous site for best reception; in remote locations, program reception is usually accomplished by means of microwave relays.

The distribution system has two major parts: the main transmission or "trunk" line, and the distribution or "feeder" line. The main trunk line consists of low-loss coaxial cable with main trunk amplifiers spaced along the cable. Bridger amplifiers are used to provide several outputs to the feeder lines from which signals are tapped off to individual subscribers. The backbone of the distribution system is the wide-band amplifier.




$C_1$ = 91 pF, silver mica
$C_2$, $C_3$, $C_8$, $C_{14}$ = 0.01 μF, ceramic
$C_4$ = 200 pF, silver mica
$C_5$ = 50 pF, silver mica
$C_6$ = 0.001 μF, ceramic
$C_7$ = 56 pF, silver mica
$C_9$, $C_{13}$ = 1000 pF, feedthrough
$C_{10}$, $C_{12}$ = 68 pF, silver mica
$C_{11}$, $C_{15}$ = 8-60 pF, Arco 404 or equiv.
$C_{16}$ = 10 pF, silver mica
$C_{17}$ = 3-35 pF, Arco 403 or equiv.
$D_1$ = 1N3193
$R_1$ = 2700 ohms, ½ W
$R_2$ = 470 ohms, ½ W
$R_3$, $R_5$ = 5.1 ohms, ½ W
$R_4$ = 51 ohms
$R_6$ = 20 ohms, ¼ W
$L_1$ = 6 turns No. 18 wire, 0.225-in. dia., 0.4-in. long; tapped at 4½ turns from collector
$L_2$, $L_4$, $L_6$ = RFC 1 turn No. 28 wire, ferrite bead, Ferroxcube No. 56-590-65/4B or equivalent
$L_3$ = 7½ turns No. 18 wire, 0.15-in. dia., 0.55-in. long; tapped at 5½ turns from collector
$L_5$ = 5 turns No. 18 wire, 0.225-in. dia., 0.35-in. long; tapped at 2½ turns from collector
$L_7$ = 3 strands No. 20 enameled wire twisted at 6 turns/in., formed in a loop of ⅜-in. dia., then cross-connected

Figure 563. 15-watt amplitude-modulated amplifier for 118-to-136 MHz operation.

## System Operation

Fig. 565 shows a simplified block diagram of a CATV system in which the TV signals are received directly off the air (no microwave relay). Elaborate arrays of stacked antenna elements in conjunction with narrow-band preamplifiers are used to receive signals in each channel; the signals are then fed into a combining network. The combined multichannel signal is then fed into the main trunk line, which brings the signal from the antenna into the community. The trunk line consists of wide-band amplifiers spaced along a 75-ohm coaxial cable. The gain of each amplifier is adjusted to compensate for cable losses and attenuation characteristics. Typical trunk-line amplifier spacing is in the order of 2500 feet. At various points along the trunk line,

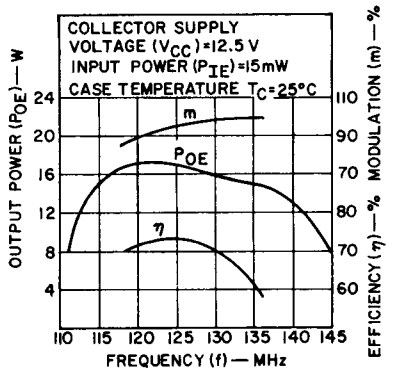



Figure 564. Typical broadband performance of the 118-to-136-MHz amplifier circuit shown in Fig. 563.

signals are supplied to the feeder lines by **bridger amplifiers.** A bridger amplifier provides several outputs to the feeder lines from which signals are tapped off to individual subscribers. One or more line-extender amplifiers may be placed along each feeder line, depending upon its length and the number of subscribers.

## Amplifier Requirements

The first requirement for CATV wide-band amplifiers is large bandwidth. The amplifiers should be able to cover a band of frequencies from 50 MHz to 300 MHz.

The next major consideration for a CATV wide-band amplifier is the required gain. The attenuation characteristic of a coaxial cable is a function of frequency; the cable losses increase logarithmically, as shown in Fig. 566. Typical loss is 0.4 dB per 100 feet at channel 2, and 1 dB per 100 feet at channel 13. The loss between trunk amplifiers is typically

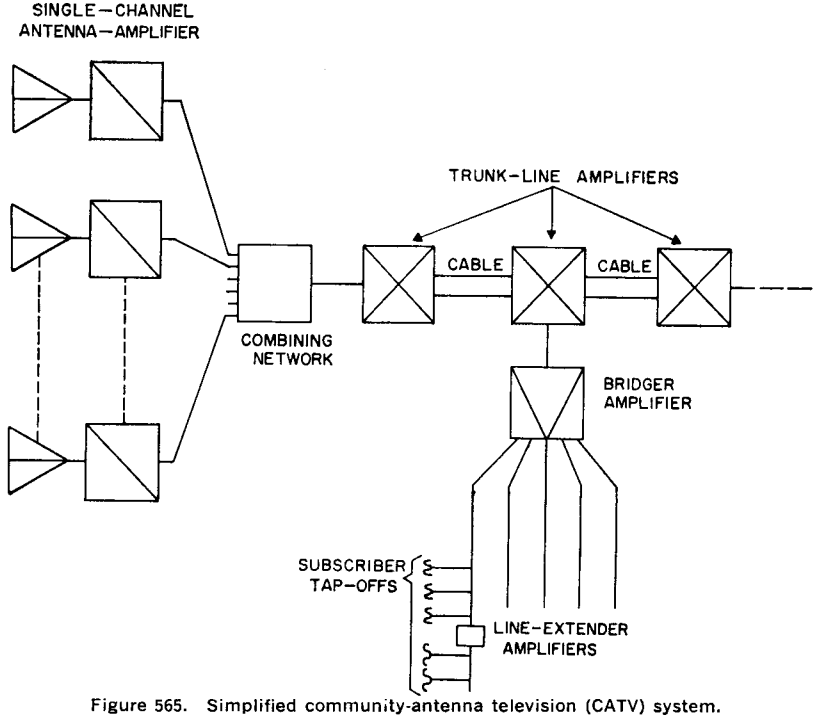



Figure 565. Simplified community-antenna television (CATV) system.

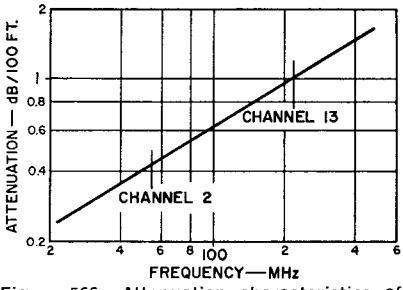



Figure 566. Attenuation characteristics of a coaxial cable as a function of frequency.

25 dB at channel 13. The gain of the trunk amplifier operating at this spacing, therefore, should be 25 dB at 216 MHz. In addition, such an amplifier must be compensated for cable-attenuation differences at each channel by controllable "slope" or "tilt." The amplifier gain must be higher at the high end of the band than at the low end.

The final requirement is for output power or voltage, which is determined by the distortion and signal-to-noise-ratio requirements. If the level of power or voltage is too high, overloading and interference between channels occur; if the level is too low, the signal-to-noise ratio decreases. The most serious distortion is cross-modulation, which produces a "windshield-wiper" effect. Cross-modulation results when several channels are passing through a wideband amplifier. The modulation of undesired interfering signals appears as modulation of the desired signal. The permissible cross-modulation level is 57 dB below the operating output-voltage level in an all-band CATV amplifier, at the end of the cable system.

"Snowy" pictures can be avoided if the signal at any point in a system is maintained at a level high enough to over-ride the noise. This relation is expressed by the signal-to-noise-ratio. The required ratios for various grades of picture quality have been determined as follows: 45 dB for excellent picture (no perceptible snow), 36 dB for fine picture (snow just perceptible), and 29 dB for passable picture (snow definitely perceptible but not objectionable). The signal-to-noise ratio always decreases when a signal passes through an amplifier. The difference between the input signal-to-noise ratio in dB and the output signal-to-noise ratio in dB is defined as the **noise figure** in dB. Noise figure, therefore, is the measure of degradation of signal-to-noise ratio in an amplifier. The noise figure in a CATV cascaded system increases 3 dB each time the length of the system is doubled; the signal-to-noise ratio decreases 3 dB under the same condition.

Typical requirements for trunk-line amplifiers to be used in a CATV cascaded system are as follows:

Frequency Band
  50 MHz to 300 MHz
Input Operating Level
  = 10 dBmV
Output Operating Level
  = 32 dBmV
Maximum Output Capability
  = 50 to 55 dBmV
Gain
  = 22 dBmV
Response
  ±0.5 dB over the band
Noise Figure
  = 12 dB at channel 13
  = 8 dB at channel 2
Tilt
  = 12 dB over the frequency range

These performance specifications must be met in outdoor tempera-

tures ranging from −40 to 140°F. The following paragraphs discuss in more detail the basic considerations for the design of single-stage amplifiers suitable for use in CATV trunk lines and distribution amplifiers.

## Transistor Wide-band Amplifier

The gain-bandwidth product of a transistor connected in a common-emitter configuration is equal to $f_T$. Thus, the bandwidth of an uncompensated common-emitter amplifier stage may be expressed as follows:

$$BW = f_T/h_{fe} = f_T\,(1-\alpha_o)/\alpha_o \quad (414)$$

where $h_{fe}$ is the low-frequency common-emitter current gain of the transistor and $\alpha_o$ is the low-frequency common-base current gain. Eq. (414) dictates the bandwidth of a transistor amplifier stage if the source impedance is large and if the load impedance is small compared to the output impedance of the transistor. In practice, the load and source impedance are such that the bandwidth of the actual amplifier is smaller than the value determined from Eq. (414). Fig. 567

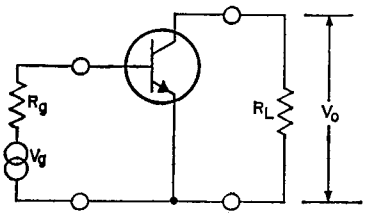



Figure 567. Common-emitter transistor amplifier.

shows a common-emitter transistor amplifier in which $R_L$ is the load resistance and $R_g$ the source resistance. The transistor can be represented by its hybrid-pi equivalent circuit, shown in Fig. 568(a), in which parasitics are not included. One difficulty with the circuit of Fig. 568(a) is the capacitance $C_c$, which prevents the circuit from being unilateral. The effect of $C_c$ may be approximated by connecting a "Miller-effect" capacitance $C_{eq}$ equal in value to $C_c\ (1 + \alpha_o\ R_L/r_e')$ from point b' to ground and omitting the capacitance $C_c$ entirely. The resulting equivalent circuit is shown in Fig. 568(b). Because, in general, $1/\omega_T r_e' >> C_c$ and $\alpha_o \approx 1$, the value for $C_{eq}$ is conveniently approximated by

$$C_{eq} \approx (1/\omega_T r_e')\ (1 + \omega_T\ C_C\ R_L) \quad (415)$$

With the aid of the simplified circuit of Fig. 568(b), the following equations for the gain and bandwidth of the amplifier are derived:

$$\text{Gain} = \frac{V_o}{V_g} = \left[\frac{\alpha_o\ RL}{(R_g+r_b')\ (1-\alpha_o) + r_e'}\right] \left[\frac{1}{1 + \dfrac{p\ (R_g+r_b')\ r_e'\ C_{eq}}{(1-\alpha_o)\ (R_g+r_b') + r_e'}}\right] \quad (416)$$

$$BW = \frac{(1-\alpha_o)\ (R_g+r_b') + r_e'}{(R_g+r_b')\ r_e'\ C_{eq}} = \frac{W_t}{(1 + W_t\ C_c\ R_L)} \left[1 - \alpha_o + \frac{r_e'}{(R_g+r_b')}\right] \quad (417)$$

Eq. (417) shows that the bandwidth is decreased by an increase

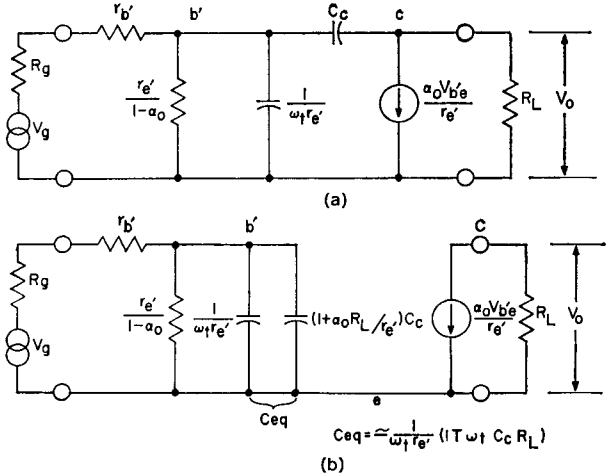



Figure 568. Equivalent circuits for common-emitter amplifier shown in Fig. 567: (a) without parasitic elements; (b) simplified equivalent circuit.

in the load resistance $R_L$ and increased by a reduction in the source resistance $R_g$. For a given $R_g$ and $R_L$, the bandwidth is also increased by an increase in $\omega_T$ and by a decrease in $r_b'$ and $C_c$. Thus, a transistor suitable for wide-band operation should have high $f_T$ (or $\omega_T$), a low collector capacitance $C_c$, and a low base resistance $r_b'$. If a transistor which has an $f_T$ of 1.5 GHz and a $C_c$ of 1.5 picofarads is used, the bandwidth calculated from Eq. (416) is 8 MHz for $R_g = 75$ ohms, $R_L = 300$ ohms, and $I_C = 50$ milliamperes. The corresponding voltage gain is 140. To obtain the bandwidth required in CATV, it is necessary to use compensation techniques that permit the trade of gain for increased bandwidth.

Transistor amplifiers cannot be designed to permit a gain-for-bandwidth trade in a 1:1 ratio. The voltage gain of a common-emitter amplifier stage, as can be determined from Eqs. (416) and (417), is not inversely proportional to the bandwidth. One of the important criteria of a wide-band transistor amplifier, therefore, is its ability to trade gain for bandwidth. Another way of stating this criterion is that degradation in gain-bandwidth should be small.

**Collector-to-Base Shunt Feedback**—One common method for trading gain for bandwidth in a common-emitter amplifier is by use of shunt resistance-inductance feedback, as shown in Fig. 569.

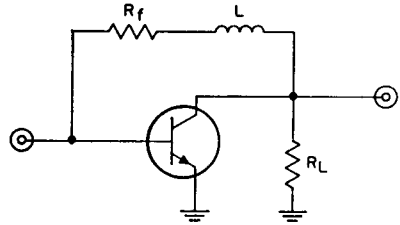



Figure 569. Common-emitter amplifier that uses shunt RL feedback to increase circuit bandwidth.

Simple feedback is provided from collector to base for trading gain and bandwidth. The current gain

at low frequencies is approximately equal to the ratio of $R_f$ to $R_L$. The feedback resistance $R_f$ should be in the range $R_L < R_f < R_L/(1-\alpha_o)$. Without the inductance L, this technique is not an efficient way to obtain wide bandwidths because the feedback resistance becomes so low that it loads both the input and output circuits and thus reduces the gain-bandwidth product. The effective input impedance is the input impedance of the transistor in parallel with a resistance equal to the feedback resistance $R_f$ divided by (1+K), where K is the voltage gain. The load presented to the collector of the transistor consists of the feedback resistance $R_f$ in parallel with resistance $R_L$. However, if an inductance L is connected in series with the feedback resistance $R_f$, the gain-bandwidth product can be restored to its value without feedback. The inductance tends to remove the feedback resistance from the circuit at frequencies above $(1-\alpha_o)\, f_T$, and thus eliminates its effect on the high-frequency current amplification. The approximate expression for determining the value of the inductance is as follows:

$$L = (R_f + r_b' + R_L)\, R_f/2\pi\, f_T\, R_L. \quad (418)$$

**Emitter Degeneration**—Another common method of trading gain for bandwidth is to use emitter regeneration or emitter peaking, as shown in Fig. 570. Simple resistance-capacitance feedback is provided from the emitter to ground for trading gain for bandwidth. The effect of the resistance $R_e$ is to reduce the gain at low frequencies. The equations for the voltage gain and the bandwidth

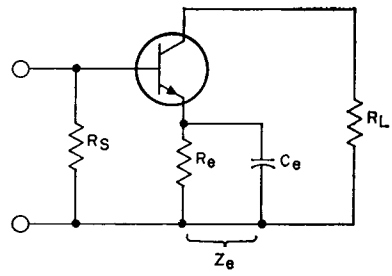



Figure 570. Transistor amplifier that uses emitter degeneration to increase circuit bandwidth.

of this amplifier, when $Z_e = R_e$, are as follows:

$$\text{Gain} = \frac{V_o}{V_g} = \left[\frac{\alpha_o R_L}{\alpha_o R_e + r_e' + (1-\alpha_o)(R_s + r_b' + R_e)}\right] \left[\frac{1}{1 + \dfrac{pC_{eq}\, r_e'\,(R_s + r_b' + R_e)}{\alpha_o R_e + r_e' + (1-\alpha_o)(R_s + r_b' + R_e)}}\right] \quad (419)$$

$$\text{BW} = \frac{\alpha_o R_e + r_e' + (1-\alpha_o)(R_s + r_b' + R_e)}{r_e'\,(R_s + r_b' + R_e)\, C_{eq}} \quad (420)$$

Comparison of Eqs. (417) and (420) shows that increased bandwidth is possible if $R_s + r_b' > R_e$. The effect of the capacitance $C_e$ in shunt with the emitter resistance $R_e$ is to decrease the degeneration at high frequencies. The required value of capacitance $C_e$ is approximately equal to $1/(15 f_T R_e)$.

Fig. 571 shows a typical amplifier that might be used as a single stage in a CATV line amplifier. Generally, four or five such stages are cascaded to provide the gain of 20 to 30 dB required for a trunk-line or extender amplifier. This amplifier uses both shunt and series feedback to achieve the simultaneous requirements of good input and output

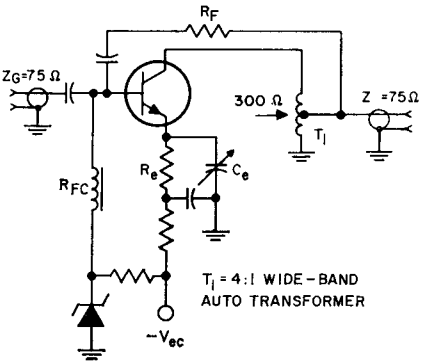



Figure 571. Typical CATV single-stage amplifier.

match, broadband performance, satisfactory gain, and good linearity.

To realize optimum linearity, most transistors require a load impedance other than the cable-line impedance of 75 ohms. A typical optimum impedance, as shown in Fig. 571, is about 300 ohms. To achieve this impedance at the transistor collector terminals, a transformer such as that shown in Fig. 572 is often used.

This transformer consists of a ferrite toroid around which a twisted pair of wires are wound. It is a transmission-line type and has excellent bandwidth. The transmission lines take the form of twisted pairs of wires. The coils are arranged so that the interwinding capacitance is a component of the characteristic impedance of the line, and forms no resonances which seriously limit the bandwidth, as in the case of a conventional transformer. For this reason, the windings can be spaced closely together to assure good coupling. Transformers of this type can provide good high-frequency response (this response is determined by the length of the windings).

The low-frequency response, on the other hand, is determined by the permeability of the core. The greater the core permeability, the fewer the turns required for a given low-frequency response and

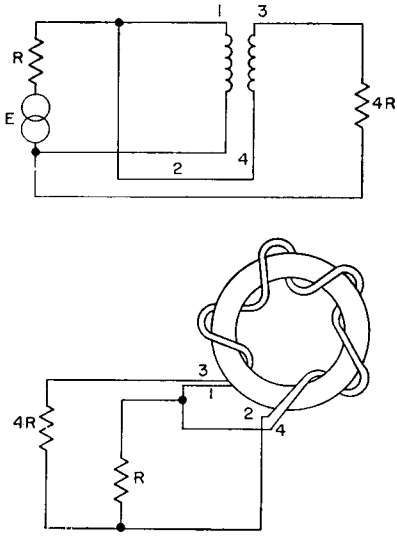



Figure 572. Wideband transformer. This transformer may be used to provide a 4:1 impedance ratio, as indicated in top diagram. The transformer is basically a twisted-pair transmission line wound about a ferrite torroid, as shown in lower diagram.

the larger the bandwidths. Thus, a good core material is desirable. Ferrite toroids have been found very satsfactory. The permeability of some ferrites is very high at low frequencies and decreases at higher frequencies. Large reactance, therefore, can be obtained with few turns at low frequencies. When the permeability decreases, the reactance is maintained by the increase in frequency, and good response is obtained over a large frequency range. It is im-

portant that coupling be high at all frequencies, or the transformer action fails.

The transformer shown in Fig. 572 has an impedance ratio of 4:1. The high-frequency response of this transformer may be calculated by use of the following equation:

$$\frac{\text{Power Available}}{\text{Power Output}} = \frac{(1+3\cos \mathrm{B}\, l)^2 + 4\sin^2 \mathrm{B}\, l}{4\,(1+\cos \mathrm{B}\, l)^2} \quad (421)$$

where B is the phase constant of the line and l is the length of the line. The response is down 1 dB when the line length is $\lambda/4$; the response is zero at $\lambda/2$. For wideband response, therefore, this transformer must be made small.

## Transistor Considerations

In selection of a transistor for CATV amplifier applications, the following performance criteria must be considered: maximum cross-modulation at a given output level, maximum IMD for a given level, noise figure, dc operating conditions, and dissipation. Obviously, optimization of all of these characteristics cannot be achieved simultaneously, so tradeoffs must be made, taking into consideration the stages in which the transistor is to be used. For instance, low noise figure and moderate output levels are associated with the input stage of a line amplifier, while extremely low distortion is required for the output stage.

One of the more serious types of nonlinearities associated with a CATV amplifier is cross-modulation.

Cross-modulation is the transfer of modulation for one AM signal to another within an amplifier. Because CATV line amplifiers process many signals simultaneously, this exchange of signal information is highly undesirable. Cross-modulation generally sets the upper limit on the output signal level at which an amplifier may operate.

Cross-modulation is caused by odd- (predominantly 3rd-) order nonlinearities in the amplifier's transfer characteristic. All the amplifier nonlinearities are, of course, attributable to the transistor. Nonlinearities within the transistor may be separated into three major areas: (1) emitter-base diode nonlinearities, (2) current gain ($h_{fe}$/Ic) nonlinearities, and (3) collector-base depletion-layer capacitance variations.

Through proper selection of bias point and amplifier circuit (primarily collector load line), a device can be optimized for minimum cross-modulation.

Because cross-modulation is a transfer of modulation from one channel to another, it can be measured by determining the degree of modulation produced on an unmodulated carrier by various combinations of interfering signals. The basic test for cross-modulation is shown in Fig. 573. A number of clean TV (modulated) signals at the various channel frequencies are combined and fed through the amplifier under test. The output signal is viewed on a good television receiver, and the output levels are increased until "windshield-wiper" (cross-modulation) effects are just visible in the picture. The level at which this condition occurs is called the maximum usable output of the amplifier.

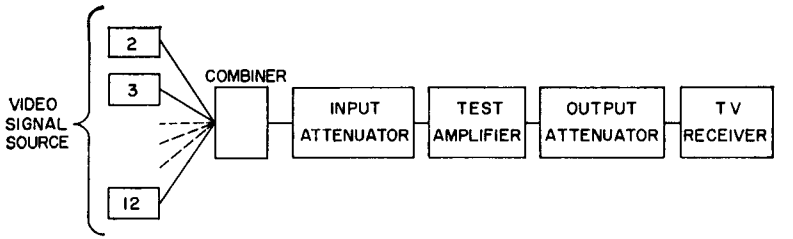



Figure 573. Block diagram of basic test setup used for cross modulation measurements.

This test is not really conclusive, however, because TV "windshield-wiper" effects can be seen much more readily on some pictures than on others. The accuracy of the test is greatly increased if an unmodulated signal is substituted for the picture signal on the viewing channel. This technique provides a white screen which does not change during the test, and allows more consistent and critical observations.

While the "white-screen" test is simple, the resulting output ratings depend somewhat on the judgment of the person carrying out the test. For more accurate results a method that gives repeatable readings is needed.

Fig. 574 shows a block diagram of a standard test set-up. Any combination of the twelve crystal-controlled carriers is available. Each one is 100-per-cent modulated with a 15-kHz square wave, except the one to which the receiver is tuned; this carrier is unmodulated, but its amplitude is the same as the peak amplitudes of the modulated carriers.

100-per-cent modulation is applied momentarily to the modulated channel and a 100-per-cent modulation reference level is noted on the 15-kHz VTVM. Modulation is then removed from the test channel. The cross-modulation is then read on the 15-kHz VTVM as a fraction of the 100-per-cent mod-

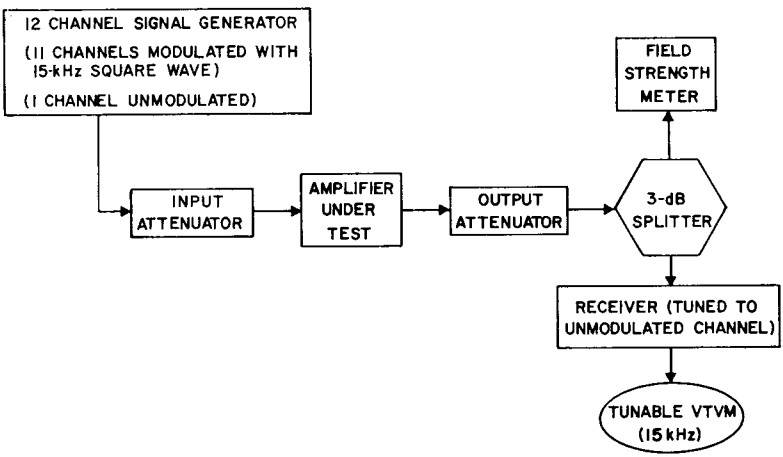



Figure 574. Block diagram of a 12-channel modulation test set.

ulated value. The receiver is tuned to each of these channels successively and the cross-modulation readings are recorded. The amplifier rating is based on that channel which shows the greatest cross-modulation. When an amplifier is being tested, its "behavior" is determined by changing the settings of the two attenuators to alter the output levels in, for example, 2-dB steps to find if the cross-modulation follows the "two-for-one" law expected of a "well-behaved" amplifier. If the amplifier does not change "two-for-one," there is a likelihood that some form of cancellation of nonlinearities is taking place.

## The 2N5109 Overlay Transistor

The RCA-2N5109 transistor, packaged in a TO-39 case, is designed to provide large dynamic range, low distortion, and low noise, and is well suited for use in a wideband amplifier in CATV applications. The 2N5109 is an epitaxial silicon overlay transistor that features low $r_b'$ and $C_C$ and high and relatively flat $f_T$ with current level. Fig. 575 shows the $f_T$ of a typical 2N5109 as a function of collector current at a $V_{CE}$ of 15 volts. The $f_T$ measured at a collector current of 50 milliamperes is 1.5 GHz; $f_T$ is within 20 per cent of its maximum value from 25 to 100 milliamperes. Fig. 576 shows the $f_T$ of a typical 2N5109 as a function of $V_{CE}$ at collector currents of 25, 50, and 100 milliamperes, respectively. The electrical characteristics of this transistor are summarized in Table XLI.

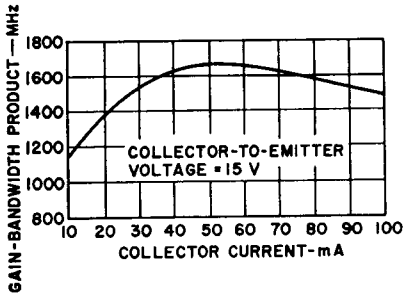



Figure 575. Gain-bandwidth product as a function of collector current for a typical RCA-2N5109 transistor.

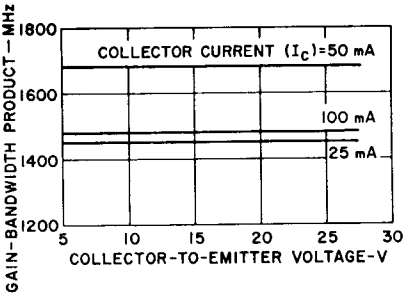



Figure 576. Gain-bandwidth product as a function of collector voltage for a typical RCA-2N5109 transistor.

Fig. 577 shows the noise figure of a typical 2N5109 as a function of collector current. The noise figure is measured with the 2N5109 operating as a narrow-band 200-MHz amplifier at a $V_{CE}$ of 15 volts. The best noise figure occurs at a collector current of less than 10 milliamperes.

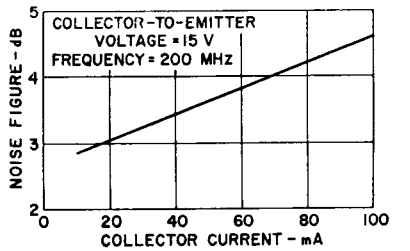



Figure 577. Noise figure as a function of collector current for a typical RCA-2N5109 transistor.

Table XLI—Electrical Characteristics of RCA-2N5109 Overlay Transistor (Case Temperature = 25°C)

| CHARACTERISTIC | SYMBOL | TEST CONDITIONS DC COLLECTOR VOLTS $V_{CB}$ | $V_{CE}$ | DC CURRENT (mA) $I_B$ | $I_C$ | LIMITS Min. | Max. | UNITS |
|---|---|---|---|---|---|---|---|---|
| Collector-Cutoff Current | $I_{CEO}$ | | 15 | | | | 20 | μA |
| Collector-to-Base Breakdown Voltage | $BV_{CBO}$ | | | | 0.1 | 40 | | V |
| Collector-to-Emitter Voltage (Sustaining) | $V_{CER}(sus)$* | | | | 5 | 40 | | |
| | $V_{CEO}(sus)$ | | | | 5 | 20 | | V |
| Emitter-to-Base Breakdown Voltage | $V_{(BR)EBO}$ | | | 0.1 | 0 | 3 | | V |
| Collector-to-Emitter Saturation Voltage | $V_{CE}(sat)$ | | | 10 | 100 | | 0.5 | V |
| Collector-to-Base Capacitance (Measured at 1 MHz) | $C_{ob}$ | 15 | | | | | 3.5 | pF |
| Small-Signal Common-Emitter Forward-Current Transfer Ratio (Measured at 200 MHz) | $h_{fe}$ | | 15 | | 20 | 4.8 | | |
| | | | 15 | | 50 | 6.0 | | |
| | | | 15 | | 100 | 4.8 | | |
| Voltage Gain (Wideband, 50 to 216 MHz) | V.G. | | 15 | | 50 | 11 | | dB |
| Cross Modulation at 54-dBmV Output | C.M. | | 15 | | 50 | | —57 (typ) | dB |
| Power Gain (Narrow-band, Measured at 200 MHz; $P_{in} = -10$ dBmV) | P.G. | | 15 | | 10 | 11 | | dB |
| Noise Figure (Measured at 200 MHz) | N.F. | | 15 | | 10 | | 3 (typ) | dB |

* With external base-to-emitter resistance $R_{BE} = 10$ ohms.

## Choice of Operating Conditions for 2N5109

The most important parameter in the input stage of a CATV system is the noise figure. Distortion is not usually important in the input stage because the voltage and current swings of the transistor are small. The dc bias of the transistor should be chosen for mimimum noise figure. An RCA-2N5109 used in the first stage should be biased at a collector current $I_C$ of 10 milliamperes and a collector-to-emitter voltage $V_{CE}$ of 10 to 15 volts. The noise figure of a typical 2N5109 measured in a CATV amplifier is 8 dB at channel 13.

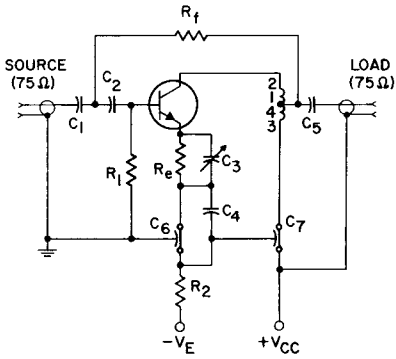
The final stage, on the other hand, should be biased so that maximum power output can be obtained with minimum cross-modulation distortion. In addition, the bias condition should be within the dissipation capability of the transistor. For example, if it is assumed that 1 volt rms (60 dBmV) is required across a load of 75 ohms, the peak-to-peak voltage swing across the 75-ohm load is 2.83 volts, and the corresponding current swing is 36.4 milliamperes. The 75-ohm load must be transformed into the collector load with the use of a wideband transformer. If a 1:1 impedance transformer is used, the collector voltage and current swings are the same as those of the 75-ohm load. The collector bias current $I_C$ can then be selected from Fig. 575 for minimum change in $f_T$

within the 36.4-milliampere current swing. The value of $I_C$ that satisfies this condition is approximately 60 milliamperes. The $V_{CE}$ corresponding to a collector load of 75 ohms is therefore 4.5 volts. Figs. 575 and 576 show that large current swings (rather than voltage swings) result in a change in $f_T$ and, therefore, in large distortion.

If a 4:1 impedance transformer is used, the collector load becomes 300 ohms and the collector voltage and current swings become 4.66 volts and 18.2 milliamperes, respectively. From Fig. 575, the value of $I_C$ can be chosen as 55 milliamperes for a minimum change of $f_T$ within this current swing. The $V_{CE}$ value corresponding to the collector load of 300 ohms is 16.5 volts. Fig. 576 shows that the $f_T$ is substantially constant within the 4.66-volt swing around 16.5 volts. The power dissipation is 0.9 watt, which is within the limit of the 2N5109. The power output for a typical 2N5109 operated at 16.5 volts and 55 milliamperes in a 12-channel system is 52 dBm with −57 dB cross modulation.

## A Wideband Amplifier Using the 2N5109

A typical single-stage wideband amplifier circuit is shown in Fig. 578. This common-emitter class A amplifier uses the RCA-2N5109 transistor and is designed for 75-ohm source and load resistances; it is suitable for the CATV application. A ferrite-toroid wideband transformer that has an impedance ratio of 4:1 is used in the output to transform a 75-ohm load into a 300-ohm collector load. Both shunt feedback and emitter degeneration are employed through $R_f$ and $R_e$, respectively. Emitter peaking is accomplished by the use of $C_3$. Two dc power supplies are used; one supply provides the collector reverse bias, and the other supply provides the emitter-to-base forward bias through resistances $R_1$ and $R_2$.




$C_1$, $C_2$, $C_5$ = 0.002 μF
$C_3$ = 8-60 pF, Arco 404 or equiv.
$C_4$ = 0.03 μF
$C_6$, $C_7$ = 1500 pF
$R_1$ = 390 ohms, ½ watt
$R_2$ = 330 ohms, 1 watt
$R_e$ = 6.8 ohms, ½ watt
$R_f$ = 200 ohms, ½ watt
T = 4-turn bifilar winding, 3/16" ID, No. 30 wire; Core: G.I. material $Q_1$ or equiv.

Figure 578. Single-stage wideband amplifier using the RCA-2N5109 transistor to provide a gain of 12 dB in the frequency band from 54 to 216 MHz.

The amplifier of Fig. 578 uses a 2N5109 operated at an $I_C$ of 55 milliamperes and a $V_{CE}$ of 16.5 volts, and can provide a minimum gain of 12 dB within the band of 54 to 216 MHz.

## Low-Noise Amplifiers

When a signal is processed by a system, a certain amount of extraneous noise is added, which degrades the original input signal-to-noise ratio.

The output signal-to-noise ratio of an amplifier with a noise figure F is 1/F times less than

the input signal-to-noise ratio. The following equation expresses this relationship:

$$\frac{P_{SO}}{P_{NO}} = \frac{1}{F}\left(\frac{P_{Si}}{P_{Ni}}\right) \quad (422)$$

Eq. (422) is valid only when the input noise ($P_{NI}$) associated with the signal is equal to $kT_oB$; here k is Boltzmann's constant, $T_o = 290°$ K, and B is the system bandwidth. If the equivalent noise temperature is not 290°K, the effect of the 1/F factor may be increased or decreased, as shown by the following more general relationship:

$$\frac{P_{SO}}{P_{NO}} = \frac{GP_{Si}}{(GP_{Ni}) + (F-1)Gk\,T_oB} \quad (423)$$

where G is the system gain.

From Eq. (423) it can be seen that the value of $P_{SO}/P_{NO}$ depends upon the level of input noise, $P_{NI}$. For high levels of $P_{NI}$ and low values of F there is little degradation in signal-to-noise ratio of the incoming signal as it is amplified:

$$\frac{P_{SO}}{P_{NO}} \approx \frac{P_{Si}}{P_{Ni}}$$

If several amplifiers are cascaded, the total noise figure, $F_{total}$, is related to the noise figures and gains of the individual stages by the following equation:

$$F_{TOTAL} = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1G_2} + \ldots + \frac{F_N - 1}{G_1G_2 \ldots G_{N-1}} \quad (424)$$

In Eq. (424), $G_1$ and $F_1$ are the gain and noise figure of the first stage of amplification. If $G_1$ is large, the noise figure of the entire amplifier chain is nearly equal to $F_1$.

**Transistor Noise Figure**—The noise figure of a single-stage transistor amplifier is a function of frequency and transistor parameters, as shown by the following equation:

$$NF = 1 + \frac{r_b'}{r_g} + \frac{r_e'}{2\,R_g} + \frac{(r_b' + r_e' + R_g)^2}{2\,\alpha_o\,R_g\,r_e'}\left[\frac{I_{co}}{I_E} + \frac{1-\alpha_o}{\alpha_o} + \left(\frac{f}{f_T}\right)\right] \quad (425)$$

At frequencies below approximately 0.1 $f_T$, the noise figure is constant with frequency and is primarily determined by $R_g$, $r_b'$, $r_e$, and $\alpha_o$. The resistance $r_e'$ is inversely related to the dc emitter current ($r_e' = 26/I_e$); therefore, there is a value of $I_e$ that corresponds to a minimum noise figure. At higher frequencies, the $(f/f_T)^2$ term in the noise-figure equation becomes predominant, with the result that the noise figure asymptotically approaches a 6-dB-per-octave slope. From the viewpoint of noise considerations, the $r_b'$ and $I_{co}$ of the transistor should be low, and $f_T$ should be high. Eq. (425) shows that the noise figure is also a function of the source resistance $R_g$ and, therefore, can be minimized by proper selection of $R_g$. The optimum source resistance can be determined if Eq. (425) is differentiated and the result is set equal to zero and solved for $R_g$. The following equation for $R_g$ is then obtained:

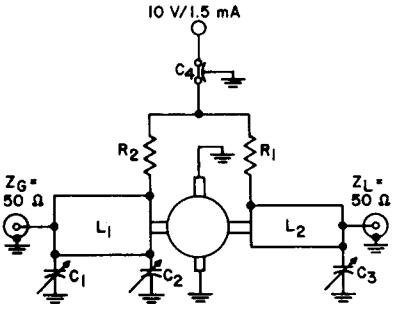
$$R_g\,(\text{optimum}) = \left[(r_e' + r_b')^2 + \frac{\alpha_o\,r_e'\,(2\,r_b' + r_e')}{\frac{1-\alpha_o}{\alpha_o} + \left(\frac{f}{f_T}\right)^2 + \frac{I_{co}}{I_E}}\right]^{\frac{1}{2}} \quad (426)$$

At low frequencies, where $(f/f_T)^2$ is small, a transistor that has high dc current gain requires a high source resistance $R_g$ for best noise performance. As the frequency approaches $f_T$, the second term of Eq. (426) becomes small, and the optimum source resistance approaches $(r_b' + r_e')$.

The circuit shown in Fig. 579 is a narrow-band 1-GHz amplifier with a gain of 10 dB and a noise figure of 3.0 dB. Microstrip construction with variable tuning allows this circuit to be used for noise-figure testing or as an amplifier in a 50-ohm system.

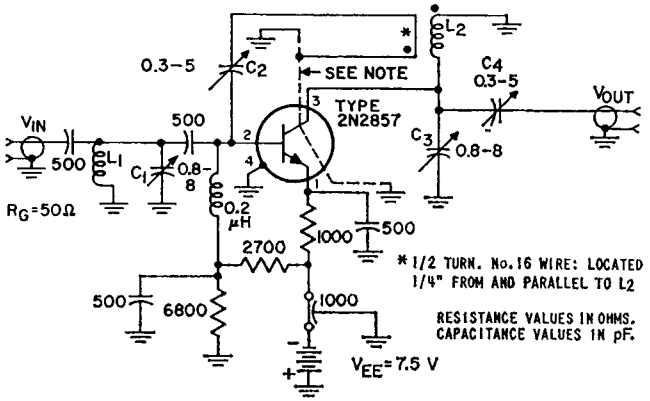
Fig. 580 shows a circuit diagram of a typical low-noise 450-MHz amplifier that uses an RCA-2N2857 transistor. This amplifier provides a gain of 12.5 dB with a noise figure of 4.5 dB.

Fig. 581 shows a typical test arrangement for measurement of transistor noise figure.




$C_1$, $C_2$, $C_3$ = 1 to 10 pF, Johanson No. 2954 or equiv.
$C_4$ = 1000 pF, feedthrough, Allen-Bradley type FAC5 or equiv.
$R_1$ = 2.7 ohms, 1/8 W, carbon
$R_2$ = 120K, 1/8 W, carbon
$L_1$ = microstrip transmission line, length = $\lambda/4$, $Z_o$ = 30 ohms
$L_2$ = microstrip transmission line, length = $\lambda/4$, $Z_o$ = 90 ohms

Figure 579. Low-noise 1-GHz amplifier.




$L_1$, $L_2$ = silver-plated brass rod, 1½ inch long-by-¼ inch diameter; install at least ½ inch from nearest vertical chassis surface

**Note:** External interlead shield to isolate the collector lead from emitter and base leads.

Figure 580. Neutralized amplifier circuit used to measure 450-MHz power gain and noise figure for type 2N2857.

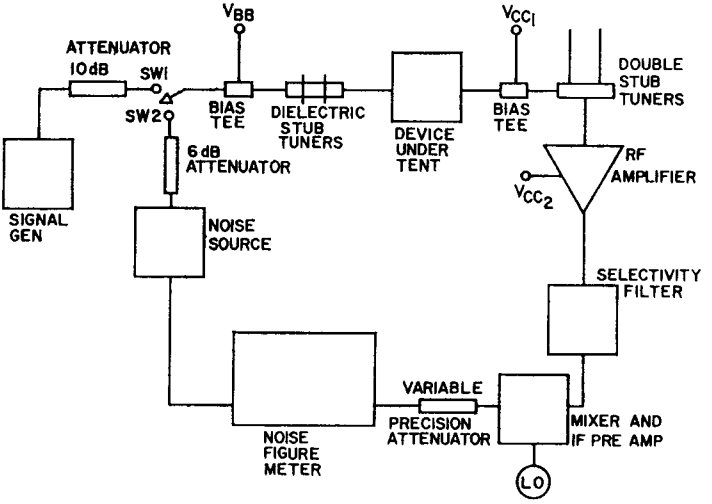



Figure 581. Typical noise-figure test set.

# Microwave Power Amplifiers and Oscillators

THIS section describes the use and the capabilities of power transistors as microwave amplifiers and oscillators, and also discusses transistor amplifier-multipliers and oscillator-multipliers with microwave output frequencies. Microwave relay links and air-traffic-control systems are then described as typical system applications of microwave power transistors.

## TRANSISTOR CONSIDERATIONS

The considerations of construction, geometry, and ratings that were discussed previously for rf power transistors are equally important at microwave frequencies. The output power, power gain, efficiency, and bandwidth are all strongly affected by package parasitic elements, which become increasingly significant at higher frequencies. Therefore, the package design is an especially important feature of microwave transistors.

### Output Power

The power-output capability of a transistor is determined by current- and voltage-handling capabilities of the device at the frequency range of interest. The current-handling capability of the transistor is limited by its emitter periphery and epitaxial-layer resistivity. The voltage-handling capability of the device is limited by the breakdown voltages, which are, in turn, limited by the resistivity of the epitaxial layer and by the penetration of the junction.

The breakdown voltage at microwave frequencies is substantially higher than the dc or static value, as indicated by the following equation:

$$V_{CEO} = V_{CBO} \Big/ \left(\frac{f_T}{f} + 1\right)^{\frac{1}{n}} \quad (427)$$

where n is the avalanche breakdown factor. Eq. (427) shows that the breakdown characteristic increases from the $V_{CEO}$ value under dc conditions to a value approaching $V_{CBO}$ at a frequency f equal to or greater than $f_T$.

Another parameter that limits the power-handling capability of the transistor is the saturation voltage. The saturation voltage at microwave frequencies, $V_{CE(sat)}$,

is significantly greater than the dc value because the active area is smaller than at dc.

In general, the operating voltage restrictions are the same for all microwave power transistors; therefore, only current-handling capability differentiates high-power transistors from lower-power units.

At high current levels the emitter current of a transistor is concentrated at the emitter-base edge; therefore, transistor current-handling capability can be increased by the use of emitter geometries which have high emitter-periphery-to-emitter-area ratios and by the use of improved techniques in the growth of collector substrate material. Transistors for large-signal applications are designed so that the peak currents do not cause base widening which would limit the current-handling capability of the device. Base-width widening is severe in transistors in which the collector side of the collector-base junction has a lower carrier concentration and higher resistivity than the base side of the junction. However, the need for low-resistivity material in the collector to handle high currents without base widening severely limits the breakdown voltages, as discussed previously. As a result, the use of a different-resistivity epitaxial layer for different operating voltages is becoming common.

## Power Gain

The power gain of a microwave transistor power amplifier is determined by the dynamic $f_T$, the dynamic input impedance, and the collector load impedance, which depends on the required power output and the collector voltage swing. The power gain, P.G., of a transistor power amplifier can be expressed as follows:

$$\text{P.G.} = \frac{(f_T/f)^2 R_L}{\text{Re}(Z_{in})} \qquad (428)$$

where $f_T$ is the dynamic gain-bandwidth product, f is the frequency of operation, $R_L$ is the real part of the collector parallel equivalent load impedance determined by the required power output, and $\text{Re}(Z_{in})$ is the real part of the dynamic input impedance when the collector is loaded with $Z_L$.

Eq. (428) shows that for high-gain operation of large-signal or power transistors, the device should have high current gain at the frequency of operation under large current swing conditions. This performance is achieved with shallow diffusion techniques.

$R_L$ is defined approximately by Eq. (429) for class B or C operation:

$$R_L \cong K \frac{[V_{CC} - V_{CE}\,(\text{sat})]^2}{2P_O} \qquad (429)$$

where K is unity or less, depending on the class of operation. The real part of the dynamic input impedance, $\text{Re}(Z_{in})$, varies considerably with signal level, and varies inversely with the power output of the device. The package parasitic inductance also are important in determining the value of $\text{Re}(Z_{in})$.

## Efficiency

The collector efficiency of a transistor amplifier is defined as the ratio of signal power output at the frequency of interest to the

dc input power. It can be calculated as:

$$\eta_c = \eta_v \eta_i \eta_{ckt} \qquad (430)$$

where $\eta_v$ is the efficiency of conversion of dc collector voltage to microwave-frequency collector voltage (determined primarily by the ratio of $V_{CC}$ to $V_{CE}$ and the class of operation), $\eta_i$ is the efficiency of conversion of dc collector current to microwave collector current (determined primarily by the class of operation and the transit time in the collector depletion region for microwave transistors), and $\eta_{ckt}$ is the circuit efficiency, which is determined by the loaded and unloaded Q's of the collector circuit.

### Bandwidth

The bandwidth of a transistor power amplifier is determined by the intrinsic frequency capability of the transistor (directly related to $f_T$), package parasitic elements, and the input and output matching circuits.

### Package Design

A suitable high-power transistor package for microwave applications must have low common-lead inductance and low shunt and feedthrough capacitance, as well as good thermal properties.

Packages such as the TO-39 and TO-60 are useful in applications above 1 GHz, but superior performance can be obtained with stripline and coaxial packages specifically designed for microwave frequencies. Fig. 479 in the preceding section showed a stripline package (HF-28) and two coaxial packages (HF-11 and HF-21) that are used at frequencies well into S-band. Best microwave-frequency amplifier performance is provided by common-base configurations in these packages. In the stripline package, the base is connected directly to the flange, and in the coaxial package, the base is connected to the flange separating the ceramics. In both types of packages the base can be rf-grounded with very low parasitic lead inductance; this grounded-base configuration minimizes out-put-to-input feedback and, therefore, facilitates stable operation.

## CIRCUIT DESIGN TECHNIQUES

In designing transistor microwave power circuits, several fundamental considerations must be taken into account: the type of circuit to be used, e.g. microstrip, coaxial, lumped element, or a combination thereof; the type of package to be used; the required size and type of heat sink; and the power output, gain, efficiency, and bandwidth required. All the above items are interrelated; for example, various heat sinks can be used with the coaxial-package devices, depending upon the circuit technique chosen.

Fig. 582 shows a coaxial package jig that uses a standard beryllium oxide ring to conduct heat from the center conductor to the outside conductor of an air-dielectric line section. This type of arrangement is useful for

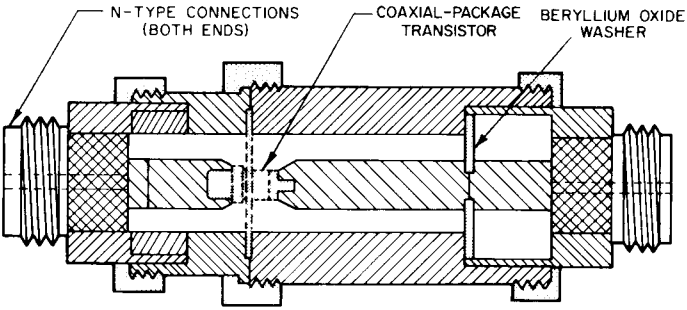



Figure 582. Heat sink for use with coaxial transistor package.

power dissipation of 5 watts or less. A more efficient heat sink is obtained by use of a boron nitride cylinder that makes intimate contact between the coaxial line conductors over the entire length of the cavity. This arrangement results in much improved heat conduction and, therefore, is more suitable for high-power microwave transistors. In addition, the boron nitride, which has electrical and thermal properties similar to those of beryllium oxide, is readily machineable and is nontoxic. Coaxial line lengths are also substantially reduced.

Fig. 583 shows a circuit mounting arrangement for coaxial-package transistors in microstripline and lumped-element circuits. The transistor is mounted vertically through a hole in the metal block which serves as both a heat sink and ground for the device. The bottom side of the metal block is counter-bored so that the base flange of the transistor is level with the surface of the block. The hole through the metal block has a somewhat larger diameter than that of the ceramic portion of the transistor which separates the base flange and the collector stud. This larger diameter permits insertion of a press-fit cylindrical sleeve of beryllium oxide or boron nitride between the transistor and the metal block to provide a

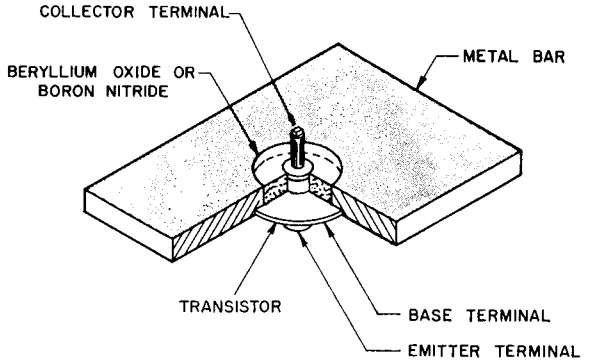



Figure 583. Mounting arrangement for a coaxial-package transistor in a microstripline circuit.

heat-conducting path from the collector stud to the block. The diameters of the hole through the metal block and the cylinder of beryllium oxide (or boron nitride) are determined by the desired characteristic impedance of the short coaxial-line section which is formed by this mounting technique. Beryllium oxide and boron nitride have excellent heat conductivity and low electrical losses and thus provide satisfactory heat dissipation from the coaxial transistor without adversely affecting the rf performance.

The circuit arrangement shown in Fig. 583 is excellent for isolation of the input and output circuits. The output circuit is constructed on the top portion of the metal block and the input circuit on the bottom portion. Fig. 584 shows the construction of the microstripline circuit. The output circuit is constructed of standard microstripline mounted to the top surface of the metal block. The input circuit is constructed of another microstripline placed directly over the bottom surface of the metal block. A stripline circuit can be formed by placing another strip of dielectric material and ground plane above the conductor strips of Fig. 584.

The design of transistor microwave power circuits involves two steps: (1) the determination of load and input impedances under dynamic operating conditions, and (2) the design of properly distributed filtering and matching networks required for optimum circuit performance. For design of the input circuit, the input impedance at the emitter-to-base terminals of the packaged transistor at the drive-power frequency under operating conditions must be known. For design of the output circuit, the load impedance presented to the collector terminal at the fundamental frequency must be known. These dynamic impedances are difficult to calculate at microwave frequencies because transistor parameters vary considerably under large-signal operation from small-signal values, and also change with power level. Small-signal equations that might serve as useful guides for transistor design cannot be applied rigorously to large-signal circuits, although it has been determined empirically that some small-signal parameters at the 10-volt level correspond rather closely

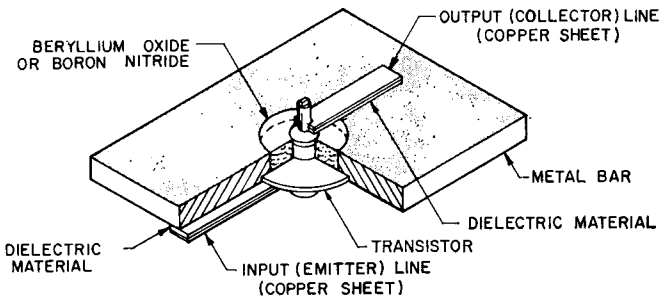



Figure 584. Construction of the microstripline circuit.

with the large-signal values at 28 volts. Because practical large-signal representation of microwave transistors has not yet been developed, transistor dynamic impedances are best determined experimentally by use of slotted-line or vector-voltmeter measurement techniques.

The system required to determine transistor impedances under operating conditions is shown in Fig. 585. This system consists of a well-padded power signal generator, a directional coupler (or reflectometer) for monitoring the input reflected power, an input triple-stub tuner, an input low-impedance line section, the transistor holder (or test jig), an output line section, a bias tee, an output triple-stub tuner, another directional coupler for monitoring the output waveform or frequency, and an output power meter. At a given frequency and input-power level, the input and output tuners are adjusted for maximum power output and minimum input reflection power. When the system has been tuned properly, the impedance across terminals 1-1, without the transistor in the system, is measured at the same frequency in a slotted-line set-up or with a vector voltmeter. The conjugate of this impedance equals the dynamic input impedance of the transistor. Similarly, the impedance across terminals 2-2, without the transistor in the system, is the load impedance presented to the collector of the transistor. Such measurements are performed at each frequency and power level.

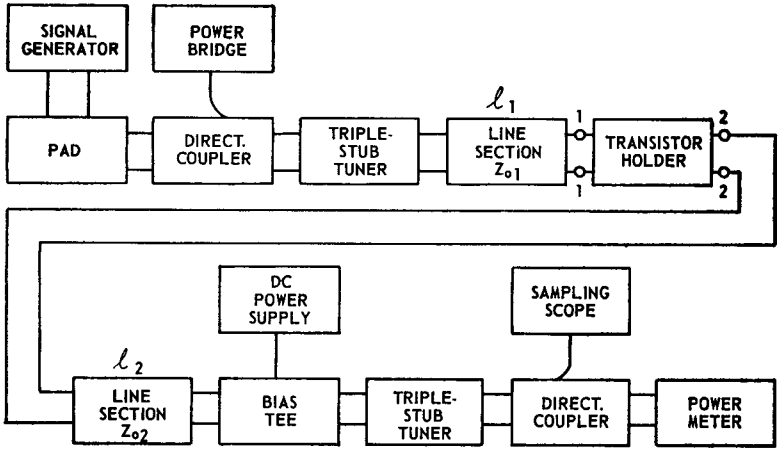



Figure 585. Block diagram of test setup used to determine input and output impedances of transistors.

In addition to determining dynamic input impedance and load impedance, the system shown in Fig. 585 is useful for determination of the performance capability of the transistor. Power output, power gain, and efficiency are readily determined. For optimum performance of the test system, careful consideration must be given to the selection of the line length and the characteristic impedance $Z_0$ of the input and output line sec-

tions $l_1$ and $l_2$, respectively). Eighth-wavelength ($\lambda/8$) line sections are preferred for $l_1$ and $l_2$ because such sections exhibit the lowest VSWR and the smallest line losses.

An alternative method of determining the dynamic input impedance is shown in Fig. 586. This method uses a well-padded, high-power signal generator connected in series with a slotted-line setup.

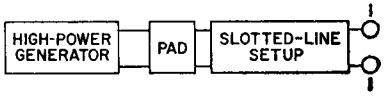



Figure 586. Block diagram of dynamic-impedance test setup that may also be used to test transistor performance capability.

The setup beyond terminals 1-1 is identical to that of Fig. 585. The high-power generator is adjusted until a desired power output is obtained. The input impedance under this condition can be measured simultaneously in a slotted-line setup. In this case, the test fixture must contain a short line section (a $\lambda/8$ section is preferred for smallest line losses) to provide a connection to the transistor.

When the dynamic input impedance and the load impedance of a packaged transistor have been established, either by direct measurement (as described in the preceding paragraph) or from published data, the input and output circuits can be designed.

Some simple designs are shown in Fig. 587. Although coaxial-line configurations are shown, the design procedures are similar for the other forms of TEM-mode distributed line sections. For the circuit shown in Fig. 587(a), the line section $l$ transforms the small input impedance of the transistor to a value closer to that of the driving-source resistance (such as a 50-ohm generator). If line section $l$ is made an eighth-wavelength long and its characteristic impedance $Z_0$ is properly determined, then the complex input impedance is transformed to a real value at the other end of this line, and the VSWR on the line section is a minimum. Capacitors $C_1$ and $C_2$, together with some lead inductance, are used as reactive dividers to step up or step down the impedance, depending on the value of the real impedance compared to the 50-ohm source. Transformation directly to 50 ohms or some other desired real impedance is also possible with this configuration. The length of the line section $l$ is less than a quarter-wavelength when the dynamic input impedance is inductive and greater than a quarter-wavelength for capacitive inputs. In this type of application, capacitors $C_1$ and $C_2$ serve to tune out imaginary components, modify imaginary components, or adjust the values of real components, depending on the frequency and the characteristics of the line section $l$.

The input circuit shown in Fig. 587(b) can be used effectively when the dynamic input impedance of the transistor is inductive. Capacitor $C_2$ is used to tune out the inductive component of the input impedance. A quarter-wave line of the proper characteristic impedance is then used for the impedance transformation between the small input resistance of the transistor and the driving-source resistance.

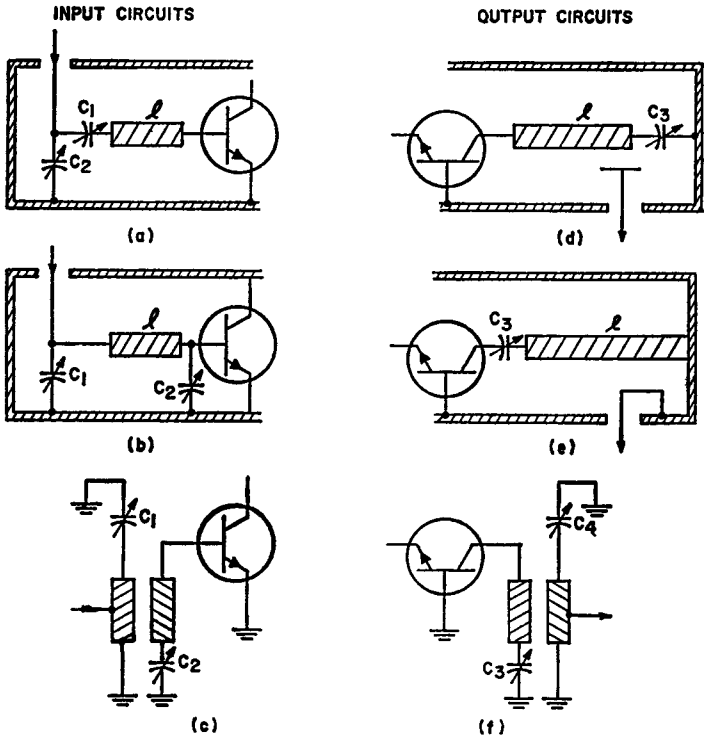



Figure 587. Transistor input and output coupling circuits suitable for use at microwave frequencies: (a) direct-coupled input network using series tuning capacitor; (b) direct-coupled input network using shunt tuning capacitor; (c) resonant-line input circuit; (d) capacitive-probe-coupled output cavity; (e) inductive-probe-coupled coaxial output cavity; (f) resonant-line output circuit.

Capacitor $C_1$ may be used to adjust for minor differences between transistors.

The output circuit shown in Fig. 587(d) is a capacitive-loaded, foreshortened quarter-wave coaxial-line cavity. A capacitive probe is used to match the output to the desired real load impedance. In the design of the circuit, the line section $l$, the capacitance $C_3$, and the dynamic output capacitance of the transistor must be resonant at the desired frequency.

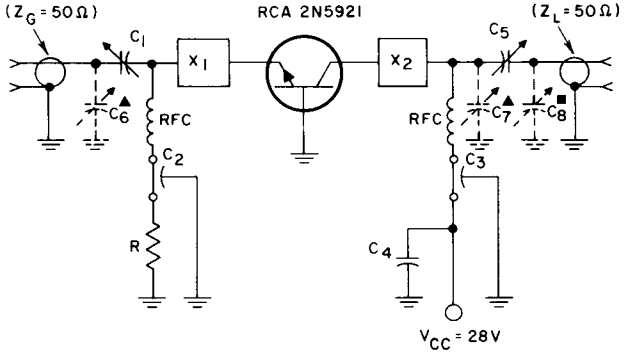
The ouput circuit shown in Fig. 587(e) is similar to that shown in Fig. 587(d) except that inductive loop coupling is used. Again, the design of the coupling loop is empirical. In general, the inductive loop is placed near the ground (high-current) end of the line; in fact, it may be tapped directly to the center conductor. Conversely, capacitive probes are generally located near the high-voltage end of the line.

The coupling networks shown in Fig. 587(a) through Fig. 587(e) can apply to either input or output circuits, and the specific illustrations are used for discussion only. The circuits

shown in Figs. 587(c) and 587(f) make use of inductive coupling and are particularly suitable for stripline circuits.

## MICROWAVE AMPLIFIER CIRCUITS

In general, transistor amplifiers are operated in the common-base configuration at microwave frequencies. As was mentioned in the previous section, either coaxial or microstrip circuit techniques can be utilized. Fig. 588 shows a typical 2N5921 coaxial-circuit amplifier for 1.2- or 2-GHz operation, and Fig. 589 shows the mechanical construction of the circuit. Figs. 590 and 591 show



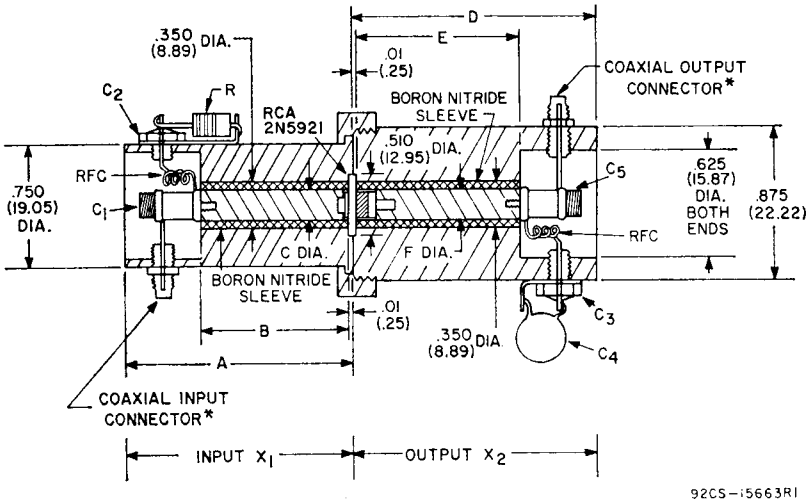

▲ Use only in the 2-GHz coaxial-line power amplifier circuit.
■ Use only in the 1.2-GHz coaxial-line test circuit.

| CIRCUIT | $C_1$ pF | $C_2$ pF | $C_3$ pF | $C_4$ μF | $C_5$ pF | $C_6$ pF | $C_7$ pF | $C_8$ pF | R ohm |
|---|---|---|---|---|---|---|---|---|---|
| 1.2 GHz (Test Circuit) | 1-10 | 1000 | 1000 | 0.01 | 1-10 | — | — | 0.3-3.5 | 0.75 |
| 2 GHz (Test Circuit) | 1-10 | 470 | 470 | 0.01 | 1-10 | — | — | — | 0.43 |
| 2 GHz (Amplifier) | 1-10 | 470 | 470 | 0.01 | 0.3-3.5 | 0.3-3.5 | 0.3-3.5 | — | 0.43 |

$C_1$, $C_5$ = 1 to 10 pF, Johanson 4581, or equivalent*
$C_5$, $C_6$, $C_7$, $C_8$ = 0.3 to 3.5 pF, Johanson 4700, or equivalent*
RFC (For 2-GHz Circuits) = 3 turns No. 32 wire 1/16 in. (1.59 mm) ID, 3/16 in. (4.76 mm) long.
(For 1.2-GHz Circuit) = 6 turns No. 32 wire 1/16 in. (1.59 mm) ID, 3/16 in. (4.76 mm) long.
$X_1$, $X_2$: Coaxial-line circuits, see Fig. 589.

