RCA Power Circuits

DC to Microwave

RCA

SUGGESTED PRICE $2.00
TECHNICAL SERIES SP-51

RCA Power Circuits

DC to Microwave

Contents

RCA

Silicon Power Circuits Manual

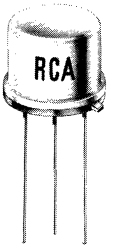
This Manual, like its preceding edition, provides design information for a broad range of power circuits using RCA silicon transistors, rectifiers, and thyristors (SCR's and triacs). It includes a brief introduction to semiconductor physics, as well as descriptions of construction, theory of operation, and important ratings and parameters for each type of device. Specific design criteria and procedures are presented for applications involving rectification and power-supply filtering, power conversion and regulation, ac line voltage controls, rf power amplifiers, and control and low-frequency power amplifiers. Design examples are given, and typical practical circuits are shown and analyzed.

Although this Manual is intended primarily for circuit and system designers working with solid-state power devices, portions of it will also be found useful by educators, students, radio amateurs, hobbyists, and others interested in the use of semiconductor devices and circuits.

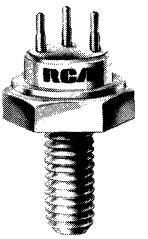
RCA | Electronic Components | Harrison, N.J. 07029

Printed in U.S.A. 1/69

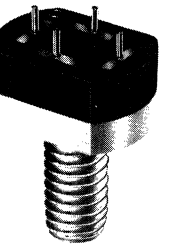
RF Power-Transistor Packages



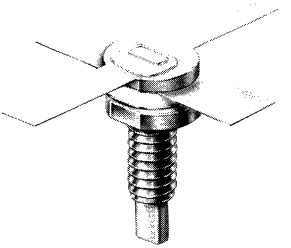
JEDEC TO-39



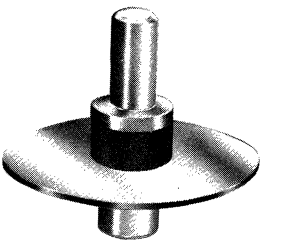
JEDEC TO-60



Molded-Silicone
Plastic Package



Hermetic Stripline
Ceramic-to-Metal
Package



Coaxial Package

Semiconductor Materials, Junctions, and Devices

SEMICONDUCTOR devices are small but versatile units that can perform an amazing 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, semiconductor 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, semiconductor 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. A relatively new material which combines the principal desirable features of both germanium and silicon is gallium arsenide. When further experience with this material has been obtained, it is expected to find much wider use in semiconductor devices.

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 germanium has a resistivity of

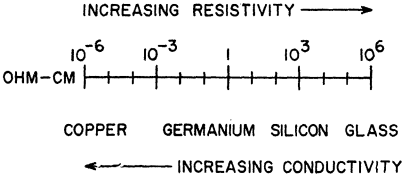


Figure 1. Resistivity of typical conductor, semiconductor, and insulator.

60 ohm-centimeters. Pure silicon has a considerably higher resistivity, in the order of 60,000 ohm-centimeters. As used in semiconductor 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 a structure has no loosely held electrons, semiconductor materials are poor conductors under normal conditions. In order to

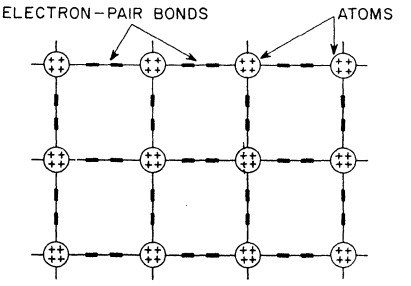


Figure 2. Crystal lattice structure.

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 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

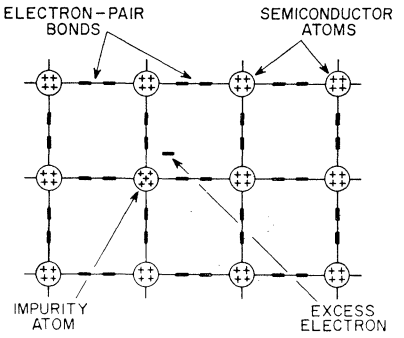


Figure 3. Lattice structure of n-type material.

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

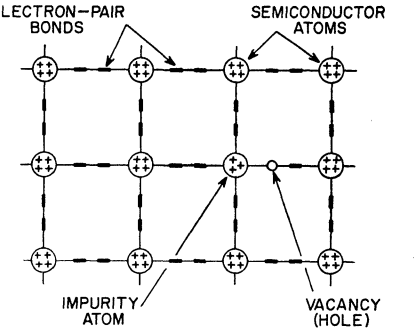


Figure 4. Lattice structure of p-type material.

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.

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 very important phenomenon occurs at the interface where the two materials meet (called the

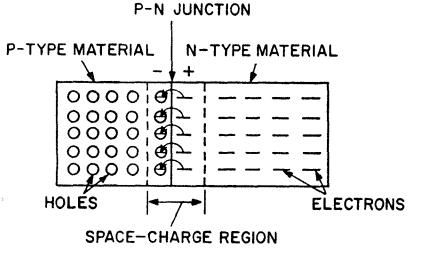


Figure 5. Interaction of holes and electrons at p-n junction.

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

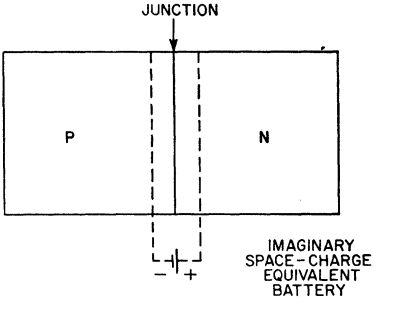


Figure 6. Potential gradient across space-charge region.

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.**

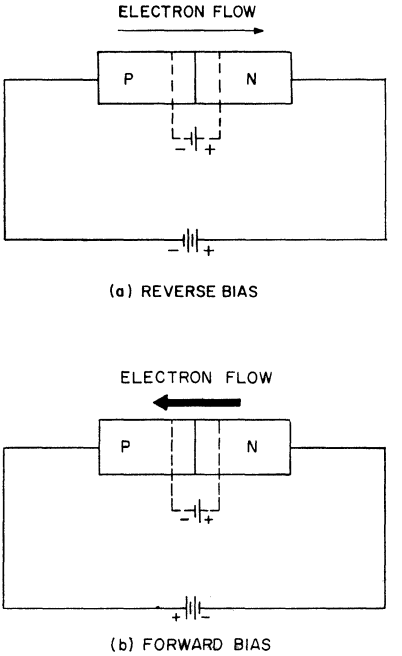


Figure 7. Electron current flow in biased p-n junctions.

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

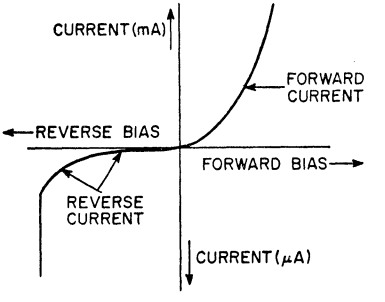


Figure 8. Voltage-current characteristic for a p-n junction.

reverse-bias and forward-bias regions. In 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 semiconductor 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

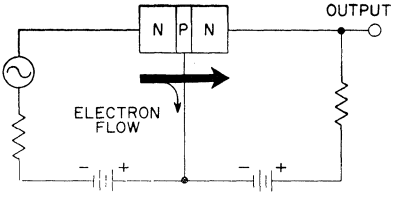


Figure 9. N-P-N structure biased for power gain.

the reverse direction to provide a high-resistance output circuit.

Electrons flow easily from the left-hand n-type region to the cen-

ter 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. (Many discussions of semiconductor theory assume that the "holes" in semiconductor material constitute the charge carriers in p-n-p devices, and discuss "hole currents" for these devices and "electron currents" for n-p-n devices. Other texts discuss neither hole current nor electron current, but rather "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. For the sake of simplicity, this discussion will be restricted to the concept of electron current flow, which travels from a negative to a positive terminal.)

Types of Devices

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

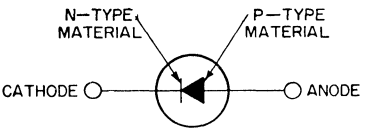


Figure 10. Schematic symbol for a semiconductor diode.

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

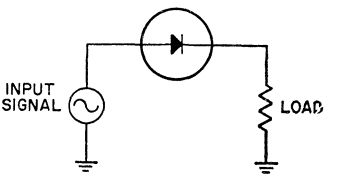


Figure 11. Simple diode rectifying circuit.

flows during the negative half cycle.

One of the most widely used types of semiconductor diode 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. 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.**

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. 12 shows the junction diagrams and schematic symbols for the SCR and triac. Such devices are discussed in the section on **Thyristors.**

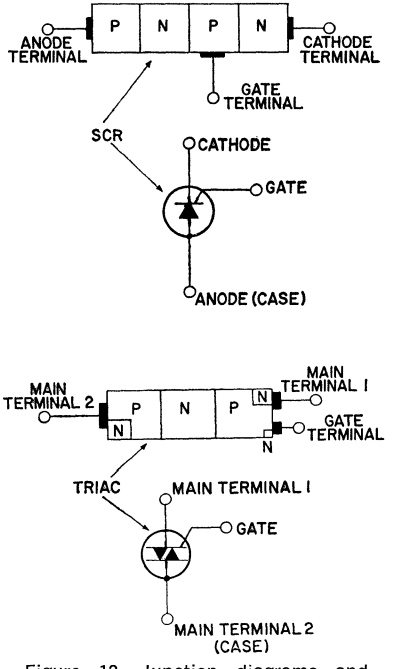


Figure 12. Junction diagrams and schematic symbols for SCR's and triacs.

Several variations of the basic junction-diode structure have been developed for use in special applications. One of the most important of these developments is the **tunnel diode,** which is used for amplification, switching, and pulse generation. This special diode is described in the RCA TRANSISTOR MANUAL SC-13 and in greater detail in the RCA TUNNEL DIODE MANUAL TD-30.

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. 13. In normal operation,

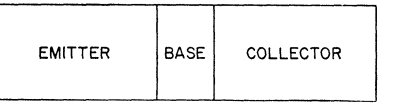


Figure 13. Functional diagram of transistor structure.

the emitter-to-base junction is biased in the forward direction and the collector-to-base junction in the reverse direction.

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. 14(a), electrons flow from the emitter to the collector. In the p-n-p transistor shown in Fig. 14(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 first two letters of the n-p-n and p-n-p designations indicate the respective polarities of the voltages applied to the emitter and the collector in normal operation. In an n-p-n transistor, the emitter is made negative with respect to both the collector and the base, and the collector is made positive with respect to both the emitter and the base. In a p-n-p transistor, the emitter is made positive with respect to both the collector and the base, and the collector is made negative with respect to both the emitter and the base.

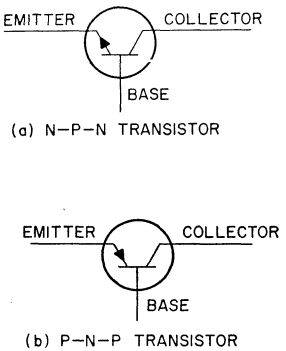


Figure 14. Schematic symbols for transistors.

The transistor, which is a three-element device, can be used for a wide variety of control functions, including amplification, oscillation, and frequency conversion. Power-transistor characteristics and ratings are discussed in detail in the section on **Silicon Power Transistors.**

Silicon Rectifiers

SILICON rectifiers can be operated at ambient temperatures up to 200°C and at current levels of hundreds of amperes and voltage levels as high as 1000 volts. In addition, parallel or series arrangements of two or more rectifiers can be used to provide even 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. The rectifiers are very small and lightweight, and can be made highly resistant to shock and other severe environmental conditions. In addition, they have excellent life characteristics which are not affected by aging, moisture, or temperature.

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. 15. In these diagrams, free electrons are represented by dots and free holes by circles; the movements of electrons and holes are indicated by arrows.

In equilibrium, as shown in Fig. 15(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

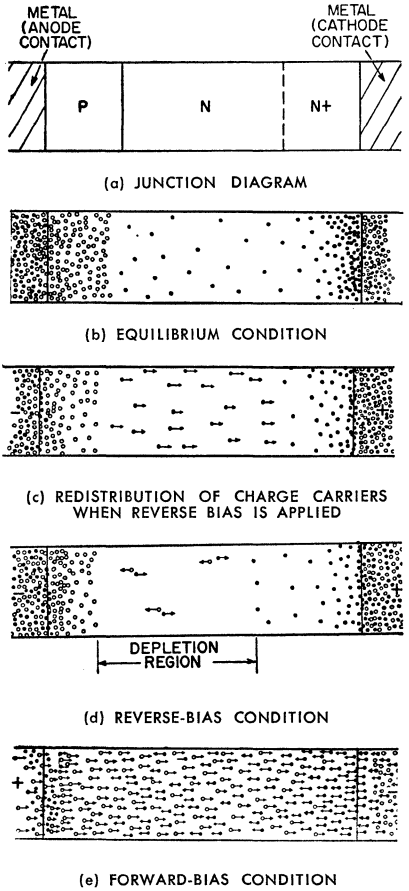


Figure 15. Concentrations of electrons (dots) and holes (circles) in a silicon rectifier.

semiconductor. The resulting behavior under forward- or reverse-bias conditions is as follows:

Reverse-Bias Operation—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. 15(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. 15(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. 15(d).

Forward-Bias Operation—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. 15(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. 16. 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. 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. 16(b).

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 elec-

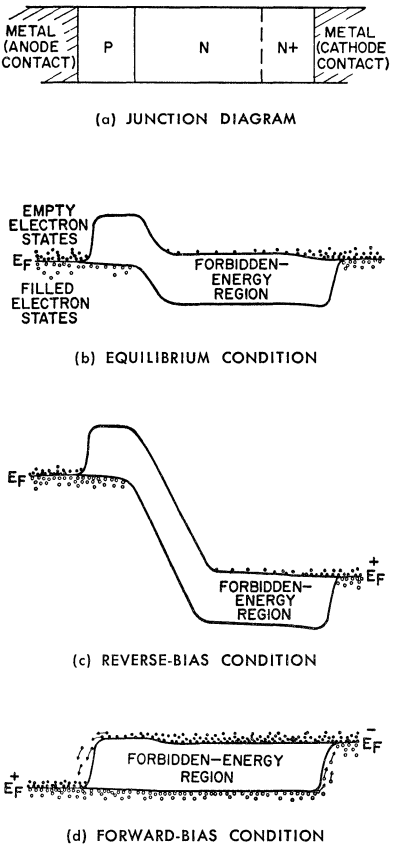


Figure 16. Potential-hill diagrams for various stages of rectifier operation (upward direction indicates increasing electron energy; downward direction indicates increasing hole energy).

trons 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. 16, 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 Operation—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. 16(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 Operation—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. 16(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. 16(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.

CHARACTERISTICS AND RATINGS

The characteristics data and the ratings given in the manufacturer's specifications on silicon rectifiers provide an important guide to the selection of the proper device for a given circuit application. Characteristics data provide the information that a circuit designer needs to predict the performance capabilities of his cir-

cuit, and form the basis for the ratings that define the safe operating limits for the rectifier.

Important Characteristics

Characteristics data given for silicon rectifiers are based on the manufacturer's determination of the inherent qualities and traits of the device. Characteristics are usually directly measurable attributes. Of the various rectifier characteristics for which data are given, four of the most important are thermal impedance, forward-voltage drop, reverse (leakage) current, and reverse recovery time. These four characteristics help determine the performance and environmental capabilities and limitations of rectifiers.

Thermal Impedance—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. 17 shows a cross-sectional diagram of a typical silicon rectifier. Because of the way in which the rectifier is constructed, there

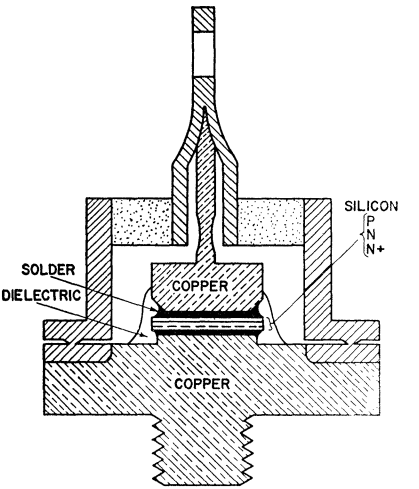


Figure 17. Cross-sectional diagram of a typical silicon rectifier.

is always a thermal "drop" between the p-n junction and the 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.

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. 18, 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

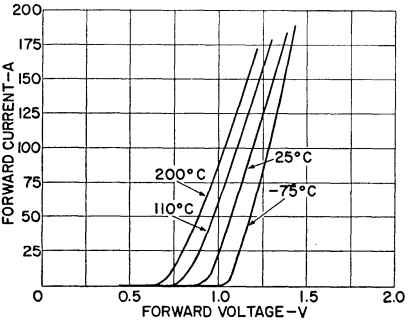


Figure 18. Typical forward characteristics of a silicon rectifier.

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. 18 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. 19, the reverse current increases slightly as the blocking voltage increases,

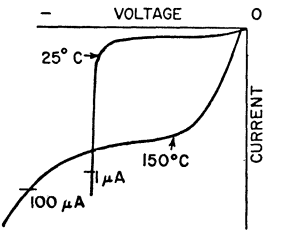


Figure 19. Typical reverse characteristics of a silicon rectifier.

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.

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, these carriers abruptly change direction and move out in the reverse direction. 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.**

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.

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 following list includes various factors for which silicon rectifiers must be rated: 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. 20.

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.

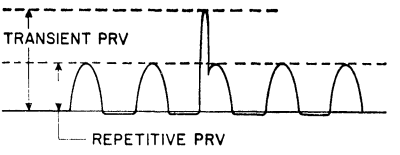


Figure 20. Typical waveform of repetitive and transient reverse voltages applied across a silicon rectifier.

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. 21.

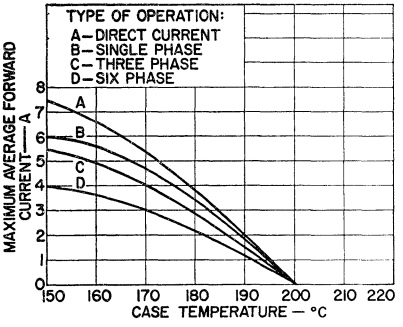


Figure 21. Current rating chart for a 12-ampere silicon rectifier.

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:

$$I_{avg} = \left[\frac{1}{\pi} \frac{\int_0^{\pi} I_{pk} \sin \omega t \, d(\omega t)}{2} \right] = \frac{1}{\pi} I_{pk} \quad (1a)$$

or

$$I_{pk} = \pi I_{avg} \quad (1b)$$

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

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

or

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

Table I summarizes the relationships expressed by Eqs. (1) and (2). 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 I—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}	$= \frac{1}{2} I_{pk}$	
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 high 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:

$$P_{(watts)} = (V_{dc} I_{dc}) + (I_{rms}^2 R_{dyn}) \quad (3)$$

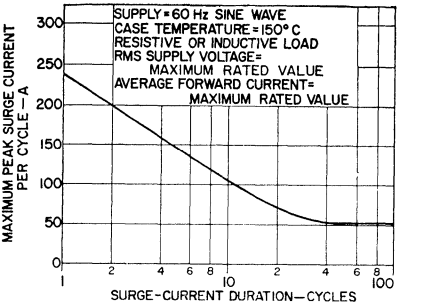
where the voltage V_{dc} is 0.4 to 0.9 volt depending upon the junction temperature; the direct current I_{dc} is equivalent to the average current I_{avg}; the current I_{rms} is the true rms current and, for a fixed average current, increases as the peak current increases; and R_{dyn} is the dynamic resistance

of the rectifier over the current range considered.

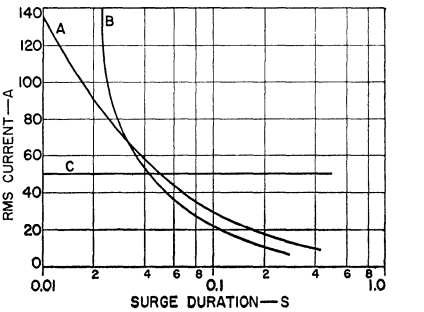
An analysis of Eq. (3) shows that if the peak current is increased and the conduction time is decreased so that the average current is held constant, the rms current and, therefore, the power dissipated in the rectifier ($I_{rms}^2 R_{dyn}$) are also increased. This behavior explains why the maximum 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. 21 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. 22(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 a higher surge rating than average current requirements indicate to be necessary may be used to meet surge requirements



(a)



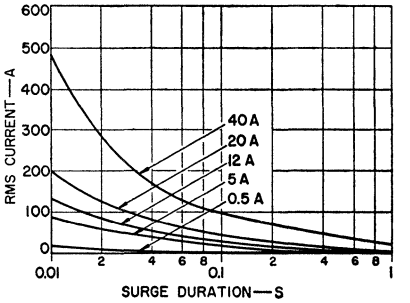
(b)

Figure 22. (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).

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. 22(b), should be constructed. This chart shows the surge rating of the rectifier (curve A), the opening char-

acteristics 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 current can be obtained from a universal surge-rating chart, such as that shown in Fig. 23. The opening characteristics of the fuse



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 23. Universal surge-current rating chart for RCA silicon rectifiers.

can be obtained from the manufacturer's published data, and the maximum surge current can be calculated.

The coordination chart shown in Fig. 22(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-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. 22(b), which is merely a reproduction of the 12-ampere curve on the universal rating chart shown in Fig. 23, 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 I 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. 22(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:

$$I^2t = \left(\frac{\text{one-cycle surge-current rating}}{2}\right)^2 \times 16.67 \times 10^{-3} \qquad (4)$$

The peak value of surge current that can be sustained by a 12-ampere silicon rectifier is given by the curve shown in Fig. 22(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:

$$I = \sqrt{\frac{I^2t}{t}} \qquad (5)$$

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

$$= 283 \text{ amperes}$$

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 I, as follows:

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

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

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 increases 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. 24, tends to cause an unequal distribution of voltage across the rectifiers. The disruptive effect of this stray capacitance is one reason for the use of a shunt capacitor C_s across each rectifier.

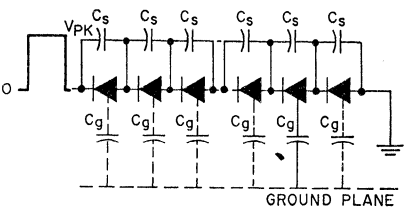


Figure 24. High-voltage rectifier assembly using shunt capacitors (C_s) to compensate for stray capacitances (C_g) and to equalize the reverse-recovery times of the rectifier units.

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 **Characteristics and Ratings,** 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.

Thyristors

THE term **thyristor** is the generic name for semiconductor devices that have characteristics similar to those of thyratron tubes. Basically, this group includes bistable semiconductor devices that have three or more junctions (i.e., four or more semiconductor layers) and that can be switched between conducting states (from OFF to ON or from ON to OFF) within at least one quadrant of the principal voltage-current characteristic. There are several different types of thyristors, which differ primarily in the number of electrode terminals and in their operating characteristics in the third quadrant of the voltage-current characteristic, as shown in Table II. Reverse-blocking triode thyristors, commonly called **silicon controlled rectifiers (SCR's)**, and bidirectional triode thyristors, usually referred to as **triacs**, are the most popular types. The discussions in this section deal primarily with these two thyristor devices.

Table II—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

A silicon controlled rectifier (SCR) is basically a four-layer p-n-p-n device that has three electrodes (a cathode, an anode, and a control electrode called the gate). Fig. 25 shows the junction diagram, principal voltage-current characteristic, and schematic symbol for an SCR. A triac also has three electrodes (main terminal No. 1, main terminal No. 2, and gate) and may be considered as two parallel p-n-p-n structures oriented in opposite directions to provide symmetrical bidirectional electrical characteristics. Fig. 26 shows the junction diagram, voltage-current characteristic, and schematic symbol for a triac. An analysis of the voltage-current characteristics of SCR's and triacs and of the charge-carrier interactions that make possible the switching transitions indicated by these characteristics provides useful information concerning the operation and possible applications of these devices.

Voltage-Current Characteristics

As shown in Fig. 25(b), the operation of an SCR under re-

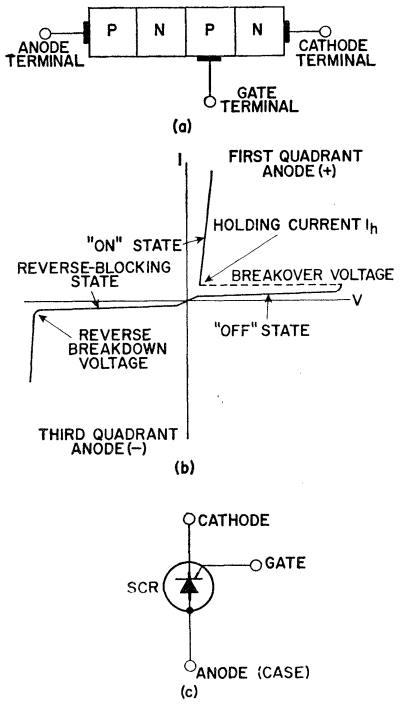


Figure 25. (a) Junction diagram, (b) principal voltage-current characteristic, and (c) schematic symbol for an SCR thyristor.

verse-bias conditions (anode negative with respect to cathode) is very similar to that of reverse-biased silicon rectifiers or other semiconductor 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.

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 OFF-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

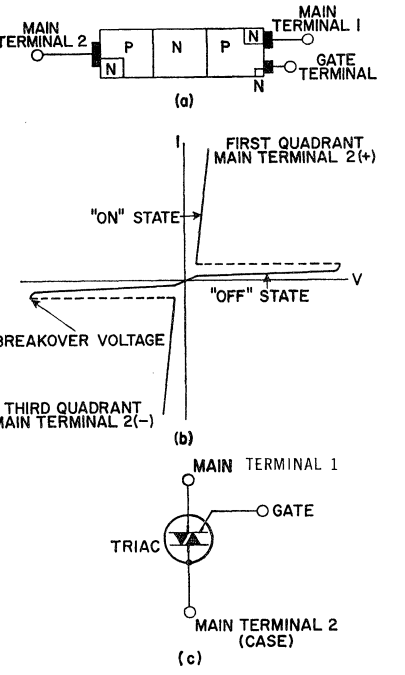


Figure 26. (a) Junction diagram, (b) principal voltage-current characteristic, and (c) schematic symbol for a triac thyristor.

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.

As shown in Fig. 26(b), a triac exhibits the forward-blocking, forward-conducting voltage-current characteristic of a p-n-p-n structure for either direction of applied voltage. 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 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.

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 III. The quadrant designations refer to the operating quadrant on the principal voltage-current characteristics, shown in Fig. 26(b) (either I or III), and the polarity symbol represents the gate-to-main-terminal-No. 1 voltage. Fig. 27 shows the flow of current in a triac for each of the four triggering modes.

Table III—Triac Triggering Modes

Gate-to-Main-Terminal-No.1 Voltage	Main-Terminal-No.2-to-Main-Terminal-No.1 Voltage	Operating Quadrant*
Positive	Positive	I (+)
Negative	Positive	I (—)
Positive	Negative	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

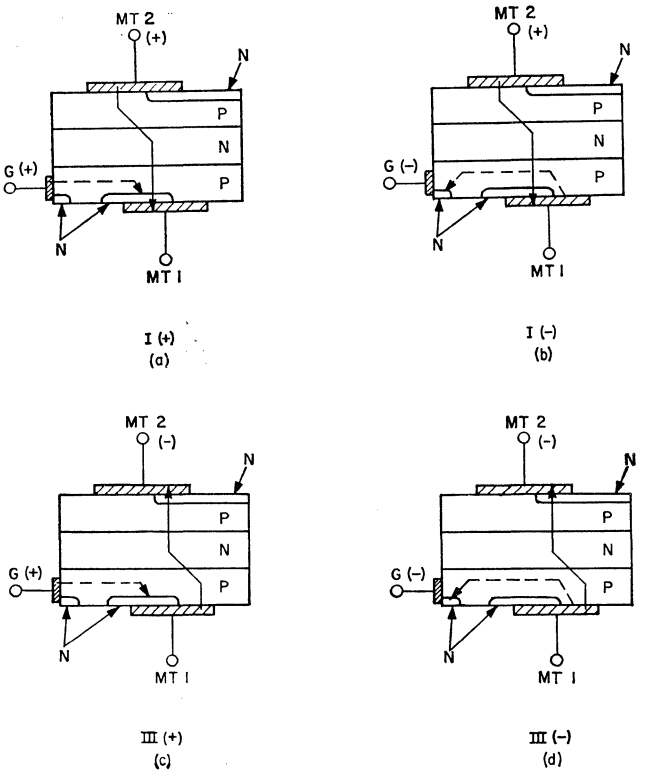


Figure 27. Current flow in the four triggering modes of a triac: (a) Mode I(+); (b) Mode I(—); (c) Mode III(+); (d) Mode III(—).

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 larger 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.

Thyristors are ideal for switching applications. When the working voltage of a 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.

The breakover voltage of a thy-

ristor can be varied, or controlled, by injection of a signal at the gate, as indicated by the family of curves shown in Fig. 28. Although this family of curves is shown in the first quadrant typical of SCR operation, a similar set

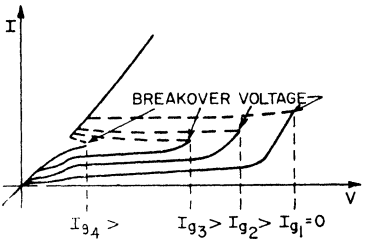


Figure 28. Curves showing breakover characteristics of a thyristor for different values of gate current.

of curves can also be drawn for the third quadrant to represent triac operation. When the gate current is zero, the principal voltage must reach the breakover value $V_{(BO)}$ of the device before breakover occurs. As the gate current is increased, however, the value of breakover voltage becomes less until the curve closely resembles that of a rectifier. In normal operation, thyristors are operated with critical values well below the breakover voltage and are made to switch ON by gate signals of sufficient amplitude to assure that the device is switched to the ON state at the instant desired.

After the thyristor is triggered by the gate signal, the current through the device is independent of gate voltage or gate current. The thyristor remains in the ON state until the principal current is reduced to a level below that required to sustain conduction.

Charge-Carrier Interactions

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. 29, 30, and 31. 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.

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 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 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. 29 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. 29(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. 29(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 this way, all free charge carriers are removed, and the hill becomes a depletion region, as shown in Fig. 29(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 displace-

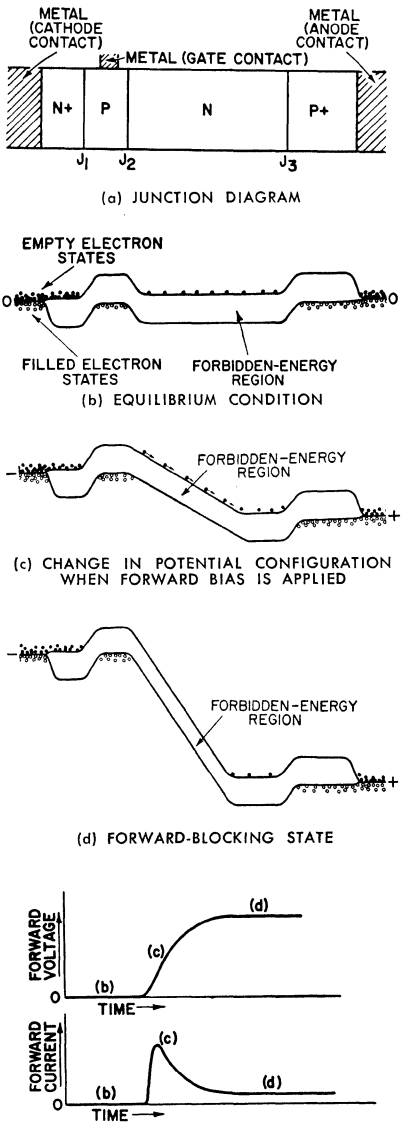


Figure 29. Potential-hill diagrams for various stages of thyristor transition from equilibrium condition to forward-blocking condition (electron energy increases upward, hole energy increases downward).

ment 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. 29(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. 30. When a thyristor is in the forward-blocking state, shown in Fig. 30(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. 30(c).

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

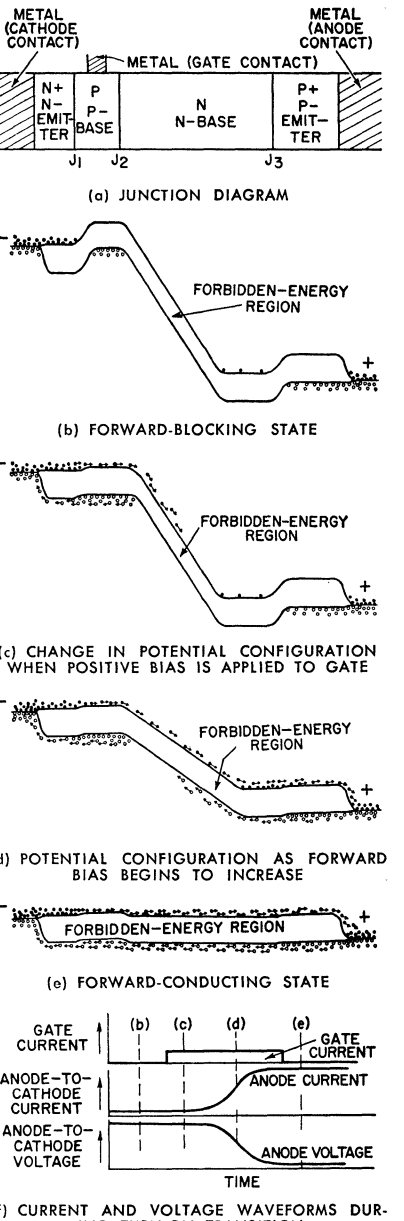


Figure 30. Potential-hill diagrams for various stages of thyristor transition from forward-blocking state to forward-conducting state.

the depletion region. These holes then immediately "climb" the potential hill at J_2, as shown in Fig. 30(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. 30(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. 30(f).

Turn-Off—The transition in the thyristor from the forward-conducting state back to the forward-blocking state is illustrated in Fig. 31. This transition is accomplished either by momentary reduction of the anode current to 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

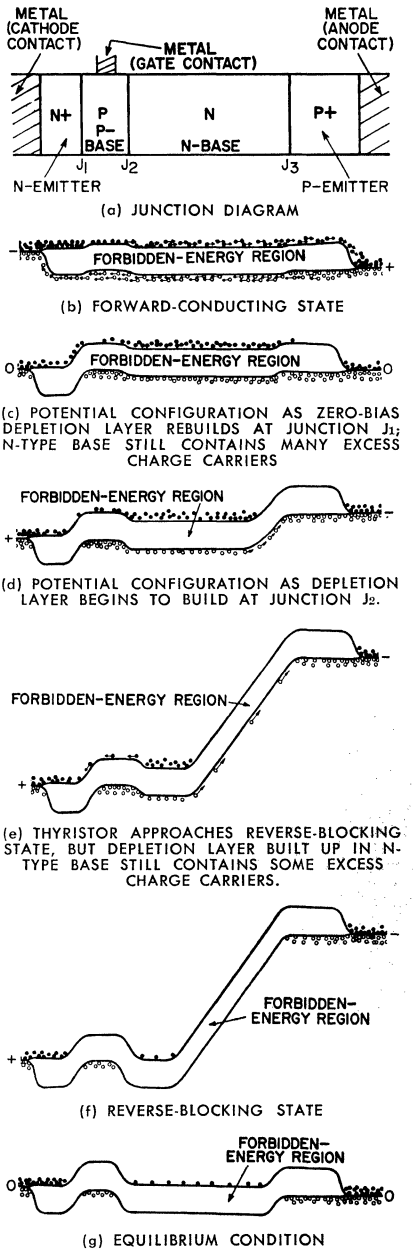


Figure 31. Potential-hill diagrams for various stages of thyristor transition from forward-conducting state to turn-off.

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. 29.

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. 31(c) to those shown in Fig. 31(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.

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. 31(d) through 31(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. 31(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. 31(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. 31(f), the excess carriers have all been removed, and device has reached its steady-state reverse-blocking condition. In Fig. 31(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 conditions described in Figs. 29, 30, and 31 are shown in Fig. 32.

CONSTRUCTION

Construction details for typical RCA thyristors are shown in Figs. 33 through 37. Fig. 33 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. A typical heat-sink arrangement for an insulating mounting of this package is shown in Fig. 34. (Various types of thyristor heat-sink arrangements are described in RCA Application Note AN-3822, "Thermal Considerations in Mounting of RCA Thyristors.") This package is used at current levels up to 7 amperes.

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. 35.