* Johanson Mfg. Corp., Boonton, N.J. 07005

Figure 588. 1.2- or 2-GHz coaxial-line amplifier circuit.




DIMENSIONS OF COAXIAL LINES $X_1$ AND $X_2$

| CIRCUIT | INPUT ($X_1$) | | | | OUTPUT ($X_2$) | | | |
|---|---|---|---|---|---|---|---|---|
| | A | B | C | Center Conductor | D | E | F | Center Conductor |
| 1.2-GHz (Test Circuit) | 1.385 (35.18) | 0.875 (22.22) | 0.282 (7.16) | 0.825 (20.95) | 1.778 (45.16) | 1.268 (32.21) | 0.213 (5.41) | 1.05 (26.67) |
| 2-GHz (Test Circuit) | 0.940 (23.88) | 0.430 (10.92) | 0.266 (6.76) | 0.380 (9.65) | 1.04 (26.42) | 0.530 (13.46) | 0.266 (6.76) | 0.370 (9.39) |
| 2-GHz (Amplifier) | 0.860 (21.84) | 0.350 (8.89) | 0.265 (6.73) | 0.300 (7.62) | 1.06 (26.92) | 0.550 (13.97) | 0.270 (6.86) | 0.385 (9.78) |

Dimensions in inches and millimeters
*Dimensions in parentheses are in millimeters and are derived from the basic inch dimensions as indicated.*
MATERIAL: Center conductor—copper
Outer conductor for input and output—brass

* Conhex 50-045-0000 Sealectro Corp., or equiv.

Figure 589. Constructional details for 1.2- or 2-GHz coaxial-line circuit.

typical performance that can be obtained with this transistor.

As can be seen from these figures, the 2N5921 can provide 5 watts output at 2 GHz with a gain of 7 dB and a collector efficiency greater than 40 per cent.

To design an amplifier such as shown in Fig. 589, the input and output impedances of the transistor must be known. Fig. 592 shows typical large-signal input impedance and collector load impedance values for the RCA-2N5921 as functions of frequency. These data are particularly useful for the design of broadband circuits.

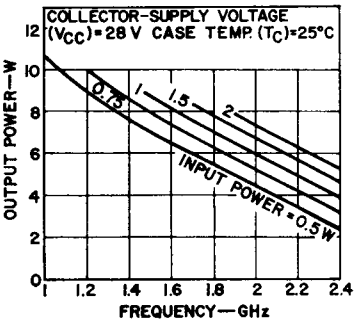



Figure 590. Typical output power as a function of frequency for the 2N5921 transistor.

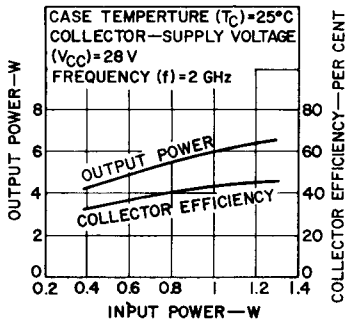



Figure 591. Typical power output or collector efficiency as a function of power input at 2 GHz for circuit shown in Fig. 589.

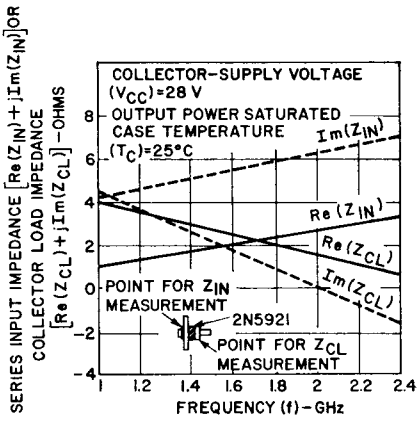



Figure 592. Typical large-signal series input impedance and large-signal collector load impedance as a function of frequency for the RCA-2N5921.

Fig. 593 shows construction details of the lower-power amplifier that uses the 2N5920, and Figs. 594 and 595 show performance characteristics of this transistor when used in the common-base configuration. The 2N5920 is particularly well suited for use as a driver for the 2N5921.

For even lower powers, the RCA-2N5470 can be utilized. Circuit and performance details of an

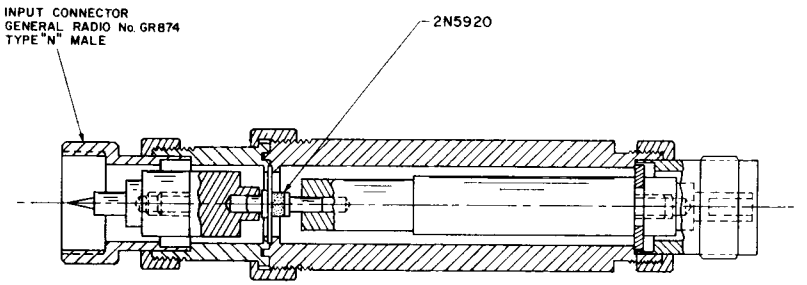



Figure 593. Constructional details of 2 GHz power amplifier.

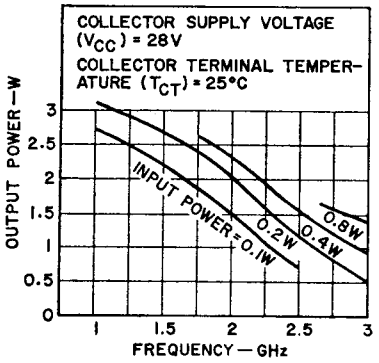



Figure 594. Typical cutput power as a function of frequency for common-base amplifier.

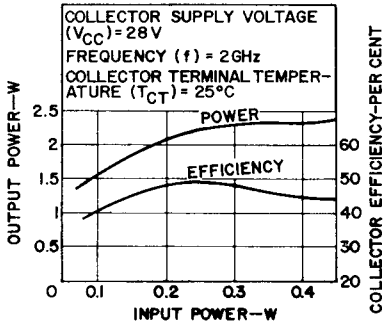
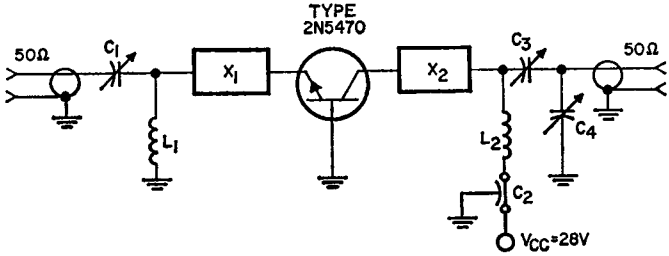
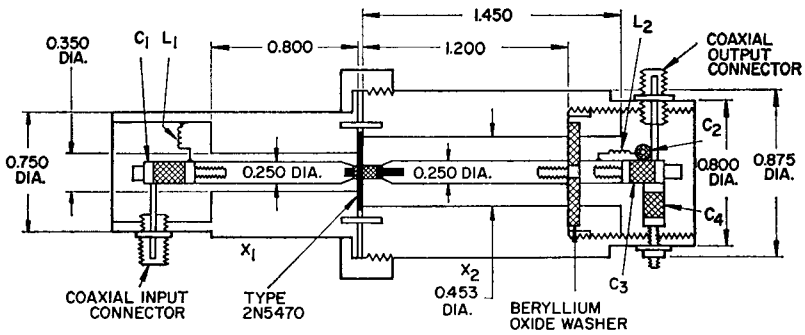



Figure 595. Typical output power and collector efficiency as a function of input power for 2-GHz common-base power amplifier.




$C_1$ = 0.8 to 10 pF, Johnson 4355, or equiv.
$C_2$ = 1000 pF, feedthru, Allen-Bradley FB2B, or equiv.
$C_3$ = 0.3 to 3.5 pF, Johnson 4701, or equiv.
$C_4$ = 0.35 to 3.5 pF, Johnson 4702, or equiv.
$L_1$, $L_2$ = RF choke, 3 turns No. 30 wire, 1/16" ID, 3/16" long
$X_1$, $X_2$ = Details given in (b)

(a)




DIMENSIONS IN INCHES

(b)

Figure 596. 2 GHz power amplifier using the RCA-2N5470 coaxial transistor: (a) circuit schematic; (b) construction details.

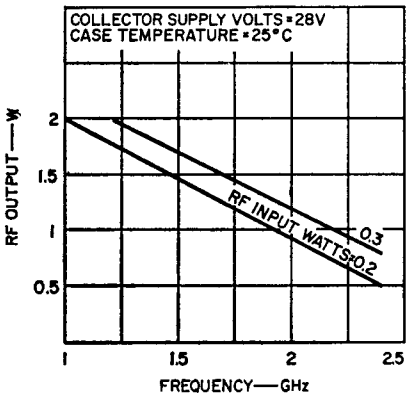



Figure 597. Power output as a function of frequency for the RCA 2N5470 transistor.

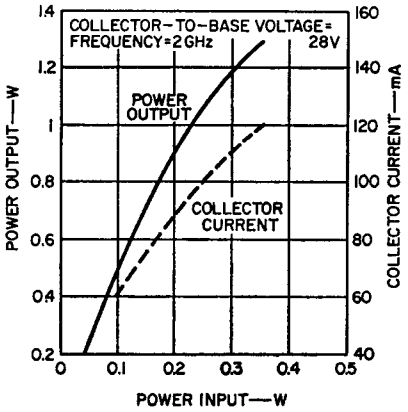
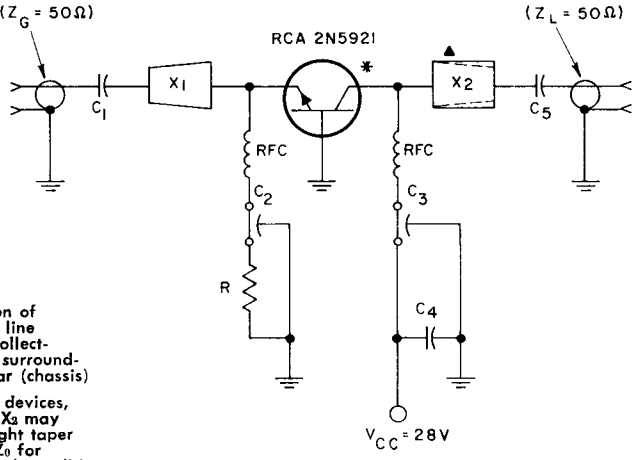



Figure 598. Power output as a function of input for amplifier shown in Fig. 596.




*Short section of transmission line formed by collector stud and surrounding metal bar (chassis)

▲with some devices, load end of $X_2$ may require a slight taper to increase $Z_0$ for optimum match condition.

$C_1$,$C_5$=300 pF, disc ceramic
$C_2$,$C_3$=470 pF, feed through, Allen-Bradley FAC5, or equivalent
$C_4$=0.01 μF, disc ceramic
R=0.43 ohm
RFC=No. 32 wire, 0.4 in. (1.02 mm) long
$X_1$=Tapered microstripline—
0.15 in. (3.81 mm) wide, input end
0.30 in. (7.62 mm) wide, output end
0.525 in. (13.33 mm) long
0.005 in. (0.13 mm) thick, copper

$X_2$=Uniform microstripline—
0.25 in. (6.35 mm) wide
0.36 in. (19.14 mm) long
0.005 in. (0.13 mm) thick, copper
DIELECTRIC MATERIAL: 0.5 in. (12.7 mm) wide
0.75 in. (19.05 mm) long
0.005 in. (0.13 mm) thick
DuPont H-Film, or equiv.

Figure 599. Typical circuit for 2-GHz grounded-base microstripline power amplifier.

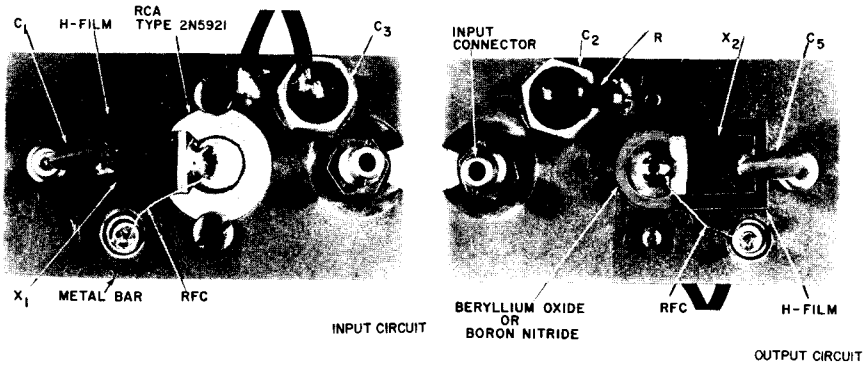



Figure 600. Suggested mounting arrangement of components for 2-GHz microstripline circuit shown in Figure 599.

RCA-2N5470 common-base amplifier are shown in Figs. 596 through 598.

The circuits in Figs. 588, 589, 593, and 596 employ coaxial resonators and matching circuits. Microstripline circuitry can also be used with excellent results. Fig. 599 shows a typical 2-GHz microstripline circuit used with the RCA-2N5921. Construction details of this circuit are shown in Fig. 600.

As discussed previously, microwave transistors are also available in stripline packages. Figs. 601 and 602 show a circuit and suggested construction for a 1- or 2-GHz amplifier that uses an RCA stripline transistor. Typical performance of this transistor is shown in Figs. 603 and 604.

The RCA stripline transistor used in the circuit shown in Fig. 601 can provide 5 watts of power output at 2 GHz with a gain of 7 dB and a collector efficiency of 40 per cent.

Another RCA stripline transistor can provide 2 watts of output power at 2 GHz with a minimum gain of 8.2 dB and collector efficiency greater than 35 per cent.

## AMPLIFIER CHAINS

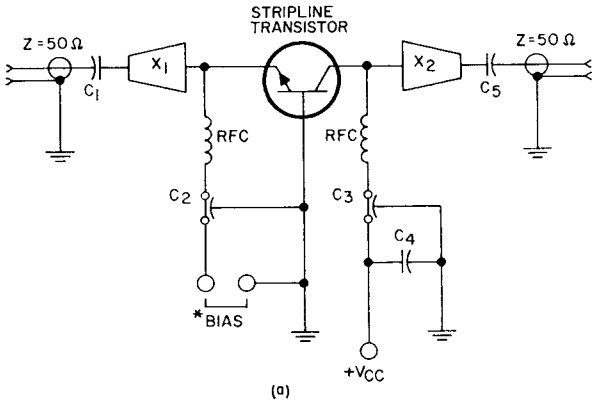
Most amplifiers include several stages of amplification. Fig. 605 shows a typical 2-GHz power amplifier chain that provides 8 watts of output power with a gain of 26 dB. The power from this chain is limited by the power capability of the output transistor.

Two or more transistors can be combined as shown in Fig. 606 to obtain higher power output. This circuit provides 15 watts at 1.8 GHz by use of quadrature hybrid couplers to parallel two output transistors. These couplers also serve to isolate the driver stage from the output stage. Other combiners, such as the Wilkinson and the simple reactive splitter, can also be used.

If interstage isolation and transistor-to-transistor isolation are not required, the transistors can be paralleled directly, provided that strict attention is given to symmetry of circuit layout.

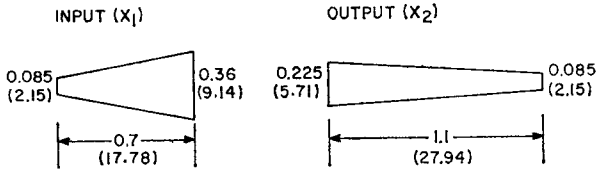
## MICROWAVE OSCILLATOR CIRCUITS

Transistors capable of power amplification are also suitable for power oscillation. The most im-

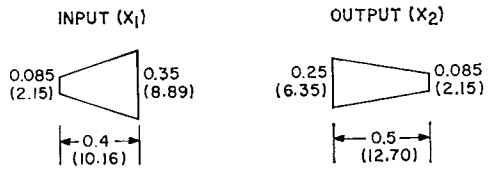



(a)

1 GHz:




2 GHz:




(b)

$C_1$, $C_5$=300 pF, ATC-100, or equivalent
$C_2$, $C_3$=Filtercon, Allen-Bradley SMFB-A1, or equivalent
$C_4$=0.001 μF, disc ceramic
RFC=2 GHz - No. 32 wire, 0.4 in. (1.02 mm) long
1 GHz - No. 32 wire, 3 turns, 1/16 in. (1.59 mm) ID, 3/16 in. (4.76 mm) long
$X_1$, $X_2$=Microstripline circuits

Note: All sections are exponentially tapered.●
Dielectric=5-mil DuPont H-Film, or equivalent
Line Material=5-mil copper
* Bias terminals normally have $R_E$ = 0.24 ohm

● Use *sparing* amount of Eastman 910 Adhesive to bond lines and dielectric to circuit board.

Figure 601. (a) Circuit configuration for 1- or 2-GHz stripline amplifier and (b) details of microstripline sections for each frequency.

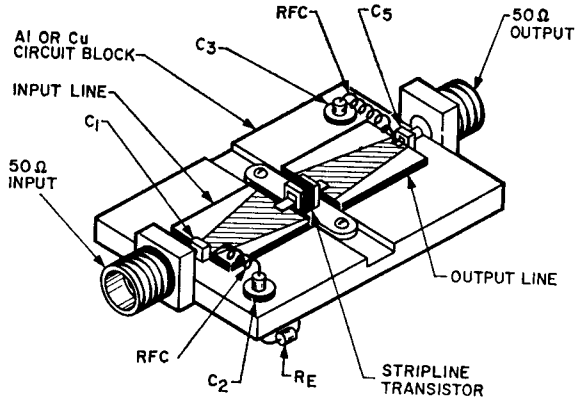



Figure 602. Suggested construction for stripline amplifier shown in Figure 601.

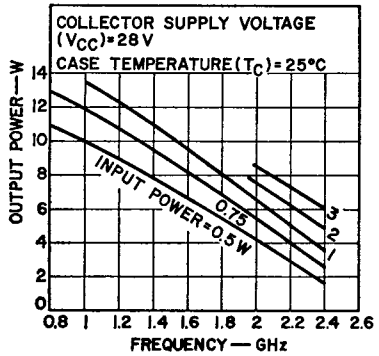



Figure 603. Typical output power as a function of frequency for amplifier shown in Figure 601.

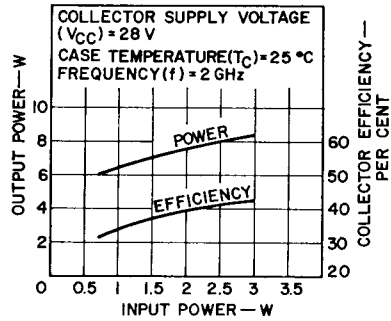



Figure 604. Typical output power or collector efficiency as a function of input power at 2 GHz for amplifier shown in Figure 601.

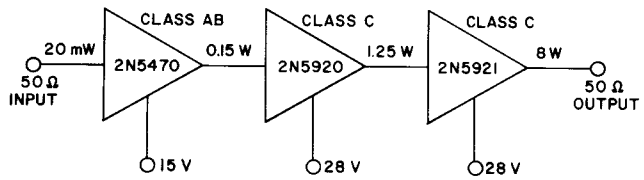



Figure 605. Typical 2-GHz power-amplifier chain.

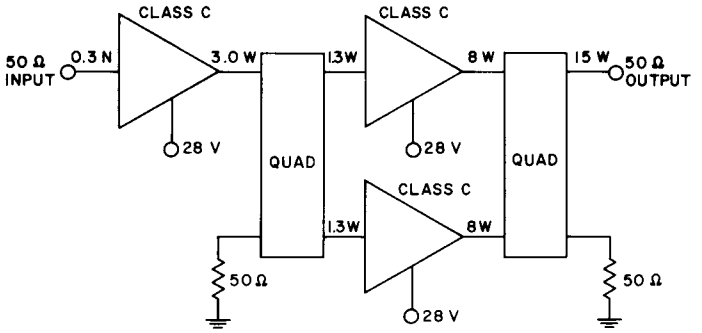



Figure 606. Typical 1.8-GHz power-amplifier chain using quadrature combiners.

portant part of every oscillator is an element of amplification. It is then necessary only to provide a path that feeds back a part of the power output to the input in the proper phase, together with a source of dc power. The maximum frequency of oscillation, which is related to $f_{max}$ in a small-signal transistor, is usually difficult to define in a uhf or microwave power transistor because of the added parasitic elements. The circuit-design approach for an oscillator circuit is similar to that discussed previously for amplifier circuits.

Fig. 607(a) shows a Colpitts transistor oscillator suitable for microwave applications. The inductance L and the capacitances $C_1$ and $C_2$ can be considered as the parasitic elements of the package. The transistor can be grounded at the collector, the base, or the emitter without effect on its performance. For example, a useful oscillator circuit can be derived from the basic Colpitts oscillator by the use of a TO-39 transistor. In Fig. 607(b), the collector of such a transistor is returned to ground through the collector parasitic inductance L. This connection is a convenient method of applying a heat sink to the collector, which is connected to the case in a TO-39 package. Output power is obtained from the base through capacitances $C_3$ and $C_4$. Fig. 607(c) shows another method of coupling power output from the oscillator.

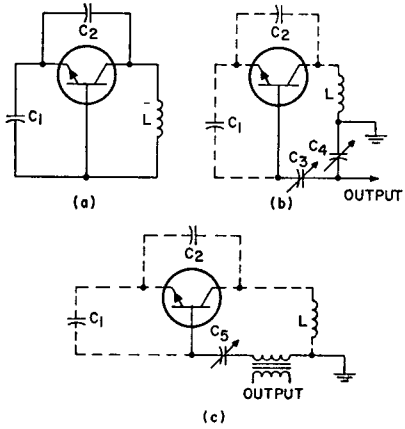
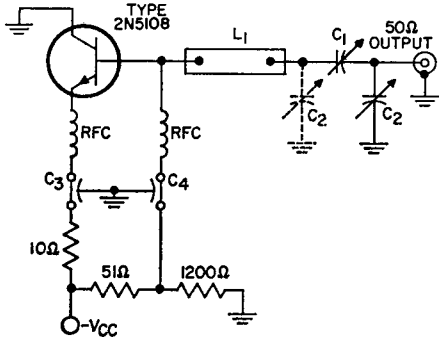



Figure 607. Colpitts oscillator for use at microwave frequencies: (a) basic ac circuit configuration; (b) basic ac circuit with collector returned to ground through parasitic inductance L and the output taken from base through capacitive voltage divider; (c) basic ac circuit with transformer-coupled output.

Fig. 608 shows the complete circuit diagram of a 1.68-GHz fundamental-frequency oscillator which makes use of the RCA-2N5108 transistor. The collector




$C_1$, $C_2$ = Variable capacitor, 0.35 to 3.5 pF, piston type
$C_3$, $C_4$ = 470 pF, feedthru
$L_1$ = Described in text
RFC = 5 turns No. 28 wire, 1/8" ID x 1/2" long

Figure 608. 1.68-GHz fundamental-frequency oscillator using the RCA-2N5108 transistor.

of this transistor, which is packaged in a TO-39 case, is grounded to the ground plane of a 1/16-inch Teflon fiberglass microstripline board. Power output is taken from the base through a 0.75-inch section of 50-ohm microstripline and the capacitor network $C_1$ and $C_2$. This oscillator can supply more than 0.3 watt of power output at 1.68 GHz and has an efficiency of 20 per cent when operated from a supply voltage of 25 volts. Fig. 609 shows the oscillator output power as a function of supply voltage.

This basic oscillator circuit is useful at frequencies from 1 to 2 GHz; only slight modifications in the length of the transmission-line $L_1$ are required to cover this range. For example, the line length is increased to 0.8 inch

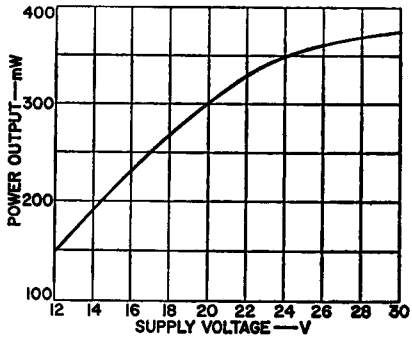
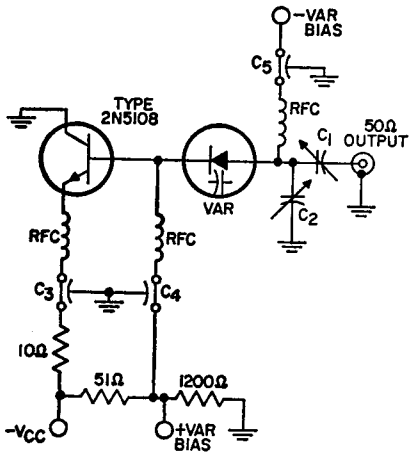



Figure 609. Power output as a function of supply voltage for the 1.68-GHz oscillator shown in Fig. 608.

to obtain optimum circuit operation at 1.5 GHz. An output power of 400 milliwatts (with a 24-volt supply) can be expected at this frequency. Another modification of interest (with the 0.8-inch line) is that optimum operation at 1.25 GHz is achieved simply by movement of capacitor $C_2$ to the position indicated by the dotted lines. Movement of this capacitor results in an improved output transformation network which can develop more than 800 milliwatts of output power at 1.25 GHz for operation from a 24-volt supply.

The inductive element introduced by the line section $L_1$ (Fig. 608) can be supplied by a high-Q varactor diode operated above its resonant frequency, as shown in Fig. 610. The bias supplied to this varactor, in effect, electrically varies this inductive component so that broadband oscillator tuning is possible. The output capacitor network, $C_1$ and $C_2$, which is used to transform a relatively small load-line impedance to the 50 ohms of the output port, could be replaced with an inductive-reactive divider-network, such as a tapped



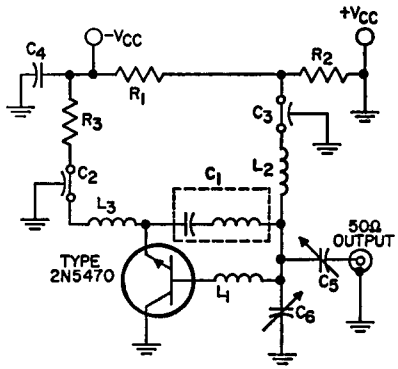

RFC = 0.1 μH
C1, C3 = 1 to 7 pF, piston capacitors
C2, C4, C5 = 470 pF, feedthru
Var. = Described in text

Figure 610. Wideband varactor-tuned L-band oscillator using the RCA-2N5108 transistor.

transmission line or helical coaxial line. Tests in which a cartridge-type silicon microwave varactor is used in this circuit show a relatively constant power output of 600 milliwatts over the range of 1.0 to 1.5 GHz. The bias on this particular varactor ranges between 0 and 22 volts for the specified tuning range, and a transistor collector supply of 28 volts is used.

The RCA-2N5470 coaxial transistor, although designed for stable operation at 2.3 GHz in the common-base amplifier mode, can also deliver a power output of 0.3 watt at 2.3 GHz as an oscillator. In oscillator applications of the 2N5470, advantage is taken of the very low parasitic elements in this transistor to simplify circuit requirements, e.g., essentially lumped-constant S-band circuits can be designed around this unit. However, because of the low feedback capacitances of this transistor, external feedback loops are generally needed for stable oscillation at S-band frequencies.

Fig. 611 shows a simple lumped-constant circuit that uses the 2N5470 transistor. The circuit is tunable over the frequency range of 1.8 to 2.3 GHz. Power output at 2 GHz is typically 0.3 watt with a 24-volt supply, and circuit efficiency is in the order of 16 per cent at this frequency. The collector is grounded, and power output is taken from the base circuit. All leads must be kept short for best high-frequency response. The "gimmick" capacitor $C_1$ forms a necessary part of the feedback loop of the circuit. The circuit is basically a Hartley type of oscillator in that inductor $L_1$ and the parasitic inductance of $C_1$ make up a tapped inductor in this feedback loop.

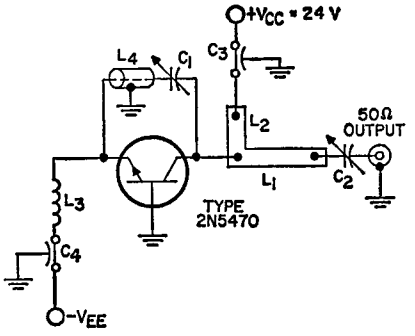



C1 = 0.82 pF, "gimmick" capacitor (manufactured by Quality Components, Inc. St. Mary's, Pa.)
C2, C3 = 100 pF, Allen-Bradley FA5C or equiv.
C4 = 0.01 pF, disc ceramic
C5, C6 = Trimmer capacitor 0.35 to 3.5 pF, Johanson Type 4702 or equiv.
L1 = 0.05" length of No. 22 wire
L2, L3 = 4 turns 7-mil wire, .062" ID x 3/16" long
R1 = 51 ohms, 1/2 watt
R2 = 1200 ohms, 1/2 watt
R3 = 5 to 10 ohms, 1/2 watt

Figure 611. Lumped-constant 2-GHz oscillator circuit using the RCA-2N5470 transistor.

Tuning is achieved largely by adjustment of capacitor $C_6$, and capacitor $C_5$ is adjusted to maintain the output match over the tuning range.

Figure 612 shows the use of the 2N5470 transistor in a Colpitts type of microstripline oscillator circuit that operates




$C_1$, $C_2$ = 0.35 to 3.5 pF, Johanson Type 4702 or equiv.
$C_3$, $C_4$ = 100 pF, Allen-Bradley Type 5A5C or equiv.
$L_1$ = microstrip line; ≅ 0.70" long x 0.30" wide strip; mounted on 1/32" Teflon fiberglass board
$L_2$ = microstrip line; ≅ 0.43" long x 0.080" wide strip; mounted on 1/32" teflon fiberglass board
$L_3$ = 5 turns 7-mil wire, 0.062" ID x 3/16" long
$L_4$ = 50-ohm miniature coaxial line, 1.5" long

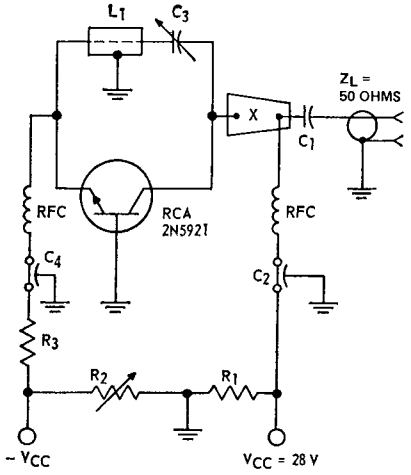
Figure 612. Microstripline 2-GHz oscillator circuit using the RCA-2N5470 transistor.

over the frequency range of 1.8 to 2.2 GHz. In this circuit, the base of the transistor is directly grounded to the ground plane of the stripline board, and collector heat is dissipated to this board through a beryllium oxide insulating washer. The necessary feedback is provided by the phase-resonant loop provided by line section $L_4$ and capacitor $C_1$. The output line section $L_2$ makes use of standard microstripline techniques to provide the necessary reactance to tune out the output capacitance; line section $L_1$ is a quarter-wave transformer which transforms the real part of the collector load impedance to about 50 ohms. This circuit can also provide about 0.3 watt of output power at 2 GHz when operated from a 24-volt supply.

For higher power outputs, the RCA-2N5921 or the RCA-2N5920 coaxial transistors and RCA stripline transistors can be used. Fig. 613 shows the 2N5921 in a power oscillator circuit that tunes from 1.2 to 1.4 GHz. In this circuit, the coaxial transistor is mounted with the collector protruding through a 0.35-inch hole in a metallic circuit board that is 0.20 inch thick. Output line X is constructed on the collector side of this circuit board, while the bias network is on the other side. Coaxial-line section $L_1$ is used to bring the collector feedback loop through the circuit board to the emitter section of the transistor. The 2N5921 in this circuit can deliver 4 watts of output power at 1.2 to 1.4 GHz. Frequency is tuned by adjustment of $C_3$.

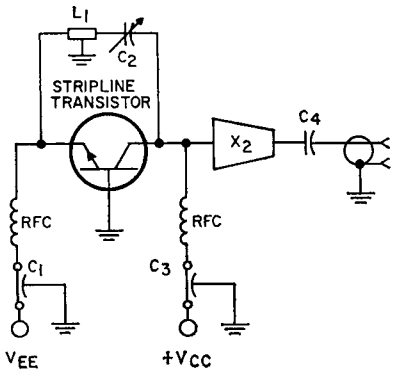
Fig. 614 shows a typical oscillator circuit that uses a stripline transistor. This oscillator can develop 3 watts of output power over the range of 1.7 to 1.8 GHz.

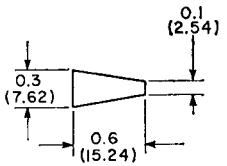
The oscillators shown so far have used transistors with the base connected to the flange. For lower-noise oscillator operation it has been found desirable to connect the emitter of the transistor to the flange. The RCA-40836 and RCA-40837 transistors are versions of the 2N5470 and 2N5920, respectively, in which the emitter, rather than the base, is connected to the flange. A simple Colpitts-type power oscillator using essentially lumped-constant circuit elements is shown in Fig. 615. This oscillator can be tuned over the range from 1.8




$C_1$=300 pF, disc ceramic
$C_2$, $C_4$=470 pF, feed-through type, Allen-Bradley FA5C, or equivalent
$C_3$=0.3-3.5 pF, Johanson 4702, or equivalent
$L_1$=1.3 in. (33.02 mm) length of 50-ohm coaxial line
$R_1$=1200 ohms
$R_2$=0-250 ohms
$R_3$=5 ohms
RFC=3 turns, No. 29 wire, 0.06 in. (1.59 mm) I.D., 0.18 in. (4.77 mm) lng.
X=Tapered microstripline—
0.1 in. (2.54 mm) wide, input end
0.24 in. (6.09 mm) wide, output end
0.475 in. (12.06 mm) long
0.005 in. (0.13 mm) thick, copper
DIELECTRIC MATERIAL=DuPont 5-MIL Kapton H, or equivalent

Figure 613. Typical circuit for tunable 1.2-to-1.4-GHz, 4-watt microstripline power oscillator.



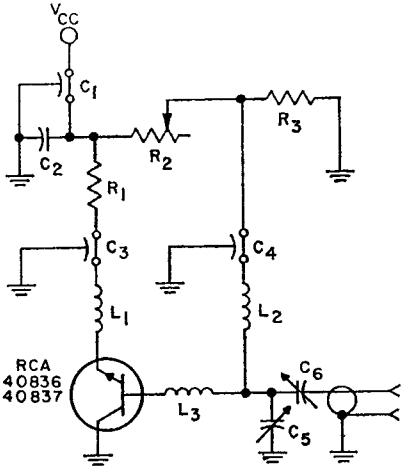




Note: Line construction similar to that of circuit shown in Fig. 613.

$C_1$, $C_3$=Filtercon, Allen-Bradley SMFB-A1, or equivalent
$C_2$=0.3-3.5 pF, Johanson 4700, or equivalent
$C_4$=300 pF, ATC-100, or equivalent
$L_1$=1.0 in. section miniature 50-ohm cable
RFC=3 turns, No. 32 wire, 1/16 in. ID, 3/16 in. long

Figure 614. Typical 1.7-GHz oscillator circuit.

to 2.1 GHz; minor circuit modifications permit tuning over a 200-MHz bandwidth in the range of 1.3 to 2.6 GHz.




$C_1$, $C_3$, $C_4$=470-pF, Feed-through Allen-Bradley FA5C, or equivalent
$C_2$=0.2 μF, disc ceramic
$C_5$, $C_6$=0.35 to 3.5-pF, Johanson 4702 or equivalent
$L_1$, $L_2$=RF choke, 0.5 in. (12.70 mm) length of No. 32 wire
$L_3$=Copper strip:
0.005 in (0.127 mm) thick
0.18 in. (0.457 mm) wide
0.3 in (0.76 mm) long
$R_1$=3 to 10 ohms, ½ W
$R_2$=0 to 500 ohms, 2 W
$R_3$=1200 ohms, ½ W

Figure 615. Typical 2-GHz grounded-collector power oscillator.

The RCA-40836 can deliver 0.5 watt of output power with a $V_{CE}$ of 21 volts. The RCA-40837 can deliver 1.25 watts at 2.1 GHz with a $V_{CE}$ of 27 volts.

This circuit can operate below 1.8 GHz if the emitter-base capacitance and inductance $L_3$ are increased. For operation in the range of 2.1 GHz to 2.6 GHz, the emitter-collector capacitance must be increased, and $L_3$ must be decreased. Capacitors $C_5$ and $C_6$ become self-resonant at 2.6 GHz and cannot be used as capacitors above this frequency. The RCA-40837 can be made to operate in the 2.6-to-3.0-GHz range by reduction of $L_3$ to only the intrinsic inductance of the base cap of the coaxial package, and use of capacitive probes for $C_5$ and $C_6$. Electronic tuning of the circuit can be achieved by use of a varactor to vary the effective inductance of $L_3$.

## FREQUENCY MULTIPLIERS

Operation of the overlay transistor in the harmonic-frequency mode can extend the upper limit of the frequency range far beyond that possible from the same transistor operating in the fundamental-frequency mode. A further advantage of the harmonic mode of operation is that frequency multiplication and power amplification can be realized simultaneously. An overlay transistor operating in this mode provides power amplification at the fundamental frequency of the input-drive power, and the nonlinear capacitance of the collector-to-base junction, acting as a varactor, generates harmonics of the input-drive frequency. It is possible, therefore, to use a single transistor to replace a transistor power amplifier and a varactor-diode frequency multiplier. In comparison with varactor frequency-multiplier circuits, the transistor multiplier is simpler, less costly, and equally efficient. It is anticipated that this mode of operation will permit extension of the available frequency spectrum for overlay transistors by a factor of two.

### Transistor Considerations

An overlay transistor used in a frequency-multiplier circuit oper-

ates simultaneously as a power amplifier to provide gain at the fundamental frequency of the input driving power and as a varactor diode to generate harmonics of the driving power frequency. Thus, two mechanisms provide amplification and frequency multiplication in overlay transistors: one capable of gain at the fundamental frequency, and the other in which the collector-base capacitance serves as a varactor capable of frequency multiplication. Transistors suitable for multiplier applications must be capable of delivering power with gain at the fundamental frequency and of converting the power from the fundamental frequency to a harmonic frequency. A good multiplier transistor, therefore, must first be a good uhf transistor capable of high power output, gain, and efficiency. In addition, its varactor section should have minimum losses to provide maximum conversion efficiency.

The figure of merit for the amplifier portion of the transistor in which parasitic elements are not included is given by the maximum frequency of oscillation $f_{max}$ as follows:

$$f_{max} = (PG)^{\frac{1}{2}} f$$
$$= [(1/8\pi)(1/r_{bb}' C_c \tau_{ec})] \quad (431)$$

where PG is the power gain, f is the fundamental frequency of operation, $r_{bb}'$ is the intrinsic base-spreading resistance, $C_C$ is the collector capacitance, and $\tau_{ec}$ is the emitter-to-collector transit or signal-delay time.

The efficiency of the varactor portion formed by the collector-base junction is determined by the cutoff frequency $f_{VCB}$ as follows:

$$f_{VCB} = 1/[2\pi C_{min} (r_b'+r_s)] \quad (432)$$

where $C_{min}$ is the minimum collector-to-base capacitance and $r_s$ is the collector series resistance.

Fig. 616 shows a cross-sectional view of an overlay transistor that indicates the capacitance and loss

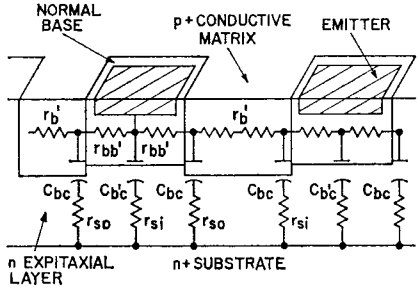



Figure 616. Cross-sectional view of an overlay transistor indicating the capacitance and loss distributions.

distributions. Fig. 617 shows how the varactor portion separates from the intrinsic transistor portion. The collector-to-base capacitance consists of two parts. The major part, which comprises the active portion of the varactor, consists of the capacitance formed by the part of the collector-to-base junction that is not opposite emitter sites. This part of the capacitance is called the outer collector

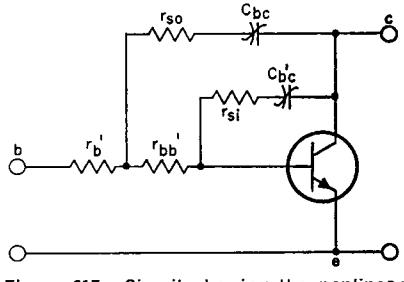



Figure 617. Circuit showing the nonlinear impedance factors that make possible frequency multiplication with overlay transistors.

capacitance $C_{bc}$. The second part consists of the part of the collector-to-base junction that is opposite the emitter-to-base junction. This part is called the inner collector capacitance $C_{b'c}$. The outer capacitance $C_{bc}$ is a much more efficient varactor than the inner capacitance $C_{b'c}$ because $C_{b'c}$ has to charge and discharge through both the intrinsic and the extrinsic base-spreading resistance $r_{bb'}$ and $r_{b'}$, as well as through the series resistance $r_{si}$, while $C_{bc}$ can charge and discharge through only $r_{b'}$ and $r_{so}$. Because the intrinsic base-spreading resistance $r_{bb'}$ is much greater than the extrinsic base-spreading resistance $r_{b'}$, the cutoff frequency $f_{VCB}$ is much larger in the active varactor portion represented by $C_{bc}$ than in the $C_{b'c}$ portions. The difference in $r_{b'}$ and $r_{bb'}$ results from the use of different sheet resistances in the two areas. Another unique feature of the overlay transistor is that the emitter area is much smaller than the base area. As a result, the inefficient portion of the varactor formed by the collector-to-base junction opposite the emitter sites is almost negligible because of the reduced emitter area.

The varactor cutoff frequency $f_{VCB}$ is also maximized by use of minimum collector series resistance $r_{so}$. This resistance is kept to a minimum by the n—n+ epitaxial structure used for the collector region. The n-type epitaxial layer forms the dominant part of the collector series resistance. The thickness of this layer is kept to the minimum value that provides the required collector-to-base breakdown voltage.

Because of the features described above, varactor loss is minimized in overlay transistors and, therefore, high conversion efficiency can be achieved. The inherent varactor frequency-multiplication ability of the collector-to-base junction capacitance, added to the excellent frequency capability of these transistors, has made possible the use of overlay devices as efficient frequency multipliers.

## Operation

The outer collector capacitance $C_{bc}$ shown in Fig. 617 varies nonlinearly with the transistor collector voltage in much the same way as the capacitance of a varactor diode varies with the voltage across the diode junction. This variable junction capacitance makes possible harmonic generation in overlay transistor circuits. The nonlinear relationship between the collector-to-base capacitance $C_{bc}$ and the collector bias voltage in overlay transistors may be expressed as follows:

$$C_{bc} = K\,(\phi - V)^{-n} \qquad (433)$$

where K is a constant determined by the area and doping of the junction, $\phi$ is the contact potential, V is the magnitude of the collector reverse-bias voltage, and the exponent n is a constant determined by the impurity distribution on both sides of the junction.

Figure 618 shows the variation in the collector-to-base capacitance

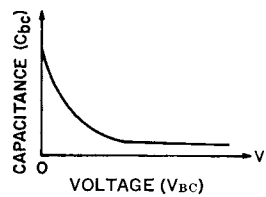



Figure 618. Collector-to-base capacitance $C_{bc}$ as a function of collector bias in overlay transistors.

$C_{bc}$ as a function of the collector bias voltage $V_{BC}$. However, this form of capacitance-voltage curve is difficult to apply directly in the analysis of high-frequency, high-power transistor circuits. Because power is the product of current and voltage swings in the transistor, the transistor current can be related to the collector-to-base capacitance if the charge Q across the junction is known. Because $dQ/dV = C(V)$, the charge Q can be determined as follows:

$$Q = \int C_{bc}\, dV \qquad (434)$$

Using the relation for capacitance $C_{bc}$ given by Eq. (433), the integration indicated in Eq. (434) can be performed with respect to the voltage V to obtain the charge. The result of this integration, shown in Fig. 619, shows the variation in the charge Q as a function of the voltage $V_{BC}$.

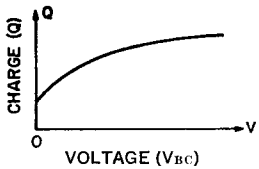



Figure 619. The charge Q in the collector-to-base junction as a function of collector-to-base voltage in an overlay transistor.

If a sinusoidal voltage such as that shown in Fig. 620(a) is developed by the amplifier section of the overlay transistor to drive the nonlinear capacitance $C_{bc}$, a highly distorted charge (or current) waveform is produced because of the nonlinear charge-voltage characteristics of the capacitance. This waveform, shown in Fig. 620(b), contains components of the fundamental frequency and of harmonic frequencies. Power output at the desired harmonic is obtained when suitable selective circuits are coupled to the collector of the transistor. In an actual circuit, the driving voltage developed by the transistor contains both fundamental-frequency and harmonic-frequency components.

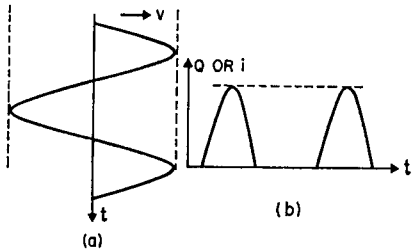



Figure 620. (a) Sinusoidal voltage developed by the amplifier section of an overlay transistor to drive the nonlinear collector-to-base capacitance; (b) distorted charge, or current, waveform produced by the nonlinear collector-to-base capacitance of an overlay transistor in the generation of harmonic power.

## Basic Transistor Frequency-Multiplier Circuits

Overlay transistors used in frequency multipliers may be connected in either common-base or common-emitter circuit configurations. In the common-base transistor frequency multiplier, harmonic generation is accomplished in essentially the same way as in a shunt-type varactor frequency multiplier because the nonlinear collector-to-base capacitance of the transistor is connected in shunt with the input circuit. In the common-emitter transistor frequency multiplier, the nonlinear capacitance is connected in series with the input; the operation of the transistor circuit is then similar to that of the series-type varactor frequency multiplier.

Fig. 621 shows the basic circuit configuration for the use of an overlay transistor in a common-

base frequency doubler. A T matching network, or other type of matching section, must be used in the input of the doubler to set up a conjugate match across the emitter-to-base terminals of the transistor at the fundamental frequency of the input driving power. This conjugate match is required to obtain a maximum transfer of

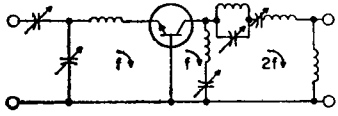



Figure 621. Basic configuration for use of an overlay transistor in a common-base frequency doubler.

power from the driving source to the transistor. Because gain at the fundamental frequency is of primary importance, an idler circuit must be connected between the collector and base of the transistor. The idler loop, which consists of a simple series LC circuit, resonates with the transistor collector-to-base capacitance at the fundamental frequency and thus enhances the flow of fundamental current through the transistor. The idler circuit also develops the driving voltage required by the nonlinear collector-to-base capacitance for the generation of harmonic power. A suitable output circuit, which is series-tuned to select output power at the second harmonic of the input frequency, completes the basic doubler circuit. In some circuits, an output trap must be added to restrict the flow of fundamental-frequency current in the output loop.

Fig. 622 shows the basic circuits for the use of an overlay transistor in the common-base frequency tripler and quadrupler, respectively. These circuits are very similar to the common-base doubler, except that an additional second-harmonic idler loop is connected in shunt with the transistor collector. The second-harmonic components produced by this idler loop beat with the fundamental-frequency components to generate additional harmonic outputs. In this way, the second-harmonic idler loop enhances the conversion efficiency. When an overlay-transistor frequency multiplier is used in a common-emitter circuit, an additional series resonant circuit must be incorporated in the input. Otherwise, the input, output, and idler circuits of common-emitter multipliers follow the considerations already described for the common-base multipliers.

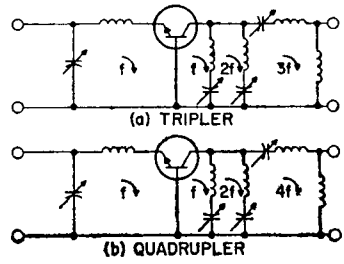



Figure 622. Basic configurations for use of an overlay transistor in a common-base frequency tripler (a) and frequency quadrupler (b).

## Design of Transistor Frequency Multipliers

The design of transistor frequency-multiplier circuits generally consists of the selection of a suitable transistor and the design of proper filtering and matching networks for optimum circuit performance.

Transistors suitable for this application must provide the desired output power and gain at

the fundamental frequency and must be able to convert the power from the fundamental frequency into power at the desired harmonic frequency. If a lossless circuit were coupled to a lossless nonlinear capacitance $C_{bc}$, power at the fundamental frequency could be converted into power at any harmonic frequency with 100-percent conversion efficiency. In practice, however, efficiency is limited by the series resistance associated with the nonlinear capacitance and the circuit losses. It can be considered that the harmonic output power of a transistor multiplier circuit, at a given input power level, is equal to the product of the power gain of the transistor at the drive frequency and the conversion efficiency that results from the varactor action of the collector-to-base capacitance $C_{bc}$. Conversion gain can be obtained only if the power gain of the transistor under consideration at the fundamental frequency is larger than the conversion loss.

In the design of such circuits, the input impedance at the fundamental frequency that exists at the emitter-to-base junction of the transistor as well as the load impedance presented to the collector at both the fundamental and harmonic frequencies must be known. Knowledge of the collector load impedance at the harmonic frequency is required for design of the output circuit. Knowledge of the collector impedance at the fundamental frequency is needed to determine the input impedance of the transistor at that frequency so that matching networks can be designed between the driving source and the transistor. The three impedances, of course, are interrelated and are functions of operating power level (i.e., are determined by voltage and current swings). Once the impedances are established, the design of the matching networks is straightforward. For the input circuit, a matching section having low-pass characteristics is preferred; for the output circuit, a matching section having high-pass or band-pass characteristics is preferred. Such arrangements assure good isolation between input and output circuits. As the frequency of operation increases above 800 MHz, the design of transistor multiplier circuits requires the use of distributed circuit techniques.

## Stability and Biasing Considerations

In general, the major problem of nonlinear devices is stability. Various types of instabilities can be incurred in transistor frequency-multiplier circuits, including hysteresis, low-frequency oscillations, parametric oscillations, and high-frequency oscillations. These difficulties can be eliminated or minimized by careful design of the bias circuit, by proper location of transistor ground connections, and by the use of packages that have minimum parasitic elements.

Hysteresis refers to discontinuous mode jumps in output power that occur when the input power or frequency is increased or decreased. This effect is caused by the dynamic detuning which results from variation in the average value of the nonlinear capacitance with rf voltage. The tuned circuit has a different resonant frequency for a strong drive input than for a weak drive input. It has been found experimentally that hysteresis effects can be mini-

mized, or sometimes eliminated, when the transistor is used in a common-emitter configuration.

Low-frequency oscillations occur because the gain of the transistor at low frequency is much higher than that at the operating frequency. This effect can be eliminated by use of a small resistance in series with the rf chokes used for the biasing circuit, as shown in Fig. 623.

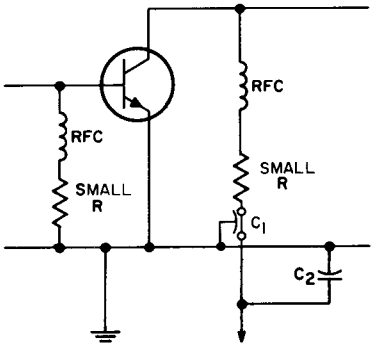



Figure 623. Circuit showing biasing techniques and bypassing capacitances used to eliminate instabilities in common-emitter frequency multipliers.

Parametric oscillations result because spurious low-frequency modulation is added to the harmonic output. This effect can be eliminated by careful selection of the bypass capacitance $C_2$ in Fig. 623 to provide a low impedance to the spurious component in addition to that provided by the rf bypass capacitance $C_1$.

High-frequency oscillation is indicated by oscillations that occur at a frequency very close to the output frequency when the input drive power is removed. With a TO-60 package transistor, common-emitter circuits are found to be less critical in this respect than common-base circuits. The high-frequency oscillations are also found to be strongly related to the input drive frequency. This type of instability can be eliminated if the input frequency is kept below certain values. The input frequency at which stable operation can be obtained seems to depend upon the method of grounding the emitter of the transistor. The highest frequency of operation can be obtained when the emitter has the shortest path to ground.

In practice, stable and reliable operation of transistors in frequency multipliers has been successfully obtained. The circuits discussed in this section are all stable frequency-multiplier circuits.

## The 2N4012 Transistor

The 2N4012 power transistor is characterized for frequency-multiplication applications and can provide a minimum power output of 2.5 watts as a frequency tripler at an output frequency of 1 GHz and a collector efficiency of 25 per cent. This overlay transistor is designed to operate in military and industrial communications equipment as a frequency multiplier in the uhf or L-band range. It can be operated as a doubler, tripler, or quadrupler to supply a power output of several watts at frequencies in the low gigahertz range.

Fig. 624 shows the power-output capabilities as a function of output frequency for a typical 2N4012 transistor used in common-emitter circuit configurations for frequency doubling, tripling, and quadrupling. In a common-emitter doubler circuit, the transistor delivers power output of 3.3 watts at 800 MHz with a conversion gain of 5 dB. In a common-emitter tripler circuit, it can

supply power output of 2.8 watts at 1 GHz with a conversion gain

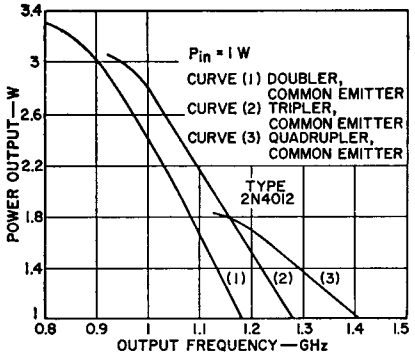



Figure 624. Power output of the RCA-2N4012 overlay transistor as a function of frequency when operated in common-emitter doubler, tripler and quadrupler circuits.

of 4.5 dB. In a common-emitter quadrupler circuit, it can provide power output of 1.7 watts at 1.2 GHz with a conversion gain of 2.3 dB.

It is of interest that the transistor frequency multipliers provide greater power outputs at higher output frequencies than the unity-gain output obtained from the transistor power amplifier at 700 MHz. When the frequency of operation is low enough so that the transistor can supply rf power with substantial gain, the output capabilities of the transistor frequency multipliers are essentially the same as those of the transistor power amplifier. For operation at the same output frequency and with the same input driving power, approximately equal amounts of power output can be obtained.

Fig. 624 shows that the amount of power output that can be supplied by a transistor frequency multiplier depends upon the order of multiplication. For a given multiplier circuit, the highest output power is obtained at the frequency for which the product of power gain and conversion efficiency has the largest value. When a 2N4012 overlay transistor is used, maximum power output is obtained at 800 MHz from a doubler circuit, at 1 GHz from a tripler circuit, and at 1.3 GHz from a quadrupler circuit.

The circuit arrangements and performance data shown in the following paragraphs illustrate several practical frequency-multiplier circuits that use the 2N4012 and other RCA overlay transistors. These circuits include a 400-to-800-MHz doubler, a 150-to-450-MHz tripler, a 367-to-1100-MHz tripler, and a 420-to-1680-MHz quadrupler. As mentioned previously, the design of multiplier circuits that have an output frequency of 800 MHz or higher requires the use of distributed-circuit techniques. All such high-frequency circuits described use coaxial-cavity output circuits. These circuits are discussed first. The low-frequency circuits, which use lumped-element output circuits, are then described.

## 400-To-800-MHz Doubler

Fig. 625 shows the complete circuit diagram of a 400-to-800-MHz doubler that uses the 2N4012 transistor. This circuit uses lumped-element input and idler circuits and a coaxial-cavity output circuit. The transistor is placed inside the cavity with its emitter properly grounded to the chassis. A pi section ($C_1$, $C_2$, $L_1$, $L_2$, and $C_3$) is used in the input to match the impedances, at 400 MHz, of the driving source and the base-emitter junction of the transistor. $L_2$ and $C_3$ provide the necessary

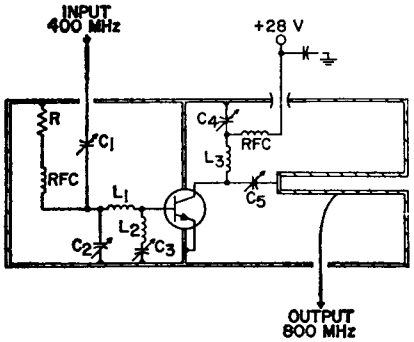



Figure 625. 400-to-800-MHz common-emitter transistor frequency multiplier.

ground return for the nonlinear capacitance of the transistor. $L_3$ and $C_4$ form the idler loop for the collector at 400 MHz. The output circuit consists of an open-ended 1¼-inch-square coaxial cavity. A lumped capacitance $C_5$ is added in series with a ¼-inch hollow-center conductor of the cavity near the open end to provide adjustment for the electrical length. Power output at 800 MHz is obtained by direct coupling from a point near the shorted end of the cavity. The bias arrangement is the same as that used in the circuit shown in Fig. 623.

Fig. 626 shows the power output at 800 MHz as a function of the power input at 400 MHz for the doubler circuit, which uses a typical 2N4012 operated at a collector supply voltage of 28 volts. The curve is nearly linear at a power output level between 0.9 and 2.7 watts. The power output is 3.3 watts at 800 MHz for an input drive of 1 watt at 400 MHz, and rises to 3.9 watts as the input drive increases to 1.7 watts. The collector efficiency, which is defined as the ratio of the rf power output to the dc power input at a supply voltage of 28 volts, is also shown in Fig. 626. The efficiency is 43 per cent measured at an input power of 1 watt. The 3-dB bandwidth of this circuit measured at a power output of 3.3 watts is 2.5 per cent. The fundamental-frequency component measured at a power-output level of 3.3 watts is 22 dB down from the output carrier. Higher attenuations of spurious components can be achieved if more filtering sections are used.

The variation of power output with collector supply voltage at an input drive level of 1 watt is

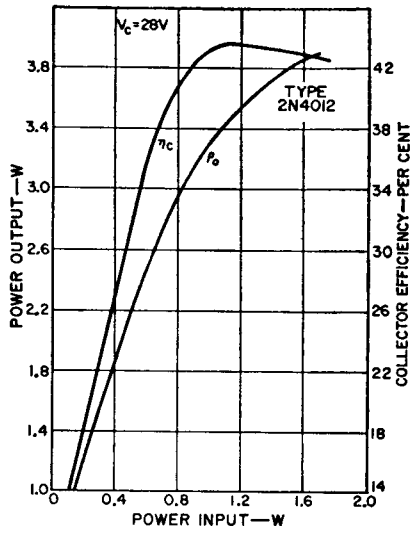



Figure 626. Output power and collector efficiency as a function of input power for the 400-to-800-MHz frequency doubler.

shown in Fig. 627. This curve is obtained with the circuit tuned at 28 volts. The curves of Figs. 626 and 627 indicate that the transistor amplifier-multiplier circuit is capable of amplitude modulation.

## 367-To-1100-MHz Tripler

The 367-to-1100-MHz tripler shown in Fig. 628 is essentially

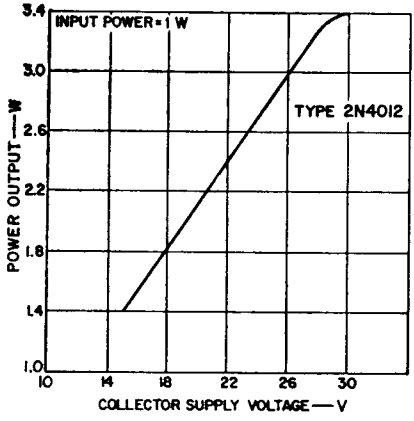



Figure 627. Power output as a function of supply voltage for the 400-to-800-MHz frequency doubler.

the same as the doubler shown in Fig. 625 except that an additional idler loop ($L_4$ and $C_6$) is added in shunt with the collector of the transistor. This idler loop is resonant with the transistor junction capacitance at the second harmonic frequency (734 MHz) of the input drive.

Fig. 629 shows the power output of the tripler at 1.1 GHz as a function of the power input at 367 MHz. This circuit also uses a typical 2N4012 transistor operated at a collector supply voltage of 28 volts. The solid-line curve shows the power output obtained when the circuit is retuned at each power-input level. The dashed-line curve shows the power output obtained with the circuit tuned at the 2.9-watt output level. A power output of 2.9 watts at 1.1 GHz is obtained with drive of 1 watt at 367 MHz. The 3-dB bandwidth measured at this power level is

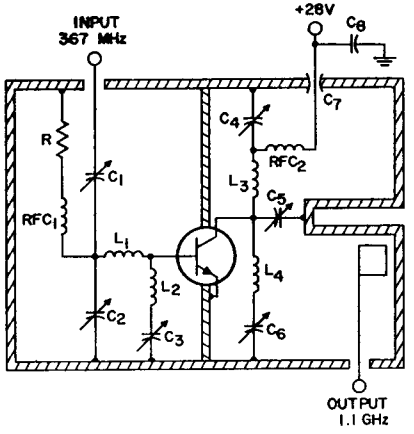



Figure 628. 367-MHz-to-1.1-GHz common-emitter transistor frequency tripler.

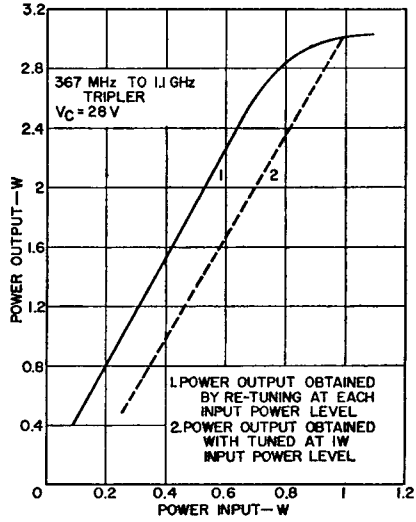



Figure 629. Power output as a function of power input for the 367-MHz-to-1.1-GHz frequency tripler.

2.3 per cent. The spurious-frequency components measured at the output are as follows: −22 dB at 340 MHz, −30 dB at 680 MHz, −35 dB at 1360 MHz.

The variation of power output with collector supply voltage at an input drive level of 1 watt is shown in Fig. 630. The variation of collector efficiency is also shown. These curves were obtained with the circuit tuned at 28 volts.

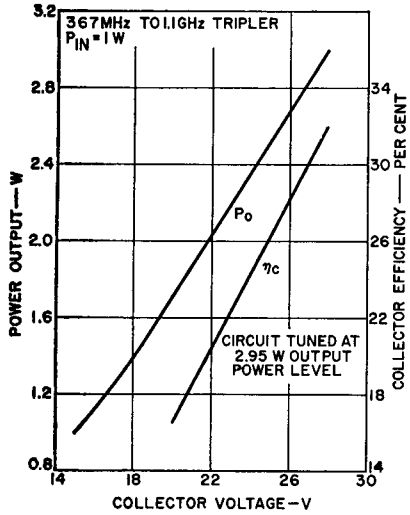



Figure 630. Power output as a function of collector supply voltage for the 367-MHz-to-1.1-GHz frequency tripler.

A 367-MHz amplifier that uses the same circuit configuration and components as those of the tripler circuit shown in Fig. 628 was constructed to compare the performance between amplifier and tripler. The conversion efficiency for a large number of tripler units was then measured. The conversion efficiency of the tripler is defined as the 1.1-GHz power obtained from the tripler divided by the 367-MHz power obtained from the amplifier at the same power-input level (1 watt). The efficiency varies between 60 and 75 per cent, and has an average value of 65 per cent; this performance is comparable to that of a good varactor multiplier in this frequency range.

A similar tripler circuit that uses a selected 2N3866 and that is operated from 500 MHz to 1.5 GHz can deliver a power output of 0.5 watt at 1.5 GHz with an input drive of 0.25 watt at 500 MHz.

## 150-To-450-MHz Tripler Circuit

Fig. 631 illustrates the use of the 2N4012 transistor in a 150-to-450-MHz frequency tripler. The input coupling network is designed to match the driving generator to the base-to-emitter circuit of the transistor. The network formed by $C_2$ and $L_2$ provides a ground return for harmonic output current at 450 MHz. The idler network in the collector circuit ($L_3$, $L_4$, and $C_4$) is designed to circulate fundamental and second-harmonic components of current through the voltage-variable collector-to-base capacitance, $C_{bc}$. The network

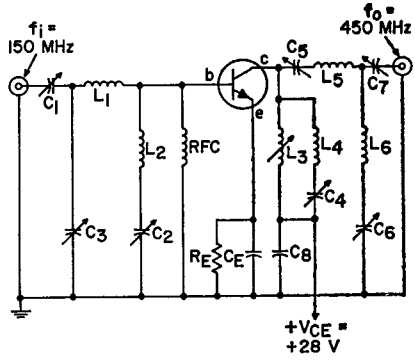



Figure 631. 150-to-450-MHz common-emitter transistor frequency tripler.

formed by $C_5$, $C_6$, $C_7$, $L_5$, and $L_6$ provides the required collector loading for 450-MHz power output. Fig. 632 shows the 450-MHz power output of the tripler as a function of the 150-MHz power input. For driving power of one watt, power output of 2.8 watts is obtained at 450 MHz. The rejection of fundamental, second, and fourth harmonics was measured as 30 dB below the 2.8-watt, 450-MHz level. The variation of power output with supply voltage is shown in Fig. 633.

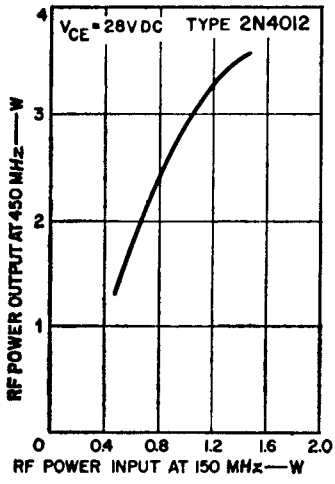



Figure 632. Power output as a function of power input for the 150-to-450-MHz frequency tripler.

## Common-Emitter and Common-Base Circuits

The performance data in this section are given for amplifier-multipliers in which the transistor is connected in a common-emitter configuration. When transistors are used in common-base circuit configurations, different results are obtained. Fig. 634 shows curves of power output and efficiency for a common-base and a common-emitter tripler circuit using a 2N4012 transistor. At low power levels, the common-base tripler provides higher gain and collector efficiency; at high power levels, higher gain and collector efficiency are provided by the common-emitter circuit. At a power input of 1 watt at 367 MHz, the common-emitter tripler delivers a power output of 2.9 watts at 1.1 GHz and the common-base circuit an output of 2.4 watts. The collector efficiencies for both circuits are approximately the same and are better than 30 per cent. The 3-dB bandwidth measured in the common-emitter tripler is 2.3 per cent, as compared to 2.5 per cent in a common-base tripler. The major difference between the two circuits is that the power output of the common-emitter tripler saturates at a much higher power-input level than that of the common-base circuit. This effect has also been observed in a straight-through amplifier. In addition, the common-emitter circuit is less sensitive to hysteresis and high-frequency oscillations, as discussed previously.

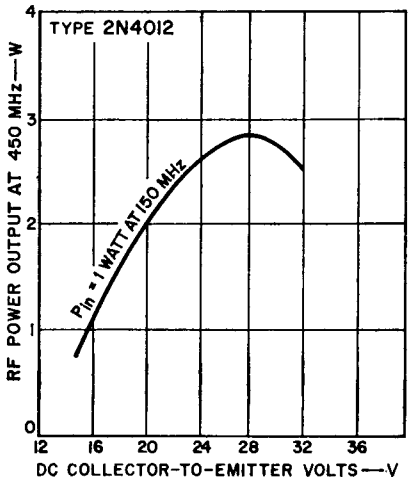



Figure 633. Power output as a function of collector supply voltage for the 150-to-450-MHz frequency tripler.

## A 420-MHz-To-1.68-GHz Oscillator-Quadrupler

The inherent varactor frequency-multiplication ability in overlay transistors also permits use of these devices as oscillator-multipliers. Fig. 635 shows an oscillator-quadrupler circuit that uses a selected 2N3866 transistor. This circuit can deliver a power output

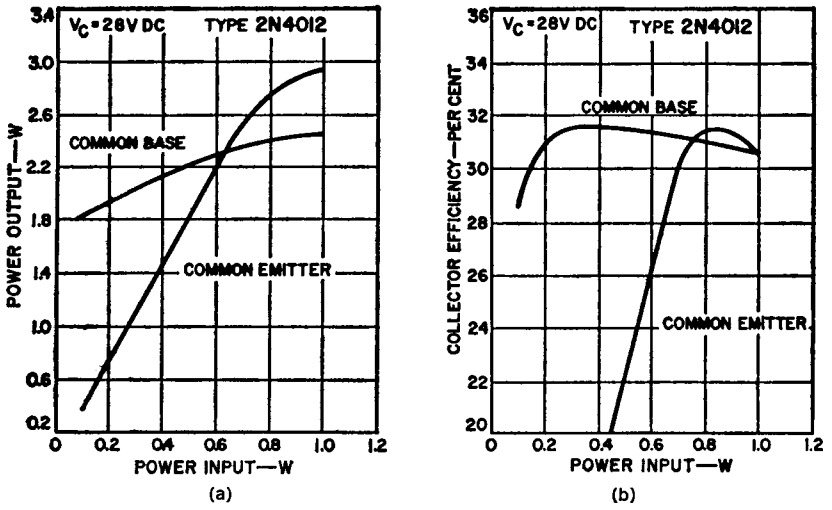



Figure 634. Comparison of performance characteristics of common-base and common-emitter tripler circuits using the RCA-2N4012 transistor: (a) power output as a function of power input; (b) collector efficiency as a function of power input.

of more than 300 milliwatts at 1.68 GHz. The first two rf chokes and the resistors $R_1$ and $R_2$ form the bias circuit. The fundamental frequency of the oscillator is 420 MHz, as determined by $C_o$, $L_1$, and $C_1$. $L_2$ and $C_2$ form the second-harmonic idler. The second-harmonic component produced by this idler circuit beats with the fundamental-frequency component to generate additional fourth-harmonic components. A series-tuned circuit consisting of $L_3$ and $C_3$ completes the output circuit.

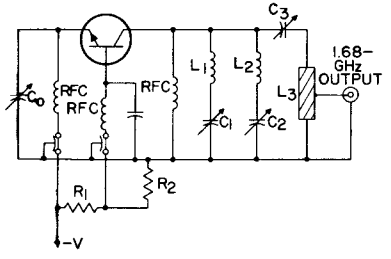



Figure 635. Oscillator-quadrupler circuit.

## SYSTEM APPLICATIONS

The major applications of microwave transistors are in the areas of relay links and air-traffic-control systems.

### Microwave Relay Links

Microwave point-to-point communication links are rapidly replacing wire-line and coaxial-cable networks. Microwave installations have inherently greater signal-handling capability and are less expensive than cable networks. The requirements for commercial relay-link systems are not as stringent as those for the military-oriented systems. However, high efficiency and high reliability are prime requirements because many of these installations are remote and unattended. The links may be used for voice, video, or data transmissions, using AM, FM or various digital modulation methods. The fre-

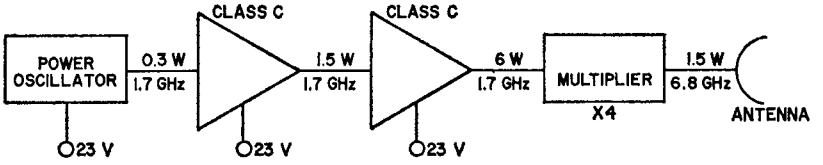
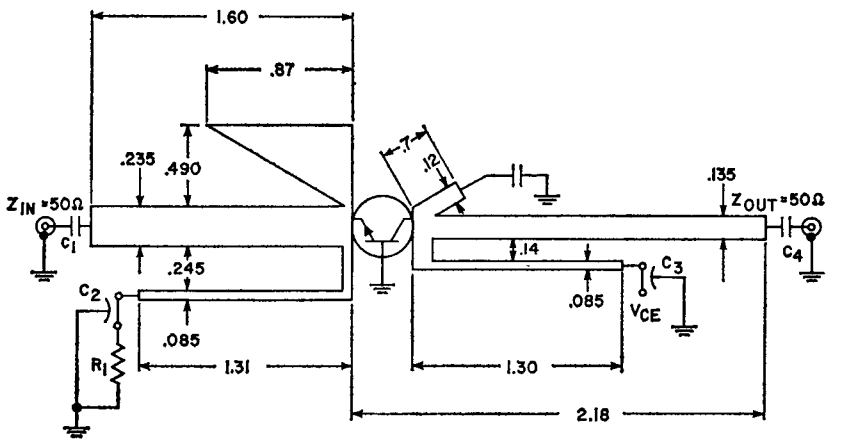



Figure 636. Microwave portion of a typical data link system.

quency spectrum used in these systems ranges from UHF to $K_u$-band, with both narrow-band and wide-band requirements. For example, some military telemetry systems operate near 1.5 GHz and 2.25 GHz, while common-carriers and other data link systems operate at C-band and X-band.

Microwave power transistors are finding extensive use in these relay link systems as signal sources, buffer-amplifier stages, and power-amplifier stages which drive either an antenna or a high-frequency multiplier stage. A typical rf power chain for a data-link system is shown in Fig. 636.

The driver and power-amplifier stages are essentially similar to those already considered in the previous sections. The passive multiplier stage will not be considered here. Because of the close spacing of channels in data-link systems, the AM and FM noise characteristics of the oscillator stage are of prime importance in these systems. These noise requirements are best met with high-Q cavity oscillators, using transistors such as the RCA 40836 or RCA 40837.




$C_1$, $C_4$ = 1000 pF, ATC-100 or equivalent
$C_2$, $C_3$ = Filtercon, low-pass filter, Allen-Bradley SMFB-A1 or equivalent

$R_1$ = 0.24 ohms, BMH, wirewound
Dielectric material = 1/32-inch-thick Teflon fiberglass double-clad circuit board ($\varepsilon$ = 2.6)

Figure 637. 2-watt, 960-to-1215-MHz microstrip amplifier.

## Air-Traffic-Control Systems

Most commercial and military air traffic control (ATC) systems operate in the frequency range from 960 to 1215 MHz. TACAN (Tactical Air Navigation System) ground stations operate from 960 to 1025 MHz, while airborne DME (Distance Measuring Equipment), TACAN, and transponders operate from 1025 to 1150 MHz. Transistors have great potential usage as pulsed amplifiers or oscillators at 1090 MHz for ATC or IFF (Identification, Friend or Foe) transponders, where peak power outputs of several hundred watts are required with short-pulse, low-duty operation. Microwave transistors are also used extensively for tube driver operation under broadband conditions. Fig. 637 shows a microstrip circuit which provides 2 watts output from 960 to 1215 MHz. Fig. 638 shows the broadband performance of this circuit.

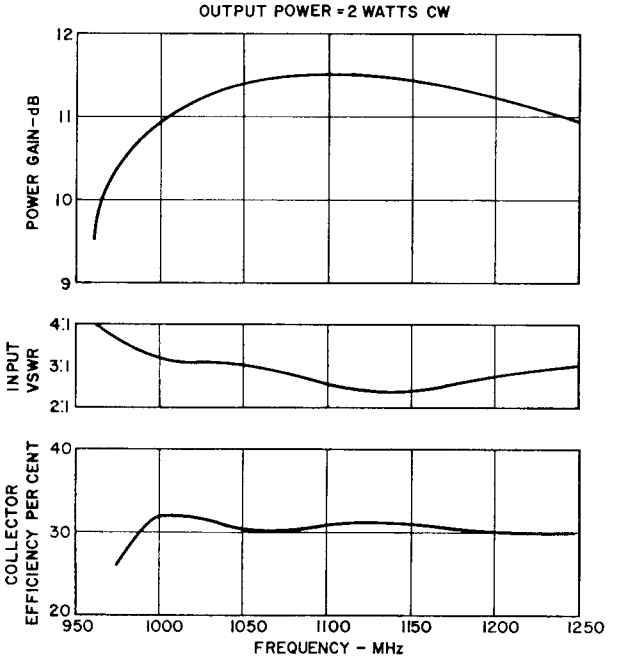



Figure 638. Performance characteristics of amplifier circuit shown in Fig. 637.

# Audio Power Amplifiers

THE quality of an audio power amplifier is measured by its ability to provide high-fidelity reproduction of audio program material over the full range of audible frequencies. The amplifier is required to increase the power level of the input to a satisfactory output level with little distortion, and the sensitivity of its response to the input signals must remain essentially constant throughout the audio-frequency spectrum. Moreover, the input-impedance characteristics of the amplifier must be such that the unit does not load excessively and thus adversely affect the characteristics of the input-signal source.

Silicon power transistors offer many advantages when used in the power-output and driver stages of high-power audio amplifiers. These devices may be used, either as discrete components or as building-block elements in power hybrid circuits, over a wide range of ambient temperatures to develop tens of watts of audio-frequency power to drive a loudspeaker system. The following paragraphs describe the basic factors that must be considered and the important concepts and techniques employed in the design of transistor audio power amplifiers.

## CLASSES OF OPERATION

A circuit designer may select any one of three classes of operation for transistors used in linear-amplifier applications. This selection is made on the basis of a combination of such factors as required power output, dissipation capability, efficiency, gain, and distortion characteristics.

The three basic classes of operation (class A, class B, and class C) for linear transistor amplifiers are defined by the operating point of the transistor. In class A operation, the active element conducts for the entire input cycle. In class B operation, the active element conducts for 180 degrees of an input cycle and is cut off during the remainder of the time. In class C operation, the active element conducts for some amount less than 180 degrees of an input cycle. The following paragraphs discuss the distinguishing features of class A and class B operation. In general, because of the

high harmonic distortion introduced as a result of the short conduction angle, class C operation is used primarily in rf-amplifier applications in which it is practical to use tuned output circuits to eliminate the harmonic components. For this reason, class C operation is not discussed further.

## Class A Operation

Class A amplifiers are used for linear service at low power levels. When power amplifiers are used in this class of operation, the amplifier output is usually transformer-coupled to the load circuit, as shown in Fig. 639. At low power levels, the class A amplifier can also be coupled to the load by resistor, capacitor, or direct coupling techniques.

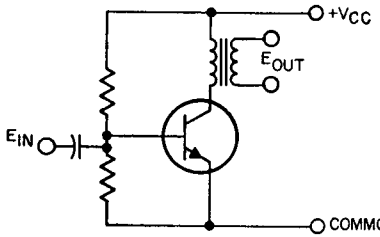



Figure 639. Basic class A, transformer-coupled amplifier.

There is some distortion in a class A stage because of the nonlinearity of the active device and circuit components. The maximum efficiency of a class A amplifier is 50 per cent; in practice, however, this efficiency is not realized. The class A transistor amplifier is usually biased so that the quiescent collector current is midway between the maximum and minimum values of the output-current swing. Collector current, therefore, flows at all times and imposes a constant drain on the power supply. The consistent drain is a distinct disadvantage when higher power levels are required or operation from a battery is desired.

## Class B Operation

Class B power amplifiers are usually used in pairs in a push-pull circuit because conduction is not maintained over the complete cycle. A circuit of this type is shown in Fig. 640. If con-

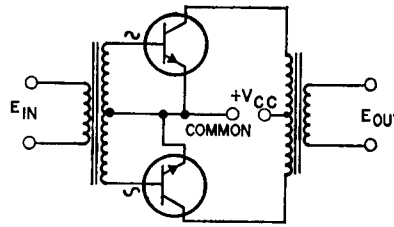



Figure 640. Basic class B, push-pull transformer-coupled amplifier.

duction in each device occurs during approximately 180 degrees of a cycle and the driving wave is split in phase, the class B stage can be used as a linear power amplifier. The maximum efficiency of the class B stage at full power output is 78.5 per cent when two transistors are used. In a class B amplifier, the maximum power dissipation is 0.203 times the maximum power output and occurs at 42 per cent of the maximum output.

Transistors are not usually used in true class B operation because of an inherent nonlinearity, called **cross-over distortion**, that produces a high degree of distortion at low power levels. The distortion results from the nonlinearities in the transistor characteristics at very low current levels. For this reason, most power stages operate in a biased condition somewhat between class A and class B.

This intermediate class is defined as **class AB**. Class AB transistor amplifiers operate with a small forward bias on the transistor to minimize the nonlinearity. The quiescent current level, however, is still low enough so that class AB amplifiers provide good efficiency. This advantage makes class AB amplifiers an almost universal choice for high-power linear amplification, especially in battery-operated equipment.

## DRIVE REQUIREMENTS

In class A amplifiers, the output stage is usually connected in a common-emitter configuration. The relatively low input impedance that generally characterizes this type of configuration may result in a severe mismatch with the output impedance of the driver transistor. Usually, at low power levels, RC coupling is used and the loss is accepted. It may be advantageous in some circuits, however, to use an emitter-follower between the driver and the output stage to obtain an improved impedance match.

Class AB amplifiers have many types of output connections. One form is the **transformer-coupled output stage** illustrated in Fig. 641. Again, the common-emitter circuit is usually employed because it provides the highest power gain. The load circuit is never matched to the output impedance of the transistor, but rather is fixed by the available voltage swing and the required power output. The transformer is designed to reflect the proper impedance to the output transistors so that the desired power output can be achieved with a specific supply voltage.

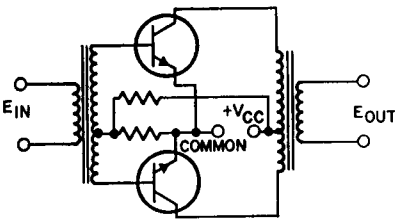



Figure 641. Basic class AB, push-pull, transformer-coupled amplifier.

The use of transformer coupling from the driver to the input of the power transistor assures that the phase split required for push-pull operation of the output stages and any necessary impedance transformation can be readily achieved. Output transformer coupling provides an easy method for matching several values of load impedance, including those encountered in sound-distribution systems. For paging service, servo motor drive, or other applications requiring a limited bandwidth, the transformer-coupled output stage is very useful. However, there are disadvantages to the use of transformer coupling. One disadvantage is the phase shift encountered at low- and high-frequency extremes, which may lead to unstable operation. In addition, the output transistors must be capable of handling twice the supply voltage because of the transformer requirements.

Another type of transistor output circuit is the **series-connected output stage.** With this type of circuit, the transistors are connected in series across the supply and the load circuit is coupled to the midpoint through a capacitor. There must be a 180-degree phase shift between the driving signals for the upper and lower transistors. A transformer can be used in this application provided that the secondary consists of two separate windings, as shown in Fig.

642. Other forms of phase splitting can be used; all have problems such as insufficient swing or poor impedance matching. Capacitor output coupling also has disadvantages. A low-frequency phase shift is usually associated with the capacitor, and it is difficult to obtain a capacitor that is large enough to produce an acceptable

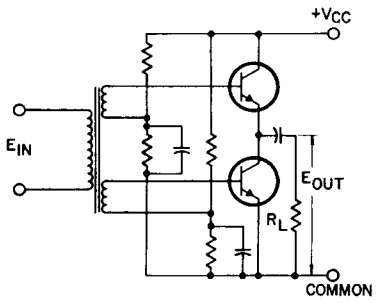



Figure 642. Class AB, push-pull amplifier with series output connection.

low-frequency output. These disadvantages can be alleviated by use of a split supply and by connection of the load between the transistor midpoint and the supply midpoint with the return path through the power-supply capacitors. The power-supply capacitors must be large enough to prevent excessive ripple.

**Complementary amplifiers** are produced when p-n-p and n-p-n transistors are used in series. A capacitor can be used to couple the amplifier output when a single supply is used, or direct coupling can be employed when a split power supply is used, as shown in Fig. 643. Because no phase inversion is needed in the driving circuit for this output configuration, there are definite advantages in the simplicity of the design. One disadvantage of this type of amplifier is that the driver must be a class A stage which may have a high dissipation. This dissipation can be reduced, however, by use of a Darlington compound connection for the output stage. This compound connection reduces the driving-stage requirement. A method of overcoming this disadvantage completely is to use a quasi-complementary configuration. In this configuration, the output transistors are a pair of p-n-p or n-p-n transistors driven by a complementary pair in the driver. In this manner the n-p-n/p-n-p drivers provide the necessary phase inversion. The availability of both n-p-n and p-n-p silicon driving transistors that have the same electrical characteristics is good.

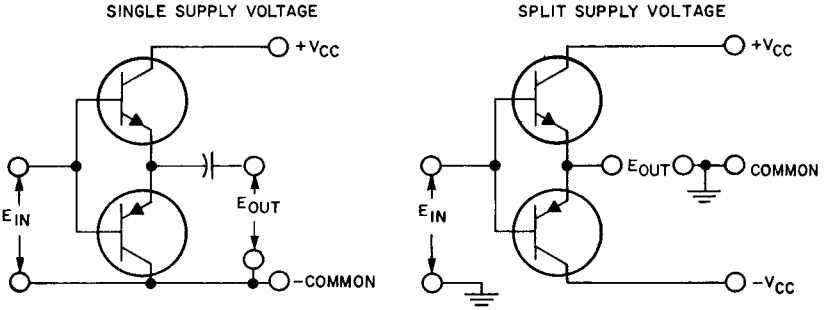



Figure 643. Circuit arrangements for operation of complementary output stages (a) from single dc supply; (b) from symmetrical dual (positive and negative) supplies.

The driving transistors are connected directly to the bases of the output transistors, as illustrated in Fig. 644.

Adequate drive may be a problem with the transistor pair shown in the upper part of the quasi-complementary amplifier unless suitable techniques are used to assure that this pair saturates. Care must also be taken when split supplies are used to assure that any ripple on the lower supply is not introduced into the predriving stages by this technique. The advantage of a split supply is that it makes possible direct connection to the load and thus improves low-frequency response.

be required to supply the drive needed. There are several complex phase-splitting circuits; a few of them are shown in Fig. 645.

## EFFECT OF OPERATING CONDITIONS ON CIRCUIT DESIGN

Some additional design problems involve the consideration of thermal stability, high line voltage, line-voltage transients, excessive drive, ambient temperature, load impedance, and other factors that may subject the transistors to abnormal high-stress conditions. A prime consideration is the

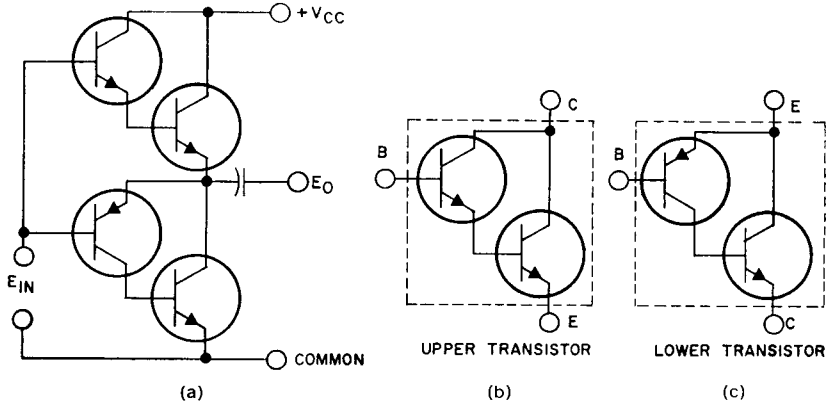



Figure 644. Compound output stage in which output transistors are driven by complementary driver transistors: (a) over-all circuit; (b) upper transistor pair; (c) lower transistor pair.

To this point, phase inversion has been mentioned but not discussed. Phase inversion may be accomplished in many ways. The simplest electronic phase inverter is the single-stage configuration. This configuration can be used at low power levels or with high-gain devices when the limited drive capability is not a drawback. At higher power levels, some impedance transformation and gain may

maximum power dissipation at high supply voltage. Thermal stability is another problem that is often difficult to control. The problem is complex because the base-to-emitter voltage $V_{BE}$ of a transistor decreases with an increase in junction temperature at a constant level of collector current. Therefore, if the $V_{BE}$ of the transistor is held constant, the collector current $I_C$ increases as the

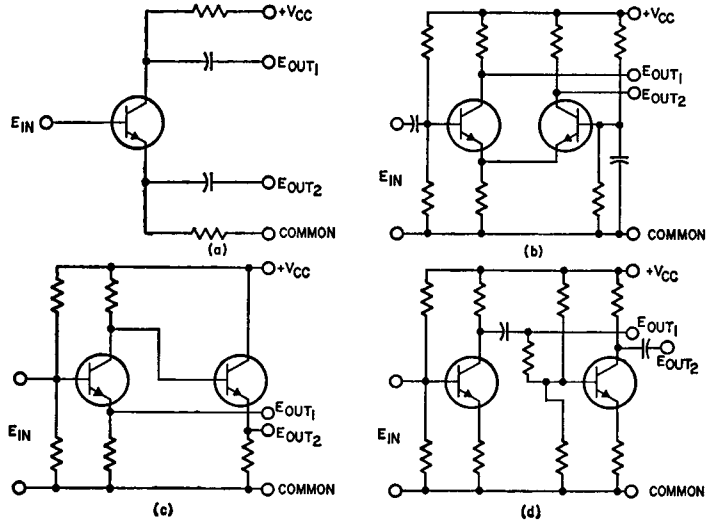



Figure 645. Basic phase-inverter circuits: (a) single-stage phase-splitter type; (b) two-stage emitter-coupled type; (c) two-stage low-impedance type; (d) two-stage similar-amplifier type.

junction temperature rises. This process is regenerative because the dissipation increases with an increase in the value of $I_C$. One solution is to place a resistor in series with the emitter lead. This approach is not the best solution to the problem, however, because the use of the resistor increases circuit losses. A decrease in the loss may be obtained if the resistor is bypassed. Another approach is to use a thermistor or similar device which, when properly connected, reduces the base drive at high temperatures. This approach improves the stability without increasing the circuit loss.

The collector-to-base leakage current $I_{CBO}$ can also be a problem because a fraction of this current is multiplied by the transistor $h_{fe}$ and appears as a component of the collector-to-emitter current. In general, the value of $I_{CBO}$ is in the order of microamperes in silicon devices and milliamperes in germanium devices. This leakage current is composed of two components. One component is caused by surface leakage and is unpredictable in its variations with temperature. It increases with voltage and may even decrease with increasing temperature. The other component is a function of the device material and geometry. This component approximately doubles with every 7°C temperature rise in silicon devices, and approximately doubles for every 10°C temperature increase in germanium devices. This component may also be voltage-dependent.

The total leakage is of interest to the circuit designer because it can be the mechanism for thermal-runaway problems. An increase in this leakage increases the total base current and thus causes an increase in collector current and dissipation. The increase in collector current and dissipation causes a rise in temperature which possibly may produce a regenerative cycle that leads to thermal

runaway. If an external resistor is connected between the base and emitter, some of this leakage current is shunted from the base, and the thermal-stability problem is reduced.

Another potential source of trouble in amplifiers is the feedback loop. Feedback is used to reduce distortion and extend the frequency range of the amplifier. The feedback loop usually encloses several if not all of the amplifier stages and can cause several problems. When transformer coupling is used, phase shifts may occur at the high- or low-frequency extremes; a positive voltage may then be fed back and cause oscillation. High-signal-level transients may cause the value of the transformer inductances and other components to change and become unstable so that they initiate oscillation. A similar condition can occur at low frequencies when capacitor-coupled transformerless designs are used.

Excessive drive levels at high frequencies can cause dissipation problems. An excessive drive level forces the output stages to saturate before the peak of the input signal is reached. This additional drive lengthens the storage time which, at high frequencies, may approach the period of the drive signal. Under this condition, two results occur: First, feedback does not increase after the point where the output stage saturates. This condition permits the drive signal to increase. Second, one transistor may not turn off until the second has been turned on. In series-type output stages, the second transistor is turned on with the full supply voltage present. This condition can lead to forward-bias second-breakdown problems. In germanium units, the excessive dissipation caused by excessive drive levels at high frequencies also contributes to the thermal-runaway problem.

Another potential source of difficulty with amplifiers occurs when the output is open- or short-circuited. Transformer-coupled output stages are particularly susceptible to operational problems with no load. Without a load, the transistors operate into a purely inductive load line and the probability of reverse-bias second breakdown must be considered. In series-type output stages, the major problem arises under short-circuit load conditions. As a result of the short circuit, feedback is removed and an open-loop gain condition exists together with the excessive-drive-condition problems previously mentioned. It is advisable to use some form of fast-acting overload protection for the power transistor; a fuse is usually not fast enough in this application.

Some frequency exists at which the gain of any transistor begins to decrease. This decrease in gain can be corrected over the required frequency range by use of feedback or a higher-frequency device. Roll-off of the frequency response of the preamplifier stages at some point prior to the limiting value of the frequency characteristics of the transistor is necessary. This technique assures that the drive is limited to a safe value by the input stage so that even the drivers are not affected by the high dissipation mentioned previously.

Several other factors that should be considered in the design of amplifiers for audio-frequency service include the frequency response desired, gain, optimum load, noise, and power output needed.

## BASIC CIRCUIT CONFIGURATIONS

The selection of the basic circuit configuration for an audio power amplifier is dictated by the particular requirements of the intended application. The selection of the basic circuit configuration that provides the desired performance most efficiently and economically is based primarily upon the following factors: power output to be supplied, required sensitivity and frequency-response characteristics, maximum allowable distortion, and capabilities of available devices.

### Class A Transformer-Coupled Amplifiers

Fig. 646 shows a three-stage class A transformer-coupled audio amplifier that uses dc feedback (coupled by $R_1$, $R_2$, $R_3$, $R_4$, and $C_1$) from the emitter of the output transistor to the base of the input transistor to obtain a stable operating point. An output capability of 5 watts with

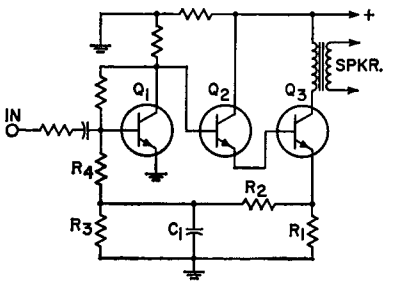



Figure 646. Three-stage transformer-coupled, class A amplifier.

a total harmonic distortion of 3 per cent is typical for this type of circuit. In general, this output level is the upper limit for class A amplifiers because the power dissipated by the output transistor in such circuits is more than twice the output power. For this reason, it is economically impractical to use class A audio amplifiers to develop higher levels of output power. A circuit such as the one shown in Fig. 646 usually requires no over-all feedback unless extremely low distortion is required. Local feedback in each stage is adequate; amplifiers of this type, therefore, are usually very stable.

### Class AB Push-Pull Transformer-Coupled Amplifiers

At power-output levels above 5 watts, the operating efficiency of the circuit becomes an important factor in the design of audio power amplifiers. The circuit designer may then consider a class AB push-pull amplifier for use as the audio-output stage.

Fig. 647 shows a class AB push-pull transformer-coupled audio-output stage. Resistors $R_1$, $R_2$, and $R_3$ form a voltage divider that provides the small amount of transistor forward bias required for class AB operation. The transformer type of output coupling used in the circuit is advantageous in that a suitable output transformer can be selected to match the audio system to any desired load impedance. This feature assures maximum transfer of the audio-output power to the load circuit, which is especially important in sound-distribution systems that use high-impedance transmission lines to reduce losses. A major disadvantage of transformer output coupling is that it tends to limit the amplifier frequency response, particularly at the low-frequency end. Variations in transformer impedance with fre-

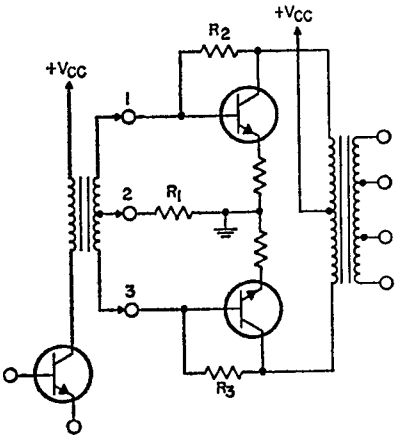



Figure 647. Class AB, push-pull, transformer-coupled audio output stage.

quency may produce significant phase shifts in the signal at both frequency extremes of the amplifier response. Such phase shifts are potential causes of amplifier instability if they occur within the feedback loop. Open-circuit stability is always a problem in designs that use output transformers because the gain increases sharply when the load is removed. If too much over-all feedback is employed, the amplifier may oscillate. The local feedback caused by the bias arrangement of $R_2$ and $R_3$ helps to eliminate this problem.

Push-pull output stages, which use identical output transistors, require some form of phase inversion in the driver stage. In the circuit shown in Fig. 647, a center-tapped driver transformer is used for this purpose. The requirements of this transformer depend upon the power levels involved, the bandwidth required, and the distortion that can be tolerated. This transformer also introduces phase-shift problems that tend to cause instabilities in the circuit when high levels of feedback are employed. Phase-shift problems are substantially reduced when the output stage is designed to operate at low drive requirements. The reduced drive requirements can be achieved by use of the Darlington circuit shown in Fig. 648. Resistors $R_1$ and $R_2$ shunt the leakage of the driver and also permit the output transistors to turn off more rapidly. Impedance levels between the class A driver and the output stage can be easily matched by the use of an appropriate transformer turns ratio.

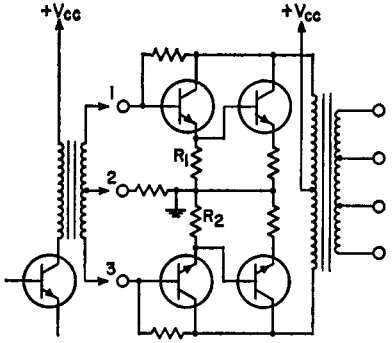



Figure 648. Class AB, push-pull, transformer-coupled audio output stage in which Darlington pairs are used to reduce drive requirements of output transistors.

An alternative method of phase inversion is to use a transistor in a phase-splitter circuit, such as those shown in Fig. 645. Unlike the center-tapped transformer method, impedance matching may be a problem because the collector of the driver, which has a relatively high impedance, operates into the low input impedance of the output stage. One solution is to reduce the output impedance of the driver stage by the use of smaller resistors. The resultant increase in collector current, however, also increases the dissipation. Moreover, very large coupling capacitors are necessary for

the achievement of good low-frequency performance. The nonlinear impedance exhibited by the input of the output transistor causes a dc voltage to be produced across the capacitor under high signal levels. An alternate solution is to use a Darlington pair to increase the input impedance of the output stage.

## Class AB Series-Output Amplifiers

For applications in which low distortion and wide frequency response are major requirements, a transformerless approach is usually employed in the design of audio power amplifiers. With this approach, the common type of circuit configuration used is the series-output amplifier.

**Push-Pull Driven Circuits**—The class-AB-operated n-p-n transistors used in the series-output circuits shown in Fig. 649 require some form of phase inversion of the drive signal for push-pull operation. A common approach is to use a driver transformer that has split secondary windings, as shown in Fig. 650. The split secondary windings are required because of the mode in which each of the series output transistors operates.

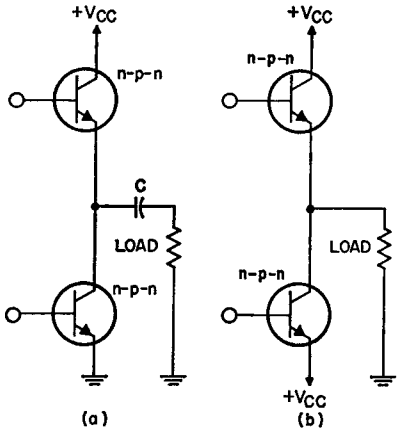



Figure 649. Circuit arrangements for operation of series output circuit from (a) a single dc supply and (b) symmetrical dual supplies.

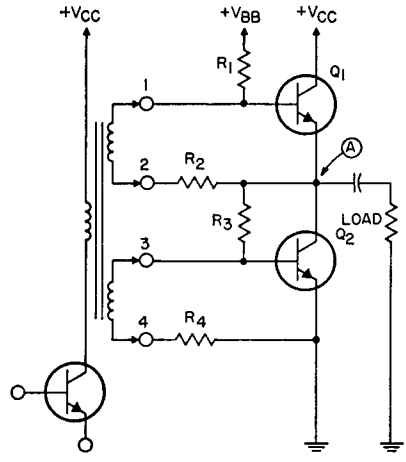



Figure 650. Circuit using a driver transformer that has split secondary windings to provide phase inversion for push-pull operation of a series-output circuit.

If ground were used as the drive reference for both secondary windings of the circuit shown in Fig. 650, transistor $Q_1$ would operate as an emitter-follower and would provide gain of somewhat less than unity. Transistor $Q_2$, however, is connected in a common-emitter configuration which can provide substantial voltage gain. For equal output-voltage swings in both directions, the drive input to transistor $Q_1$ is applied directly across the base and emitter terminals. Transistor $Q_1$ is then effectively operated in a common-emitter configuration (although there is no phase reversal from input to output) and has a voltage gain equal to that of transistor $Q_2$.

The disadvantages of a driver

transformer discussed previously also apply to the circuit shown in Fig. 650. In addition, coupling through interwinding capacitances can adversely affect the performance of the circuit. Such coupling is particularly serious because at both ends of the upper secondary (terminals 1 and 2) the ac voltage with respect to ground is approximately equal to the output voltage. During signal conditions, when output transistor $Q_1$ is turned on, this coupling provides an unwanted drive to $Q_1$. The forward transistor bias required to maintain class AB circuit operation is provided by the resistive voltage divider $R_1$, $R_2$, $R_3$, and $R_4$. These resistors also assure that the output point between the two transistors (point A) is maintained at one-half the dc supply voltage $V_{CC}$.

As in the case of the transformer-coupled output, phase inversion can be accomplished by use of an additional transistor. Fig. 651 shows a circuit in which the transistor phase inverter is used, together with a Darlington output stage to minimize loading on the phase inverter. It should be noted that capacitor C provides a drive reference back to the emitter of the upper output transistor. In effect, this arrangement duplicates the drive conditions of the split-winding transformer approach. A disadvantage of this circuit is the high quiescent dissipation of the phase inverter $Q_1$ which is necessary to obtain adequate drive at full power output. An unbypassed emitter resistor R is necessary because a signal is derived from this point to drive the lower output transistor. When transistor $Q_1$ is driven into saturation, the minimum collector-to-ground voltage that can be obtained is limited primarily by the peak emitter voltage under these conditions. To obtain the necessary voltage swing at this collector (a voltage swing that is also approximately equal to the output voltage swing), it is necessary to use a quiescent collector-to-emitter voltage higher than that required in a stage that uses a bypassed emitter resistor.

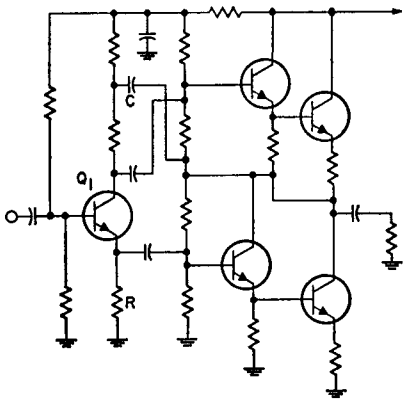



Figure 651. Push-pull series-output amplifier in which driver and output transistors are connected as Darlington pairs and drive-signal phase inversion is provided by phase-splitter stage $Q_1$.

**Complementary Amplifiers** — When a complementary pair of output transistors (n-p-n and p-n-p) is used, it is possible to design a series-output type of audio power amplifier which does not require push-pull drive. Because phase inversion is unnecessary with this type of configuration, the drive circuit for the amplifier is simplified substantially. Fig. 652 shows a basic complementary type of series-output circuit together with a simple class A driver stage. The voltage drop across resistor R provides the small amount of forward bias required for class AB operation of the complementary pair of output transistors.

In practice, a diode is employed in place of resistor R. The purpose of the diode is to maintain the quiescent current at a reasonable value with variations in junction temperatures. It is usually

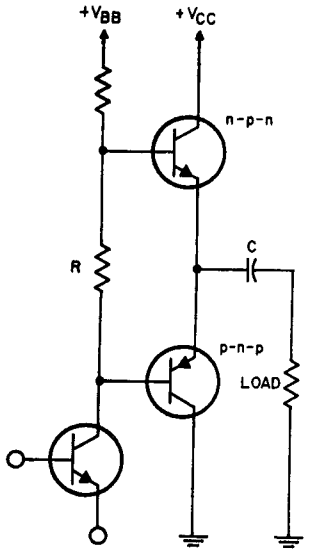



Figure 652. Basic complementary type of series-output circuit.

thermally connected to one of the output transistors and tracks with the $V_{BE}$ of the output transistors.

The complementary circuit is by far the most thermally stable output circuit. It places the output transistors in a $V_{CES}$ mode because both transistors are operated with a low impedance between base and emitter. Therefore, the $I_{CBO}$ leakage is the only component of concern in the stability criteria. At power-output levels from 3 to 20 watts, a complementary-symmetry amplifier offers advantages in terms of circuit simplicity. At higher power levels, however, the class A driver transistor is required to dissipate considerable heat, the quiescent power-supply current drain becomes significant, and excessively large filter capacitors are required to maintain a low hum level. For these reasons, the maximum practical output for a true complementary-symmetry amplifier is considered to be about 20 watts; at higher power levels, this type of amplifier is usually replaced by the quasi-complementary circuit.

**Quasi-Complementary Amplifiers**—In the quasi-complementary amplifier, shown in Fig. 653, the driver transistors provide the necessary phase inversion. A single but descriptive way to analyze the operation of a quasi-complementary amplifier is to consider the result of connecting a p-n-p transistor to a high-power n-p-n output transistor, as shown in Fig. 654. The collector current of the p-n-p transistor becomes the base current of the n-p-n transistor. The n-p-n transistor, which is operated as an emitter-follower, provides additional current gain

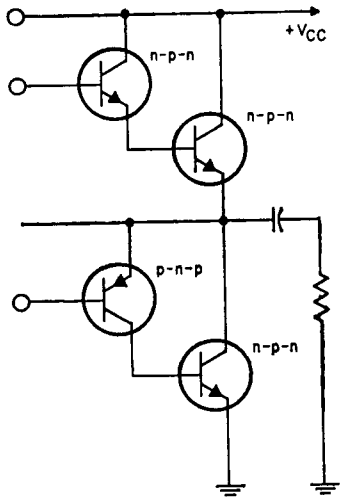



Figure 653. Basic quasi-complementary type of series-output circuit.

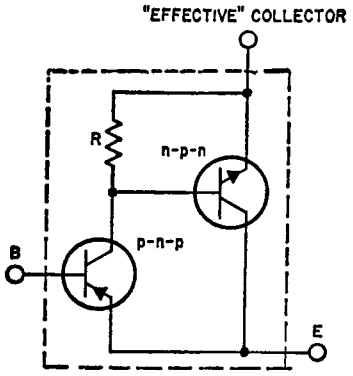



Figure 654. Connection of n-p-n driver transistor to n-p-n output transistor.

without inversion. If the emitter of the n-p-n transistor is considered as the "effective" collector of the composite circuit, it becomes apparent that the circuit is equivalent to a high-gain, high-power p-n-p transistor. The output characteristics of the p-n-p circuit shown in Fig. 654 and of a high-gain, high-power n-p-n circuit formed by the connection of the same type of n-p-n output transistor and an n-p-n driver transistor in a Darlington configuration, such as shown in Fig. 655, are compared in Fig. 656.

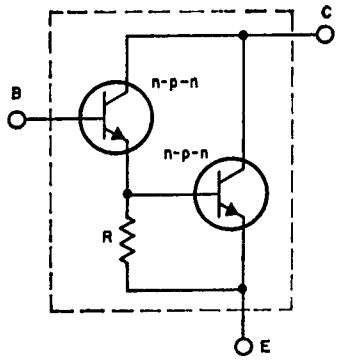



Figure 655. Darlington connection of n-p-n driver transistor to n-p-n output transistor.

The saturation characteristics of the over-all circuit in both cases are the combination of the base-to-emitter voltage $V_{BE}$ of the output transistor and the collector saturation voltage of the driver transistor. Moreover, in both cases the current gain is the product of the individual betas of the transistors used. A quasi-complementary amplifier, therefore, is effectively the same as a simple complementary output circuit such as that shown in Fig. 652, and is formed by the use of high-gain, high-power n-p-n and p-n-p equivalent transistors. In both cases, the resistor R between the emitter and base of the output transistor places the device in a $V_{CER}$ mode. This mode is not

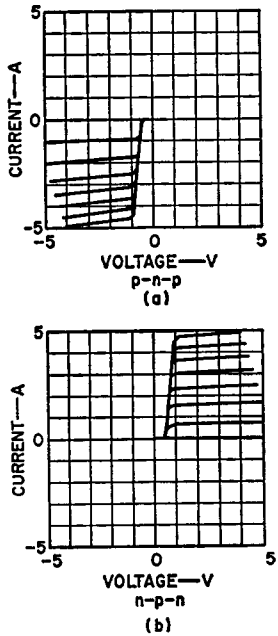



Figure 656. Output characteristics for (a) p-n-p/n-p-n driver-output transistor pair shown in Fig. 654 and for (b) Darlington pair of n-p-n transistors shown in Fig. 655.

as stable as that of the complementary amplifier, but presents no problem for silicon transistors.

A typical quasi-complementary amplifier is shown in Fig. 657. Capacitor C performs two functions essential to the successful operation of the circuit. First, it acts as a bypass to decouple any power-supply ripple from the driver and predriver stages. Second, it is connected as a "bootstrap" capacitor to provide the drive necessary to pull the upper Darlington pair of transistors into saturation. This latter function results from the fact that the stored voltage of the capacitor, with reference to the output point A, provides a higher voltage than the normal collector-supply voltage to drive transistor $Q_2$. This

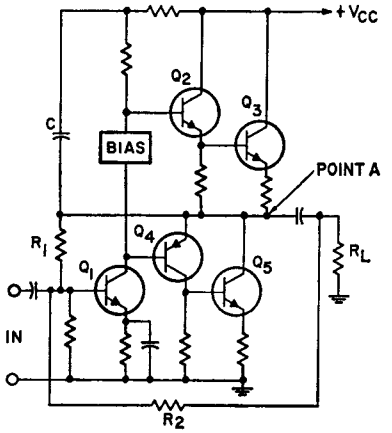



Figure 657. Quasi-complementary audio power amplifier that operates from a single dc supply.

higher voltage is necessary during the signal conditions that exist when the upper transistors are being turned on because the emitter voltage of transistor $Q_2$ then approaches the normal supply voltage. An increase in the base voltage to a point above this level is required to drive the transistor into saturation. Resistor $R_1$ provides the necessary dc feedback to maintain point A at approximately one-half the nominal supply voltage. Over-all ac feedback from output to input is coupled by resistor $R_2$ to reduce distortion and to improve low-frequency performance.

As indicated in Fig. 649(b), series-output circuits can be employed with separate positive and negative supplies; no series output capacitor is then required. The elimination of this capacitor may result in an economic advantage, even though an additional power supply is used, because of the size of the series output capacitor necessary in the single-supply case to obtain good low-frequency performance (e.g., a 2000-microfarad capacitor is required to provide a 3-dB point at 20 Hz for a 4-ohm load impedance). Split supplies, however, pose certain problems which do not exist in the single-supply case. The output of the amplifier must be maintained at zero potential under quiescent conditions for all environmental conditions and device parameter variations. Also, the input ground reference can no longer be at the same point as that indicated in Fig. 657, because this point is at the negative supply potential in a split-supply system.

If the ground-point reference for the input signal were a common point between the split supplies, any ripple present on the negative supply would effectively drive the amplifier through transistor $Q_1$, with the result that this stage would operate as a common-base amplifier with its base grounded through the effective impedance of the input signal source.

To avoid this condition, the amplifier must include an additional p-n-p transistor as shown in Fig. 658. This transistor ($Q_6$) reduces the drive effects of the negative supply ripple because of the high collector impedance (1 megohm or more) that it presents to the base of transistor $Q_1$, and effectively

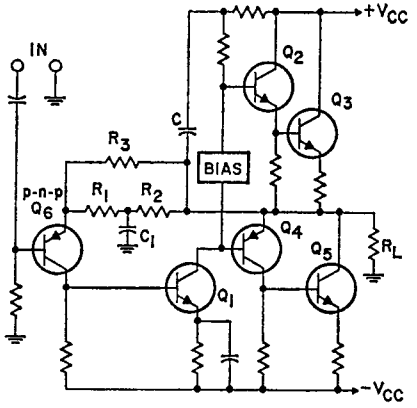



Figure 658. Quasi-complementary audio power amplifier that operates from symmetrical dual dc power supplies. The p-n-p transistor input stage is required to prevent ripple component from driving amplifier.

isolates the input source impedance from transistor $Q_1$. In practice, transistor $Q_1$ may be replaced by a Darlington pair to reduce the loading effects on the p-n-p predriver.

Negative dc feedback is applied from the output to the input stage by $R_1$, $R_2$, and $C_1$ so that the output is maintained at about zero potential. Actually, the output is maintained at approximately the forward-biased base-emitter voltage of transistor $Q_6$, which may be objectionable in a few cases, but which can be eliminated by a method discussed later. Capacitor $C_1$ effectively bypasses the negative dc feedback at all signal frequencies. Resistor $R_3$ provides ac feedback to reduce distortion in the amplifier.

## POWER OUTPUT IN CLASS B AUDIO AMPLIFIERS

For all cases of practical interest, the power output ($P_o$) of an audio amplifier is given by the following equation:

$$P_o = I(\text{rms}) \times E(\text{rms}) = (I_p E_p)/2$$
$$= (I_p^2 R_L)/2 = E_p^2/2R_L \quad (435)$$

where $I_p$ and $E_p$ are the peak load current and voltage, respectively, and $R_L$ is the load impedance presented to the transistor. Fig. 659 shows the relationship among these various factors in graphic form. Obviously, the peak load current is the peak transistor current, and the transistor breakdown-voltage rating must be at least twice the peak load voltage. The vertical lines that denote 4-ohm, 8-ohm, and 16-ohm resistances are particularly useful for transformerless designs in which the transistor operates directly into the loudspeaker.

### Rating Methods

The **Institute of High Fidelity** (IHF) and the **Electronic Industries Association** (EIA) have attempted to standardize power-output ratings to establish a common reference of comparison and to provide a solid definition of the capabilities of audio power amplifiers. Obviously, an audio power amplifier using an unregulated supply can deliver more output power under transient conditions than under steady-state conditions. The rating methods which have been standardized for this type of operation are the **IHF Dynamic Output Rating (IHF-A-201)** and

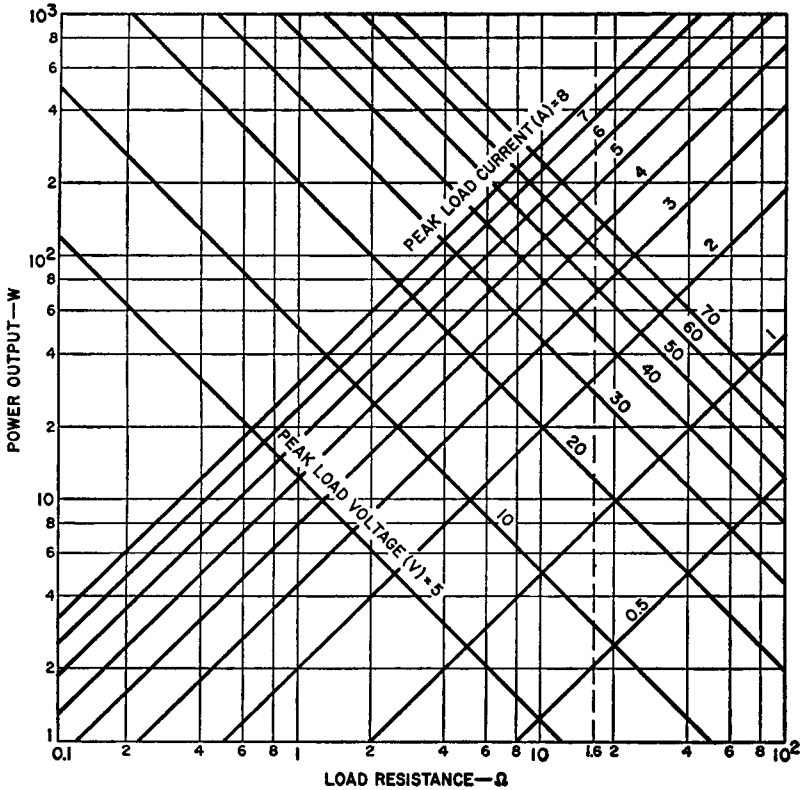



Figure 659. Peak transistor currents and load voltages for various output powers and load resistances.

the **EIA Music Power Rating (EIA RS-234-A).**

Both of these measurement methods allow the use of regulated supply voltage to simulate transient conditions. Because the regulated supply has no source impedance or ripple, the results do not completely represent the transient conditions, as will be explained later.

**Measurement Techniques**—The EIA standard is used primarily by manufacturers of packaged equipment, such as portable phonographs, packaged stereo hi-fi consoles, and packaged home-entertainment consoles. The **EIA music power output** is defined as the power obtained at a total harmonic distortion of 5 per cent or less, measured after the "sudden application of a signal during a time interval so short that supply voltages have not changed from their no-signal values." The supply voltages are bypassed voltages. These definitions mean that the internal supply may be replaced with a regulated supply equal in voltage to the no-signal voltage of the internal supply. For a stereo amplifier, the music power rating is the sum of both channels, or twice the single-channel rating.

The IHF standard provides two methods to measure dynamic output. One is the **constant-supply method.** This method assumes that under music conditions the amplifier supply voltages undergo only insignificant changes. Unlike the EIA method, this measurement is made at a reference distortion. The constant-supply method is used by most high-fidelity component manufacturers. The reference distortion chosen is normally less than one per cent, or considerably lower than the EIA value of 5 per cent used by packaged-equipment manufacturers.

A second IHF method is called the **"transient distortion" test.** This method requires a complex setup including a low-distortion modulator with a prescribed output rise time and other equipment. The modulator output is required to have a rise time of 10 to 20 milliseconds to simulate the envelope rise time of music and speech. This measurement is made using the internal supply of the amplifier and, consequently, includes distortion caused by voltage decay, power-supply transients, and ripple. This method tends to be more realistic and to yield lower power-output ratings than the constant-supply method. Actually, both IHF methods should be used, and the lowest power rating obtained at reference distortion with both channels operating, both in and out of phase, should be used as the power rating. (There is some question concerning unanimity among high-fidelity manufacturers on actually performing both IHF tests.)

Because music is not a continuous sine wave, and has average power levels much below peak power levels, it would appear that the music power or dynamic power ratings are true indications of a power amplifier's ability to reproduce music program material. The problem is that all three methods described have a common flaw. Even the transient-distortion method fails to account for the ability of the audio amplifier to reproduce power peaks while it is already delivering some average power. The amplifier is almost never delivering zero output when it is called on to deliver a transient. For every transient that occurs after an extremely quiet passage or zero signal, there are hundreds that are imposed on top of some low but non-zero average power level.

This condition can best be clarified by consideration of the power supply. Many amplifiers have regulated supplies for the front-end or low-level stages, but almost none provides a regulated supply for the power-output stages because regulation requires extra transistors or other devices; it becomes costly, especially at high power levels. The power supply for the output stages of power ampliers is commonly a nonregulated rectifier supply having a capacitive input filter. The output voltage of such a supply is a function of the output current and, consequently, of the power output of the amplifier.

**Effect of Power-Supply Regulation**—Power-supply regulation is dependent on the amount of effective internal series resistance present in the power supply. The effective series resistance includes such things as

the dc resistance of the transformer windings, the amount and type of iron used in the transformer, the amount of surge resistance present, the resistance of the rectifiers, and the amount of filtering. The internal series resistance causes the supply voltage to drop as current is drawn from the supply.

Fig. 660 shows a typical regulation curve for a rectifier power supply that has a capacitive input filter. The voltage is a linear

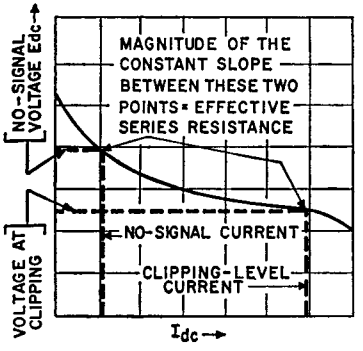



Figure 660. Regulation curve for capacitive rectifier power supply.

function of the average supply current over most of the useful range of the supply. However, a rapid change in slope occurs in the regions of both very small and very large currents. In class B amplifiers, the no-signal supply current normally occurs beyond the low-current knee, and the current required for the amplifier at the clipping level occurs before the high-current knee. The slope between these points is nearly linear and may be used as an approximation of the equivalent series resistance of the supply.

The amount of power lost depends on the quality of the power supply used in the amplifier. Accordingly, rating amplifier power output with a superb **external** power supply (that is, not using the built-in amplifier power supply) provides false music power outputs. Under actual usage the output is lower, as illustrated by the following example.

Figs. 661 and 662 show equivalent circuits for capacitive-input rectifier supplies. In these circuits, $I_{dc}$ is the average supply current, $R_s$ is the effective equivalent series resistance of the power supply, $E_0$ is the no-signal voltage, and $E_s$ is the steady-state supply voltage. The steady-state voltage, $E_s$, is related to the no-signal voltage, $E_0$, as follows:

$$E_s = E_0 - R_s I_{dc} \quad (436)$$

Eq. (436) shows that the supply voltage, $E_s$, is equal to the no-signal supply voltage, $E_0$, only when there is no current other than the no-signal current being drawn from the supply. As soon as the amplifier begins to deliver some power to the load, the power supply is called upon to deliver some current. A single-ended

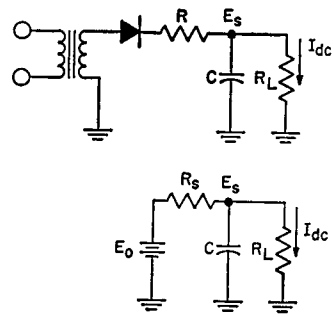



Figure 661. Equivalent circuit for single-ended capacitive-input rectifier supply.

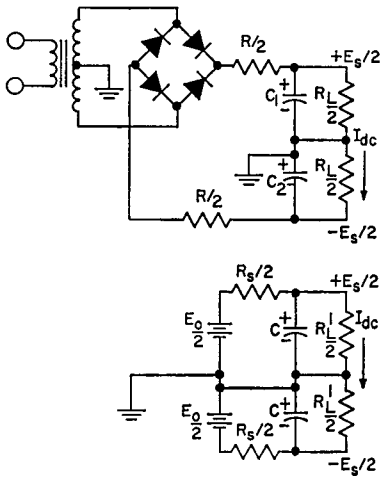



Figure 662. Equivalent circuit for split capacitive-input rectifier supply.

power supply delivers current on alternate half-cycles, and each half of a split supply delivers current on alternate half-cycles. Therefore, in each case, the supply current, $I_{dc}$, is related to the peak output current, as follows:

$$I_{dc} = \frac{I_{pk}}{\pi} \qquad (437)$$

The power output is related to the peak output current, as follows:

$$P_o = \frac{I_{pk}^2}{2} R_L \qquad (438)$$

where $R_L$ is the speaker load resistance. Consequently, the supply current is related to the power output by

$$I_{dc} = \left(\frac{2 P_o}{\pi^2 R_L}\right)^{1/2} \qquad (439)$$

Eqs. (436) and (439) are combined to obtain the following result:

$$E_s = E_o - R_s \left(\frac{2 P_o}{\pi^2 R_L}\right)^{1/2} \qquad (440)$$

In relating average current and power output, it is assumed that sine-wave signals are included and that no parasitic losses exist.

This relationship can be simplified if it is assumed that $R_L$ is 8 ohms and $\pi^2$ is 10. Eq. (440) then becomes

$$E_s = E_o - 0.158\, R_s\, (P_o)^{1/2} \qquad (441)$$

To illustrate the inability of the no-signal supply voltage to indicate the transient power capability of an amplifier, it is assumed that an amplifier power supply has the regulation characteristics shown in Fig. 663. Using the values for voltage and current from Fig. 663, the music power rating, based on the no-signal voltage (44.8 volts) and an 8-ohm load, is given by

$$P_{music} = \frac{E_o^2}{8 R_L} = 31.5 \text{ watts/channel} \qquad (442)$$

If no parasitic losses are assumed, therefore, the stereo music power would be 63 watts.

[The factor 8 in Eq. (442) is derived in converting peak-to-peak volts to rms volts.]

The effective series resistance, according to Fig. 660, is approximately 6 ohms (for the values of $E_o$ and $R_L$ specified above). If the amplifier is delivering an average power of 2 watts per channel, or a total of 4 watts, the supply voltage decreases from 44.8 volts to a smaller value, determined from the following equation:

$$E_s = 44.8 - 0.158\,(6)\,(2) = 42.9$$

The music power rating is then given by $E_s^2/8R_L = 28.7$ watts per channel, or 57.4 watts for stereo. The difference in these ratings represents a 10-per cent decrease in the actual transient power capability.

The preceding calculations prove that an amplifier with a 31.5-watt-per-channel music power rating may, in fact, have an actual power-output capability of only 28.7 watts per channel.

This decrease of 10 per cent in measured transient capability may be as much as 20 per cent in some cases. One such case is where the no-signal load is less than that shown in Fig. 663. The

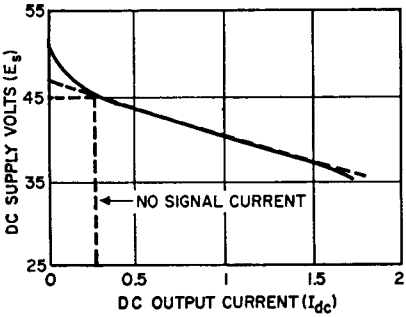



Figure 663. Typical power supply regulation curve.

no-signal load includes the class AB bias current of the output stages and all of the current drawn by the preceding stages and their associated bias networks. When this total current is below the 250-milliampere value shown, the no-signal voltage is located on the steep portion of the regulation curve. As a result, there is a greater decline in supply voltage when the amplifier is called on to deliver 2 watts or more of average power.

It should be emphasized that, while there is a discrepancy between the actual power available and the power measured under the EIA Music Power or the IHF Dynamic Power methods, these methods are not without merit. The IHF dynamic power rating, in conjunction with the continuous power rating, produces an excellent indication of how the amplifier will perform. The EIA music power rating, which is measured at a total harmonic distortion of 5 per cent with a regulated power supply, provides a less adequate indication of amplifier performance because there is no indication of how the amplifier power-supply voltage reacts to power output.

Some important factors considered by packaged-equipment manufacturers, the primary users of the EIA music power rating, are mostly economic in nature and affect many aspects of the amplier performance. Because there is no continuous power output rating required, two amplifiers may receive the same EIA music power rating but have different continuous power ratings. The ratio of music power to continuous power is, of course, a function of the regulation and effective series resistance of the supply.

One reason for the difference between ratings used by the console or the packaged-equipment manufacturer and those used by the hi-fi component manufacturer is that the latter does not always know just what will be required of the amplifier. The console manufacturer always designs an amplifier as part of a system, and consequently knows the speaker impedances and the

power required for adequate sound output. The console manufacturer may use high-effciency speakers requiring only a fraction of the power needed to drive many component-type acoustic-suspension systems. The difference may be such that the console may produce the same sound pressure level with an amplifier having one-tenth of the power output. High ratios of music-power to continuous-power capability are common in these consoles. A typical ratio of IHF music power to continuous power may be 1.2 to 1 in component amplifiers, whereas a typical ratio of EIA music power to continuous power in a console system may be 2 to 1. Console manufacturers use the EIA music power rating to economic advantage as a result of the reduced regulation requirement of the power supply. A high ratio of music power to continuous power means higher effective series resistance in the power supply. This resistance, in turn, means less continuous dissipation on the output transistors, smaller heat sinks, and a lower-cost power supply.

## Basic Power-Dissipation Relationships

Under ideal conditions (i.e., with a perfectly regulated dc power supply), maximum transistor power dissipation in a class B audio output stage is approximately 20 per cent of the maximum unclipped sine-wave power output and occurs when the output stage is delivering approximately 40 per cent of the maximum output power to the load. In the following paragraphs, this statement is verified by analysis of a typical complementary-symmetry amplifier, such as those shown in Fig. 664. The effect of a nonregulated power supply on transistor dissipation is then examined.