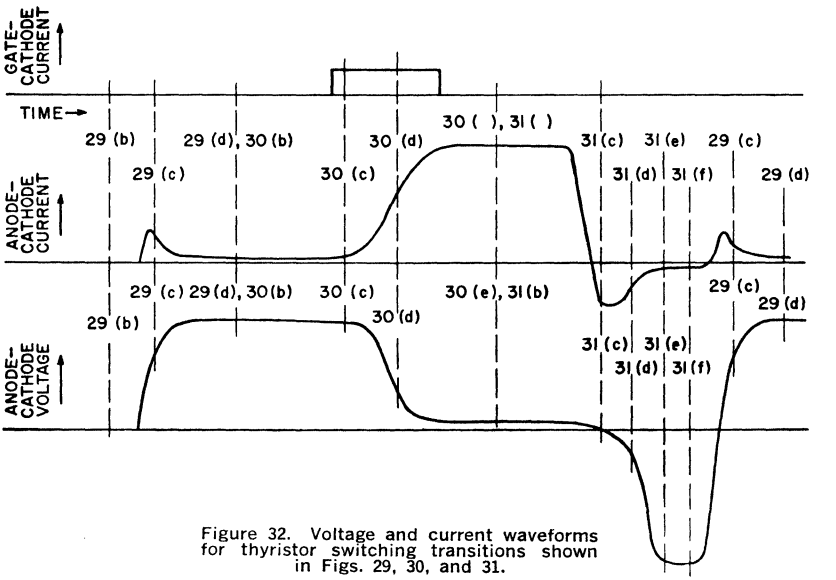


Figure 32. Voltage and current waveforms for thyristor switching transitions shown in Figs. 29, 30, and 31.

Construction details of a typical SCR pellet are shown in Fig. 36. 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. 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

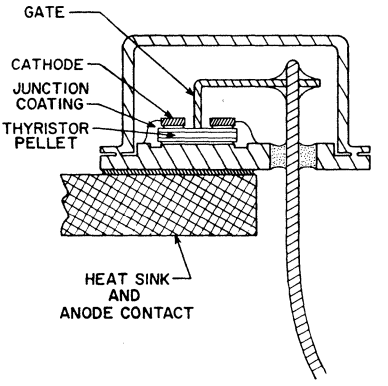


Figure 33. Cross-section of RCA two-lead TO-5 thyristor package.

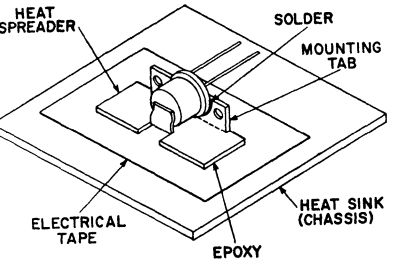


Figure 34. Typical heat-sink isolation technique for a chassis-mounted two-lead TO-5 thyristor.

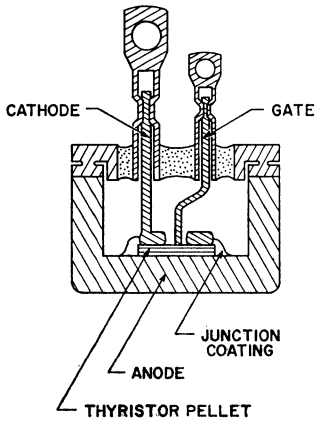


Figure 35. Cross-section of RCA press-fit thyristor package.

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.

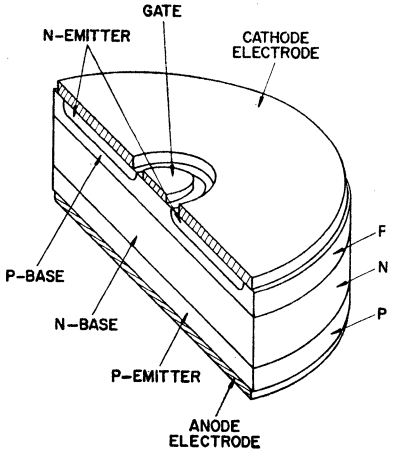


Figure 36. Cross-section of a typical SCR pellet.

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. 37. 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.

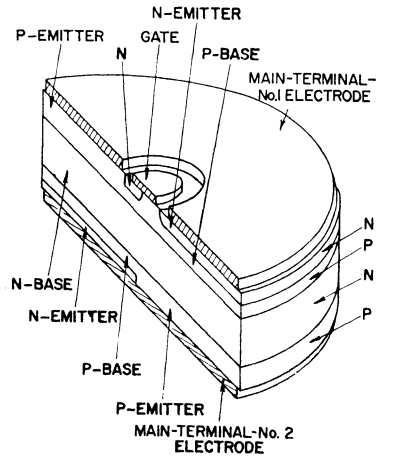


Figure 37. Cross-section of a typical triac pellet.

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 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. 36 and 37, 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 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 parameters, called 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.

Voltage and Temperature Ratings

The voltage ratings of thyristors are given for both steady-state and transient operation and for both forward- and 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 V_{DRM} is the maximum value of OFF-state voltage, either transient or steady-state, 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 V_{DRM} 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 specified with no external gate-to-cathode impedance 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 OFF-state voltage rating V_{DRM}. 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 **peak OFF-state voltage** V_{DM}. 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. 38 shows a simple, inexpensive test circuit that may be used to evaluate the OFF-state voltage capabilities of thyristors.

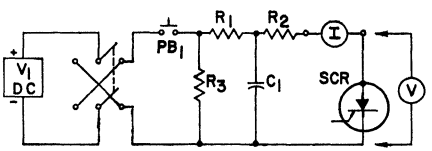


Figure 38. Test circuit used to determine dc forward- and reverse-voltage-blocking capabilities and leakage current of thyristors.

(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 Reverse-Blocking Thyristors)—Reverse-voltage ratings are given for SCR's to provide operating guidance in the third quadrant, or reverse-blocking mode. There are two voltage ratings for SCR's in the reverse-blocking mode: **repetitive peak reverse voltage** (V_{RRM}) and **nonrepetitive peak reverse voltage** (V_{RSM}).

The repetitive peak reverse voltage 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 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. 38 may be used for reverse-voltage tests of an SCR.

ON-State Voltage—When a thyristor is in a high-conduction state, the voltage drop across the device is no different in nature from the forward-conduction voltage drop of a rectifier, although the magnitude may be slightly higher. As in rectifiers, the ON-state voltage-drop characteristic is the major source of power losses in the operation of the thyristor, and the temperature produced

becomes a limiting factor in the rating of the device.

Thermal Resistance

The thermal resistance of a thyristor is an indication of the ability of the device to remove heat that is generated internally within the pellet. If a thyristor has a low thermal resistance, the junction temperature does not rise as high for a given conduction current or junction power dissipation. The most common or useful thyristor thermal resistance specified is the value from the junction to a particular point on the case. This value is referred to as the **junction-to-case thermal resistance,** $\theta_{J\text{-}C}$. The next value of thermal resistance that is necessary for proper use of the thyristor is the thermal resistance from the heat sink, which is usually attached to the thyristor, to the ambient air, $\theta_{C\text{-}A}$. If the thermal resistance of the total path from the junction to the ambient air and the power dissipated within the device are known, the average temperature can be calculated.

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. 39 shows curves of the maximum average forward power dissipation for the RCA-2N3873

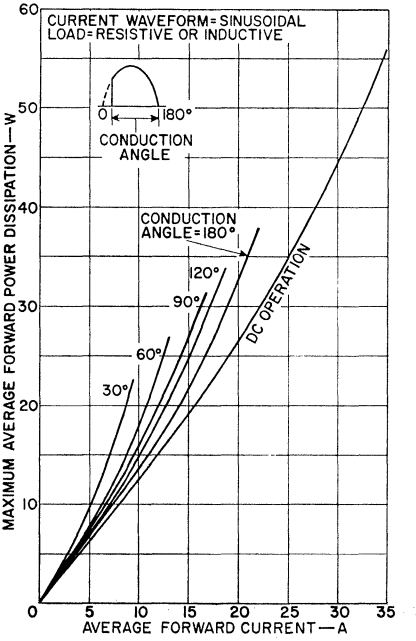


Figure 39. Power-dissipation rating chart for the RCA-2N3873 SCR.

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 temperature $T_{C(max)}$ is assumed to be 65°C, the maximum average forward power dissipation can be determined as follows:

$$P_{AVG(max)} = \frac{T_{J(max)} - T_{C(max)}}{\Theta_{J-C}} \quad (6)$$

$$= \frac{(100 - 65)\ °C}{0.92\ °C/watt}$$

$$= 38\ watts$$

The maximum average forward current rating for the specified conditions can then be determined from the rating curves shown in Fig. 39. 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.

ON-State Current—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.

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 **maximum average ON-state current rating** is usually specified for a half-sine-wave current at a particular frequency. Fig. 40 shows curves of the maximum allowable average ON-state current $I_{T F(avg)}$ for the RCA-2N3873 SCR family as a function of case temperature. Because peak and rms currents may be high for small conduction angles, the curves in Fig. 40 also show maximum allowable average currents as a function of conduction angle. The maximum operating junction temperature for the 2N3873 is 100°C.

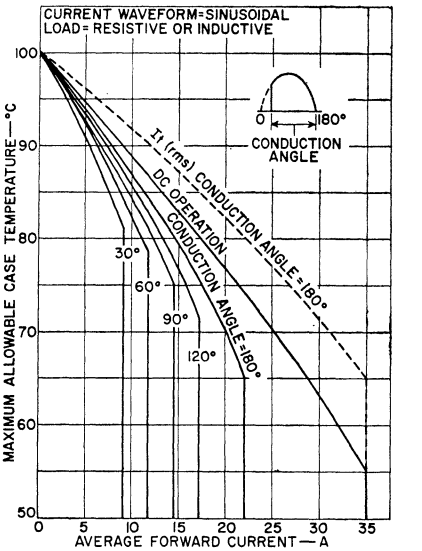


Figure 40. Current rating chart for the RCA-2N3873 SCR.

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 capacitive 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 $I_{TF(avg)}$, therefore, can be readily converted to the maximum rms ON-state current rating $I_{TF(rms)}$. For example, as may be determined from Fig. 40, 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:

$$\begin{aligned} I_{TF(rms)} &= I_{TF(avg)} \times 1.57 \\ &= 22 \text{ amperes} \times 1.57 \\ &= 35 \text{ amperes} \end{aligned}$$

The dashed-line curve in Fig. 40 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. 41 shows an rms ON-state current rating curve for a typical triac as

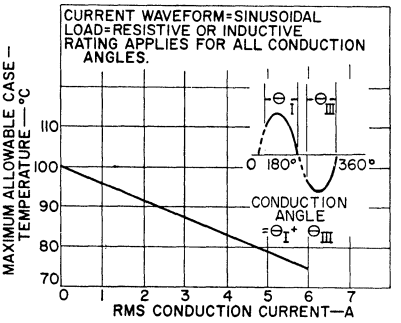


Figure 41. Current rating curve for a typical RCA triac.

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.

The **surge ON-state current rating** $I_{TF(surge)}$ 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. 42 shows a surge-current rating curve for the 2N3873 SCR. This

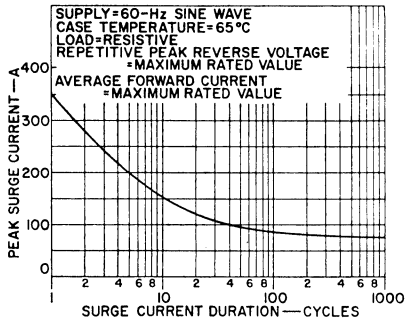


Figure 42. Surge-current rating curve for the RCA-2N3873 SCR.

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. 43 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.

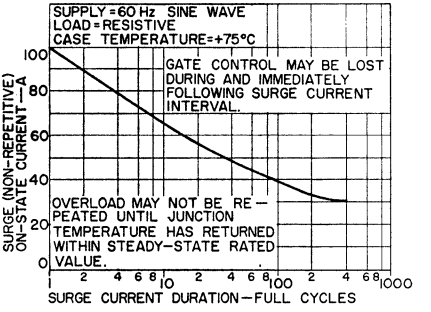


Figure 43. Surge-current rating chart for a typical triac.

Critical Rate of Rise of ON-State Current (di/dt)—In a thyristor, the load current is initially concentrated in a small area of the pellet when load current first begins to flow. This small area effectively limits the amount of current that the device can handle and results in a high voltage drop across the pellet in the first microsecond after the thyristor is triggered. If the rate of rise of current is not maintained within the rating of the thyristor, localized hot spots may occur within the pellet and permanent damage to the device may result. The waveshape for testing the di/dt capability of the RCA 2N3873 is shown in Fig. 44. The critical rate of rise of ON-state current is dependent upon

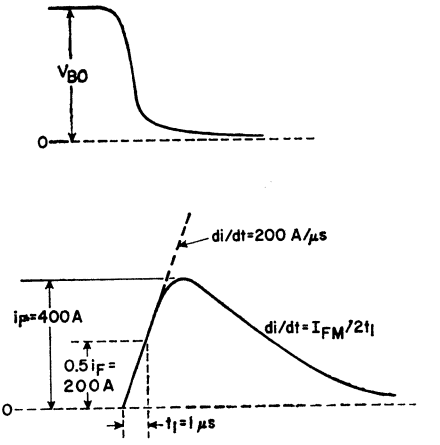


Figure 44. Voltage and current waveforms used to determine di/dt rating of the RCA-2N3873 SCR.

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.

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 rating of a thyristor specifies 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 latching-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 such 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. 45 shows a simple test circuit that may be used to determine the holding and latching currents of thyristors. For the holding-current tests, the value of potentiometer R_1 is adjusted to approximately 50 ohms, and the spring-loaded push-button switch PB_1 is momentarily depressed to turn on the thyristor. The value of R_1 is then gradually increased to the point at which the thyristor turns off.

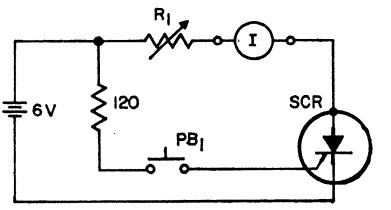


Figure 45. Test circuit used to determine holding and latching currents of thyristors.

For the latching-current test, the value of potentiometer R_1 is initially adjusted so that the main-terminal current is less than the holding level. The value of R_1 is then decreased, as push-button switch PB_1 is alternately depressed and released, until the thyristor latches on.

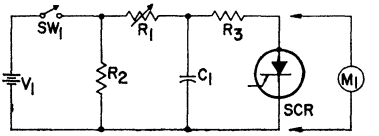
Critical Rate of Rise of OFF-State Voltage (dv/dt)—Because of the internal capacitance of a thyristor, the forward-blocking capability of the device is sensitive to the rate at which the forward voltage is applied. A steep rising voltage impressed across the main terminals of a thyristor causes a capacitive charging current to flow through the device. This charging current ($i = Cdv/dt$) is a function of the rate of rise of the OFF-state voltage.

If the rate of rise of the forward voltage exceeds a critical value, the capacitive charging current may become large enough to trigger the thyristor. The steeper the wavefront of applied forward voltage, the smaller the value of the breakover voltage becomes.

The use of the shorted emitter construction in RCA SCR's has

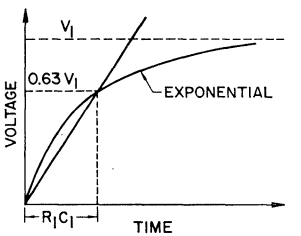
resulted in a substantial increase in the dv/dt capability of these devices by providing a shunt path around the gate-to-cathode junction. Typical units can withstand rates of voltage rise up to 200 volts per microsecond under worst-case conditions. The dv/dt capability of a thyristor decreases as the 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. 46(a) shows a simple test circuit that may be used to determine the dv/dt capability of a thyristor. The curves in Fig. 46(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 46. (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:

$$\frac{dv}{dt} = \frac{\text{rated value of thyristor voltage } (V_{BO})}{\text{RC time constant}} \times 0.632 \quad (7)$$

The dv/dt rating 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.

It has been found in many applications that simple circuit additions, such as shown in Fig. 47, can be used to reduce the dv/dt

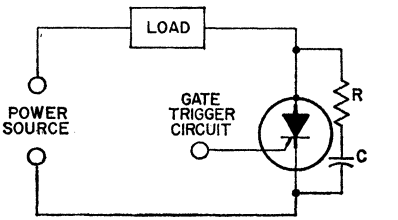


Figure 47. Diagram showing use of RC network to improve the dv/dt capability of an SCR.

stress on the thyristor. The dv/dt capability is also increased by application of reverse bias to the gate during the rise of OFF-state voltage.

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-performance 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. 48. The total turn-on time t_{gt} is defined as the time interval between the initiation of the gate

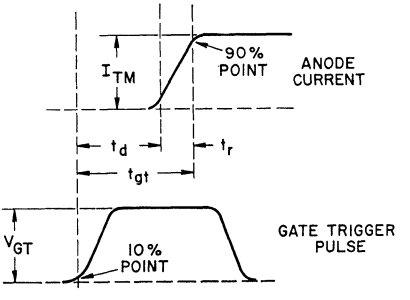


Figure 48. Gate current and voltage turn-on waveforms for a thyristor.

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. 49 shows the variation in turn-on time with gate-trigger current for the RCA-2N3873 SCR.

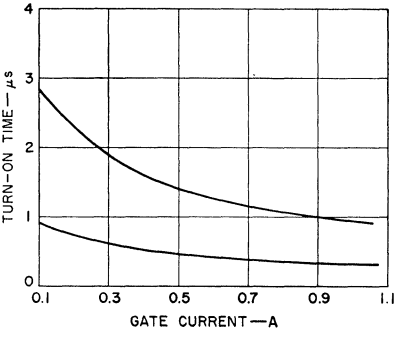


Figure 49. Turn-on time characteristics for the RCA-2N3873 SCR.

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. 50 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.

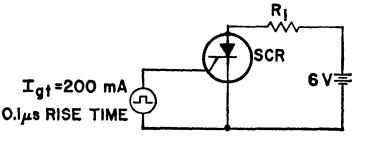


Figure 50. Test circuit used to determine turn-on time of thyristors.

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. 51. 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.

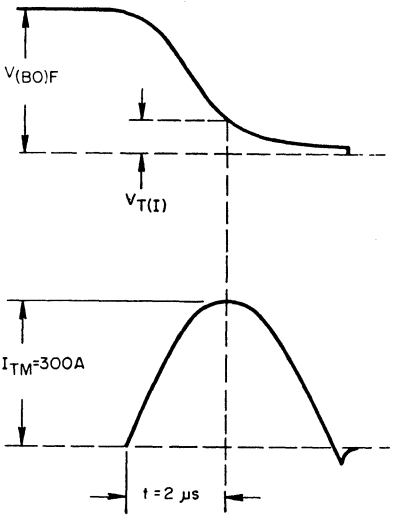


Figure 51. Initial ON-state voltage and current waveforms for the 2N3873 SCR.

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. 52. 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.

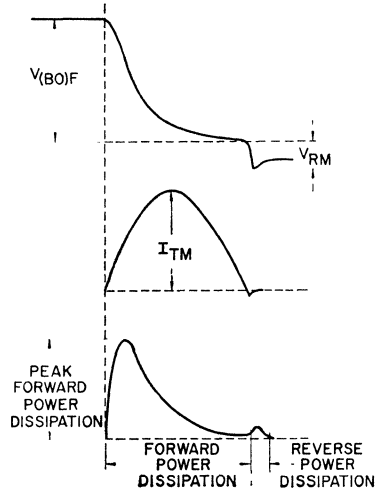


Figure 52. Instantaneous power dissipation in a thyristor during turn-on.

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—The turn-off time of an SCR also consists of two stages, a reverse-recovery time and a gate-recovery time, as shown in Fig. 53. When the forward current of an SCR is reduced to zero at the end of a conduction period, application of reverse voltage between the anode and cathode terminals causes reverse current to flow in the SCR until the reverse-blocking junction establishes a depletion region. The time interval between the application of reverse voltage and the time that the reverse current passes its peak value to a steady-state level is called the **reverse-recovery time** t_{rr}. A second recovery period, called the **gate-recovery time** t_{gr}, must then elapse for the forward-blocking junction to

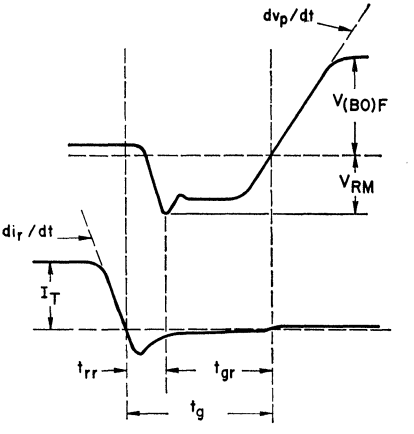


Figure 53. Circuit-commutated turn-off voltage and current waveforms for a thyristor.

establish a forward-depletion region so that forward blocking voltage can be re-applied and successfully blocked by the SCR.

The gate-recovery time of an SCR is usually much longer than the reverse-recovery time. The total time from the instant reverse-recovery current begins to flow to the start of the re-applied forward-blocking voltage is referred to as the circuit **commutated turn-off time** t_g. The turn-off time is dependent upon a number of circuit parameters, including the ON-state current prior to turn-off, the rate of change of current during the forward-to-reverse transition, the reverse-blocking voltage, the rate of change of the re-applied forward voltage, the gate trigger level, the gate bias, and the junction temperature. The junction temperature and the ON-state current, however, have a more significant effect on turn-off time than any of the other factors. 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. 54 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

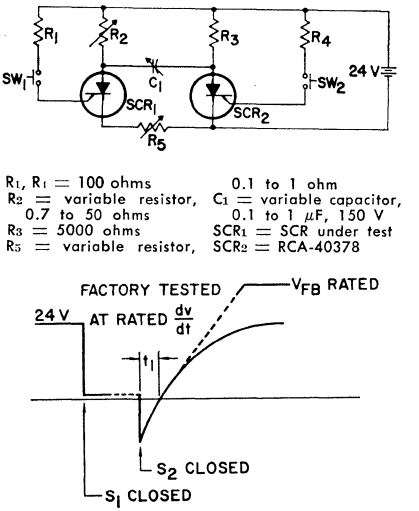


Figure 54. Test circuit and voltage waveforms used to determine turn-off times of thyristors.

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 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. 54.

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 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. 55 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.

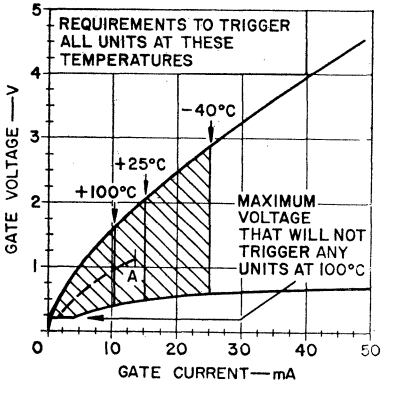


Figure 55. Gate-characteristics curves for a typical RCA SCR.

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. 56. 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-

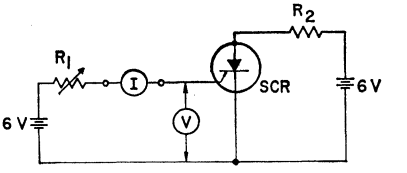


Figure 56. Test circuit used to determine gate-trigger-pulse requirements of thyristors.

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.

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—When very precise triggering of a thyristor is desired, the thyristor gate must be overdriven by a pulse of current much larger than the dc gate current required to trigger the device. The use of a large current pulse reduces variations in turn-on time, minimizes the effect of temperature variations on triggering characteristics, and makes possible very short switching times.

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 **Switching Characteristics,** 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. 57, 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 application of gate power. The curves shown in Fig. 57(a) are for RCA SCR's that have relatively small current ratings (2N4101, 2N4102, and 40379 families), and the curves shown in Fig. 57(b) are for RCA SCR's that have larger current ratings (2N3670, 2N3873, and 2N3899 families). Because the higher-current thyristors have larger pellets, they also have greater thermal capacities 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. 58. These curves indicate that reverse dissipations do not exceed the maximum allowable power dissipation for the device.

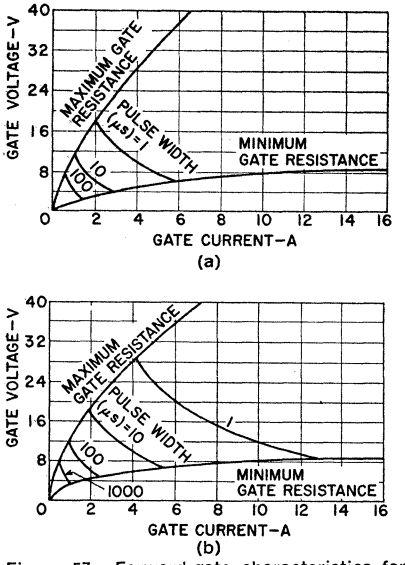


Figure 57. Forward-gate characteristics for pulse triggering of RCA SCR's: (a) low-current types, (b) high-current types.

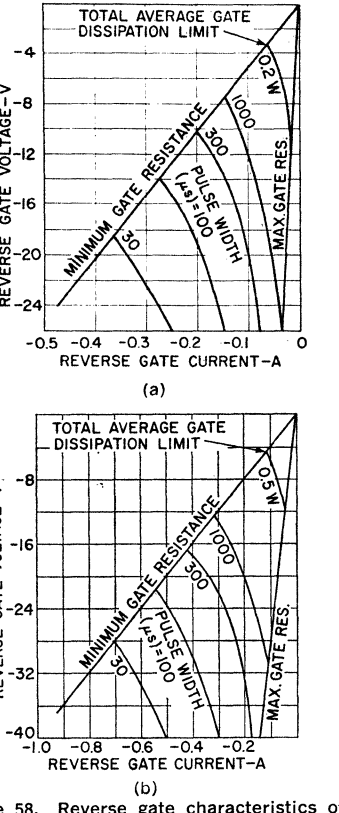


Figure 58. Reverse gate characteristics of RCA SCR's: (a) low-current types, (b) high-current types.

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. 59. 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.

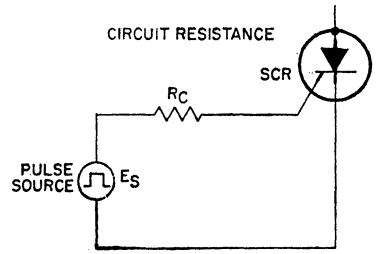


Figure 59. Equivalent diagram of the basic gate-trigger circuit for a thyristor.

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.

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. 60. 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

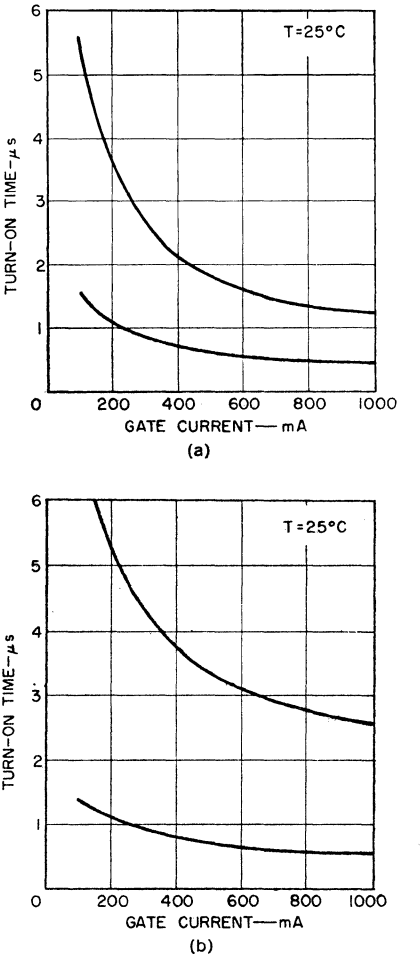


Figure 60. Turn-on time distribution among RCA SCR's: (a) low-current types, (b) high-current types.

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.

The following example, in which an RCA-2N3873 SCR is to be turned on in 2.5 microseconds, demonstrates the use of the various characteristics in the solution of a typical triggering problem:

The turn-on characteristics shown in Fig. 60 indicate that a gate-trigger current of 1 ampere is required to insure that all devices of this type will turn on in 2.5 microseconds (the 2.5-microsecond ordinate level intersects the upper curve at 1000 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. 61. 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 (1 ampere) 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

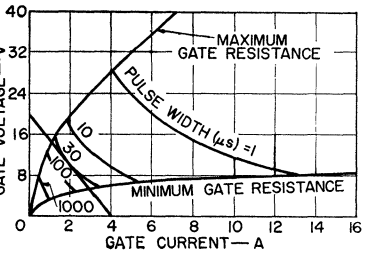


Figure 61. Forward gate characteristics of typical RCA SCR's showing load line for a source of 20 volts and a required gate current of 1 ampere.

2.5-microsecond curve). The maximum allowable pulse width is obtained by estimation of the pulse-width curve tangent to the load 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:

$$f = P_{g(avg)}/P_{g(pk)} \times PW_g \qquad (8)$$

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 of the RCA-2N3873, or 0.25 watt, is used to trigger the device, then the maximum allowable operating frequency is determined as follows:

$$f = \frac{0.25\ \text{W}}{20\text{W} \times 2.5 \times 10^{-6}\,\text{second}}$$
$$= 5000\ \text{Hz}$$

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. 62 represents a reverse gate-source 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:

$$PW = \frac{0.25\ \text{W}}{5000\ \text{Hz} \times 10\text{W}} \qquad (9)$$
$$= 5\ \text{microseconds}$$

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. 62. In this example, the maximum allowable pulse width for reverse dissipation is 100 microseconds.

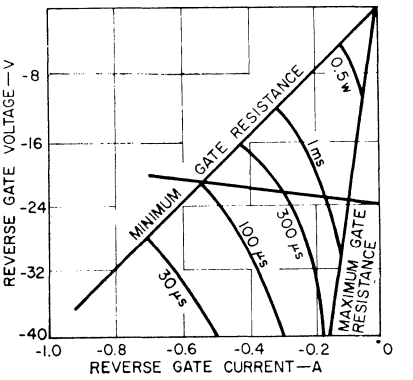


Figure 62. Reverse gate characteristics for typical RCA SCR's showing load line for a reverse gate-source voltage of 24 volts and an external circuit resistance of 5 ohms.

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.

SERIES AND PARALLEL OPERATION

The voltage or current capabilities of a single thyristor can be extended by use of two or more thyristors of the same type in series or parallel arrangements, respectively. The following paragraphs discuss basic considerations important to the successful operation of thyristors used in multiple connections.

Series Connections

When thyristors are connected in series for higher-voltage operation, certain procedures should be followed. These procedures usually depend upon the typical electrical characteristics of the thyristors used and the requirements of the circuit application.

The most important consideration in series connections of thyristors is to assure that voltages are divided equally across the individual units in the series string. One technique that may be used to obtain the desired voltage distribution is to select units that are matched with respect to such characteristics as OFF-state voltage breakdown, reverse voltage and current, and temperature effects. The use of a resistor across each unit in the stack is also recommended for improved series operation of thyristors. The value of the resistors should be some fraction of the maximum OFF-state resistance of the thyristor to force equal voltage division among the devices.

When SCR's are used in series arrangements, differences in the reverse-recovery times of the units have an important bearing on the voltage division. Variation in internal capacitances of thyristors and in stray capacitances between thyristors and ground can also result in an unequal voltage division among the various units. The use of capacitive voltage dividers is recommended to eliminate the effects of such conditions. When capacitive voltage dividers are used, however, a damping resistor should be connected in series with each capacitor to restrict peak-current values when the thyristors are switched to the ON state while the OFF-state voltage is present on the capacitor.

When thyristors are connected in series, the gate-trigger circuit used to turn on the various units requires special consideration. Because of differences in the delay and rise times of thyristors, gate-trigger currents that have a fast rise time must be used to turn on the units in the series string. The use of large gate-trigger currents minimizes turn-on differences. If large, quick-rising trigger currents are not used, the voltage across units that have longer turn-on times may exceed peak-voltage ratings.

Parallel Connections

Thyristors are connected in parallel to obtain output currents higher than the current ratings of an individual thyristor. The main consideration for this type

of operation is that the current must be divided equally among the parallel thyristors. One technique that is used to assure proper current division is to connect an identical balancing resistor in series with each thyristor. The value of these resistors should be several times larger than the maximum ON-state impedance of the thyristors so that the current through each thyristor will be essentially the same even though the ON-state impedances of the thyristors are different. The addition of a balancing resistor to each conduction path, however, increases the power dissipation and, consequently, decreases the over-all efficiency of the circuit. The efficiency of the circuit is improved if reactors, rather than resistors, are used to achieve balanced currents.

Another technique used in the parallel connection of thyristors is to select matched devices on the basis of specific conduction characteristics. When this technique is employed, circuit and load impedances must be carefully designed to assure an equal impedance for each conduction path in the parallel array. When factory-matched units are employed, care must be taken to insure that all units are operated at essentially the same case temperature. Because the forward voltage drop of a thyristor is temperature-dependent, differences in case temperature can result in unequal current sharing. All thyristors in the parallel array, therefore, should be mounted on a common heat sink to assure that the operating junction temperature of each device is the same.

When thyristors are connected in parallel, it is preferable to use a continuous gate drive to turn on the devices because of the differences in the latching levels of individual thyristors. Continuous gate drive is particularly important when inductive loads are used because such loads produce slow-rising output currents, and the continuous drive assures that gate current is present throughout the full conduction period. If pulse triggering is employed, the duration of the gate-trigger pulse must be sufficient to allow the conduction currents through all the thyristors to build up to values greater than the latching values to assure that all thyristors are completely turned on. The gate-trigger pulses should be fast-rising, high-amplitude pulses to assure good current sharing among the parallel thyristors during the turn-on interval.

Consideration should be given in parallel arrays to the possibility that one thyristor may be inadvertently turned on from some extraneous source, e.g., a high rate of rise of OFF-state voltage (dv/dt). Under such conditions, it is possible that an excessive amount of current may flow through this thyristor.

TRANSIENT PROTECTION

Voltage transients occur in electrical systems when some disturbance disrupts the normal operation of the system. These disturbances may be produced by various sources (such as lightning surges, energizing transformers, and load switching) and may generate voltages which exceed the rating of the thyristors. In addition, transients generally have a fast rate of rise that is usually greater than the critical

value for the rate of rise of the thyristor OFF-state voltage (static dv/dt).

If transient voltages have magnitudes far greater than the device rating, the thyristor may switch from the OFF state to the ON state; the excess voltage is then transferred from the thyristor to the load. Because the internal resistance of the thyristor is high during the OFF state, the transients may cause considerable energy to be dissipated in the thyristor before breakover occurs. In such instances, the transient voltage exceeds the maximum allowable voltage rating, and irreversible damage to the thyristor may occur.

Even if the magnitude of a transient voltage is within the maximum allowable voltage rating of the thyristor, the rate of rise of the transient may exceed the static dv/dt capability of the thyristor and cause the device to switch from the OFF state to the ON state. In this case, thyristor switching from the OFF state to the ON state occurs because of the fast rate of rise of OFF-state voltage (dv/dt) and the thyristor capacitance (C), which result in a turn-on current $i = Cdv/dt$. Thyristor switching produced in this way is free from high-energy dissipation, and turn-on is not destructive provided that the load current that results is within the device capability.

In either case, transient suppression techniques are employed to minimize the effects of turn-on because of overvoltage or because the thyristor dv/dt capability is exceeded.

One of the obvious solutions to insure that transients do not exceed the maximum allowable voltage rating is to provide a thyristor with a voltage rating greater than the highest transient voltage expected in a system. This technique, however, does not represent an economical solution because, in most cases, the transient magnitude, which is dependent on the source of transient generation, is not easily defined. (Transient voltages as high as 2600 volts have resulted from lightning disturbances on a 120-volt residential power line.) Usually, the best solution is to specify devices that can withstand voltage from 2 to 3 times the steady-state value. This technique provides a reasonable safety factor. The effects of voltage transients can further be minimized by use of external circuit elements, such as a series RC network, across the thyristor terminals, as shown in Fig. 63.

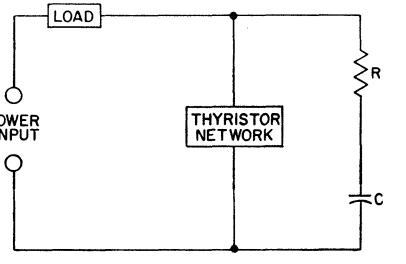


Figure 63. Diagram showing use of RC network for transient suppression in thyristor circuits.

The rate at which the voltage rises at the thyristor terminal is a function of both the load impedance and the values of the resistor R and the capacitor C. Because the load impedance is usually variable, the preferred approach is to assume a worst-case condition for the load and, through actual transient mea-

surement, to select a value of C that provides the minimum rate of rise at the thyristor terminals. The resistance R should be selected to minimize the capacitor discharge currents during turn-on.

For applications in which it is necessary to minimize false turn-on because of transients, the addition of a coil in series with the load, as shown in Fig. 64, is very effective for reduction of the rate of voltage rise at the thyristor terminals. For example, if a transient of infinite rise time is assumed to occur at the input terminals and if the effects of the load impedance are neglected, the rise time of the transient at the thyristor terminals is approximately equal to $E_{pk}/\sqrt{LC}$. If the value of the added inductor L is 100 microhenries and the value of the snubber capacitor C is 0.1 microfarad, the infinite rate of rise of the transient at the thyristor terminals is reduced by a factor of 3. For a filter network consisting of $L = 100$ microhenries, $C = 0.22$ microfarad, and $R = 47$ ohms, a 1000-volt-per-microsecond transient that appears at the input terminals is suppressed by a factor of 6 at the thyristor terminals.

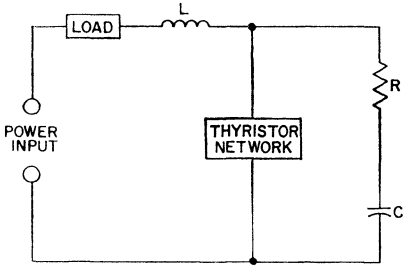


Figure 64. Diagram showing use of RC network for transient suppression and an inductance to prevent false turn-on because of transients in thyristor circuits.

COMMUTATING dv/dt CONSIDERATIONS

In ac power control, a triac must switch from the conducting state to the blocking state at each zero current point, or twice per cycle. This action is called commutation. If the triac fails to block the circuit voltage following the zero-current point, control of the load power is lost. This action is not damaging to the triac. 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 by the triac without reverting to the ON state is termed the critical rate of rise of commutation voltage for the triac or the commutating dv/dt.

SCR's do not experience commutation limitations because turn-on is not possible for the polarity of voltage opposite to current flow.

The triac may be considered as two SCR's in an inverse-parallel connection with the exception that the high-voltage blocking function is common to both SCR's. In the circuit shown in Fig. 65, during the zero-current crossover (point A), the half of the triac in conduction starts to commutate when the main current flow falls below the holding current. At the instant the conducting half of the triac turns off, an applied voltage opposite

to the previous current polarity is applied across the triac. When this voltage is applied, a displacement current results from the formation of a depletion region at the junction. The portion of this displacement current which crosses the n-type emitter junction of the other side of the triac may be sufficient to trigger the device into the ON state in that direction. The over-all result is loss of power control to the load. The rate of decrease of current prior to the zero-current point and the rate of application of voltage in the opposite polarity determine the commutating duty on the triac.

The most economical approach to reduction of the dv/dt stress so that it is within the capability of the triac is the use of a series RC network across the triac. The rate of change of re-applied voltage is then dependent on the inductance and capacitance of the load and the impedance of the network. The magnitude of added capacitance is determined by the load impedance and the dv/dt limitation of the triac. The value of the added resistance R should be sufficient both to damp the LC oscillation and to maintain the capacitor discharge currents during triac turn-on within acceptable limits.

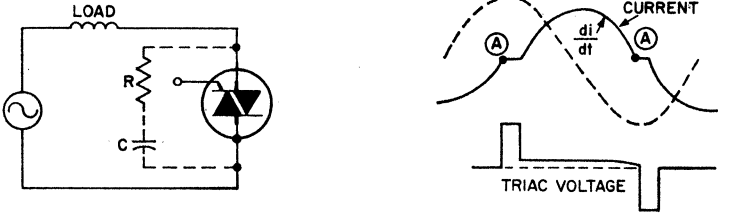


Figure 65. Simplified diagram and voltage and current waveforms for operation of a triac in a circuit that has an inductive load and a lagging-current power factor.

Silicon Power Transistors

THE performance of power transistors in electronic equipment depends on many factors besides the basic characteristics of the semiconductor material. The two most important factors are the design and fabrication of the transistor structure and the manner in which power is dissipated from the device. Other factors that must also be considered are maximum ratings, basic parameters, reliability, and types of service in which power transistors are used.

DESIGN AND FABRICATION

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 has been developed in the course of transistor evolution.

The earliest transistors made were of the **point-contact** type. In this type of structure, two pointed wires are placed next to each other on an n-type block of semiconductor material, and the p-n junctions are formed by electrical pulsing of the wires. This type has been superseded by junction transistors, which are fabricated by various alloy, diffusion, and crystal-growth techniques.

In **grown-junction** transistors, the impurity content of the semiconductor material is changed during the growth of the original crystal ingot to provide the p-n-p or n-p-n regions. The grown crystal is then sliced into a large number of small-area devices, and contacts are made to each region of the devices. Fig. 66 shows a cross section of a grown-junction transistor.

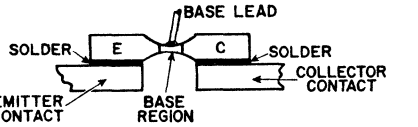


Figure 66. Cross-section of grown-junction transistor.

In **alloy-junction** transistors, two small "dots" of a p-type or n-type impurity element are placed on opposite sides of a thin wafer of n-type or p-type semiconductor material, respectively, as shown in Fig. 67. After proper heating, the impurity "dots" alloy with the semiconductor material to form the regions for the emitter and collector junctions. The base connection in this structure is made to the original semiconductor wafer.

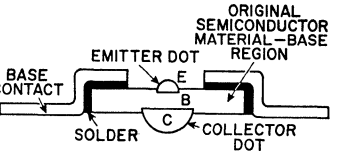


Figure 67. Cross-section of alloy-junction transistor.

The **drift-field** transistor is a modified alloy-junction device in

which the impurity concentration in the base wafer is diffused or graded, as shown in Fig. 68. Two advantages are derived from this structure: (a) the resultant built-in voltage or "drift field" speeds current flow, and (b) the ability to use a heavy impurity concentration in the vicinity of the emitter

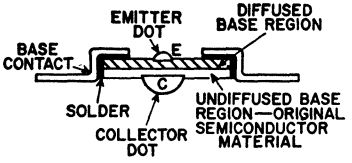


Figure 68. Cross-section of drift-field transistor.

and a light concentration in the vicinity of the collector makes it possible to minimize capacitive charging times. Both these advantages lead to a substantial extension of the frequency performance over that of the alloy-junction device.

The **diffused-junction** transistor represents a major advance in transistor technology. The increased control over junction spacings and impurity levels made possible by the diffusion technique has led to significant improvements in transistor performance capabilities.

The first diffused-junction silicon power transistor was the **single-diffused** "hometaxial" structure shown in Fig. 69. The hometaxial transistor is fabricated

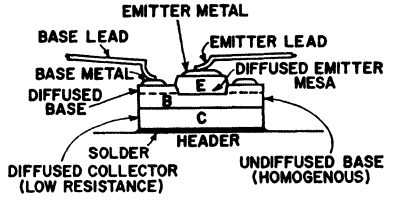


Figure 69. Cross-section of hometaxial-base silicon power transistor.

by a simultaneous diffusion of impurity from each side of a homogeneously doped slice of silicon. A **mesa** is etched on one side of the slice in an intricate design to define the transistor emitter and expose the base region to facilitate application of metal contacts to the semiconductor. Large amounts of heat can be dissipated from a hometaxial structure through the highly conductive solder joint between the semiconductor material and the device header. This structure results in a high-performance, rugged power transistor that has a very low collector resistance.

The **double-diffused** transistor provides a structure which has an additional degree of freedom for selection of the impurity levels and junction spacings of the base, emitter, and collector. This structure makes possible high voltage capability through a lightly doped collector region without compromise of the junction spacings which determine device frequency response and other important characteristics. Fig. 70 shows a typical double-diffused silicon transistor. In this type of transistor, the emitter and base junctions are formed on the same side of the silicon slice by photolithographic and silicon-dioxide mask-

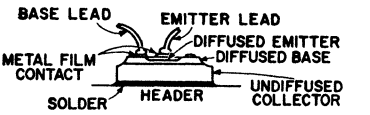


Figure 70. Cross-section of double-diffused silicon power transistor.

ing and solid-state diffusion. A mesa is usually etched through the base region to reduce the collector area at the base-to-collector junction and to provide a rugged, stable semiconductor surface.

The double-diffused silicon **planar** transistor provides the inherent advantages of the double-diffused design together with the added advantage of protection or passivation of the emitter-to-base and collector-to-base junction surfaces. Fig. 71 shows a typical double-diffused silicon planar transistor. The base and emitter regions terminate at the top surface of the semiconductor slice under

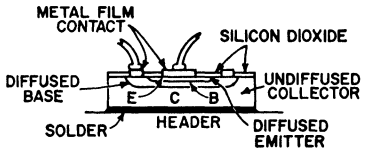


Figure 71. Cross-section of double-diffused silicon planar power transistor.

the protection of a layer of silicon dioxide. Photolithographic techniques and silicon-dioxide masking are used to provide for diffusion of both base and emitter impurities in selective areas of the silicon slice.

The **triple-diffused** transistor structure provides the advantages of the double-diffused design without the disadvantage of high collector resistance. This structure has a heavily doped region diffused from the bottom of the silicon slice which effectively reduces the thickness of the lightly doped collector region to a value dictated only by electric-field considerations. This design thus minimizes the thickness of the lightly doped or high-resistivity portion of the collector to obtain a low collector resistance. A low collector resistance is a particularly important advantage in high-current applications. Fig. 72 shows a section of a triple-diffused planar structure. A triple-diffused mesa structure could also be fabricated.

The **epitaxial** double-diffused transistor structure also has a low collector resistance. The structure and characteristics of this type of transistor are similar to those of the triple-diffused transistor shown in Fig. 72. The epitaxial structure differs fundamentally only in the way the collector region is fabricated. The lightly

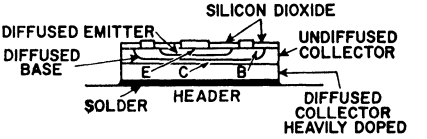


Figure 72. Cross-section of triple-diffused silicon planar power transistor.

doped collector portion of the epitaxial structure is grown on top of a heavily doped silicon slice in a high-temperature reaction chamber. This growth proceeds atom by atom and is a perfect extension of the crystal lattice of the heavily doped silicon slice on which it is grown. In contrast, the triple-diffused structure starts with a lightly doped region and has a heavily doped region diffused into a portion of it. Both techniques provide the low collector resistance required for high-current or high-power circuit applications.

The epitaxial-base transistor has a structure in which a lightly doped base region is deposited by epitaxial techniques on a heavily doped silicon slice of opposite-type dopant. Photolithographic and silicon-dioxide masking and a single impurity diffusion are used to define the emitter region, as shown in Fig. 73. This structure offers the advantage of low collector resistance and ease of controlling

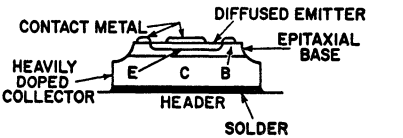


Figure 73. Cross-section of epitaxial-base silicon power transistor.

impurity spacings and emitter geometry for high-current and moderate-frequency performance. A variation of this structure uses two epitaxial layers. A thin lightly doped epitaxial layer, which is used for the collector, is deposited over the heavily doped silicon slice prior to the epitaxial deposition of the base region. The collector epitaxial layer is of opposite-type dopant to the epitaxial base layer. This structure, shown in Fig. 74, has the advantages of the epitaxial-base device, with an added advantage of higher voltage ratings provided by the epitaxial collector layer.

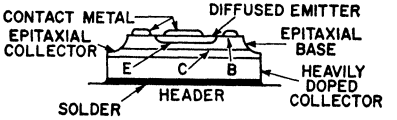


Figure 74. Cross-section of dual-epitaxial-layer silicon power transistor.

The **overlay** transistor is a double-diffused epitaxial device which employs a unique emitter structure. In this structure, a large number of separate emitters are tied together by diffused and metalized regions, as shown in Fig. 75. This design concept increases the emitter edge-to-area ratio and reduces the charging time constants without compromise of the transistor current and power-handling capability. The overlay transistor is fabricated by exceptionally well controlled diffusions and very precise photolithographic and silicon-dioxide masking operation. Fig. 76 shows a section through a typical overlay emitter region.

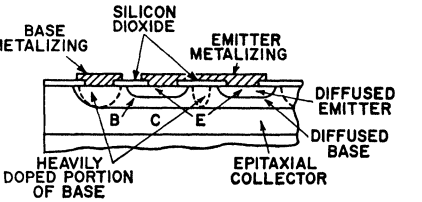


Figure 76. Cross-section of overlay transistor.

Power-transistor designs differ fundamentally from signal-level transistors in the way that the semiconductor element is packaged to provide for high thermal conductivity and low-resistance electrical contacts. The power semiconductor element is usually soldered or gold-alloyed to a solid metal header, as shown in Fig. 77. For the high-power types, the header is generally constructed from copper or laminated copper and steel for improved heat transfer. Low-resistance contacts are soldered or metal-bonded from the

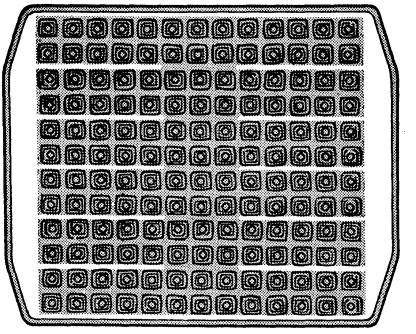


Figure 75. Emitter structure of an overlay transistor.

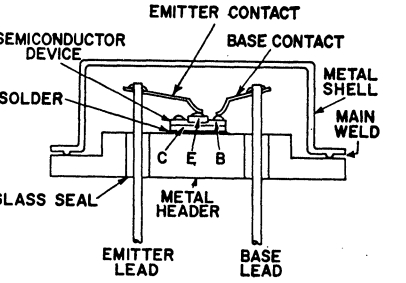


Figure 77. Assembly details for high-power silicon transistor.

emitter or base metalizing contacts to the appropriate package leads. This packaging concept results in a simple structure that can be readily attached to a variety of circuit heat sinks and that can safely withstand power dissipations of hundreds of watts and currents of tens of amperes.

A few high-performance power transistors are packaged with electrical isolation of the collector. This isolation is achieved, without compromise of thermal dissipation, by gold-alloying of the semiconductor element to a metalized ceramic disc. This disc, which is usually beryllium oxide, is brazed to the package header to provide a low thermal-resistance path to the circuit heat sink.

BASIC TRANSISTOR PARAMETERS

A transistor is usually employed to obtain a power gain by use of a small control signal to produce larger signal variations in the output current. In a vacuum tube, the most common gain parameter is the voltage amplification factor (μ) from the grid to the plate. In a transistor, the most commonly specified gain parameter is the current gain (β) from the base to the collector. Power gain of a transistor operated in the common-emitter circuit configuration is proportional to the square of the current gain multiplied by the load resistance divided by the input resistance, as indicated in Fig. 78.

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:

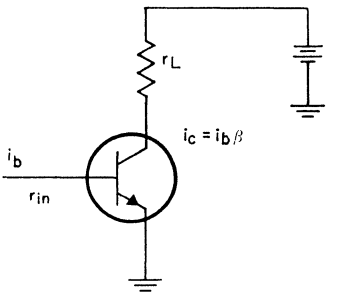
1. beta (β)—general term for current gain from base to collector.

2. alpha (α)—general term for current gain from emitter to collector.

3. h_{fe}—ac gain from base to collector.

4. h_{FE}—dc gain from base to collector.

The input impedance r_{in} of a transistor also affects the power output, as indicated by the equations shown in Fig. 78. The input impedance is not usually specified directly because of the large number of components of which it is comprised, but is usually specified as a maximum base-to-emitter voltage V_{be} under specified input-current conditions.



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

Other low-frequency electrical characteristics commonly specified for transistors are those needed to verify the maximum ratings and the leakage currents. Leakage currents are important because they affect biasing in amplifier applications and represent the "OFF" condition for transistors used in switching applications. Several different leakage currents are commonly specified. The most basic specification is I_{CBO}, which indicates the leakage from collector to base with the emitter open. This leakage, which is simply the reverse current of the collector-to-base diode, is composed of two components, a saturation current that doubles in value for approximately every 8°C increase in temperature and a surface-leakage component that is not directly related to temperature. In a silicon transistor, the saturation current is normally small; at room temperature, only the surface leakage is measurable. High-temperature leakage currents are usually specified in the published data for a transistor.

In addition to the I_{CBO} ratings, I_{CEV}, I_{CEO}, and I_{CER} ratings are often specified for transistors. I_{CEV} is the leakage from the collector to emitter with the base-to-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}.

When transistors are used at higher frequencies, their gain decreases. This condition is discussed more completely in the section on **High-Frequency Power Amplifiers.** The h_{fe} decreases with frequency in a predictable way, as shown in Fig. 79.

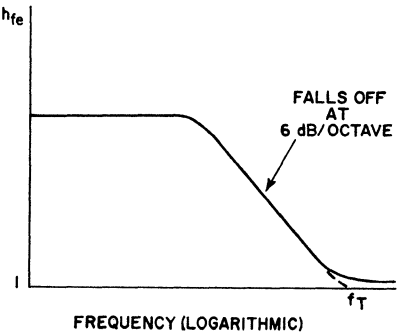


Figure 79. Transistor current-gain parameter h_{fe} as a function of frequency.

Because of the regular decrease at high frequencies (6 dB per octave), a measurement of the gain at any frequency on the 6-dB-per-octave slope multiplied by the frequency at which beta is measured results in approximately the same value. This value, called the **gain-bandwidth product** f_T, is indicative of the high-frequency capability of a transistor. Other parameters which affect high-frequency performance are the capacitance or resistance which shunts the load and the input impedance, the effect of which is shown by the equations of Fig. 78.

The specification of all the characteristics which affect high-frequency performance is so complex

that often a manufacturer does not specify all the parameters but instead specifies transistor performance in an rf-amplifier circuit. This information is very useful when the transistor is operated under conditions very similar to those of the test circuit, but is difficult to apply when the transistor is used in a widely different application. Some manufacturers also specify transistor performance characterisetics as a function of frequency, which alleviates these problems.

The power-dissipation capability of a transistor is usually given by a dissipation derating curve, a thermal-resistance specification, and a maximum junction temperature, or a safe-area curve. **Thermal resistance** is the increase in temperature of the junction of a transistor (with respect to some reference) divided by the power dissipated. In power transistors, the thermal resistance is normally measured from the pellet of the transistor to the case. The user is then required to determine the remainder of the heat-flow path.

Thermal resistance is usually measured by operation of the transistor in a pulsed type of service in which power is dissipated in the transistor during most of the operating cycle. During the remainder of the cycle, a temperature-sensitive parameter of the transistor is monitored to determine the junction temperature. The parameter usually monitored to determine junction temperature is the forward-voltage drop across the base-to-emitter diode junction. For optimum accuracy, the thermal resistance of each transistor must be determined individually. The test, therefore, is very slow and expensive. There is an additional disadvantage to thermal-resistance testing in that the measurement is based on the average junction temperature and does not indicate the maximum temperature attained within the device if temperature distribution across the pellet is not uniform.

Safe-area curves are normally verified with a power rating test. In this type of test, the transistor is subjected to a high-level pulse of power (usually equal to the power ratings of the transistor at 25°C) for a period of time that is long compared to the transistor internal time constant, but short compared to the thermal time constant of the case. The voltage across the transistor is monitored. A failing unit momentarily shorts because of localized overheating. During the test, this overheating is detected by a sensing circuit, and power is removed. In an actual application, the transistor would be destroyed. A safe-area rating curve indicates whether a transistor can withstand a given power level without excessive localized overheating.

MAXIMUM 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 storage temperature, 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 emitter-to-base 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 reverse 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 reverse 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.

Current Ratings

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.

Storage and Operating Temperature Ratings

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 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 discussed in the following sections.

THERMAL CONSIDERATIONS

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 (power) is 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.

Thermal Resistance

The methods of rating power transistors under steady-state conditions are embodied in the following definition of thermal resistance: The thermal resistance of a semiconductor device is the quotient of the temperature drop and the heat generated through internal power dissipation **under steady-state conditions,** the temperature drop being measured between the region of heat generation and some reference point.

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 design is unnecessarily conservative and expensive. In the next section, the problem of transient thermal response is investigated and the concept of thermal capacitance is introduced.

Junction-to-Case Thermal Impedance—The heat-flow problem in a transistor may be analyzed in terms of the simple electrical analog shown in Fig. 80. The model uses an energy-storage element C, which introduces the concept of thermal capacitance, to explain the transient thermal properties of transistors. Although this model may be conveniently used to predict the rise of junction temperature that results from a unit step or pulse input of power, the two-element equivalent circuit is an extreme over-simplification.

For example, in the double-diffused epitaxial planar transistor shown in Fig. 81, the major

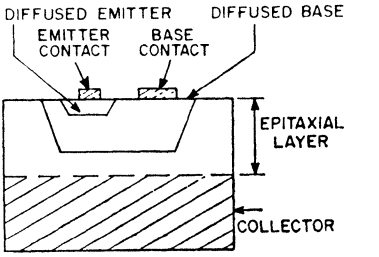
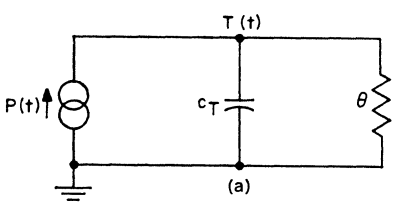


Figure 81. Structure of double-diffused epitaxial planar transistor.

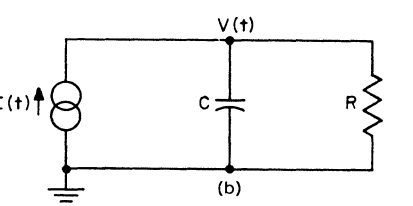
physical source of heat is the collector-to-base junction. Fig. 82 shows the thermal equivalent circuit for this transistor. In general, the thermal-resistance elements shown to the left of the junction are so large that the power flow through this path is negligible. In addition, the heat energy introduced by the bulk-resistance terms for moderate current levels may be considered small in comparison to the heat injected at the collector-to-base junction.

If these simplifications are made, the thermal equivalent circuit may be reduced to the form



THERMAL CIRCUIT

ELEMENT	SYMBOL	UNITS
TIME	t	SECONDS
HEAT FLOW = POWER	P(t)	WATTS = JOULES/SECOND
TEMPERATURE	T(t)	°C
THERMAL RESISTANCE	θ	°C/WATT
THERMAL CAPACITANCE	C_T	WATT-SECONDS/°C



ELECTRICAL EQUIVALENT

ELEMENT	SYMBOL	UNITS
TIME	t	SECONDS
CURRENT	I(t)	AMPERES
VOLTAGE	V(t)	VOLTS
RESISTANCE	R	OHMS
CAPACITANCE	C	FARADS

Figure 80. Electrical analog representation of the heat-flow path of a transistor: (a) thermal circuit; (b) electrical equivalent of thermal circuit.

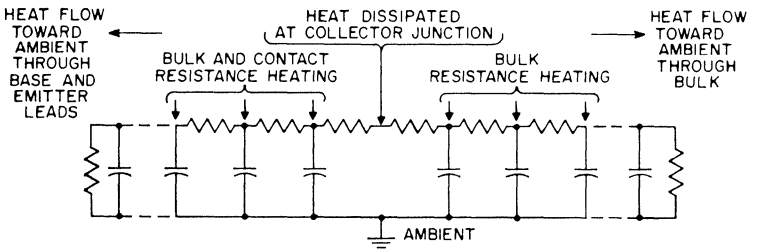


Figure 82. Thermal equivalent circuit for transistor structure shown in Fig. 81.

shown in Fig. 83. It should be noted that the elements of this circuit are complex quantities which are dependent upon the operating conditions of the transistor. At high current levels, current-crowding effects cause the current-flow paths from emitter to collector junction to be modified. The modified current-flow paths tend to concentrate the heat in a small portion of the collector junction area and thus to create hot spots. As these conditions develop, the thermal equivalent circuit must also be modified to account for the restricted heat-flow paths. It should be obvious from this discussion that a complete thermal equivalent circuit is very complex.

However, the response of this complex system to a given input can be determined as follows: If the model is restricted to sufficiently low currents and the dependence of the thermal resistance on the operating point is neglected, the simple equivalent shown in Fig. 84 may be used

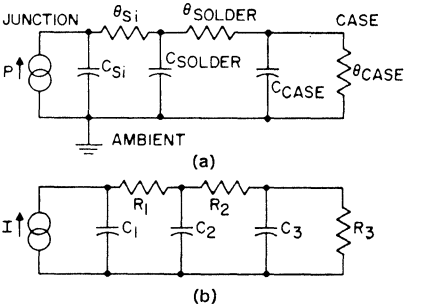


Figure 84. Simplified equivalent circuits for a small-signal medium-power transistor: (a) thermal circuit; (b) electrical analog of thermal circuit.

for a small-signal medium-power transistor. Representative values for the thermal parameters in a typical transistor are as follows:

$$\Theta_{Si} = 15\ ^\circ C/W$$
$$\Theta_{solder} = 0.4\ ^\circ C/W$$
$$\Theta_{case} = 19.6\ ^\circ C/W$$
$$C_{Si} = 0.18 \times 10^{-3}\ (W\text{-sec})/^\circ C$$
$$C_{solder} = 0.04 \times 10^{-3}\ (W\text{-sec})/^\circ C$$
$$C_{case} = 285 \times 10^{-3}\ (W\text{-sec})/^\circ C$$

Because of the difference in the magnitude of these elements, the equivalent circuit may be further simplified in the time domain.

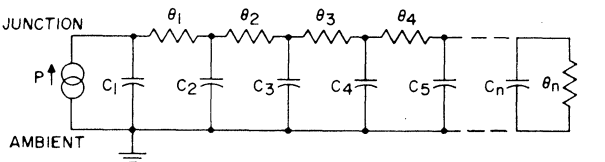


Figure 83. Simplified equivalent circuit for transistor structure shown in Fig. 81.

For very short times ($t << \theta_{case} \times C_{case}$), the thermal impedance to the right of C_{solder} in Fig. 84 is approximately an open circuit, and the equivalent circuit appears as shown in Fig. 85. The electrical

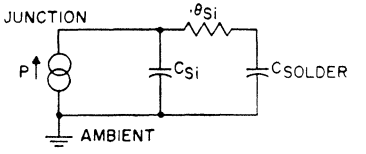


Figure 85. Simplified thermal equivalent circuit for a transistor when the duration of applied voltage and current is very short (i.e., $t << \theta_{case}\ C_{case}$).

analog of such a circuit can be used to show that the response to a step input of power P is given by the following expression:

$$T_J(t) = \frac{P}{(C_{Si} + C_{solder})} \left[\theta_{Si} (C_{solder} - C_E) \left(1 - \exp \frac{-t}{\theta_{Si} C_E} + t \right) \right] + T_{amb} \quad (10)$$

$$C_E = \frac{C_{Si}\, C_{solder}}{C_{Si} + C_{solder}} \quad (11)$$

where T_J is the junction temperature, and the other parameters are evident from Fig. 85.

Eq. (10) shows that initially the junction temperature rises faster than an exponential because of the linearly increasing t term. For time durations greater than five thermal time constants ($t \geq 5\ \theta_{Si}C_E$), the exponential term $\exp(-t/\theta_{Si}C_E)$ approaches zero, and the junction temperature rises linearly.

If the response is analyzed a long time after the application of the step of power, some definite statements can be made. All of the thermal capacitors are charged, and the junction temperature is determined as follows:

$$T_J(t) = T_J = P(\Theta_{Si} + \Theta_{solder} + \Theta_{case}) + T_{amb} \quad (12)$$

There is no time dependence implicit in this equation, which is identical to the solution for the steady-state condition.

At intermediate times such that $\theta_{Si}C_E << t < \theta_{case}\ C_{case}$, the effective equivalent circuit can be represented as shown in Fig. 86. The electrical analog can again be used to show that the response

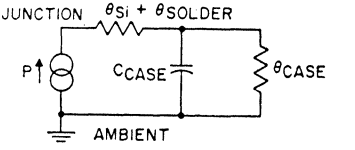


Figure 86. Simplified thermal equivalent circuit for a transistor for $\theta_{Si}\ C_E << t < \theta_{case}\ C_{case}$.

of this circuit to a step input of power, under these conditions, is given by the following equation:

$$T_J(t) = [P(\Theta_{Si} + \Theta_{solder}) \times U_1(t)] + P\,\Theta_{case} \times [1 - \exp(-t/\Theta_{case}\, C_{case})] \quad (13)$$

where the unit step function $U_1(t)$ is zero for $t < 0$ and unity for $t > 0$.

Eq. (13) indicates that when $\theta_{Si}C_E << t < \theta_{case}\ C_{case}$, the junction temperature rises exponentially from a constant value, and reaches another constant value in approximately five time constants.

The response of the circuit in Fig. 84 has been predicted during two transient periods for steady-state conditions. The response must then be determined for intermediate times between the two transient responses given by Eqs. (10) and (13). Because it is assumed that the analog model

contains only linear elements, superposition may be used. The response of the circuit in Fig. 84, therefore, may be determined by addition of the responses defined by Eqs. (10) and (13) with appropriate summing factors. This summation can be performed graphically; the results of such an analysis based upon the numerical values given previously are shown in Fig. 87. The ordinate value is the thermal impedance, or transient thermal resistance

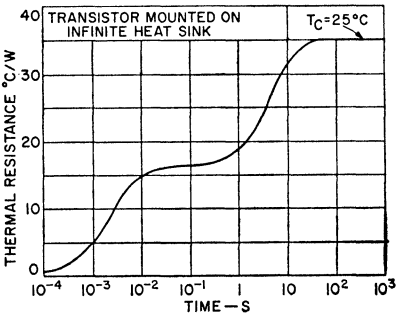


Figure 87. Graphical representation of transient thermal resistance curve.

θ_T. The value of θ_T is determined by the pulse width and may be used to calculate the peak pulse power P_p, as follows:

$$P_p = \frac{T_J\,(max) - T_{amb}}{\Theta_T\,(T_p)} \qquad (14)$$

where T_p is the pulse width.

The heating and cooling curves are conjugates, and may be related by the following equation:

$$T_c = \Delta T_n - T_n = PR_T - T_h \qquad (15)$$

where T_c is the cooling response and T_n is the heating response. It is apparent, therefore, that the θ_T for a given device can be determined if the junction temperature is monitored as a function of time after the removal of a steady-state load.

In a power transistor, it is often desirable to mount the silicon pellet on a copper pedestal so that the heat spreads over a wider area before it reaches the steel case. This dispersal of heat introduces two more elements in the equivalent circuit to account for the thermal resistance and capacitance of the copper pedestal. In practice, θ_{Cu} is very small and C_{Cu} is comparable to C_{case}. As a first approximation, θ_{Cu} may be neglected, and C_{Cu} and C_{case} can be lumped together. The equivalent circuit is then shown in Fig. 84, and the thermal response is shown in Fig. 87.

In this analysis, a simple thermal equivalent circuit has been used to predict the junction-temperature response to a suddenly applied pulse of power. Although the elements of the equivalent circuit are dependent upon the operating conditions, this dependence has been neglected to simplify the problem to obtain some insight to the response. Fig. 87 does not present an actual transient thermal response of a transistor, but is a graphical presentation of the result obtained from the application of a step function of input power to the approximate model developed.

Thermal Impedances External to the Transistor—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 IV 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

Table IV—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

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.

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. 88. From the electrical analog, the effective resistance of the two parallel thermal paths is smaller than

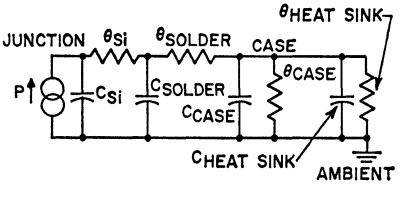


Figure 88. Thermal equivalent circuit for a transistor mounted on a heat sink.

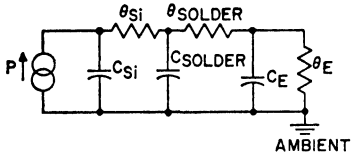
that of either of the paths. The effect of the heat sink is to provide an additional low-thermal-resistance path from case to ambient air. The heat-sink thermal resistance actually consists of two series elements, the thermal resistance from case to heat sink that results from conduction ($\theta_{C\text{-}S}$) 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 transistor, 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 transistor 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 θ_{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. 89.

The form of the equivalent circuits shown in Figs. 84 and 89 is the same, and their response characteristics are similar.



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

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

Figure 89. 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:

$$\Theta_{cond} = t/4.186\ kA\ ^\circ C \text{ per watt} \quad (16)$$

where t is the length of the thermal path in centimeters, k is the thermal conductivity in CAL/(sec) (cm) (°C), A is the area perpendicular to the thermal path t in square centimeters, and the conversion factor 4.186 is given in (watt) (sec)/CAL.

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:

$$\Theta_{conv} = (2300/A)\,(L/T_s - T_{amb})^{\frac{1}{4}} \quad (17)$$

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:

$$\Theta_{rad} = \frac{1793 \times 10^8}{AE\,(T_s^2 - T_{amb}^2)}\,(T_s - T_{amb}) \quad (18)$$

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.

Selection and Use of External Heat Sinks

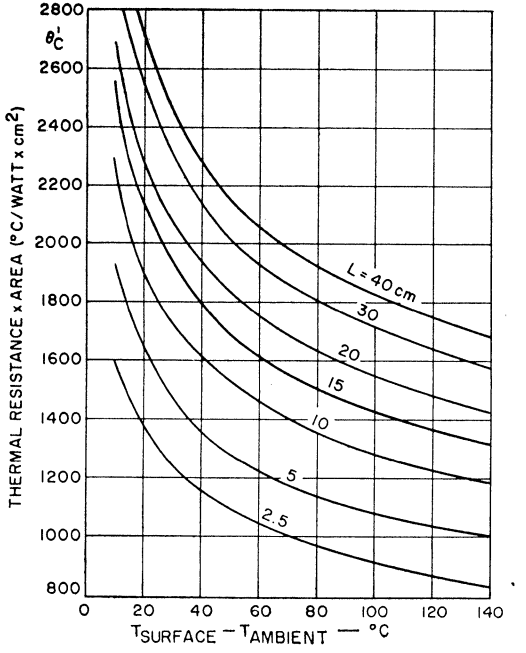
The sources of thermal resistance both internal and external to the transistor have been discussed, and the processes which may be used for heat removal have been explained briefly. A power transistor is normally designed to be used with an external heat sink. This section discusses the factors involved in the selection of an external heat sink.

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. 90, 91, and 92.

A useful nomograph which considers heat removal by both convection and radiation is given in Fig. 93. This nomograph applies for natural bright finish on the copper or aluminum.



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

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

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

Figure 90. 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.)

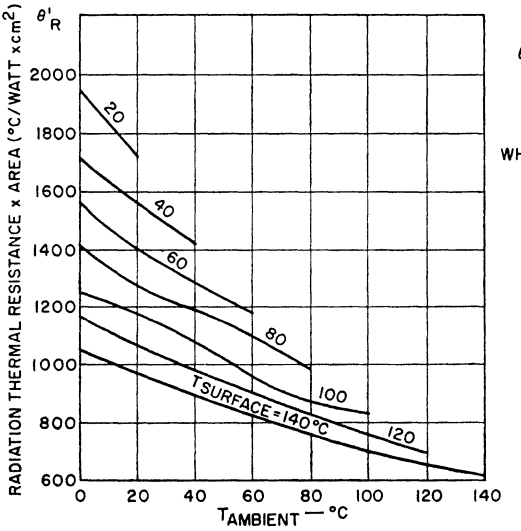


Figure 91. Radiation thermal resistance as a function of ambient temperature for various heat-sink surface temperatures. (Reprinted from **Control Engineering,** October 1956.)