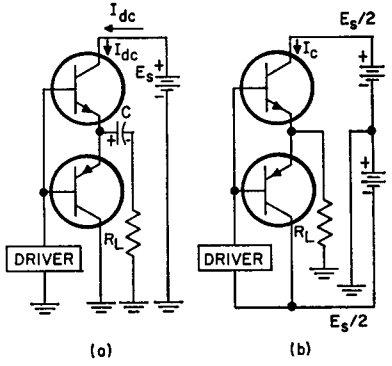



Figure 664. Typical complementary-symmetry circuits.

**Regulated Supply**—When the amplifier circuit shown in Fig. 664(a) is operated from a regulated supply, the capacitor C, under no-signal conditions, is charged to a voltage equal to one-half the supply voltage (i.e., $E_c = E_s/2$) at the clipping level. The maximum peak load current $I_{pk}(\max)$ is given by

$$I_{pk}(\max) = E_s/2\,R_L \quad (443)$$

Because the supply delivers current on alternate half-cycles, the average supply current $I_{dc}$ is given by

$$I_{dc} = I_{pk}/\pi \quad (444)$$

The power $P_o$ delivered by the supply can then be expressed as follows:

$$P_s = (I_{pk}\,E_s)/\pi \quad (445)$$

The power delivered to the load $P_o$ is given by

$$P_o = (I_{pk}^2\,R_L)/2 \quad (446)$$

The dissipation $P_T$ for each transistor is equal to half the difference between the supply power delivered $P_s$ and the power dissipated in the load $P_o$, as follows:

$$P_T = (P_s - P_o)/2$$

$$P_T = \frac{I_{pk} E_s}{2\pi} - \frac{I_{pk}^2 R_L}{4} \quad (447)$$

If Eq. (447) is differentiated and solved for the peak load current $I_{pk}$ at maximum average transistor dissipation, the following expression is obtained:

$$I_{pk} = E_s/(\pi R_L) \quad (448)$$

When this value is substituted in Eq. (447), the ratio of maximum average transistor dissipation $P_T(max)$ to power delivered to the load at full power output $P_o(max)$ can be expressed as follows:

$$\frac{P_T(max)}{P_o(max)} = \frac{2}{\pi 2} \quad (449)$$

Eq. (333) indicates that maximum transistor dissipation is approximately 20 per cent of full power output. At the point of maximum dissipation, the power output is given by

$$P_o(\text{max diss.}) = \frac{E_s^2}{2\pi^2 R_L} \quad (450)$$

The ratio of the power output at maximum dissipation $P_o$ (max diss.) to maximum power output $P_o(max)$ is then given by

$$\frac{P_o(\text{max diss.})}{P_o(max)} = \frac{4}{\pi^2} \quad (451)$$

**Nonregulated Supply**—In the case of a nonregulated supply that has an internal resistance $R_s$, the supply voltage $E_s$ is expressed by Eq. (436). If this value for $E_s$ is substituted in Eq. (437), the following result is obtained:

$$P_T = \frac{I_{pk} E_o}{2\pi} - \frac{R_s (I_{pk})^2}{2\pi^2} - \frac{(I_{pk})^2 R_L}{4} \quad (452)$$

The partial derivative of Eq. (452) with respect to $I_{pk}$ is set equal to zero, tested for a maximum value, and solved for $I_{pk}$. This value of $I_{pk}$ is then used in Eq. (452) to determine the maximum transistor dissipation $P_T(max)$, as follows:

$$\frac{dP_T}{dI_{pk}} = \frac{E_o}{2\pi} - I_{pk}\frac{2R_S + \pi^2 R_L}{2\pi^2} \quad (453)$$

$$I_{pk} = \frac{E_o \pi}{2R_S + \pi^2 R_L} \quad (454)$$

$$P_T(max) = \frac{E_o^2}{8R_S + 4\pi^2 R_L} \quad (455)$$

Clipping begins at the point where the peak collector current $I_{pk}$ is given by

$$I_{pk} = \frac{E_o \pi}{R_S + 2\pi R_L} \quad (456)$$

Power output at clipping can then be expressed as follows:

$$P_o(\text{clipping}) = \frac{E_o^2 \pi^2 R_L}{2(R_S + 2\pi R_L)^2} \quad (457)$$

If $R_s = 0$ is substituted in Eq. (457), the power output may be expressed as follows:

$$P_o = E_o^2/8R_L \quad (458)$$

This value is equivalent to the power output just prior to clipping with a fully regulated supply and, for the remainder of this discussion, is referred to as the music power output. [This definition of music power output, i.e., as the maximum unclipped sine-wave power output, differs from the EIA standard

(RS-234-A), which defines the music power output as the point at which the total harmonic distortion is 5 per cent when a regulated supply is used. The EIA value is about 10 per cent greater.]

Maximum average transistor dissipation is related to the music power output by the following expression:

$$\frac{P_T\,(max)}{P_o\,(music)} = \left[\frac{\pi^2}{2} + \frac{R_S}{R_L}\right]^{-1} \quad (459)$$

The power output at which maximum average transistor dissipation occurs $P_o$ (max diss) is related to the music power output as follows:

$$\frac{P_o\,(max\ diss.)}{P_o\,(music)} = \left[\frac{\pi^2}{4} + \frac{R_S}{R_L} + \frac{R_S^2}{\pi^2 R_L^2}\right]^{-1} \quad (460)$$

The continuous power output at the clipping level, $P_o$ (clipping), is related to the music power output by the following expression:

$$\frac{P_o\,(clipping)}{P_o\,(music)} = \left[1 + \frac{R_S}{\pi R_L} + \frac{R_S^2}{4\pi^2 R_L^2}\right]^{-1} \quad (461)$$

Eq. (459), (460), and (461) are plotted in Fig. 665. Power levels are normalized with respect to the music power output and are plotted as a function of $R_S/R_L$.

The equations plotted in Fig. 665 suggest some interesting possibilities. Transistor power dissipation is only a small fraction of the clipping power output for higher ratios of $R_S/R_L$. For

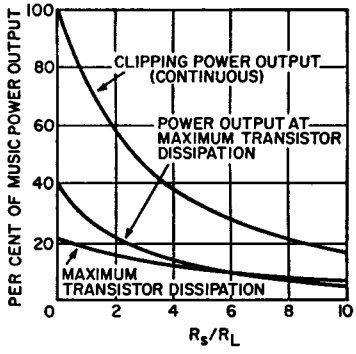



Figure 665. Power output and dissipation as functions of $R_S$ and $R_L$.

example, a 100-watt amplifier could be built using transistors and associated heat sinks capable of only about 7 watts of maximum dissipation each.

The equations presented, however, do not consider high line voltage or effects of ripple voltage. Calculations for average transistor dissipation should also include no-signal bias dissipation and the increase in bias dissipation with increasing ambient and junction temperatures in class AB circuits. Storage effects, phase shift, and thermal tracking should also be considered.

Of the above factors, bias dissipation probably contributes the greatest percentage of average worst-case transistor dissipation. The output stage is usually biased on slightly (class AB) to reduce cross-over distortion.

It is possible, however, to design amplifiers for which bias dissipation is not a problem. One such amplifier is shown in Fig. 666. The bias dissipation in this amplifier is negligible at all practical temperatures. One side is cut off and the other conducts less than one milliampere. Thermal runaway cannot be ini-

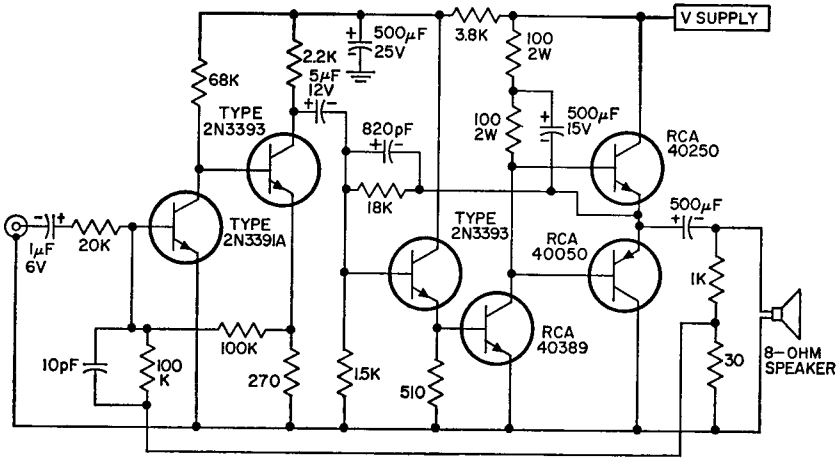



Figure 666. Class B complementary-symmetry power amplifier.

tiated in the output stage at any junction temperature below its maximum rating. Consequently, thermal tracking may also be neglected so long as the ambient temperature plus the product of the instantaneous dissipation times the junction-to-ambient thermal resistance is less than the maximum junction temperature rating.

Storage effects are also reduced as a result of the reverse bias provided for the off-transistor by the on-transistor in complementary symmetry. This circuit, then, is one practical example of an amplifier capable of achieving the characteristics shown in Fig. 665.

## Ratio of Music Power To Continuous Power

Some advantages of high values of the ratio $R_S/R_L$ and correspondingly high ratios of music power output to transistor dissipation are as follows:

1. Reduced heat sink or transistor cost: Because the volt-ampere capacity of the transistor is determined by the music power output, it is not likely that reduced thermal-resistance requirements will produce significant cost reductions. Alternatively, the heat-sink requirements may be reduced.

2. Reduced power supply costs: Transformer and/or filter-capacitor specifications may be relaxed.

3. Reduced speaker cost: Continuous power-handling capability may be relaxed.

These cost reductions may be passed along to the consumer in the form of more music power per dollar.

The question arises as to how high the ratio $R_S/R_L$ and the corresponding ratio of music power output to continuous power output may go before the capability of the amplifier to reproduce program material is impaired.

The objective is to provide the

listener with a close approximation of an original live performance. Achievement of this objective requires the subjective equivalents of sound pressure levels that approach those of a concert hall. Although the peak sound pressure level of a live performance is about 100 dB, the average listener prefers to operate an audio system at a peak sound pressure level of about 80 dB. The amplifier, however, should also accommodate listeners who desire higher-than-average levels, perhaps to peaks of 100 dB.

A sound pressure level of 100 dB corresponds to about 0.4 watt of acoustic power for an average room of about 3,000 cubic feet. If speaker efficiencies are considered to be in the order of 1 per cent, a stereophonic amplifier must be capable of delivering about 20 watts per channel. Higher power outputs are required for lower-efficiency speakers. The peak-to-average level for most program material is between 20 and 23 dB. A system capable of providing a continuous level of 77 dB and peaks of 100 dB would satisfy the power requirements of nearly all listeners. For this performance to be attained, the power-supply voltage cannot drop below the voltage required for 100 dB of acoustic power while delivering the average current required for 77 dB. Moreover, because sustained passages that are as much as 10 dB above the average may occur, the power-supply voltage cannot drop below the value required for 100 dB of acoustic power while delivering 87 dB of acoustic power (87 dB of acoustic power corresponds to about 1 watt per channel). This performance means that for 8-ohm loads, with output-circuit losses neglected, the power-supply voltage must not decrease to a value less than 36 volts, while delivering the average current required for 1 watt per channel (0.225 ampere dc).

It should be noted that the power-output capability for peaks while the amplifier is delivering a total of 2 watts is not the music power rating of the amplifier because the power-supply voltage is below its no-signal value by an amount depending on its effective series resistance.

## Maximum Effective Series Resistance

There is a relationship between the maximum effective series resistance of the power supply and the music power rating of the amplifier if it is to perform to the standards as outlined above.

The power-supply series resistance $R_s$ may be expressed as a function of music-power output, as follows:

$$R_s = \left[\frac{(8R_L)\ P_o\ (\text{music})}{\bar{I}^2}\right]^{1/2} - \frac{E_s\ (\text{min})}{\bar{I}} \qquad (462)$$

where $E_s(\text{min})$ is the minimum voltage required for 100 dB of acoustical power output and $\bar{I}$ is the current required for 87 dB of acoustical power output, less the idle current. [$E_s(\text{min})$ should, in practice, be increased by peak output-circuit voltage losses.]

Eq. (462) is plotted in Fig. 667. The value of $R_s$ is the absolute maximum value of effective

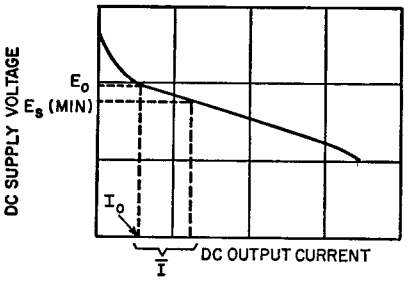



Figure 667. Power-supply regulation curve.

supply resistance for each music-power value that will allow the amplifier to deliver a minimum of 100 dB of acoustical power output as described above.

Use of Fig. 668 in conjunction with Fig. 665 shows that very high ratios of music power output to continuous power output may be employed without sacrifice of the subjective ability of

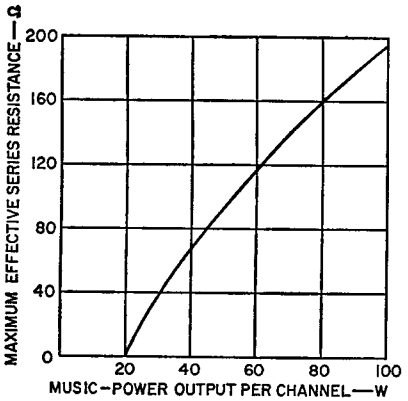



Figure 668. Maximum effective series resistance as a function of music power output.

the amplifier to reproduce program material. This technique provides economic advantages while adhering to a minimum "power margin" for the faithful reproduction of program material, even at loud levels.

## THERMAL-STABILITY REQUIREMENTS

One serious problem that confronts the design engineer is the achievement of a circuit which is thermally stable at all temperatures to which the amplifier might be exposed. As previously discussed, thermal runaway may be a problem because the $V_{BE}$ of all transistors decreases at low current. It should be noted, however, that at high current levels the base-to-emitter voltage of silicon transistors increases with a rise in junction temperature. This characteristic is the result of the increase in the base resistance that is produced by the rise in temperature. The increase in base resistance helps to stabilize the transistor against thermal runaway. In high-power amplifiers, the emitter resistors employed usually have a value of about 1 ohm or less. The size of the capacitor required to bypass the emitter adequately at all frequencies of interest makes this approach economically impractical. A more practical solution is to increase the value of the emitter resistor and shunt it with a diode. With this technique, sufficient degeneration is provided to improve the circuit stability; at low currents, however, the maximum voltage drop across the emitter resistor is limited to the forward voltage drop of the diode.

The quasi-complementary amplifier shown in Fig. 669 incorporates the stabilization techniques described. A resistor-diode network is used in the emitter of transistor $Q_3$, and another such network is used in the collector of transistor $Q_5$, with the emitter of

transistor $Q_4$ returned to the collector of transistor $Q_5$. Previous discussion regarding the p-n-p driver and n-p-n output combination ($Q_4$ and $Q_5$) showed that the collector of the output device becomes the "effective" emitter of the high-gain, high-power p-n-p equivalent, and vice versa. For maximum operating-point stability, therefore, the diode-resistor network should be in the "effective" emitter of the p-n-p equivalent. Quasi-complementary circuits employing the stabilization resistor in the emitter of the lower output transistor, as shown in Fig. 657, do not improve the operating-point stability of the over-all circuit.

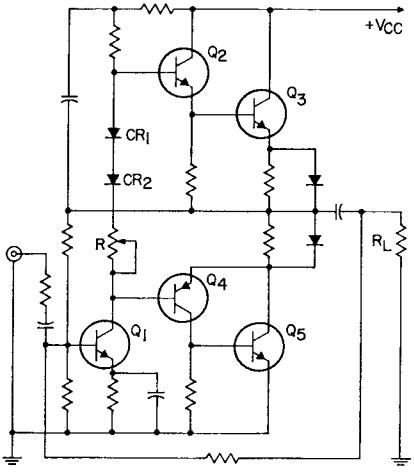



Figure 669. Quasi-complementary amplifier that incorporates two stabilization networks.

The circuit shown in Fig. 669 is biased for class AB operation by the voltage obtained from the forward drop of two diodes, $CR_1$ and $CR_2$, plus the voltage drop across potentiometer R, which affords a means for a slight adjustment in the value of the quiescent current. The current necessary to provide this voltage reference is the collector current of driver transistor $Q_1$. The diodes may be thermally connected to the heat sink of the output transistors so that thermal feedback is provided for improvement of further thermal stability. Because the forward voltage of the reference diodes decreases with increasing temperature, these diodes compensate for the decreasing $V_{BE}$ of the output transistors by reducing the external bias applied. In this way, the quiescent current of the output stage can be held relatively constant over a wide range of operating temperatures.

## EFFECTS OF LARGE PHASE SHIFTS

The amplifier frequency-response characteristic is an important factor with respect to the ability of the amplifier to withstand unusually severe electrical stress conditions. For example, under certain conditions of input-signal amplitude and frequency, the amplifier may break into high-frequency oscillations which can lead to destruction of the output transistors, the drivers, or both. This problem becomes quite acute in transformer-coupled amplifiers because the characteristics of transformers depart from the ideal at both low and high frequencies. The departure occurs at low frequencies because the inductive reactance of the transformer decreases, and at high frequencies because the effects of leakage inductance and transformer winding capacitance become appreciable. At both frequency extremes, the effect is to introduce a phase shift between input and output voltage.

Negative feedback is used almost universally in audio amplifiers; the voltage coupled back to the input through the feedback loop may cause the amplifier to be potentially unstable at some frequencies, especially if the additional phase shift is sufficient to make the feedback positive. Similar effects can occur in transformerless amplifiers because reactive elements (such as coupling and bypass capacitors, transistor junction capacitance, stray wiring capacitance, and inductance of the loudspeaker voice coil) are always present. The values of some of the reactive elements (e.g., transistor junction capacitance and transformer inductance as the core nears saturation) are functions of the signal level; coupling through wiring capacitance and unavoidable ground loops may also vary with the signal level. As a result, an amplifier that is stable under normal listening levels may break into oscillations when subjected to high-level signal transients.

A large phase shift is not only a potential cause of amplifier instability, but also results in additional transistor power dissipation and increases the susceptibility of the transistor to forward-bias second-breakdown failures. The effects of large-signal phase shifts at low frequencies are illustrated in Fig. 670, which shows the load-line characteristics of a transistor in a class AB push-pull circuit similar to that shown in Fig. 659 for signal frequencies of 1000 Hz and 10 Hz. The phase shift is caused primarily by the output capacitor. In both cases, the amplifier is driven very strongly into saturation by a 5-volt input signal. The increased dissipation at 10 Hz, compared to that obtained at 1000 Hz, results from simultaneous high-current high-voltage operation. The transistor is required to handle safely a current of 0.75 ampere at a collector voltage of 40 volts for an equivalent pulse duration of about 10 milliseconds; it must be free from second-breakdown failures under these conditions.

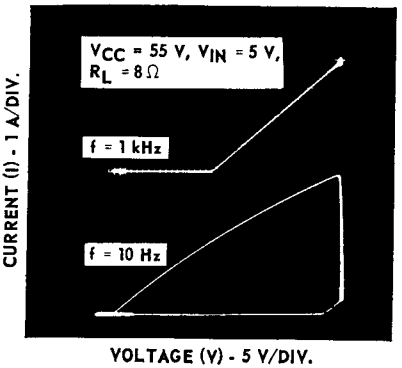



Figure 670. Effect of large signal phase shift on the load-line characteristics of a transistor at low frequencies.

## EFFECT OF EXCESSIVE DRIVE

Simultaneous high-current high-voltage operation may also occur in class B amplifiers at high frequencies when the amplifier is overdriven to the point that the output signals are clipped. For example, if the input signal applied to the series-output push-pull circuit shown in Fig. 671(a) is large enough to drive the transistors into both saturation and cutoff, transistor A is driven into saturation, and transistor B is cut off during a portion of the input cycle. Fig. 671(b) shows the collector-current waveform for transistor A under these conditions.

During the interval of time from $t_2$ to $t_3$, transistor A operates in the saturation region, and the

output voltage is clipped. The effective negative feedback is then reduced because the output voltage does not follow the sinusoidal input signal. Transistor A, therefore, is driven even further into saturation by the unattenuated input signal. When transistor B starts to conduct, transistor A cannot be turned off immediately because the excessive drive results in a large storage time. As a result, transistor B is required to support almost the full supply voltage (less only the saturation voltage of transistor A and the voltage drop across the emitter resistors, if used) as its current is increased by the drive signal.

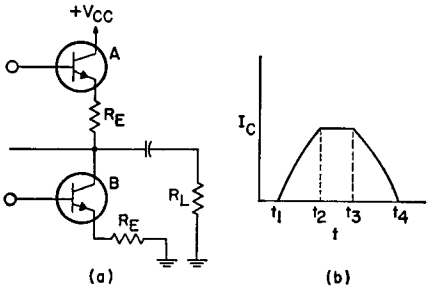



Figure 671. (a) Class B series-output stage, (b) collector-current waveform under overdrive (clipping) conditions.

For this condition to occur, a large input signal is required at a frequency high enough so that the storage time is greater than one-quarter cycle.

Because of the charging current through the output coupling capacitor, transistor A in Fig. 671(a) is also subject to forward-bias second-breakdown failure if the dc supply voltage and a large input signal are applied simultaneously.

All of these conditions point to the need for a good "safe area" of operation. Fig. 672 shows the

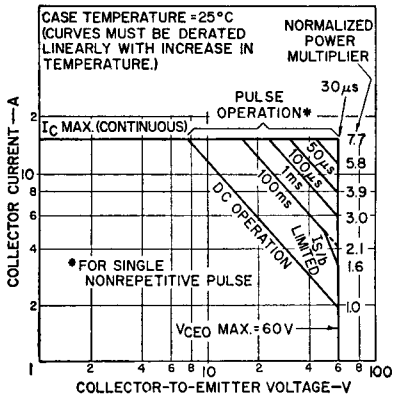



Figure 672. Safe-area-of-operation rating chart for the RCA-2N3055 homotaxial-base transistor.

safe area for the RCA-2N3055. In all cases, the load lines fall within the area guaranteed safe for this transistor.

## SHORT-CIRCUIT PROTECTION

Another important consideration in the design of high-power audio amplifiers is the ability of the circuit to withstand short-circuit conditions. As previously discussed, overdrive conditions may result in disastrously high currents and excessive dissipation in both driver and output stages. Obviously, some form of short-circuit protection is necessary. One such technique is shown in Fig. 673. A current-sampling resistor R is placed in the ground leg of the load. If any condition (including a short) exists such that higher-than-normal load current flows, diodes $CR_1$ and $CR_2$ conduct on alternate half-cycles and thus provide a high negative feedback which effectively reduces the drive of the amplifiers. This feedback should not exceed the stability

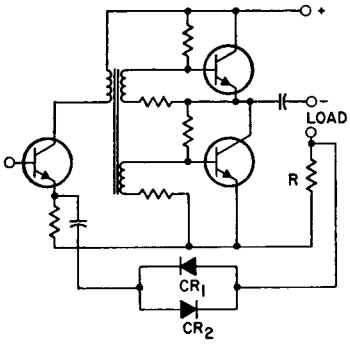



Figure 673. Push-pull power amplifier with short-circuit protection.

margin of the amplifier. This technique in no way affects the normal operation of the amplifier.

A second approach to current limiting is illustrated by the circuit shown in Fig. 674. In this circuit, a diode biasing network is used to establish a fixed current limit on the driver and output transistors. Under sustained short-circuit conditions, however, the output transistors are required to support this current limit and one-half the dc supply voltage.

The circuit shown in Fig. 675 illustrates a dissipation-limiting technique that provides positive protection under all loading conditions. The limiting action of this circuit is shown in Fig. 676. This safe-area limiting technique permits use of low-dissipation driver and output transistors and of smaller heat sinks in the output stages. The use of smaller heat sinks is possible because the worst-case dissipation is normal 4-ohm operation instead of short-circuit conditions. With this technique, highly inductive or capacitive loads are no longer a problem, and thermal cut-outs are unnecessary. In addition, the technique is inexpensive.

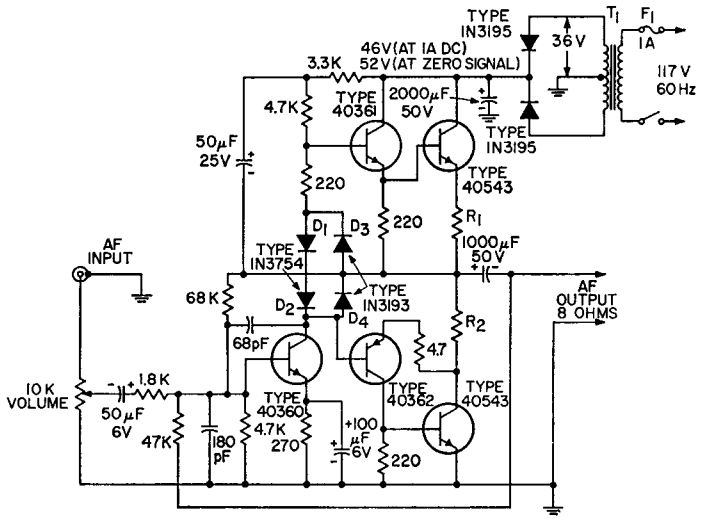



Figure 674. 25-watt (rms) quasi-complementary audio amplifier using current-limiting diodes ($D_3$ and $D_4$).

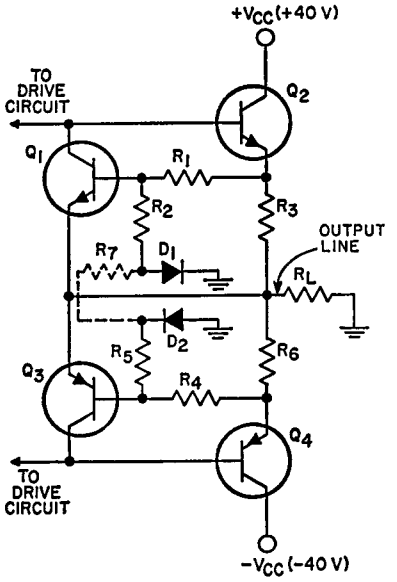



Figure 675. Quasi-complementary audio output in which diode-resistor biasing network is used to prevent complementary transistors $Q_1$ and $Q_2$ from being forward-biased by the output voltage swing.

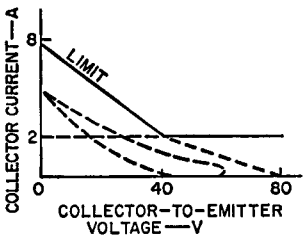



Figure 676. Load lines for the circuit of Fig. 675. Load lines showing effect of the inclusion of high-resistance diode-resistor network in the forward-biasing path of $Q_1$ are shown dotted.

## UNIVERSAL-AMPLIFIER DESIGN APPROACH

A novel approach to the design of high-quality audio power amplifiers is demonstrated by the RCA "Universal Audio-Amplifier Program." In this program, the objective is to devise basic circuit configurations that can be used for a broad family of practical audio amplifiers to cover a wide range of output-power levels. The program has provided two basic amplifier configurations, one a true-complementary-symmetry circuit and the other a quasi-complementary-symmetry circuit, that provide virtually all the information on circuit and component requirements that a manufacturer needs to produce a comprehensive line of predesigned high-quality audio amplifiers for power outputs of 3 to 70 watts. Each amplifier is designed to provide a maximum performance-to-cost ratio.

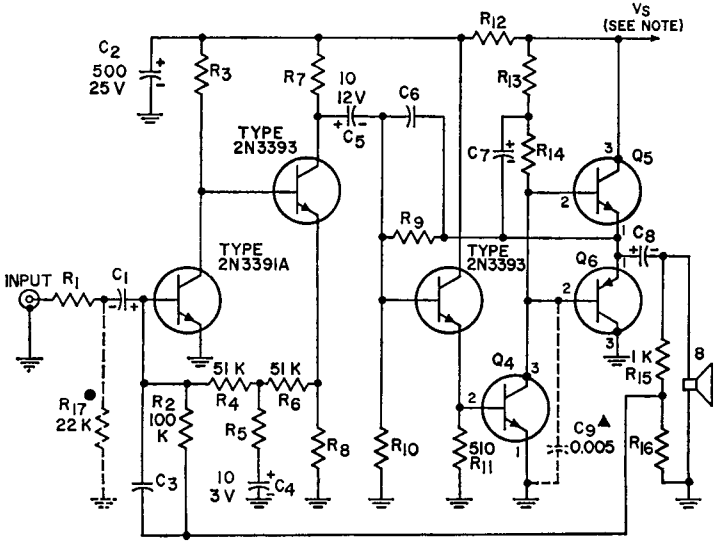
With just minor changes in component values, transistor complement, and supply voltage, the two basic circuit configurations offer the manufacturer of either packaged audio systems or audio components a choice of 13 separate amplifier designs. The need for expensive and time-consuming breadboarding is eliminated, and only two printed-circuit boards are required for all circuits. In addition, the manufacturer may choose between the true-complementary-symmetry and the quasi-complementary circuit.

The basis for the Universal Audio-Amplifier Program is a completely new family of low-cost, high-quality RCA transistors specifically designed for use in audio-amplifier circuits. This family includes a broad complement (27 types) of n-p-n and p-n-p transistors, both silicon and germanium, and features RCA plastic transistors and RCA hometaxial-base types, the industry's best solution to second-breakdown problems.

## Universal True-Complementary-Symmetry Circuit

Fig. 677 shows the basic configuration of a true-complementary circuit that can be used for nine separate audio amplifiers. Three of these amplifiers feature an all-silicon complement of transistors; the other six amplifiers use a combination of silicon and germanium transistors. By selection of the proper combination of component values, transistors, and supply voltage, the basic "universal" configuration can be modified to provide continuous sine-wave audio-power outputs of 3 (two circuits), 5 (two circuits), 7 (two circuits), 12, 16, or 20 watts (rms). For each output-power level, the circuit is designed to provide the full rated output to frequencies well beyond 20 kHz with a total harmonic distortion of 1 per cent.

Component values shown on the circuit schematic and the transistors used in the input and predriver stages remain fixed for all output-power levels. The values of all other resistors and capacitors, the driver and out-




| Note: | $P_{OUT}$ (W) | $V_S$ (V) |
|---|---|---|
| | 3 | 20 |
| | 3 | 20 |
| | 5 | 24 |
| | 5 | 24 |
| | 7 | 29 |
| | 7 | 29 |
| | 12 | 36 |
| | 16 | 40 |
| | 20 | 44 |

All resistances are in ohms and are ½-watt types unless otherwise specified.
All capacitance values are in microfarads unless otherwise specified.
• Used in all 3-, 5-, and 7-watt amplifiers.
▲ Used in the 3-, 5-, and 7-watt amplifiers with all silicon transistors.

Figure 677. Universal true-complementary-symmetry amplifier. This basic amplifier configuration can be used for nine separate audio amplifiers capable of power output of 3 to 20 watts (rms).

put-stage transistors, and the supply voltage must be changed to alter the output-power level. Table XLII lists the specific values of these other resistors and capacitors and indicates the transistor requirements for each amplifier. The supply voltage required for different output-power levels is listed in the voltage chart shown with the circuit schematic in Fig. 677.

For all practical purposes, the universal complementary-symmetry circuit may be considered as a true class B amplifier. No emitter resistors are used in the output stage, and the bases of the complementary pair of output transistors are connected together. If there were no dc feedback current (i.e., current from output voltage divider $R_{I5}$ and $R_{I6}$ back to the base of the predriver transistor $Q_1$ that is required to establish the center voltage of the output stage), both transistors would be reverse-biased and, therefore, cut off. Because of this dc feedback, the n-p-n output transistor $Q_5$ is required to conduct a small amount of current, as established by the output voltage-divider network.

Table XLII—Transistor Requirements and Component Values for Universal True-Complementary-Symmetry Audio-Amplifier Design Package

| POWER OUTPUT watts (rms) | 3 | 3 | 5 | 5 | 7 | 7 | 12 | 16 | 20 |
|---|---|---|---|---|---|---|---|---|---|
| TRANSISTOR COMPLEMENT | (All Si) | (Si/Ge) | (All Si) | (Si/Ge) | (All Si) | (Si/Ge) | (Si/Ge) | (Si/Ge) | (Si/Ge) |
| **Transistor types for driver and output stage:** | | | | | | | | | |
| $Q_4$ (driver) | 40611 | 40611 | 40616 | 40616 | 40616 | 40616 | 40389 | 40625 | 40628 |
| $Q_5$ (n-p-n output) | 40613 | 40610 | 40615 | 40618 | 40621 | 40620 | 40622 | 40624 | 40627 |
| $Q_6$ (p-n-p output) | 40612 | 40609 | 40614 | 40022 | 40022 | 40619 | 40050 | 40623 | 40626 |
| **Capacitance values:** | | | | | | | | | |
| $C_1$ ($\mu$F/V) | 0.1/6 | 0.1/6 | 0.25/6 | 0.25/6 | 0.5/6 | 0.5/6 | 1/6 | 1/6 | 2/6 |
| $C_3$ (pF) | 10 | 10 | 5 | 5 | 5 | 5 | 10 | 22 | 10 |
| $C_6$ (pF) | 100 | 100 | 150 | 150 | 150 | 150 | 220 | 470 | 270 |
| $C_7$ ($\mu$F/V) | 500/12 | 500/12 | 500/12 | 500/12 | 500/12 | 500/12 | 500/15 | 500/15 | 500/15 |
| $C_8$ ($\mu$F/V) | 500/25 | 500/25 | 500/25 | 500/25 | 500/25 | 500/25 | 500/50 | 500/50 | 500/50 |
| $C_9$ ($\mu$F) | 0.005 | — | 0.005 | — | 0.005 | — | — | — | — |
| **Resistance values (all resistors are ½-watt types unless otherwise specified):** | | | | | | | | | |
| $R_1$ (kilohms) | 91 | 91 | 51 | 51 | 27 | 27 | 16 | 10 | 8.2 |
| $R_3$ (kilohms) | 68 | 68 | 68 | 68 | 91 | 91 | 91 | 91 | 91 |
| $R_5$ (kilohms) | 2.7 | 2.7 | 3.3 | 3.3 | 5.1 | 5.1 | 7.5 | 6.8 | 8.2 |
| $R_7$ (kilohms) | 3.9 | 3.9 | 3.9 | 3.9 | 3.9 | 3.9 | 2.7 | 2.7 | 2.2 |
| $R_8$ (ohms) | 620 | 620 | 620 | 620 | 620 | 620 | 390 | 430 | 360 |
| $R_9$ (kilohms) | 33 | 18 | 27 | 12 | 27 | 22 | 18 | 27 | 22 |
| $R_{10}$ (kilohms) | 5.6 | 3 | 3.6 | 1.5 | 3.3 | 2.4 | 1.8 | 1.8 | 1.3 |
| $R_{12}$ (kilohms) | 1 | 1 | 1.8 | 1.8 | 2.7 | 2.7 | 3.3 | 6.8 | 4.7 |
| $R_{13}$ (ohms) | 120 | 120 | 110/1W | 110/1W | 120/1W | 120/1W | 91/2W | 120/2W | 100/2W |
| $R_{14}$ (ohms) | 150 | 150 | 110/1W | 110/1W | 120/1W | 120/1W | 91/2W | 120/2W | 100/2W |
| $R_{16}$ (ohms) | 22 | 22 | 27 | 27 | 47 | 47 | 56 | 100 | 100 |
| $R_{17}$ (ohms) | 22 | 22 | 22 | 22 | 22 | 22 | — | — | — |

This current is at least an order of magnitude less than the idle current normally required in an output stage to reduce crossover distortion. Because the n-p-n output transistor ($Q_5$) is on, the p-n-p output transistor ($Q_6$) is reverse-biased by the amount of the forward base-to-emitter voltage required by the n-p-n transistor to supply the dc feedback current. The dissipation in transistor $Q_5$ produced by the dc feedback current is negligible. In addition, the dc feedback current is inversely proportional to temperature because the base-to-emitter voltage of the driver and predriver transistors decreases with temperature. As a result, the already negligible dissipation decreases as temperature increases.

The use of a complementary output stage operating at zero bias with a single class A driver results in exceptional dc stability in the universal complementary-symmetry amplifier circuit. Each transistor of the complementary output pair ($Q_5$ and $Q_6$) is reverse-biased during the off portion of its operating cycle; common-mode conduction, therefore, is held to a small value. Diodes are not required in the circuit, except in the dc power supply.

The idling current in the driver stage ($Q_3$), which must at least equal the maximum peak base current required by the n-p-n output transistor, is established by the two bias resistors $R_9$ and $R_{10}$. The driver collector current is equal to the difference between the supply voltage and the center voltage (voltage at the common emitter connection of the complementary output transistors) divided by the series resistance ($R_{13} + R_{14}$).

For proper operation of the amplifier, the current through the resistor $R_{14}$ must remain essentially constant during ac excursions of the output voltage. For this purpose, a "bootstrap" capacitor $C_7$ is connected from the junction of resistors $R_{13}$ and $R_{14}$ to the emitters (center point) of the output transistors. Because the voltage across this capacitor does not change during ac output-voltage excursions, the change in voltage is the same at both connection points of the capacitor. The change in voltage at the base of the n-p-n output transistor ($Q_5$) is almost the same as that at the junction of resistors $R_{13}$ and $R_{14}$, and differs only by the small change in the base-to-emitter voltage of transistor $Q_5$. The voltages on both sides of resistor $R_{14}$, therefore, change by essentially the same amount so that the voltage across and the current through this resistor remain essentially constant.

It is important that a reverse base-to-emitter bias voltage be applied to the output transistors during the off half-cycles. When reverse bias is applied to the output transistors during the off half-cycle, they are turned off in a very short time (the stored base-emitter charge is removed rapidly). As a result, high-frequency operation remains in the highly efficient class B mode. When the output transistors are not reverse-biased, but instead are allowed to drift off, the stored charge holds the transistors on after forward drive has been removed. Operation then shifts toward class A push-pull.

Therefore, efficiency is reduced, and current drain, dissipation, and operating temperatures are increased at high frequencies.

Bias for the base of the driver stage is derived from the center point of the output stage. As a result, dc and ac feedback proportional to the center voltage is fed to the base of the driver stage. The actual dc center voltage which the feedback establishes depends on the ratio of resistors $R_9$ and $R_{10}$ and on the base-to-emitter voltage and base current of $Q_3$. If a heavy direct current is bled through $R_7$, changes in the base current of transistor $Q_3$ become insignificant. The dc voltage at the center point is then determined by the base-to-emitter voltage of transistor $Q_3$. Because the percentage of variation in base-to-emitter voltage of a silicon transistor is small, the center-point dc voltage is held close to the desired value. The values of resistors $R_9$ and $R_{11}$ are chosen so that (1) the bleeder current in $R_7$ is large compared to the base current in transistor $Q_3$, (2) the ratio of the resistors provides the desired center-point voltage, and (3) the desired amount of ac feedback current is obtained.

The predriver stages of the universal complementary-symmetry amplifier circuit employ a pair of n-p-n silicon transistors ($Q_1$ and $Q_2$) in a direct-coupled circuit. The feedback from the emitter of transistor $Q_2$ to the base of transistor $Q_1$ serves primarily to hold the dc operating point of transistor $Q_2$ within the limits necessary to prevent a dynamic-range limitation, despite variations in individual transistors or in temperature. The loop feedback resistor $R_2$ also serves as the dc bias resistor from the base of transistor $Q_1$ to ground (the resistor returns to ground through the output voltage-divider resistors $R_{15}$ and $R_{16}$).

For high-frequency stabilization of loop feedback, it is necessary that one element within the loop have relatively poorer high-frequency response than the rest. Because of the local feedback established by resistor $R_9$, the response of the output section is very high. Therefore, the response must be limited somewhere in the front-end section. The collector-to-base feedback capacitance $C_{ob}$ of transistor $Q_1$ is used for this purpose. The high collector load impedance of transistor $Q_1$ causes a high-frequency local feedback current to flow through this capacitance and thus limits the high-frequency response of transistor $Q_1$.

The same rules hold for low-frequency loop-feedback stabilization. It is necessary that one element within the loop have relatively poorer low-frequency response than the rest. The low-frequency response is limited by $C_8$, the speaker coupling capacitor.

Because the value of resistor $R_2$ is established by the dc bias considerations for the front end, the proper amount of ac loop feedback is established by deriving the feedback from a voltage divider across the output. Resistors $R_{13}$ and $R_{14}$ divide the output voltage down to a level which provides the desired feedback current through $R_2$. Resistors $R_{13}$ and $R_{14}$ also serve as an output termination when there is no speaker load.

When feedback is applied to the base of a transistor, the tran-

sistor must be driven by a current source. If a resistor is used to provide a current source, the resistor also isolates the feedback from the rest of the circuit. The resistor $R_1$ serves as a current source for transistor $Q_1$ and prevents the feedback from interacting with the predriver circuit.

Some advantages of the class B mode and the universal complementary-symmetry amplifier design in particular are:

1. Very low dissipation in the output stage at zero signal.
2. Reduced hum and noise at the output. The p-n-p half of the output stage is off and is not turned on by the hum and noise. Only that hum and noise amplified by the n-p-n output transistor appears at the output.
3. Small total harmonic distortion as a direct result of the large amounts of feedback necessary to reduce the crossover distortion. An additional gain stage is required for this amplifier because of the feedback. This additional gain stage is traded for the output-stage bias diode and emitter resistors. Although cross-over distortion is always larger for class B amplifiers than for class AB amplifiers, the class B amplifier can be designed to have an acceptable intermodulation distortion (which is predominantly cross-over distortion).
4. Reduced storage effect in the output transistors. This advantage is inherent in all true-complementary-symmetry amplifiers; when one of the output transistors turns on, it automatically reverse-biases the other transistor and thus pulls the stored charge out of the base region of the off transistor.
5. Reduced power-supply voltage requirements because of the removal of the emitter resistors in the output stage. There are no voltage drops between the power supply and the speaker except the transistors; therefore, the power-supply voltage can be reduced by $2(R_E I_p)$.

Typical performance characteristics for all the amplifiers derived from the basic universal complementary-symmetry circuit configuration are shown in Table XLIII. Fig. 678 shows curves of distortion as a function of power output and output response as a function of frequency for the 20-watt amplifier. With the exception of the power-output level, these curves are also representative of the performance obtained with the other complementary-symmetry amplifiers.

## Universal Quasi-Complementary-Symmetry Circuit

Fig. 679 shows the basic configuration of a quasi-complementary-symmetry circuit that can provide outstanding performance for the most stringent requirements of manufacturers of high-fidelity amplifiers. This basic circuit can be used for four separate audio amplifier circuits that provide continuous sine-wave power ouputs of 12, 25, 40, and 70 watts (rms) with just minor changes in component values, supply voltage, and transistors. Values of some components (shown on the circuit schematic) and the input, predriver, and protection circuit transistors remain the same for all output-power levels. Table XLIV lists the other component requirements for the four circuits. The voltage chart

Table XLIII—Typical Performance Data for Complementary-Symmetry Circuits (3 to 20 Watts)

(Measured at a line voltage of 120V, $T_A = 25°C$, and a frequency of 1 kHz, unless otherwise specified.)

| Circuit Description | Power Output (W) (8-ohm load) Continuous (1% THD, unregulated supply) | Music (5% THD, regulated supply) | Dynamic (1% THD, regulated supply) | Hum and Noise (dB) (Below continuous $P_{OUT}$ Input Shorted | Input Open | Sensitivity (mV) (For continuous $P_{OUT}$) | Input Resistance (kΩ) | 10 dB below continuous $P_{OUT}$ IMD (%) (60 Hz and 7 kHz, 4:1) |
|---|---|---|---|---|---|---|---|---|
| 3-W (All Si) | 3 | 4.5 | 3.6 | 62 | 59 | 110 | 91 | 0.5 |
| 3-W (Si/Ge) | 3 | 4.5 | 3.6 | 62 | 59 | 110 | 91 | 0.5 |
| 5-W (All Si) | 5 | 6.5 | 5.8 | 71 | 69.5 | 100 | 51 | 0.6 |
| 5-W (Si/Ge) | 5 | 6.5 | 5.8 | 71 | 69.5 | 100 | 51 | 0.6 |
| 7-W (All Si) | 7 | 11 | 8 | 71 | 69 | 110 | 27 | 0.8 |
| 7-W (Si/Ge) | 7 | 11 | 8 | 71 | 69 | 110 | 27 | 0.8 |
| 12-W (Si/Ge) | 12 | 15 | 13 | 77 | 75 | 100 | 16 | 0.6 |
| 16-W (Si/Ge) | 16 | 20 | 19 | 84 | 66 | 120 | 10 | 0.2 |
| 20-W (Si/Ge) | 20 | 30 | 26 | 82.3 | 82.3 | 100 | 8.2 | 0.25 |

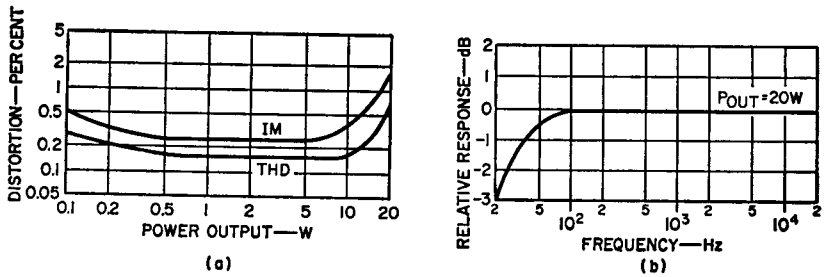
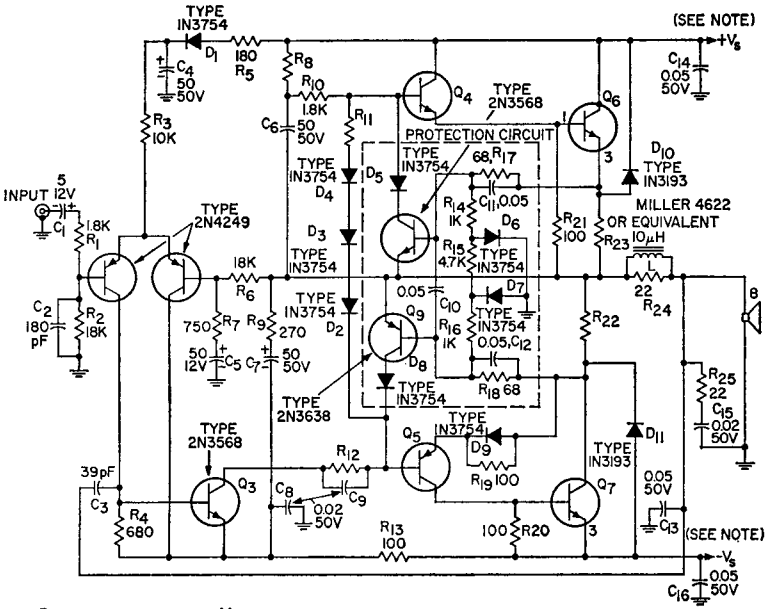



Figure 678. Performance characteristics for the 20-watt complementary-symmetry amplifier: (a) distortion as a function of power output; (b) relative response as a function universal complementary-symmetry design.)

shown with the circuit schematic indicates the supply voltage required for each output-power level.

The universal quasi-complementary-symmetry amplifiers feature rugged hometaxial-base silicon n-p-n output transistors. These transistors and the complementary driver transistors are operated class AB in an arrangement that ensures a small zero-signal current drain. Other features of the circuit include direct-coupled preamplifier and predriver stages and short-circuit protection or safe-area limiting.

The preamplifier stage consists of a balanced-bridge circuit ($Q_1$ and $Q_2$) that maintains quiescent zero dc voltage at the output. Feedback is coupled through resistor $R_6$, and ground reference is provided through resistor $R_2$ and capacitor $C_2$. The common emitters are returned to the positive supply through resistor $R_3$, diode $D_1$, and resistor $R_5$. Diode $D_1$ and capacitor $C_4$ minimize




| Note: $P_{OUT}$ (W) | $V_S$ (±V) |
|---|---|
| 12 | 19 |
| 25 | 26 |
| 40 | 32 |
| 70 | 42 |

All resistances are in ohms and are ½-watt types unless otherwise specified.
All capacitance values are in microfarads unless otherwise specified.

Figure 679. Universal quasi-complementary-symmetry audio-amplifier circuit. This circuit configuration may be used for four separate audio amplifiers capable of power outputs of 12 to 70 watts (rms).

turn-off transients and provide power-supply decoupling. The bridge circuit is direct-coupled to a class A predriver stage ($Q_3$), which is coupled to the complementary drivers ($Q_4$ and $Q_5$) through $R_{12}$. The dissipation-limiting protection circuit is also connected at this point. The purpose of this circuit, as previously indicated, is to prevent the output stage from being driven into conduction if abnormally high dissipation occurs. The dissipation-limiting circuit provides a shunt path for the drive current from the associated driver and output devices. Resistor $R_{12}$ provides some limiting of the current that transistor $Q_9$ must support during overload conditions. Capacitor $C_9$ bypasses $R_{12}$ to improve transient response. Diodes $D_2$, $D_3$, and $D_4$ and resistor $R_{11}$ provide a controlled forward bias on the drivers and output devices so that class AB operation is maintained. The "bootstrap" capacitor $C_6$ supplies the extra voltage swing necessary to saturate the upper output pair ($Q_4$ and $Q_6$) through resistors $R_8$ and $R_{10}$. Capacitor $C_7$ supplies a controlled voltage swing through $R_9$ and $R_{13}$ to overcome the normal losses introduced by resistor $R_{12}$. Resistor $R_{13}$ and capacitor $C_8$ provide high-fre-

Table XLIV—Transistor Requirements and Component Values for Universal Quasi-Complementary-Symmetry Audio-amplifier Design Packages

| POWER OUTPUT watts (rms) | 12 | 25 | 40 | 70 |
|---|---|---|---|---|
| TRANSISTOR COMPLEMENT | All Silicon | All Silicon | All Silicon | All Silicon |
| **Transistor types for driver and output stage:** | | | | |
| $Q_4$ (n-p-n driver) | 2N3568 | 2N3568 | 40635 | 40594 |
| $Q_5$ (p-n-p driver) | 2N3638 | 2N3638 | 40633 | 40595 |
| $Q_6$, $Q_7$ (output) | 40631 | 40632 | 40633 | 40636 |
| **Resistance values (all resistors are ½-watt unless otherwise specified)** | | | | |
| $R_3$ (kilohms) | 10 | 12 | 15 | 18 |
| $R_7$ (ohms) | 750 | 680 | 560 | 470 |
| $R_8$ (ohms) | 1000 | 1800 | 2200 | 2700 |
| $R_{10}$ (ohms) | 1800 | 2200 | 2700 | 3300 |
| $R_{11}$ (ohms) | 40 | 47 | 47 | 47 |
| $R_{12}$ (ohms) | 180 | 270 | 390 | 470 |
| $R_{22}$, $R_{23}$ (ohms) | 0.47/5W | 0.43/5W | 0.39/5W | 0.33/5W |

quency decoupling for the negative dc supply line. Resistors $R_{20}$ and $R_{21}$, with $R_{22}$ and $R_{23}$, provide the necessary stabilization for the output transistors ($Q_6$ and $Q_7$). Current is sampled across resistor $R_{23}$ for positive-cycle sensing and coupled to transistor $Q_8$ through resistor $R_{17}$. Simultaneous voltage sampling is provided by resistor $R_{14}$ and diode $D_6$. Current is sampled across resistor $R_{22}$ for negative-cycle limiting and coupled to transistor $Q_9$ through resistor $R_{18}$. Voltage sampling by resistors $R_{15}$ and $R_{16}$ and diode $D_7$ causes a change in the slope of the limiting characteristics. Resistors $R_{24}$ and $R_{25}$, capacitors $C_{13}$ and $C_{15}$, and inductor $L_1$ provide high-frequency roll-off (above 50 kHz) so that a good margin of stability can be maintained under any loading conditions; capacitors $C_{10}$, $C_{11}$, and $C_{12}$ provide additional stability during limiting. Diodes $D_5$ and $D_8$ prevent forward-biasing of the collector-base junctions of transistors $Q_8$ and $Q_7$ during alternate half-cycle signal swings. Capacitors $C_{14}$ and $C_{16}$ provide parasitic suppression. Diode $D_9$ and resistor $R_{18}$ assure a transconductance match between the upper and lower Darlington pairs to minimize low-level distortion.

Table XLV summarizes the performance characteristics of the four amplifiers that can be designed from the universal quasi-complementary-symmetry design package. Each amplifier provides the full rated power output to frequencies well beyond 20 kHz at a total harmonic distortion of 1 per cent. The curves of distortion as a function of power output and of output response as a function of frequency for the 60-watt amplifier, shown in Fig. 680, are indicative of the performance capabilities of these amplifiers.

Table XLV—Typical Performance Data for Quasi-Complementary-Symmetry Circuits (12 to 70 Watts)

(Measured at a line voltage of 120V, $T_A = 25°C$, and a frequency of 1 kHz, unless otherwise specified.)

| Circuit Description | Power Output (W) (8Ω load) Continuous (1% THD, unregulated supply) | Music (5% THD, regulated supply) | Dynamic (1% THD, regulated supply) | Hum and Noise (dB) (Below continuous $P_{OUT}$ Input Shorted | Input Open | Sensitivity (mV) (For continuous $P_{OUT}$ | Input Resistance (kΩ) | 10 dB below continuous $P_{OUT}$ IMD (%) (60 Hz and 7 kHz, 4:1) |
|---|---|---|---|---|---|---|---|---|
| 12-W (All Si) | 12 | 18 | 16.5 | 75 | 70 | 500 | 20 | 0.2 |
| 25-W (All Si) | 25 | 38 | 33 | 80 | 75 | 600 | 20 | 0.1 |
| 40-W (All Si) | 40 | 55 | 50 | 80 | 75 | 600 | 20 | 0.1 |
| 70-W (All Si) | 70 | 100 | 88 | 85 | 80 | 700 | 20 | 0.1 |

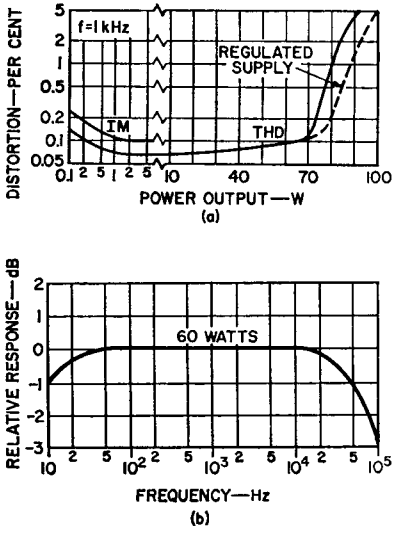



Figure 680. Performance curves for the 70-watt quasi-complementary-symmetry audio amplifier: (a) distortion as a function of power output; (b) relative response as a function of frequency. (These curves are representative of the performance provided by the universal quasi-complementary-symmetry design.)

## BRIDGE-AMPLIFIER DESIGN APPROACH

Fig. 681 shows the block diagram of an audio-amplifier configuration that, for a given dc supply voltage, transistor voltage-breakdown capability, and load, can provide four times the power output obtainable from a conventional push-pull audio-output stage. Alternatively, given power-output and load requirements may be achieved from this circuit configuration with half the supply voltage and transistor voltage-breakdown capabilities required of conventional circuits. This performance is possible because the load can swing the full supply voltage on each half-cycle. The load is direct-coupled between the center point of two series-connected push-pull stages. This bridge type of arrangement eliminates the need for expensive coupling capacitors or transformers. These features are very attractive in applications for which the sup-

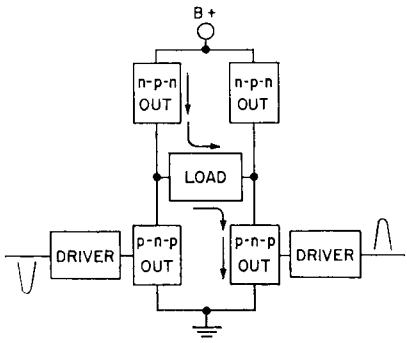



Figure 681. Block diagram of bridge type of audio-amplifier circuit.

ply voltage is fixed, such as automotive or aircraft supplies.

The bridge-amplifier configuration consists essentially of two complementary-symmetry amplifiers with the load direct-coupled between the two center points. Each amplifier section is driven by a class A driver stage that uses a transistor Darlington pair. The amplifiers must be driven 180 degrees out of phase. This dual-phase drive is provided by a differential-amplifier type of input stage, which also provides the advantage of a high input impedance.

Fig. 682 shows the basic configuration of an experimental breadboard circuit designed to evaluate the bridge-amplifier approach to audio-amplifier design. The major difference between this type of circuit and the conventional complementary-symmetry circuit, besides the increased output power, is the higher current requirement of the class A driver stages. This current is twice the value normally required because the peak value of the output current is doubled. The feedback network from each complementary-symmetry output

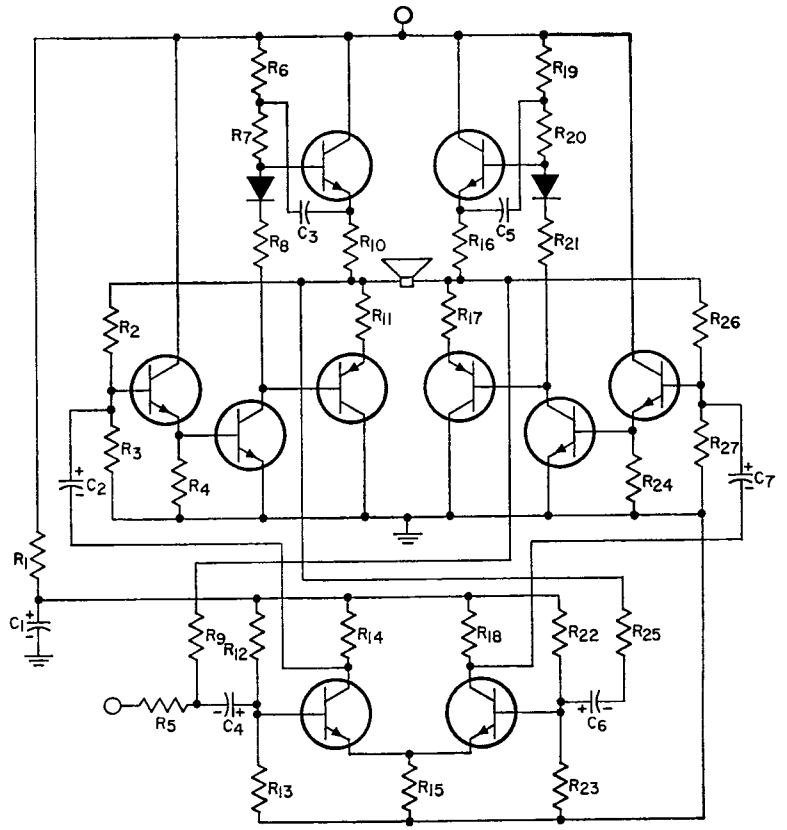



Figure 682. Basic circuit configuration for a bridge type of audio amplifier.

section back to the base of the corresponding class A driver stage, which establishes the center-point voltage in the output stage, also provides a minimum of 22 dB of ac feedback.

The differential-amplifier input stage operates at ten times the required value of peak input current to assure linear operation. Balanced feedback is taken from each side of the load and coupled back to the separate bases of the differential-amplifier transistors. Fig. 683 shows curves of total harmonic distortion as a function of power output for operation of the bridge amplifier with 0 dB, 20 dB, and 28 dB of balanced feedback. Figs. 684 and 685 show total harmonic distortion and relative response as functions of frequency for the bridge amplifier operated with 20 dB of balanced feedback.

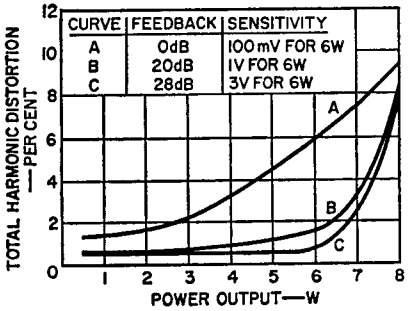



Figure 683. Total harmonic distortion (at 1 kHz) of the bridge audio amplifier as a function of power output for different values of balanced loop feedback. (Distortion performance is comparable to that of a single-ended amplifier that provides one-quarter of the power output for the same dc supply voltage.)

One problem encountered in the bridge amplifier is the achievement of a zero center-point (offset) voltage. The load circuit conducts a direct current proportional to the difference (offset) between the voltage at the center points of the two complementary-symmetry output stages. The dc dissipation in the load circuit is, of course, proportional to the square of the offset voltage. In this breadboard circuit, two potentiometers are used to balance the center-point voltage of the two output-stage sections.

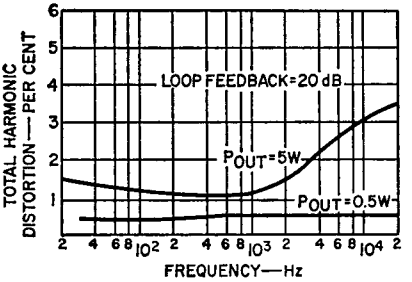



Figure 684. Total harmonic distortion of the bridge audio amplifier as a function of frequency.

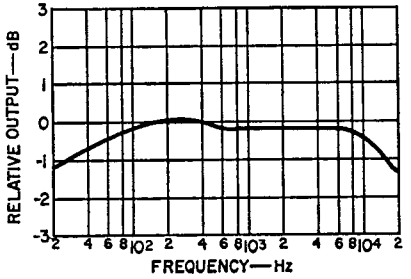



Figure 685. Relative response of the bridge audio amplifier.

## HYBRID-CIRCUIT APPROACH

The RCA-HC1000 hybrid linear power amplifier is a complete, highly versatile amplifier system that is well suited for use as an audio power amplifier, as well as for a wide variety of other applications at frequencies from dc to 50 kHz. This hybrid circuit combines high power-dissipation capabilities, complex circuitry,

and a built-in load protection network in a small, easy-to-handle package that can deliver an rms power output of 100 watts into a 4-ohm load with a peak current of 7 amperes. The HC1000 operates from a total supply voltage of 75 volts, and can deliver 60 watts of power output at a frequency of 30 kHz. The outstanding performance capabilities of the HC1000 are achieved through the use of hybrid construction techniques together with several design innovations that take advantage of previously incompatible processes to extend present technological capabilities. (The internal structure and special features of the HC1000 were described previously in the general discussion of **Power Hybrid Circuits.**)

## Circuit Description

The schematic diagram of the HC1000 hybrid linear power amplifier is shown in Fig. 686. The circuit consists of a differential-amplifier input stage ($Q_1$ through $Q_3$) followed by a bidirectional current source ($Q_4$ and $Q_5$) which drives the class B output stages ($Q_8$ through $Q_{11}$). The bias resistor $R_4$ shunts the input to ground and sets the input impedance at 18000 ohms. The capacitor $C_3$ causes no significant reduction of input impedance at frequencies below 50 kHz. Resistors $R_2$, $R_3$, and $R_4$ provide dc bias for transistor $Q_1$.

The input differential amplifier consists of transistors $Q_1$ and $Q_2$. The input signal is delivered to the base of $Q_1$ (i.e., to the noninverting input terminal of the

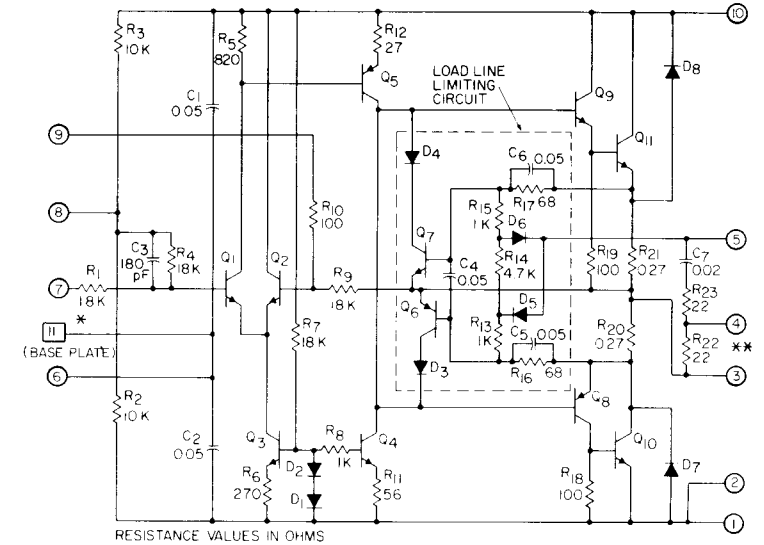



Figure 686. Schematic diagram of the RCA–HC1000 hybrid linear power amplifier.

amplifier). The base of $Q_2$ receives a feedback signal from the output through resistor $R_9$. The constant-current source in the emitter circuit, $Q_3$, permits operation of the differential amplifier to be independent of supply voltage.