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. 90, 91, and 92 provides the basis for the design of flat vertical plates for use as heat 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. 92. 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

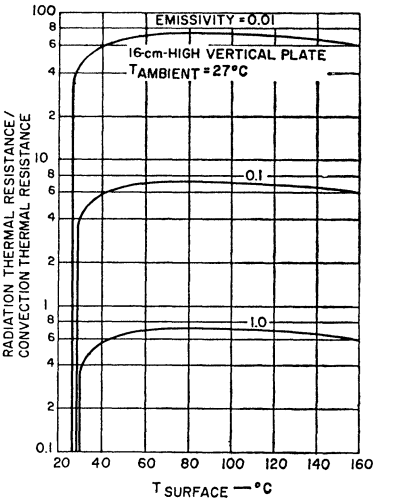
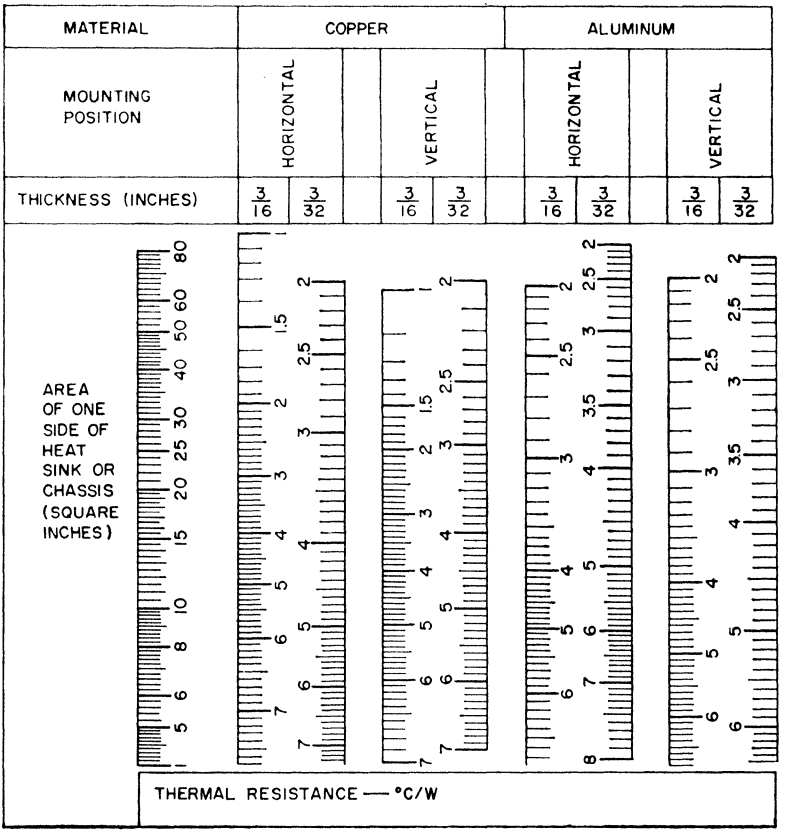


Figure 92. Ratio of radiation thermal resistance to convection thermal resistance as a function of heat-sink surface temperature for various surface emissivities. (Reprinted from **Control Engineering,** October 1967.)



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

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.

Selection and Use of Insulators

As pointed out previously, when transistors 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 V. 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

Table V—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

$\theta_{C\text{-}S}$ can be readily calculated by use of Eq. (16). 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.

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. 94 illustrates both types of mounting.

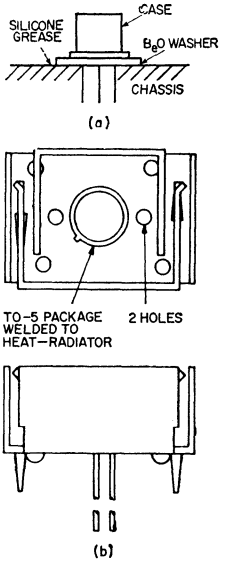


Figure 94. Suggested mounting arrangements for transistors having a JEDEC TO-5 package: (a) without heat sink; (b) with fin-type heat sink.

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.

Effect of Thermal Factors on Dissipation Capability

The effects of the heat sink and the insulating washer between case and heat sink can be clarified by sample calculations of the maximum allowable power dissipation in a transistor. In these calculations, 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:

$$\Theta_{J-A} = \Theta_{J-C} + \Theta_{C-S} + \Theta_{S-A} \quad (19)$$

The maximum value of the junction-to-case thermal resistance $\theta_{J\text{-}C}$ 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:

$$P_{ss}(max) = \frac{T_J(max) - T_{amb}}{\Theta_{J-A}} \quad (20)$$
$$= (200-50)/4.5$$
$$= 33.3 \text{ watts}$$

The case temperature of the 2N3055 transistor that results from the maximum steady-state dissipation is calculated from the following equation:

$$T_C = P_{ss}(\Theta_{C-A}) + T_{amb} \quad (21)$$

The term θ_{C-A} represents the total thermal resistance from case to ambient air (i.e., $\theta_{C\text{-}A} = \theta_{C\text{-}S} + \theta_{S\text{-}A} = 3$°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. 95, in the form of a normalized power multiplier M. For a 1-millisecond

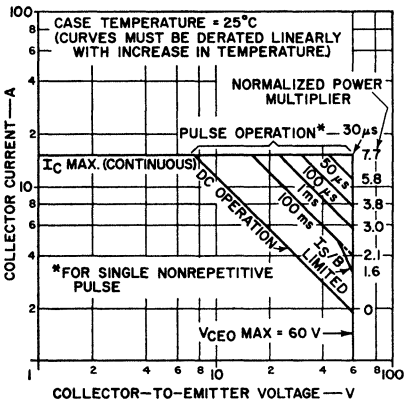


Figure 95. Maximum-operating-area curves for the 2N3055 silicon power transistor.

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:

$$TDF = 1 - \frac{T_C - 25°C}{T_J\ (max) - 25°C} \quad (22)$$

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:

$$P_{sp} = M\ (TDF)\ (P_{max}) \quad (23)$$

$$= M\left(1 - \frac{T_C - 25}{T_J\ (max) - 25}\right)\ (P_{max})$$

$$= 3\left(1 - \frac{50 - 25}{200 - 25}\right)\ (115)$$

$$= 296 \text{ watts}$$

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:

$$P_{avg} = P_{pk}(d) \quad (24)$$

where P_{pk} is the peak pulse power dissipated in the transistor and d is the duty cycle.

The effective case temperature that results from this average dissipation is determined from the following expression:

$$T_C(\text{eff}) = T_{amb} + P_{avg}\,\theta_{J-A} = T_C + P_{avg}\,\theta_{J-C} \quad (25)$$

Substitution of Eq. (24) into Eq. (25) yields the following result:

$$T_C(\text{eff}) = T_{amb} + P_{pk}(d)\,\theta_{J-A} \quad (26)$$

If the effective case temperature $T_C(\text{eff})$, as defined by Eq. (26), is substituted for the case temperature T_C in Eq. (23), the following expression is obtained for the maximum allowable power dissipation P_{rp} for repetitive pulses:

$$P_{rp} = \frac{M\,[T_J(\max) - T_{amb}]\,P_{max}}{T_J(\max) - 25 + MdP_{max}\,\theta_{J-A}} \quad (27)$$

On the basis of the definition given in the section on **Thermal Resistance,** junction-to-case thermal resistance may be expressed by the following equation:

$$\theta_{J-C} = \frac{T_J(\max) - 25}{P_{max}} \quad (28)$$

If the relationship expressed by Eq. (28) is used, Eq. (27) can be simplified to the following form:

$$P_{rp} = M\left(\frac{T_J(\max) - T_{amb}}{\theta_{J-C} + M\,d\,\theta_{J-A}}\right) \quad (29)$$

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
θ_{J-C}	= 1.5°C/watt
M	= 3
θ_{J-A}	= 4.5°C/watt

When these values are substituted in Eq. (27), 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:

$$P_{avg} = P_{ss}(\max)\,(T_p/t) \quad (30)$$

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:

$$P_{tr} = \frac{T_J(\max) - T_C(\text{avg})}{\theta_{J-C}'} \quad (31)$$

where θ_{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. 96.

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. 97 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 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. 97) may be formed within the silicon pellet. If unchecked, these hot spots initiate a regenerative cycle of high-

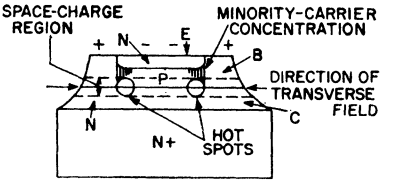


Figure 97. Cross-section of power transistor under forward-bias conditions.

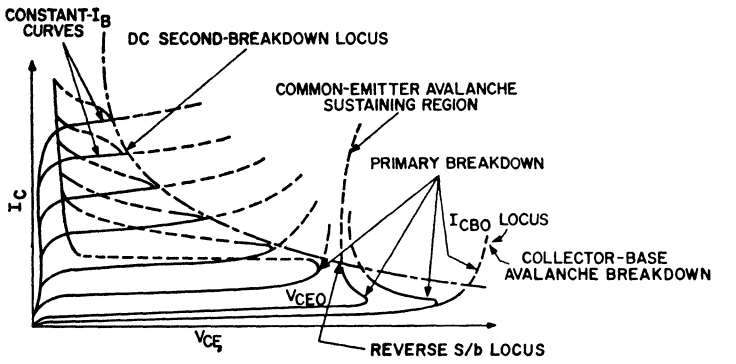


Figure 96. Transistor collector characteristics.

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.

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. 98 and 99 show variation in $I_{S/b}$ capability as a function of frequency capability f_T and collector voltage V_{CE}, respectively. The curves shown in these figures are graphical representations of the following empirical relationships:

1. Frequency relationship:

$$I_{S/b} = K_1/\sqrt{f_T} \qquad (32)$$

2. Voltage relationship:

$$I_{S/b} = K_2/V_{CE}^{n} \qquad (33)$$

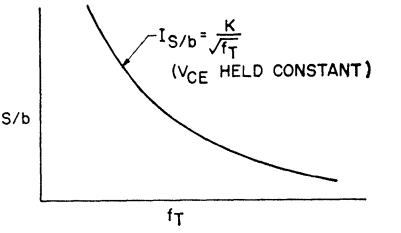


Figure 98. Variation in forward-bias second-breakdown energy level as a function of transistor frequency capability.

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. (32) 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. (33) indicates that supply voltage should be as small as possible and that high-voltage transients should be restricted as much as possible to

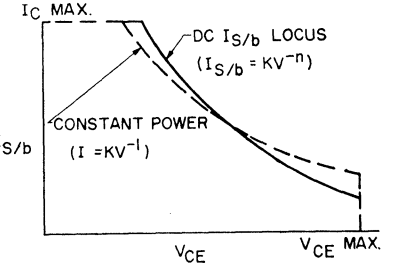


Figure 99. Variation in forward-bias second-breakdown energy level as a function of collector voltage.

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. 100 shows the variation in $I_{S/b}$ capability with pulse width.

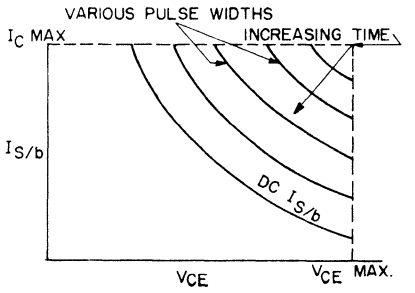


Figure 100. Variation in forward-bias second-breakdown energy level as a function of pulse width.

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. 101 shows a cross section of a typical power-transistor pellet under reverse-bias conditions.

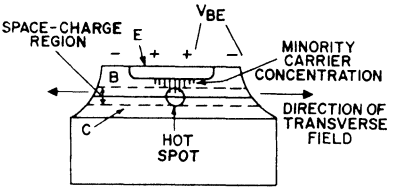


Figure 101. Cross-section of power transistor under reverse-bias conditions.

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. 102 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 in-

verse function of base currents, the circuit designer must compromise between transistor turn-off speed and second-breakdown considerations.

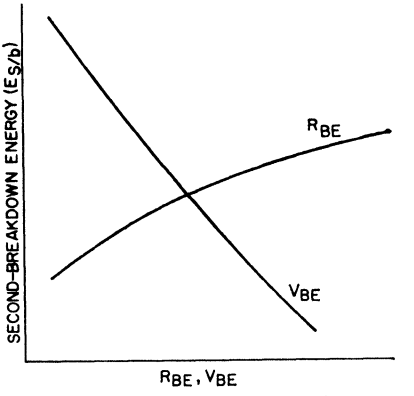


Figure 102. Variation in reverse-bias second-breakdown energy level as a function of V_{BE} and R_{BE}.

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. 103 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 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

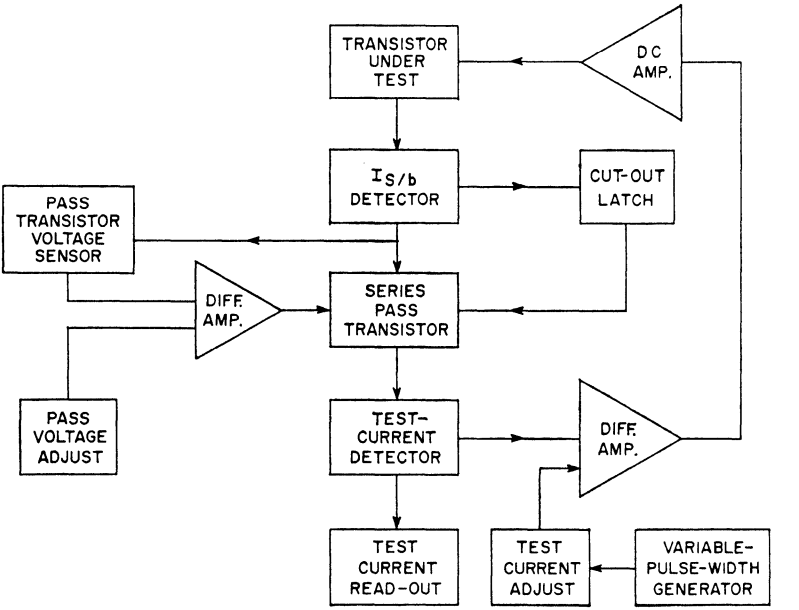


Figure 103. Block diagram of typical $I_{S/b}$ test set.

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. 104 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. If the drop across R_E is neglected, the energy absorbed by the transistor is given by the following equation:

$$E_C = C\,V_C^2/2 \qquad (34)$$

Current and voltage waveforms of the 2N3442 silicon power transistor tested in the capacitance

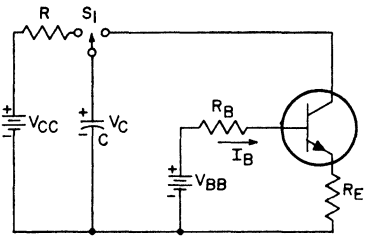


Figure 104. Capacitive-discharge test circuit used to determine forward-bias second-breakdown energy levels in transistors.

discharge test set are shown in Fig. 105.

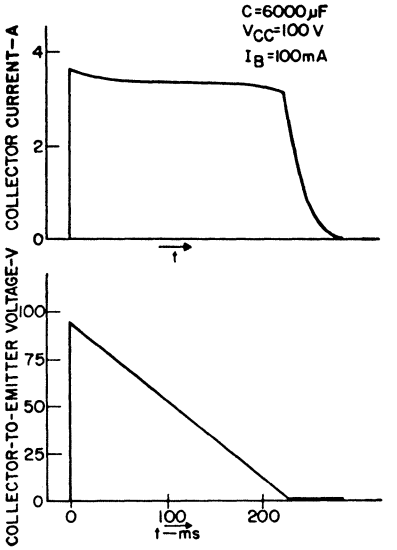


Figure 105. Test waveforms for a 2N3442 transistor subjected to capacitive-discharge second-breakdown energy test.

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:

$$t_{off} \ll L\, I_C(max)/V_{CEX}\,(sus) \quad (35)$$

The energy that the transistor absorbs during the $E_{S/b}$ test can be determined from the following equation:

$$E_{S/b} = \frac{L\, I_C(max)^2}{2} \left[\frac{V_{CEX}\,(sus)}{V_{CEX}(sus) - V_{CC}}\right] \quad (36)$$

Voltage and current waveforms for a 2N3442 silicon power transistor subjected to the $E_{S/b}$ test are shown in Fig. 106.

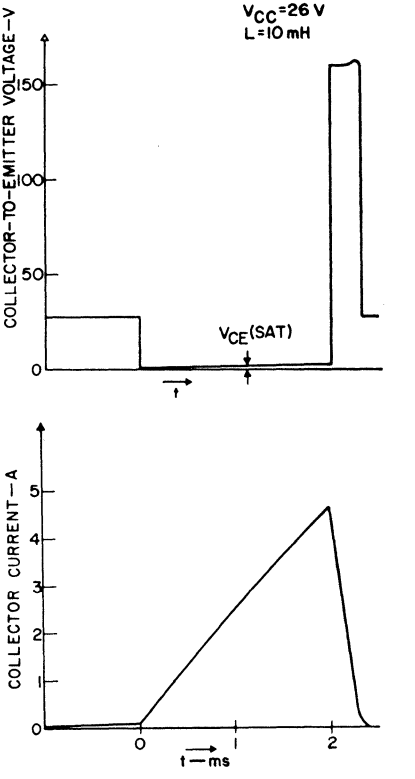


Figure 106. Test waveforms for a 2N3442 transistor subjected to $E_{S/b}$ inductive energy test.

Fig. 107 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-

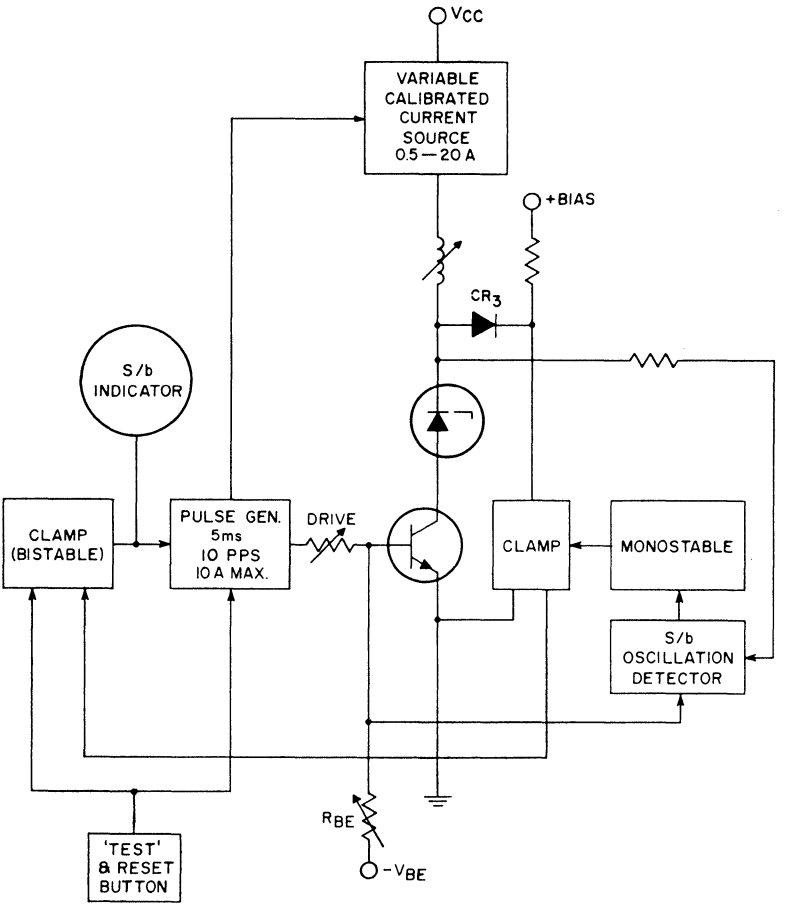


Figure 107. Block diagram of $E_{S/b}$ test set.

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 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. 108.

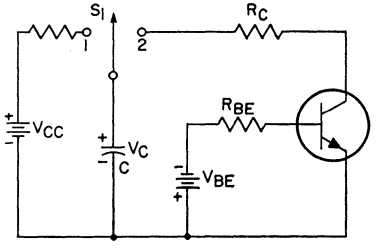


Figure 108. Capacitive-discharge test circuit used to determine reverse-bias second-breakdown energy levels in transistors.

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:

$$E_{C_{S/b}} = \int_0^\infty (V_{CEX}/R)\,(V_{CC} - V_{CEX})\,\epsilon^{-t/RC}dt \qquad (37)$$

$$= C\,[V_{CEX}(sus)\,V_{CC} - V_{CEX}(sus)^2] \qquad (38)$$

These equations assume that the $V_{CEX}(sus)$ breakdown value of the transistor does not change with current. Variations in $V_{CEX}(sus)$ 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. (34) 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. 109 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 horizontal regions of the curve

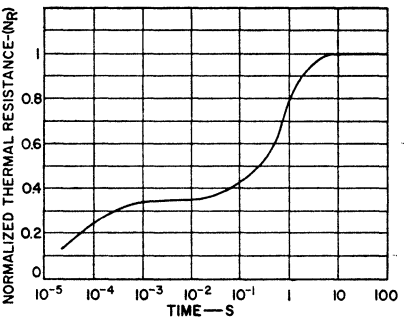


Figure 109. Normalized thermal resistance as a function of the duration of applied power for the 2N3442 silicon power transistor.

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 Θ_{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:

$$P_{diss} = \frac{T_J(max) - T_C}{\Theta_{J-C}(N_R)} \qquad (39)$$

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. 110 shows a forward-bias safe-area rating chart for a typical high-speed silicon power transistor, RCA-2N3585, which has a gain-bandwidth product f_T in the

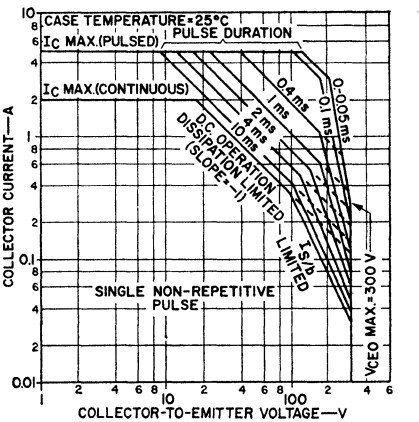


Figure 110. Safe-area rating chart for the 2N3585 silicon power transistor.

order of 20 MHz. Fig. 111 shows a similar rating chart for a lower-frequency silicon power transistor, 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

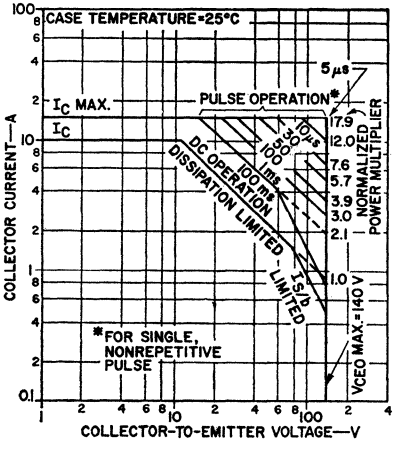


Figure 111. Safe-area rating chart for the 2N3442 silicon power transistor.

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 $V_{CEO}(sus)$], and the thermal and second-breakdown ratings of the transistors.

As shown in Fig. 110, 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 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. 110 and 111 shows that the second-breakdown ratings of the 2N3442 are substantially higher than those of the 2N3585, as can be predicted from Eq. (32) 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. 110 and 111, safe-area ratings are normally given for single nonrepetitive 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. 112 shows temperature derating for the 2N3585 safe-area chart of Fig. 110. These curves show that thermal ratings are affected far more by increases in

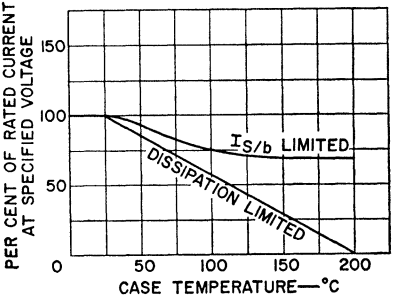


Figure 112. Safe-area temperature-derating curves for the 2N3585 silicon power transistor.

case temperature than are second-breakdown ratings. The thermal (dissipation-limited) derating curve is a graphic representation

of Eq. (22) for the temperature derating factor given in the section on **Thermal Considerations.** 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 **Thermal Considerations.** 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 θ_{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—A typical high-speed, high-voltage, 100-watt, two-transformer inverter is shown in Fig. 113. Fig. 114 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,

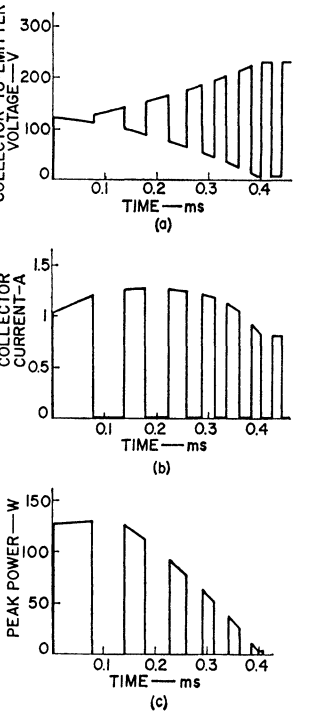


Figure 114. Waveforms for inverter circuit shown in Fig. 113.

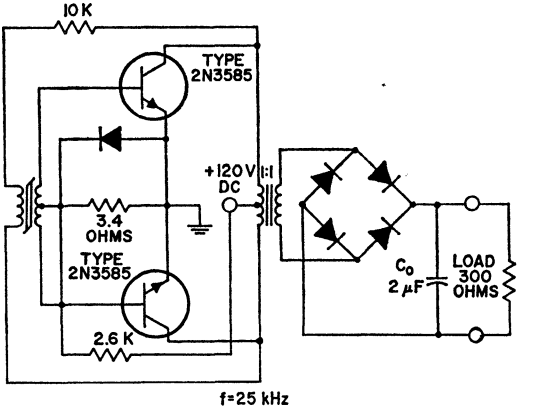


Figure 113. High-speed inverter using RCA 2N3585 transistors.

as shown in Fig. 115. Because the circuit being analyzed is a switching circuit, each pulse has a different load line associated

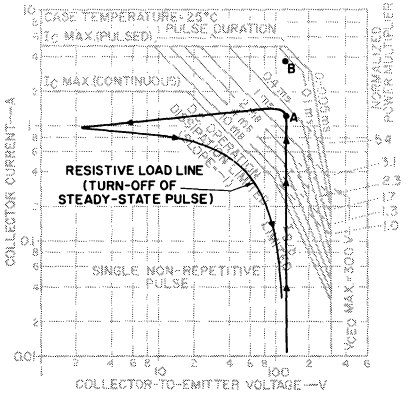


Figure 115. Inverter turn-on load line on 2N3585 safe-area curve.

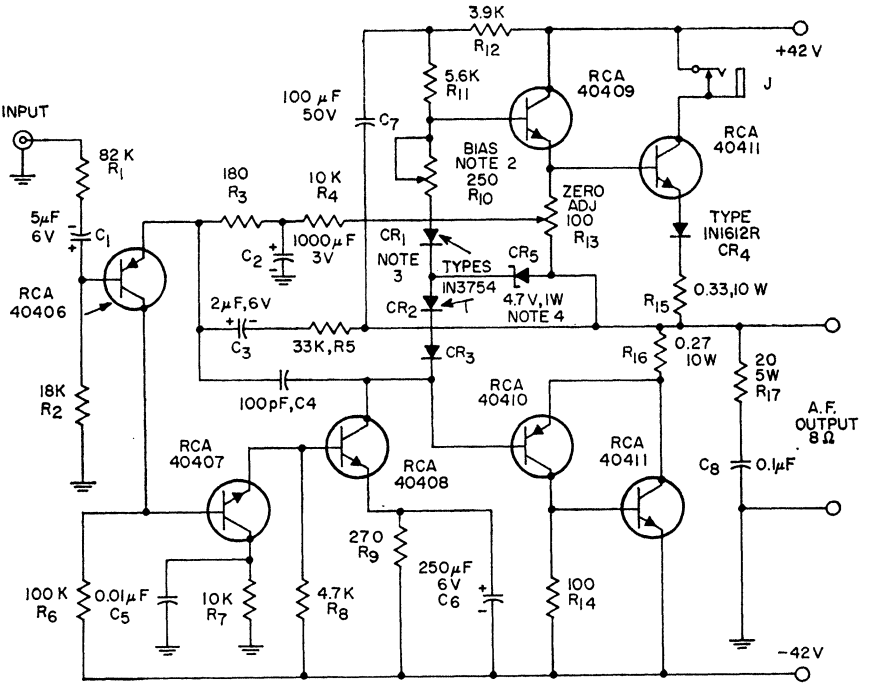
with it. Initially, the pulses follow a highly capacitive load line, but become more resistive as the output capacitor charges. This change of pulse character presents no problem in this analysis because it is being performed for initial turn-on only. For this reason, the designer need only plot the turn-on load line of the first pulse and the locus of peak voltage and current of the remaining pulses, as shown in Fig. 115. 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. 114(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. 115. 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. 112 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



NOTE 1: RESISTORS ARE 1/2-W TYPES UNLESS OTHERWISE SPECIFIED.
NOTE 2: SET BIAS FOR 20 mA QUIESCENT CURRENT (MEASURED AT J WITHOUT LOAD).
NOTE 3: THERMALLY CONNECTED TO HEAT SINK FOR OUTPUT TRANSISTORS.
NOTE 4: ZENER DIODE - IN1519.

Figure 116. 70-watt, silicon-transistor audio amplifier.

power amplifier is shown in Fig. 116. This amplifier is terminated in a resistance-inductance series 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. 117.

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. 118. 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. 117(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. (25), 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.

The temperature derating curves for the 40411 are given in Fig. 119. 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. 119. 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. 118 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

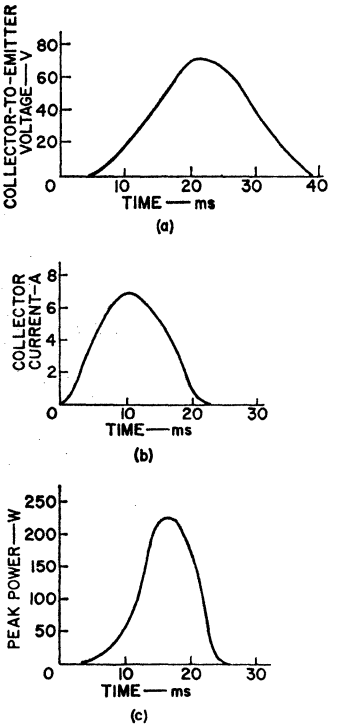


Figure 117. Audio power-amplifier waveforms.

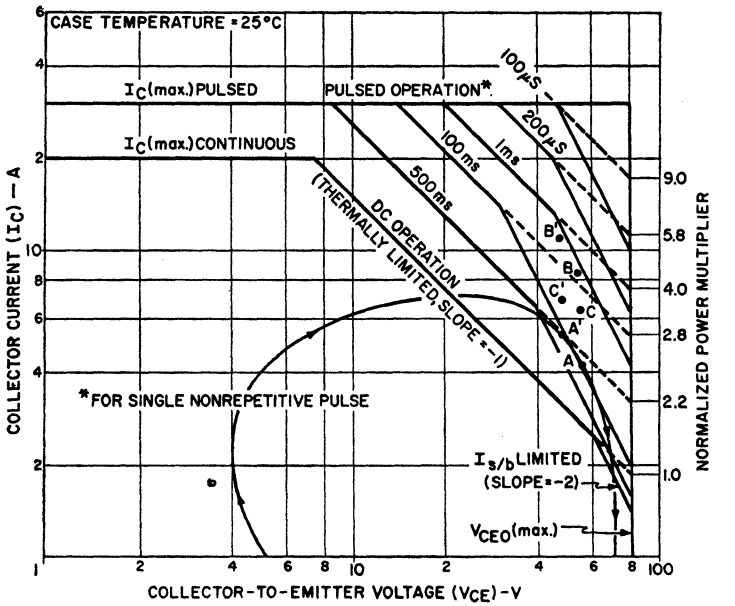


Figure 118. Safe-area rating chart for an RCA 40411 transistor.

Fig. 119. 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 greater distance between points A and C than between points A′ and C′ in Fig. 118 indicates that there is a greater margin of safety in the second-breakdown region than in the dissipation region.

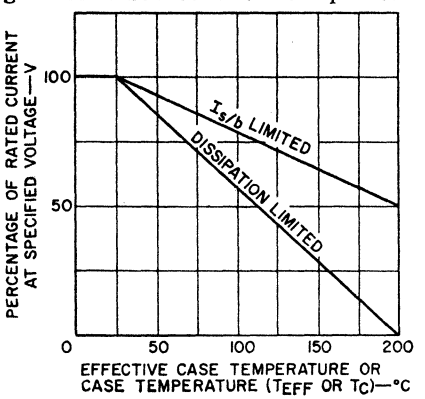


Figure 119. Temperature derating for the RCA 40411 transistor.

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. (36).

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. 120. These curves, expressed in terms of peak current I_{pk}, can be readily converted to energy E through the use of the following relationship:

$$E = \tfrac{1}{2} L I_{pk}^2 \qquad (40)$$

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. 112 is used.

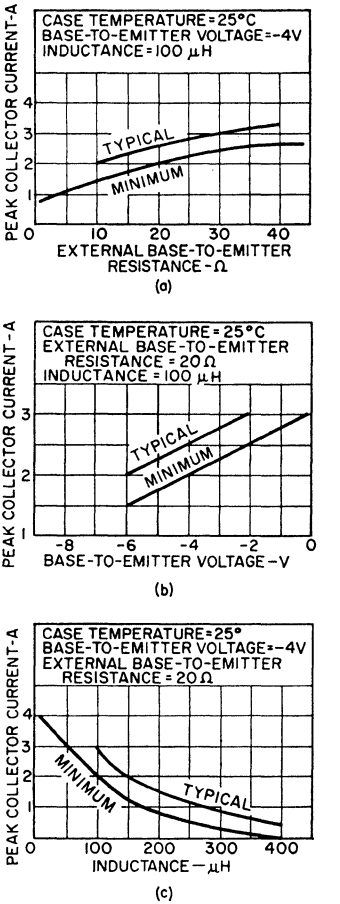


Figure 120. Reverse-bias energy of the RCA 2N3585 transistor.

The use of the reverse-bias second-breakdown rating curves of Fig. 120 is illustrated below by analysis of the inverter circuit shown in Fig. 113. 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 trans-

former in the inverter shown in Fig. 113 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. (36). However, the rating curves in Fig. 120 are based on measurements made in the $E_{S/b}$ test set with constant current drive, and on results calculated by use of Eq. (40). Therefore, an effective series inductance for the circuit is obtained by setting Eq. (36) equal to Eq. (40) and solving in terms of the inductance L or, in this case, L_{eff}.

$$L_{eff} = L\left[\frac{V_{CEX(sus)}}{V_{CEX(sus)} - V_{CC}}\right] \quad (41)$$

Eq. (41) is valid for all circuits that operate from a constant supply voltage. For the inverter circuit of Fig. 113, 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. (41), 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. 120. 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. 120(a) and 120(b) as 2 amperes.

To permit the application of the curves of Fig. 120 to the R_{BE} and V_{BE} of the inverter circuit, translation ratios must be calculated from the slope of the curves in Figs. 120(a) and 120(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. 120(a). The result is as follows:

$$R_{BE}\ (trans) = \frac{0.7}{2.0} = 0.35$$

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. 120(b). This ratio is given by

$$V_{BE}\ (trans) = \frac{1.5}{2.0} = 0.75$$

The minimum peak current for the series inductance of 25 microhenries is determined from Fig. 120(c) as 3.4 amperes. The equivalent minimum peak current for the inverter circuit is obtained by translating this value as follows:

$$I_{pk} = (3.4A)\ (0.35)\ (0.75)$$
$$= 0.89 \text{ amperes}$$

This peak current can then be converted to the second-breakdown energy of the inverter circuit, as follows:

$$E_{s/b} = \tfrac{1}{2} LI_p^2 = (\tfrac{1}{2})\ (25\ \mu H)\ (0.89)^2 = 10 \text{ microjoules}$$

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

$$E_{S/b} = \tfrac{1}{2} L_{eff} I_{pk}^2 = (\tfrac{1}{2})\ (25\ \mu H)\ (0.8)^2 = 8 \text{ microjoules}$$

The average-power contribution to this turn-off energy is determined by dividing the energy by the pulse period, as follows:

$$P_{avg} = 8\ \mu J/40\ \mu S = 0.2 \text{ W}.$$

Substitution of this value in Eq. 7 yields

$$T_c\,(\text{eff}) = T_c + P_{avg}\,\theta_{J/c} = T_c + (0.2)\ (5°C/W) = T_c + 1°C$$

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 $I_{pk}(act)$ by the maximum safe-area operation value of $I_{pk}(S.A.)$, as follows:

$$\frac{I_{pk}(act)}{I_{pk}(S.A.)} = \frac{0.8}{0.89} = 0.9, \text{ or } 90\%$$

From Fig. 112, 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.

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. 121. The terminal parameters for the two-port net-

work are defined by six sets of equations which differ only in the

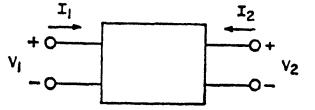


Figure 121. Linear two-port network.

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 $V_1 = z_{11}I_1 + z_{12}I_2$ (42)

CE $V_i = z_{ie}i_i + z_{re}i_o$ (43)

General $V_2 = z_{21}I_1 + z_{22}I_2$ (44)

CE $V_o = z_{fe}i_i + z_{oe}i_o$ (45)

2. y-parameter equations:

General $I_1 = y_{11}V_1 + y_{12}V_2$ (46)

CE $i_i = y_{ie}V_i + y_{re}V_o$ (47)

General $I_2 = y_{21}V_1 + y_{22}V_2$ (48)

CE $i_o = y_{fe}V_i + y_{oe}V_o$ (49)

3. h-parameter equations:

General $V_1 = h_{11}I_1 + h_{12}V_2$ (50)

CE $V_i = h_{ie}i_i + h_{re}V_o$ (51)

General $I_2 = h_{21}I_1 + h_{22}V_2$ (52)

CE $i_o = h_{fe}i_i + h_{oe}V_o$ (53)

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

$$h_{ie} = \left.\frac{V_i}{i_i}\right|_{V_o=0} \quad (54)$$

short-circuit forward-current transfer ratio

$$h_{fe} = \left.\frac{i_o}{i_i}\right|_{V_o=0} \quad (55)$$

open-circuit reverse-voltage transfer ratio

$$h_{re} = \left.\frac{V_i}{V_o}\right|_{i_i=0} \quad (56)$$

open-circuit output conductance

$$h_{oe} = \left.\frac{i_o}{V_o}\right|_{i_i=0} \quad (57)$$

The equivalent circuit for these equations is shown in Fig. 122.

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

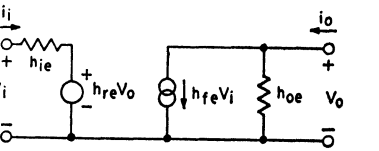


Figure 122. h-parameter equivalent circuit for a transistor used in a common-emitter configuration.

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. 123:

$$E_{r1} = S_{11}E_{i1} + S_{12}E_{i2} \qquad (58)$$

$$E_{r2} = S_{21}E_{i1} + S_{22}E_{i2} \qquad (59)$$

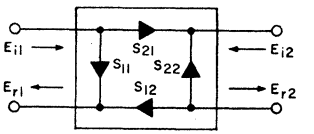


Figure 123. Scattering parameters for linear two-port network.

If Eqs. (58) and (59) 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

$$S_{11} = \left.\frac{E_{r1}}{E_{i1}}\right|_{E_{i2}=0} \qquad (60)$$

reverse-transmission voltage ratio with port 1 terminated in a matched load

$$S_{12} = \left.\frac{E_{r1}}{E_{i2}}\right|_{E_{i1}=0} \qquad (61)$$

forward-transmission voltage ratio with port 2 terminated in a matched load

$$S_{21} = \left.\frac{E_{r2}}{E_{i1}}\right|_{E_{i2}=0} \qquad (62)$$

voltage-reflection coefficient at port 2 with port 1 terminated in a matched load

$$S_{22} = \left.\frac{E_{r2}}{E_{i2}}\right|_{E_{i1}=0} \qquad (63)$$

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 VI.

When the transistor is connected in a practical circuit, such as that shown in Fig. 124, the

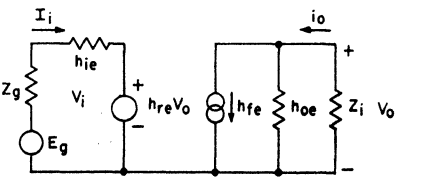


Figure 124. Terminated h-parameter equivalent circuit for a transistor used in a common-emitter configuration.

source and load impedances affect the terminal properties of the network. This effect can best be shown by calculation of the network input impedance from the h-parameter equations, as follows:

$$V_i = h_{ie} i_i + h_{re} V_o \quad (51)$$

$$i_o = h_{fe} i_i + h_{oe} V_o \quad (53)$$

Table VI—Parameter Equivalencies

	z		y		h		S	
z	z_{11}	z_{12}	$\frac{y_{22}}{\Delta y}$	$\frac{-y_{12}}{\Delta y}$	$\frac{\Delta h}{h_{22}}$	$\frac{h_{12}}{h_{22}}$	$\frac{1+S_{11}-S_{22}-\Delta S}{1-S_{11}-S_{22}+\Delta S}$	$\frac{2\,S_{12}}{1-S_{11}-S_{22}+\Delta S}$
	z_{21}	z_{22}	$\frac{-y_{21}}{\Delta y}$	$\frac{y_{11}}{\Delta y}$	$\frac{-h_{21}}{h_{22}}$	$\frac{1}{h_{22}}$	$\frac{2\,S_{21}}{1-S_{11}-S_{22}+\Delta S}$	$\frac{1-S_{11}+S_{12}-\Delta S}{1-S_{11}-S_{22}+\Delta S}$
y	$\frac{z_{22}}{\Delta z}$	$\frac{-z_{12}}{\Delta z}$	y_{11}	y_{12}	$\frac{1}{h_{11}}$	$\frac{-h_{12}}{h_{11}}$	$\frac{1-S_{11}+S_{22}-\Delta S}{1+S_{11}+S_{22}+\Delta S}$	$\frac{-2\,S_{12}}{1+S_{11}+S_{22}+\Delta S}$
	$\frac{-z_{21}}{\Delta z}$	$\frac{z_{11}}{\Delta z}$	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}	$\frac{1-S_{11}+S_{22}-\Delta S}{1-S_{11}-S_{22}+\Delta S}$	$\frac{2\,S_{12}}{1-S_{11}+S_{22}-\Delta S}$
	$\frac{-z_{21}}{z_{22}}$	$\frac{1}{z_{22}}$	$\frac{y_{21}}{y_{11}}$	$\frac{\Delta y}{y_{11}}$	h_{21}	h_{22}	$\frac{-2\,S_{21}}{1-S_{11}-S_{22}-\Delta S}$	$\frac{1-S_{11}-S_{12}+\Delta S}{1-S_{11}+S_{22}-\Delta S}$
S	A	C	E	G	I	K	S_{11}	S_{12}
	B	D	F	H	J	L	S_{21}	S_{22}

$$A=\frac{\Delta z+z_{11}-z_{22}-1}{\Delta z+z_{11}+z_{22}+1} \quad C=\frac{2\,z_{12}}{\Delta z+z_{11}+z_{22}+1} \quad E=\frac{(1-y_{11})(1+y_{22})-y_{12}y_{21}}{1+y_{11}+y_{22}+\Delta y}$$

$$B=\frac{2\,z_{21}}{\Delta z+z_{11}+z_{22}+1} \quad D=\frac{\Delta z-z_{11}+z_{22}-1}{\Delta z+z_{11}+z_{22}+1} \quad F=\frac{-2\,y_{21}}{1+y_{11}+y_{22}+\Delta y}$$

$$G=\frac{-2\,y_{12}}{1+y_{11}+y_{22}+\Delta y} \quad I=\frac{\Delta h+h_{11}-h_{22}-1}{\Delta h+h_{11}+h_{22}+1} \quad K=\frac{2\,h_{12}}{\Delta h+h_{11}+h_{22}+1}$$

$$H=\frac{(1+y_{11})(1-y_{22})-y_{12}y_{21}}{1+y_{11}+y_{22}+\Delta y} \quad J=\frac{-2\,h_{21}}{\Delta h+h_{11}+h_{22}+1} \quad L=\frac{-\Delta h+h_{11}-h_{22}+1}{\Delta h+h_{11}+h_{22}+1}$$

$$\Delta z = z_{11}\,z_{22} - z_{12}\,z_{21} \qquad \Delta h = h_{11}\,h_{22} - h_{12}\,h_{21}$$
$$\Delta y = y_{11}\,y_{22} - y_{12}\,y_{21} \qquad \Delta S = S_{11}\,S_{22} - S_{12}\,S_{21}$$

For the terminated network shown in Fig. 124, V_o may be expressed by the following relationship:

$$V_o = -i_o z_L = -i_o/y_L \qquad (64)$$

Eq. (53) then becomes

$$i_o = h_{fe} i_i + h_{oe}(-i_o/y_L) \qquad (65)$$

or

$$i_o = \frac{h_{fe}(y_L) i_i}{y_L + h_{oe}} \qquad (66)$$

Substitution of Eq. (64) in Eq. (51) yields the following result:

$$V_i = \left(h_{ie} + \frac{h_{fe}\, h_{re}}{y_L + h_{oe}}\right) i_i \qquad (67)$$

If both sides of Eq. (67) are divided by i_i, the following equation for the input impedance Z_{in} is obtained:

$$Z_{in} = \frac{V_i}{i_i} = h_{ie} + \frac{h_{fe}\, h_{re}}{y_L + h_{oe}} \qquad (68)$$

In a similar manner, the equation for the output impedance Z_{out} may be derived to obtain the following relationship:

$$Z_{out} = \frac{V_o}{i_o} = \frac{Z_g + h_{ie}}{Z_g h_{oe} + h_{ie} h_{oe} - h_{re} h_{fe}} \qquad (69)$$

Eq. (66) is rewritten to obtain the following expression for the current gain K_i of the circuit shown in Fig. 124:

$$K_i = \frac{i_o}{i_i} = \frac{h_{fe}\, y_L}{h_{oe} + y_L} \qquad (70)$$

The voltage gain K_v of the circuit, as determined from Eqs. (64), (66) and (67), may be expressed by the following relationship:

$$K_v = \frac{V_o}{V_i} = \frac{-h_{fe}/h_{ie}}{y_L + h_{oe}\left(1 - \frac{h_{fe}\, h_{re}}{h_{ie}\, h_{oe}}\right)} \qquad (71)$$

The power gain PG, which is the product of the current and voltage gains, may be expressed as follows:

$$PG = K_i K_v \qquad (72)$$

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 VII.

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 VIII.

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. 125. 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 VII—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 VIII—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}	$h_{ob} = \frac{h_{oe}}{1 + h_{fe}}$	$h_{oc} = h_{oe}$

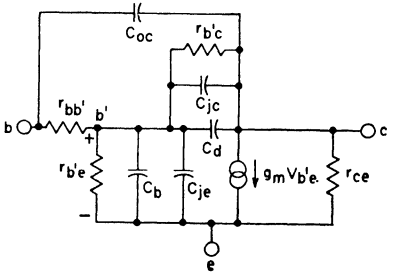


Figure 125. Complete hybrid-pi equivalent circuit for a transistor used in a common-emitter configuration.

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:

$$i_{b1} = g_{b'e} V_{be} = (1/r_{b'e}) V_{be} \quad (73)$$

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.

$$i_{b2} = C_b (dV_{be}/dt) \quad (74)$$

The total base current i_b then is defined by the following equation:

$$i_b = (1/r_{b'e}) V_{be} + C_b (dV_{be}/dt) \quad (75)$$

where C_b is the base charging capacitance.

Eq. (75) provides the basis for a first-order equivalent-circuit approximation of the transistor. This equivalent circuit, shown in Fig. 126, 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. 126. Under conditions of large voltage gain the load impedance is large. The width of the

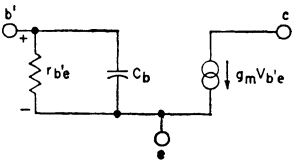


Figure 126. Elementary small-signal transistor equivalent circuit.

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. 127.

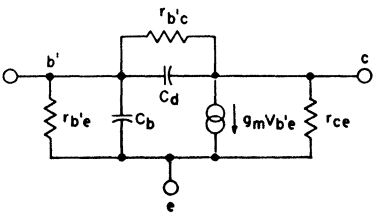


Figure 127. Elementary small-signal transistor equivalent circuit which includes components to represent effects of base-width modulation.

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

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. 125, 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. 128.

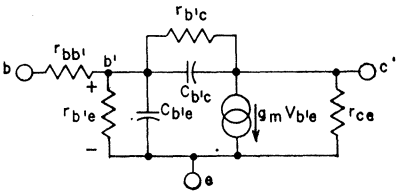


Figure 128. Conventional form of the complete hybrid-pi equivalent circuit shown in Fig. 125.

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. 129. As the first step in the estimation of the frequency response for this circuit, it is necessary to

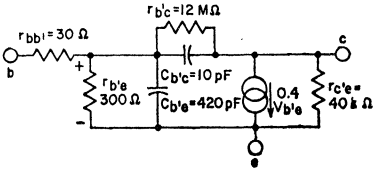


Figure 129. Hybrid-pi equivalent circuit for the 2N2102 triple-diffused silicon planar transistor.

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:

$$f_1 = \frac{1}{2\pi\, r_{b'e}\, C_{b'e}} \qquad (76)$$

$$= \frac{1}{6.28 \times 300 \times 420 \times 10^{-12}}$$

$$= 1.3 \text{ MHz}$$

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:

$$f_2 = \frac{1}{2\pi\, r_{b'c}\, C_{b'c}} \qquad (77)$$

$$= \frac{1}{6.28 \times 12 \times 10^6 \times 10 \times 10^{-12}}$$

$$= 1.3 \text{ kHz}$$

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. 130 is applicable. At frequencies above f_2 but much below f_1, $r_{b'e}$ and $C_{b'e}$ are negligible, and the middle-frequency equivalent circuit shown in Fig. 131 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. 132 should be used to represent the transistor.

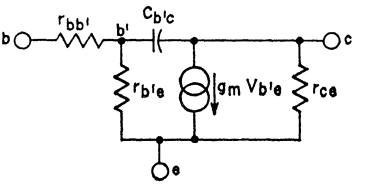


Figure 130. Simplified low-frequency hybrid-pi transistor equivalent circuit.

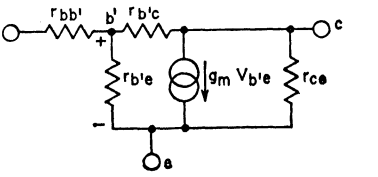


Figure 131. Simplified medium-frequency hybrid-pi transistor equivalent circuit.

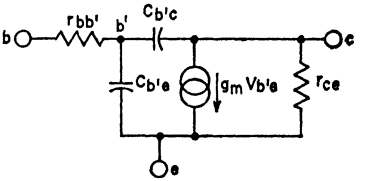


Figure 132. Simplified high-frequency hybrid-pi transistor equivalent circuit.

In practice, the mid-frequency model shown in Fig. 131 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. 133, is useful over the middle- and high-frequency ranges.

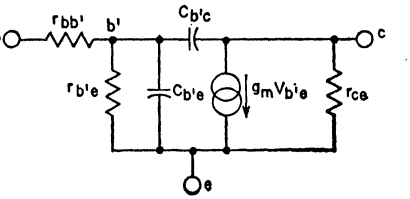


Figure 133. Simplified hybrid-pi transistor equivalent circuit for use at medium and high frequencies.

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:

$$g_m = \frac{q}{kT} \mid I_C \mid \text{mhos} \qquad (78)$$

where

q = electronic charge = 1.6×10^{-19} coulomb
k = Boltzmann's constant = 1.38×10^{-23} watt-sec/°K
T = absolute temperature in °K
I_C = collector current in amperes.

At room temperature, T = 290°K. Eq. (78) can then be rewritten as follows:

$$g_m = 0.04 \mid I_C \mid \text{mhos} \qquad (79)$$

where the collector current I_C is given in milliamperes.

The 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. 130, the relationship between the resistance $r_{b'e}$ and the low-frequency current gain may be expressed as follows:

$$h_{fe}|_{flo} = i_o/i_i = g_m r_{b'e} \qquad (80)$$

When this equation is solved for the resistance $r_{b'e}$, the following result is obtained:

$$r_{b'e} = \frac{h_{fe}|_{flo}}{g_m} \qquad (81)$$

At low frequencies, the 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 measurement, 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:

$$h_{ie}|_{flo} = V_i/i_i = r_{bb'} + r_{b'e} \qquad (82)$$

The low-frequency value for the resistance $r_{bb'}$, therefore, may be determined from the following relationship:

$$r_{bb'} = h_e|_{flo} - r_{b'e} \qquad (83)$$

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:

$$1/r_{bb'} = R_e(y_{ie}) \qquad (84)$$

Eq. (84) is valid for the following condition:

$$\omega(C_{b'e} + C_{b'c}) > > 1/r_{bb'} \qquad (85)$$

In the calculation of the 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. 130, the open-circuit reverse-voltage transfer ratio can be related to the resistance $r_{b'e}$ as follows:

$$h_{fe|flo} \simeq \frac{V_i}{V_o} = \frac{r_{b'e}}{r_{b'e} + r_{b'c}} \quad (86)$$

From practical considerations, it is known that $r_{b'c}$ is much greater than $r_{b'e}$. Eq. (86), therefore, may be rewritten as follows:

$$h_{re|flo} = r_{b'e}/r_{b'c} \quad (87)$$

or

$$r_{b'c} = \frac{r_{b'e}}{h_{re|flo}} \quad (88)$$

In most practical applications of power transistors, the 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:

$$h_{oe|flo} \approx 1/r_{ce} \quad (89)$$

or

$$r_{ce} \approx \frac{1}{h_{oe|flo}} \quad (90)$$

Before the capacitance $C_{b'c}$ can be calculated, the common-base output capacitance C_{ob} must be measured. Fig. 125 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:

$$C_{ob} = C_{b'c} + C_{oc} + C_H \quad (91)$$

$$C_{b'c} = C_{ob} - C_{oc} - C_H \quad (92)$$

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:

$$C_{b'c} \approx C_{ob} \quad (93)$$

The remaining capacitance $C_{b'e}$ may be determined by use of the high-frequency equivalent circuit shown in Fig. 133. For this circuit, the short-circuit current gain is determined from the following relationship:

$$h_{fe|fhi} = \frac{i_o}{i_i} = \frac{g_m}{\omega(C_{b'e} + C_{b'c})} \quad (94)$$

If the gain-bandwidth product f_T is defined as the frequency at which $h_{fe} = 1$, the following frequency relationships become apparent:

$$h_{fe}f = f_T \quad (95)$$

$$h_{fe}\omega = \omega_T \quad (96)$$

$$\omega = \omega_T/h_{fe} \quad (97)$$

If the relationship expressed by Eq. (96) is substituted into Eq. (94), the following result is obtained:

$$\omega_T = g_m/(C_{b'e} + C_{b'c}) \quad (98)$$

$$C_{b'e} = (g_m/\omega_T) - C_{b'c} \quad (99)$$

Eqs. (78) through (99) demonstrate how the hybrid-pi parameters may be calculated for a known operating condition. For convenience and reference, these relations are summarized in Table IX.

Table IX—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 $\omega (C_{b'e} + C_{b'c}) >> \frac{1}{r_{bb'}}$		
$r_{b'c}$	$\cong$	$\frac{r_{b'e}}{h_{re}\lvert f_{lo}}$
r_{ce}	$\cong$	$h_{oe}\lvert f_{lo}$
$C_{b'c}$	=	$C_{ob} - C_{oc} - C_H \cong C_{ob}$
$C_{b'e}$	=	$\frac{g_m}{\omega_T} - C_{b'c}$

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. 134 are both very large, their effects are negligible, and the equivalent circuit is as shown in Fig. 135. The node equations for this circuit can be written as follows:

$$V_{b'e}[(1/r_{b'e}) + sC_{b'e} + sC_{b'c}] - V_o(sC_{b'c}) = I_i \quad (100)$$

$$V_{b'e}(g_m - sC_{b'c}) + V_o[(1/r_{ce}) + (1/R_L) + sC_{b'c}] = 0 \quad (101)$$

where $s = j\omega C$.

When conductance terms are substituted for the resistance terms, the node equations become

$$V_{b'e}[g_{b'e} + s(C_{b'e} + C_{b'c})] - V_o(sC_{b'c}) = I_i \quad (102)$$

$$V_{b'e}(g_m - sC_{b'c}) + V_o(g_{ce} + G_L + sC_{b'c}) = 0 \quad (103)$$

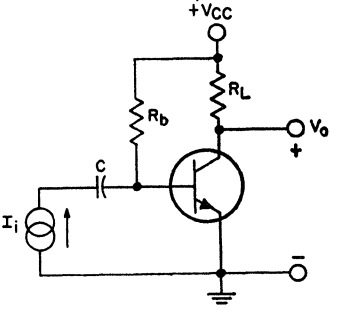


Figure 134. Common-emitter transistor amplifier.

Eqs. (102) and (103) may be further simplified by practical considerations. In the second term of Eq. (102), the $sC_{b'c}$ element represents feedback from output to input which is necessary and critical. In Eq. (103), the $sC_{b'c}$ element in

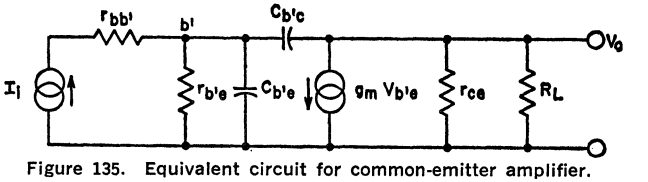


Figure 135. Equivalent circuit for common-emitter amplifier.

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. (103), 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. (102) and (103) may be rewritten in the following form:

$$V_{b'e}\,[g_{b'e} + s(C_{b'e} + C_{b'c})] - V_o\,(sC_{b'c}) = I_i \qquad (104)$$

$$V_{b'e}\,(g_m) + V_o\,G_L = 0 \qquad (105)$$

Eqs. (102) and (103) can be solved for the internal input admittance $I_i/V_{b'e}$ as follows:

$$V_o = (g_m/V_{b'e})/G_L \qquad (106)$$

$$V_{b'e}\,[g_{b'e} + s(C_{b'e} + C_{b'c})] + [(V_{b'e}\,g_m)/G_L]\,(sC_{b'c}) = I_i \qquad (107)$$

$$V_{b'e}\,\{g_{b'e} + s[C_{b'e} + C_{b'c}\,(1 + g_m\,R_L)]\} = I_i \qquad (108)$$

$$y = \frac{I_i}{V_{b'e}} = g_{b'e} + s\,[C_{b'e} + C_{b'c}\,(1 + g_m R_L)] \qquad (109)$$

Eq. (109) represents an equivalent input circuit that consists of a resistor $r_{b'e}$ in shunt with an equivalent capacitance expressed by the following equation:

$$C_{eq} = C_{b'e} + C_{b'c}\,(1 + g_m\,R_L) \qquad (110)$$

When the series spreading resistance $r_{bb'}$ is added, the total input circuit can be represented as shown in Fig. 136. Eq. (109) was derived for the case of a purely

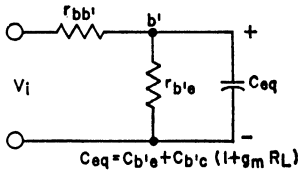


Figure 136. Equivalent input circuit for common-emitter amplifier.

resistive load; this equation, however, can be generalized so that it is applicable to a complex load impedance Z_L as follows:

$$y = g_{b'e} + s\,[C_{b'e} + C_{b'c}\,(1 + g_m\,Z_L)] \qquad (111)$$

A circuit representation of the generalized input impedance defined by Eq. (111) is shown in Fig. 137. The equivalent admit-

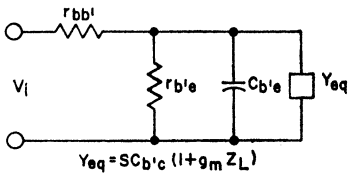


Figure 137. General equivalent input circuit for a common-emitter amplifier.

tance for this circuit is expressed by the following equation:

$$y_{eq} = sC_{b'c}\,(1 + g_m\,Z_L) \qquad (112)$$

Eqs. (109) and (111) 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 capaci-

tance may be calculated from Eq. (110) 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. 128 can be redrawn for the common-base configuration as shown in Fig. 138. 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. 139. In this circuit, the resistance r_e and

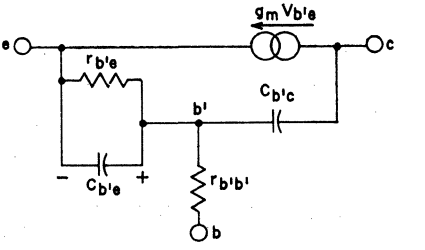


Figure 138. Hybrid-pi equivalent circuit for a transistor used in a common-base configuration (effects of base-width modulation are neglected).

the current-gain parameter α are defined as follows:

$$r_e = r_{b'e} \parallel \frac{1}{g_m} = \frac{1}{(1/r_{b'e}) + g_m} \approx \frac{1}{g_m} \qquad (113)$$

$$\alpha = \frac{g_m}{g_m + (1/r_{b'e})} \approx 1 \qquad (114)$$

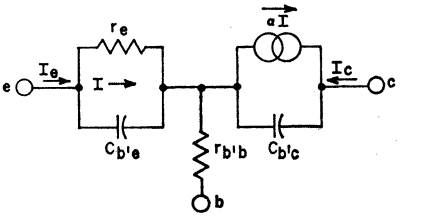


Figure 139. "T" equivalent circuit for a transistor used in a common-base configuration when the effects of base-width modulation are neglected.

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. 140. It can be shown that if the base-modulation, base-spreading, output-loading, and input-loading terms are neglected, the

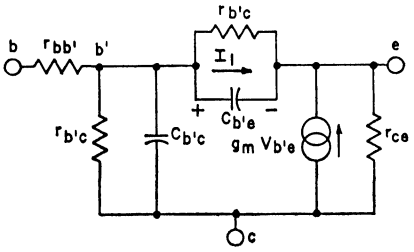


Figure 140. Hybrid-pi equivalent circuit for a transistor used in a common-collector configuration.

common-collector equivalent circuit can be simplified to the form shown in Fig. 141, in which the dependent generator is converted into a current-dependent source.

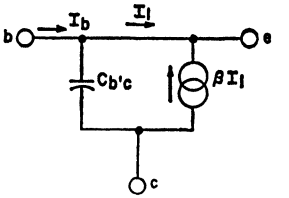


Figure 141. Simplified form of hybrid-pi equivalent circuit shown in Fig. 140.

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 VII) 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 VII. The calculations are laborious, and the results are readily available in the literature. The general qualitative results are shown in Table X. 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. 142, the short-circuit current gain h_{fe} can be determined from the following equations:

Table X—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

$$V_{b'e} = \frac{i_i}{y_i|_{s.c.}} = \frac{i_i}{(1/r_{b'e})+s(C_{b'e}+C_{b'c})} \quad (115)$$

$$i_o = g_m V_{b'e} \quad (116)$$

$$h_{fe} = \frac{i_o}{i_i} = \frac{g_m}{(1/r_{b'e}) + s(C_{b'e} + C_{b'c})} \quad (117)$$

The s term ($s = j\omega C$) in Eq. (117) shows that h_{fe} is a complex number that has both magnitude

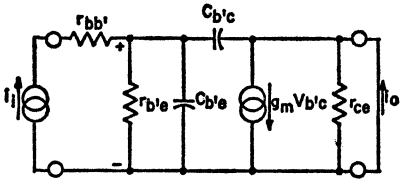


Figure 142. Medium- and high-frequency common-emitter hybrid-pi transistor model used for calculation of h_{fe}.

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:

$$(1/r_{b'e}) = 2\pi f_\beta (C_{b'e} + C_{b'c}) \quad (118)$$

$$f_\beta = \frac{1}{2\pi r_{b'e} (C_{b'e} + C_{b'c})} \quad (119)$$

The beta-cutoff frequency f_β is the frequency at which the short-circuit current ratio h_{fe} (or β) 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. 143 and 144. For frequencies much above f_β, the response curve becomes asymptotic to a line that has a slope of -1 on the log-log

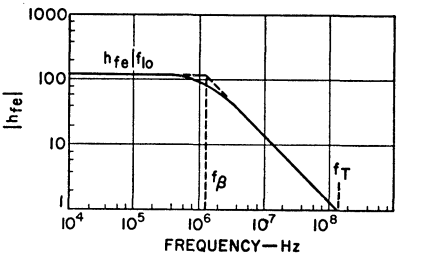


Figure 143. Magnitude of h_{fe} as a function of frequency for the 2N2102 silicon power transistor.

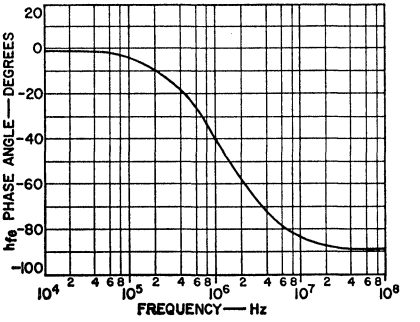


Figure 144. Phase of h_{fe} as a function of frequency for the 2N2102 silicon power transistor.

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:

$$f_T = \frac{g_m}{2\pi(C_{b'e} + C_{b'c})} \quad (120)$$

The term f_T is called the gain-bandwidth product. At any frequency along the -1 asymptote (i.e., $f > 3 f_\beta$), Eq. (120) can be rewritten in terms of the operating frequency, as follows:

$$|h_{fe}| \omega = \omega_T \quad (121)$$

$$|h_{fe}| f = f_T \quad (122)$$

For the hybrid-pi model for the common-base configuration shown in Fig. 145, the common-base short-circuit current gain h_{fb} can

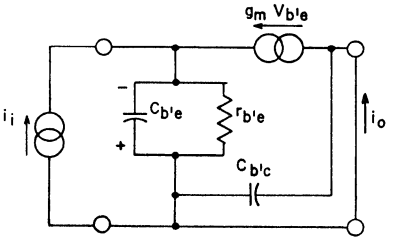


Figure 145. High-frequency common-base hybrid-pi transistor model used for calculation of h_{fb} (transverse voltage drops in the base are neglected).

be calculated, for the case when $r_{bb'} = 0$, from the following equation:

$$h_{fb} = \frac{i_o}{i_i} = \frac{g_m}{(1/r_{b'e}) + sC_{b'e}} \qquad (123)$$

The s term in Eq. (123) 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_α, referred to as the alpha cutoff frequency, at which h_{fb} is reduced to 0.707 of its low-frequency value. The frequency f_α is determined as follows:

$$f_\alpha = \frac{1}{2\pi\, r_{b'e}\, C_{b'e}} \qquad (124)$$

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 IX), 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. 125 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 IX shows that the transconductance g_m may be defined as follows:

$$g_m = \frac{q}{kT} \mid I_C \mid \qquad (125)$$

Eq. (125) indicates that g_m is directly dependent upon the collector current and is independent of the collector voltage V_C. Table IX also shows that $r_{b'e}$ may be determined from the following relationship:

$$r_{b'e} = \frac{h_{fe}|_{flo}}{g_m} = \frac{h_{fe}|_{flo}}{\frac{q}{kT}|I_C|} \quad (126)$$

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:

$$i_{b2} = (C_b \, d \, V_{be})/dt \quad (127)$$

Eq. (127) may be rewritten in the following form:

$$dQ_b = i_{b2}dt = C_b dV_{be} \quad (128)$$

or

$$C_b = dQ_b/d\,V_{be} \quad (129)$$

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. 146. The total base charge can be found by integration of Eq. (128), 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:

$$Q_b = (q\,n'_{b(o)}\,WA)/2 \quad (130)$$

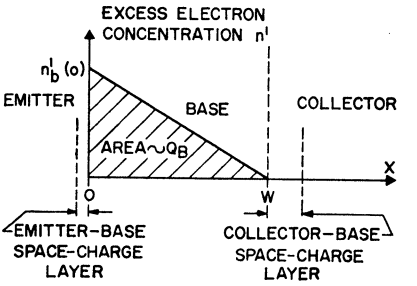


Figure 146. Excess minority-carrier concentration in the base region of a uniform-base n-p-n transistor.

where q is the electronic charge, $n'_{b(o)}$ is the concentration of excess electrons at the edge of the base region (X = O), 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:

$$I_C = qD_nA\left[\frac{n'_{b(o)}}{W}\right] \quad (131)$$

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:

$$T_F = Q_b/I_C = W^2/2D_n \quad (132)$$

Eqs. (129) through (132) are combined to obtain the following equation, which can be used to determine the variation in C_b with incremental changes in Q_b:

$$C_b = \frac{dQ_b}{dI_C}\left(\frac{d\,I_C}{d\,V_{be}}\right) = \left(\frac{W^2}{2\,D_n}\right)g_m = T_F g_m \quad (133)$$

It has been shown previously, by Eq. (125), that g_m is directly dependent upon I_C; Eq. (133), therefore, shows that the same dependence applies for Q_b.

The relationships expressed by Eqs. (125) through (133) 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. (125) and (126) 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. (132) and (133) 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. 125, 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:

$$C_J = K/(V'/n) \qquad (134)$$

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. 147 and a knowledge of resistance effects, r_{bb}' can be expressed as follows:

$$r_{bb}' = P_b X/A = P_b X/WY \qquad (135)$$

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.

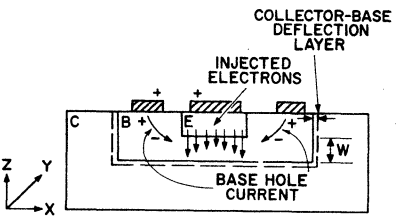


Figure 147. Model of n-p-n transistor showing transverse base field and current-crowding effect.

Temperature Dependence—The parameters shown in the hybrid-pi equivalent circuit are also temperature-dependent. Eq. (113) shows that g_m is inversely proportional to temperature. It can be determined from Eq. (126) that $r_b'_e$ is proportional to $h_{fe}|f_{lo}(T)$. Because the parameter $h_{fe}|f_{lo}$ in-

Table XI—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 β_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 β_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. (125), (132), and (133) 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. (135) with the knowledge that, for the normal resistivities used, ρ 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 XI summarizes these discussions, and Figs. 148 and 149 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

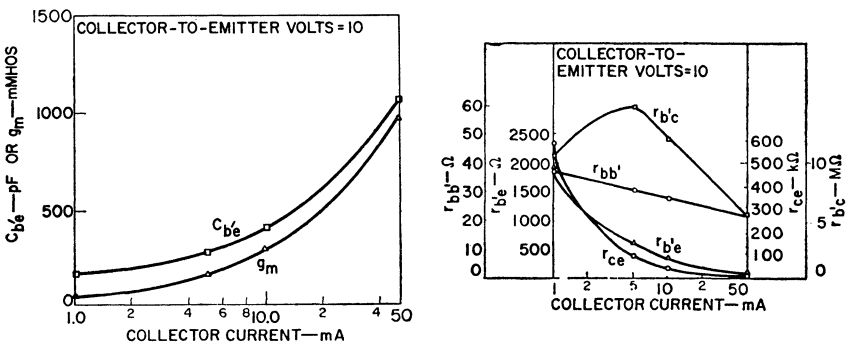


Figure 148. Hybrid-pi parameters as a function of collector current for the 2N2102 silicon power transistor.

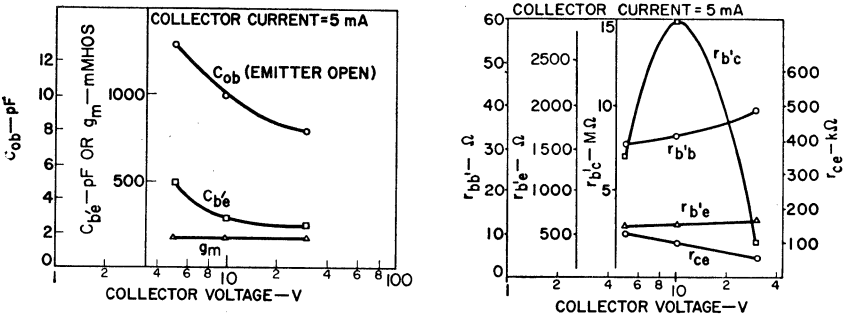


Figure 149. Hybrid-pi parameters as a function of collector voltage for the 2N2102 silicon power transistor.

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. 150. 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:

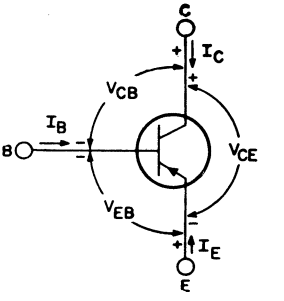


Figure 150. Terminal parameters for a p-n-p transistor.

$$I_{EF} = I_{ES}[\exp(qV_{EB}/kT) - 1] \quad (136)$$

$$I_{CF} = -\alpha_F I_{EF} \quad (137)$$

where the subscript F denotes forward operation, α_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:

$$I_{CR} = I_{CS}[\exp(qV_{CB}/kT) - 1] \quad (138)$$

$$I_{ER} = -\alpha_R I_{CR} \quad (139)$$

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:

$$I_E = I_F + I_{ER} = I_{ES}[\exp(qV_{EB}/kT) - 1] - \alpha_R I_{CS}[\exp(qV_{CB}/kT) - 1] \quad (140)$$

$$I_C = I_{CF} + I_{CR} = -\alpha_F I_{ES}[\exp(qV_{EB}/kT) - 1] + I_{CS}[\exp(qV_{CB}/kT) - 1] \quad (141)$$

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. 151. 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.

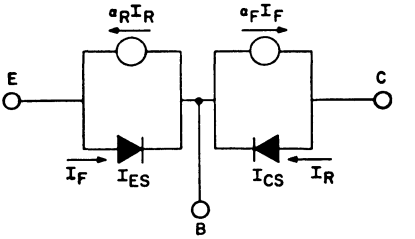


Figure 151. Ebers-Moll model of a p-n-p transistor controlled by diode currents.