The bidirectional current source consists of a current-source transistor, $Q_4$ and a class A amplifier, $Q_5$. Because both the differential amplifier and the bidirectional current source are independent of supply-voltage variations, the complete amplifier can be operated over a supply-voltage range of 30 to 75 volts without bias adjustments. The high output impedance of $Q_4$ and $Q_5$ causes the output transistors to operate on their current-gain characteristics and allows class B operation with minimal crossover distortion.

The output stages use emitter-follower Darlington configurations of the quasi-complementary-symmetry form. The power transistors $Q_{10}$ and $Q_{11}$ are rugged single-diffused hometaxial devices with excellent second-breakdown capabilities. Resistors $R_{20}$ and $R_{21}$ provide stability and sensing points for the short-circuit protection network. Diodes $D_7$ and $D_8$ protect the output transistors from reverse potentials that occur during switching of a transformer-coupled load.

## Typical Audio-Amplifier Connections

Typical connections for the HC1000 for use as an audio amplifier are shown in Fig. 687. Fig. 687(a) illustrates the method for using a split supply with a direct-coupled load, and Figs. 687(b) and 687(c) show a single-ended supply with a capacitive-coupled load. Gain can be adjusted by varying the gain-control resistance $R_G$. The input impedance is

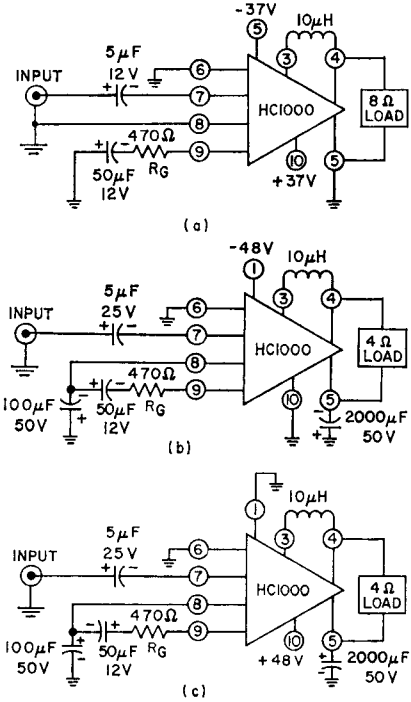



Figure 687. Typical connections for the HC1000 for use as an audio amplifier: (a) with split power supply and direct-coupled load; (b) and (c) with single-ended supply and capacitive-coupled load.

nominally 18000 ohms; the output impedance varies with frequency, as shown in Fig. 688, and can be varied by adjusting $R_G$, as shown in Fig. 689.

When the HC1000 is connected as shown in Fig. 687 and operated at 1 kHz, total harmonic distortion is less than 0.5 per cent at an output of 60 watts and drops to below 0.15 per cent at 1 watt. Low-frequency capability can be improved by increasing the value of the coupling capacitors at the input and in series

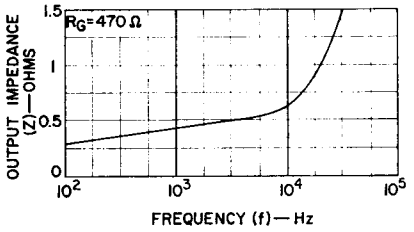



Figure 688. Output impedance as a function of frequency for the circuits of Fig. 687.

with the gain-control resistor. Roll-off at high frequencies depends on power dissipation to a small degree; at a particular power level, however, the frequency must be limited if power dissipation becomes excessive.

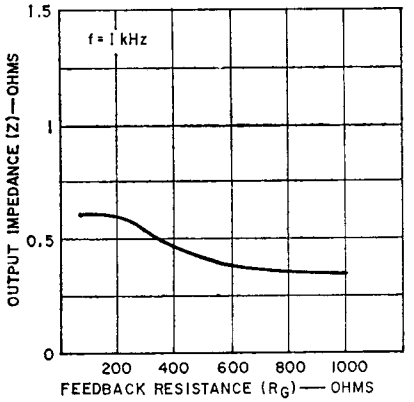



Figure 689. Output impedance as a function of gain-control resistance $R_G$.

## Transformer-Coupled Audio Amplifier

In some applications, it is necessary to deliver considerable power to a high-impedance load, and a transformer must be used, as shown in Fig. 690. When this configuration is used, however, several precautions must be taken.

When the amplifier is in a quiescent mode, the offset voltage (quiescent load voltage), which is typically 75 millivolts, can be as high as 250 millivolts while the dc resistance of the transistor primary is several milliohms. The resulting offset current may be sufficient to activate the short-circuit protection network and cause considerable

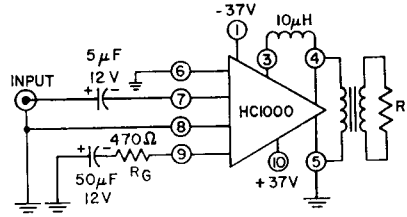



Figure 690. Typical connections for the HC1000 with a split power supply and a transformer-coupled load.

power dissipation in one output device. The following methods can be used to correct this condition:

1. Use a transformer with a high-resistance primary winding.
2. Add resistance in series with the primary.
3. Add a suitable electrolytic capacitor in series with the primary.
4. Balance out the offset voltage by use of the balancing network shown in Fig. 691. The

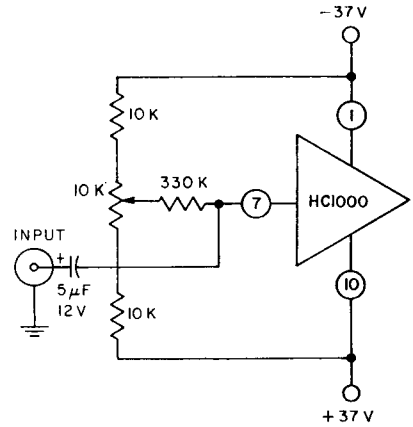



Figure 691. Balancing network for use in the circuit of Fig. 690.

typical temperature coefficient of offset voltage is 0.4 millivolt per °C up to case temperatures of 100° C.

Another problem may be encountered when the amplifier is operated below the low-frequency capability of the transformer. At such frequencies, the transformer presents an inductive load which may activate the protection circuit. The resulting transient condition can distort the waveform, as shown in Fig. 692. The solution to this problem

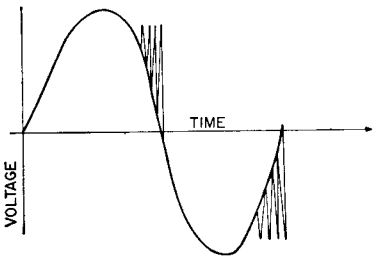



Figure 692. Waveform distortion caused by low-frequency limitations of transformer.

is to design the transformer to be compatible with the lowest frequency at which the amplifier will be used.

## Bridge Amplifiers

Two HC1000 amplifiers can be used in the configuration shown in Fig. 693 to provide amplified outputs in excess of 100 watts. Maximum power output is 200 watts because the effective load voltage is doubled while the maximum load current remains the same. In this circuit, the protection-network terminals are connected to opposite sides of the load instead of to ground to increase the slope of the protection-network characteristic. However, the characteristic lies within the safe operating area of the protection-network transistors for voltages up to ± 23 volts, and the short-circuit protection remains the same as for a single amplifier.

Because the amplifiers must be driven 180 degrees out of phase, one amplifier receives its input signal at the inverting input. Resistances $R_1$ and $R_2$ are selected to provide the proper input impedance of each amplifier for the same voltage gain into a fixed load.

A bridge amplifier with a sin-

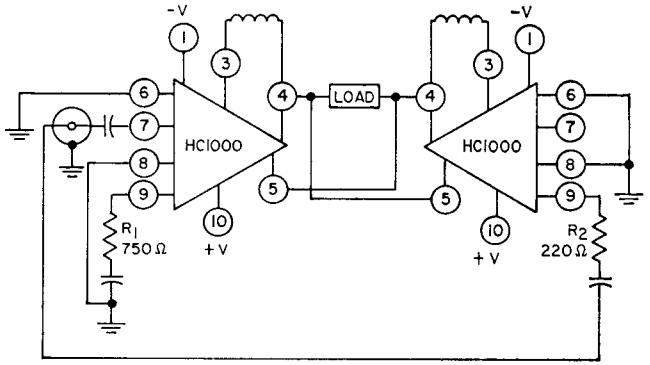



Figure 693. Bridge amplifier using two HC1000 hybrid modules to provide more than 100 watts.

gle-ended supply is shown in Fig. 694. In this configuration, the signal source must be separated from ground, as shown. If it is necessary to ground the generator, the supplies can be isolated. If a grounded load is required, a transformer must be utilized.

Distortion is not appreciably greater in the bridge configurations of Figs. 693 and 694 than in a single amplifier. Measurements with a direct-coupled load yielded total-harmonic-distortion levels below 0.2 per cent at 1 kHz and an output of 40 watts. Power dissipation levels are calculated by use of an apparent load resistance which is half the actual load resistance.

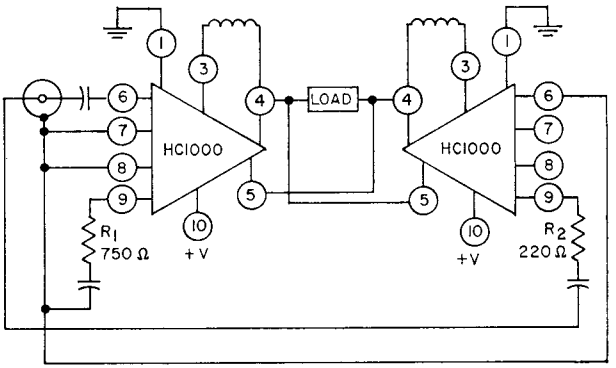



Figure 694. Bridge amplifier using a single-ended power supply.

# Ultrasonic Power Sources

ULTRASONICS is a term applied to the relatively new field of engineering in which high-frequency acoustical energy is used to effect an ultimate improvement in a product or process. The improvement may take place in cleaning, soldering, welding, drilling, defoaming, and degassing, or in control, measurement, detection, and medical diagnostics.

The frequency range used in ultrasonics is typically between 15 kHz and 10 MHz. A few applications employ lower frequencies to achieve maximum particle displacement; at these lower frequencies, however, the power level must be kept low to avoid painful discomfort to those working in the vicinity. In testing applications, higher frequencies are required because the smaller the wavelength, the smaller the flaw that can be detected.

The power level used in ultrasonic engineering depends upon the application. Large-scale industrial-cleaning operations may require many kilowatts, while measuring and testing applications may require only a few microwatts. Table XLVI lists some of the general industrial applications of ultrasonics, together with a brief description of the various applications and the typical power level and frequency required for each.

## CHARACTERISTICS OF ULTRASONIC TRANSDUCERS

Many devices can be used to produce ultrasonic energy; these devices are called transducers. All transducers can be classified in one of three groups: mechanical, magnetostrictive, or electrostrictive. Mechanical transducers are applied for the most part to the production of acoustic and ultrasonic oscillations in air or other gaseous media. Mechanical transducers used as sources of ultrasonic waves in air include whistles, gas-jet generators, and sirens. The power sources used in these devices usually incorporate a type of pressurized gas or fluid. The gas and liquid transducers convert a steady mechanical force into a vibratory mechanical force.

In solids, however, the same effect is not possible. In this case a source of electrical energy at the required operating frequency

Table XLVI—Ultrasonic Applications

| APPLICATION | DESCRIPTION | POWER RANGE (Watts) | FREQUENCY RANGE (kHz) |
|---|---|---|---|
| Ultrasonic cleaning and degreasing. | Cavitated cleaning solution scrubs parts immersed in solution. | 50 to 25,000 (Typically 100 watts per gallon of solution). | 20 to 40 |
| Drilling, cutting, and polishing of hard and brittle materials. | Abrasive slurry between vibrating tool and work piece cuts into material. | 50 to 2,000 | 16 to 30 |
| Soldering and brazing. | Ultrasonically vibrating solder removes oxide film eliminating the need for flux. | 0.5 to 250 | 16 to 30 |
| Welding metals and plastics. | Vibrating tool generates high temperature at interface of the two materials. | 10 to 1,000 | 16 to 30 |
| Emulsification, dispersion, and homogenization. | Mixing and homogenizing of liquids, slurries and creams. | 100 to 2,000 | 16 to 1,000 |
| Control and measurement, alarm systems, counting. | Interruption or deflection of beam, damping of transducer. | 0.1 to 50 | 16 to 45 |
| Flaw detection. | Determination of size and location of flaws in solids by the pulse-echo technique. | 0.5 to 20 | 1,000 to 10,000 |
| Medical: surgery and diagnostics. | Ultrasonic surgical knife cuts through tissue. Locating tumors and other flaws using the pulse-echo technique. | 1 to 1,000 | 100 to 10,000 |

is converted into a vibrating mechanical force. This conversion is accomplished through the use of special materials which have magnetostrictive or electrostrictive properties.

**Magnetostriction** is the name applied to the change in length of a magnetic material under the influence of an external magnetic field. Whether a magnetic material (such as iron, nickel, cobalt, or a magnetic alloy) lengthens or shortens depends on a property of the material and is not dependent on the direction of the magnetic field. Fig. 695 shows the strain (change in length per unit length) as a function of magnetic field strength for several magnetostrictive materials. It can be seen that nickel gets shorter as the magnetic

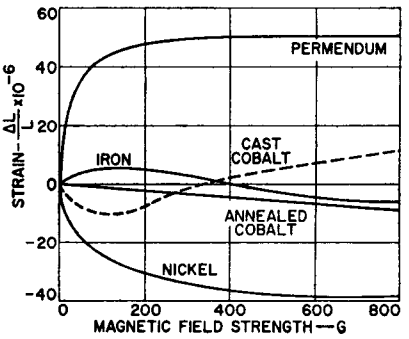



Figure 695. Strain as a function of magnetic field strength for several magnetostrictive materials.

field is increased, while Permendum gets longer. Fig. 696 shows how a bar of material that has a positive strain coefficient (lengthens with increased magnetic field) would react to an alternating magnetic field with no static biasing field. It can be seen that the bar vibrates at twice the generator frequency and that the amplitude is $\Delta L$ peak to peak.

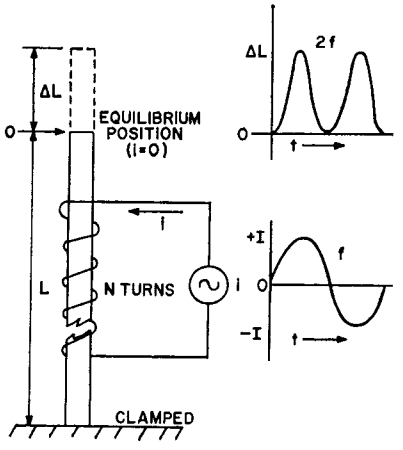



Figure 696. Reaction of a bar of material that has a positive strain coefficient to an alternating magnetic field when no static biasing field is used. Waveforms show change in length of bar (top) and alternating current (bottom) used to produce the magnetic field.

Fig. 697 shows the effect of adding a static biasing magnetic field. This bias could also be supplied by a permanent magnet. The dc bias field yields an initial displacement $\Delta L$. Under these conditions, the bar oscillates about its equilibrium position at the frequency of the generator with a peak-to-peak amplitude of $2\Delta L$.

The **piezoelectric effect** is a phenomenon that occurs in certain crystals; the crystals are deformed when subjected to an electric field. The converse is also true; i.e., if the crystal

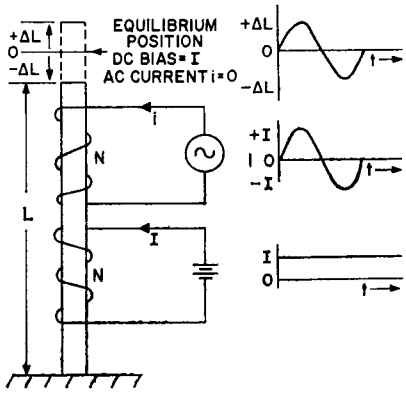



Figure 697. Reaction of bar of material that has a positive strain coefficient to an alternating magnetic field when static biasing is employed. Waveforms show change in length of bar (top), alternating current used to produce alternating field (center), and direct current (bottom) used to produce the bias field.

(quartz, Rochelle salt, barium titanate) is strained, an electric charge appears at its edges.

The piezoelectric effect in the first mode is used in the generation of high-frequency sound waves. This effect is accomplished by application of an alternating voltage of the desired frequency to the crystal. Fig. 698 shows an example of this method.

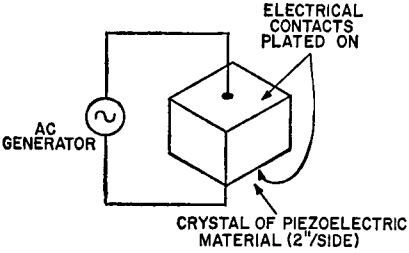



Figure 698. Application of an alternating voltage to a piezoelectric crystal to produce high-frequency sound waves.

In the design of equipment that uses electromechanical transducers, a useful equivalent circuit

for the transducer must be available. Fig. 699(a) shows the equivalent of a magnetostrictive transducer in which $Z_A$, $Z_B$, and N depend upon the magnetic and physical properties of the core material. Fig. 699(b) is an approximate equivalent circuit for the transducer. The reactive component of the input impedance is attributed primarily to the inductance of the winding. This induc-

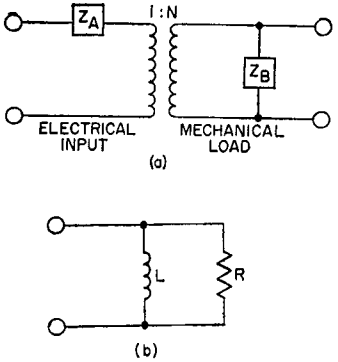



Figure 699. (a) Actual equivalent circuit and (b) simplified approximation of a magnetostrictive transducer.

tance is a function of the number of turns and the transducer core material. The resistance R represents the mechanical load. To obtain mechanical energy, it is necessary to provide electrical power to this resistance. Because magnetostrictive transducers usually operate with a static bias field, a dc component of current must be supplied to the transducer. Fig. 700 shows a typical circuit.

In the circuit, the choke is used to prevent the high-frequency signal from shorting through the low-impedance dc supply. The capacitor C is required to prevent dc from flowing through the generator. In addition, the value of C can be chosen so that the induc-

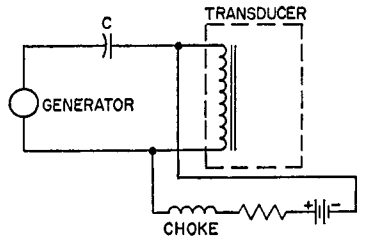



Figure 700. Circuit showing application of electrical power to a magnetostrictive transducer.

tive reactance of the transducer is cancelled.

Fig. 701(a) is the equivalent circuit for a piezoelectric crystal; $Z_A$, $Z_B$, and N are functions of the electrical and physical properties of the crystal. Fig. 701(b) shows the approximate equivalent circuit used to represent a piezoelectric transducer for the purpose of making calculations. The capacitance is usually tuned out by use of either a parallel or series inductor in the matching circuit between the generator and transducer.

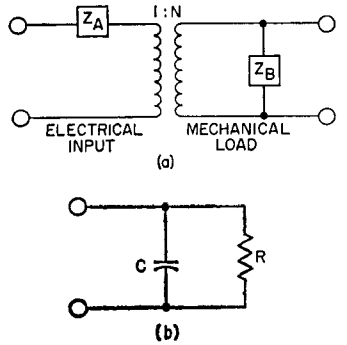



Figure 701. (a) Actual equivalent circuit and (b) simplified approximation of a piezoelectric crystal.

## ULTRASONIC GENERATORS

The majority of ultrasonic applications employ a continuously oscillating power source. In fact,

the only application listed in Table XLVI that does not make use of a continuous wave is flaw detection by the pulse-echo technique. For this reason, the following discussion of ultrasonic power sources is limited to the continuous-wave type. Table XLVI shows that most of the frequencies and power levels required are such that transistors can be used in the power generators. Therefore, the power sources discussed below are of the solid-state type.

The waveform delivered to the transducer can be of the square or sinusoidal type. As a result, there are four basic methods of power generation:

1. a low-power square-wave inverter followed by a class B push-pull power amplifier,
2. a square-wave power inverter that drives the load directly,
3. a low-power sine-wave oscillator followed by a class B push-pull amplifier,
4. a self-oscillating power amplifier that drives the load directly.

The detailed explanation of circuit operation and design procedures for each of these circuits is given in other parts of this Handbook.

If the transducer used can operate with a square-wave power source, then an inverter should be used because it affords very high efficiency. However, if the electromechanical transducer is required to deliver sinusoidal power to its load (cleaning solution, abrasive slurry, and the like), sinusoidal electrical power must be delivered to the resistor representing the load in the equivalent circuit of the transducer.

## Inverter Circuits

Fig. 702 shows one method of obtaining a voltage sine wave across $R_L$. In this circuit, the generator supplies a square-wave voltage; the matching network filters out the harmonics so that only the fundamental component remains. The matching network includes the reactive component of the transducer as a shunt inductor or capacitor, depending upon whether the transducer is of the magnetostrictive or electrostrictive type. In other words, the reactive component of the transducer is used as part of the filter.

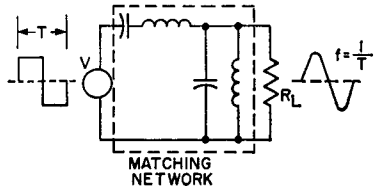



Figure 702. Use of a transducer and resonant matching network to convert a square-wave input to a sinusoidal output. Reactive component of transducer is used as the shunt inductor or capacitor of the matching network depending upon whether a magnetostrictive or electrostrictive type of transducer is used.

With this type of network, a transistorized inverter can be used to drive the transducer. The Q of the series tuned matching circuit should be at least 5.

The simplicity of this type of system is shown in Fig. 703. In the push-pull inverter with a series tuned load, each transistor provides current half of the time. The current flows only during the time that the transistor collector-to-emitter voltage is near zero [$V_{CE}(sat)$]. During the half-cycle when the voltage across the transistor is equal to $2V_{CC}$, there is no current flow. During both half-cycles, the dissipation in the device is essentially zero. Theoreti-

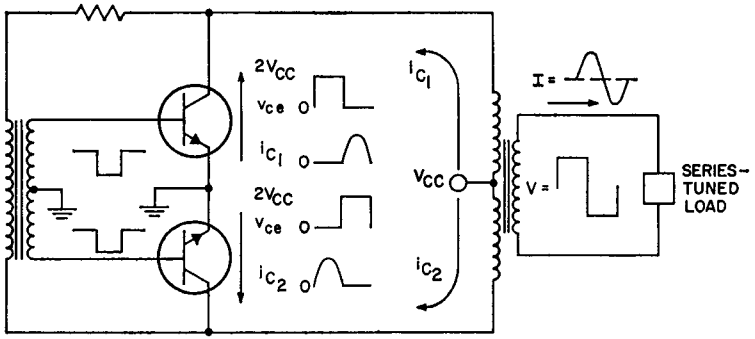



Figure 703. Use of a push-pull switching inverter to drive a transducer that forms part of a series-tuned load circuit.

cally, then, the efficiency could approach 100 per cent. A thorough analysis and detailed design procedure for inverters is given in the section on **Power Conversion.**

## Class C Oscillators

One disadvantage of the inverter approach is that the fundamental frequency is determined by the feedback network. Any time there is a change in the reactance of the load, its resonant frequency changes and the operating frequency of the inverter must be adjusted to the new resonant frequency. If the frequency is not adjusted, the power delivered to the load decreases and the power dissipated in the transistor increases. With most practical transducers, the reactive component is continually changing.

One method used to overcome this problem is to let the load determine the frequency by use of a tuned-load class C oscillator, such as that shown in Fig. 704. With this arrangement, the operating frequency is always the resonant frequency of the load.

Fig. 705 shows that the class C oscillator provides a pulse of current to the load. The load is parallel tuned; the voltage across the load, therefore, is sinusoidal. The period (T) of the current pulse is

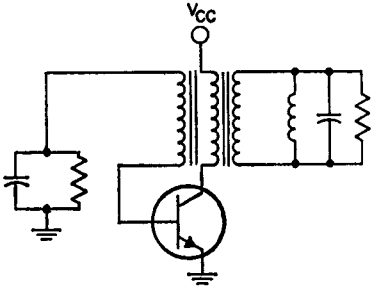



Figure 704. Class C oscillator that operates into a tuned load circuit.

equal to the reciprocal of the resonant frequency $f_r$ of the load. Therefore, if $f_r$ changes, there is a corresponding change in T. Fig. 706 shows the collector voltage and collector current for the class C oscillator.

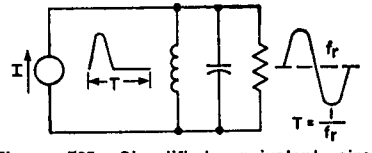



Figure 705. Simplified equivalent circuit for the class C oscillator shown in Fig. 704.

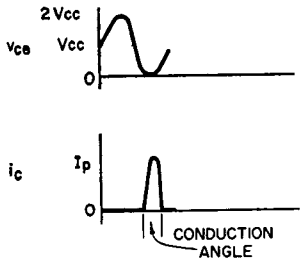



Figure 706. Collector voltage and current waveforms for the class C oscillator shown in Fig. 704.

The magnitude of the collector-current pulse is determined by the load power. The current peak occurs at $V_{CE}(sat)$, which is approximately zero. As the conduction time of $i_C$ is made smaller, the efficiency increases; however, $i_C$ must also increase to maintain the same power output. In the limit, an infinite pulse of zero width would yield 100-per-cent efficiency. However, this limit would require an infinite circuit Q. It can easily be shown that, for a fixed $V_{CC}$, the power output is proportional to the area under the current pulse shown in Fig. 706, where the area is determined by the magnitude and conduction angle of the current pulse. The maximum value of $i_C$ is limited by the maximum current rating of the transistor used. The maximum power output [for a given $V_{CC}$ and $I_{C(max)}$], therefore, is proportional to the conduction angle. However, because the efficiency is inversely proportional to the conduction angle, it is obvious that some sort of compromise must be made. The following examples should help to determine the best compromise:

**Example No. 1**—In class C oscillators, the maximum collector voltage rises to a value equal to twice the supply voltage [i.e., $V_{CE(max)} = 2V_{CC}$], as indicated in Fig. 707. This condition occurs when the transistor is reverse-biased. The $V_{CEV}(sus)$ rating of the transistor used, therefore, should be equal to, or greater than, $2V_{CC}$. The relationship between dc input power $P_s$, power delivered to the load $P_L$, transistor dissipation $P_d$, and circuit efficiency $\eta$ can be calculated for a typical transistor operated in a circuit of this type.

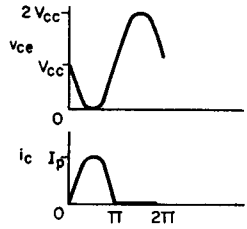



Figure 707. Collector voltage and current waveforms for an oscillator circuit that has a conduction angle of 180 degrees.

The parameters assumed for the transistor are as follows:

$V_{CEV}(sus) = 100$ volts
$I_C(max) = 20$ amperes
$T_d(max) = 200°C$
$TR_{J-C}$ (includes heat sink) $= 3°C/W$
$T_A = 80°C$ (ambient)

For these parameters, $P_d$ should not exceed $(200 - 80)/3$, or 40 watts. For $V_{CC} = 100/2 = 50$ volts, $I_p = I_C(max) = 20$ amperes, and the conduction angle $\theta = \pi$ (maximum power output), the quantities $P_s$, $P_L$, $P_d$, and $\eta$ are calculated as follows:

$$P_s = \frac{1}{2\pi}\int_0^{2\pi} V_{CC}\, i_C\, d\theta$$

$$= \frac{V_{CC}}{2\pi}\int_0^{\pi} I_p \sin\theta\, d\theta$$

$$= \left[\frac{V_{CC}\, I_p}{2\pi}(-\cos\theta)\right]_0^{\pi}$$

$$= \frac{V_{CC}\, I_p}{2\pi}\,(1+1)$$

$$= \frac{V_{CC}\, I_p}{\pi} = 0.317\, V_{CC}\, I_p = 320 \text{ watts}$$

$$P_L = \frac{1}{2\pi}\int_0^{\pi} V_{CC} \sin\theta\, I_p \sin\theta\, d\theta$$

$$= \frac{V_{CC}\, I_p}{2\pi}\int_0^{\pi} \sin^2\theta\, d\theta = \frac{V_{CC}\, I_p}{4}$$

$$= 0.25\, V_{CC}\, I_p = 250 \text{ watts}$$

$$P_d = P_s - P_L = 0.067\, V_{CC}\, I_p$$

$$= 70 \text{ watts}$$

$$\eta = P_L/P_s = 78\%$$

The calculated value for the transistor dissipation ($P_d$ = 70 watts) exceeds the maximum allowable value (40 watts). This condition indicates the value calculated for the maximum power output ($P_L$ = 250 watts) cannot be obtained because of thermal limitations.

**Example No. 2**—If the conditions $V_{CC}$ = 50 volts and $\theta = \pi$ are maintained, then the efficiency $\eta$ is still 78 per cent. The peak current $I_p$, therefore, must be reduced so that the transistor dissipation $P_d$ does not exceed 40 watts. (The same heat sink and thermal temperature used in example No. 1 are assumed.) The new value of $I_p$ is calculated as follows:

$$P_d = 0.067\, V_{CC}\, I_p = 40 \text{ watts}$$

$$I_p = 40/(0.067 \times 50) = 11.5 \text{ amperes}$$

The power delivered to the load $P_L$ then becomes

$$P_L = (0.25)(50)(11.5) = 142 \text{ watts}$$

Although the transistor current is only slightly more than one-half the maximum current rating, the dissipation is equal to the maximum allowable value under the given conditions. In other words, the junction temperature is at its maximum rating.

**Example No. 3**—If the conduction angle is decreased to ⅓ of the cycle (i.e., $\theta = 2\pi/3 = 120°$), the transistor dissipation is substantially reduced. Fig. 708 shows the collector current and voltage waveforms for this condition. If

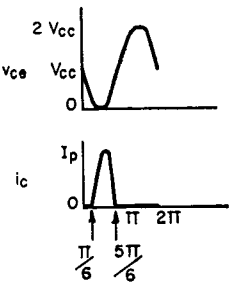



Figure 708. Collector voltage and current waveforms for an oscillator circuit that has a conduction angle of 120 degrees.

all other conditions are assumed to be the same as for example No. 1, the dc input power, load power, transistor dissipation, and efficiency are calculated as follows:

$$P_s = \frac{1}{2\pi}\int_{\pi/6}^{5\pi/6} V_{CC}\, I_p \sin\frac{3}{2}\theta\, d\theta$$

$$= \left[\frac{V_{CC}\, I_p}{2\pi}\left(-\frac{2}{3}\cos\frac{3}{2}\theta\right)\right]_{\pi/6}^{5\pi/6}$$

$$= \frac{-V_{CC}\, I_p}{3\pi}\left[\cos\frac{3}{2}\left(\frac{5\pi}{6}\right) - \cos\frac{3}{2}\left(\frac{\pi}{6}\right)\right]$$

$$= \frac{V_{CC}\, I_p}{3\pi}\,(2)\left(-\frac{1}{\sqrt{2}}\right)$$

$$= 0.15\, V_{CC}\, I_p = 150 \text{ watts}$$

$$P_L = \frac{1}{2\pi}\int_{\pi/6}^{5\pi/6} V_{CC} \sin\theta\, I_p \sin\frac{3}{2}\theta\, d\theta$$

$$= \frac{V_{CC}\, I_p}{2\pi}\int_{\pi/6}^{5\pi/6} \sin\theta \sin\frac{3}{2}\theta\, d\theta$$

$$= \frac{V_{CC}\, I_p}{2\pi}\left[\frac{\sin\left(\frac{3}{2}-1\right)}{2\left(\frac{3}{2}-1\right)} - \frac{\sin\left(\frac{3}{2}+1\right)}{2\left(\frac{3}{2}+1\right)}\right]_{\pi/6}^{5\pi/6} = \frac{V_{CC}\, I_p}{2\pi}$$

$$\left[\sin\frac{1}{2}\theta - \frac{\sin\frac{5}{2}\theta}{5}\right]_{9/\pi}^{5\pi/6}$$

$$= \frac{V_{CC}\, I_p}{2\pi}\,(0.966-0.05-0.26+0.193)$$

$$= \frac{0.85}{2\pi}\, V_{CC}\, I_p = 0.135\, V_{CC} I_p$$

$$= 35 \text{ watts}$$

$$P_d = P_s - P_L = 0.015\, V_{CC}\, I_p$$

$$= 15 \text{ watts}$$

$$\eta = P_L/P_s = 90 \text{ per cent}$$

For a conduction angle of one-third of a cycle, therefore, the transistor is not limited by power dissipation under the conditions stated. The transistor can operate at full voltage and current ratings. If the heat sink used in examples Nos. 1 and 2 is employed, the junction temperature is maintained well below the rated level.

**Example No. 4**—The design of a practical class C oscillator which has a conduction angle $\theta$ of 120° and an over-all circuit efficiency $\eta$ of about 80 per cent is illustrated by the following example:

The design conditions are as follows:

$V_{CC}$ = 50 volts; $P_L$ = 125 watts
$R_L$ = 1000 ohms in parallel with a 0.005-microfarad capacitor
f = 25 kHz
$TR_{HS}$ = 2°C/W
$T_A$ = 80°C
$\Theta = 2\pi/3$

For these conditions, the following values are calculated:

$P_L = (0.135)(V_{CC})(I_p)$
$125 = (0.135)(50)(I_p)$
$I_p = 18.5$ amperes
$P_d = (0.015)(50)(18.5) = 14$ watts

The Q of the load circuit, which is equivalent to $R_L/2\pi fL$ for a parallel tuned network, is 2.5. The value of the load-circuit inductance L, therefore, may be calculated as follows:

$$L = 1000/(2\pi)(25)(10^3)(2.5)$$
$$= 2.5 \text{ millihenries}$$

The load-circuit capacitance then is determined as follows:

$2\pi f = 1/(LC)^{\frac{1}{2}}$
C = 0.01 microfarad

Because the load resistance $R_L$ is shunted by a capacitance of 0.005 microfarad, the actual value of the capacitor used in the output tuned circuit is 0.015 − 0.005, or 0.01 microfarad.

The transistor requirements are as follows:

$V_{CEV}(sus) \geq 2\, V_{CC} = 100$ volts
$I_C(max) \geq 18.5$ amperes
$P_d(max) \geq 14$ watts at $T_C = 108°C$ [80°C ambient + (14)(2°C/W)]

Therefore, the thermal resistance from junction to case $\theta_{J\text{-}C} \leq 7°/$ watt.

Information on the selection of core size and material is given in the section on inverters. For this design, a toroid of linear material (Arnold Engineering No. A438381-2 or equivalent) is used. Use of 100 turns of No. 24 wire for the secondary winding provides 2.7 millihenries of open-circuit inductance. This secondary provides the inductance of the matching network.

The power output $P_L$ is equal to 125 watts, and the load resistance $R_L$ is equal to 1000 ohms. The peak voltage across the load, therefore, is 500 volts. The transformer turns ratio then becomes

$$N = 500/50 = 10:1$$

Ten turns of No. 22 wire, therefore, are required for the primary. The remainder of the circuit design procedure is covered under the design of class C oscillators in the section on **RF Power Amplifiers.** Fig. 709 shows the schematic diagram of the completed circuit, and Fig. 710 shows the circuit waveforms.

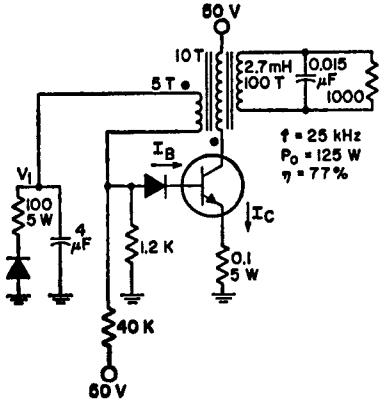



Figure 709. 125-watt, 25-kHz, class C oscillator.

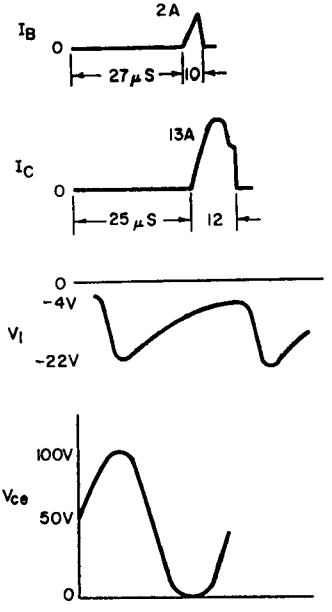



Figure 710. Current and voltage waveforms for the class C oscillator shown in Fig. 709.

## ULTRASONIC POWER AMPLIFIERS

In general, the power amplifiers used to drive ultrasonic transducers are the same as those used to drive the loudspeakers in audio-amplifier applications. The basic design considerations and circuit configurations described in the section on **Audio Power Amplifiers** are applicable, therefore, to the design of power amplifiers for ultrasonic applications. The frequency range of the basic amplifier configurations can be readily extended into the range of 10 kHz to 100 kHz normally used in ultrasonic systems by selection of higher-frequency power transistors, use of smaller inductive and capacitive coupling components, and a proper choice of values for feedback elements.

The RCA line of linear high-

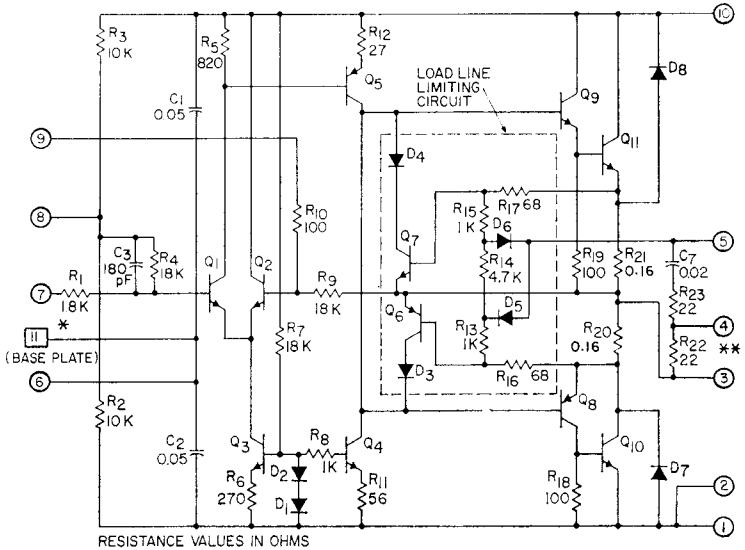



Figure 711. Schematic diagram of the RCA–HC2000 power hybrid operational-amplifier circuit.

current hybrid amplifiers includes a versatile 7-ampere operational-amplifier circuit that is ideally suited for use in applications that require large amounts of power amplification at ultrasonic frequencies. The basic configuration of this amplifier, shown in Fig. 711, is similar to the quasi-complementary-symmetry universal-amplifier configuration described in the section on **Audio Power Amplifiers.**

In the hybrid-circuit amplifier, medium-frequency multiple-diffused power-transistor chips are used in the quasi-complementary-symmetry class B output stage to develop up to 60 watts of average power output at frequencies from 10 kHz to 100 kHz. The amplifier operates from either single-ended or split dc power supplies at dc supply voltages from 30 to 75 volts. Fig. 712 shows the circuit connections for use of the 60-watt hybrid-circuit amplifier to drive an ultrasonic transducer.

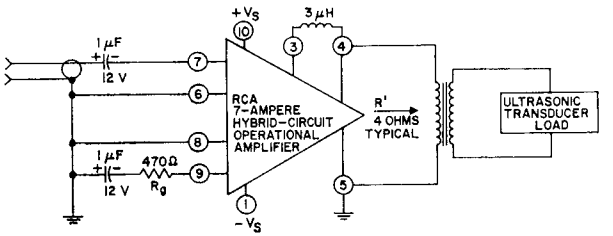



Figure 712. RCA–HC2000 operational amplifier used to drive an ultrasonic transducer.

# TV Deflection Systems

FOR reproduction of a transmitted picture in a television receiver, the face of a cathode-ray tube is scanned with an electron beam while the intensity of the beam is varied to control the emitted light at the phosphor screen. The scanning is synchronized with a scanned image at the TV transmitter, and the black-through-white picture areas of the scanned image are converted into an electrical signal that controls the intensity of the electron beam in the picture tube at the receiver.

## SCANNING FUNDAMENTALS

The scanning procedure used in the United States employs horizontal linear scanning in an odd-line interlaced pattern. The standard scanning pattern for television systems includes a total of 525 horizontal scanning lines in a rectangular frame having an aspect ratio of 4 to 3. The frames are repeated at a rate of 30 per second, with two fields interlaced in each frame. The first field in each frame consists of all odd-number scanning lines, and the second field in each frame consists of all even-number scanning lines. The field repetition rate is thus 60 per second, and the vertical scanning rate is 60 Hz. (For color systems, the vertical scanning rate is 59.94 Hz.)

The geometry of the standard odd-line interlaced scanning pattern is illustrated in Fig. 713. The scanning beam starts at the upper left corner of the frame at point A, and sweeps across the frame with uniform velocity to cover all the picture elements in one horizontal line. At the end of each trace, the beam is rapidly returned to the left side of the frame, as shown by the dashed speed. The slope of the horizontal-line. The horizontal lines slope downward in the direction of scanning because the vertical deflecting signal simultaneously produces a vertical scanning motion, which is very slow compared with the horizontal scanning speed. The slope of the horizontal-line trace from left to right is greater than the slope of the retrace from right to left because the shorter time of the retrace does not allow as much time for vertical deflection of the beam.

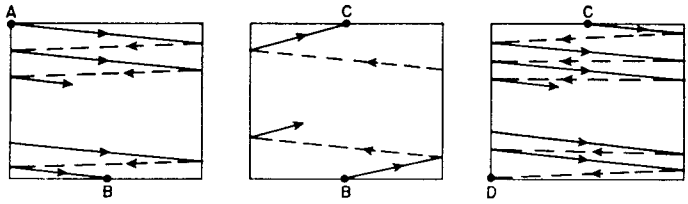



Figure 713. The odd-line interlaced scanning procedure.

Thus, the beam is continuously and slowly deflected downward as it scans the horizontal lines, and its position is successively lower as the horizontal scanning proceeds.

At the bottom of the field, the vertical retrace begins, and the beam is brought back to the top of the frame to begin the second or even-number field. The vertical "flyback" time is very fast compared to the trace, but is slow compared to the horizontal scanning speed; therefore, some horizontal lines are produced during the vertical flyback.

All odd-number fields begin at point A in Fig. 713 and are the same. All even-number fields begin at point C and are the same. Because the beginning of the even-field scanning at C is on the same horizontal level as A, with a separation of one-half line, and the slope of all lines is the same, the even-number lines in the even fields fall exactly between the odd-number lines in the odd field.

## Sync Pulses

In addition to picture information, the composite video signal from the video detector of a television receiver contains timing pulses to assure that the picture is produced on the faceplate of the picture tube at the right instant and in the right location. These pulses, which are called sync pulses, control the horizontal and vertical scanning generators of the receiver.

Fig. 714 shows a portion of the detected video signal. When the picture is bright, the amplitude of the signal is low. Successively deeper grays are represented by higher amplitudes until, at the "blanking level" shown in the diagram, the amplitude represents a complete absence of light. This "black level" is held constant at a value equal to 75 per cent of the maximum amplitude of the signal during transmission. The remaining 25 per cent of the signal amplitude is used for syn-

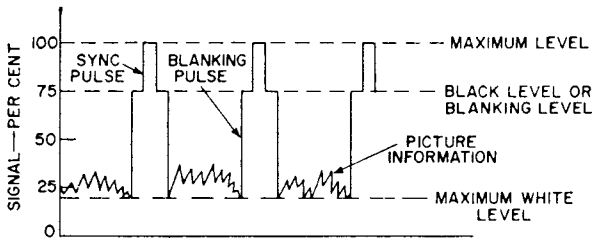



Figure 714. Detected video signal.

chronization information. Portions of the signal in this region (above the black level) cannot produce light.

In the transmission of a television picture, the camera becomes inactive at the conclusion of each horizontal line and no picture information is transmitted while the scanning beam is retracing to the beginning of the next line. The scanning beam of the receiver is maintained at the black level during this retrace interval by means of the blanking pulse shown in Fig. 714. Immediately after the beginning of the blanking period, the signal amplitude rises further above the black level to provide a horizontal-synchronization pulse that initiates the action of the horizontal scanning generator. When the bottom line of the picture is reached, a similar vertical-synchronization pulse initiates the action of the vertical scanning generator to move the scanning spot back to the top of the pattern.

## Sync Separation

The sync pulses in the composite video signal are separated from the picture information in a **sync-separator stage,** as shown in Figs. 715 and 716. This stage is biased sufficiently beyond cutoff so that current flows and an output signal is produced only at the peak positive swing of the input signal. In the diode circuit of Fig. 715, negative bias for the diode is developed by R and C as a result of the flow of the diode current on the positive extreme of signal input. The bias automatically adjusts itself so that the peak positive swing of the input signal drives the anode of the diode positive and allows the flow of current only for the sync pulse. In the circuit shown in Fig. 716, the base-emitter junction of the transistor functions in the same manner as the diode in Fig. 715, but in addition the pulses are amplified.

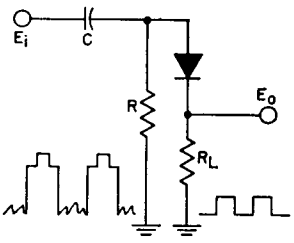



Figure 715. Diode sync-separator circuit.

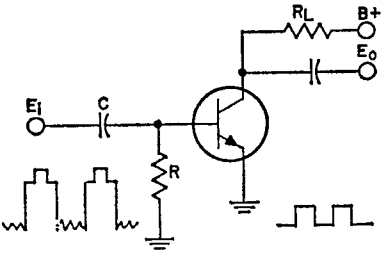



Figure 716. Transistor sync-separator circuit.

After the synchronizing signals are separated from the composite video signal, it is necessary to filter out the horizontal and vertical sync signals so that each can be applied to its respective deflection generator. This filtering is accomplished by RC circuits designed to filter out all but the desired synchronizing signals. Although the horizontal, vertical, and equalizing pulses are all rectangular pulses of the same amplitude, they differ in frequency and pulse width, as shown in Fig. 717. The horizontal sync pulses have a repetition rate of 15,750 per second (one for each horizontal line) and a pulse

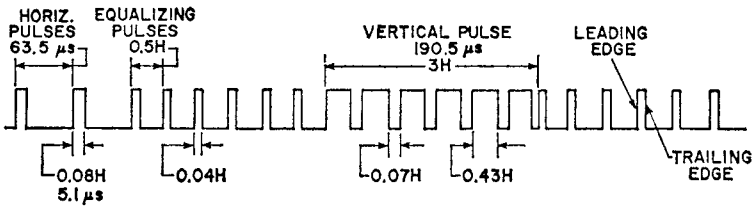



Figure 717. Waveform of TV synchronizing pulses (H = horizontal line period of 1/15,750 seconds, or 63.5 μs).

width of 5.1 microseconds. (For color systems, the repetition rate of the horizontal sync pulses is 15,734 per second.) The equalizing pulses have a width approximately half the horizontal pulse width, and a repetition rate of 31,500 per second; they occur at half-line intervals, with six pulses immediately preceding and six following the vertical synchronizing pulse. The vertical pulse is repeated at a rate of 60 per second (one for each field), and has a width of approximately 190 microseconds. The serrations in the vertical pulse occur at half-line intervals, dividing the complete pulse into six individual pulses that provide horizontal synchronization during the vertical retrace. (Although the picture is blanked out during the vertical retrace time, it is necessary to keep the horizontal scanning generator synchronized.)

All the pulses described above are produced at the transmitter by the synchronizing-pulse generator; their waveshapes and spacings are held within very close tolerances to provide the required synchronization of receiver and transmitter scanning.

The horizontal sync signals are separated from the total sync in a differentiating circuit that has a short time constant compared to the width of the horizontal pulses. When the total sync signal is applied to the differentiating circuit shown in Fig. 718, the capacitor charges completely very soon after the leading edge

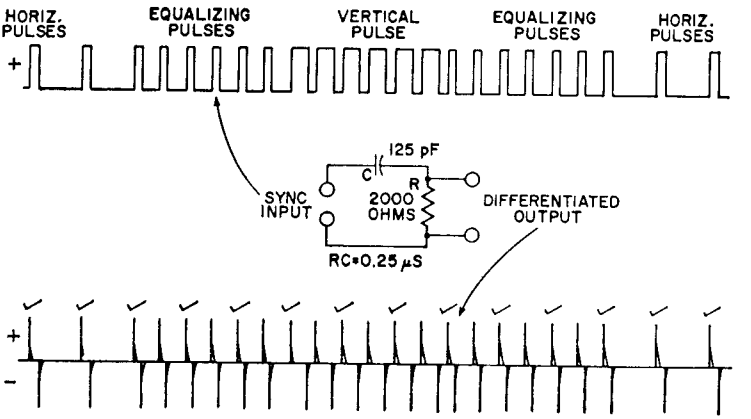



Figure 718. Separation of the horizontal sync signals from the total sync by a differentiating circuit.

of each pulse, and remains charged for a period of time equal to practically the entire pulse width. When the applied voltage is removed at the time corresponding to the trailing edge of each pulse, the capacitor discharges completely within a very short time. As a result, a positive peak of voltage is obtained for each leading edge and a negative peak for the trailing edge of every pulse. One polarity is produced by the charging current for the leading edge of the applied pulse, and the opposite polarity is obtained from the discharge current corresponding to the trailing edge of the pulse.

As mentioned above, the serrations in the vertical pulse are inserted to provide the differentiated output needed to synchronize the horizontal scanning generator during the time of vertical synchronization. During the vertical blanking period, many more voltage peaks are available than are necessary for horizontal synchronization (only one pulse is used for each horizontal line period). The check marks above the differentiated output in Fig. 718 indicate the voltage peaks used to synchronize the horizontal deflection generator for one field. Because the sync system is made sensitive only to positive pulses occurring at approximately the right horizontal timing, the negative sync pulses and alternate differentiated positive pulses produced by the equalizing pulses and the serrated vertical information have no effect on horizontal timing. It can be seen that although the total sync signal (including vertical synchronizing information) is applied to the circuit of Fig. 718, only horizontal synchronization information appears at the output.

The vertical sync signal is separated from the total sync in an integrating circuit which has a time constant that is long compared with the duration of the 5-microsecond horizontal pulses, but short compared with the 190-microsecond vertical pulse width. Fig. 719 shows the general circuit configuration used, together with the input and output signals for both odd and even fields. The period between horizontal pulses,

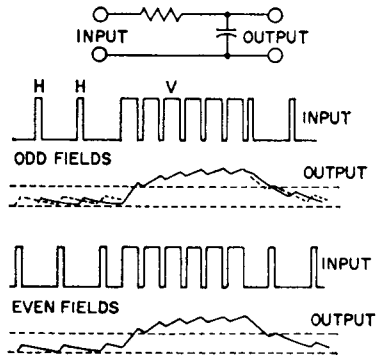



Figure 719. Separation of vertical sync signals from the total sync for odd and even fields with no equalizing pulses. (Dashed line indicates triggering level for vertical scanning generator.)

when no voltage is applied to the RC circuit, is so much longer than the horizontal pulse width that the capacitor has time to discharge almost down to zero. When the vertical pulse is applied, however, the integrated voltage across the capacitor builds up to the value required for triggering the vertical scanning generator. This integrated voltage across the capacitor reaches its maximum amplitude at the end of the vertical pulse, and then declines practically to zero, producing a pulse of the triangular wave

shape shown for the complete vertical synchronizing pulse. Although the total sync signal (including horizontal information) is applied to the circuit of Fig. 716, therefore, only vertical synchronization information appears at the output.

The vertical synchronizing pulses are repeated in the total sync signal at the field frequency of 60 per second (59.94 per second in color systems). Therefore, the integrated output voltage across the capacitor of the RC circuit of Fig. 719 can be coupled to the vertical scanning generator to provide vertical synchronization. The six equalizing pulses immediately preceding and following the vertical pulse improve the accuracy of the vertical synchronization for better interlacing. The equalizing pulses that precede the vertical pulses make the average value of applied voltage more nearly the same for even and odd fields, so that the integrated voltage across the capacitor adjusts to practically equal values for the two fields before the vertical pulse begins. The equalizing pulses that follow the vertical pulse minimize any difference in the trailing edge of the vertical synchronizing signal for even and odd fields.

## HORIZONTAL-DEFLECTION SYSTEM

The main functions of the horizontal-deflection system in a television receiver are to deflect the electron beam linearly (from left to right) across the picture-tube screen, return the beam rapidly to the left side of the screen, and then repeat the process. Fig. 720 shows an idealized waveform of the current that passes through the horizontal-deflection-yoke windings during one complete scanning cycle. As pointed out previously, 525 such "scanning lines" are required to produce each picture in a United States television system.

In addition to beam deflection, the over-all horizontal system performs a number of auxiliary functions. These functions include:

1. Generation of the high voltage for the picture tube.
2. High-voltage regulation.
3. Scan-linearity correction.
4. Retrace blanking.
5. Gating signal for automatic gain control (agc).
6. Timing reference for automatic frequency control (afc).
7. Bias voltage for grids 2 and 4 of the picture tube.

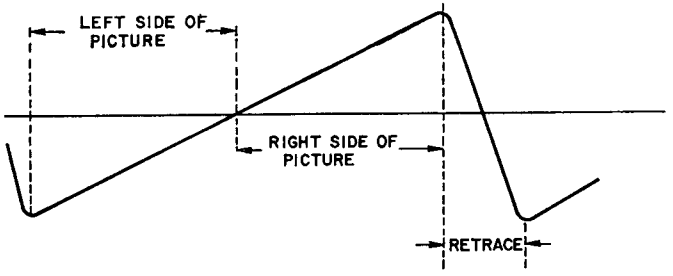



Figure 720. Current waveform applied to the horizontal-yoke windings during one complete scanning cycle.

## Basic Analysis of Horizontal-System Switching

In the horizontal-deflection system of a television receiver, a current I that varies linearly with time and has a sufficient peak-to-peak amplitude must be passed through the horizontal-deflection-yoke winding to develop a magnetic field adequate to deflect the electron beam of the television picture tube. After the beam is deflected completely across the face of the picture tube, it must be returned very quickly to its starting point. (As explained previously, the beam is extinguished during this retrace by the blanking pulse incorporated in the composite video signal, or in some cases by additional external blanking derived from the horizontal-deflection system.)

The simplest form of a deflection circuit is shown in Fig. 721(a). In this circuit, the yoke impedance L is assumed to be a perfect inductor. When the switch S is closed, the yoke current starts from zero and increases linearly. The rate of increase in current (di/dt) is determined as follows:

$$\frac{di}{dt} = \frac{E}{L} \qquad (461)$$

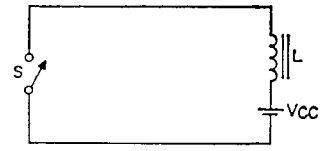
Integration of Eq. (461) yields the following expression for the instantaneous value of current i at any time t:

$$i = \frac{Et}{L} \qquad (462)$$

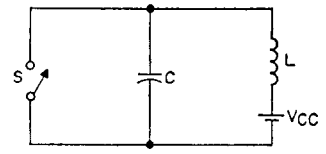
If the switch is opened at $t = t_1$, the current I instantly drops to zero from an initial value determined as follows:

$$I = \frac{Et_1}{L} \qquad (463)$$

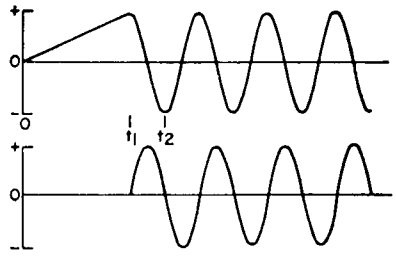
Although the simple circuit shown in Fig. 721(a) satisfies the basic requirement for hori-



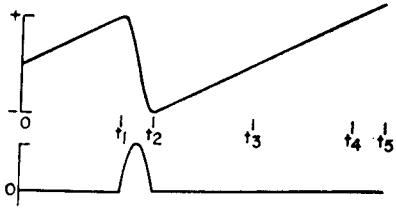

(a) SIMPLE DEFLECTION CIRCUIT



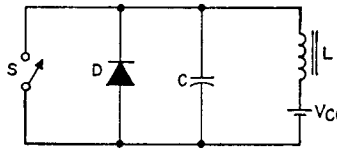

(b) ADDITION OF CAPACITOR




(c) YOKE CURRENT (top) AND SWITCH VOLTAGE (bottom) FOR CIRCUIT (b)




(d) YOKE CURRENT (top) AND SWITCH VOLTAGE (bottom) FOR SWITCH CLOSED AT $t_2$




(e) ADDITION OF DAMPER DIODE

Figure 721. Development of horizontal-deflection circuit.

zontal beam deflection, it presents some serious problems and limitations.

The voltage across the switch is given by the following equation:

$$e = L\frac{di}{dt} \qquad (464)$$

Because the rate of change in current di/dt is infiinite, the voltage across the switch also is infinite.

In addition, if very little of the total time were spent at zero current, the average supply current would be I/2. This current would require a tremendous amount of dc power because the voltage-current product (E × I) for standard deflection systems (E = 18 kilovolts for 110-degree U.S. black-and-white systems) is in the order of 300 watts. (The product of the peak inverse voltage and the peak-to-peak current, which is often used to describe a system, is closer to 2500 watts, as explained later.) In addition, the operation of the switch would be rather critical with regard to both its opening and its closing.

A final limitation would be the fact that the deflection field would be poled or phased in one direction only, so that the beam would have to be centered at the extreme left of the screen for zero yoke current.

If a capacitor is placed across the switch, the yoke current still increases linearly when the switch is closed at time t = 0. When the switch is opened, however, at time $t = t_1$, a parallel resonant circuit is formed by the parallel combination of L and C, as shown in Fig. 721(b). The initial conditions of this simple transient network are as follows:

$$i_y = \frac{Et_1}{L}$$

$$e_c = 0$$

where $i_y$ is the yoke current and $e_c$ is the capacitor or switch voltage. The resulting yoke currents and switch voltages are shown in Fig. 721(c). The current is maximum when the voltage equals zero, and the voltage is maximum when the current equals zero. The ringing frequency (if zero losses are assumed) is given by

$$f_{osc} = \frac{1}{2\pi\sqrt{LC}} \qquad (465)$$

If the switch is closed again at any time the capacitor voltage is not equal to zero, an infinite switch current flows as a result of the capacitive discharge. However, if the switch is closed at the precise moment that the capacitor voltage equals zero, the capactior current effortlessly transfers to the switch and a new transient condition results, as described below.

At the time of the proposed switching, time $t = t_2$ and the yoke current $i_y = -I$. The yoke current increases at the rate dictated by Eq. (461), but it starts from −I instead of from zero. Fig. 721(d) shows the yoke current and switch voltage waveforms.

If the switch is again opened at $t_4$, closed at $t_5$, and so on, the desired sweep will result, the peak switch voltage will be finite, and the average supply current will be zero. The deflection system is then lossless and efficient. Because the average yoke current is also zero, beam decentering is avoided.

The only fault in the final circuit is the critical timing of the

switch, particularly at time $t = t_2$. If the switch is shunted by a properly poled diode (the damper diode), as shown in Fig. 721(e), the diode acts as a closed switch as shown as the capacitor voltage reverses slightly. The switch may then be closed any time from $t = t_2$ to $t = t_3$.

The horizontal scanning rate for television systems in the United States is 15,750 scans per second for black-and-white types and 15,734 scans per second for color types. Obviously mechanical switches cannot operate at such high rates. In practice, the switch S will be an active device, such as an electron tube, a transistor, or a thyristor.

Fig. 722 shows a basic horizontal deflection circuit that uses a transistor in place of switch S.

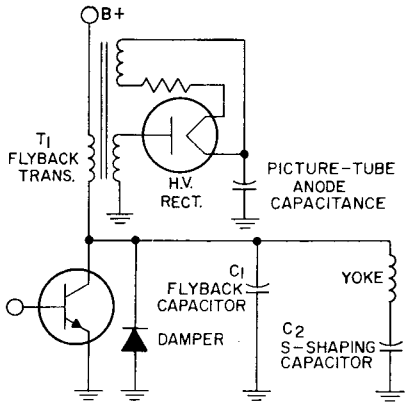



Figure 722. Simple transistor horizontal-deflection circuit.

High voltage is generated by use of the step-up transformer $T_1$ in parallel with the yoke. This step-up transformer is designed so that its leakage inductance, distributed capacitance, and output stray capacitance complement the yoke inductance and retrace tuning capacitance in such a manner that the peak voltage across the primary winding is reduced and the peak voltage across the secondary winding is increased, as compared to the values that would be obtained in a perfect transformer. This technique, which is referred to as **third-harmonic tuning,** yields a voltage ratio of secondary-to-primary peak voltage of approximately 1.7 times the value expected in a perfect transformer.

The following paragraphs describe the use of power transistors and thyristors (SCR's) as the main switching element in practical horizontal deflection systems.

## Transistor Horizontal-Deflection Circuits

Fig. 723 shows the functional relationship among the various circuit elements of a horizontal-deflection circuit that uses a power transistor to generate the sawtooth of current through the deflection yoke and to develop the beam accelerating voltage for the picture tube. The high-voltage transformer shown across the output stage may be used as a slight step-up or step-down transformer for the picture-tube high-voltage supply, the yoke, the damper diode, the capacitor, or any combination of these elements.

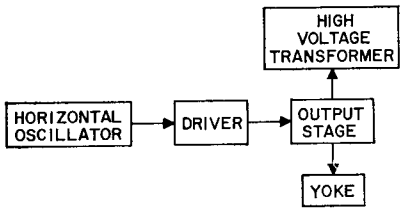



Figure 723. Block diagram of a transistor horizontal-deflection system.

In the following paragraphs, the design factors and technical considerations used in the development of a typical horizontal-deflection-system circuit are explained. This system is assumed to provide the deflection energy and high voltage required for a 19-inch, 20-kilovolt, 114-degree monochrome receiver from a power supply having a 12-microsecond retrace time. Basic circuit configurations for practical horizontal-deflection systems for both monochrome and color television receivers are then shown and analyzed.

**Voltage Considerations** — For an idealized horizontal-deflection circuit, the peak voltage $E_{max}$ across the transistor is given by

$$E_{max} = \left(1.79 + 1.57\frac{T_t}{T_R}\right) E_{dc} \quad (466)$$

where $T_t$ is the scanning or trace time, $T_R$ is the retrace time, and $E_{dc}$ is the supply voltage. If third-harmonic tuning is employed, the peak voltage is reduced by approximately 20 per cent.

The highest anticipated value of $E_{max}$ is determined by use of the value of $E_{dc}$ obtained at high ac line voltage and at the lowest horizontal-oscillator frequency, i.e., the longest trace time. (For these conditions, of course, the receiver is out of sync.) The tolerances on the inductors and capacitors alter the trace time only slightly and usually may be ignored if a 10-per-cent tolerance is used for the tuning capacitor.

When a capacitor is used in series with the yoke for linearity correction, the peak-to-peak yoke current and the flyback voltage are both increased by about 10 per cent. In a first-order approximation, this effect may be ignored if the system is designed without S-shaping. If shaping is employed, however, the supply voltage must be reduced by 5 to 10 per cent to restore the scan conditions originally observed.

An abnormality that must be considered is high-voltage arcing. Fig. 724 shows the normal transistor load for third-harmonic tuning of the flyback transformer in which the leakage inductance, secondary-winding capacitances, and anode stray capacitances are reflected to the primary. In a properly designed system, the leakage inductance is about one-half the shunt inductance (yoke plus flyback primary inductance).

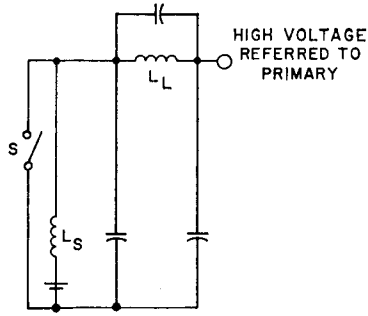



Figure 724. Equivalent output circuit for third-harmonic tuning (referred to primary side).

When a high-voltage arc occurs, the secondary is momentarily shorted, placing the leakage inductance in parallel with the shunt inductance. As a result, the peak collector current is increased by a factor of about three, and the retrace time is decreased by a factor of about two (if that transistor is still operating as an ideal switch).

Because the flyback voltage would then be increased by a factor of 2.5, avalanche breakdown occurs at a high current level, second breakdown is initiated, and

the transistor is destroyed. Since occasional high-voltage arcing is unavoidable in the picture-tube gun, the output transistor must be protected.

If a diode and capacitor are connected in series and placed across the transistor, the flyback pulse is clamped at a level equal to the normal peak value when a high-voltage arc occurs. When the arcing is sustained long enough for an appreciable increase in the capacitor voltage, the increased drain (caused by the very high peak collector current) opens a fuse in the B-supply, and the transistor is adequately protected. A bleeder resistor is placed across the capacitor to protect against intermittent arcs. This circuit also protects against several other types of high-voltage short circuits.

Another method used to reduce the effect of high-voltage arcing is to make the leakage inductance of the secondary very high compared to the shunt inductance by designing the secondary to resonate at the fundamental frequency (15 kHz). In addition to protection during high-voltage arcs, this method reduces peak collector current (caused by higher primary inductance) and also facilities manufacture of the flyback transformer.

There are several disadvantages, however. Because the flyback primary current is very high and circulates at all times (as opposed to the case of third-harmonic tuning), very high primary and secondary losses occur. In addition, the magnetic field of the transformer is quite high and causes interference problems in the rest of the receiver. It also becomes difficult to enclose the transformer in a cage without causing an excessive shorted-turn problem because the cage is magnetically coupled.

A third and rather significant disadvantage is the high peak-to-peak secondary voltage developed for a given value of dc high voltage. In third-harmonic tuning, the secondary voltage waveform exhibits a narrow spike for approximately 10 per cent of the cycle and a low constant voltage for the remainder of the cycle. As a result, the peak-inverse rating on the high-voltage rectifier is approximately 1.1 times the dc high voltage developed. In the "fundamental-tuned" arrangement, the secondary voltage is nearly a sine wave and results in a peak-inverse rating on the high-voltage rectifier approximately twice that of the third-harmonic-tuned system.

**Choice of Retrace Time**—The choice of a slightly longer retrace time offers the following significant advantages for circuit design:

1. As retrace time is lengthened, the product of peak voltage and peak current is reduced directly.
2. The peak stored energy, as well as the voltage-current product, is reduced because more primary inductance can be used in the flyback transformer.
3. The retrace losses are reduced with the square of the retrace time.
4. Losses in the yoke and flyback that result from skin effect are reduced.
5. The core losses in the flyback transformer are reduced because of the greater inductance.
6. The supply voltage may be

increased because of the lower flyback pulse.

7. The flyback transformer secondary becomes easier to wind.

**High-Voltage Power** — High voltage is obtained by means of a tertiary winding on the flyback transformer which through autotransformer action steps up the yoke pulse to a high value. The energy typically extracted is seldom greater than 0.3 millijoule (for 5 watts of beam power) and results in a typical circuit Q of approximately 60, if other degenerative losses are neglected. When an LC network is damped to a Q of 60, the voltage and current waveshapes for the first $\pi$ radians show very little change (except for phase relationship) over the infinite-Q condition. Therefore, the losses, which are determined by the voltage and current waveshapes, do not increase when beam current flows; i.e., 5 watts of beam power reflects only an added demand of 5 watts in the power supply.

A further point of interest in transistor deflection circuits is the excellent high-voltage regulation encountered. This improvement is the result of the high efficiency of these circuits, which keeps the extracted energy to a minimum and results in a fairly high circuit Q. As noted, the anode-voltage amplitude does not change much as energy is extracted and thus accounts for good high-voltage regulation.

**Scan Linearity**—For accurate reproduction of pictures on the picture-tube screen, the electron beam must move at a linear rate across the faceplate of the picture tube. If the faceplate were a section of a sphere with its center at the center of deflection, a linear sawtooth of current through the deflection yoke would be required for linear deflection. Although the faceplate is a section of a sphere, its radius is much greater than the distance to the center of deflection. For all practical purposes, the faceplate can be considered as a flat plane. The distance from the center of deflection to the faceplate, therefore, is greater at the edges than at the center of the picture tube. Consequently, a given amount of deflection of the beam at the deflection center produces a greater movement of the electron beam on the face of the picture tube. For this reason, the required current waveform through the deflection yoke should be somewhat "S"-shaped rather than an absolutely linear sawtooth. Much of this S-shaping is accomplished by the capacitor connected in series with the yoke. At the start of trace (left side of the picture), maximum energy is stored in the yoke. Current flows from the yoke into the S-shaping capacitor and causes the capacitor to charge at an exponential rate. Yoke current rises linearly during the first half of trace and gradually decreases near the right-hand side of the screen as the S-shaping capacitor accumulates charge. However, because the first part of trace is linear, the left side of the picture will be stretched in comparison to the right side, and additional linearity correction is required.

Methods for eliminating this nonlinearity include the use of a saturable reactor in series with the yoke, the use of a permanent magnet near the yoke to distort the field, and the use of a damped series-resonant circuit

connected in parallel with the S-shaping capacitor and in series with the yoke.

**Deflection Energy Requirement**—The peak deflection energy required by the yoke for complete scanning of picture tubes varies directly with the high voltage, the 5/2 power of the deflection angle (approximately), and the neck diameter (where all geometries of the yoke are adjusted in direct proportion). The peak energy required for minimum scanning of a 114-degree picture tube having a 1-1/8-inch neck diameter and an anode voltage of 20 kilovolts is 2.4 millijoules (scan from center to either side).

When full scan is obtained at low line voltage and at an anode voltage that corresponds to low line, the peak stored energy equals $\epsilon$. When the line voltage is increased, the peak energy increases in proportion to the square of the voltage. If the low line voltage is 105 volts and high line is 135 volts, the increase in energy is a factor of 1.65.

If the receiver is adjusted out of sync by 2 microseconds (or a 480-cycle pullout range), the energy, which is proportional to the square of the trace time in a fixed circuit, increases by a factor of 1.08. If the yoke is shunted by a practical flyback transformer, the inductance is reduced by a factor of approximately 1.3 and, therefore, the peak stored energy in the system is increased by a factor of 1.3. When all three items are considered, the transistor must handle 2.3 times the peak stored energy normally expected.

**Transistor Drive Considerations**—Transformer drive is usually employed for the output transistor. When this type of drive is used, the collector load may be placed in series with either the collector or the emitter because, in either case, the transformer secondary appears from base to emitter. If the load is in series with the emitter (emitter loading), the collector is directly at the supply-voltage potential. If a positive supply is used, the transistor case is at chassis potential. The damper diode is constructed with its anode at case potential so that it is also at chassis potential.

This method has a disadvantage in that a high potential is placed between the primary and secondary windings of the driver transformer. Because the driver transformer is very tightly coupled, insulation breakdown must be carefully considered.

While the output stage is cut off, the driver stage should be conducting; the transformer secondary can then provide any current demanded. (The current, however, is limited by the leakage inductance.) When the driver stage is cut off, the energy stored in the transformer flows in the secondary in the form of a constant current. If this mode of drive is employed, and if the base-to-emitter voltage of one transistor varies from that of another, the turn-on current still starts at the same value but decays at a different rate. If all charge is removed from the base of the output transistor during turn-off, no more transformer current is required and the transistor stays at a reverse-bias mode.

No impedance should be placed in the base. Transistor interchangeability is thus improved because the voltage levels remain

low enough to prevent breakdown of the base-emitter junction during the turn-off period. The primary and secondary windings of the driver transformer must be very tightly coupled to obtain a large spike of current during the turn-off period (for a fast turn-off time).

The circuit shown in Fig. 725 is used to develop the all-important waveshaping. The 560-ohm

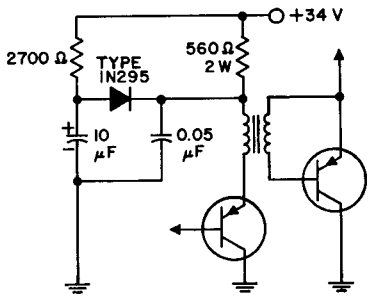



Figure 725. Waveforming circuit.

resistor in combination with the 0.05-microfarad capacitor increases the amplitude and rise time of the turn-off base current for the first few microseconds. The 1N295 diode, together with the 2700-ohm resistor and 10-microfarad capacitor, serves as a clamp circuit which assures that the output transistor is always reverse-biased during the entire turn-off period, even in the presence of high $I_{CBO}$ at several hundred volts and elevated temperatures. With this circuit, the 560-ohm, 0.05-microfarad combination can be optimized for the best turn-off time without regard for the remainder of the off signal.

The turn-off pulse developed by this circuit is 3 amperes for approximately 2 microseconds, followed by a constant voltage of approximately 1/2 volt for 18 microseconds. The on-pulse then initiates at 650 milliamperes and, 45 microseconds later, decreases to 500 milliamperes.

**Deflection Circuit for Monochrome Receiver**—The following paragraphs describe a practical horizontal-deflection system for a 19-inch black-and-white (monochrome) television receiver. The deflection system operates from a regulated dc supply of 100 volts.

The power-supply voltage of 100 volts is decoupled to 85 volts for raster regulation with brightness. A retrace time of 14 microseconds is selected to present the maximum usable picture, although a value of 17 microseconds could have been used with no sacrifice in performance as compared to present-day receivers.

The picture tube used, the 19DQP4, has minimum usable screen dimensions of 15⅛ inches horizontally and 12 inches vertically. These dimensions establish the front mask size for the cabinet and fix the aspect ratio at 1.26. The diagonal deflection angle of the 19DQP4 is 114 degrees, and the neck diameter is a nominal 1⅛ inches. The zero-beam accelerating potential is 20 kilovolts. The horizontal circuit should be capable of providing an average beam current of 400 microamperes with virtually no change in raster height or width at any brightness setting between zero and full current. An overscan of 4 per cent is desired.

Energy requirements for horizontal deflection show that the peak stored energy in the yoke must be 2.4 millijoules to fulfill the requirements for the 19DQP4.

If a trace time of 49.5 microseconds and a power-supply voltage of 85 volts are used, and if it is assumed that the use of "S" shaping (by use of a capacitor in series with the yoke) has the effect of increasing the scan by 5 to 10 per cent, the yoke inductance must be 1 millihenry and the peak-to-peak yoke current must be 4.4 amperes.

**Driver and output circuit:** The horizontal-drive-and-output circuit for the receiver is shown in Fig. 726. The output circuit ($Q_5$) is basically a self-oscillator which requires the 27,000-ohm resistor to initiate oscillation. Drive current is obtained through the 33-ohm resistor in parallel with the picture-tube heater and through diode $D_5$, and is applied from the feedback winding of the transformer $T_1$ through the 50-microfarad capacitor in parallel with the 10-ohm resistor to the base of the output transistor $Q_5$. If this drive circuit is correctly designed, transistor $Q_5$ does not come out of saturation during normal operation. When retrace is to be initiated, the driver transistor $Q_6$ is driven heavily into saturation. The drive current is then shunted to ground, and the base of transistor $Q_5$ is simultaneously reverse-biased by means of the charge stored on the 50-microfarad capacitor. If the resistor shunting this capacitor is large compared to 2 ohms, most of the drive current flows through the capacitor. When transistor $Q_6$ is saturated, therefore, a capacitor current of opposite polarity results through $Q_6$ and $D_4$. As a result, the turn-off drive to transistor $Q_5$ is always a reverse-bias voltage equal to the forward drop across diode $D_4$. If the value of the parallel resistor is made extremely large, the circuit may not start.

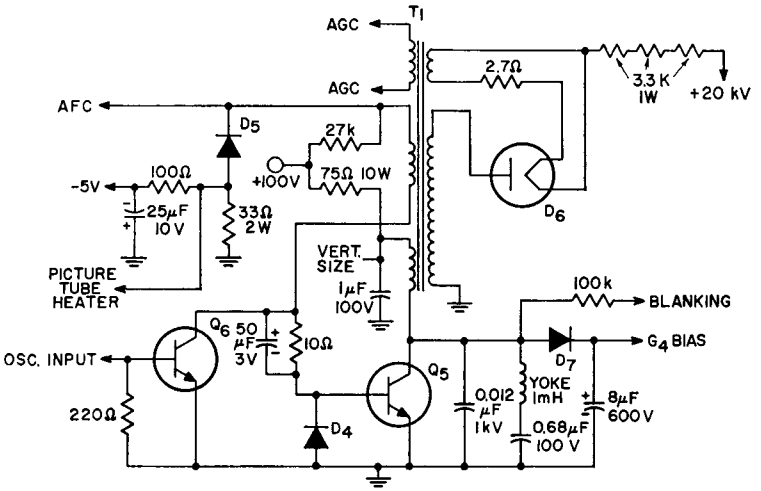



Figure 726. Horizontal driver and output circuit for a black-and-white television receiver.

When the output transistor $Q_5$ is turned off, the collector voltage comes out of saturation and goes through the normal flyback pulse. During this time, the feed-

back drive (already shunted to ground) decreases and, as the collector voltage passes the supply voltage, a heavy reverse drive current results. However, the driver diode $D_5$ blocks this reverse current flow. $D_5$ also permits the starting current (through the 27,000-ohm resistor) to flow through the base of $Q_5$.

The damper current flows through the collector-base diode of $Q_5$ in series with $D_4$ to ground. Diode $D_4$ is a silicon type that has a low forward drop at 2 amperes and a minimum breakdown requirement of only 1 volt. It must be capable of dissipating 300 milliwatts. $D_4$ is called a damper diode, even though it only partially fulfills this function. The 50-microfarad coupling capacitor must have a low series resistance to obtain proper turn-off.

The 100-volt supply voltage is reduced 15 per cent at zero beam current by means of the 75-ohm decoupling resistor. When the beam current is increased to 400 microamperes, the demand for extra power of 7 or 8 watts causes the decoupled voltage to drop. As a result, the high voltage and the scanning current decrease linearly with the decoupled voltage. The high voltage also decreases because of the lack of perfect high-voltage regulation. If the circuit is designed correctly, the high voltage decreases with the square of decoupled voltage so that the scanning-energy requirement approximately tracks the scanning energy provided. This decoupled voltage is also fed back to the vertical circuit in the size-determining portion of the circuit so that the vertical scan energy also tracks the high voltage as a function of picture-tube average brightness setting.

A separate winding on the flyback transformer $T_1$ provides gating for the agc circuit. A signal taken from the driver diode $D_5$ provides a timing reference for the horizontal phase circuit (afc). A positive voltage of approximately 500 volts is available from the clamp circuit provided by diode $D_7$ to supply bias to grid No.2 or grid No.4 of the picture tube. (The current drain should be kept below 1 milliampere.)

Picture-tube heater power is also derived from the horizontal-driver circuit. When the receiver is first turned on, the base drive current to transistor $Q_5$ is larger than normal because of the thermally nonlinear characteristic of the heater. This method of providing heater power should prove to be satisfactory for long picture-tube life. However, excessive heater-to-cathode capacitance may cause a video modulation in the form of a vertical line similar to a drive line in tube deflection. No such problem has been experienced with the approach shown. A more conservative control of heater power may be obtained by means of a separate winding or by incorporation of the heater with the agc winding.

The video-blanking circuit must be gated from the flyback transformer $T_1$. A 100,000-ohm resistor is fed from the collector of the output transistor $Q_5$ for this function. This resistor provides blanking whenever the collector voltage is more positive than 25 volts.

The picture-tube heater has a dc voltage across it, together with a large ac voltage. After adequate decoupling, this dc volt-

age provides a convenient source of negative potential to power the agc and sync-separator circuits.

Various forms of arcing protection are provided in the horizontal output circuit.

A voltage-clamp circut is provided by the clamp diode $D_7$ in conjunction with the 8-microfarad capacitor. Sufficient current drain must be provided across the capacitor to discharge it between arcs. The capacitor must be large enough to absorb most of the energy stored by the picture-tube capacitance. The purpose of this clamp circuit is to assure that the transistor does not go into voltage breakdown during high-voltage arcs.

The 75-ohm decoupling resistor provides raster regulation, as mentioned previously, and also limits the maximum power that may be delivered to the entire horizontal-scanning circuit to 30 watts.

If the output transistor $Q_5$ is pulled out of saturation at a high collector-current level as a result of high-voltage arcing, the feedback drive circuit turns off $Q_5$ and thus controls the transistor load line. The transistor turns off fairly fast under this condition because it is in an unsaturated state. If the driver transistor $Q_6$ is turned on or off when $Q_5$ is reverse-biased, no change in state occurs because the drive is basically self-oscillating and transistor $Q_6$ functions merely as a gate. If drive is available, $Q_6$ may exclude it. If drive is being applied, $Q_6$ may turn it off. However, $Q_6$ may not provide drive if $Q_5$ is not saturated.

Although drive is available at the beginning of trace time, it should be excluded by $Q_6$ until about 5 microseconds prior to normal need. As a result, $Q_6$ should receive a drive pulse that saturates it for approximately 30 microseconds, and turns it off for the remaining 34 microseconds.

The clamp diode $D_7$ must not fail. If it does, destruction of $Q_6$ is almost assured.

**Horizontal oscillator:** Fig. 727 shows a simplified diagram of a multivibrator type of horizontal-oscillator circuit. It should be noted that a gated dc feedback signal is provided from the power supply. If the 100-volt supply becomes excessively high as a result of a fault, the horizontal-oscillator frequency is raised to such a point that the flyback voltage remains within specifications.

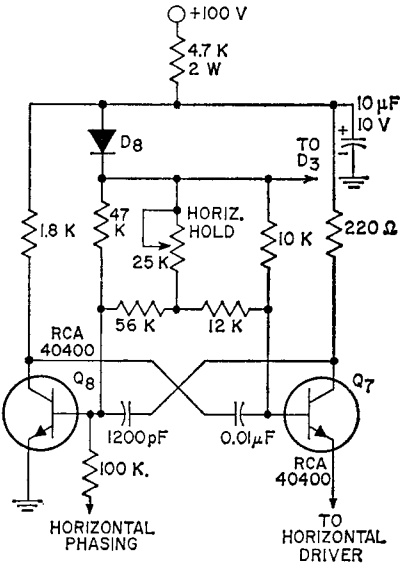



Figure 727. Horizontal oscillator circuit.

**Horizontal phasing** (afc): The horizontal phasing used is novel. Gating is obtained from a 1-milliampere sync pulse that is only 2

microseconds wide (and can be much narrower if desired). Nearly any desired average control current up to several milliamperes can be provided. When the picture is correctly phased, the circuit is open to receive a sync pulse for only 4 microseconds, and thus is relatively immune to noise. Because the circuit functions on the leading edge of the sync pulse, rather than on the entire area of the differentiated pulse, the effects of the vertical equalizing pulses and the serrations in the vertical pulse are greatly minimized. As a result, the top of the picture exhibits proper synchronization at essentially all settings of the hold control (as is not the case with normal afc).

The horizontal-phasing circuit is shown in Fig. 728. A control current of only 20 microamperes is required for the low-level oscillator. Transistors $Q_9$ and $Q_{10}$ are connected in a latching configuration that resembles a thyristor. (A thyristor could be used if it were fast enough and sensitive enough.) A replica of the flyback pulse is applied to the emitter of transistor $Q_9$ from diode $D_5$ through the voltage divider consisting of the 10,000-ohm and 3300-ohm resistors. During trace time, this voltage is slightly negative, and any signal appearing on the base of transistor $Q_{10}$ is ineffective. When the flyback voltage appears across transistor $Q_9$, however, the gate is open to receive a sync pulse. When the pulse fires transistor $Q_{10}$, the combination of transistors $Q_9$ and $Q_{10}$ becomes regenerative; the transistors then become heavily saturated and pass any amount of current the voltage divider will permit. Transistors $Q_9$ and $Q_{10}$ remain in saturation until the voltage reverses and resets the latch. Control current of the opposite polarity may be obtained from the emitter of transistor $Q_9$ if desired. Diode $D_{16}$ serves to block reverse current.

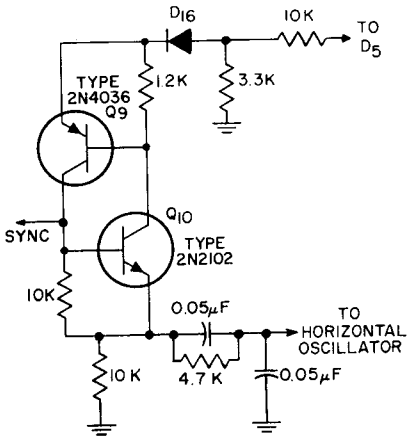



Figure 728. Horizontal-phasing (afc) circuit.

**Deflection Circuit for Color Receiver**—Fig. 729 shows a schematic of a transistorized horizontal-deflection circuit for a color TV receiver. The horizontal output transistor, $Q_4$, is a high-voltage silicon transistor. The normal collector-to-emitter pulse voltage across $Q_4$ includes an ample safety factor that allows for any increased pulse that may result from out-of-sync operation, line surges, and other abnormal conditions.

A unique feature of the horizontal-deflection circuit is the low-voltage supply of approximately 23 volts that is derived from it. This feature makes it possible to eliminate the power transformer in the power supply. The low-voltage power is used to

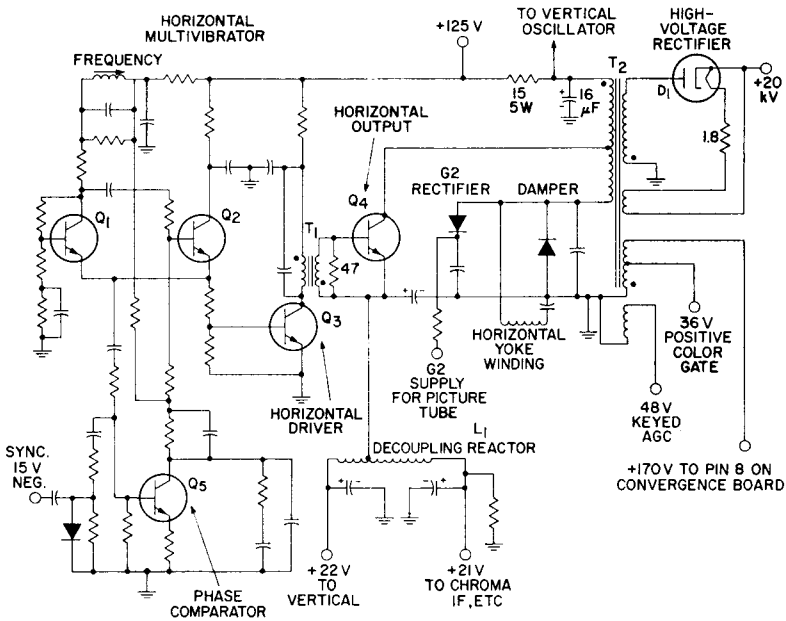



Figure 729. Horizontal-deflection system for a color television receiver.

operate all but the high-voltage receiver stages, such as the video-output stage, the audio-output stage, and the horizontal oscillator and driver. The vertical oscillator is supplied from the same point which supplies the horizontal output in such a way that the actual voltage is a function of beam current; this connection compensates for the tendency for picture height to change with brightness settings.

The transistorized deflection circuit achieves commercially acceptable high-voltage regulation without the use of the high-voltage shunt regulator used with tube-type deflection circuits. With a flyback transformer of normal design and a low-voltage power supply with about 3-per-cent regulation, high-voltage regulation from zero beam to full load of 750 microamperes is about 3 kilovolts and is accompanied by a considerable increase in picture width. Improvement of this behavior with brightness changes is achieved by utilizing the accompanying changes of direct current to the deflection circuit in two ways. First, the air gap of the transformer is reduced to permit core saturation to decrease the system inductance as the high-voltage load is increased. When this method is used, regulation is improved to about half that of the normal transformers with no circuit instabilities, but picture-width change is still greater than desired. Second, series resistance is added to the B supply to decrease power input at full load and thereby reduce the change in picture width (at some sacrifice in high-voltage regulation). The net result of both changes is a regulation of

about 2.8 kilovolts for the high voltage, with very little variation in picture size.

A secondary benefit of the inherently good regulation of the transistor deflection system is a reduction in the size of the flyback transformer. The size reduction is accomplished by a reduction in the area of the "window" in the flyback core. A reduction in the size of the high-voltage cage required to maintain adequate isolation of the high-voltage winding from ground is possible because of the smaller flyback transformer.

The transformer-coupled driver stage takes advantage of the high-voltage capability and switching speed of the horizontal driver transistor which is designed primarily for video-output use. A sine-wave stabilized multivibrator type of horizontal oscillator is used. This type of oscillator is especially useful in experimental work with deflection systems because it permits on-time and off-time periods to be easily varied.

The afc phase detector operates on the principle of pulse-width variation of combined sync and reference pulses. In the circuit shown in Fig. 729, timing information is related to the leading edges of the sync pulses, and the retrace process is initiated prior to the leading edge of the sync pulse; performance of the circuit is very satisfactory.

## SCR Horizontal-Deflection System

Until recently, solid-state horizontal deflection has been limited to small-screen monochrome receivers with relatively low energy requirements. Solid-state devices which would compete with receiving tubes for both cost and performance were not available. The development of silicon controlled rectifiers (SCR's) and fast-recovery diodes capable of operating at the horizontal scanning rates used in television has made possible the design of a horizontal-deflection system that is economically competitive and, at the same time, provides greater reliability than any other known deflection system.

In this system, the switching action required to generate the scan current in the horizontal-yoke windings and the high-voltage pulse used to derive the dc operating voltages for the picture tube is controlled by two SCR's that are used in conjunction with associated fast-recovery diodes to form bipolar switches.

The SCR's used to control the trace current and to provide the commutating action to initiate trace-retrace switching exhibit high voltage- and current-handling capabilities together with the excellent switching characteristics required for reliable operation in deflection-system applications. The switching diodes, (trace and commutating diodes), provide fast recovery times, high reverse-voltage blocking capabilities, and low turn-on voltage drops. These features, together with the fact that, with the exception of one noncritical triggering pulse, all control voltages, timing, and control polarities are supplied by passive elements within the system (rather than by external drive sources), contribute substantially to the excellent reliability of the SCR deflection system.

The system operates directly from a conventional, unregulated dc power supply of +155 volts, and provides full-screen deflection at angles up to 90 degrees at full beam current. The current and voltage waveforms required for horizontal deflection and for generation of the high voltage are derived essentially from LC resonant circuits. As a result, fast and abrupt switching transients which would impose strains on the solid-state device are avoided.

A regulator stage is included in the SCR horizontal-deflection circuit to maintain the scan and the high voltage within acceptable limits with variations in the ac line voltage or picture-tube beam current. The system also contains circuits that provide full protection against the effects of arcs in the picture tube or the high-voltage rectifier, and linearity and pincushion correction circuits.

**Basic Deflection Circuit**—The essential components in the SCR horizontal-deflection system required to develop the scan current in the yoke windings are shown in Fig. 730. Essentially the trace-switch diode $D_T$ and the trace-switch controlled rectifier $SCR_T$ provide the switching action which controls the current in the horizontal yoke windings $L_y$ during the picture-tube beam-trace interval. The commutating-switch diode $D_C$ and the commutating-switch controlled rectifier $SCR_C$ initiate retrace and control the yoke current during the retrace interval. Inductor $L_R$ and capacitors $C_R$, $C_A$, and $C_y$ provide the necessary energy storage and timing cycles. Inductor $L_{CC}$ supplies a charge path for capacitor $C_R$ from the dc supply voltage (B+) so that the system can be recharged from the receiver power supply. The secondary of inductor $L_{CC}$ provides the gate trigger voltage for the trace-switch SCR. Capacitor $C_R$ establishes the optimum retrace time by virtue of its resonant action with inductor $L_R$.

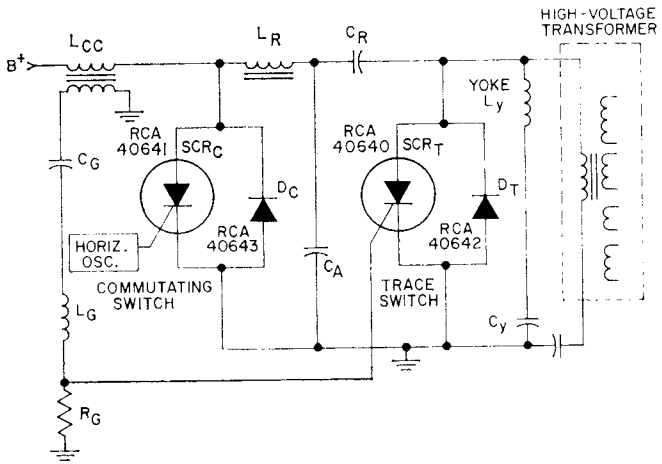



Figure 730. Basic circuit for generation of the deflection-current waveform in the horizontal yoke winding.

The complete horizontal-deflection cycle may best be described as a sequence of discrete intervals, each terminated by a change in the conduction state of a switching device. In the following discussion, the action of the auxiliary capacitor $C_A$ and the flyback high-voltage transformer are initially neglected to simplify the explanation.

**First half of the trace interval:** Fig. 731 shows the circuit elements involved and the voltage and current relationships during the first half of the trace deflection-current interval, the period from $T_0$ to $T_2$. At time $T_0$, the magnetic field has been established about the horizontal yoke windings $L_y$ by the circuit action during the retrace period of the preceding cycle (explained in the subsequent discussion of retrace intervals). This magnetic field generates a decaying yoke current $i_y$ that decreases to zero when the energy in the yoke winding is depleted (at time $T_2$). This current charges capacitor $C_y$ to a positive voltage $V_{Cy}$ through the trace-switch diode $D_T$.

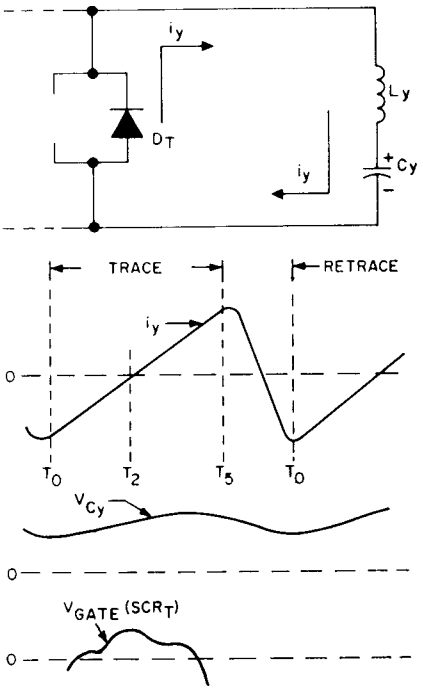



Figure 731. Effective configuration of the deflection circuit during the first half of the trace interval, time $T_0$ to $T_2$, and operating voltage and current waveforms for the complete trace-retrace cycle.

During the first half of the trace interval (just prior to time $T_2$) the trace controlled rectifier $SCR_T$ is made ready to conduct by application of an appropriate gate voltage pulse $V_{GATE}$. $SCR_T$ does not conduct, however, until a forward bias is also applied between its anode and cathode. This voltage is applied during the second half of the trace interval.

**Second half of the trace interval:** At time $T_2$, current is no longer maintained by the yoke inductance, and capacitor $C_y$ begins to discharge into this inductance. The direction of the current in the circuit is then reversed, and the trace-switch diode $D_T$ becomes reverse-biased. The trace-switch controled rectifier $SCR_T$, however, is then forward-biased by the voltage $V_{Cy}$ across the capacitor, and the capacitor discharges into the yoke inductance through $SCR_T$, as indicated in Fig. 732. The capacitor $C_y$ is sufficiently large so that the voltage $V_{Cy}$ remains essentially constant during the entire trace and retrace cycle. This constant voltage results in a linear rise in current through the yoke inductance $L_y$ over the entire scan interval from $T_0$ to $T_5$.

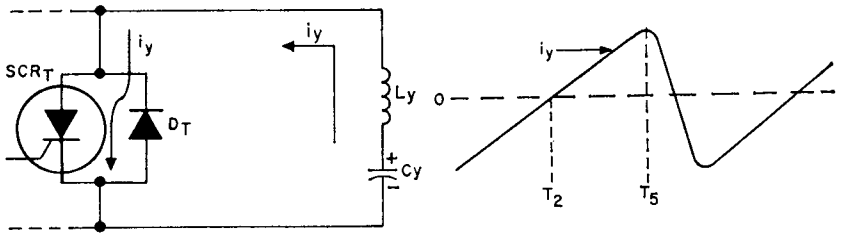



Figure 732. Effective configuration of the deflection circuit during the second half of the trace interval, time $T_2$ to $T_5$, and the complete scan-current waveform.

**Start of the retrace interval:** The circuit action to initiate retrace starts before the trace interval is completed. Fig. 733 shows the circuit elements and the voltage and current waveforms required for this action. At time $T_3$, prior to the end of the trace period, the commutating-switch controlled rectifier $SCR_C$ is turned on by application of a pulse from the horizontal oscillator to its gate. Capacitor $C_R$ is then allowed to discharge through $SCR_C$ and inductor $L_R$. The current in this loop, referred to as the commutating circuit, builds up in the form of a half-sine-wave pulse. At time $T_4$, when the magnitude of this current pulse exceeds the yoke current, the trace-switch diode $D_T$ again becomes forward-biased. The excess current in the commutating pulse is then bypassed around the yoke winding by the shunting action of diode $D_T$. During the time from $T_4$ to $T_5$, the trace-switch controlled rectifier $SCR_T$ is reverse-biased by the amount of the voltage drop across diode $D_T$. The trace-switch controlled rectifier, therefore, is turned off during this interval and is allowed to recover its ability to block the forward voltage that is subsequently applied.

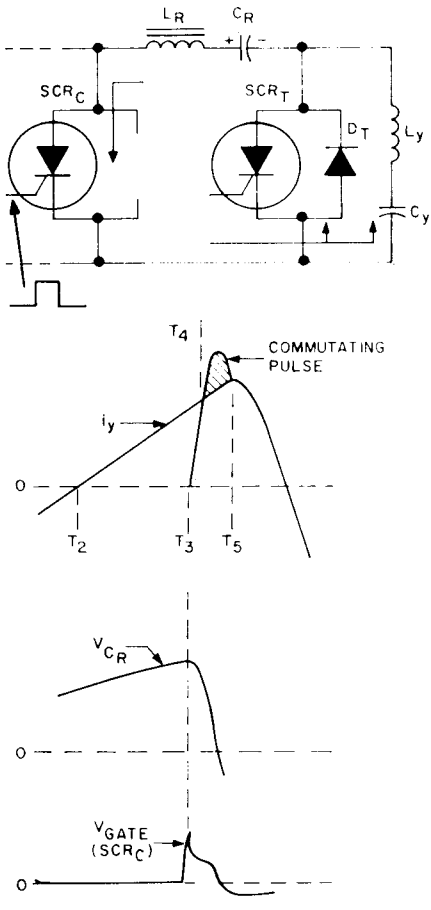



Figure 733. Effective configuration of the deflection circuit and significant voltage and current waveforms for initiation of retrace, time $T_3$ to $T_5$.

**First half of the retrace interval:** At time $T_5$, the commutating pulse is no longer greater than the yoke current, as shown in Fig. 734; trace-switch diode $D_T$ then

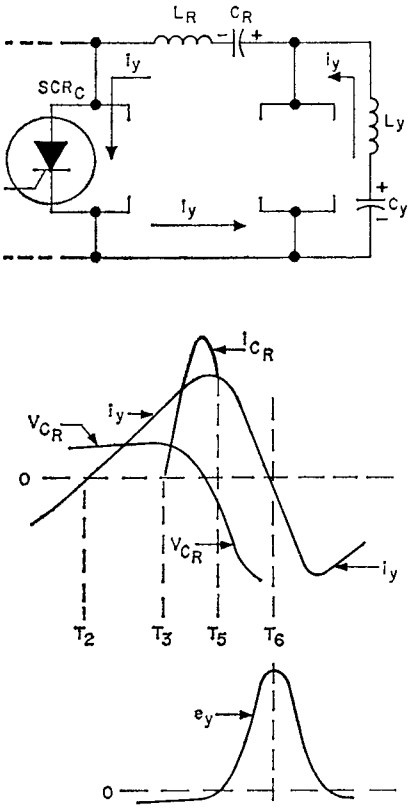



Figure 734. Effective configuration of the deflection circuit and operating voltage and current waveforms during the first half of retrace, time $T_5$ to $T_6$.

ceases to conduct. The yoke inductance maintains the yoke current but, with $SCR_T$ in the off state, this current now flows in the commutating loop formed by $L_R$, $C_R$, and $SCR_C$. Time $T_5$ is the beginning of retrace.

As the current in the yoke windings decreases to zero, the energy supplied by this current charges capacitor $C_R$ with an opposite-polarity voltage in a resonant oscillation. At time $T_6$, the yoke current is zero, and capacitor $C_R$ is charged to its maximum negative-voltage value. This action completes the first half of retrace.

**Second half of the retrace interval:** At time $T_6$, the energy in the yoke inductance is depleted, and the stored energy on the retrace capacitor $C_R$ is then returned to the yoke inductance. This action reverses the direction of current flow in the yoke. During the reversal of yoke current, the commutating-switch diode $D_C$ provides the return path for the loop current, as indicated in Fig. 735. The commutating-

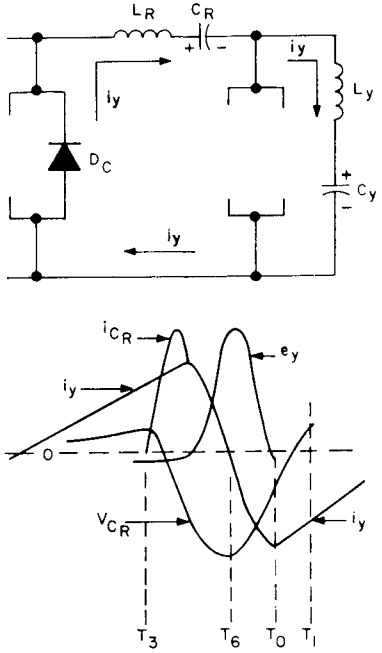



Figure 735. Effective configuration of the deflection circuit and operating voltage and current waveforms during the second half of retrace, time $T_6$ to $T_0$.

switch controlled rectifier $SCR_C$ is reverse-biased by the amount of the voltage drop across diode $D_C$. The commutating-switch controlled rectifier, therefore, turns off and recovers its voltage-blocking capability. As the yoke current builds up in the negative direction, the voltage on the retrace capacitor $C_R$ is decreased. At time $T_0$, the voltage across capacitor $C_R$ no longer provides a driving voltage for the yoke current to flow in the loop formed by $L_R$, $C_R$, and $L_y$. The yoke current finds an easier path up through trace-switch diode $D_T$, as shown in Fig. 736. This action represents the beginning of the trace period for the yoke current (i.e., the start of a new cycle of operation), time $T_0$.

Once the negative yoke current is decoupled from the commutating loop by the trace-switch diode, the current in the commutating circuit decays to zero. The

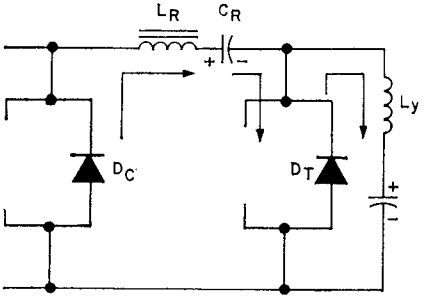



Figure 736. Effective configuration of the deflection circuit during the switchover from retrace to trace, time $T_0$.

stored energy in the inductor $L_R$ charges capacitor $C_R$ to an initial value of positive voltage. Because the resonant frequency of $L_R$ and $C_R$ is high, this transfer is accomplished in a relatively short period, $T_0$ to $T_1$, as shown in Fig. 735.

**Recharging and resetting actions:** The action required to restore energy to the commutating circuit and to reset the trace SCR are also very important considerations in the operation of the basic deflection circuit. Both actions involve the inductor $L_{CC}$.

During the retrace period, inductor $L_{CC}$ is connected between the dc supply voltage (B+) and ground by the conduction of either the commutating-switch SCR or diode ($SCR_C$ or $D_C$), as indicated in Fig. 737. When the

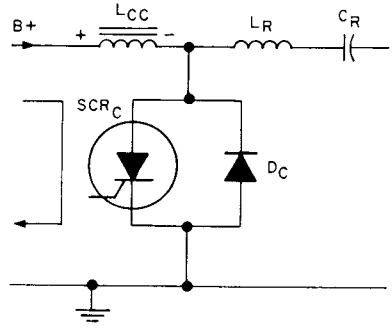



Figure 737. Circuit elements and current path used to supply energy to the charging choke $L_{CC}$ during period from the start of retrace switching action to the end of the first half of the retrace interval, time $T_3$ to $T_1$.

diode and the SCR cease to conduct, however, the path from $L_{CC}$ to ground is opened. The energy stored in inductor $L_{CC}$ during the retrace interval then charges capacitor $C_R$ through the B+ supply, as shown in Fig. 738. This charging process continues through the trace period until retrace is again initiated. The resultant charge on capacitor $C_R$ is used to resupply energy to the yoke circuit during the retrace interval.

The voltage developed across inductor $L_{CC}$ during the charging

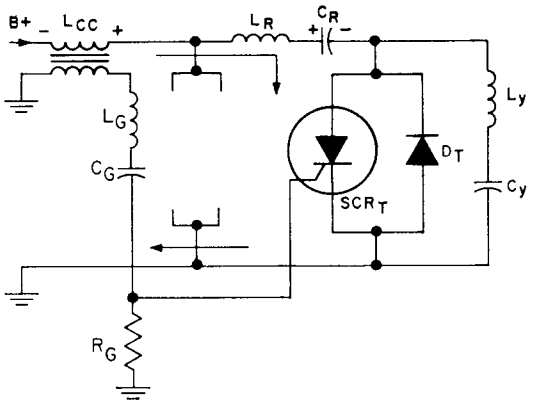



Figure 738. Effective configuration of the deflection circuit for resetting (application of forward bias to) the trace SCR and recharging the retrace capacitor $C_R$, during time interval from $T_1$ to $T_3$.

of capacitor $C_R$ is used to forward-bias the gate electrode of the trace SCR properly so that this device is made ready to conduct. This voltage is inductively coupled from $L_{CC}$ and applied to the gate of $SCR_T$ through a wave-shaping network formed by inductor $L_G$, capacitor $C_G$ and resistor $R_G$. The resulting voltage signal applied to the gate of $SCR_T$ has the desired shape and amplitude so that $SCR_T$ conducts when a forward bias is applied from anode to cathode, approximately midway through the trace interval.

**Effect of auxiliary capacitor $C_A$:** In the preceding discussions of the operation of the deflection circuit, the effect of capacitor $C_A$ was neglected. Inclusion of this capacitor affects some of the circuit waveforms, as shown in Fig. 739, aids in the turn-off of the trace SCR, reduces the retrace time, and provides additional energy-storage capability for the circuit.

During most of the trace interval (from $T_0$ to $T_4$), including the interval ($T_3$ to $T_4$) during which the commutating pulse occurs, the trace switch is closed, and capacitor $C_A$ is in parallel with the retrace capacitor $C_R$. From the start of retrace at time $T_4$ to the beginning of the next trace interval at time $T_0$, the trace switch is open. For this condition, capacitor $C_A$ is in series with the yoke $L_y$ and the retrace capacitor $C_R$ so that the capacitance in the retrace circuit is effectively decreased. As a result, the resonant frequency of the retrace is increased, and the retrace time is reduced.

The auxiliary capacitor $C_A$ is also in parallel with the retrace inductor $L_R$. The waveshapes in the deflection circuit are also affected by the resultant higher-frequency resonant discharge around this loop. The voltage and current waveforms shown in Fig. 739 illustrate the effects of the capacitor $C_A$.

The auxiliary capacitor $C_A$ also helps to prevent the fast-rise-

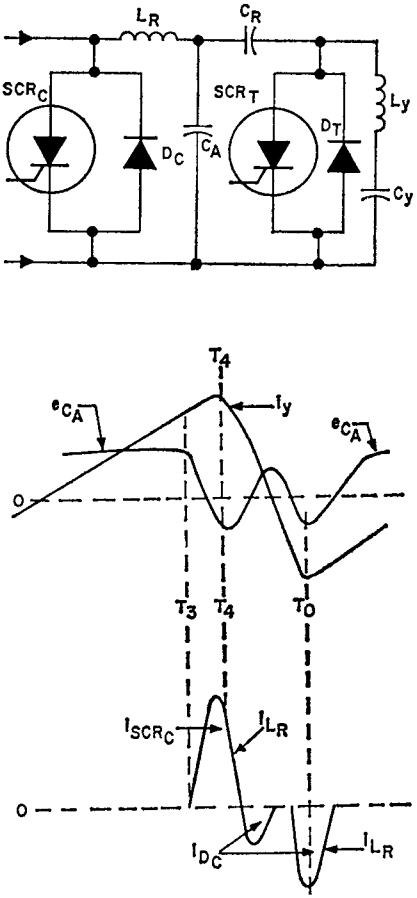



Figure 739. Circuit configuration showing the addition of auxiliary capacitor $C_A$ and current and voltage waveforms showing the effect of this capacitor.

time voltages developed by the flyback transformer from appearing across the trace switch. Fig. 740 shows the basic trace-switch circuit and the waveforms developed across this circuit with and without the auxiliary capacitor. These waveforms show that the integrating action of the auxiliary capacitor eliminate the initial steep rise of the retrace voltage pulse.

**High-Voltage Generation**—The SCR horizontal-deflection system generates the high voltage for the picture tube in essentially the same manner as has been used for many years in television receivers, i.e., by transformation of the horizontal retrace pulse to a high voltage with a voltage step-up transformer and subsequent rectification of this stepped-up voltage. In common with other solid-state receiver designs, a solid-state voltage multiplier is used as the high-voltage rectifier. A high-voltage rectifier tube such as the 3CZ3 could also be used although the increased source impedance in the high-voltage transformer would result in slightly poorer high-voltage regulation.

**High-Voltage Regulation**—The use of a silicon voltage multiplier for the high-voltage rectifier, to-

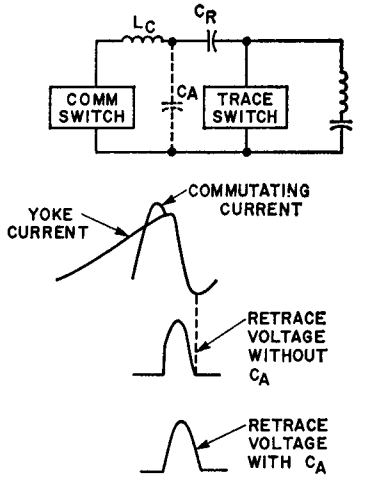



Figure 740. Simplified schematic of trace-switch circuit and waveforms showing effect of auxiliary capacitor on the rise time of the retrace voltage pulse.

gether with very tight coupling between the primary and secondary of the high-voltage transformer, results in a high-voltage system that has very low internal impedance. As a result, it is necessary to regulate the high voltage only against changes in line voltage. The regulator, shown in Fig. 741, is reactive (non-dissipating) and provides good reliability at low cost.

A supplementary winding on the high-voltage transformer provides a pulse voltage proportional

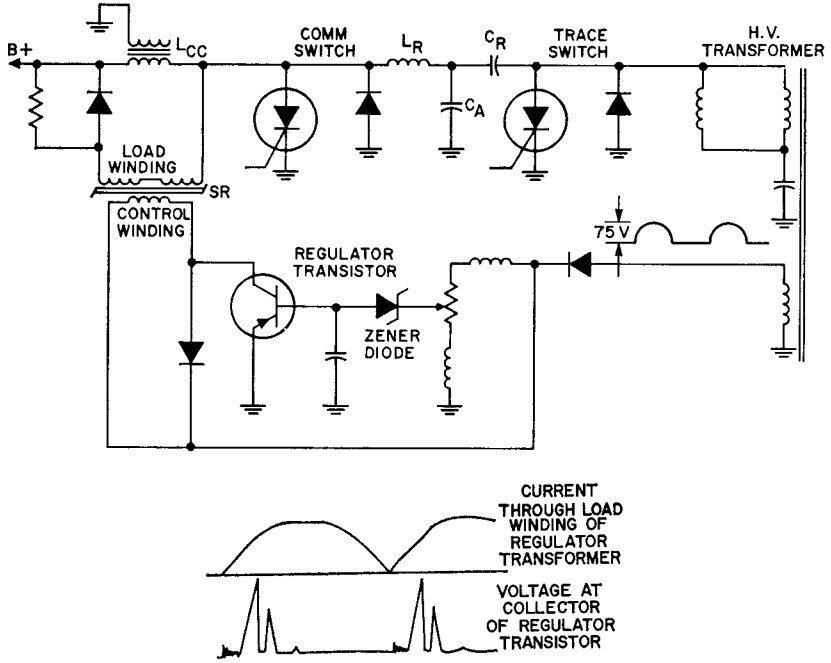



Figure 741. High voltage regulator circuit and operating voltage and current waveforms.

to the supply voltage. Because the supply voltage varies directly with line voltage, the voltage pulse provides an excellent reference for sensing variations in the line voltage. This voltage pulse is rectified and used as the collector voltage source for the regulator transistor. By means of a resistive voltage divider and the zener diode, it also provides base bias to the regulator transistor. When the voltage, as determined by the resistive divider, exceeds the zener voltage, the transistor conducts and current flows through the control winding of the saturable reactor. This current saturates the core of the saturable resistor (to a degree dependent upon the base voltage applied to the regulator transistor) and the inductance of the load winding drops sharply. Because the load winding is in parallel with the input reactor, $L_{CC}$, it limits the amount of energy that can be stored in $L_{CC}$ and, therefore, the amount of

energy that can be stored in $L_R$, $C_R$, and $C_A$. Consequently, as line voltage increases, the amount of energy stored in these components is limited, and the increase in high-voltage is limited accordingly. If line voltage decreases, the pulse voltage applied to the regulator circuit is reduced, less current is drawn by the regulator transistor, and the degree of core saturation of the saturable reactor is reduced. Consequently, more "relative" energy can be stored in $L_{CC}$, even though the input voltage to the system is reduced, and the high voltage remains constant.

The high-voltage regulator system, as mentioned previously, dissipates very little energy and keeps the high-voltage constant with variations in horizontal-oscillator frequency or with component values. The response time of the system is very short so that essentially every horizontal line is regulated. Although the major function of the system is to maintain a constant high voltage (and scan) with variations in line voltage, it does provide some supplemental regulation of high voltage with picture-tube beam current, because there is some variation in supply voltage with the power drawn from the high-voltage supply.

**Linearity Correction**—Some S-shaping can be obtained by the use of a capacitor in series with the yoke. The nonlinearity caused by the yoke resistance results in left-hand stretch and right-hand compression.

Fig. 742 illustrates two methods of linearity correction. In the circuit shown in Fig. 742(a), a damped series-resonant circuit is

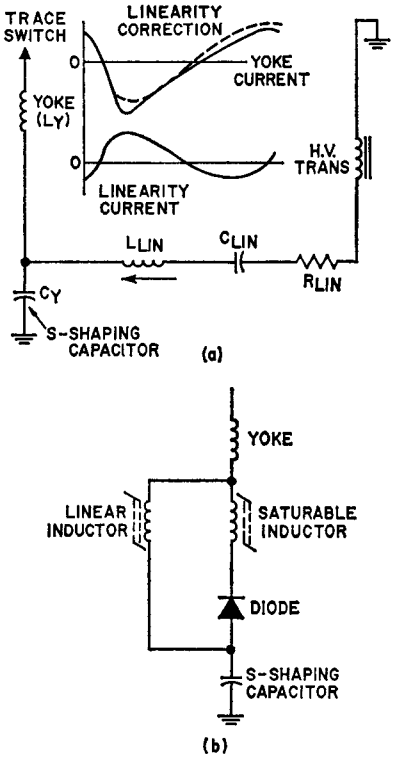



Figure 742. Two methods of linearity correction.

connected between an auxiliary winding on the high-voltage transformer and the ungrounded side of the S-shaping capacitor. This circuit produces a damped sine-wave of current which effectively adds to and subtracts from the charge on the S-shaping capacitor, thus altering the yoke current to correct for any trace-current nonlinearity. The circuit shown in Fig. 742(b) acts as a variable inductance in series with the yoke. Yoke current is blocked by the diode during the first part of trace and flows through the linear inductor. During the second part of trace, the diode becomes forward-biased and yoke

current is gradually shunted through the self-saturable inductor. With the proper values for the two inductors, the equivalent inductance in series with the yoke varies just the right amount to produce the proper degree of linearity correction.

**Raster Correction**—The distance from the center of deflection to the outside edge of the raster on the picture tube is greatest at the corners of the raster, decreasing to a minimum at the center. Because the electron beam must travel a greater distance to reach the corners of the raster, a given deflection of the beam produces a greater movement on the faceplate of the picture tube, and a type of distortion known as "pincushion" is produced. This effect is shown in Fig. 743. The degree of pincushion distortion increases with deflection angle. Correction of this type of raster distortion, therefore, is of greater importance with wide-angle picture tubes.

Correction of pincushion distortion can be accomplished by decreasing the deflection (yoke current) at the corners of the raster or by increasing the deflection at the center of the raster. The usual method is to reduce the yoke current as the beam approaches the corners of the raster. One method of pincushion correction is shown in Fig. 744. In this circuit, the collector supply for transistor $Q_2$ is the voltage across the 0.68-microfarad capacitor $C_1$ in the primary of the high-voltage transformer. Loading of this capacitor by transistor $Q_2$ increases the energy being drawn from the high-voltage transformer and thus reduces scan. Transistor $Q_2$ is driven by a vertical sawtooth. During the second half of vertical trace, the 10-microfarad capacitor $C_2$ is discharged by the collector current of $Q_2$, thereby loading the capacitor $C_1$ with an increasing current from the middle to the end of vertical scan. When transistor

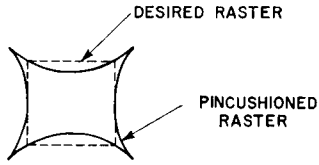



Figure 743. Effect of pinchushion distortion.

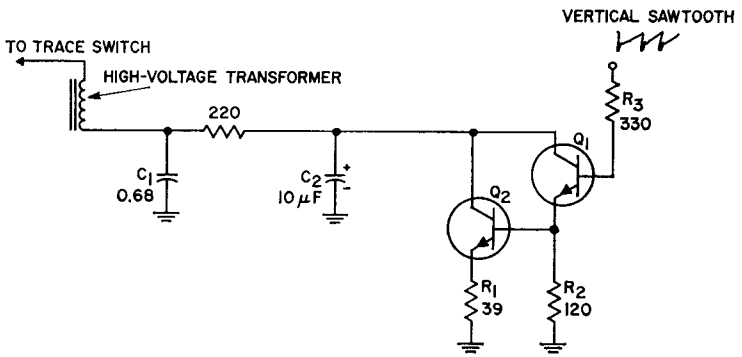



Figure 744. Active side-pincushion correction circuit.

$Q_2$ is turned off at the end of vertical scan, capacitor $C_2$ is again charged by the energy stored in the capacitor $C_1$. The loading of capacitor $C_1$ decreases toward zero from the top toward the middle of vertical scan as soon as the capacitor $C_2$ charges.

Another method of correcting pincushion distortion is shown in Fig. 745. In this circuit, the "control winding" of the saturable transformer is supplied with a vertical sawtooth (preferably somewhat parabolic in shape) which determines the degree of core saturation of the transformer and thus the impedance of the secondary windings, which

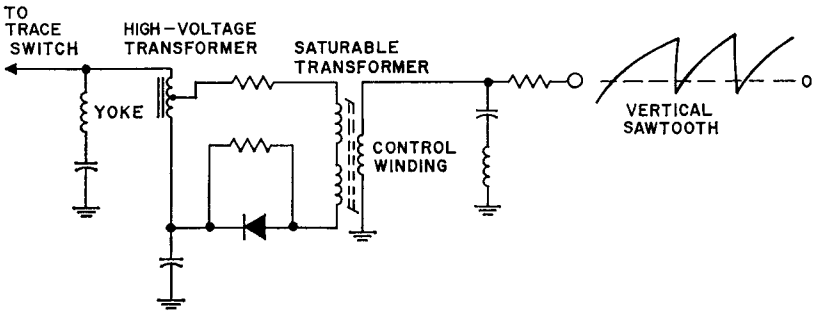



Figure 745. Saturable reactor side-pincushion correction circuit.

are shunted across a portion of the primary of the high-voltage transformer. During the second half of vertical scan, the current through the control winding gradually increases from zero at the center of scan to maximum at the end of scan. The degree of core saturation also increases, and the inductance of the windings in shunt with the high-voltage transformer (and yoke) decreases accordingly. The loading of the high-voltage transformer thus gradually increases from a minimum at the center of vertical scan to a maximum at the end of scan. At the beginning of scan, the current through the control winding is at a maximum, and the degree of loading of the yoke is maximum. The current in the control winding (and the yoke loading) gradually decreases to a minimum toward the center of scan.

**Auxiliary Power Supplies**—An important area of potential cost reduction in the SCR deflection system is the power supply. SCR's have much greater current-carrying capability than that required for deflection. This extra capability can be used to derive the operating power for other portions of the receiver from the horizontal-deflection system. Fig. 746 shows several possible methods of deriving power from the deflection system.

An advantage of the circuits shown in (c) through (e) of Fig. 746 is that they are regulated against changes in line voltage by the same regulator circuit used for regulating scan against changes in line voltage.

Rectifiers used to obtain dc power from the horizontal-deflection system operate at 15.75 kHz. Therefore, fast-recovery types must be used.

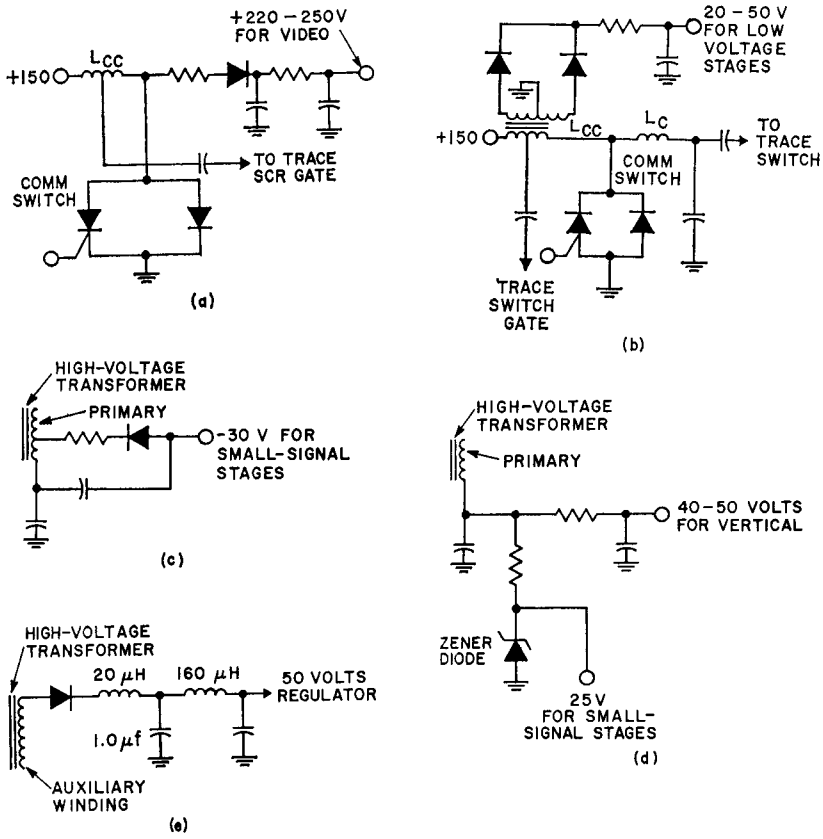



Figure 746. Various circuit arrangements for developing an auxiliary low voltage from the SCR deflection system.

### Over-all SCR Deflection System

—Fig. 747 shows the circuit diagram of a complete SCR horizontal deflection system. This system is designed to operate directly from the rectified line voltage. A 250-volt unregulated supply for the video amplifier is obtained by rectifying the voltage obtained from the input reactor ($L_{CC}$). A 40-volt supply (regulated) for the small-signal stages is obtained by rectifying a pulse obtained from an auxiliary winding on the high-voltage transformer.

## VERTICAL DEFLECTION

The vertical-deflection circuit in a television receiver is essentially a class A audio amplifier with a complex load line, severe low-frequency requirements (much lower than 60 Hz), and a need for controlled linearity. The equivalent low-frequency response for a 10-per-cent deviation from linearity is 1 Hz. A simple circuit configuration is shown in Fig. 748.

The required performance can be obtained in a vertical-deflec-

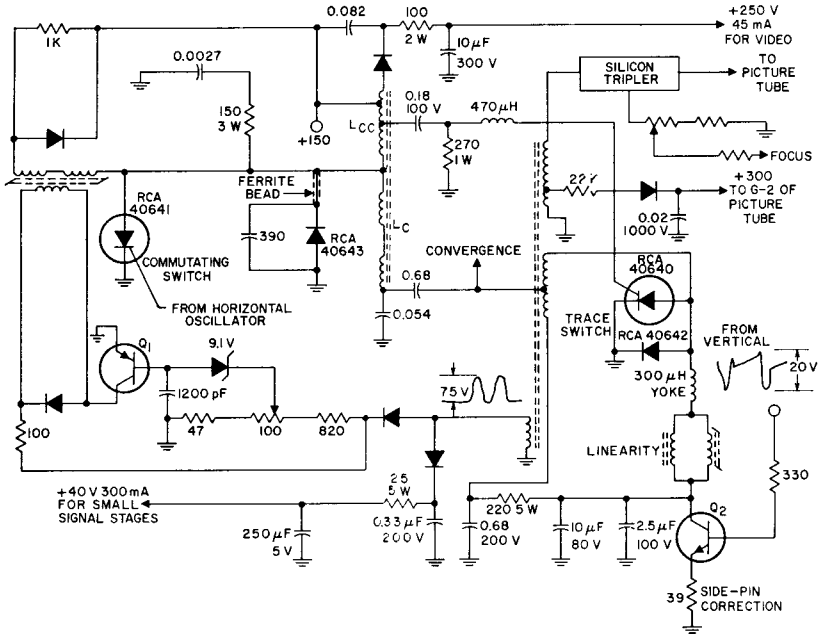



Figure 747. Line-operated, SCR horizontal-deflection system for 90-degree color picture tube.

tion circuit in any of three ways. The amplifier may be designed to provide a flat response down to 1 Hz. This design, however, requires an extremely large output transformer and immense capacitors. Another arrangement is to design the amplifier for fairly good low-frequency response and predistort the generated signal.

The third method is to provide extra gain so that feedback techniques can be used to provide linearity. If loop feedback of 20 or 30 dB is used, transistor gain variations and nonlinearities become fairly insignificant. The feedback automatically provides the necessary "predistortion" to correct low-frequency limitations. In addition, the coupling of miscellaneous signals (such as power-supply hum or horizontal-deflection signals) in the amplifying loop is suppressed.

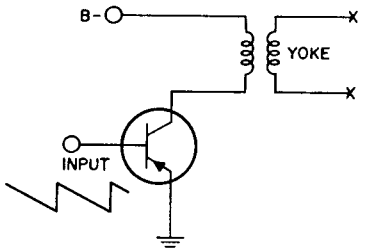



Figure 748. Simple vertical-deflection circuit.

The inductance of the output transformer must be fairly low for maximum efficiency. When a circuit is designed for maximum efficiency, the transistor dissipation must be at least three times the yoke power. When interchangeability, line-voltage varia-

tions, and bias instability are considered, the dissipation may reach high levels (e.g., 14 watts in a 25-inch color receiver); as a result, expensive bias techniques and extruded-aluminum heat sinks must be used.

Use of a toroid yoke having an L/R time constant of 3.2 milliseconds reduces the maximum dissipation to 3 or 4 watts and allows the plated steel chassis to be used as the heat sink for the transistor. The output transformer may also be reduced in size.

The higher Q of the toroid yoke normally results in a long retrace time or a very high flyback voltage.

## Basic Design Approach

In recent commercial television receivers, the Miller-integrator concept is employed in the generation of the linear ramp of current required in the vertical-deflection yoke. Fig. 749 shows the basic configuration of a Miller-integrator type of vertical-deflection circuit. In this circuit, a high-gain amplification system is used to develop the drive current for the yoke winding, and the integrating capacitor is connected in shunt with the yoke and the amplifier system. In effect, the Miller circuit multiplies the capacitor charging current by a factor equal to the gain of the amplifier without feedback. This technique results in an extremely linear output current waveform. In addition, variations in supply voltage, amplifier gain, and other factors that drastically affect the output of conventional vertical-deflection circuits have but slight adverse effects in the Miller circuit because of the large degenerative feedback.