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:

$$I_E = I_{EF} + I_{ER} = -I_{ES}[\exp(-qV_{EB}/kT) - 1] + \alpha_R I_{CS}[\exp(-qV_{CB}/kT) - 1] \quad (142)$$

$$I_C = I_{CF} + I_{CR} = \alpha_F I_{ES}[\exp(-qV_{EB}/kT) - 1] - I_{CS}[\exp(-qV_{CB}/kT) - 1] \quad (143)$$

Figs. 152 and 153 show the Ebers-Moll models for the n-p-n transistor.

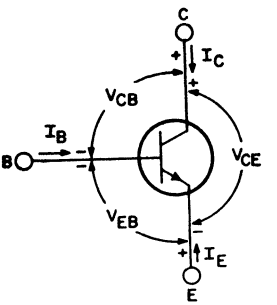


Figure 152. Terminal parameters for an n-p-n transistor.

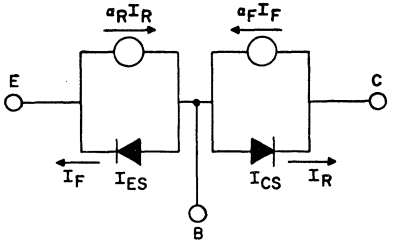


Figure 153. Ebers-Moll model of an n-p-n transistor controlled by diode currents.

Large-Signal Characteristics

Fig. 154 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

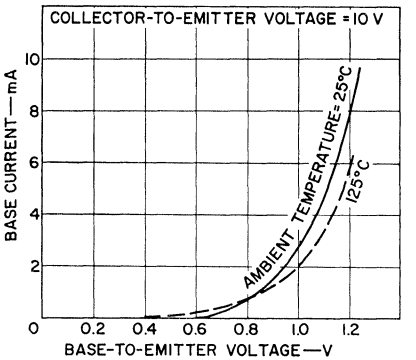


Figure 154. Typical input characteristics for the 2N3053 silicon power transistor.

that the resistance defined by this characteristic may be determined from the following calculation:

$$r_B = d\,V_{BE}/dI_B = r_{b'e} + r_{bb'} \qquad (144)$$

$$= \frac{h_{fe}|_{flo}}{g_m} + r_{bb'} \qquad (145)$$

$$= \frac{h_{fe}|_{flo}}{0.04|I_C|} + r_{bb'} \qquad (146)$$

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. 155 shows typical output characteristics for a 2N3053 transistor. These characteristics display the static or dc value of

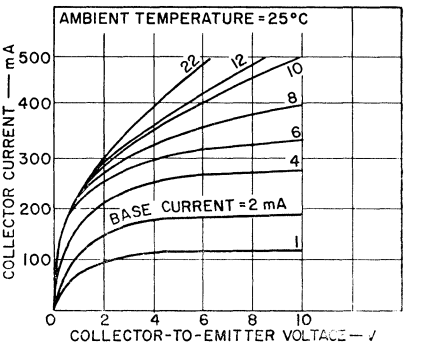


Figure 155. Typical output characteristics for the 2N3053 silicon power transistor.

h_{FE}, which is not the same as the low-frequency small-signal $h_{fe}|_{flo}$ used in previous calculations. The dc h_{FE} is useful primarily for establishment of dc bias conditions.

Fig. 156 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:

$$h_{FE} = \alpha/(1-\alpha) \qquad (147)$$

$$\alpha = \beta\,\gamma\,\mu \qquad (148)$$

where α is the h_{FB} or common-base short-circuit current gain, β is the transport factor, γ is the

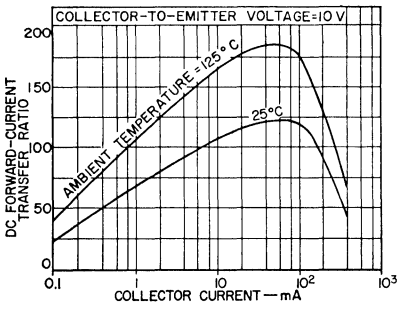


Figure 156. Typical dc forward-current transfer ratio (h_{FE}) as a function of collector current for the 2N3053 silicon power transistor.

emitter efficiency, and μ is the collector multiplication factor. For silicon transistors at low current, γ 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, γ increases because the recombination current remains constant. β 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 β and, therefore, increases h_{FE}.

The effect of temperature on h_{FE} is also shown in Fig. 156. 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. 157. If leakage current is considered, the equation for V_{BE} is as follows:

$$V_{BE} = (I_B - I_{CBO})\, r_{bb}' + V_{b'e} \quad (149)$$

$$= (I_C/h_{FE})\, r_{bb}' - I_{CBO}\, r_{bb}' + V_{b'e} \quad (150)$$

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

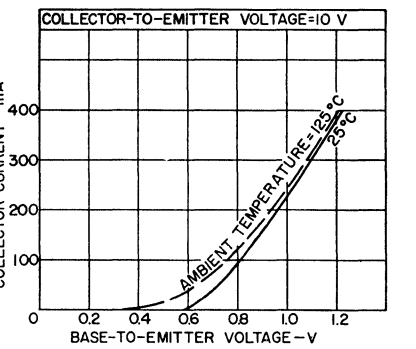


Figure 157. Typical transfer characteristics for the 2N3053 silicon power transistor.

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. 158.

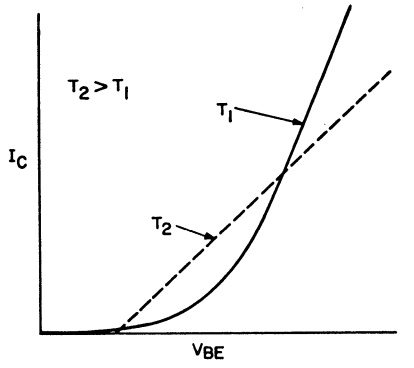


Figure 158. Typical transconductance characteristics for a silicon power transistor at high currents.

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. 159 shows a typical test circuit used to measure the switch-

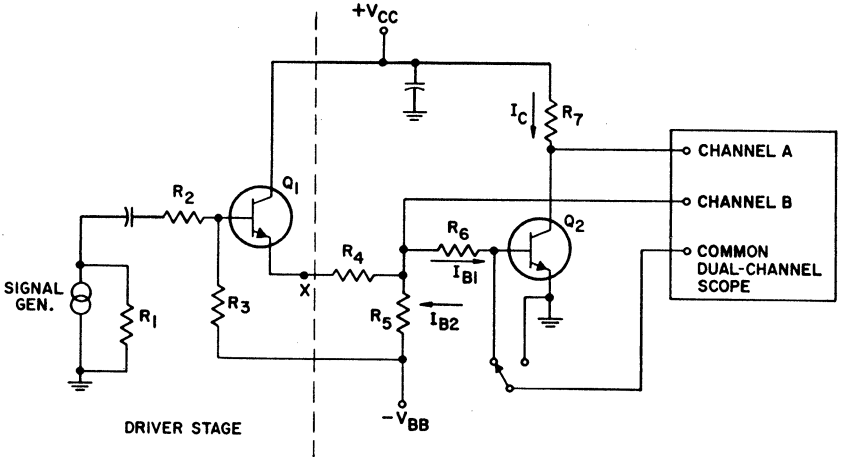


Figure 159. Test circuit used to measure switching times in power transistors.

ing times of a power transistor, and Figs. 160 and 161 show 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

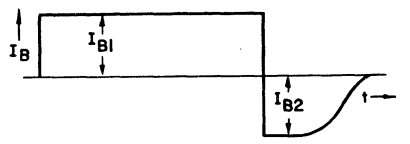


Figure 160. Base-current waveform that results from change in conducting state of a power transistor.

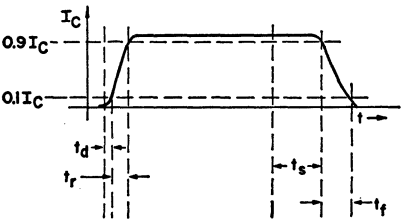


Figure 161. Collector-current waveform that results from change in conducting state of a power transistor.

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. 160. When the common point of the oscilloscope is connected to ground, the collector current I_C can be viewed. Fig. 161 shows the collector waveform. (It is, of course, inverted.) The waveforms for I_B and I_C shown in Figs. 160 and 161 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 of 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:

$$T_{on} = t_d + t_r \quad (151)$$

$$T_{off} = t_s + t_f \quad (152)$$

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,
$$\tau_E = (Q_B/I_E) \quad (153)$$

2. Base time constant,
$$\tau_B = (Q_B/I_B) \quad (154)$$

3. Collector time constant,
$$\tau_C = (Q_B/I_C) \quad (155)$$

The time constant τ_B represents the effective minority-carrier lifetime in the base. This time constant is related to the time constant τ_E as follows:

1. For a uniform base,
$$\tau_E = \tau_B(1-\alpha) \approx 1.2/\omega_b \quad (156)$$

2. For a graded base,
$$\tau_E = \tau_B(1-\alpha) \approx 0.6/\omega_b \quad (157)$$

where ω_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 τ_C can be defined in terms of τ_E by the following equation:

$$\tau_C = \tau_E/\alpha \quad (158)$$

The basic equation of charge continuity for transistors may be written as follows:

$$I_B = (dQ_B/dt) + (Q_B/\tau_B) \quad (159)$$

where I_B is the base current, dQ_B/dt is equal to the change in base charge, and Q_B/τ_B can be interpreted as the recombination rate.

Eq. (159) can be integrated with respect to time to obtain the following result:

$$\int_0^t I_B dt = \int_{Q_o}^{Qt} dQ_B + \int_0^t (Q_B dt/\tau_B) \quad (160)$$

Eq. (160) 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. (159)]:

$$\int_0^t I_B dt = \int_{V_{BE1}}^{V_{BE2}} C_{Te}\, dV_{BE} + \int_{V_{CB1}}^{V_{BC2}} C_{Tc}\, dV_{CB} + \int_{Q_{B1}}^{Q_{B2}} dQ_B + \int_0^t (Q_B dt/\tau_B) \quad (161)$$

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:

$$C_{Te} = C'_{Te}\, V_{BE}^{-\frac{1}{3}} \quad (162)$$
(step-junction assumption)

$$C_{Tc} = C'_{Tc}\, V_{CB}^{-\frac{1}{3}} \quad (163)$$
(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. (162) and (163) are used and the indicated

integrations in Eq. (161) are performed to obtain the following expression:

$$t_d = (2/I_{B1}) \{C'_{Te} [V_{BE(off)}]^{\frac{1}{2}} + C'_{Tc} [(V_{CE} + V_{BE(off)})^{\frac{1}{2}} - V_{CC}^{\frac{1}{2}}]\} \quad (164)$$

The important conclusions that can be drawn from Eq. (164) 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

$$I_B = \frac{dQ_{Te}}{dt} + \frac{dQ_{Tc}}{dt} + \frac{dQ_B}{dt} + \frac{Q_B}{\tau_B} \quad (165)$$

After a number of simplifying assumptions are made the following equation can be obtained:

$$t_r \approx h_{FE} \left(\frac{1}{\omega_T} + 1.7\, R_L\, C_{Tc}\right) \ln \frac{h_{FE}\, I_{B1}}{h_{FE} I_{B1} - 0.9\, I_C} \quad (166)$$

where ω_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:

$$I_B = \frac{Q_B}{\tau_B} + \frac{Q_{BS}}{\tau_s} + \frac{dQ_B}{dt} + \frac{dQ_{BS}}{dt} \quad (167)$$

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. (164), the equation may be rewritten in the following form:

$$I_{B2} = \frac{\tau_C\, I_C}{\tau_B} + \frac{Q_{BS}}{\tau_S} + \frac{dQ_{BS}}{dt} \quad (168)$$

If Eq. (168) is integrated with respect to time, the following equation is obtained:

$$\int_{I_{B1}-I_{B2}}^{0} \frac{dI_B}{I_{B2} - I_C\,(\tau_C/\tau_B) - (Q_{BS}/\tau_S)} = \int_0^{t_S} \frac{dt}{(dQ_{BS}/dI_B)} \quad (169)$$

With the additional information that τ_C/τ_B is equal to $1/h_{FE}$ and that dQ_{BS}/dI_B is equal to τ_S, the final expression for t_s becomes

$$t_S = \tau_S \ln \frac{I_{B1} - I_{B2}}{(I_C/h_{FE}) - I_{B2}} \quad \text{where } I_{B2} \text{ is negative} \quad (170)$$

The term τ_S in this equation is the storage time constant, and its value depends on the structure of the transistor. The equations used to determine τ_S for different transistor types are tabulated below:

1. For alloy-type transistors having no collector storage charge,

$$\tau_S = 2.4/\omega_b(1-\alpha) \quad (171)$$

where ω_b is the base cutoff frequency and α is the grounded-base current gain, both measured at the edge of saturation.

2. For diffused-base alloy transistors,

$$\tau_S = 3/\omega_b(1-\alpha) \quad (172)$$

3. For graded-base mesa or planar transistors,

$$\tau_S = (0.6/\omega_b) + (\tau_{mc}/2) \quad (173)$$

where τ_{mc} is the minority-carrier lifetime in the collector.

From the equation for τ_s, it is apparent that an increase in I_{B2} causes a decrease in τ_s, an increase in I_{B1} (more overdrive) causes an increase in τ_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:

$$t_f = h_{FE}\left(\frac{1}{\omega_T} + 1.7\ R_L\ C_{Tc}\right) \ln \frac{I_C - h_{FE}\ I_{B2}}{0.1\ I_C - h_{FE} I_{B2}} \quad (174)$$

Because I_{B2} is negative, Eq. (174) 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. 162. 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

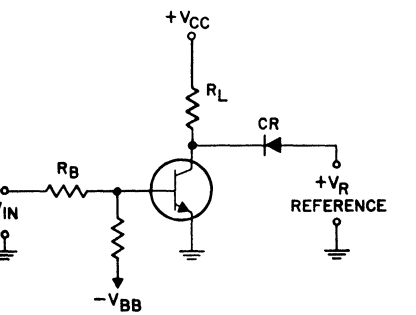


Figure 162. "Collector-catcher" circuit.

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.

The difficulty with this circuit is that the maximum I_C is essentially beta-dependent and, be-

cause 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. 163(a) and 163(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.

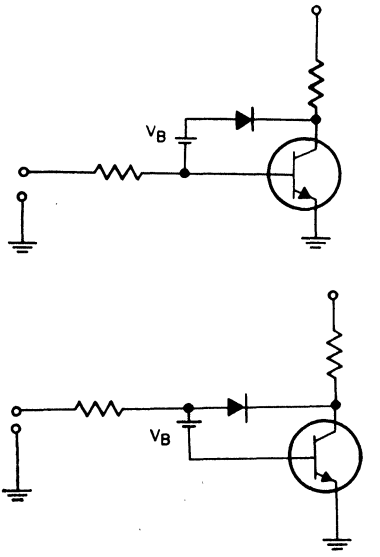


Figure 163. "Collector-catcher" circuits in which feedback is used.

A practical circuit is shown in Fig. 164. 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

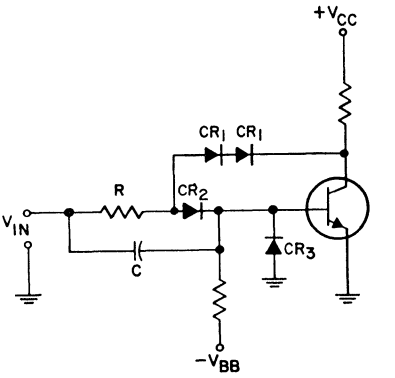


Figure 164. Practical "collector-catcher" circuit.

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 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. 165. Waveforms for this circuit are shown in Fig. 166.

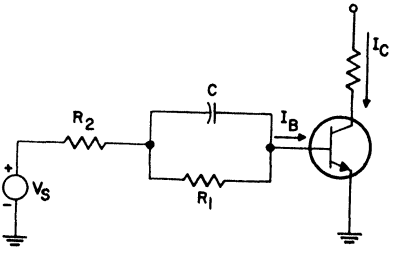


Figure 165. Circuit that uses speed-up capacitor to reduce transistor switching times.

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.

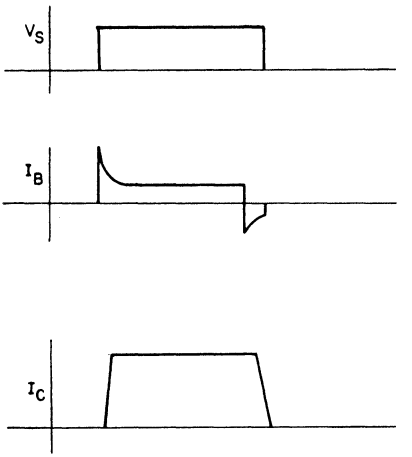


Figure 166. Waveforms for circuit shown in Fig. 165: (a) source voltage V_S; (b) base current I_B; (c) collector current I_C.

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:

$$P_{avg} = (1/T) \int_0^{t_1} V_{CE(sat)}\, I_{C(sat)}\, dt + (1/T) \int_{t_1}^{t_2} V_{CE}\,(t)_{off}\, I_C\,(t)_{off}\, dt + (1/T) \int_{t_2}^{t_3} V_{CE(off)}\, I_{C(off)}\, dt + (1/T) \int_{t_3}^{T} V_{CE}\,(t)_{on}\, I_C\,(t)_{on}\, dt \qquad (175)$$

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. (174) 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. 167), 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:

$$P_{avg} = \underbrace{\frac{1}{\tau}\int_0^{t_1} I_P\, V_{CE(sat)}\, dt}_{\text{ON}} + \underbrace{\frac{1}{T}\int_{t_1}^{t_2} I_p\left(1-\frac{t}{T_{sw}}\right) V_P\left(\frac{t}{T_{sw}}\right) dt}_{\text{TURN-OFF}}$$

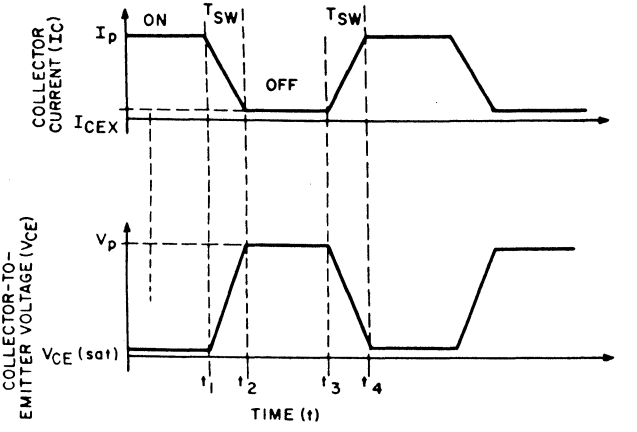


Figure 167. Idealized collector current and voltage waveforms for a transistor that has equal turn-on and turn-off currents.

$$+\underbrace{\frac{1}{T}\int_{t_2}^{t_3} I_{CEX}\, V_P dt}_{\text{OFF}}$$

$$+\underbrace{\frac{1}{T}\int_{t_3}^{t_4} I_p\left(\frac{t}{T_{sw}}\right) V_P\left(1-\frac{t}{T_{sw}}\right) dt}_{\text{TURN-ON}}$$

$$= \frac{I_P\, V_{CE(sat)}\, T_{on}}{\tau} + \frac{I_{CEX}\, V_P\, T_{off}}{\tau}$$

$$+ \frac{I_P V_P\, T_{sw}}{3\tau} \qquad (176)$$

From Eq. (176), 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. 168. The current-sensing resistor in the collector circuit should be non-inductive and should have a resistance value much

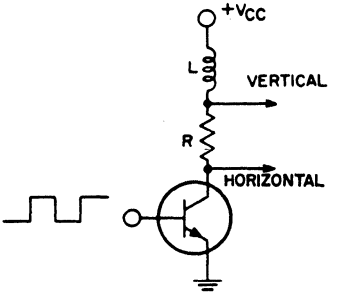


Figure 168. Test circuit used to determine transistor load lines.

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. 169(a). If the inductance is reduced so that $L(di/dt) < V_{CEX(sus)}$, the circuit has a load line as illustrated in Fig. 169(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. 169(a) is applicable.

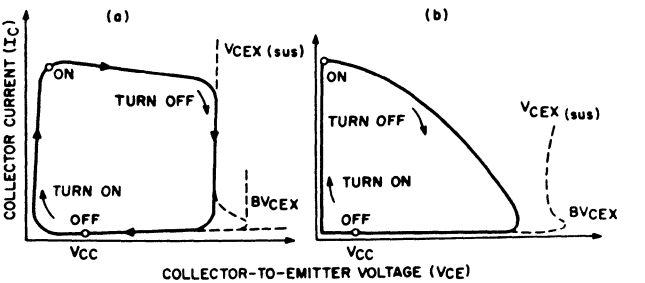


Figure 169. Inductive load line for the test circuit shown in Fig. 168; (a) typical load line; (b) load line when $L(di/dt) < V_{CEX(sus)}$.

Fig. 170 shows typical voltage and current curves as a function of time for this switch, with the load line indicated in Fig. 169(a). From these curves, the peak and average power dissipation, voltage limitations, and secondary-voltage-breakdown energy can be determined, as follows:

$$P_{pk} = V_{CEX}\,(sus)\,I_{pk} \qquad (177)$$

$$P_{avg} = \underbrace{\frac{1}{\tau}\int^{t_1} I_{CEX}\,V_{CC}\,dt}_{ON}$$

$$+ \underbrace{\frac{1}{\tau}\int_{t_1}^{t_2} \frac{I_{pk}}{2}\left(\frac{I_{pk}\,R_{sat}}{2}\right) dt}_{TURN\text{-}ON}$$

$$+ \underbrace{\frac{1}{\tau}\int_{t_2}^{t_3} I_{pk}\,(I_{pk}R_{sat})\,dt}_{OFF}$$

$$+ \underbrace{\frac{1}{\tau}\int_{t_3}^{t_4} \frac{I_{pk}}{2}\,(V_{CEX\,(sus)})\,dt}_{TURN\text{-}OFF}$$

$$\approx \frac{1}{\tau}\int_{t_3}^{t_4} \frac{I_{pk}}{2}\,(V_{CEX(sus)})\,dt \qquad (178)$$

Because I_{CEX} and R_{sat} are small, the following approximation is valid:

$$t_4 - t_3 = t_{off}$$

$$= \int_{I_{CBX}}^{I_P} \frac{L\,di}{V_{CEX(sus)} - V_{CC}}$$

$$\approx \frac{LI_P}{V_{CEX(sus)} - V_{CC}} \qquad (179)$$

The average power dissipated in the switch shown in Fig. 168 is then given by

$$P_{avg} = \frac{t_{off}}{\tau}\left(\frac{I_P}{2}\right) V_{CEX(sus)}$$

$$= \frac{1}{2}\left(\frac{LI_P^2}{\tau}\right)\left[\frac{V_{CEX(sus)}}{V_{CEX(sus)} - V_{CC}}\right] \text{watts} \qquad (180)$$

The turn-off energy is expressed by the following equation:

$$Et_{off} = \frac{1}{2}\,(LI_p^2)\left[\frac{V_{CEX(sus)}}{V_{CEX(sus)} - V_{CC}}\right] \qquad (181)$$

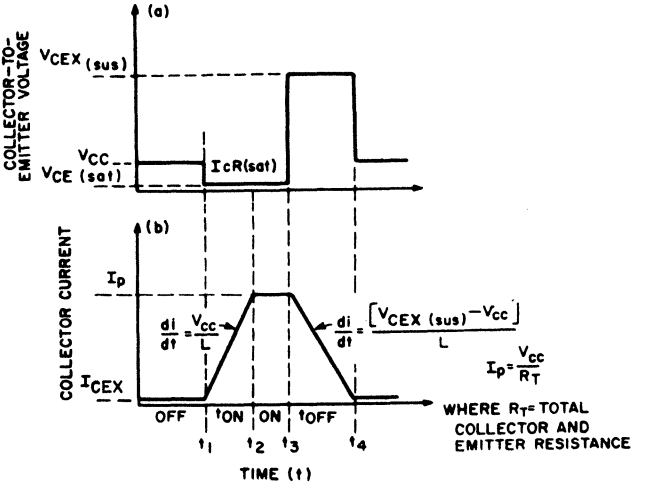


Figure 170. Voltage and current waveforms for the switching circuit shown in Fig. 168.

This energy must not exceed the $E_{S/b}$ rating of the transistor.

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 case 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. 171(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. 171(b), 171(c) and 171(d).

In the circuits of Fig. 171, 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. 171(b), 171(c), and 171(d)].

Five theoretical cases—The five cases to be analyzed are shown in Fig. 172. The energy equations for each case and corresponding current and voltage

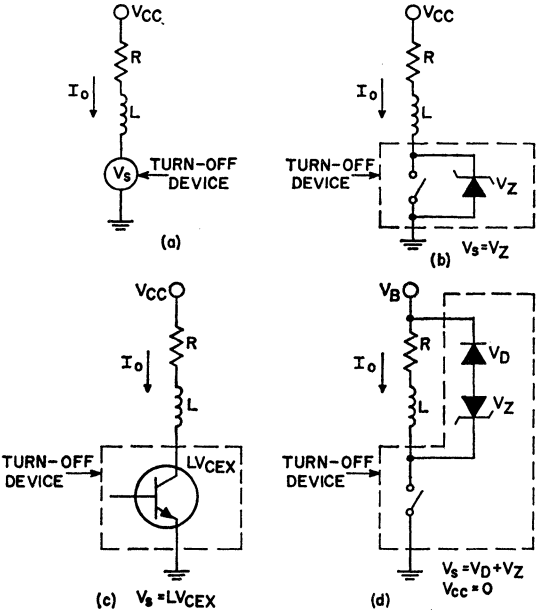


Figure 171. Basic switching circuits with inductive loads.

waveforms for the switching device are also shown.

In case 1, shown in Fig. 172(a), the energy is given by the familiar expression for energy stored in an inductor, as follows:

$$E_1 = (1/2)\, L\, I_0^2 \qquad (182)$$

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. 172(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:

$$E_2 = (1/2)\, L\, I_0^2 \left(\frac{V_s}{V_s - V_{CC}}\right) \qquad (183)$$

To show that case 2 is a modification of case 1, a multiplying factor k_2 can be defined as follows:

$$E_2 = (1/2)\, L\, I_0^2\, k_2 \qquad (183a)$$

where the factor k_2 is given by

$$k_2 = \frac{V_s}{V_s - V_{CC}} = \frac{1}{1 - \rho} \qquad (183b)$$

Fig. 173 shows a curve of k_2 as a function of the ratio V_{CC}/V_s.

Case 3, shown in Fig. 172(c), causes an energy dissipation E_3 given by the following expression:

$$E_3 = I_0\, V_s \frac{L}{R} \left[\frac{\ln\left(\dfrac{1}{1 + (I_0 R)/(V_s - V_{CC})}\right)}{(I_0 R)/(V_s - V_{CC})} + 1\right] \qquad (184)$$

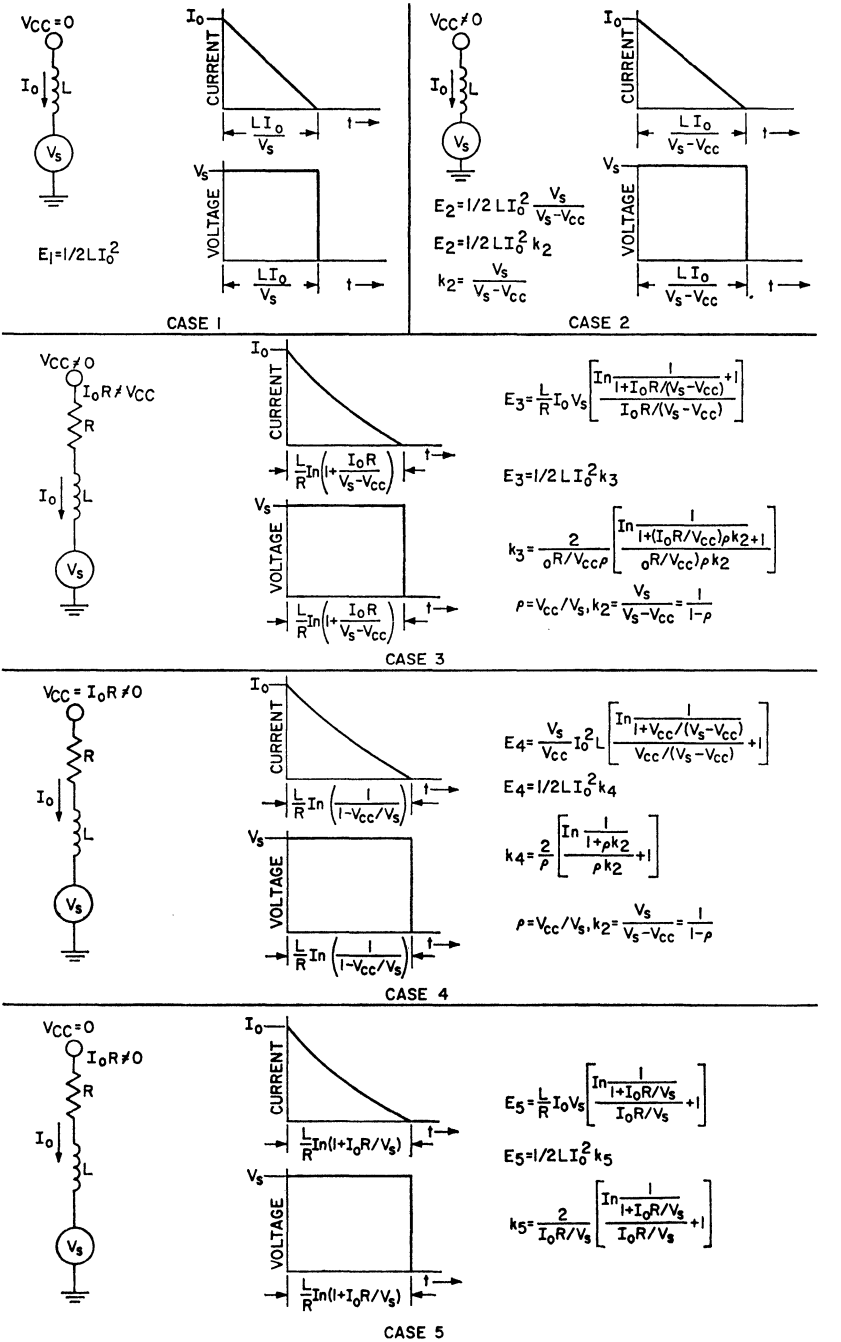


Figure 172. Equivalent circuits, energy equations, and voltage and current waveforms for the five basic circuit configurations.

Eq. (184) can be rearranged, as was Eq. (183), 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_0R/V_{CC}. Eq. 184 then becomes:

$$E_3 = (1/2)\, L\, I_0^2\, k_3 \qquad (184a)$$

where the factor k_3 is defined as follows:

$$k_3 = f\,(I_0R/V_{CC}\,,\ V_{CC}/V_s) = \frac{2}{(I_0R/V_{CC})_\rho}\left[\frac{\ln\left(\frac{1}{1+(I_0R/V_{CC})_\rho\, k_2}\right)}{(I_0R/V_{CC})_\rho\, k_2}+1\right] \qquad (184b)$$

Fig. 173 shows curves of k_3 as a function of the ratio V_{CC}/V_s for several values of I_0R/V_{CC}. When the ratio I_0R/V_{CC} becomes less than 0.01, k_3 can be approximated by k_2. Case 3 then reduces to case 2.

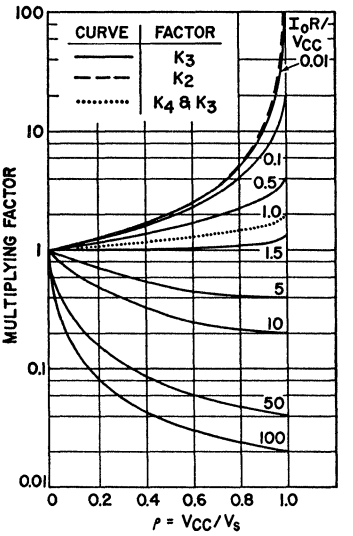


Figure 173. Multiplying factors (k_2, k_3, k_4) as a function of the ratio V_{CC}/V_s.

Case 4, shown in Fig. 172(d), is a special form of case 3. The condition $I_0R = V_{CC}$ 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:

$$E_4 = \frac{V_s}{V_{CC}}\, I_0^2\, L \left[\frac{\ln\left(\frac{1}{1+V_{CC}/(V_s-V_{CC})}\right)}{V_{CC}/(V_s-V_{CC})}+1\right] \qquad (185)$$

Again, it is convenient to define a multiplying factor k_4 which is a function of the ratio V_{CC}/V_s, as follows:

$$E_4 = (1/2)\, L\, I_0^2\, k_4 \qquad (185a)$$

where

$$k_4 = f\,(\rho) = \frac{2}{\rho}\left[\frac{\ln\frac{1}{(1+\rho\, k^2)}}{\rho\, k_2}+1\right] \qquad (185b)$$

Fig. 173 includes a curve of k_4 plotted as a function of ρ. This curve corresponds to the curve of k_3 for $I_0R/V_{CC} = 1$. As V_{CC} approaches V_s, the multiplying factor approaches its maximum value of two. Therefore, for any inductive switch in which $V_{CC} = I_0R$, the maximum energy that must be handled by the switching device is $L\, I_0^2$, or twice the energy of $(1/2)\, L\, I_0^2$ for case 1.

Case 5, shown in Fig. 172(e), can also be regarded as a special form of case 3. A practical example of case 5 is shown in Fig. 171(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

$$E_5 = V_s I_0 \frac{L}{R} \times \left[\frac{\ln \frac{1}{1 + (I_0R/V_s)}}{(I_0R/V_s)} + 1 \right] \quad (186)$$

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. 173. 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. 186 becomes

$$E_5 = (1/2)\ L\ I_0^2\ k_5 \quad (186a)$$

where

$$k_5 = f\ (I_0R/V_s) = \frac{2}{(I_0R/Vs)} \left[\frac{\ln \frac{1}{1 + (I_0R/V_s)}}{(I_0R/V_s)} + 1 \right] \quad (186b)$$

Fig. 174 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_3 could also have been defined in terms of ρ and I_0R/V_s, instead of ρ 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

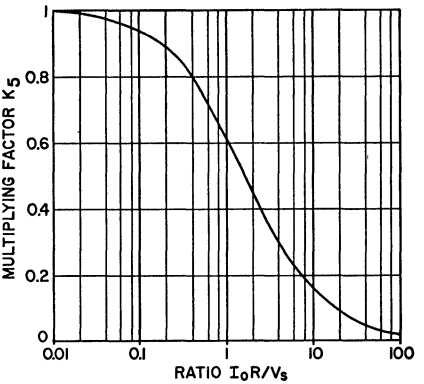


Figure 174. Multiplying factor k_5 as a function of the ratio I_0R/V_s.

constants k_2, k_3, k_4 and k_5 can be easily determined from Figs. 173 and 174 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. 175 are identical except that R_2 in Fig. 175(b) is greater than R_1 in Fig. 175(a).

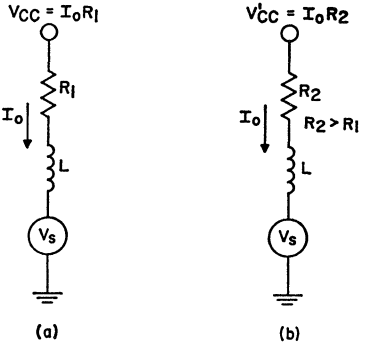


Figure 175. Circuits used in Example 1 to demonstrate the use of case-4 equations.

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. 173 indicates that as V_{CC}/V_s increases (i.e., as V_{CC} increases), k_4 also increases. Because $E_4 = (1/2)\ L\ I_o^2\ k_4$, 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. 175, 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. 176 demonstrate the difference between cases 3 and 4.

The circuit of Fig. 176(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.

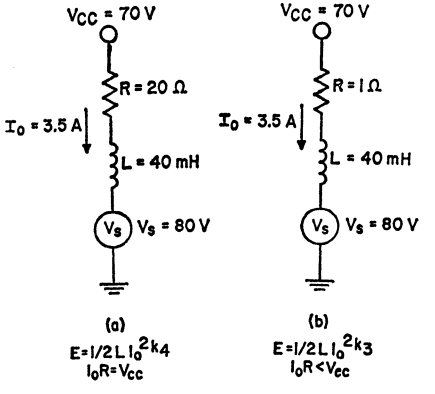


Figure 176. Circuits used in Example 2 to show the difference between cases 3 and 4.