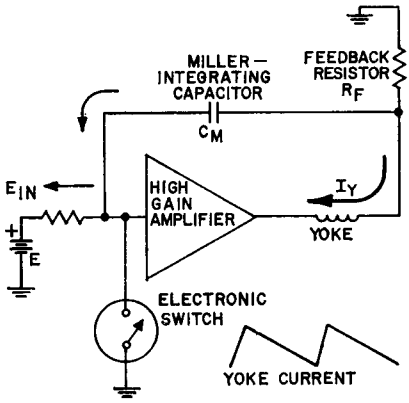



Figure 749. Basic Miller Sweep Circuit.

At the beginning of the vertical-trace interval, the integrating capacitor $C_M$ is charged from a voltage source E. The resulting voltage across the capacitor causes the amplifier to supply current to the yoke winding and to the feedback resistor $R_F$, which is directly coupled to the integrating capacitor. The feedback action of the integrating capacitor tends to maintain a constant input to the amplifier so that the voltage across the capacitor builds up (integrates) at a constant rate. Because the voltage across the feedback resistor, which is essentially the same as the voltage across the integrating capacitor, is directly proportional to the yoke current, the yoke current increases at a constant rate, and a linear scan results. The sweep rate is determined by an electronic switch which discharges the integrating capacitor at the end of each scan period.

The amplifiers used in the vertical-deflection system are similar to those used in any high-gain

audio-amplifier system. Either conventional transformer-coupled types or transformerless true-complementary-symmetry or quasi-complementary-symmetry types may be used. The following paragraphs describe the use of different types of output amplifiers and their associated circuitry in vertical-system applications.

## Vertical Circuit that Uses a Conventional Output Stage

Fig. 750 shows the basic functional relationship among the various stages of a Miller-integrator vertical-deflection circuit

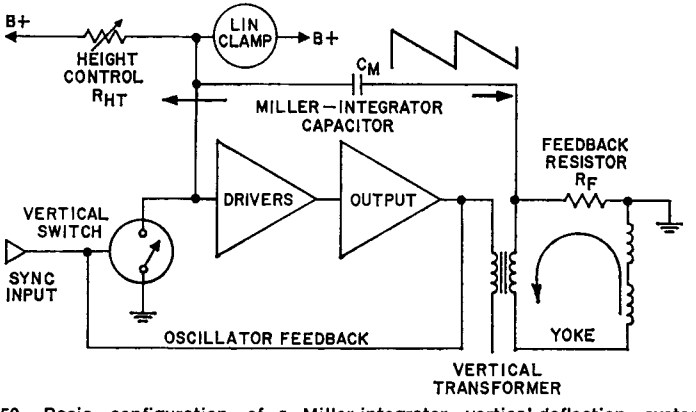



Figure 750. Basic configuration of a Miller-integrator vertical-deflection system that uses a conventional transformer-coupled output stage.

used in a recent commercial color-television receiver. The vertical-switch circuit controls the trace and retrace times and, therefore, the over-all operating frequency of the circuit. The switching action of the vertical switch is made self-sustaining by use of positive feedback from the output stage. Vertical synchronizing pulses applied to the switch from the sync separator determine the exact instant at which the switch is triggered on and, in this way, synchronize the switching action with the transmitted scanning interval. The Miller high-gain amplification system includes predriver and driver stages in addition to a conventional transformer-coupled output power-amplifier stage. The Miller-integrator capacitor is connected between the yoke winding and the input to the predriver so that it shunts the gain stages. The linearity-clamp circuit provides the initial charging current for this capacitor.

**Vertical Switch**—The vertical switch discharges the Miller-integrator capacitor at the end of the vertical scanning interval and, in this way, causes beam retrace and prepares the circuit for a subsequent scanning interval. Fig. 751 shows the schematic diagram and operating waveforms for the vertical-switch circuit. The operation of the circuit is made self-sustaining by two feedback signals.

One feedback signal is applied to the base of the vertical-switch

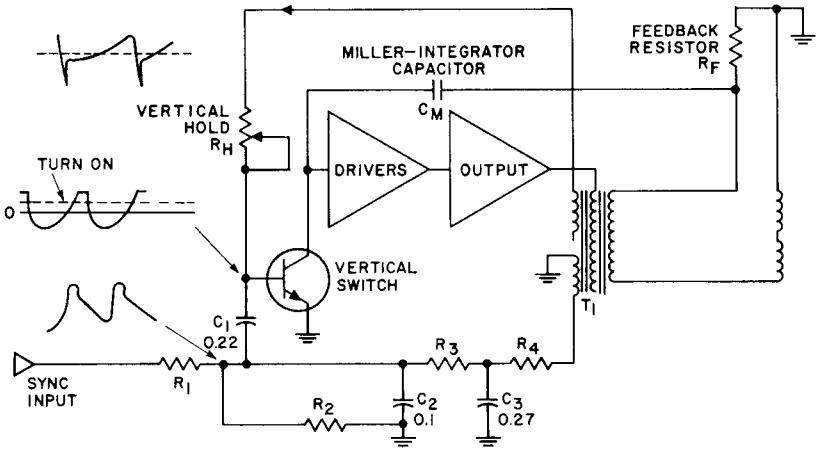



Figure 751. Vertical-switch circuit.

transistor from a secondary winding on the vertical-output transformer through resistors $R_3$ and $R_4$. This feedback signal is referred to as the triggering or turn-on pulse. The vertical synchronizing pulses from the sync separator are integrated by resistors $R_1$ and $R_2$ and capacitor $C_2$ and added to the triggering pulse.

Another feedback signal from a different secondary winding on the vertical-output transformer is applied to the base of the switch transistor through the vertical-hold potentiometer $R_H$. The addition of this waveform to the turn-on waveform causes the voltage at the base of the switch transistor to pass very quickly through the transistor turn-on voltage. As a result, the turn-on action of the vertical switch is very stable and relatively immune to noise voltages. The vertical-hold potentiometer provides some control over the shape of the latter feedback waveform and, therefore, offers limited control over the exact point at which the switch turns on.

**Driver Stages**—Two common-emitter stages (predriver and driver) provide the amplification required to increase the amplitude of the vertical-switch output sufficiently to drive the vertical-output stage. Fig. 752 shows a simplified circuit diagram of the driver stage.

The vertical predriver employs an n-p-n transistor $Q_3$ that is directly coupled to the p-n-p transistor $Q_2$ used in the driver stage. The emitter supply voltage for the driver is obtained from the voltage-divider network formed by resistors $R_5$ and $R_6$. The collector load of the driver consists of the parallel combination of the 680-ohm resistor $R_4$ and the base-emitter junction of the output-stage transistor $Q_1$. The service switch $S_1$ included in the emitter circuit of the driver can be used to cut off the vertical scanning during set-up adjustments of the picture tube if desired. When this switch is closed, the emitter of the driver is shorted to ground, and no vertical-deflection signals are developed.

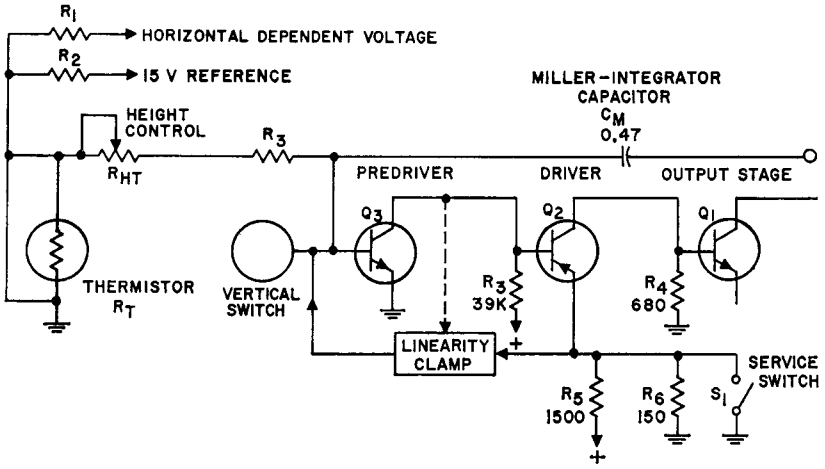



Figure 752. Vertical predriver and driver stages.

The predriver input waveform is supplied by the charging action of the Miller-integrator capacitor $C_1$, which is charged through the height-control potentiometer $R_{HT}$. The height-control supply voltage is made relatively immune to temperature-caused variations by the thermistor $R_T$. This supply also receives some dynamic regulation from a voltage supplied from the horizontal-deflection system. The addition of this regulating voltage helps to maintain a constant vertical height with respect to horizontal-scan and high-voltage variations.

**Vertical Output Stage**—Fig. 753 shows the circuit details for the output stage of the vertical system. This stage, which is directly driven by the driver circuit, uses a transistor operated in a common-emitter amplifier configuration to develop the power necessary to produce the required vertical deflection of the picture-tube beams. The collector load circuit consists of the vertical-output transformer $T_1$ and the vertical convergence circuitry. The secondary of the vertical-output transformer is loaded by the vertical yoke windings, two feedback paths, and the pincushion-correction circuitry. The Miller-integrator capacitor $C_M$ is coupled to the 5.6-ohm feedback resistor $R_F$, which is connected in series with the output-transformer secondary and the windings of the vertical-deflection yoke. Two feedback waveforms are provided from the output stage (from separate secondary windings on the output transformer) to the vertical switch to assure stable, self-sustaining switch operation.

The diode $D_1$ and the filter network formed by resistor $R_2$ and capacitor $C_1$ form a protective clamp circuit for the output transistor. Positive-going retrace pulses cause the diode $D_1$ to conduct and capacitor $C_1$ charges rapidly through the short-time-constant path provided by diode $D_1$ and resistor $R_2$. After the retrace pulse is removed, the capacitor attempts to discharge

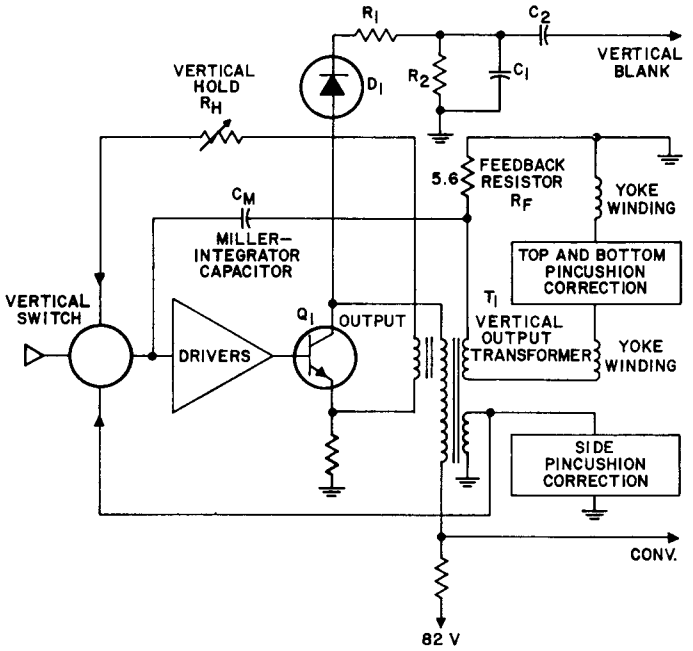



Figure 753. Transformer-coupled vertical output circuit.

through the resistor $R_2$. Because of the long-time-constant path provided by this resistor, the capacitor is only allowed to discharge an amount sufficient to assure a voltage differential across the diode when the retrace pulses occur. This action effectively clamps the collector output of transistor $Q_1$ to the voltage across capacitor $C_1$. The pulses that appear across this capacitor during the conduction of the diode are coupled by capacitor $C_2$ to the television-receiver video-amplifier circuit for use in vertical-retrace blanking.

**Linearity Clamp**—A circuit referred to as the linearity clamp is included in the vertical-deflection system to assure that sufficient initial-scan charging current is provided for the Miller-integrator capacitor. Fig. 754 illustrates the action of this circuit.

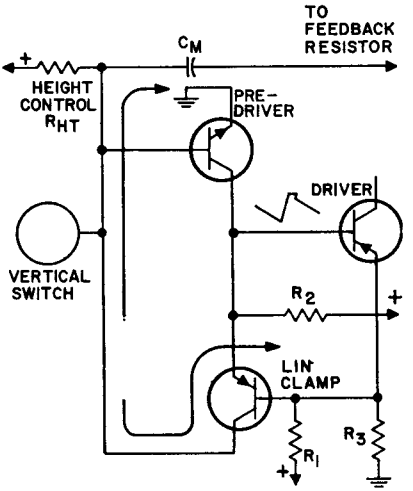



Figure 754. Linearity clamp.

When the Miller-integrator capacitor $C_M$ is discharged by the vertical switch at the end of a vertical-scan interval, the capacitor discharges into the base circuit of the predriver stage, and the predriver transistor is cut off. The positive voltage that then appears at the collector of the predriver transistor forward-biases the p-n-p linearity-clamp transistor, and current flows through this transistor, resistor $R_2$, and the vertical switch. After approximately 700 microseconds, the vertical switch turns off, and the current through the linearity clamp is used to provide rapid initial charging of the Miller-integrator capacitor $C_M$. As the initial charge quickly builds up on the capacitor, the predriver and driver stages start to conduct, and the base-emitter junction of the linearity-clamp transistor is reverse-biased by the voltage drop across the base-emitter junction of the driver transistor. This action cuts off the linearity-clamp circuit and initiates another vertical-scan interval. The Miller-integrator capacitor continues to charge through the height-control potentiometer $R_{HT}$ for the duration of the scan interval.

## Vertical Circuit that Uses Complementary-Symmetry Output Stage

The introduction of complementary pairs of power transistors has led to the development of class B transformerless output stages that are both economical and efficient. In vertical-output applications, such circuits may be capacitively coupled to the yoke, and the output transformer, together with the problems of nonlinearity, low-frequency phase shift, and excessive retrace pulse amplitudes associated with it, can be eliminated. Regardless of the type of output stage used, the generation of a linear sawtooth by use of the Miller-integrator circuit has become widespread.

Fig. 755 shows a block diagram of a vertical-deflection system of this type that uses a true-complementary-symmetry output stage.

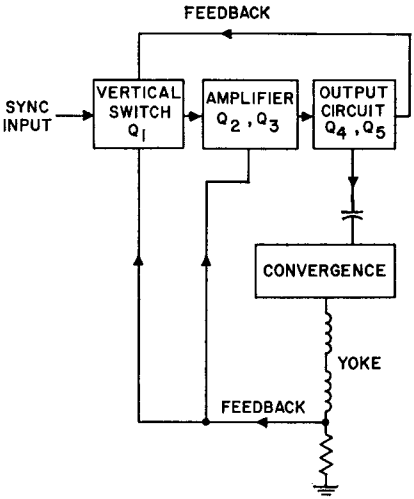



Figure 755. Block diagram of a vertical-deflection system that uses a true-complementary-symmetry output stage.

The vertical switch controls the free-running frequency of the vertical system. The high-gain amplifier consists of a direct-coupled predriver and driver, in addition to the true-complementary-symmetry output stage. The output stage is capacitively coupled to the convergence circuitry and the vertical-deflection yoke.

**Vertical-Switch and Predriver Circuit**—Fig. 756 shows the circuit configuration of the predriver circuit and its interconnection with the vertical-switch circuit. An increase in the posi-

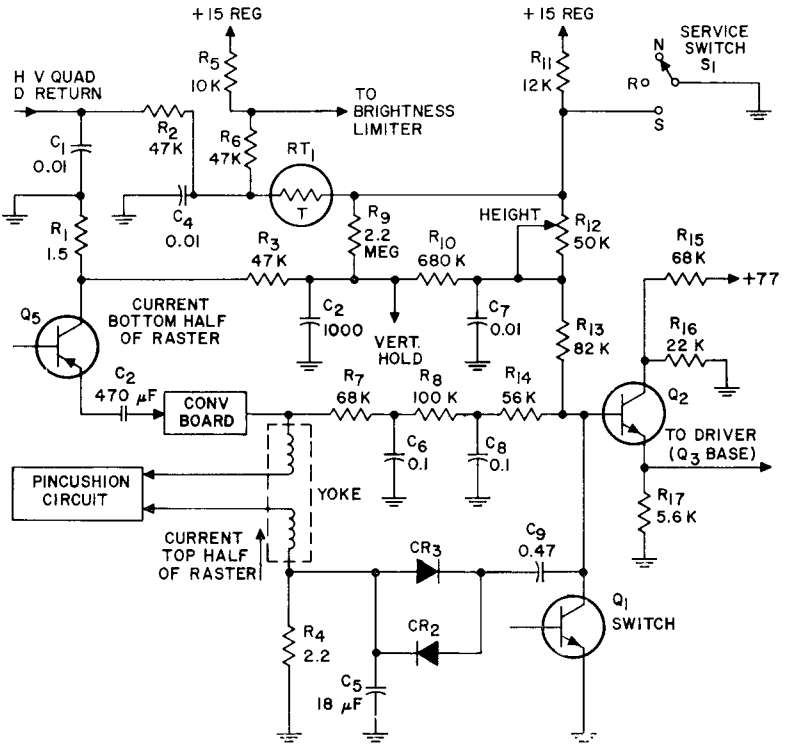



Figure 756. The vertical predriver with its inputs.

tive voltage at the junction of resistors $R_{12}$ and $R_{13}$ will increase the raster height. Because the major source of the voltage to $R_{13}$ is obtained from the +15-volt regulated supply through resistors $R_{11}$ and $R_{12}$, as the resistance of $R_{12}$ is decreased, raster height increases. If the setup switch $S_1$ is closed to the service position, the supply to $R_{13}$ is diminished practically to zero, and the raster collapses.

Because the output-stage transistor $Q_5$ is cut off while the top half of the raster is scanned, the voltage at the collector of this transistor is zero until vertical scan reaches the center. During the bottom half of scan, the collector current of output transistor $Q_5$ increases linearly, so that the voltage fed back to $R_{13}$ tends to "stretch" the lower part of the raster, to overcome some tendency towards bottom compression. Feedback to the vertical-switch transistor also is derived from resistor $R_1$.

The remaining input to resistor $R_{13}$ is obtained from the horizontal system. As high-voltage return current to the brightness limiter increases or decreases, the voltage at the junction of $R_5$ and $R_6$ also varies. An increase in picture-tube current, therefore, reduces slightly the voltage to

the height control and causes a slight decrease in vertical deflection. This action causes scanning height to track scanning width.

A feedback signal is fed to capacitor $C_9$ from the junction of the system feedback resistor $R_4$ and the yoke. If the effect of capacitor $C_5$ is ignored, the voltage at this point reaches its maximum positive value at the beginning of scan, passes through zero, and reaches maximum negative just before vertical retrace. Therefore, the feedback to the predriver transistor $Q_2$ is degenerative, because voltage at the base of $Q_2$ tends to rise throughout the scanning interval. Capacitor $C_5$ is used to filter out any horizontal-deflection voltage which may be present.

The transistor vertical-switch circuit shown in Fig. 757 performs three functions. It controls the free-running frequency of the vertical-deflection system, allows synchronization with the received signal, and determines the duration of vertical retrace. The overall vertical system may be considered as a free-running oscillator. The base of switch transistor $Q_1$ is returned to the supply voltage (height-control B+) through resistor $R_7$, the hold control, and a 680-ohm resistor $R_8$. If no sync pulses are present at the moment after the end of retrace, capacitor $C_6$ begins to charge, and the base of $Q_1$ begins to swing positive. When $Q_1$ begins conducting (about 17 milliseconds later), predriver and driver transistors $Q_2$ and $Q_3$, shown in Figs. 756 and 758, respectively, conduct less, and the output transistor $Q_4$, which was cut off during the lower half of vertical scan, resumes conduction. Because the voltage across the yoke inductance leads the current through it, a sharp positive pulse appears at the input to re-

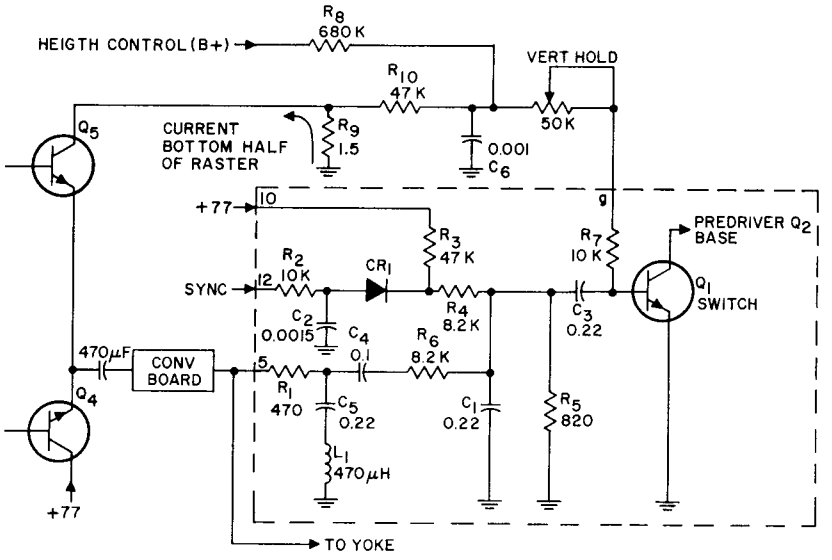



Figure 757. Vertical-switch circuit.

sistor $R_1$, and this pulse, coupled to the base of $Q_1$, drives $Q_1$ into saturation. This transistion of $Q_1$ from cutoff to saturation is very rapid.

Capacitor $C_5$ and inductor $L_1$, connected from the junction of resistor $R_1$ and capacitor $C_4$ to ground, are series resonant at the horizontal-scan frequency, and shunt to ground any 15.734-kHz energy which may be present. The presence of horizontal ripple at the vertical switch tends to synchronize the vertical scan with the horizontal scan and causes a degradation of interlace. Resistor $R_6$ and capacitor $C_1$ shape the feedback pulse so that the transition of $Q_1$ from cutoff to saturation is as rapid as possible.

When $Q_1$ saturates, $Q_4$ reaches maximum conduction, and the yoke current rises to maximum in the direction which produces maximum upward deflection. During retrace, the base current of transistor $Q_1$ charges capacitor $C_3$ negatively. The duration of the scanning is determined by the length of time required for the base of $Q_1$ to become forward-biased once more.

A second feedback circuit improves the frequency stability of the oscillator circuit. During the top half of scan, output transistor $Q_5$ is cut off, and the voltage at the junction of resistors $R_9$ and $R_{10}$ is essentially zero. Therefore, the voltage rise at the base of switch transistor $Q_1$ is exponential. But, as scan nears the bottom of the raster, transistor $Q_4$ conducts, and causes a positive voltage to be developed across resistor $R_9$. This voltage sharpens the voltage rise at the base of $Q_1$, so that its transistion from cutoff to saturation is more rapid. Similarly, the sharp drop in voltage across $R_9$ (from maximum to zero during the first half of retrace) enhances the cutoff characteristics of the $Q_1$ circuit.

The composite sync signal is introduced into the vertical system at terminal 12. Resistor $R_2$ and capacitor $C_2$ integrate the input so that the horizontal sync pulses are reduced in amplitude to about 8 volts and the vertical pulses about twice this amplitude. Since diode $CR_1$ has about 12 volts of positive bias on its cathode, only the vertical sync pulse can pass to the switch transistor. If the free-running frequency of the vertical system is slightly less than the vertical-sync rate, $Q_1$ is at the threshold of conduction when each sync pulse arrives, so that the vertical system is synchronized at the vertical-sync pulse rate.

**Vertical Driver and Output Stage**—Fig. 758 shows the schematic diagram of the vertical-system driver and of the output stage with the yoke circuit simplified. The circuit configuration is very similar to that of a high-quality audio power amplifier. The yoke itself is analogous to the speaker voice coil, $C_C$ is the coupling capacitor, and $R_y$ is the equivalent of the total resistance of the yoke and convergence circuit. The value of capacitor $C_C$ is selected to provide maximum energy transfer at the vertical scanning frequency. Feedback to the Miller capacitor is developed across resistor $R_{13}$, and capacitor $C_{10}$ is a filter.

During retrace, transistor $Q_3$ is cut off, and its collector voltage rises towards the supply voltage;

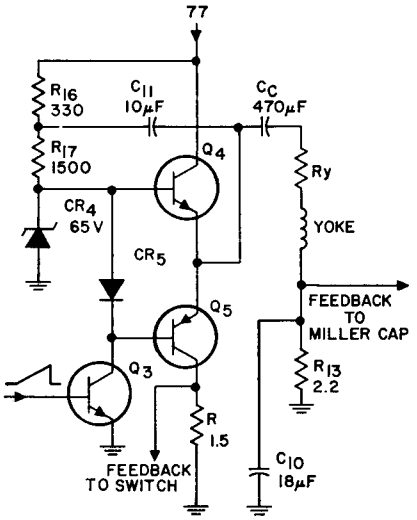



Figure 758. Vertical driver, output, and simplified yoke circuit.

however, the 65-volt zener diode $CR_4$ limits the maximum base bias of transistor $Q_4$ and, in this way, limits the yoke retrace current. During the scanning interval, the bases of transistors $Q_4$ and $Q_5$ are driven progressively less positive at a linear rate. Conduction is through $Q_4$ during most of the retrace time and as scan passes from the top of the raster to center. The voltage across capacitor $C_C$ at vertical scan center has reached maximum (90° out of phase with the current), and during the lower half of scan, capacitor $C_C$ discharges back through the yoke and transistor $Q_4$. This current increases at a linear rate, because the forward bias on the base of transistor $Q_4$ is increasing at a linear rate.

The diode connected between the bases of transistors $Q_4$ and $Q_5$ improves the switching characteristics of the transistors at mid-scan. $Q_5$ has zero bias as long as $Q_4$ is conducting. Therefore, only slight voltage swings are necessary to cut off $Q_4$ and turn on $Q_5$ at the center of the raster. If the diode were shorted or bypassed, reverse bias would exist between base and emitter of $Q_5$ while $Q_4$ was conducting, and consequently there would be appreciably more disturbance in the circuit during transition time.

## Vertical Circuit that Uses a Quasi-Complementary-Symmetry Output Stage

A disadvantage of the true-complementary-symmetry vertical-output circuit is the higher cost of p-n-p power transistors in comparison to n-p-n power transistors. Because control of the base diffusion is more difficult in p-n-p devices, their cost is generally 25 per cent more than comparable n-p-n devices.

Fig. 759 shows the complete circuit diagram for a vertical-deflection system that uses a quasi-complementary-symmetry output stage to drive a low-impedance toroidal yoke (L = 950 microhenries, R = 1.5 ohms). This system is basically the same as described earlier in which a true-complementary-symmetry output stage is used, with the exception of some minor modifications necessary to supply the higher deflection current required for the toroidal yoke.

Transistors $Q_3$ and $Q_5$ are functionally equivalent to the n-p-n output device in the true-complementary-symmetry circuit and transistors $Q_4$ and $Q_6$ function as an equivalent p-n-p device.

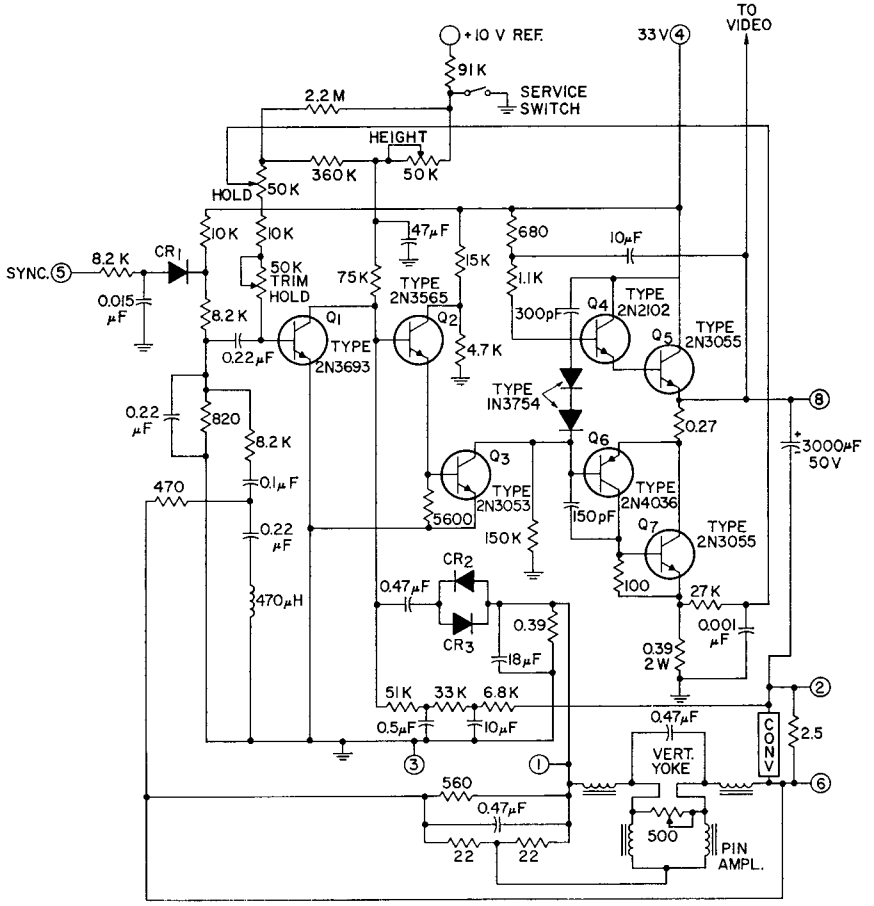



Figure 759. Complete transistor vertical-deflection system that uses a quasi-complementary-symmetry output stage.

# Ignition Systems

THE increasing use of transistor and thyristor ignition systems in the automotive industry has stimulated demands for improved performance, reduction in emissions that results because more accurate spark timing is achieved with magnetic-pickup distributors, and greater reliability. The following discussion covers the requirements for automobile ignition systems and compares the relative merits of capacitive and inductive types. Both systems are described in terms of operation, performance, and limitations; practical circuits are shown. The application of the capacitive-discharge principle in small-engine ignition circuits and in gas-system igniters is also described.

## BASIC CONSIDERATIONS FOR AUTOMOTIVE SYSTEMS

Under worst-case conditions, about 22 kilovolts are required to ignite the combustible mixture in the cylinder of an automobile engine. In addition, a minimum energy of about 20 millijoules must be available in the spark to assure propagation of a stable flame front originating at the spark. The exact values of voltage and energy required under all operating conditions depend on many factors, including those described in the following paragraphs.

### Condition of Spark Plugs

Fouled plugs reduce both the voltage and the energy available for ignition. The plug gap also affects both the voltage and the energy required. As the plug gap is increased, the required voltage increases, but the required energy decreases.

### Cylinder Pressure

The cylinder pressure depends on both the compression at the point of ignition and the air-fuel mixture. The minimum breakover voltage in any gas is a function of the product of gas pressure and electrode spacing (Paschen's Law). In automobile engines, the minimum voltage increases as this product increases. Therefore, higher pressures also require higher voltages. However, the energy required decreases as the pressure increases and increases as the fuel-air mixture

moves away from the optimum ratio. Worst-case conditions occur when the engine is started, at idle speeds, and during acceleration from a low speed because carburetion is poor and the fuel-air mixture is lean. The combination of a lower cylinder pressure and a dilute fuel-air mixture results in a high energy requirement under these conditions.

## Plug Polarity

The center electrode is hotter than the outside electrode because of the thermal resistance of the ceramic sleeve that supports it. If the center electrode is made negative, the effect of thermionic emission from this electrode can reduce the required ignition voltage by 20 to 40 per cent.

## Spark-Plug Voltage Waveshape

The spark-plug voltage waveshape is shown qualitatively in Fig. 760. The voltage starts to rise at point A, and reaches ignition at point B. The region from B′ to C represents the sustaining voltage for ionization across the spark plug. When there is insufficient energy left to maintain the discharge (at point C), current flow ceases and the remaining energy is dissipated by ringing.

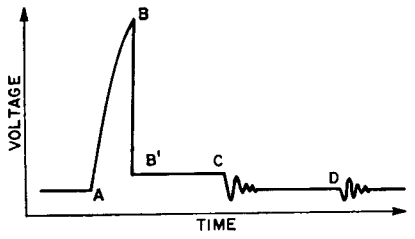



Figure 760. Ignition-voltage waveshape.

The final small spike at point D occurs when the ignition coil again starts to pass current.

The two most important characteristics of the voltage waveshape are its rise time (from A to B) and the spark duration (from B′ to C). A rise time that is too long results in excessive energy dissipation with fouled plugs; a rise time that is too short can lead to radiation losses of the high-frequency voltage components through the ignition harness. The minimum rise time should be about 10 microseconds; a 50-microsecond rise time is acceptable. Conventional systems have a typical rise time of about 100 microseconds. It should be noted that, at an engine speed of 5000 revolutions per minute, one revolution takes 12 milliseconds. Engine timing accuracy is usually no better than 2 degrees, which corresponds to 67 microseconds. The error caused by the rise time is therefore comparable to normal timing errors. At normal cruising speeds (about 2000 revolutions per minute), the 2-degree timing error corresponds to about 165 microseconds and rise-time effects are negligible.

## Energy Storage

The energy delivered to the spark plug can be stored in either an inductor or a capacitor. Although the inductive storage method is the more common approach, both are used; both are discussed. One requirement common to both methods is that, after the storage element is discharged by ignition, it must be recharged before the next spark plug is fired. For an eight-cylinder engine that has a dwell angle of 30 degrees, the time $\tau$ between ignition

pulses (in milliseconds) is equal to 15000 divided by the engine rpm, and the time $\tau_{ON}$ during which the points are closed is equal to 10000/rpm. When the engine rpm is 5000, $\tau_{ON}$ is 2 milliseconds. For either an inductive or a capacitive storage system, therefore, the charging-time constant should be small compared to 2 milliseconds.

## INDUCTIVE-DISCHARGE AUTOMOTIVE SYSTEM

Fig. 761 shows the basic circuit for an inductive-discharge system.

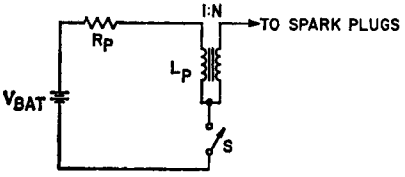



Figure 761. Basic inductive-discharge ignition circuit (Kettering system).

The total primary circuit resistance (ballast plus coil) is represented by $R_p$; the coil primary inductance is represented by $L_p$. Switch S represents the points in a conventional system. The transformer step-up turns ratio is N. When the points close, current increases exponentially with a time constant $\tau_L$ equal to $L_p/R_p$. The maximum primary current $I_p$ is equal to $V_{BAT}/R_p$, and the energy $\epsilon_L$ stored in the coil is equal to $L_pI_p^2/2$. When the points open, a voltage $V_P$ is generated across the primary terminals; this voltage is equal to $-L_p(dI_p/dt)$, where $I_p$ is the primary current as a function of time t. The secondary voltage $V_S$, which is delivered to the spark plugs through the distributor, is equal to $NV_p$.

The maximum current is limited to about 4 amperes by possible burn-out of the points. The total energy stored in the coil must be about 50 millijoules to provide for energy losses by radiation, fouled plugs, and the like. For a battery voltage of 12 volts and a primary circuit resistance of 3 ohms, $L_p$ must have a value of about 6 millihenries. The time constant $\tau_L$ is then about 2 milliseconds; the coil current does not reach its maximum value at high engine speeds. Fig. 762 shows primary current and secondary voltage as a function of engine speed for a typical nontransistorized ignition circuit. The degradation in secondary voltage follows the primary current. The available energy decreases even more rapidly because it is proportional to the square of the current. This problem can be even more severe than indicated because some conventional ignition coils have inductances as high as 12 millihenries, and the time constant is correspondingly longer.

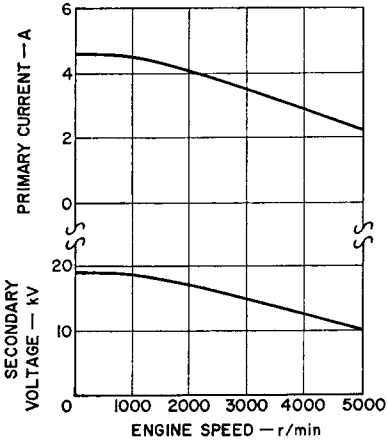


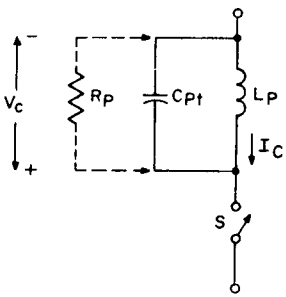

Figure 762. Performance of conventional inductive-discharge ignition circuit.

## Transistor Ignition System

In an inductive-discharge transistor ignition system, a transistor switch is used instead of the points to control the coil current. The points control the base current of the transistor, which may be only a few hundred milliamperes. As a result, burnout of points is eliminated. In addition, because a transistor can switch higher currents, a low-inductance coil can be used for the same coil energy. For example, a 10-ampere coil current (fairly typical for some systems) requires a coil inductance of only 1 millihenry. The corresponding primary resistance is then only about 1.2 ohms, and the value of $\tau_L$ is about 0.8 millisecond. This value is sufficiently short for good high-speed performance. Typical results are given later.

**Limitations**—One disadvantage of this type of transistorized ignition system is the higher current drain from the battery or alternator. However, this disadvantage can be more than offset by easier starting, better high-speed performance, and reduced maintenance. Another important consideration is temperature capability. Ambient temperatures under the hood of modern automobiles may reach 125°C, and the engine itself may reach 150°C. The junction temperature of a transistor may rise 10 or 15°C above ambient or case temperature. Available low-cost silicon transistors can operate at junction temperatures up to 200°C. In addition to increased reliability, these transistors also make feasible the use of active-mode switching. This factor in itself offers some important advantages over the saturated-mode switching techniques currently used in transistorized auto ignition systems.

**Ignition-Coil and Transistor Characteristics** — Because the transistor acts as a simple switch in series with the ignition coil, its required characteristics are largely determined by the characteristics of the coil itself. The most important coil parameters are shown in the equivalent circuit of Fig. 763. In this circuit, $L_p$ is the primary inductance, and $C_{pt}$ is the total primary capacitance (including all capacitance in the secondary circuit, referred to the primary circuit). For simplicity,




$$\epsilon = \frac{1}{2} L_P I_C^2$$

$$\epsilon = \frac{1}{2} C_{Pt} V_C^2$$

$$\therefore V_C I_C = \frac{2\epsilon}{\sqrt{L_P C_{Pt}}} = 4\pi f_o \epsilon$$

$$\text{WHERE } f_o = \frac{1}{2\pi\sqrt{L_P C_{Pt}}}$$

Figure 763. Simplified equivalent circuit of ignition coil primary.

this analysis neglects the coil primary series resistance, the leakage inductance, and the shunt resistance reflected into the primary

(that is, $1/N^2$ times the resistance across the secondary winding caused by fouled plugs, where N is the turns ratio). Switch S represents the transistor.

With switch S closed, the energy $\epsilon$ stored in the inductance is given by

$$\epsilon = \tfrac{1}{2} L_p I_c^2 \qquad (467)$$

where $I_C$ is the transistor collector current. When the switch is opened, this energy is transferred to the capacitor and is given by

$$\epsilon = \tfrac{1}{2} (C_p V_c^2) \qquad (468)$$

where the capacitor voltage $V_c$ also appears directly across the switch (transistor terminals). Therefore, the transistor must have a current-voltage rating given by

$$V_C I_C = 2 \epsilon/(I_p C_{pt})^{\frac{1}{2}} = 4\pi f_o \epsilon \qquad (469)$$

where $f_o$ is the resonant frequency of the circuit.

Eq. (469) represents a worst-case requirement because it assumes no shunt resistance across the primary. In practice, when the secondary voltage reaches the ionization potential, the plug fires and introduces an effective shunt resistance $R_p$ across the primary, as shown in Fig. 763. This damping resistance reduces the voltage developed across $C_{pt}$ because the energy available for charging $C_{pt}$ is less than that stored initially in $L_p$ by the energy dissipated in the spark plug. Shunt resistance introduced by fouled plugs also reduces the voltage developed across $C_{pt}$, and thus the output voltage.

The amount by which the voltage is reduced depends on the critical damping resistance of the ignition coil $R_{crit}$. This parameter is the value of resistance which, if placed in parallel with $L_p$ and $C_{pt}$ in Fig. 763, just prevents the voltage from oscillating. If the resistance caused by fouled plugs equals the critical value, the output voltage is reduced by a factor of 2.7. To avoid excessive reduction in output voltage, it is necessary that the actual shunt resistance be large compared with the critical resistance. Therefore, the coil must be designed for low critical resistance. The value of $R_{crit}$ is governed by the coil inductance and total capacitance, as shown by the following expression:

$$R_{crit} = [(L_p/4) C_{pt}]^{\frac{1}{2}} \qquad (470)$$

This relationship is another argument for a low value of primary inductance.

The discussion thus far has assumed that the transistor can support the full voltage $V_C$ given by Eq. (469). The $V_{CE}(sus)$ of the transistor (the sustaining value of the common-emitter avalanche breakdown voltage for the particular operating conditions used) must therefore be greater than $V_C$. If $V_{CE}(sus)$ is less than $V_C$, the transistor voltage is clamped to the $V_{CE}(sus)$ value until the secondary capacitance discharges sufficiently for the voltage to decrease below $V_{CE}(sus)$. The output voltage is limited to $NV_{CE}(sus)$.

Even more important is the possible destruction of the transistor. Although operation in the sustaining breakdown region is not necessarily destructive, the transistor may enter a second-breakdown region in which it can be destroyed. Operation in a particular circuit may bring even high-voltage transistors into the sustained break-

down region for a brief period. Therefore, the ability of a transistor to withstand second-breakdown destruction becomes an important device characteristic. Because the second-breakdown phenomenon is energy-dependent, the transistor can be characterized as safe for operation in the second-breakdown region in terms of energy, current, and circuit inductance, or in terms of power and pulse duration. Second breakdown can occur under conditions of either forward or reverse bias. In ignition systems, only the reverse-bias case is of interest. For reliable operation, the transistor must be able to withstand the total energy stored by the coil safely.

The current-voltage capability given in Eq. (469) can also be represented by the transistor load-line characteristic, as shown in Fig. 764. The transistor is driven into saturation (point A) and held there until an ignition pulse is required. When the transistor is turned off, a very fast voltage spike occurs initially (point B) because of the small amount of energy stored in the coil-primary leakage inductance. Leakage inductance was ignored in the derivation of Eq. (469).

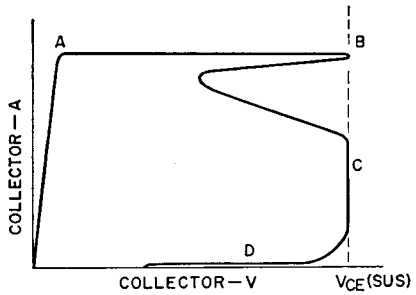



Figure 764. Typical transistor load line characteristic.

The initial pulse, which usually contains insufficient energy to damage the transistor, does not appear at the spark plug. However, the energy stored in the coil magnetizing inductance results in a second voltage pulse (point C) which is coupled to the ignition-coil secondary. The time required to switch from A to C may be in the order of microseconds. After the transistor current decreases to its leakage value (essentially zero, compared with the on current of several amperes), the collector voltage continues to oscillate (point D). The voltage may even exceed $V_{CE}(sus)$ and be limited instead by the transistor breakdown voltage $V_{(BR)CE}$. Most of the stored energy is transferred to the spark plug. Ignition takes place during this time because the build-up in coil voltage is determined by the resonant frequency of the parallel-tuned circuit of Fig. 763 and not by the transistor turn-off time. When ignition coils designed for transistorized operation are used, voltage rise times in the order of 50 to 100 microseconds can be achieved. The much shorter transistor turn-off time (1 or 2 microseconds) is then negligible.

In addition to the transistor characteristics discussed thus far (current, voltage, and reverse-bias second-breakdown capabilities; reverse breakdown voltage; and switching speed), there are two other characteristics of importance in ignition systems. The first is the saturation voltage $V_{CE}(sat)$ which, together with the current $I_p$, governs the transistor power dissipation. The other is the dc common-emitter forward-current transfer ratio $h_{FE}$, which determines the required base-drive

current and thus the current through the points in a one-transistor circuit.

## Practical Transistor Inductive-Discharge Circuits

Fig 765 shows a **two-transistor ignition circuit** for negative-ground automobiles. Current flows from the 12-volt dc supply through the coil and 1-ohm resistor, through $Q_2$ and $CR_1$, and then to ground. Transistor $Q_2$ is turned off when $Q_1$ is turned on, and vice versa. When the points close, $Q_1$ is turned off ($V_{BE\text{-}1} = 0$), and $Q_2$ is turned full on (is in saturation). The base-drive current for $Q_2$ is set by the 7.5-ohm resistor.

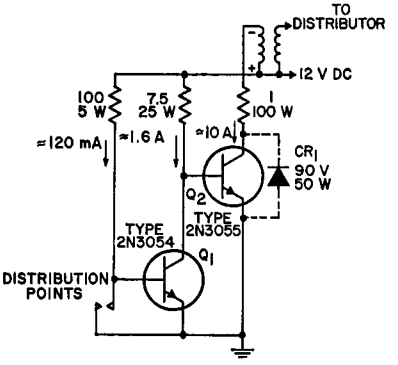



Figure 765. Transistor ignition circuit.

When the points open, base current through the 100-ohm resistor turns $Q_1$ on, drives it into saturation, and turns $Q_2$ off. The $V_{CE}(sat)$ of $Q_1$ is less than 0.25 volt, which is much less than the $V_{BE}$ turn-on threshold for $Q_2$ (about 0.6 volt). This relationship is necessary to assure that $Q_2$ is turned off. The zener diode $CR_1$ should be used to protect $Q_2$ from second breakdown.

The primary current and secondary voltage for the circuit of Fig. 765 are shown as functions of engine speed in Fig. 766. Two points should be noted. First, coil current and secondary voltage are nearly constant with engine speed, especially when compared with the curves for a nontransistorized system shown in Fig. 762. This improvement is a result of the much shorter inductive time constant $\tau_L$ for the transistorized system. Second, the output voltage decreases significantly under simulated fouled-plug conditions, although stored energy or coil current is independent of the shunt resistance. The decrease is caused by the energy lost in the 1-megohm shunt resistor used to simulate the fouled plugs.

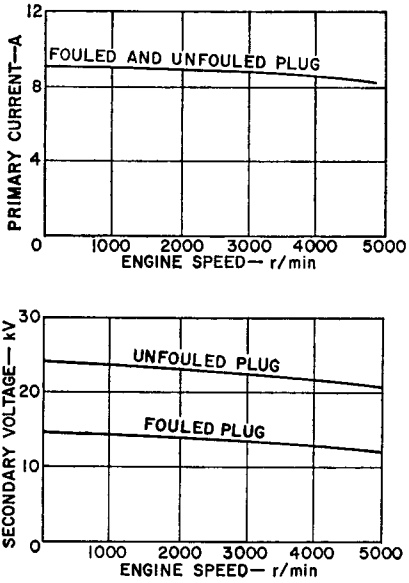



Figure 766. Performance of a transistorized ignition circuit.

Fig. 767 shows an **active-mode, current-regulated transistor ignition circuit** in its simplest form.

The circuit shown can be applied directly to a positive-ground system, or it can be modified for use with a negative-ground system. Because of its simplicity, this circuit has been used to test performance of experimental RCA transistors under varied conditions.

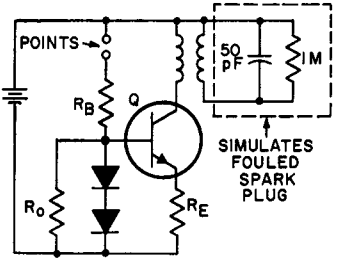



Figure 767. Active-mode, current-regulated circuit.

The operation of the circuit shown in Fig. 767 differs from that of most transistor ignition circuits in one important respect. For normal battery voltages and engine speeds, the transistor switches from the active region rather than from saturation. Current regulation is accomplished by means of current feedback through the emitter resistor. The two diodes in the base circuit provide a reference voltage that remains constant at 1.5 volts ± 5 per cent when the battery voltage varies from 8 to 16 volts. When the points close, current builds up in the coil and the transistor along the saturation line. The negative feedback voltage developed across the emitter resistor reduces the net base-to-emitter drive voltage. Thus, the transistor comes out of saturation and reaches a steady-state current in the active region. Fig. 768 shows waveforms for this type of high-current operation.

The upper curve in Fig. 768 shows the collector-current waveform. The center curve shows the collector-voltage waveform. When the points close, $V_C$ drops from the upper flat portion (equal to the battery voltage) to zero. The voltage then rises to a constant value during the on condition. When the points open, a large oscillating voltage is generated. The lower curve shows the operating load line on an expanded scale. For the particular test conditions used, approximately two milliseconds are required for the transistor to come out of saturation. This time corresponds to an engine speed of 5000 revolutions per minute in an eight-cylinder engine with a 30-degree dwell angle. For lower speeds, the coil current and, therefore, the output voltage remain constant, as discussed later.

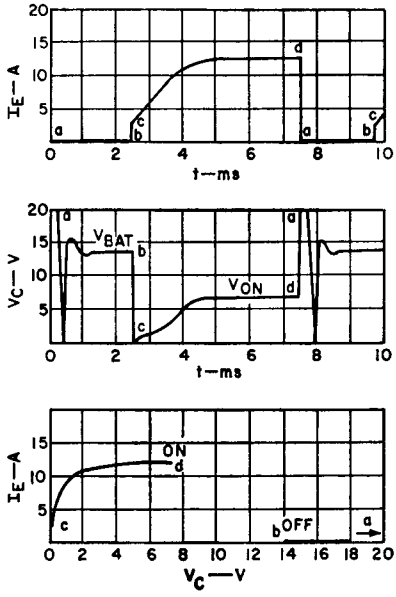



Figure 768. Typical waveforms for active-mode switching.

Advantages of this type of operation include excellent regulation with engine speed, reduction of transistor storage time during switching to the off condition, and great simplicity and design flexibility. Performance of the basic circuit shown in Fig. 767 is good for battery voltages down to 3.5 volts. Further improvement can be obtained if the base resistor $R_B$ is replaced by a second transistor which acts as a base-drive voltage regulator to compensate for battery-voltage variations.

The major disadvantage of this type of operation is the higher average power dissipation in the transistor. Although the higher temperature rise is not as serious for silicon transistors as for germanium transistors, it is still a drawback for normal active-mode switching. With some additional circuit modifications, however, it is possible to make the transistor on time a controlled function of engine speed. The advantages of active-mode switching can then be obtained without the chief disadvantage, the high average power dissipation.

The basic circuit of Fig. 767 has been used for both high-current and low-current operation. The choice is determined by the selection of the ignition coil and by the values used for $R_E$ and $R_B$, as follows:

| Component | High-Current | Low-Current | |
|---|---|---|---|
| $R_E$ | 0.02 | 0.1 | ohm |
| $R_B$ | 5 | 10 | ohms |
| $R_D$ | 220 | 220 | ohms |
| Coil: $L_p$ | 0.9 | 4 | mH |
| N | 0.3 | 0.9 | |
| $R_p$ | 240 | 113 | ohm |

Fig. 769 shows characteristic curves for the output voltage of the high-current circuit of Fig. 767 into a 50-picofarad, 1-megohm load as a function of engine speed for several different battery voltages. For normal cruising speeds

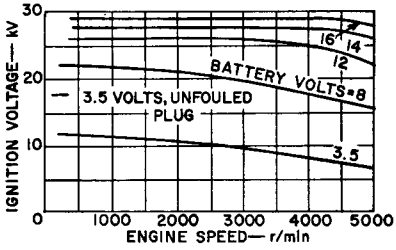



Figure 769. Ignition voltage as a function of engine speed for the high-current system.

and a battery voltage of 13.5 volts, the peak voltage is about 27 kilovolts. Even at 5000 revolutions per minute, the voltage is still about 25 kilovolts, or substantially greater than the normal 12 to 14 kilovolts required at this speed. Even at a battery voltage of 8 volts, the output voltage exceeds 15 kilovolts and satisfies normal requirements at this speed. Under extreme conditions (such as cranking the engine on a very cold morning), the voltage available from a normal 12-volt battery may be as low as 3.5 volts. At this low voltage level, the high-current circuit provides almost 12 kilovolts for a fouled plug (1-megohm load) and 17.5 kilovolts for an unfouled plug.

Fig. 770 shows the output-voltage waveform at two different time scales for a battery voltage of 16 volts and an engine speed of 5000 revolutions per minute. The output voltage is about 28 kilovolts and the rise time (measured between the 10- and 90-percent points) is only 60 microseconds. This short rise time minimizes timing errors and permits

the firing of fouled plugs. Fouled plugs require short rise times to reduce the amount of energy that is dissipated by the shunt resistance and thus made unavailable for ignition.

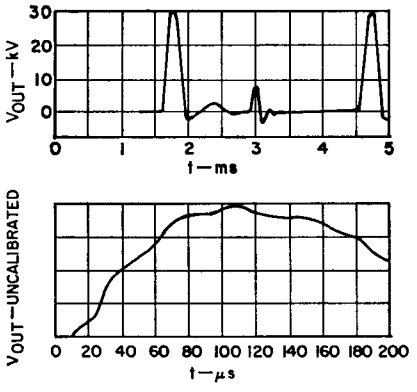



Figure 770. Output voltage for high-current operation.

Fig. 771 shows the collector current (top), collector voltage (middle), and load-line characteristic (bottom) for the same conditions. The peak voltage-current product is 4300 volt-amperes (250 volts and 18 amperes), as compared with a typical product of 1000 volt-amperes for most other auto-ignition transistors. At idling speeds, the average current drain and transistor power dissipation are excessively high in the simple test circuit of Fig. 767. However, controlled timing can be used to limit the transistor on time, for all engine speeds, to the minimum value required to just reach the desired current level. Such timing reduces both current drain and power dissipation to acceptable levels.

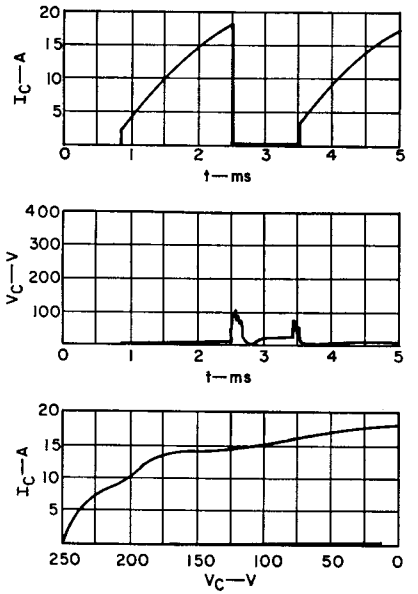



Figure 771. Collector current and voltage for high-current operation.

Fig. 772 shows the performance characteristics of the low-current circuit of Fig. 767 under fouled-plug conditions. At normal cruising speeds and a battery voltage of 13.5 volts, the peak voltage is about 25 kilovolts. At 5000 revolutions per minute, the voltage is about 16 kilovolts, which is still adequate for ignition. For the extreme case of the 3.5-volt supply, the output voltage is greater than 10 kilovolts for a fouled plug and is about 16.5 kilovolts for an unfouled plug.

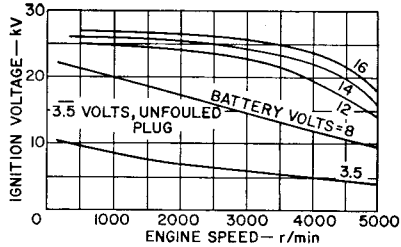



Figure 772. Ignition voltage as a function of engine speed for the low-current system.

The somewhat poorer high-speed performance of the low-current circuit results directly from the use of an ignition coil that has

higher primary inductance, as discussed previously. However, these performance characteristics are still superior to those of a conventional ignition system, which uses a coil of even higher inductance. A contributing factor to the improved performance is the use of current regulation in this transistorized system. Further improvements can be achieved by use of ignition coils with improved characteristics.

Fig. 773 shows the output-voltage waveform at two different time scales for a battery supply of 14 volts and an engine speed of 1000 revolutions per minute. The output voltage is 26 kilovolts, and the rise time is 66 microseconds.

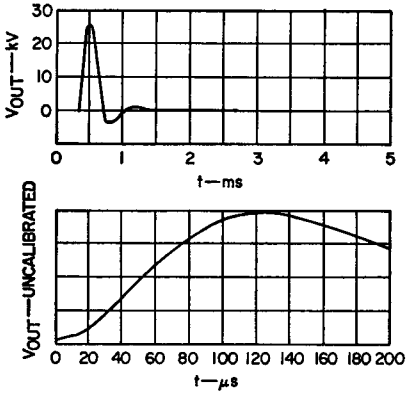



Figure 773. Output voltage for low-current operation.

The somewhat longer rise time, which results from the higher coil inductance, is still much shorter than the 100-to-150-microsecond rise time typical of conventional ignition systems.

Fig. 774 shows the emitter current (top), collector voltage (middle), and load-line characteristics (bottom) for the same conditions. The peak voltage-current product is only 1400 volt-amperes (220 volts and 7.6 amperes). This value, which is determined by the circuit operation, is well within the 4000-volt-ampere capability of the transistor.

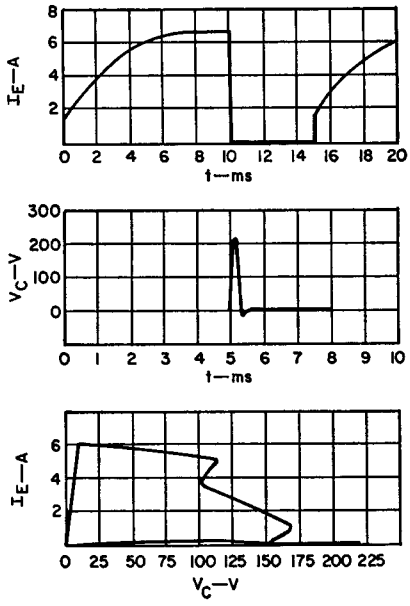



Figure 774. Emitter current and collector voltage for low-current operation.

## CAPACITIVE-DISCHARGE SYSTEMS

The basic capacitive-discharge system is illustrated in Fig. 775. It is important to note that the transformer serves simply as a pulse transformer. Therefore, performance at high engine speeds is not affected by the transformer primary inductance but, instead, is governed by the time required to charge capacitor C to the desired voltage level.

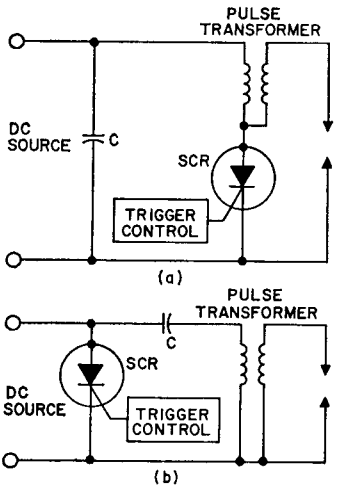



Figure 775. Basic configuration for capacitive-discharge ignition circuits: (a) storage capacitor connected across input voltage source; (b) storage capacitor connected in series with input voltage source and pulse transformer.

## Basic Circuit Operation

The trigger-control circuit (which can be a transistor switch) is controlled by the distributor points. More sophisticated distributor control, such as that available from distributors in which the voltage pulses are derived magnetically or photooptically, can also be used to control the trigger circuit. (Such distributors are also useful in the inductive-discharge transistor ignition circuits previously discussed.) The capacitor is charged to the dc voltage; the stored energy $\epsilon$ is equal to $C(V_c)^2/2$, where $V_c$ is the capacitor voltage. At the appropriate time, the trigger-control circuit fires the silicon controlled rectifier (SCR). The capacitor discharges through the transformer, which steps up the voltage to a value $V_s$ equal to $KNV_c$, where N is the transformer turns ratio and K is a constant that is dependent mainly on the value of the capacitance and of the transformer leakage inductance and generally ranges between 1 and 1.5. The stored energy is thus delivered to the spark plug in the form of a high-voltage pulse. Typical values for $V_c$ and C are about 350 volts and 1 microfarad, respectively. Thus the energy $\epsilon$ is about 60 millijoules.

Because the energy dissipated in the spark gap is equal to the energy stored in the capacitor minus the losses in the transformer and SCR, the energy available in the system is relatively easy to calculate. Examination of the basic circuits shows that the energy is transferred only when the SCR is forward conducting with the gate biased on. However, part of the energy is not available in the basic circuit because the capacitor and inductor form a tuned circuit when the SCR is on, and the energy that would normally flow back from the inductor to the capacitor is stopped by the high reverse impedance of the SCR. This energy is, therefore, lost as available spark energy. The duration of the spark is limited, then, to approximately one-half cycle of the natural LC frequency of oscillation. Some of the energy lost can be regained and used to increase the spark duration by installing a diode in the basic circuits of Fig. 775 as shown in Fig. 776. The diode not only bypasses the reverse impedance of the SCR but eliminates the possibility that the SCR might conduct in the reverse direction should the gate of the SCR be biased on at this time.

Thus, in addition to improving system low-temperature performance by increasing spark duration, the diode reduces the possibility of excessive heating and damage to the SCR that could accompany

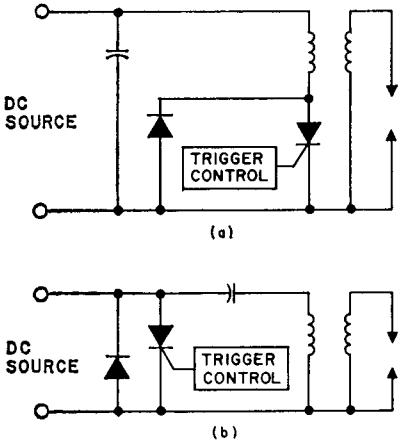



Figure 776. Basic circuit configurations shown in Fig. 775 modified by the addition of a diode in shunt with the SCR switching device.

reverse conduction, and thereby reduces the over-all cost of the system by reducing the reverse blocking requirement of the SCR. The ration of spark duration to charging time decreases with increasing RPM so that in some applications an RPM limit may be reached below the desired maximum because of the charging-time requirements.

## Economic Considerations

The capacitive-discharge system is generally acknowledged to be technically superior to the inductive-discharge system. Its chief drawback has been (and continues to be) economic. Although the transformer may be less expensive than an ignition coil, the capacitor must be of fairly high quality. The SCR and its associated trigger circuit are generally more expensive than comparable transistors and trigger circuits for inductive-discharge systems. Finally, the capacitor charging circuit in an automotive ignition system is a dc-to-dc converter, which represents an additional cost element. Such converter circuits typically require a transformer, two or four diodes for rectification, and one or two transistors. With the increasing use of the ac alternators in modern automobiles, it may eventually be possible to tap off the ac voltage, step it up to the desired voltage by use of a simple transformer, and rectify it. However, it is not clear at this time that this approach is desirable or less costly. Despite these considerations, there are a number of capacitive-discharge systems available on the replacement market.

## Technical Considerations

One significant feature of the capacitive-discharge system is that the input power increases directly with the increased spark-plug power required as engine speed increases. In the inductive-discharge system, on the other hand, the opposite is true, as shown in Fig. 777. The required power is the product of the energy required per ignition pulse and the number of ignition pulses per second. The upper curve of Fig. 777 was determined experimentally for the circuit of Fig. 761 for a battery supply of

14 volts. In the capacitive-discharge system, input power can be made proportional to the required power because the capacitor is charged once per ignition

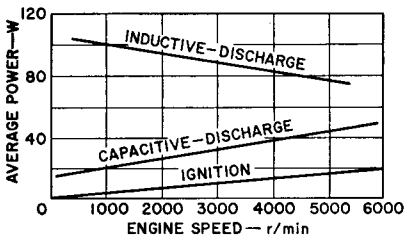



Figure 777. Ignition power requirements.

pulse and holds this charge until needed. In addition, feedback can be used to turn the converter off after the capacitor is charged and thus cut off the input power to the system. The input power can therefore be made proportional to engine speed, and higher efficiency can be achieved at all speeds. The curve shown in Fig. 777 applies for a commercially available capacitive-discharge system. The higher efficiency of this system is apparent.

A second important advantage of the capacitive-discharge system is that faster rise times are more readily obtained because the transformer acts only as a pulse transformer and not as an energy-storage element. Therefore, its high-frequency response characteristic is governed by its leakage inductance, which is much smaller than the primary magnetizing inductance. This advantage is obtained even when a conventional ignition coil is used as the pulse transformer. Secondary-voltage rise times of about 15 to 30 microseconds are readily obtained. As discussed previously, the shorter rise time greatly enhances the ability of this system to fire fouled plugs.

A major operating point that must be considered in the capacitive-discharge system is when to charge the capacitor. In some systems the capacitor is charged soon after discharge, in others, just before discharge. The second method is the better in that it minimizes the losses resulting from leakage, but this advantage is somewhat negated because of the precise timing required to institute the charge just prior to discharge. This requirement can result in complex mechanical or electrical arrangements.

## Component Requirements

In a capacitive-discharge ignition system, the forward blocking voltage of the SCR must be greater than about 400 volts, and its current-handling capability must be about 5 amperes. The SCR firing characteristics as a function of temperature are important, and must be taken into account in the design of the trigger circuit. Because the specifications of the spark coil are usually known, the design of an ignition circuit is usually begun there; capacitor size and voltage are determined from these known quantities. The capacitor chosen determines the spark-gap duration and the charging time required; the reliability of the system is almost directly related to the reliability of this component. No recommendations for this capacitor are made here except that the user is reminded that the high peak current which flows through the SCR also flows through this capacitor, necessitating that the capacitor be as carefully chosen as the SCR.

The limit on charging time is the dv/dt rating of the SCR. By

using the capacitor size and the dv/dt rating of the SCR, the charging current required is determined by applying the formula $L = Cdv/dt$. This formula can also be used to determine the peak current and its duration, and the rate of rise of current. The rate of rise of the current should be checked against the limiting circuit value, the di/dt rating of the SCR.

Table XLVII—SCR Parameter Values of Importance in Ignition Circuits

| | |
|---|---|
| $I_{DROM}$ | when $V_D = V_{Capacitor} +$ 20% at a case temperature $T_C$ of 100°C. |
| $I_{RRDM}$ | when $V_R = 25$ volts when using a diode as in the circuit of Fig. 776 or when $V_R = V_{peak\ reverse}$ due to flywheel effect in the flywheel charged system. In both cases, $T_C$ is 100°C. |
| $v_T$ | when $I = I_{peak}$, the value of $v_T$ is approximately 2.5 to 4 volts depending on current pulse amplitude, repetition rate, and case temperature. |
| $V_{gate}$ | at 12 volts with $R_A = 30$ ohms. $V_{gate}$ and $I_{gate}$ will be maximum or minimum limits depending on trigger-circuit requirements. Higher limits help prevent spurious firing as a result of noise. |

The SCR parameter values that must be specified to assure reliable device operation in ignition circuit applications are shown in Table XLVII.

The ignition coil used in the capacitive-discharge circuit can be a specially wound low-inductance unit or the existing coil. The existing coil has an advantage in that it can be used with a breaker-point distributor to provide the ignition function in the event of an electronic failure. The major disadvantage of the use of the existing coil is that the benefits of ignition pulses with sharp rise times, the type of pulses needed to fire fouled spark plugs, are reduced because of the inductance of the coil. Because of the ease of obtaining a trigger pulse and the circuit simplicity, the SCR capacitor-discharge system is used almost exclusively on small engines.

## Types of Capacitor-Discharge Systems

There are three systems in which capacitor-discharge ignition circuits can be used to good advantage; the flywheel-charged small-engine system, the line-charged ignitor used with gas-operated appliances such as dryers and furnaces, and the inverter-charged system such as that used in automotive and stationary engine systems. All of these systems are operated similarly: energy stored in a capacitor is transferred to a spark gap through a transformer and SCR; the SCR assures a short-duration spark.

The circuit of Fig. 778 is typical of that used in the three systems. The ac potential across transformer $T_1$ is rectified by diode $D_2$, and charges capacitor $C_1$ to the required voltage. Resistor

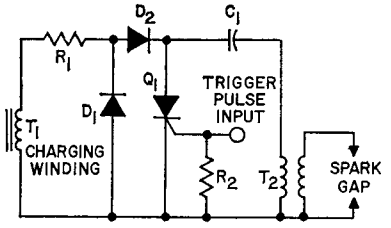



Figure 778. Typical circuit configuration for a capacitor-discharge ignition system.

$R_1$ limits the current and prevents the SCR from firing as a result of the imposition of a dv/dt value in excess of the capability of the SCR. The combination of diodes $D_1$ and $D_2$ prevents the charging winding of transformer $T_1$ from impressing a high reverse voltage across the SCR. Resistor $R_2$ damps variations in the input impedance of the SCR. The SCR is triggered at the appropriate time, and the energy stored in the capacitor is transferred into the primary of $T_2$, thus causing a spark at the spark gap. The voltage required to break down the spark gap is a function of the spacing of the electrodes and pressure in the cylinder in the vicinity of the gap. The spark in the gap lasts until the value of current passing through the SCR is below its holding current. When the SCR stops conducting, $D_1$ and $D_2$ start conducting in the reverse direction and lengthen spark duration. After the SCR turns off, $C_1$ is discharged, and the circuit is ready to repeat the cycle.

**Flywheel-Charged Systems** — Some of the simplest ignition systems are constructed using the flywheel-charging method as this method affords a reliable circuit with a minimum of active components. The system makes use of a rotating magnetic field to charge the capacitor and to trigger the SCR; mechanical position determines timing. The designer has several options in the determination of when the charging of the capacitor takes place in the flywheel system. The most advantageous time occurs just before the capacitor is to be discharged. However, some voltage regulation problems must be considered. Because $V = Nd\phi/dt$ where $d\phi$ and N are constant, the voltage produced across the charging winding varies with RPM. At low flywheel speeds, there may not be enough voltage available to produce the energy required; at high flywheel speeds, it is possible to have too high a voltage and therefore to exceed the voltage breakdown rating of the SCR and cause premature triggering. If the breakdown voltage rating of the capacitor is also exceeded, the capacitor will be damaged. Therefore, some means of accommodating or regulating the voltage must be considered.

The design of the trigger coil is also important. It must be capable of providing voltages and currents high enough to gate the SCR into conduction at all temperatures. In addition, consideration should be given to the fact that the gate pulse should end before the current through the SCR ceases to flow so that the device is not gated during the period of reverse voltage.

As is evident in the above discussion, a major factor in the performance of the flywheel ignition circuit is the design of the magnetic components used for the triggering and charging functions. A typical example of a fly-

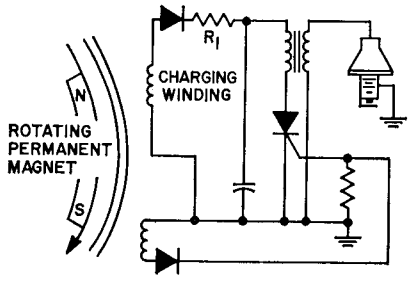



Figure 779. Flywheel-charged igniter circuit

wheel-charged ignition circuit is shown in Fig. 779.

**Line-Charged Igniter** — The line-charged ignition circuit finds greatest use in heating systems and large gas-operated appliances. The line-charged igniter has small regulation or charging problems as it uses the power lines as its charging source. Normally, circuits of this type charge at the line voltage that produces the lowest-cost circuit. The gate trigger pulse is derived by using a diac and an RC phase-shift network such as that shown in Fig. 780.

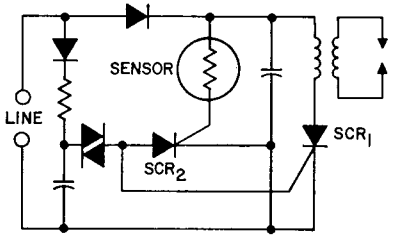



Figure 780. Line-charged ignition circuit.

On each positive half-cycle, the circuit operates once until the flame is burning; the sensor then turns on $SCR_2$ which prevents $SCR_1$ from being triggered and causing further sparks. This arrangement will cause the igniter to operate whenever the flame goes out and thus provides a valuable safety feature. If a bimetallic element were used in series with the line to sense the flame, $SCR_2$ could be eliminated. The use of the bimetallic element would also contribute to a low standby power drain; however, at higher energy levels, the bimetallic strip method of gating is probably only marginally reliable as the gate current produced is not sufficient for the anode-current pulse-amplitude required.

**Inverter-Charged Systems**—A system that eliminates the need for flywheel magnetics is the inverter-charged system. This system is used where a battery is available, such as in an automobile. The design of the inverter is discussed in the section of this Handbook concerning **Power Conversion.**

There are some practical considerations which limit the use of the inverter system. The first limitation is starting under low temperature. At an ambient temperature of −40°C, the available battery voltage in a "12-volt" automotive system (12-volts nominal at 25°C) may be as low as 6 volts dc because of the starter current required at this temperature and the reduced battery capability. In addition, at this temperature, the fuel-air mixture is wet, particularly in a two-cycle engine. For reliable starting, the full spark energy must be available immediately. This means that the inverter must be capable of producing the full energy at low supply voltages. The voltage step-up ratio of the system transformer is constant and therefore cannot be increased as the temperature decreases; such an action would assure sufficient volt-

age at low temperatures, but would subject the capacitor and SCR to voltages in excess of their ratings under normal conditions and after starting. The problem of starting at low temperature may be circumvented by regulating the voltage on the capacitor or by using a transformer with a higher step-up ratio than required and then shutting down or removing the inverter, with its transformer, from the circuit at a time that will prevent any voltages from becoming a problem.

As the maximum RPM of the engine increases, the demands on the inverter also increase; this variation in demand can be alleviated by ballasting. When ballasting of the ignition is accomplished by means of a regulator circuit, external ballasts are not needed. A typical example of an inverter-type ignition system with regulator ballasting is shown in Fig. 781. The trigger circuit shown in the figure is subject to the same variations in potential as the inverter circuit in addition to others arising from the need to gate the SCR with a high-current at low temperature when the available voltage is low. This gating problem can only be solved by a compromise between overdriving of the gate at high temperature and maintaining only an adequate drive at low temperatures; there are many circuits that can be used to achieve this compromise.

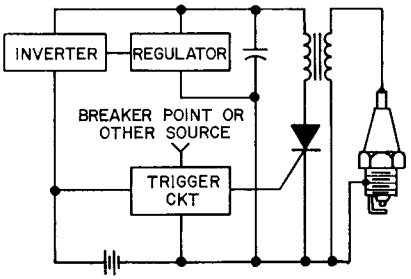


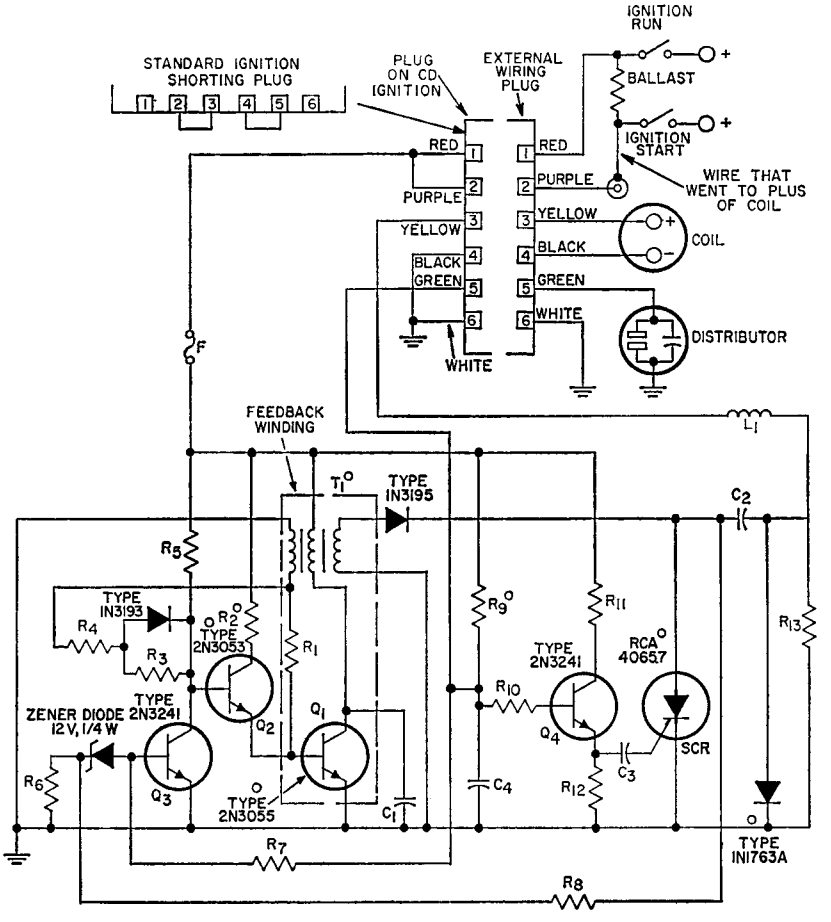

Figure 781. Block diagram of an inverter-charged ignition system.

The inverter must be capable of handling a power level, typically between 20 and 50 watts, representing an energy level of 80 millijoules per pulse, for a 4-cycle, 8-cylinder engine. The inverter circuit should operate at a frequency high enough to make use of smaller transformer-core sizes and yet be able to incorporate low-cost power devices. The RCA line of hometaxial power transistors is generally very reliable and economical in inverter ignition systems.

## Capacitive-Discharge Automotive Ignition System

Fig. 782 shows the circuit diagram for a low-cost transistor/SCR capacitor-discharge ignition system for passenger automobiles. This system offers the advantages of reduced maintenance, smaller current drain on the automobile battery, full output voltage at low battery voltage (down to 4 volts), and a high-voltage output pulse that has a rapid rate of rise. As pointed out previously, this latter factor provides greater assurance of the firing of fouled spark plugs.

The SCR ignition system is essentially a combination of eight basic circuit units, as follows: (1) A single-ended, self-oscillating swinging-choke inverter is used to provide the dc-to-ac inversion and the step-up of the battery voltage. (2) An output circuit that includes an SCR, a storage capacitor, an ignition coil (a standard automotive ignition coil is used), and a com-




$C_1$ = 0.25 μF, 200V
$C_2$ = 1 μF, 400V
$C_3$ = 1 μF, 25V
$C_4$ = 0.25 μF, 25V
F = 5A
$L_1$ = 10 μH, 100T of No. 28 wire are wound on a 2W resistor (100 ohms or more)
$R_1$ = 1000 ohms
$R_2$ = 50 ohms, 5W
$R_3$ = 22000 ohms
$R_4$ = 1000 ohms
$R_5$ = 10000 ohms
$R_6$ = 15000 ohms
$R_7$ = 8200 ohms
$R_8$ = 0.39 megohm
$R_9$ = 220 ohms, 1W
$R_{10}$ = 1000 ohms
$R_{11}$ = 68 ohms
$R_{12}$ = 4700 ohms
$R_{13}$ = 27000 ohms

$T_1$ = Transformer, wound as follows: A ½ in. bobbin and o El stack of grain oriented silicon steel are used; first, 150 turns of No. 28 wire are wound and labeled start 1 and finish 1 on the winding; second, 50 turns of No. 24 and No. 30 wires are wound bifilar and labeled start 2 and finish 2; third, 150 turns of No. 28 wire are wound and labeled start 3 and finish 3. All windings are wound in the same direction. A total air gap of 70 mil (35 mil spacer) is used. Connections are made as shown in Fig. 783.

All the resistors are ½W unless otherwise indicated

° These components are subject to excessive temperature rise unless provision is made to transfer the heat to the ambient air by means of an appropriate heat sink.

Figure 782. An SCR capacitor-discharge automobile ignition circuit.

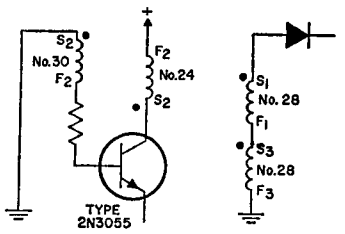



Figure 783. Details of inverter transformer ($T_1$) shown in Fig. 782.

mutating diode develops the fast-rising high-voltage pulse for the spark plugs. (3) The commutating diode and a single rectifying diode form a capacitor-charging circuit to transfer energy from the inverter to the output-circuit storage capacitor. (4) A regulator stage controls the frequency of the inverter stage to provide efficient regulation of the voltage across the storage capacitor. (5) A protection circuit (limiting inductance and resistance) prevents damage to the system by transients that may be developed in the case of an open or shorted ignition coil or because of high-voltage arcing to either primary terminal of the ignition coil. (6) A shutdown circuit holds the inverter inoperative when the ignition breaker points are open. (7) A trigger circuit suppresses the normal bounce of the breaker points and also prevents SCR triggering by the residual voltage across the closed points. (8) A method of SCR commutation is used that involves the interplay of several parts of the overall system.

**Inverter, Regulator, and Capacitor-Charging Circuit** — The inverter uses a 2N3055 transistor ($Q_1$) in a single-ended output stage, a 2N3053 transistor ($Q_2$) in an emitter-follower driver stage, and a 2N3241 transistor ($Q_3$) in a control stage that is part of the shutdown circuit which holds the inverter inoperative when the ignition breaker points are open. Regenerative feedback is coupled from the feedback winding of the inverter output transformer $T_1$ back to the bases of the driver and output transistors. The high gain provided by the combination of transistors $Q_1$ and $Q_2$ assures oscillation and low drive-power requirements for the inverter. The starting resistor $R_5$ provides a forward bias that drives transistors $Q_1$ and $Q_2$ into conduction to initiate oscillation in the inverter. The regenerative action of the circuit very quickly drives the output transistor $Q_1$ into conduction, and essentially the full battery voltage is then applied across the primary of the inverter output transformer $T_1$. The resultant current increase in the transformer primary winding induces a voltage across the feedback winding that supplies sufficient current through resistors $R_1$ and $R_4$ and diode $D_1$ to maintain the output transistor $Q_1$ in saturation. During this part of the operating cycle (i.e., during the conduction of transistor $Q_1$, the voltage across the secondary winding of transformer $T_1$ reverse-biases the rectifying diode $D_2$ in the capacitor-charging circuit, and no energy is transferred to the output circuit of the ignition system.

With transistor $Q_1$ operating with fixed base current (in saturation), its collector current rises to a value beyond which it cannot increase. As a result,

the feedback voltage is decreased, and no longer maintains base drive to transistor $Q_1$, and the transistor starts to turn off. The regenerative action of the inverter circuit causes a rapid reversal of the base drive for transistors $Q_1$ and $Q_2$. These transistors, therefore, are quickly cut off, and a "flyback" voltage pulse is generated at the collector of the output transistor $Q_1$. Diode $D_3$ blocks the reverse voltage and limits the reverse base drive. The reverse-bias current that turns off transistors $Q_1$ and $Q_2$ is applied through resistors $R_3$ and $R_1$, respectively. The flyback-pulse voltage is stepped up across the secondary of transformer $T_1$. The polarity of this pulse, however, is such that the rectifying diode $D_2$ becomes forward-biased. As a result, the energy previously stored in the primary winding of transformer $T_1$ is transferred through the secondary winding, rectifying diode $D_2$, and commutating diode $D_3$ to charge the output-circuit storage capacitor $C_2$. The capacitor $C_1$ connected across transistor $Q_1$ reduces the amplitude of the leakage-inductance pulse and restricts the rate of rise of the collector voltage of transistor $Q_1$. The charging current for capacitor $C_2$ is shunted around the ignition coil by the commutating diode $D_3$ so that no energy is transferred into the ignition coil and from there to the spark plugs.

When the collector voltage of transistor $Q_1$ decreases to a value less than the battery voltage, it again begins to conduct, and the cycle is repeated to charge the storage capacitor $C_2$ to a higher voltage. Until the voltage across the storage capacitor rises above a predetermined value, the voltage applied to the zener diode $D_4$ from the voltage divider formed by resistors $R_6$ and $R_8$ is insufficient to cause the zener diode to conduct. If the ignition breaker points are closed during this time, resistor $R_7$ is returned to ground and transistor $Q_3$ cannot conduct. When the capacitor voltage rises to a level high enough to cause zener diode $D_4$ to conduct, transistor $Q_3$ turns on and shunts base drive current from transistor $Q_2$. This effect reduces the base drive of transistor $Q_1$ and causes this transistor to pull out of saturation at a lower collector-current level which, in turn, increases the frequency of oscillation. The cutback in peak primary current reduces the charging rate of the storage capacitor $C_2$ to the level required to replenish circuit losses and prevents further rise in the output voltage.

Transistor $Q_3$ also holds the inverter inoperative when the ignition breaker points are open. When these points open, the current fed from the voltage at the breaker points through resistor $R_7$ causes transistor $Q_3$ to conduct heavily. This effect shorts the base of transistor $Q_2$ and stops the oscillation.

Fig. 784 shows that the collector voltage of transistor $Q_1$ swings alternately between the saturation level and the peaks of flyback pulses of increasing amplitude. The change in frequency that results from regulator action is apparent in the voltage waveform. The collector voltage then decreases to the supply voltage when the ignition breaker points open and shut down the

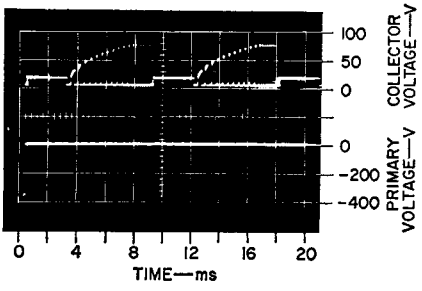



Figure 784. Collector voltage of the inverter output transistor $Q_1$ (top) and ignition-coil primary voltage (bottom) as functions of time (2000 rpm; $V_{CC}$ = 12V.)

inverter. Fig. 785 shows an expanded view of the turn-off and flyback characteristics of transistor $Q_1$ at a point when the storage capacitor is being charged.

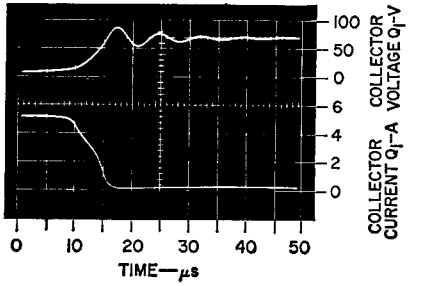



Figure 785. Expanded collector voltage (top) and current (bottom) of the inverter output transistor ($Q_1$) as functions of time during turn off when the storage capacitor ($C_2$) is being charged ($V_{CC}$ = 12V).

**Output Circuit**—When a high-voltage pulse is required, the RCA-40657 SCR in the output circuit is gated on. As a result, the anode voltage of the SCR decreases to approximately zero, and the voltage across the charged storage capacitor is applied to the primary of the ignition coil. (The value of the inductance $L_1$ is negligible in comparison to the inductance of the ignition coil and is not considered in this analysis.) The ungrounded (+) side of the ignition-coil primary (terminal 3 on the connecting plug) is driven negative with respect to the capacitor potential. Diode $D_3$, in parallel with the coil, is reverse-biased at this time. The discharge of the capacitor into the primary of the ignition coil generates a high-voltage pulse across the secondary.

The capacitor discharges into the primary inductance of the ignition coil and builds up the primary current in the coil. When the voltage across the capacitor (and coil primary) decreases to zero and starts to reverse, the commutating diode $D_3$ becomes forward-biased and begins to conduct. The current through the primary of the ignition coil is at a peak at the time the diode begins to conduct. The current then suddenly switches out of the SCR and into the diode. The primary-coil voltage remains clamped at zero, and the primary current decays at a rate determined by the L/R ratio of the coil. Because of the clamping action of the commutating diode $D_3$, the duration of the spark in the spark plug is lengthened.

When the SCR is on, it effectively places a short across the secondary of the inverter transformer. However, the inverter is off when the SCR is on (because of the shut-down circuitry); the inverter, therefore, does not operate into the short.

Figs. 786 and 787 show the SCR voltage and current as a function of time. The starting point of the waveform shown in Fig. 786 occurs at the instant the ignition points open. The

anode voltage of the SCR decreases to zero and the anode current builds up to the peak value in a quarter cycle. The current is then switched out of the SCR, and SCR current decreases suddenly almost to zero.

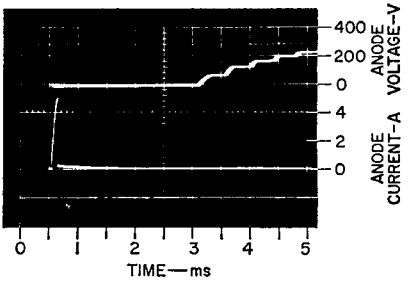



Figure 786. SCR voltage (top) and current (bottom) as a function of time (2000 rpm; Vcc = 12V).

The small residual current is a result of the energy stored in the inverter transformer during the period that the inverter is inoperative. This stored energy causes a current to circulate from the secondary of the transformer through the SCR. When the ignition points close and the capacitor recharges, the SCR blocks the voltage on the capacitor.

The starting point for the waveform shown in Fig. 787 occurs at the instant that the ignition points close. The period between the instant at which the points close and that at which the voltage first begins to rise is the time during which the collector current of transistor $Q_1$ builds up to the switching level. The significance of this time is explained subsequently during the discussion on commutation of the SCR.

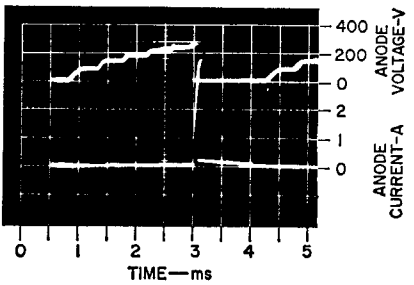



Figure 787. SCR voltage (top) and current (bottom) as a function of time (oscilloscope sweep triggered at instant ignition breaker points are closing; 4000 rpm; Vcc = 12V).

Fig. 788 shows the waveforms for voltage and current in the primary of the ignition coil that result when a 7-millihenry inductor is used to simulate the primary of the ignition coil. The primary voltage increases rapidly to a peak negative value, then decays sinusoidally to zero as the current builds to a peak. The primary voltage is then clamped at zero, and the current decays exponentially.

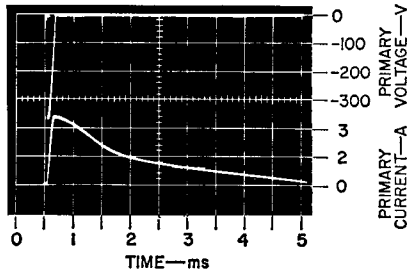



Figure 788. Primary voltage (top) and current (bottom) as a function of time (1-mH coil used in place of standard ignition coil; 2000 rpm; Vcc = 12V).

It should be noted that an actual ignition coil, operated with the secondary open, will reflect a tuned circuit into the primary. This operation causes a ringing on top of the waveforms, as shown in Fig. 788. The anode voltage of the SCR may actually reverse for a short time because of this ringing. The SCR essentially blocks this reverse volt-

age except for a small current that flows because of the presence of positive gate signal. As a result, some instantaneous dissipation occurs in the reverse-blocking function of the SCR. The SCR can safely withstand this dissipation for the short period of time required. The gate signal is kept positive through the ringing cycle so that the SCR continues to conduct when the anode voltage rings back positive. This ringing does not occur when the secondary voltage fires a plug because the ionized plug shorts the secondary winding.

**Protection Circuit**—Inductance $L_1$ is used to protect the system against a shorted primary in the ignition coil. The limiting inductance controls the rate of change of current (di/dt) and peak current that occurs when the SCR is turned on with a short across the primary of the ignition coil. Resistance $R_{13}$ is used to assure that the voltage across the primary of the ignition coil is not negative when the coil is open. If this voltage were not clamped, the regulator would not operate properly, and the peak collector voltage of transistor $Q_1$ would exceed the limits specified for the device.

**Trigger Circuit**—The triggering circuit performs the following functions: (1) triggers and holds the SCR on when the ignition points open (at battery voltage down to 4 volts), (2) applies a signal back into the inverter shutdown circuitry when the ignition points are open, (3) suppresses the inverter signal that rides on the power supply so that it does not trigger the SCR, (4) prevents the residual voltage across the closed points from triggering the SCR, (5) prevents normal point bounce that occurs when the points close, and (6) maintains proper operation whether or not the capacitor is present across the breaker points. The 2N3241 transistor $Q_4$ is used to perform these functions.

The trigger current for the gate of the SCR is initiated when the base voltage of transistor $Q_4$ reaches approximately 0.6 volt above the emitter voltage. The trigger current flows from the supply through resistor $R_{11}$, transistor $Q_4$, and capacitor $C_3$ to the gate of the SCR.

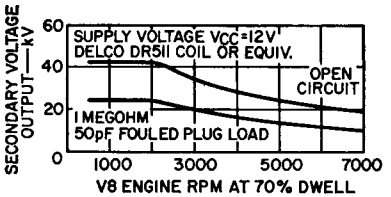
When the ignition points open, capacitor $C_4$ (and the capacitor across the points) charges because of the current through resistor $R_3$. If the points are open long enough (without bouncing), the voltage across capacitor $C_4$ becomes high enough to turn on transistor $Q_4$. The voltage required to turn on transistor $Q_4$ is the sum of the gate-cathode voltage of transistor $Q_5$, the voltage across capacitor $C_3$, the emitter-base voltage of transistor $Q_4$, and the voltage drop across resistor $R_{10}$. (Resistor $R_{10}$ ensures that the voltage across the open ignition points rises to a value high enough to supply sufficient current through resistor $R_7$ to shut down the inverter.) At normal engine speeds, the average voltage level across capacitor $C_3$ keeps both the gate-cathode junction of the SCR and the emitter-base junction of the transistor reverse-biased until capacitor $C_4$ charges high enough to turn on transistor $Q_4$. Because transistor $Q_4$ is off and the gate of the SCR is reverse-biased

when the points are closed, the desired suppression of inverter signal and residual point voltage is achieved.

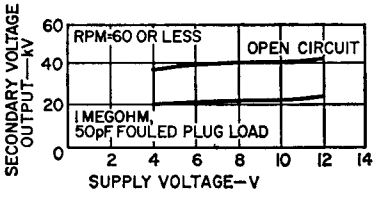
If the points bounce during normal operation, they discharge $C_4$ (and the distributor capacitor) almost instantly each time they close. Thus, each time they bounce open, these capacitors much recharge from zero toward the triggering level. With normal bouncing, the points do not stay open long enough for the triggering level to be reached, so the SCR is not triggered. If severe bouncing occurs at very high speeds, the points can stay open long enough to cause triggering of the SCR.

Better filtering is achieved when the automobile distributor capacitor is retained, but satisfactory operation is achieved without this capacitor. With the capacitor left in the distributor, it is possible to switch back to standard ignition by switching the plug shown in Fig. 782.

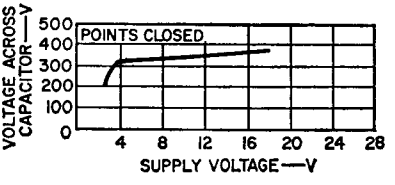
**Commutating the SCR**—All the parts of the system work together in such a manner as to cause the SCR current to go to zero for a sufficient length of time to cause commutation (turn off). As explained earlier, the current through the primary of the ignition coil is switched from the SCR and into the commutating diode when the primary voltage decreases to zero. From that time on, the diode keeps the coil current clamped out of the SCR. The SCR then conducts only the small current that results from the energy stored in the inverter transformer. When the points close again, the inverter restarts and, as explained previously, the rectifying diode $D_2$ is reverse-biased during the time the collector current of transistor $Q_1$ builds back up to the switching level. No current then flows in the SCR, and the SCR is allowed to turn off. Fig. 786 shows that the current is zero for about 300 microseconds before anode voltage is reapplied at a rate of 0.5 volt per microsecond. A worst-case SCR



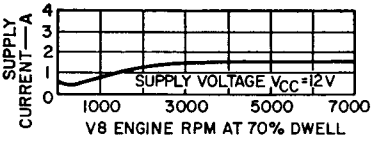

(a) Output voltage as a function of engine rpm at 12 volts for both an open secondary and a fouled-plug load.




(b) Output voltage as a function of battery voltage at cranking speeds for both an open secondary and a fouled-plug load.




(c) Regulation curve showing peak capacitor (and SCR) voltage as a function of battery voltage.




(d) Battery drain as a function of engine rpm at a battery voltage of 12 volts.

Figure 789. Performance of the capacitor-discharge ignition circuit.

commutates in less than 100 microseconds at a temperature of 100°C under these operating conditions.

**Performance**—Fig. 789 (on preceding page) shows several performance curves for the SCR capacitor-discharge ignition system. Fig. 790 shows the open-circuit output voltage as a function of time, and Fig. 791 shows the output voltage when the load on

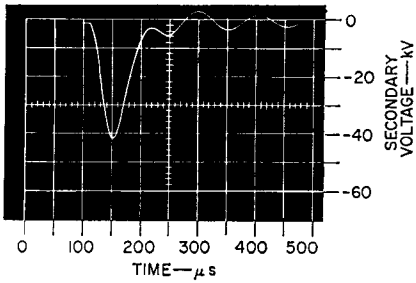



Figure 790. Open-circuit output voltage (standard ignition coil, Delco D511 or equivalent; 2000 rpm; Vcc = 12V).

the secondary consists of a 1-megohm resistor in parallel with a 50-picofarad capacitor.

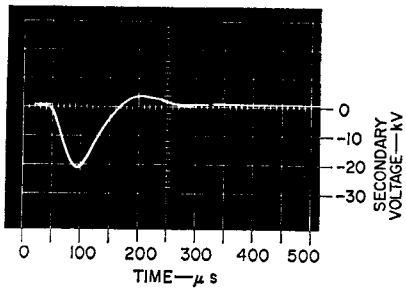



Figure 791. Output voltage with fouled spark plugs (standard ignition coil, Delco D511 or equivalent; 50-pF load in parallel with 1-megohm resistance; 2000 rpm; Vcc = 12 V).

Figs. 792 and 793 show the secondary voltage under sparking conditions. Arc duration, as

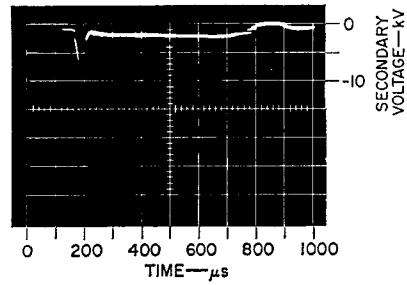



Figure 792. Output voltage showing duration of spark arc (standard ignition coil, Delco D511 or equivalent; 2000 rpm; Vcc = 12 V).

shown in Fig. 792, is 300 microseconds (single polarity) under wide-gap conditions. The narrower-gap conditions shown in Fig. 793 result in an arc duration of 400 microseconds.

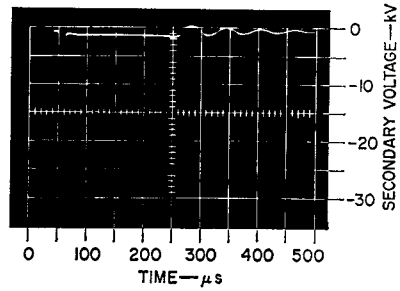



Figure 793. Output voltage with spark gap shortened (standard ignition coil, Delco D511 or equivalent; 2000 rpm; Vcc = 12V).

**Mounting Considerations**—The SCR ignition circuit must be protected from moisture. Heat-generating components, however, should not be enclosed in non-circulating atmosphere. Still air has a thermal-resistance 12,000 times that of copper, and generation of heat in this high thermal resistance could cause the inside ambient temperature to rise above the specified limits. All components subject to high temperature rise (marked in Fig. 782 with a small circle) should be

thermally connected to a low-thermal-resistance path to the outside environment. For example, the SCR may be mounted to an aluminum plate on a mica insulating washer. This plate should then be fastened to the inside of the chassis wall that provides the thermal path to the outside environment. The resistors and diodes which require a heat sink should be attached to the chassis with a thermally conductive epoxy.

# Index


| | Page |
|---|---|
| AC line-voltage controls | 339 |
| AC voltage regulator | 400 |
| Admittance parameters | 151 |
| Aircraft communications, military | 489 |
| Aircraft radio, commercial | 505 |
| Design considerations | 509 |
| Design problems | 507 |
| Desirable features | 506 |
| Output amplifier | 513 |
| Performance and adjustment | 512 |
| Power and modulation | 508 |
| Air-rescue beacons | 497 |
| Air-traffic-control systems | 564 |
| Alpha | 88 |
| Amperes squared-seconds rating | 69 |
| Amplifier chains, microwave | 539 |
| Amplifier circuits, microwave | 536 |
| Amplifiers, linear | 480 |
| Amplifiers, low-noise | 525 |
| Amplifiers, wide-band | 517, 525 |
| Anode | 9, 195 |
| Audio-Amplifier circuit configurations: | |
| Class A transformer-coupled | 572 |
| Class AB push-pull transformer-coupled | 572 |
| Class AB series-output | 574 |
| Quasi-complementary-symmetry | 576 |
| True-complementary-symmetry | 575 |
| Audio-amplifier design: | |
| Bridge-circuit approach | 595 |
| Hybrid-circuit approach | 607 |
| Universal-amplifier approach | 595 |
| Audio amplifier, direct-coupled, low-frequency safe-area analysis of | 139 |
| Audio power amplifiers | 565 |
| Basic circuit configurations | 572 |
| Basic power-dissipation relationships | 585 |
| Classes of operation | 565 |
| Design approach | 595 |
| Drive requirements | 566 |
| Effect of excessive drive | 592 |
| Effect of large phase shifts | 591 |
| Effect of operating conditions on circuit design | 569 |
| Power output in class B types | 579 |
| Rating methods | 579 |
| Short-circuit protection | 593 |
| Thermal-stability requirements | 590 |
| Automotive ignition systems, basic considerations | 667 |
| Condition of spark plugs | 667 |
| Cylinder pressure | 667 |
| Energy storage | 669 |
| Plug polarity | 668 |
| Spark-plug voltage waveshape | 668 |
| Avalanche breakdown | 114, 116, 209 |
| Average current | 66 |
| **B**allast circuits for mercury-arc lamps | 409 |
| Ballasting circuits, solid-state | 413 |
| Ballasting methods (for mercury-arc lamps): | |
| Conventional ballasting | 412 |
| Solid-State ballasting | 413 |
| Base | 11 |
| Base-conductivity modulation | 95 |
| Base-spreading resistance | 159 |
| Base-to-emitter voltage | 93 |
| Base widening | 97 |
| Basis for device ratings | 16 |
| Beta | 88 |
| Bilateral switch | 13 |
| Bridge-circuit audio-amplifier design approach | 604 |
| Bridge-circuit inverter | 303 |
| Broadband circuit design | 482 |
| **C**apacitor-input filter | 254 |
| Capacitive-discharge ignition systems | 677 |
| Basic circuit operation | 678 |
| Component requirements | 680 |
| Economic considerations | 679 |
| Flywheel charged systems | 682 |
| Inverter-charged systems | 683 |
| Technical considerations | 679 |
| Types | 681 |
| Capacitive-load circuits | 255 |
| Carrier concentration | 61 |
| Carrier-flow analysis (of silicon rectifiers) | 58 |
| Equilibrium condition | 59 |
| Forward-bias condition | 60 |
| Cathode | 9, 195 |
| Case-to-ambient thermal capacitance | 45 |
| Case-to-ambient thermal resistance | 45 |
| Charge-carrier flow (in transistors) | 81 |
| Choke-input filter | 254 |
| Circuit turn-off time | 234 |
| Class A operation | 566 |
| Class B operation | 566 |
| Class C oscillators (ultrasonic) | 617 |
| Classes of operation (for audio amplifiers) | 565 |
| Collector | 11 |
| Collector-delay time constant | 92 |
| Collector depletion-layer transit time | 92 |
| Collector punch-through voltage | 114 |
| Commercial aircraft radio | 505 |
| Common-base avalanche breakdown | 114 |
| Common-base current gain | 89 |
| Common-base equivalent circuit | 163 |
| Common-collector equivalent circuit | 116 |
| Common-emitter avalanche breakdown | 116 |
| Common-emitter current gain | 89 |
| Common-emitter equivalent circuit | 154 |
| Community-antenna television | 513 |
| Amplifier requirements | 515 |
| Low-noise amplifiers | 525 |
| System operation | 514 |
| Transistor considerations | 521 |
| Wide-band amplifier | 517, 525 |
| Commutating dv/dt capability | 238 |
| Compensating diode | 11 |
| Conduction angle | 339 |
| Construction (of Thyristors) | 206 |
| Controls, Thyristor ac line-voltage | 339 |
| Controls, thyristor turn-on | 339 |
| Converter | 300 |
| Critical rate of rise of commutation voltage | 238 |
| Critical rate of rise of off-state voltage | 209, 220 |
| Critical rate of rise of on-state current | 210 |
| Crowbar trigger circuit | 279 |
| Current crowding | 96 |
| Current flow | 7 |
| Current-gain parameters | 88 |
| Current gain, physical basis, for | 88 |
| Current-limited regulated power supply | 270 |
| Current-limiting techniques | 266 |
| Current ratings | 20, 122 |
| Current-regulating power supplies | 263 |
| Current-voltage relationships (for transistors) | 83 |
| Cyclic thermal stresses, effect of | 55 |
| **D**amper diode | 631 |
| DC power supplies | 249 |
| Dc supply, line-voltage-regulated | 406 |
| Deflection circuit for color receiver | 640 |
| Deflection circuits for monochrome receiver | 636 |
| Driver and output stages | 637 |
| Horiztontal oscillator | 639 |
| Horizontal-phasing (afc) circuit | 639 |
| Delay time | 176, 231 |
| Depletion layer | 6 |
| Depletion region | 59 |
| Design, processing, and packaging (of power transistors | 99 |
| Geometries | 108 |
| Packaging | 105 |
| Structures | 99 |
| Device ratings, explanation of | 16 |
| Diacs | 13, 352 |
| di/dt capability | 220 |
| Diffused transistors | 102 |
| Diffusion current | 6 |

| | Page |
|---|---|
| Diode | 9 |
| Distributed base resistance | 159 |
| Doping level | 61 |
| Double-time-constant triggering network | 348 |
| Drive requirements (for audio amplifiers) | 567 |
| dv/dt capability | 209 |
| EIA music-power rating | 580 |
| Electronic Industries Association | 579 |
| Emitter | 11 |
| Emitter efficiency | 89 |
| Energy barrier | 7 |
| Epitaxial transistors | 104 |
| Equivalent-model analyses (of power transistors) | 150 |
| Large-signal analysis (linear service) | 170 |
| Large-signal analysis (switching service) | 175 |
| Small-signal analysis (linear service) | 150 |
| Equivalent-model analysis (of thyristors) | 195 |
| Equivalent transistor input circuits | 161 |
| Excessive drive (in audio amplifiers), effects of | 592 |
| Explanation of device ratings | 16 |
| Fabrication (of power hybrid circuits) | 243 |
| Fall time | 177 |
| Fast-recovery rectifiers | 70 |
| Feedback capacitance, collector-to-base | 160 |
| Feedback resistance, collector-to-base | 159 |
| Fermi energy level | 199 |
| Ferrite cores | 484 |
| Filtering | 253 |
| Firing angle | 339 |
| Flanged packages | 27 |
| Flashover | 381 |
| Flash at turn off | 387 |
| Flywheel-charged ignition system | 682 |
| Foldback current limiting | 276 |
| Foldback current-limiting circuit | 279 |
| Foldback-limited regulated power supplies | 279 |
| Forbidden-energy region | 199 |
| Forward-bias capacitance-discharge test | 131 |
| Forward-bias $I_s/_b$ test | 129 |
| Forward-bias safe-area ratings | 134 |
| Forward-bias second breakdown | 126 |
| Forward breakover voltage | 205 |
| Forward current | 66 |
| Forward-current ratings | 20 |
| Forward-transmission voltage | 152 |
| Forward voltage drop | 63 |
| Frequency considerations for power transistors | 164 |
| Frequency doubler, microwave | 557 |
| Frequency multipliers | 550 |
| Basic circuits | 553 |
| Circuit design | 554 |
| Operation | 552 |
| Practical circuits | 556 |
| Stability and biasing | 555 |
| Transistor considerations | 550 |
| Frequency tripler, microwave | 558, 560 |
| Full-wave rectifier circuits | 250 |
| Single-phase types | 250 |
| Three-phase types | 251 |
| Gas-discharge lighting systems | 409 |
| Gate | 195 |
| Gate characteristics (of thyristors) | 223 |
| Effect of gate signal on breakover voltage | 224 |
| Pulse triggering | 225 |
| Trigger-circuit requirements | 227 |
| Trigger level | 225 |
| Generators, ultrasonic | 615 |
| Geometries, transistor | 108 |
| Half-wave rectifier circuits | 250 |
| Single-phase type | 250 |
| Three-phase Y type | 251 |
| Heat removal | 47 |
| Heat sink | 46 |
| Effect on basic thermal circuit | 46 |
| Insulators | 51 |
| Performance | 49 |
| Requirements | 48 |
| Types | 49 |
| Heat sink, use of | 46 |
| Heater regulation | 403 |
| Heating controls | 366 |
| Comparison of | 372 |
| General design considerations | 367 |
| Phase-control type | 372 |
| Zero-voltage switched types | 373 |
| High-frequency power transistors | 430 |
| Case-temperature effects | 435 |
| Characterization of large-signal types | 441 |
| DC safe area | 432 |
| Emitter ballasting | 433 |
| Multiple connection of | 443 |
| Packages | 436 |
| Physical-design parameters | 430 |
| Reliability considerations | 436 |
| RF operation | 434 |
| Special ratings concept | 431 |
| High-power plastic packages | 32 |
| Cleaning | 39 |
| Lead-forming techniques | 34 |
| Mounting | 35 |
| High-voltage power (for television receiver) | 635 |
| Holding current | 218 |
| Hometaxial-base transistors | 101 |
| Horizontal deflection system | 628 |
| Basic analysis of | 629 |
| SCR type | 643 |
| Transistor type | 631 |
| Horizontal phasing | 639 |
| Horizontal oscillator | 639 |
| h-parameter equations | 151 |
| Hybrid circuits | 14, 243 |
| Hybrid-circuit audio-amplifier design approach | 607 |
| Bridge circuit | 610 |
| Transformer-coupled circuit | 609 |
| Typical connections | 608 |
| Hybrid-circuit linear power amplifier | 246 |
| Hybrid-circuit series voltage regulator | 281 |
| Hybrid combiner/dividers | 484 |
| Hybrid-parameter relationships | 155 |
| Ignition systems | 667 |
| Basic considerations | 667 |
| Capactive-discharge type | 677 |
| Inductive-discharge type | 669 |
| IHF dynamic output rating | 579 |
| Impedance-admittance chart | 450 |
| Impurities | 5 |
| Induction-motor reversing controls | 398 |
| Inductive-discharge automotive ignition system | 669 |
| Active-mode-switched type | 673 |
| Circuits | 673 |
| Ignition-coil and transistor characteristics | 670 |
| Limitations | 670 |
| Saturation-mode-switched type | 673 |
| Inductive-load switching, analysis of | 186 |
| Inrush current | 382 |
| Intermodulation distortion | 474 |
| Institute of High Fidelity | 579 |
| Integrated-circuit zero-voltage switch | 358 |
| Circuit operation | 359 |
| Effect on thyristor load characteristics | 361 |
| Fail-safe feature | 361 |
| Half-cycling and hysteresis characteristics | 363 |
| Introduction | 1 |
| Inverter | 300 |
| Inverter-charged ignition system | 683 |
| Inverter circuits (ultrasonic) | 616 |
| Inverter initial-turn-on safe-area analysis | 137 |
| Inverter turn-off safe-area analysis | 143 |
| Isolated trigger circuits | 355 |
| Isolated heat controls | 376 |

| | Page |
|---|---|
| Junctions | 3, 6 |
| Junction-temperature ratings | 21, 212 |
| Junction-to-case thermal impedance | 42 |
| Lamp dimmers | 383 |
| Double-time-constant type | 384 |
| Flash at turn off | 387 |
| Hysteresis effect | 385 |
| Initial brightness | 384 |
| Single-time-constant type | 384 |
| Trigger-device characteristics | 386 |
| Zero-voltage-switched type | 388 |
| Large-signal transistor analysis: | |
| Linear service | 170 |
| Switching service | 175 |
| Large-signal transistor characteristics | 173 |
| Large-signal transistor equivalent circuits | 171 |
| Latching current | 218 |
| Lighting controls, incandescent | 381 |
| Lighting systems, relative merits of | 409 |
| Linear amplifiers, circuit design | 486 |
| Linear amplifiers, typical | 480 |
| Linear voltage regulator | 261 |
| Line-charged igniter | 683 |
| Line-voltage controls | 339 |
| Line-voltage-regulated dc supply | 406 |
| Load-line analysis | 183 |
| Load regulation | 265 |
| Magnetostriction | 613 |
| Marine radio | 499 |
| DC operating voltages | 500 |
| Instabilities in | 501 |
| Matching networks | 501 |
| Package considerations | 500 |
| Practical circuits | 503 |
| Reliability | 502 |
| Transistor requirements | 500 |
| Matching networks | 447 |
| Design objectives | 447 |
| Input-circuit matching | 441 |
| Network design | 448 |
| Output-circuit matching | 448 |
| Materials, Junctions, and Devices | 3 |
| Mercury-arc lamps | 409 |
| Advantages of | 410 |
| Characteristics of | 411 |
| Conventional ballasting | 412 |
| Peak starting current | 422 |
| Solid-state ballasting | 413 |
| Microwave amplifier circuits | 537 |
| Microwave oscillator circuits | 540 |
| Miscrowave power amplifiers | 529 |
| Chains of | 539 |
| Circuits | 536 |
| Design techniques | 531 |
| Transistor considerations | 529 |
| Microwave power oscillators | 529 |
| Circuits | 540 |
| Design techniques | 531 |
| Transistor considerations | 529 |
| Microwave systems | 562 |
| Mobile radio | 499 |
| DC operating voltages | 500 |
| Instabilities in | 501 |
| Package considerations | 500 |
| Practical circuits | 503 |
| Matching networks | 501 |
| Reliability | 502 |
| Transistor requirements | 500 |
| Molded-plastic packages | 32 |
| Cleaning | 39 |
| Lead-forming techniques | 34 |
| Mounting | 35 |
| Motor-control circuit, half-wave | 355 |
| Motor controls | 393 |
| For universal motors | 393 |
| Full-wave types | 397 |
| Inductive-motor reversing types | 399 |
| Neon bulbs | 351 |
| Network design | 448 |
| Network properties | 164 |
| Network terminal properties | 155 |
| Nonerepetitive peak off-state voltage | 211 |
| Nonrepetitive peak reverse voltage | 212 |
| N-P-N and p-n-p structures | 8 |
| Off-state voltage (of thyristors) | 208 |
| On-off heater control | 373 |
| On-off lamp controls, photocell-operated | 390 |
| On-state current ratings | 214 |
| On-state voltage | 205 |
| Operating conditions for hybrid circuits | 247 |
| Oscillator circuits, microwave | 540 |
| Oscillator, quadrupler, microwave | 561 |
| Oscillator, low-power | 498 |
| Packages, hermetically sealed | 23 |
| Press-fit types | 30 |
| Stud types | 28 |
| Types with mounting flanges | 27 |
| Packages, molded-plastic | 32 |
| Cleaning | 39 |
| Lead-forming techniques | 34 |
| Mounting | 35 |
| Packaging, handling, and mounting | 22 |
| General considerations | 22 |
| Hermetically sealed packages | 23 |
| Molded-plastic packages | 32 |
| Packaging, transistor | 110 |
| Lead attachment | 110 |
| Metal-ohmic contacts | 110 |
| Pellet attachment | 110 |
| Pass element | 261 |
| Peak reverse voltage | 65 |
| Phase control (of thyristors) | 339 |
| Phase-controlled regulated power supply | 297 |
| Phase shifts (in audio amplifiers), effect of | 591 |
| Photocell trigger circuits | 357 |
| Piezoelectric effect | 614 |
| Pincushion distortion | 652 |
| Potential-energy analysis (of thyristors) | 199 |
| Basic energy states | 199 |
| Forward-blocking state | 200 |
| Forward-conducting state | 201 |
| Turn off | 203 |
| Potential-hill analysis (of silicon rectifiers) | 60 |
| Equillibrium condition | 60 |
| Forward-bias conditions | 61 |
| Reverse-bias conditions | 61 |
| Power amplifiers: | |
| Audio | 565 |
| Microwave | 529 |
| RF | 430 |
| Ultrasonic | 621 |
| Power conversion | 300 |
| Power dissipation | 182 |
| Power-dissipation ratings | 123 |
| Repetitive-pulse operation | 125 |
| Single-pulse operation | 124 |
| Steady-state operation | 124 |
| Power hybrid circuits | 14, 243 |
| Fabrication | 243 |
| Hybrid-circuit linear power amplifier | 246 |
| Operating conditions | 247 |
| Power output in class B audio amplifiers | 579 |
| Effect of nonregulated supply | 586 |
| Effect of power-supply regulation | 581 |
| Effect of series resistance | 589 |
| Power-dissipation relationships | 585 |
| Rating methods | 579 |
| Ratio of music power to continuous power | 588 |
| Power sources, ultrasonic | 612 |
| Power supplies: | |
| Current-regulating types | 263 |
| Voltage-regulating, current-limiting types | 263 |
| Voltage-regulating, current, regulating types | 264 |
| Voltage-regulating type | 262 |
| Power-transistor analysis | 150 |
| Power transistors | 11, 80 |
| Power transistors, basic design considerations | 80 |
| Design, processing, and packaging | 99 |

| | Page |
|---|---|
| General physical theory | 80 |
| Special physical considerations | 95 |
| Power transistors, high-frequency | 430 |
| Power transistors in Switching service | 175 |
| Inductive-load switching | 186 |
| Load-line anlysis | 183 |
| Switching speed | 175 |
| Switching-time reduction techniques | 180 |
| Switching times | 180 |
| Power-transistor ratings, physical basis for | 113 |
| Current and temperature ratings | 122 |
| Safe-area ratings | 134 |
| Second breakdown | 126 |
| Power-dissipation ratings | 123 |
| Thermal-cycling ratings | 144 |
| Voltage ratings | 113 |
| P-N Junctions | 6 |
| Forward-biased | 8 |
| Reverse-biased | 7 |
| Press-fit packages | 30 |
| Proportional heating control | 374 |
| Pulse transformers | 357 |
| Pulse triggering (of thyristors) | 225 |
| Punch-through voltage | 114, 209 |
| Push-pull switching converter | 302, 314 |
| Push-pull switching inverters | 321 |
| Ramp generators | 377 |
| Rating chart, thermal-cycle | 147 |
| Ratings and limiting characteristics (of thyristors) | 208 |
| Ratings, basis for | 16 |
| Ratings, types of | 17 |
| RC-compensated rectifier assemblies | 76 |
| RC triggering networks | 348 |
| Recovery-time test circuit | 70 |
| Rectification | 249 |
| Rectifiers: | |
| Controlled-avalanche types | 74 |
| Fast-recovery types | 70 |
| Reed relays | 357 |
| Regulation | 261 |
| Regulator, ac voltage | 400 |
| Relay links, microwave | 562 |
| Repetitive peak forward current | 67 |
| Repetitive peak off-stage voltage | 211 |
| Repetitive peak reverse voltage | 212 |
| Resistivity | 3 |
| Reverse-bias capacitance-discharge test | 132 |
| Reverse-bias $E_{s/b}$ test | 13 |
| Reverse-bias safe area ratings | 142 |
| Reverse-bias second breakdown | 128 |
| Reverse blocking current | 205 |
| Reverse current | 64 |
| Reverse-recovery time | 64, 234 |
| Reverse-transmission voltage ratio | 152 |
| Reverse voltages (of SCR's) | 212 |
| Reversing controls for induction motors | 398 |
| RF amplifier for military applications | 489 |
| RF power amplifiers | 430 |
| Circuit considerations | 445 |
| Class of operation | 439 |
| Design considerations | 439 |
| Matching networks | 447 |
| Modulation | 441 |
| Transistor selection | 444 |
| Ringing choke converter | 301, 306 |
| Rise times | 176, 231 |
| Safe-area design analysis | 137 |
| Inverter initial-turn-on analysis | 137 |
| Low-frequency analysis of direct-coupled audio amplifier | 139 |
| Safe-area ratings | 134 |
| Scan linearity | 634, 657 |
| Scanning Fundamentals | 623 |
| Scattering parameters | 152 |
| SCR capacitive-discharge automotive ignition system | 684 |
| Charging circuit | 686 |
| Output circuit | 688 |
| Performance | 692 |
| Protection circuit | 690 |
| SCR commutation circuit | 691 |
| SCR horizontal-deflection system | 642 |
| Auxiliary power supplies | 653 |
| Basic circuit | 643 |
| High-voltage generation | 649 |
| High-voltage regulation | 649 |
| Linearity correction | 651 |
| Over-all circuit | 654 |
| Raster correction | 652 |
| SCR inverter | 334 |
| Applications | 337 |
| Circuit operation | 334 |
| Gate-trigger-pulse generator | 336 |
| SCR motor-control, circuit, half-wave | 355 |
| Second breakdown | 126 |
| Evaluation techniques | 129 |
| Forward-bias capacitance-discharge test | 131 |
| Forward-bias $I_{s/b}$ test | 129 |
| Reverse-bias capacitance-discharge test | 132 |
| Reverse-bias $E_{s/b}$ test | 131 |
| Semiconductor Materials | 3 |
| Series regulators | 262 |
| S-shaping | 634, 651 |
| Short-circuit protection | 593 |
| Shunt regulators | 262, 283 |
| Silicon controlled rectifiers | 11, 194 |
| Construction | 206 |
| Ratings and characteristics | 208 |
| Two-transistor analogy | 195 |
| Voltage-current characteristics | 204 |
| Silicon power transistors | 11, 80 |
| Silicon rectifiers | 10, 58 |
| Controlled-avalanche type | 74 |
| Electrical characteristics | 63 |
| Fast-recovery types | 70 |
| High-voltage assemblies | 76 |
| Maximum ratings | 65 |
| Parallel arrangements of | 75 |
| Series arrangements of | 76 |
| Theory of operation | 58 |
| Thermal considerations | 62 |
| Single-sideband systems | 472 |
| Analysis of SSB signal | 473 |
| Bias control | 477 |
| Compensating diode | 479 |
| Intermodulation distortion | 474 |
| Linearity test | 473 |
| Transistor requirements | 475 |
| Single-time-constant triggering network | 348 |
| Six-phase star rectifier | 252 |
| Small-signal analysis of power transitors | 150 |
| Small-signal parameters, variation of | 166 |
| Bias dependence | 166 |
| Temperature dependence | 168 |
| Small-signal transistor equivalent circuits | 150 |
| Snubber network | 240 |
| Solid-state ballasting, advantages of | 429 |
| Solid-state ballasting circuits | 413 |
| Design procedure | 423 |
| Switching-regulator types | 416, 420 |
| Types of | 414 |
| Solid-state devices, types of | 9 |
| Sonobuoy transmitters | 494 |
| Space-charge region | 59 |
| S parameters | 152 |
| Speed controls for universal motors | 393 |
| Staged heat control | 380 |
| Step-down transformers | 355 |
| Step-down transistor switching regulator | 299 |
| Storage time | 176 |
| Structures, transistor | 99 |
| Stud packages | 28 |
| Subcycle surge-current rating | 69 |
| Surge current | 67 |
| Surge-current considerations | 381 |
| Surge ratings (of thyristors) | 217 |
| Switching characteristics (or thyristors) | 230 |
| Commutating capability (triacs) | 238 |
| Turn-on time | 231 |
| Turn-off time (SCR's) | 233 |
| Switching regulator | 287 |
| Switching-regulator ballast circuits | 416, 420 |
| Circuit components | 426 |
| Circuit design | 422 |
| Reactor element | 424 |
| Switching transistor | 426 |
| Switching speed | 175 |

| | Page |
|---|---|
| Switching times | 177 |
| Switching-time relationships | 177 |
| Switching, zero-voltage | 345 |
| Sync separation | 625 |
| Sync pulses | 624 |
| Temperature ratings | 21, 122 |
| Thermal capacitance | 41, 62 |
| Case-to-ambient | 45 |
| Junction-to-case | 42 |
| Thermal-circuit analog | 43 |
| Thermal conductance | 42 |
| Thermal conductivity | 42 |
| Thermal-cycling capability, improvement of | 146 |
| Thermal-cycling ratings | 56, 145 |
| Thermal-cycling rating chart | 147 |
| Thermal factors | 40 |
| Thermal fatigue | 40 |
| Therlmal-fatigue testing | 148 |
| Thermal impedance | 40, 62 |
| Basic relationships | 41 |
| Case-to-ambient | 45 |
| Junction-to-case | 42 |
| Thermal resistance | 41, 62 |
| Case-to-ambient | 45 |
| Junction-to-case | 42 |
| Thermal-stability requirements (for audio amplifiers) | 590 |
| Third-harmonic tuning | 631 |
| Three-layer diode | 13 |
| Three-phase double Y rectifier | 252 |
| Three-phase inductive load | 365 |
| Three-phase resistive loads | 364 |
| Three-phase triac controls | 363 |
| Three-phase Y full-wave rectifier | 251 |
| Three-phase Y half-wave rectifier | 251 |
| Thyristors | 11, 194 |
| Construction | 206 |
| Equivalent model analysis | 195 |
| Gate characteristics | 223 |
| Limiting characteristics | 208 |
| Potential-energy analysis | 199 |
| Ratings | 208 |
| Switching characteristics | 230 |
| Theory of operation | 194 |
| Voltage-current characteristics | 204 |
| Thyristors ac line-voltage controls | 339 |
| Thyristor current relationships | 341 |
| Thyristor power-control circuit configurations | 341 |
| Thyristor triggering circuit: | |
| Basic form of | 349 |
| Integrated-circuit type | 358 |
| Isolated types | 355 |
| Load line for | 350 |
| Neon-bulb type | 352 |
| Photocell types | 357 |
| Two-transistor type | 354 |
| Typical gate-current waveform | 351 |
| Unijunction-transistor type | 353 |
| Thyristor turn-on controls, types of | 339 |
| Total gain factor | 115 |
| Total turn-on time | 231 |
| Traffic-control flasher | 392 |
| Traffic-signal-lamp controls | 391 |
| Transformer, wide-band | 520 |
| Transformer, pulse | 357 |
| Transformers, step-down | 355 |
| Transistor horizontal-deflection circuits | 631 |
| Deflection circuit for color receiver | 640 |
| Deflection circuit for monochrome receiver | 636 |
| Drive considerations | 635 |
| Energy requirement | 635 |
| High-voltage power | 634 |
| Retrace time | 633 |
| Scan linearity | 634 |
| Voltage considerations | 632 |
| Transistor inverters and converters | 300 |
| Basic circuit configurations | 301 |
| Design of practical circuits | 306 |
| Feedback resistance | 327 |
| Second-breakdown considerations | 326 |
| Starting circuits | 327 |
| Transformer considerations | 303 |
| Transistor, injunction | 353 |
| Transistors, power | 11, 80 |
| Transition region | 6 |
| Transmitters, sonobuoy | 494 |
| Eighth-wave line sections | 468 |
| Half-wave line sections | 467 |
| Quarter-wave line sections | 468 |
| Tapered line section | 469 |
| Triacs | 12, 194 |
| Construction | 206 |
| Ratings and characteristics | 208 |
| Two-SCR analogy | 197 |
| Voltage-current characteristic | 205 |
| Trigger diodes | 352 |
| Triggering devices | 12, 349 |
| Basic operation | 349 |
| Types of | 351 |
| Voltage-current characteristic | 349 |
| Triggering configurations, basic | 347 |
| Triggering techniques | 346 |
| Turn-on time | 177 |
| Turn-on-time test circuit | 237 |
| TV deflection systems | 623 |
| Two-transistor trigger circuit | 354 |
| Types of Devices | 9 |
| Types of ratings | 17 |
| Ultrasonic generators | 615 |
| Ultrasonic power amplifiers | 621 |
| Ultrasonic power sources | 612 |
| Ultrasonic transducers, characteristics of | 612 |
| Unijunction transistors | 353 |
| Universal audio-amplifier design approach | 595 |
| Universal quasi-complementary-symmetry audio amplifier | 600 |
| Universal true-completary-symmetry audio amplifier | 596 |
| Vertical circuit that uses a complementary-symmetry output stage | 661 |
| Vertical driver and output stage | 664 |
| Vertical-switch and predriver circuit | 661 |
| Vertical circuit that uses a quasi-complementary-symmetry output stage | 665 |
| Vertical-deflection circuit that uses a conventional output stage | 657 |
| Driver stage | 658 |
| Linearity clamp | 660 |
| Output stage | 659 |
| Vertical switch | 657 |
| Vertical deflection system | 654 |
| Basic design approach | 656 |
| Simple deflection circuit | 655 |
| Versawatt plastic packages | 32 |
| Cleaning | 39 |
| Lead-forming techniques | 34 |
| Mounting | 35 |
| VHF/UHF transistor amplifiers, instabilities | 501 |
| Voltage-current characteristics (of thyristors) | 204 |
| Voltage ratings | 17, 113 |
| Voltage-reference diodes | 10 |
| Voltage reflection coefficient | 152 |
| Voltage-regulating, current-limiting power supplies | 263 |
| Voltage-regulating, current-regulating power supplies | 264 |
| Voltage-regulating power supplies | 262 |
| Wide-band transformer | 520 |
| Y-parameter equations | 151 |
| Zener diodes | 10 |
| Zero-voltage-switched heater controls | 373 |
| On-off type | 373 |
| Proportional type | 374 |
| Zero-voltage switch, integrated-circuit | 358 |
| Zero-voltage-switched lamp dimmer | 389 |
| Zero-voltage switching | 345 |
| Z-parameter equations | 151 |