The circuit of Fig. 176(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. 176(a) is calculated as follows: The ratio $V_{CC}/V_s = 70/80 = 0.875$. The corresponding k_4 multiplying factor (from Fig. 173) is approximately 1.6. The energy E_4 is then given by

$$E_4 = (1/2)\ L\ I_o^2$$
$$= (1/2)\ (40 \times 10^{-3})\ (3.5)^2\ (1.6)$$
$$= 400 \text{ millijoules}$$

[Calculation of E_4 directly from Eq. (185) yields a value of 393 millijoules.]

To find the energy requirement for the circuit of Fig. 176(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. 173 the interpolated value for k_3 is approximately 6.5. From Eq. (184a), the energy for this circuit is determined as follows:

$$E_3 = (1/2)\ L\ I_o\ k_3$$
$$= 1/2\ (40 \times 10^{-3})\ (3.5)^2\ (6.5)$$
$$= 1590 \text{ millijoules}$$

[Calculation of E_3 directly from Eq. (184) yields a value of 1570 millijoules.]

It should be noted that the energy requirement for the circuit of Fig. 176(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. 177 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.

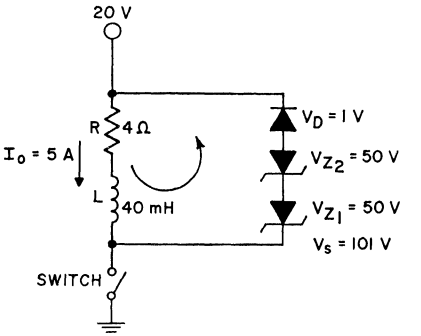


Figure 177. Circuits used in Example 3 to demonstrate the use of case-5 equations.

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. 174. The ratio $I_0R/V_s = 20/101$, or approximately 0.2; from Fig. 174, 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. 178(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. 178(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. 178(b) is similar to case 3. The equivalent circuit is shown in Fig. 178(c).

First the ratios V_{CC}/V_s and I_0R/V_{CC} are calculated as follows:

$$V_{CC}/V_s = 80/100 = 0.8$$

$$I_0R/V_{CC} = \frac{(15)\ (0.2)}{80}$$

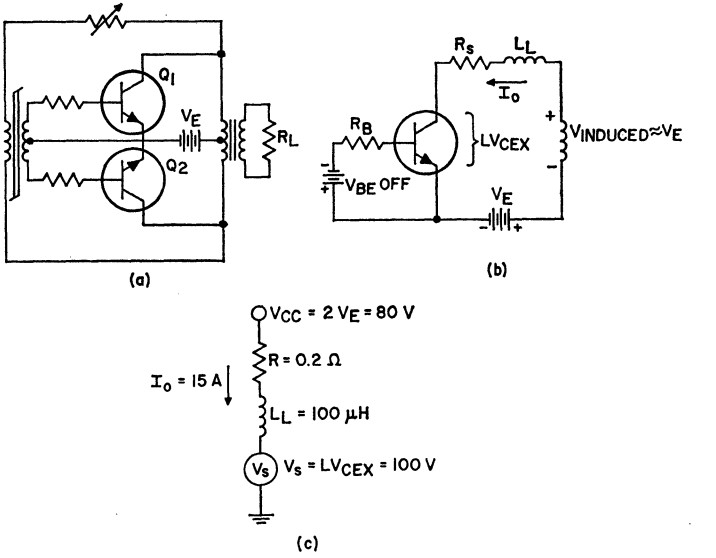


Figure 178. Typical power inverter circuit and equivalent circuits used in Example 4 as an application of case-3 equations.

$$I_0R/V_{CC} = \frac{3.0}{80} = 0.0375$$

Then interpolating from Fig. 173, 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. (184) gives 46 millijoules.]

The energy requirement for the transistor Q_I in Fig. 178(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.

Rectification

THE current and voltage relationships for silicon rectifiers vary for different types of circuit configurations. The particular circuit in which a rectifier is used is chosen on the basis of the requirements for a specific application.

Silicon rectifiers are used in a continually broadening range of applications. Originally developed for use in such equipments as dc-to-dc converters, battery chargers, mobile power supplies, radio and TV transmitters, and electroplating service, silicon rectifiers are also used in power supplies for radio and television receivers and phonograph amplifiers, as well as in such applications as line-type modulators, hold-off and charging diodes, pulse-forming networks, and brushless alternators. They are also being used in many aircraft applications because of their small size, light weight, and high efficiency.

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. 179 through Fig. 185 show seven basic rectifier configurations. (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. 179 delivers only one phase of current for each cycle of ac input voltage. As shown by the current waveform, the single rectifier conducts the entire current flow. This type of circuit contains a very high percentage of output ripple.

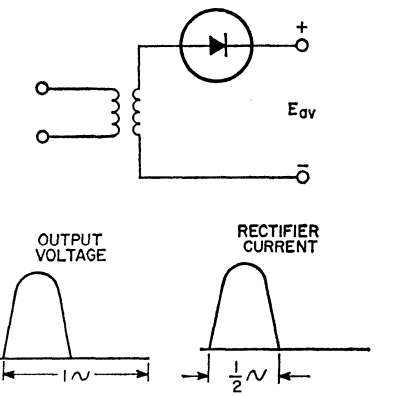


Figure 179. Single-phase half-wave circuit.

Fig. 180 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. 179 and about 65 per cent less ripple. Only 50 per cent of the total current flows through each rectifier. This type of circuit is widely used in television receivers and large audio amplifiers.

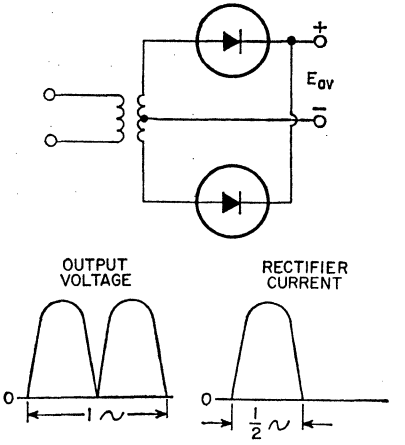


Figure 180. Single-phase full-wave circuit with center-tapped power transformer.

The single-phase full-wave bridge circuit shown in Fig. 181 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 as the circuit of Fig. 180 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. 182 through 185 are usually found in heavy industrial equipment such as high-power transmitters. The three-phase

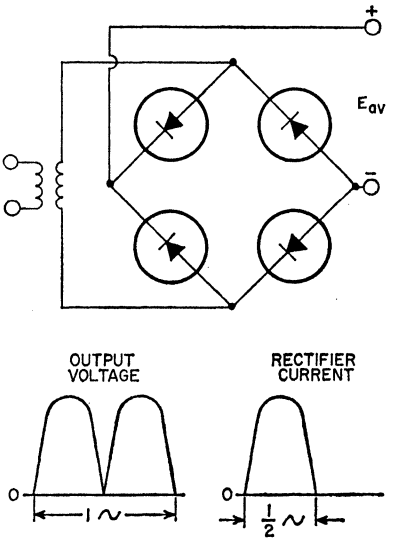


Figure 181. Single-phase full-wave circuit without center-tapped power transformer (i.e., bridge-rectifier circuit).

Y half-wave circuit shown in Fig. 182 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. 183 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. 182 for the same transformer conditions. In addition, this circuit, as well as those shown in Figs. 184 and 185, has an extremely small percentage of ripple.

In the six-phase "star" circuit shown in Fig. 184, which also uses six rectifiers, the least amount of the total output current (one-sixth) flows through each output rectifier. The three-phase double-Y and interphase transformer circuit shown in Fig. 185 uses six half-wave rectifiers in parallel.

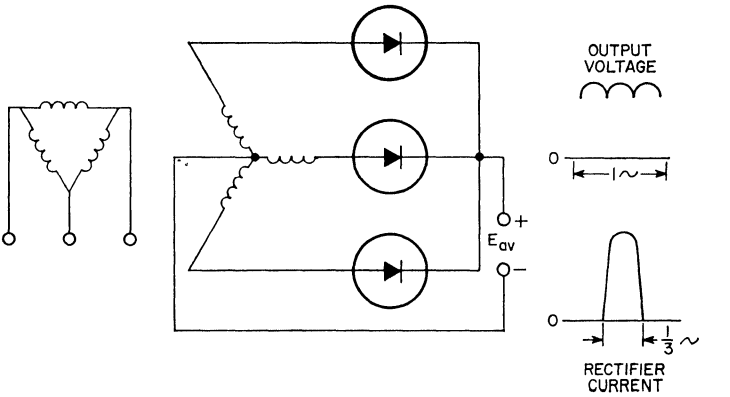


Figure 182. Three-phase "Y" half-wave circuit.

This arrangement delivers six current pulses per cycle and twice as much output current as the circuit shown in Fig. 182.

Table XII lists voltage and current ratios for the circuits shown in Figs. 179 through 185 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. 179. 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 XII can be used to determine the parameters and characteristics of the circuit.

In Table XII, all ratios are shown as functions of either the average output voltage E_{av} or the

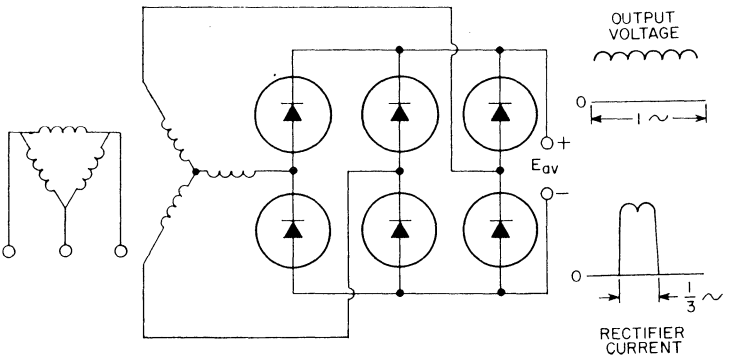


Figure 183. Three-phase "Y" full-wave circuit.

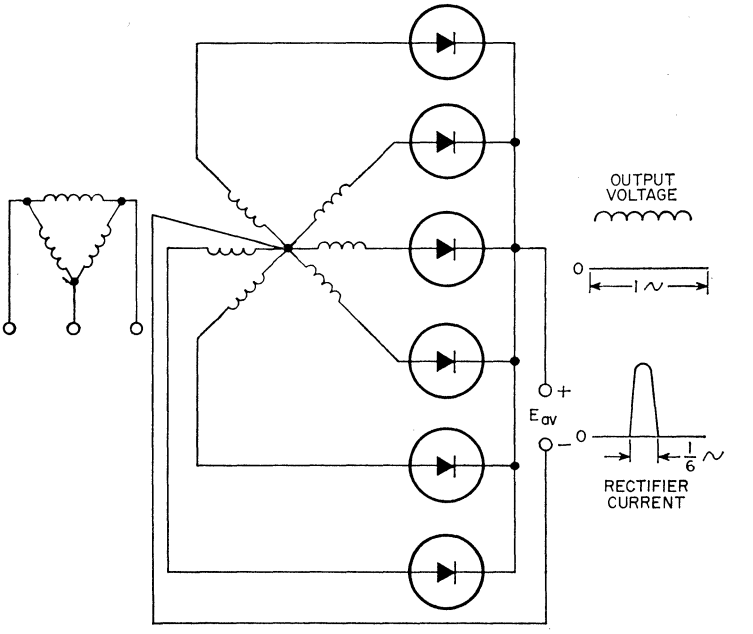


Figure 184. Six-phase "star" circuit.

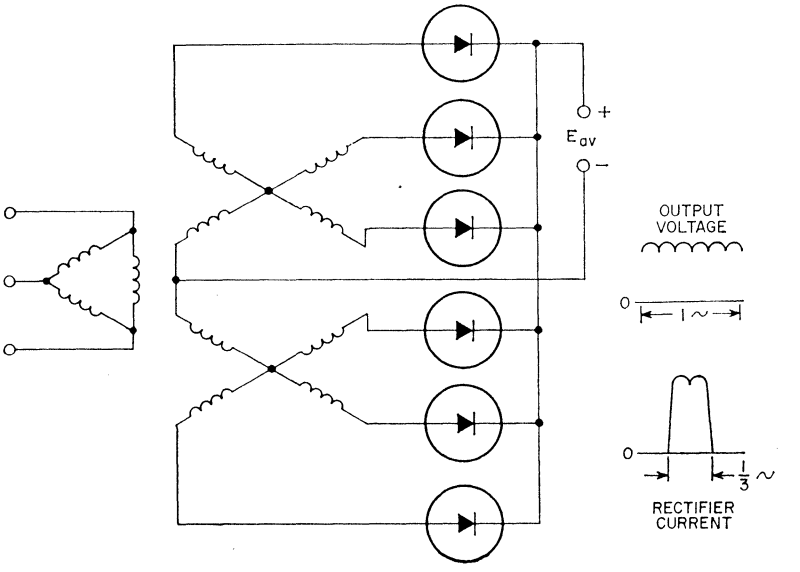


Figure 185. Three-phase "double-Y" and interphase-transformer circuit.

Table XII—Voltage and Current Ratios for Rectifier Circuits Shown in Figs. 179 Through 185. Fig. 179 Uses a Resistive Load, and Figs. 180 Through 185 an Inductive Load

CIRCUIT RATIOS	Fig. 179	Fig. 180	Fig. 181	Fig. 182	Fig. 183	Fig. 184	Fig. 185
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

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 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. 186.

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. 187.

Higher average dc output volt-

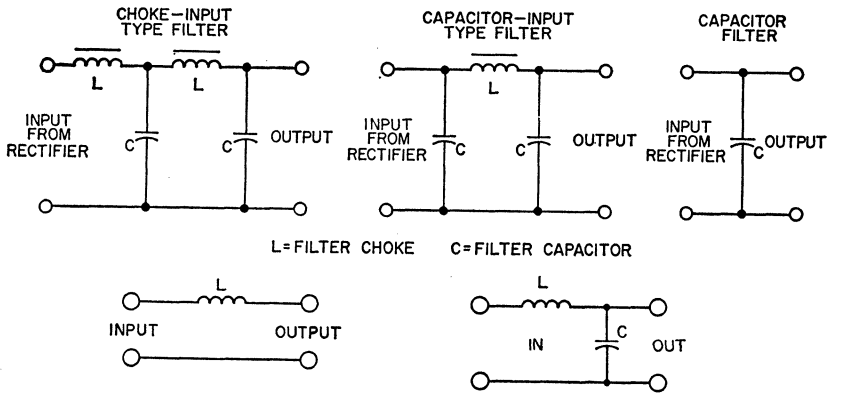


Figure 186. Typical filter circuits.

ages 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 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. 186 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.

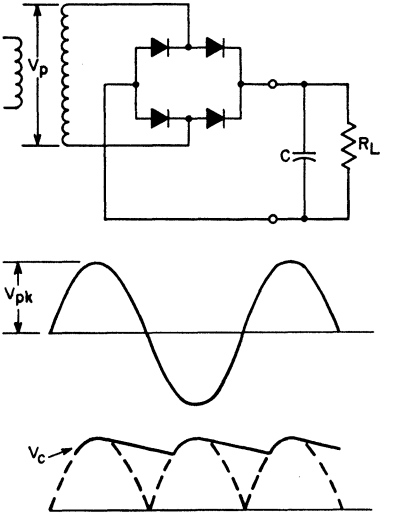


Figure 187. Bridge-rectifier circuit with capacitor input filter.

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. 188 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. 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. 189 and 190. 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.

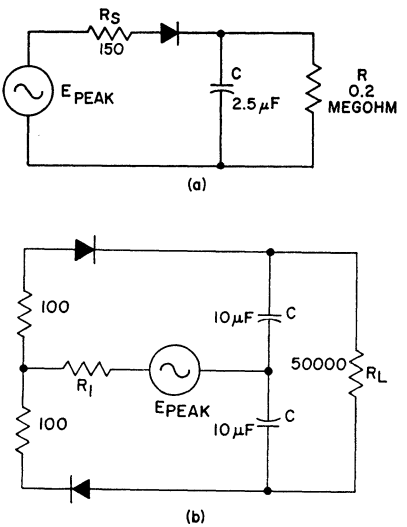


Figure 188. Typical rectifier circuits using capacitive loads: (a) half-wave rectifier circuit; (b) voltage doubler.

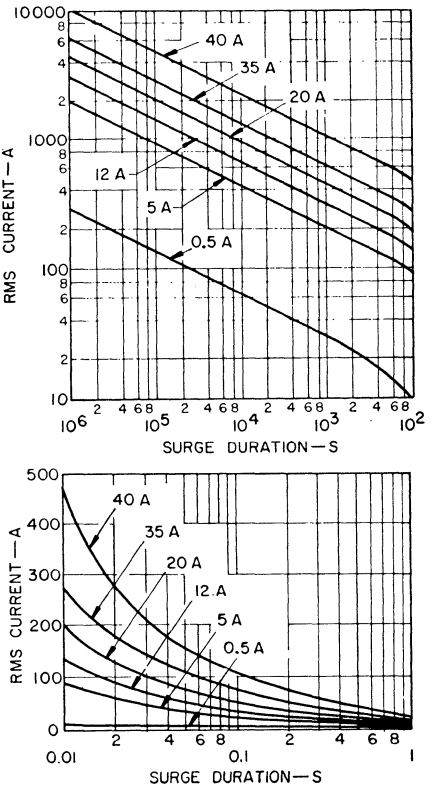


Figure 189. Universal surge rating charts for RCA rectifiers.

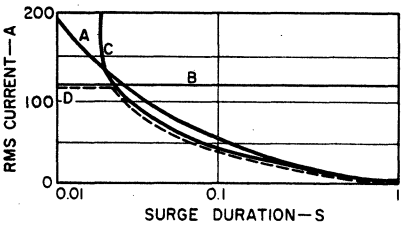


Figure 190. Typical coordination chart for determination of fusing requirements: Curve A—surge rating for 20-ampere rectifier; Curve B—expected surge current in half-wave circuit; Curve C—opening characteristics of protective device; Curve D—resulting surge current in modified circuit.

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:

$$I_{rms} = 0.7\ (E_{peak}C/R_sC) = 0.7\ (E_{peak}C/t) \qquad (187)$$

and

$$I_{rms}t = 0.7\ E_{peak}\ C \qquad (188)$$

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. 188(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:

$$I_{rms}t = 0.7\ (4950)\ (2.5 \times 10^{-6}) = 0.0086$$

When this value is plotted on the surge-rating chart of Fig. 191, the resulting line intersects the rectifier rating curve at 3.3×10^{-4} second. The minimum limiting resistance which affords adequate surge protection is then calculated as follows:

$$R_sC \geq 3.3 \times 10^{-4}$$

$$R_s \geq \frac{3.3 \times 10^{-4}}{2.5 \times 10^{-6}} = 132 \text{ ohms}$$

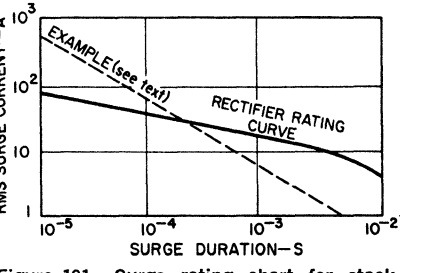


Figure 191. Surge rating chart for stack rectifier CR210.

Therefore the value of 150 ohms shown for R_s in Fig. 188(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. 192 and 193 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. 177 as functions of the circuit constants $n\omega CR_L$ and R_s/nR_L. The quantity ωCR_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. 193, 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. 193 merely indicates a higher peak and higher rms current than will actually flow in the circuit, i.e., the rectifiers will operate more conservatively

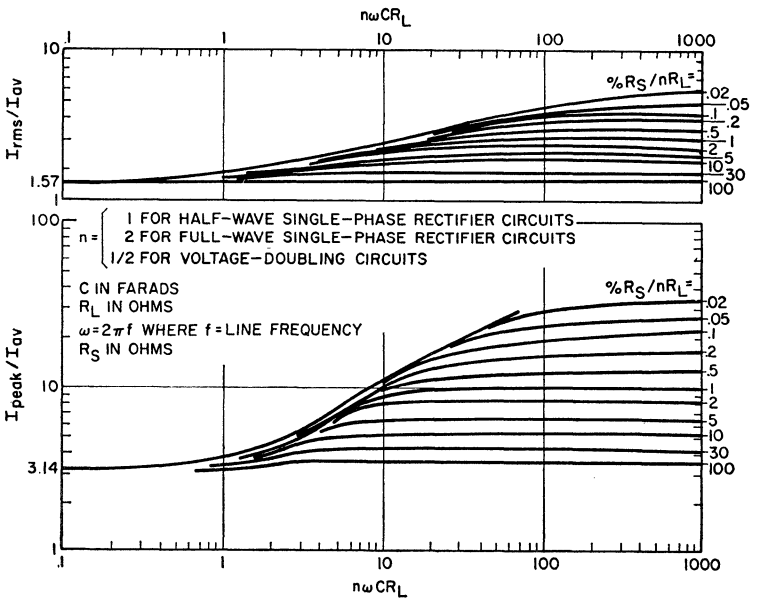


Figure 192. Relationship of peak, average, and rms rectifier currents in capacitor-input circuits.

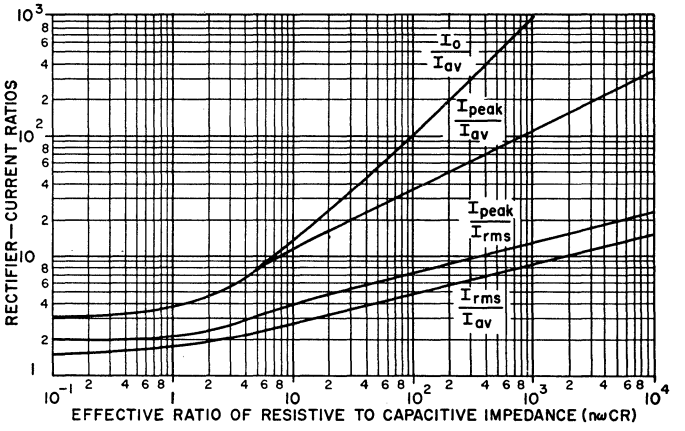


Figure 193. Forward-current ratios for rectifiers in capacitor-input circuits in which R_s is much less than 1/C.

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. 192 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. 194, 195, and 196, respectively. Fig. 197 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. 188(a), the resistive-to-capacitive reactance ωCR_L is given by

$$\omega CR_L = 2\pi \times 60 \times 2.5 \times 10^{-6} \times 200{,}000 = 189$$

For an exact solution using Fig. 192, the ratio of R_s to R_L is first calculated as follows:

$$\frac{R_s}{R_L} = \frac{150}{200{,}000} = 0.075$$

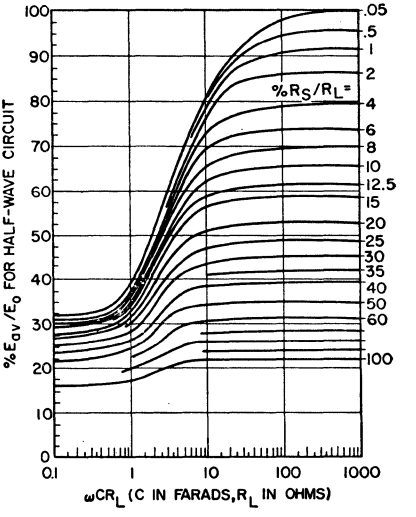


Figure 194. Relationship of applied ac peak voltage to dc output voltage in half-wave capacitor-input circuit.

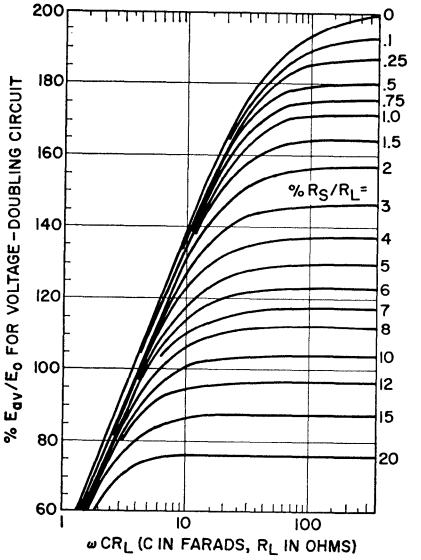


Figure 195. Relationship of applied ac peak voltage to dc output voltage in capacitor-input voltage-doubler circuit.

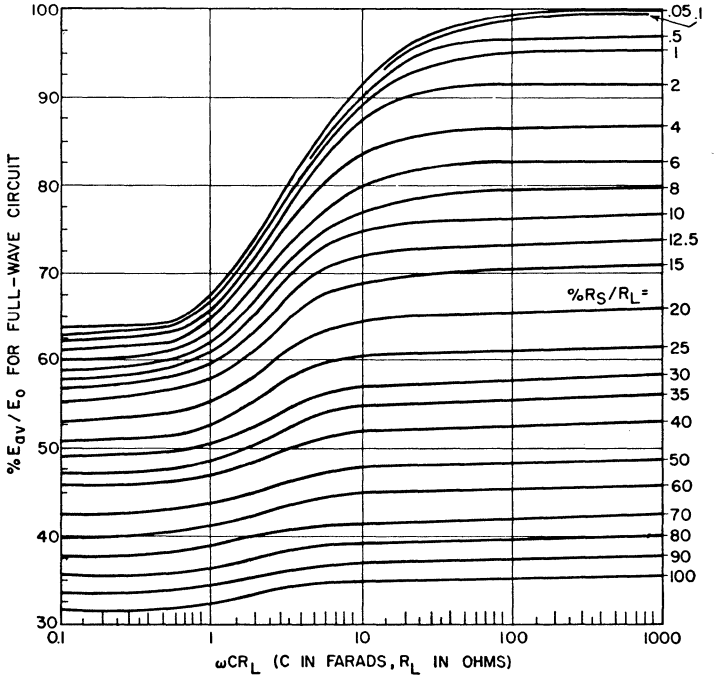


Figure 196. Relationship of applied ac peak voltage to dc output voltage in full-wave capacitor-input circuit.

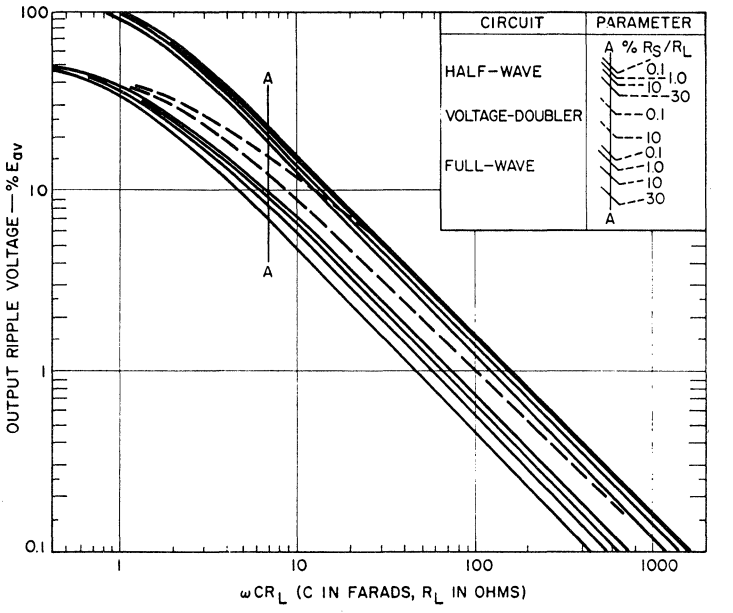


Figure 197. RMS ripple voltage in capacitor-input circuits.

The values for ωCR_L and R_s/R_L are then plotted in Fig. 194 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. 192, 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. 193, 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. 193, 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 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 ωCR on Figs. 192 and 193), 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. 198 shows an example of this conversion for RCA stack rectifier rating curves.

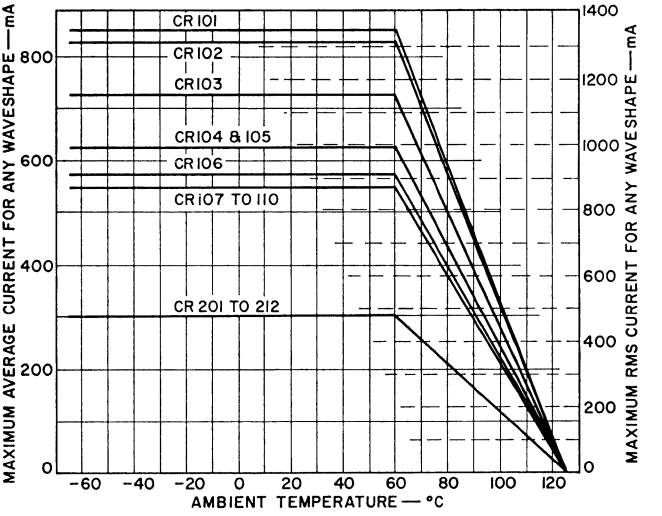


Figure 198. Current-temperature ratings for silicon stack rectifiers.

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

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. 199 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

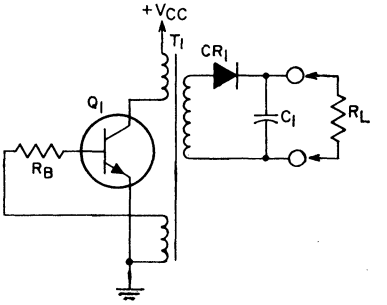


Figure 199. Basic circuit configuration for a ringing-choke dc-to-dc converter.

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_I 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. 200 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

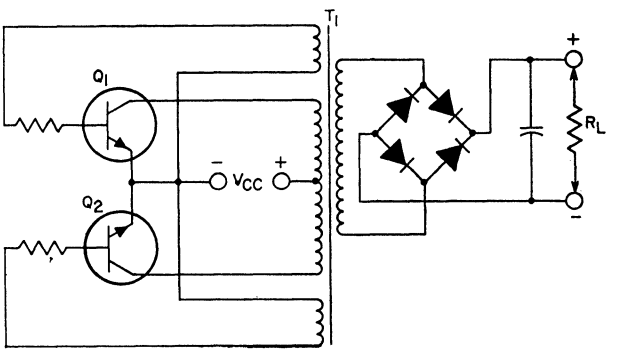


Figure 200. Basic circuit configuration of a single-transformer push-pull switching converter.

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. 201 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. 200.

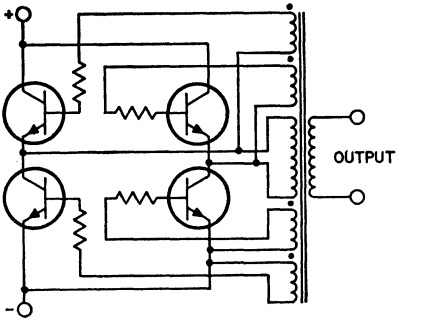


Figure 201. Basic circuit configuration of a four-transistor, single-transformer bridge inverter.

Fig. 202 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. 200. The saturable-transformer technique may also be applied in the design of a bridge converter, as shown in Fig. 203.

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

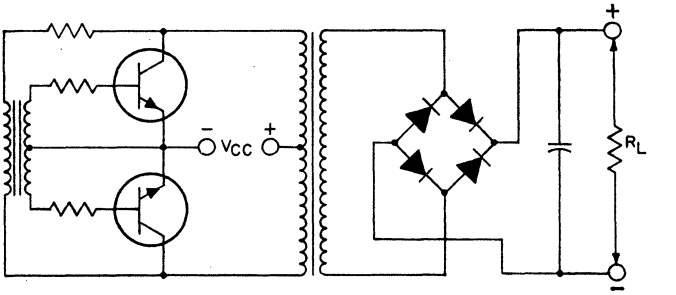


Figure 202. Basic circuit configuration of a two-transformer push-pull switching converter.

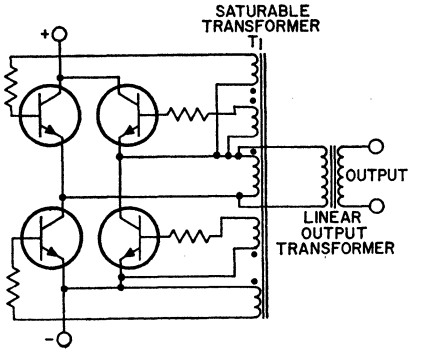


Figure 203. Basic configuration of a four-transistor bridge inverter that uses a saturable output transformer.

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 ensure 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 XIII through XVI 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:

$$T_{J(AV)} = T_A + P_{AV}\,\theta_{J-C} + \theta_{C-A} \quad (189)$$

The average junction temperature calculated by use of Eq. (189) 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. 204(a) shows the basic configuration for a practical circuit of a ringing-choke converter, which is basically a one-transistor, one-transformer circuit. Fig. 204(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. 204(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.

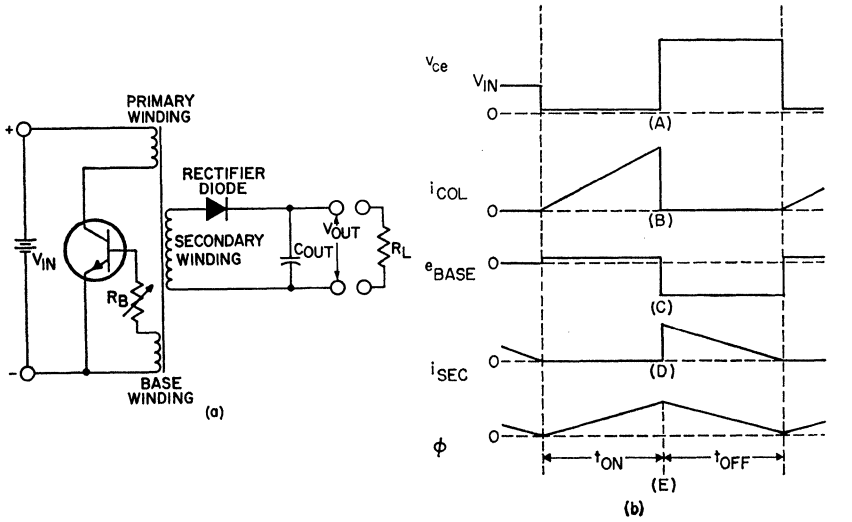


Figure 204. Ringing-choke converter circuit: (a) Schematic diagram; (b) Typical operating waveforms in a ringing-choke converter—(A) primary voltage; (B) primary current; (C) base-to-emitter voltage; (D) secondary current; (E) magnetic flux in transformer core.

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. 204(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. 204(a). All references to time periods and waveforms apply to those shown in Fig. 204(b).

The average value of the current in the primary winding of the transformers $I_{P(AV)}$ may be expressed as follows:

$$I_{P(AV)} = I_{out} (V_{out}/\eta V_{in}) \quad (190)$$

where I_{out} is the load current in amperes, V_{out} is the dc output voltage in volts, η is the circuit-efficiency factor as given in Table XIII, 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. 204(b), is expressed by the following equation:

$$\hat{I}_P = 2I_{P(AV)} (T/t_{on}) = (2T/\eta\, t_{on}) (V_{out}/V_{in})\, I_{out} \quad (191)$$

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:

$$L_P (dI_P/dt) = V_{in} \quad (192)$$

For the triangular current waveform, waveform B in Fig. 204(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}:

$$dI_P/dt = \hat{I}_P/t_{on} \quad (193)$$

When this relationship [Eq. (193)] is used, Eq. (192) may be rewritten in the following form:

$$L_P = (V_{in}/\hat{I}_P)\, t_{on} \quad (194)$$

With $\hat{I}_p$ defined as indicated in Eq. (191), the primary inductance L_p may then be expressed as follows:

$$L_P = \frac{\eta\, V_{in}^2\, t_{on}^2}{2\, V_{out}\, I_{out}\, T} = \left(\frac{\eta\, V_{in}^2}{2\, P_{out}}\right)\left(\frac{t_{on}^2}{T^2}\right)\left(\frac{1}{f}\right) \quad (195)$$

where f is the operating frequency in hertz, and is equal to 1/T.

Before Eq. (195) can be evaluated, the ratio t_{on}/T must be determined. From Eq. (194) the conduction period of the transistor t_{on} is given by

$$t_{on} = L_P\hat{I}_P/V_{in} \quad (196)$$

An equation for the OFF period (t_{off}) can be derived from the following relationship:

$$V'_{out} = L_P\, (dI_s'/dt) \quad (197)$$

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:

$$t_{off} = L_P\, (\hat{I}_P/V'_{out}) \quad (198)$$

The total period of oscillation ($T = t_{on} + t_{off}$), as determined from the summation of Eqs. (196) and (198), is given by the following equation:

$$T = L_P\hat{I}_P\left(\frac{1}{V_{in}} + \frac{1}{V'_{out}}\right) \quad (199)$$

The ratio of ON time to total period of oscillation, therefore, can be expressed as follows:

$$\frac{t_{on}}{T} = \frac{L_P\hat{I}_P\left(\frac{1}{V_{in}}\right)}{L_P\,\hat{I}_P\left(\frac{1}{V_{in}} + \frac{1}{V'_{out}}\right)} = \frac{V'_{out}}{V_{in} + V'_{out}} \quad (200)$$

This relationship is used in Eq. (195) to obtain the following expression for the required primary inductance:

$$L_P = \left(\frac{\eta\, V_{in}^2}{2\, P_{out}\, f}\right)\left[\frac{(N_P/N_S)^2\, V_{out}^2}{V_{in}^2 + (N_P/N_S)^2 + 2\, V_{in}\, V_{out}\, (N_P/N_S)}\right] \quad (201)$$

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

$$\frac{\Delta\phi}{A} = \hat{I}_P\left(\frac{4\pi\, N_P}{10}\right)\left[\frac{1}{l_i/\mu_i + l_a/\mu_a}\right] \quad (202)$$

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, μ_i is the permeability of the core material, and μ_a is the permeability of air ($\mu_a = 1$).

The maximum permissible flux-density swing may also be expressed as follows:

$$\frac{\Delta\phi}{A} = \frac{\phi_{(max)}}{A} - \frac{\phi_{(res)}}{A} \qquad (203)$$

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:

$$\Delta\phi = (L_P \hat{I}_P/N_P) \times 10^8 \qquad (204)$$

or

$$L_P = (N_P \Delta\phi \times 10^{-8})/\hat{I}_P \qquad (205)$$

Eqs. (202), (204), and (205) can be combined to provide the following expression for the required primary inductance:

$$L_P = \left(\frac{4\pi N_P^2}{10^9}\right)\left[\frac{1}{(l_i/\mu_i) A + (l_a/\mu_a)}\right] \qquad (206)$$

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. (204).

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:

$$V_P = V_{in} - \hat{I}_P (R_{sat} + R_P + R_{supply}) \qquad (207)$$

The required number of turns for the base winding is given by

$$N_B = \frac{N_P V_B}{V_{in} - \hat{I}_P (R_{sat} + R_P + R_{supply})} \qquad (208)$$

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. (208) may be rewritten to express N_B in terms of V_{BE} as follows:

$$N_B = \frac{2N_P V_{BE(max)}}{V_{in} - \hat{I}_P (R_{sat} + R_P + R_{supply})} \qquad (209)$$

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:

$$V_{CB(max)} = V_{in} + (N_P/N_S) V_{out} \qquad (210)$$

or

$$N_S = \frac{N_P V_{out}}{V_{CB(max)} - V_{in}} \qquad (211)$$

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:

$$R_B = V_{BE(max)}/I_B \quad (212)$$

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:

$$\hat{I}_S = \hat{I}_P (N_P/N_S) \quad (213)$$

$$\hat{V}_S = (V_{CB(max)} - V_{in}) (N_S/N_P) \quad (214)$$

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:

$$R_L = V_{out}^2/P_{out}$$
$$C_{out}R_L = 10\ t_{on}$$
$$C_{out} = (10\ t_{on}\ P_{out})/V_{out}^2 \quad (215)$$

The optimum ratio of primary-to-secondary winding space K^2 for the transformer is determined from the following expression:

$$K = \frac{1}{\sqrt{(N_S/N_P)\ (V_{in}/V_{out})+1}} \quad (216)$$

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 XIII 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 XIII 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 XIII—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_{CE}(max) (V)	Min. I_C(pk) (A)	Min. P_D* (W)	Area A (cm^2)	Length l_i (cm)	Factor η
1	250	6-10	5	25	0.5	0.1	0.5-1.5	2.5-10	0.75
	500	10-15	10	35	0.4	0.08			
	750	15-20	20	45	0.3	0.07			
5	250	6-12	1	30	3	1.5	0.5-5	2.5-12	0.75
	500	12-20	2	45	2	1			
	750	20-28	8	60	1	0.5	0.5-5	2.5-12	0.7
10	300	6-12	0.8	30	6	3	1-7.5	2.5-15	0.7
	500	12-18	1	45	4	2			
	750	18-28	1.2	60	2	1			
25	400	10-18	0.5	45	10	10	1-10	5-15	0.65
	600	18-26	0.8	60	6	5			
	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 XIII 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 η. The data in Table XIV 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 XIV—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. (211) 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 XIII reduces the required step-up ratio.]

2. The value determined for N_s/N_P and Eq. (201) 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. (206) 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. (200), 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. (191), 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. (204) 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 XV 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 XV—Typical Values Of Magnetic Parameters For Commercial Transformer-Core Materials.

Material	Maximum Permeability (μ_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. (209), 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. (212), 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. (211), 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. (213) and (214), 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. (215)]:

$$C_{out} \geq (10\ t_{on}\ P_{out})/V_{out}^2$$

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. (216), 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. 205 shows a push-pull, transformer-coupled, dc-to-dc converter that uses one transformer and two transistors. Fig. 206 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. 206. At the start of the

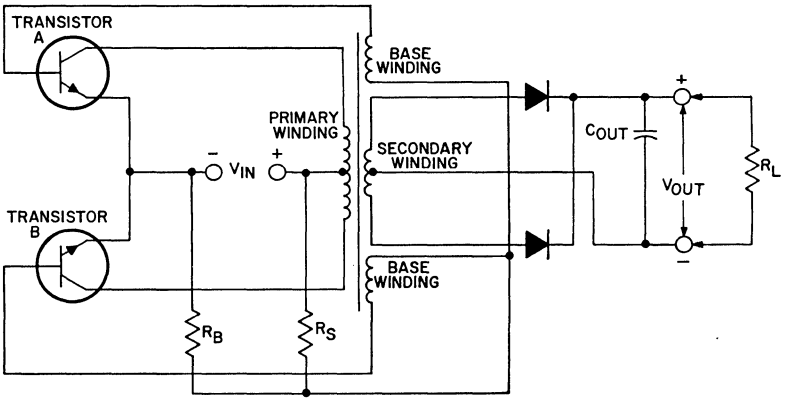


Figure 205. Two-transistor, one-transformer push-pull switching converter.

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 ensures 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. 206, 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

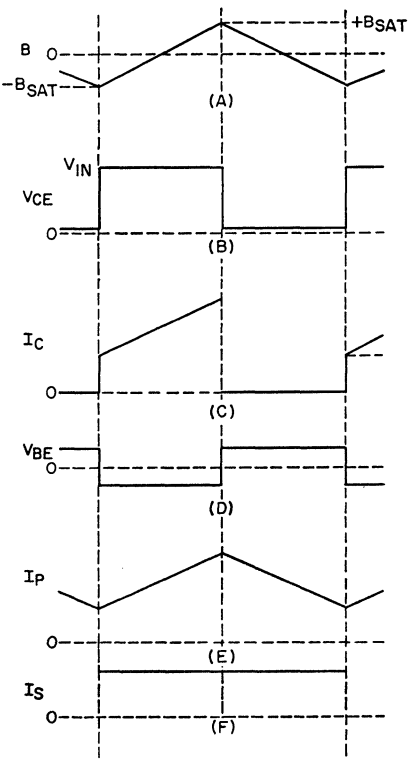


Figure 206. Typical operating waveforms for a two-transistor, one-transformer switching converter: (A) flux density in transformer core; (B) collector voltage of one transistor; (C) collector current of one transistor; (D) base voltage of one transistor; (E) primary current; (F) secondary current.

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. 206, to flow through transistor A. As in the ringing-choke converter, the linearly increasing primary current induces substantially constant voltages, waveform D in Fig. 206, 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:

$$L_P\,(dI_P/dt) = V_{in} \qquad (217)$$

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. 206, the instantaneous rate of change of current can be approximated as follows:

$$dI_P/dt = 2\hat{I}_P/0.5T = 4I_P/T = 4f\hat{I}_P \qquad (218)$$

where T is the total period of oscillation in seconds and f is the operating frequency in hertz ($f = 1/T$).

Eqs. (217) and (218) 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:

$$\hat{I}_P = V_{in}/4f\,L_P \qquad (219)$$

The average value of collector current in the conducting transistor is given by:

$$I_{AV} = P_{out}/\eta\,V_{in} \qquad (220)$$

The maximum primary current then can be expressed by the following equation:

$$I_{P(max)} = \hat{I}_P + I_{AV} = (V_{in}/4f\,L_P) + (P_{out}/\eta\,V_{in}) \qquad (221)$$

The required inductance for one-half of the primary is de-

termined from the following relationship:

$$L_P = (N_P \Delta\phi / \hat{I}_P) \times 10^{-8} \quad (222)$$

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 θ_{res}/A, therefore, is zero, and the maximum permissible magnetic swing is determined as follows:

$$\frac{\Delta\phi}{A} = \frac{\phi_{max}}{A} - \frac{\phi_{res}}{A} = \frac{\phi_{max}}{A} = B_{max} \quad (223)$$

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:

$$L_P = (N_P B_{max} A \times 10^{-8}) / \hat{I}_P \quad (224)$$

where B_{max} is expressed in gauss.

If Eq. (219) is substituted into Eq. (224), the following result can be obtained:

$$V_{in} = 4 N_P f B_{max} A \times 10^{-8} \quad (225)$$

Eq. (225) may be rewritten to obtain the following expression for the required number of turns in the primary:

$$N_P = (V_{in} \times 10^8) / 4f B_{max} A \quad (226)$$

Because no air gap is required, the required inductance for one-half the primary is given by

$$L_P = (4\pi N_P^2 / 10^9) (\mu_i A / l_i) \quad (227)$$

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:

$$N_B = N_P [2 V_{BE(max)} / V_{in}] \quad (228)$$

The required number of turns for each half of the secondary winding is expressed as follows:

$$N_S = (N_P / V_{in}) (V_{out} + R_{out} I_{out}) \quad (229)$$

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

$$V_{in(max)} \leq [V_{CB(max)} - V_{BE(wind)}]/2 \quad (230)$$

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. (230) 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. (230).

The peak secondary current is approximately equal to the dc load current, i.e.,

$$\hat{I}_S = I_{RL} = P_{out}/V_{out} \qquad (231)$$

and the peak secondary voltage is given by

$$V_S = V_{in}\,(N_S/N_P) \qquad (232)$$

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

$$C_{out} = 10/R_L\,f \qquad (233)$$

where the load resistance R_L is determined as follows:

$$R_L = (\hat{V}_S/\hat{I}_S) \qquad (234)$$

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.,

$$R_S = \left(\frac{V_{in} - 0.6}{0.6}\right) R_B \qquad (235)$$

Slightly higher starting voltages may be required for operation at low temperatures.

SAMPLE DESIGN: The design data shown in Table XVI 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 XVI 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 XIV indicates a ferrite-core type of transformer should be used. From Table XVI 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 XVI—Design Data For Push-Pull, Transformer-Coupled DC-To-DC Converters.

	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. $V_{CB(max)}$ (V)	Min. $I_{C(pk)}$ (A)	Min. P_D* (W)	Area A (cm^2)	Length l_i (cm)	Factor η
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 μ_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 η, given in Table XVI, 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. (230), 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. (226), 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. (218) 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. (221), 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. (228), the required number of turns for each section of the base winding then may be determined from the following calculation:

$$N_B = N_p \left[\frac{2\ V_{BE(max)}}{V_{in}}\right]$$
$$= \frac{8\ (2.5)^2}{13.6} = 4.1 \text{ turns}$$

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 primary current I_p of 11.5 amperes for the 40251 transistor. The resistance R_B, therefore, is calculated as follows:

$$R_B = 2.5/0.8 = 3 \text{ ohms}$$

8. From Eq. (229), 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. (231) and (232), 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. (233), 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. (234) to be

$$R_s = \frac{(13.6 - 0.6)(3)}{0.6} = 65 \text{ ohms}$$

Fig. 207 shows the circuit schematic and Fig. 208 shows the performance curves for the 13.6-to-110-volt converter.

Two-Transistor, Two-Transformer Converter—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,

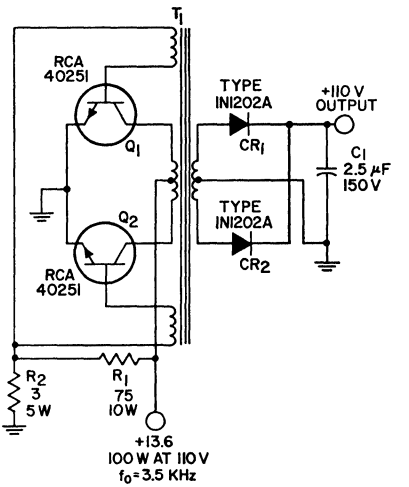


Figure 207. Schematic diagram of 13.6-to-110-volt transformer-coupled push-pull converter.

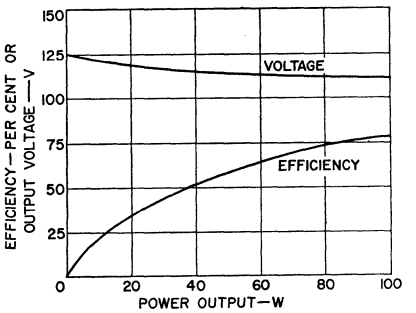


Figure 208. Output voltage and efficiency as a function of power output for the converter circuit shown in Fig. 207.

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. 209.

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 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

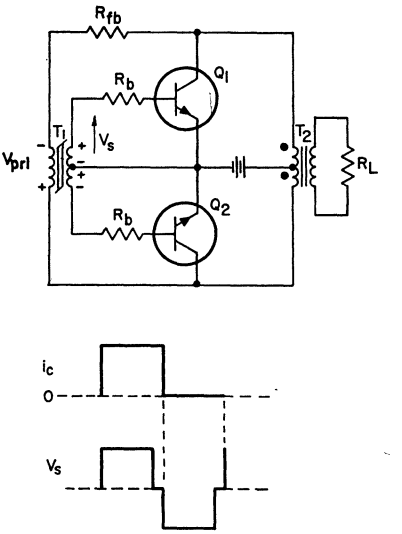


Figure 209. Two-transistor, two-transformer push-pull switching converter.

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. 209, 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. 209, 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. 210. In Fig. 210(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. 210(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:

$$V_{BE} = V_s - V_{rb} - V_T \qquad (236)$$

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

$$V_{BE} = V_{rb} - V_T \qquad (237)$$

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 dissipation, 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.

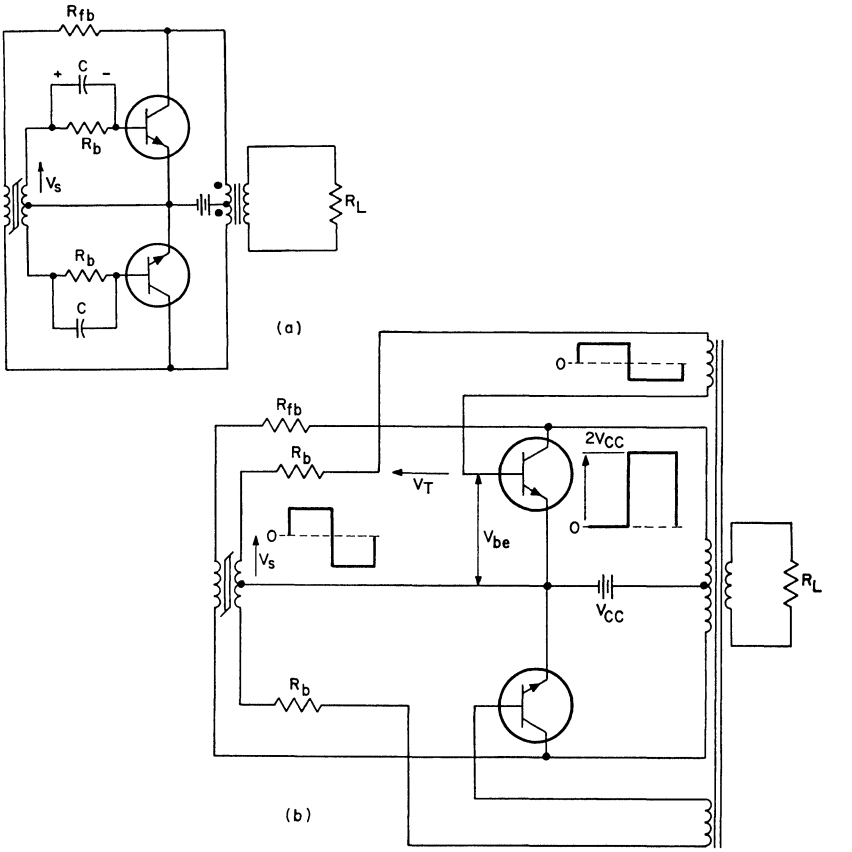


Figure 210. Two-transistor, two-transformer push-pull switching converters in which transistor-storage times are reduced: (a) Capacitor in parallel with each base resistor assures sharp turn-off of associated transistor; (b) Feedback winding from output transformer in series with base of each transistor assures sharp cutoff characteristics.

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 pro-

cedure to design the converter circuit:

1. The power input to the output transformer T_2 (P'_{out}) is computed as follows:

$$P'_{out} = P_{out}/\eta_2 \quad (238)$$

where P_{out} is the required output power of the inverter circuit and η_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:

$$I'_C = P'_{out}/V_s \quad (239)$$

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:

$$I_C'' = P'_{out}/[V_s - V_{CE(sat)}] \quad (240)$$

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:

$$P_{in} = V_{BE}\left(\frac{I_C''}{h'_{FE}}\right) + \left(\frac{I_C''}{h_{FE}}\right)^2 R_B \quad (241)$$

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 η_1, as follows:

$$P'_{in} = P_{in}/\eta_1 \quad (242)$$

6. The collector current can then be approximated on the basis of the total power developed in the converter circuit:

$$I_C = \frac{P_{out}' + P_{in}'}{V_S - V_{CE(sat)}} \quad (243)$$

If the collector current given by Eq. (243) is significantly higher than that given by Eq. (240), 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:

$$Z_L' = (V_s - V_{sat})/I_C \quad (244)$$

Thus, the turns ratio N_2 for T_2 is determined from the following equation:

$$N_2^2 = Z_L/Z_L' \quad (245)$$

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:

$$I_{pri1} + I_{m1} = (P_{in}'/V_{pri1}) + I_{m1} \quad (246)$$

9. The turns ratio for transformer T_1 is given by

$$N_1 = (V_{BE} + I_B R_B)/V_{pri1} \quad (247)$$

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

$$N_p = (V_{pri} \times 10^8)/4fAB \quad (248)$$

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. (248) 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:

$$I_{m1} = (H_s l_i)/(1.26 N_p) \quad (249)$$

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. (248).

This value of I_{m1} must be added to I_{pri1} when the value of R_{fb} is determined (step 8). Eq. (249) 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. (248) 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

$$I_C = P_{out}\, \eta/[V_S - V_{CE(sat)}] \quad (250)$$

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 η 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):

$$P_D = (T_1/T)\,(V_{CE(sat)}\,I_C + 2I_{CEX}\,V_S) + [(t_{on}+t_f)/T]\,(V_S I_C/3) \quad (251)$$

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. (250); 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. (250) and h'_{FE} given in step 4 of the general procedure]; t_f is the transistor "fall" time (at I_C given by Eq. (250) 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. (251) 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:

$$\Theta_{C-A} = (\Delta T/P_D) - \Theta_{J-C} \quad (252)$$

where ΔT is the permissible junction temperature rise ($\Delta T = T_{J(max)} - T_A$); P_D is the transistor dissipation; and $\theta_{C\text{-}A}$ 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 manual) 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_{pri}. 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. (248), 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. 209 and 210 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. 211 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. (235). With this circuit, a compromise of reliable starting and tolerable bleeder current must be reached.

Fig. 212 shows a diode starting circuit in which the bases of the two inverter transistors are supplied by a resistance R_1, which is determined as follows:

$$R_1 = V_{CC}/2I_B \qquad (253)$$

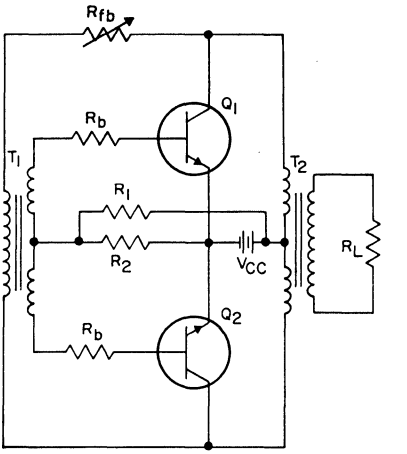


Figure 211. Two-transistor, two-transformer push-pull inverter that uses a resistive voltage-divider network to provide starting bias.

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.

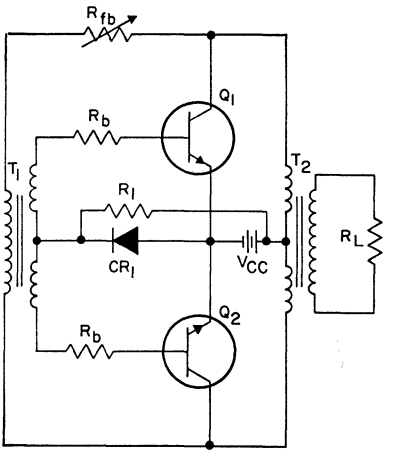


Figure 212. Two-transistor, two-transformer push-pull inverter that uses a diode starting circuit.

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. (238),

$$P'_{out} = 250/0.95 = 262.5 \text{ watts}$$

Eq. (239) 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. (240) 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 ρ 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, ρ', 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. (248) as follows:

$$N_p = \frac{27.25 \times 10^8}{(4)(5)(10^4)(2.04\times10^3)} = 6.55 \text{ turns}$$

If N_p = 6 turns, then N_s = (6)(2.08) = 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

$$\rho = 63 \ \mu\text{W/cm}^3 \text{ Hz}$$

The core loss at 50 kHz is then calculated as follows:

$$\rho' = 63 \times 50 \times 10^3 = 3.15 \text{ W/cm}^3$$

On the basis of a transformer efficiency of 95 per cent, the magnetically dissipated power in the drive transformer is given by

$$P_m = P_{in}\ (1.00-0.95) = (0.86)(0.05) = 0.43 \text{ watt}$$

The maximum volume is then calculated as follows:

$$\text{Vol} = P_m/\rho' = 0.43/3.15 = 0.136 \text{ cm}^3$$

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 three 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. (248) as follows:

$$N_p = \frac{(27.25)(10^8)}{(4)(5\times10^4)(0.119)(3000)} = 37.5 \text{ turns}$$

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. (249) as follows:

$$I_{m1} = (3.9\times1)/(1.26\times37) = 0.084 \text{ ampere}$$

The total primary current is then

$$I_{pri} = 0.033 + 0.084 = 0.117 \text{ ampere}$$

The feedback resistance R_{fb} is then calculated as follows:

$$R_{fb} = 27.25/0.117 = 235 \text{ ohms}$$

4. Thermal-Resistance Calculations. From Eq. (251), the average transistor dissipation is given by

$$P_D = [(20-1)/(2\times20)][(0.75\times9.65) + 2(0.020\times28)] + (1/20)[(28\times9.65)/3] = 8.65 \text{ watts}$$

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 θ_{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 θ_{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. 213 shows the schematic diagram for the completed circuit.

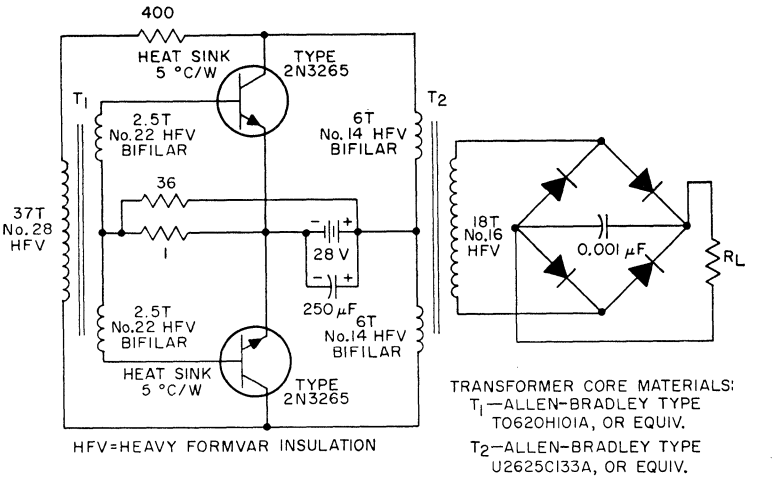


Figure 213. Schematic diagram of 250-watt, 50-kHz push-pull dc-to-dc converter.

The values for the feedback resistance and the bias-starting resistance were arrived at experimentally with the calculated values used as a beginning.

Fig. 214 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. 215 shows the experimental transistor load line for a load resistance of 25.6 ohms and a sup-

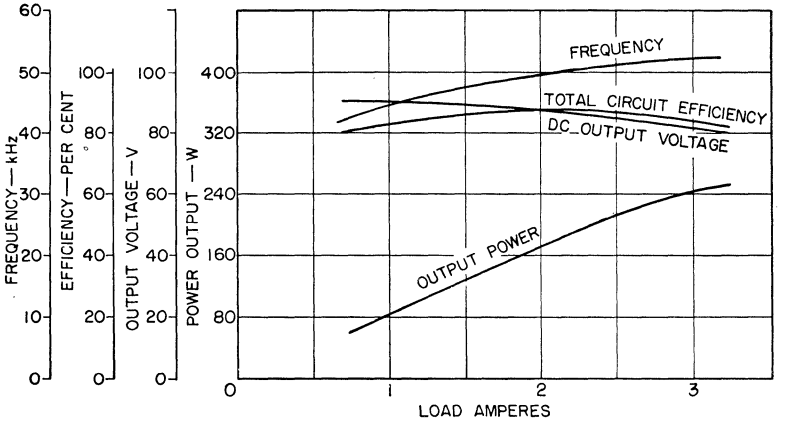


Figure 214. Output characteristics (i.e., frequency, efficiency, voltage, and power) of the 250-watt converter as a function of the load.

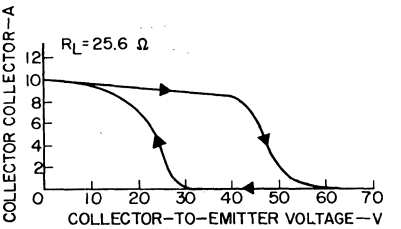


Figure 215. Experimental load line for the 2N3265 transistor using a load impedance of 25.6 ohms and a supply voltage of 28 volts.

ply 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.

Fig. 216 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. 217 shows the collector current on an expanded time scale to illustrate the current rise and fall times.

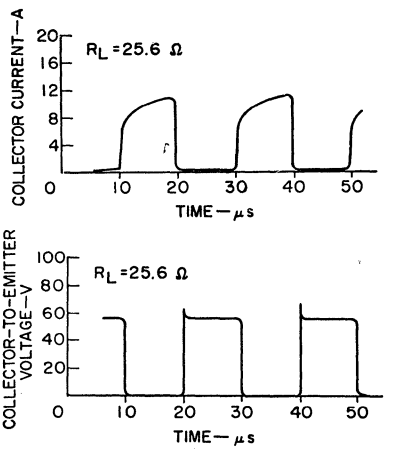


Figure 216. Collector-current and voltage waveforms for the 2N3265 transistors used in the 250-watt converter.

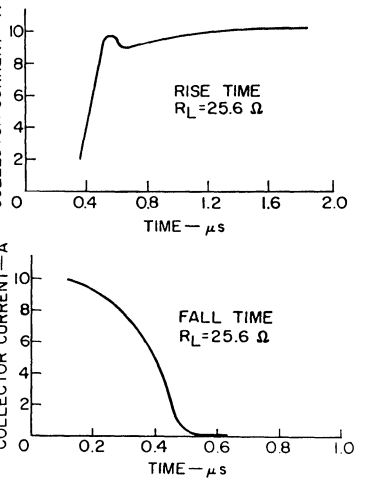
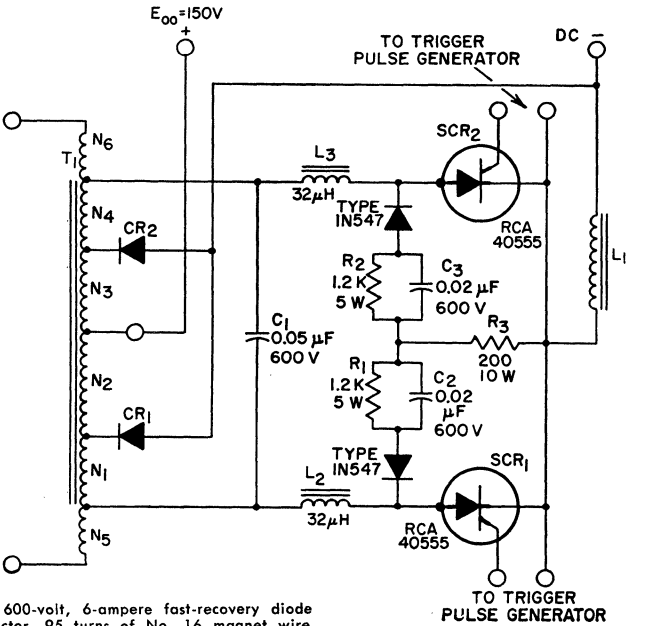


Figure 217. Collector-current rise and fall times.

SCR Inverter

Fig. 218 shows a typical high-frequency SCR switching inverter; Fig. 219 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 8 kHz provided that a suitable output transformer is used and the timing capacitors are changed in 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 simi-



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 266215-A0000-R026 (or equiv.) with 4-mil air gap

Figure 218. High-frequency (10-kHz) SCR push-pull switching inverter.

lar to that of the two-transistor push-pull inverter except that external gate-trigger signals are required to initiate the SCR switching action.

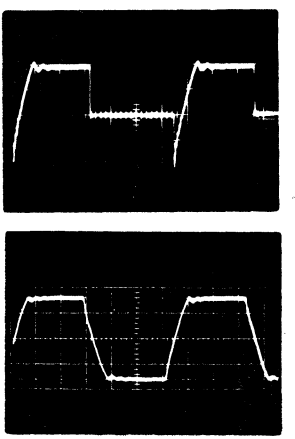


Figure 219. Typical operating waveforms for SCR inverter shown in Fig. 218.

Circuit operation—Fig. 218 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. 220 to produce an alternating current in the primary of the power transformer.

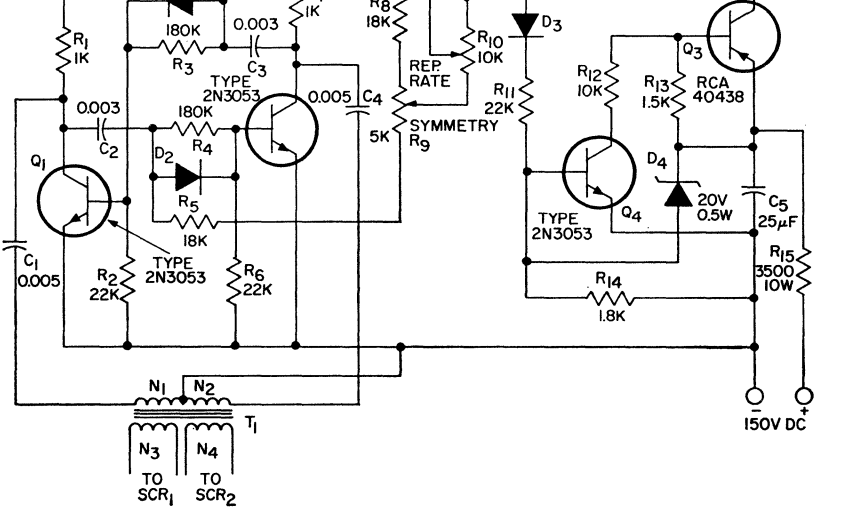
The thyristors are commutated by capacitor C_1, which is connected between the anodes of SCR_1 and SCR_2. The flow of current through the circuit can be traced more easily if it is assumed that initially SCR_1 is conducting and SCR_2 is cut off and that the common cathode connection of the SCR's is the 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, 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 $-2\ E_{oo}$, 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



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, or equiv.

Figure 220. Gate-trigger pulse generator for SCR inverter shown in Fig. 218.

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. 220, 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 required 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.

Power Regulation

THE performance of oscillators, high-gain amplifiers, and other electronic circuits that have exacting frequency, stability, or output requirements can be critically affected by wide variations in dc supply voltages. Large supply-voltage variations may also result in voltage levels that exceed acceptable circuit limits. Moreover, laboratory tests and measurements of electronic devices and circuits often require the use of constant, precisely controlled dc voltages. For these reasons, 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.

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 pass element (transistor or 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 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.

LINEAR VOLTAGE 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. 221 shows the response characteristics for each type of series-regulated power supply.

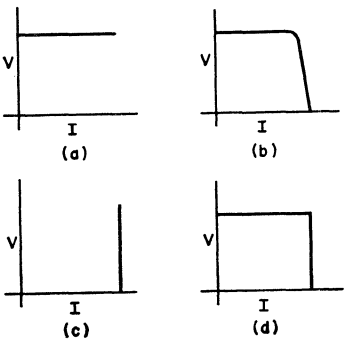


Figure 221. Typical response characteristics for series-regulated power supplies: (a) voltage-regulating types; (b) voltage-regulating current-limiting types; (c) current-regulating types: (d) voltage-regulating current-regulating types.

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.

Fig. 222 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, 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, independ-

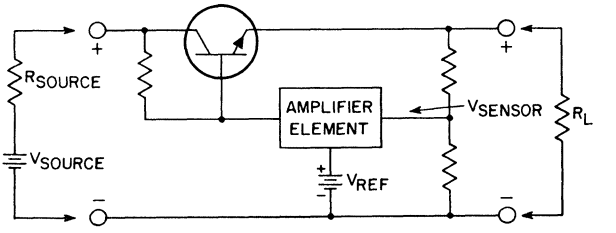


Figure 222. Basic series voltage regulator.

ent of the load current, as shown in Fig. 221(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. 221(b).

Fig. 223 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 output characteristics as shown in Fig. 221(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. 221(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.

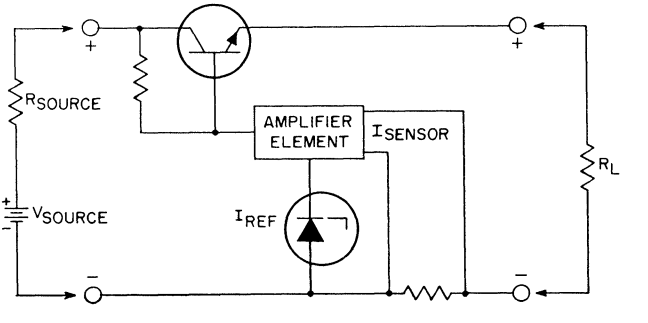


Figure 223. Basic series regulator modified for current sensing.

Fig. 224 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.**)

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

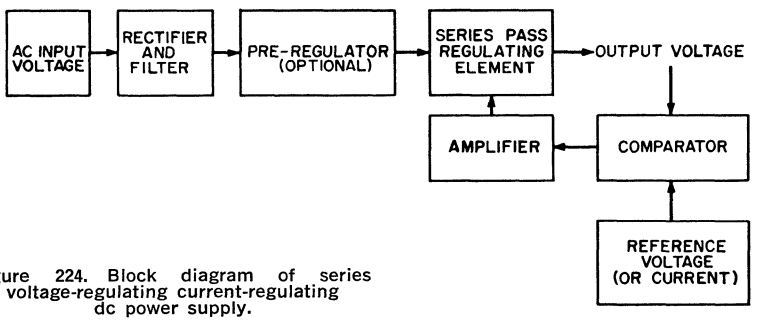


Figure 224. Block diagram of series voltage-regulating current-regulating dc power supply.

defined as 100 (V_o'/V_o), or as the change in output voltage ΔV_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. 225. 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

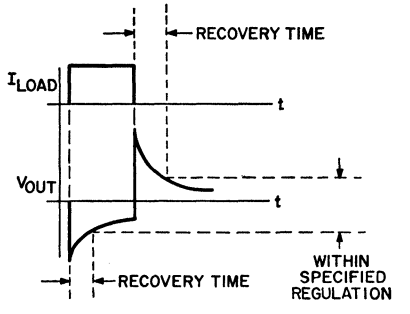


Figure 225. Typical recovery-time characteristics for regulated dc power supplies.

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.

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 S/b.

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 regu-

lator 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. 226 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.

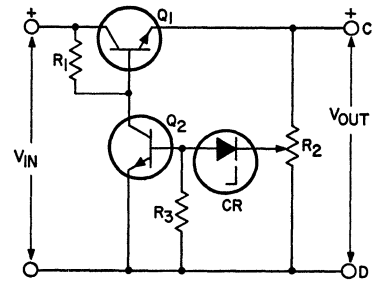


Figure 226. Series voltage regulator without current limiting.

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 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.

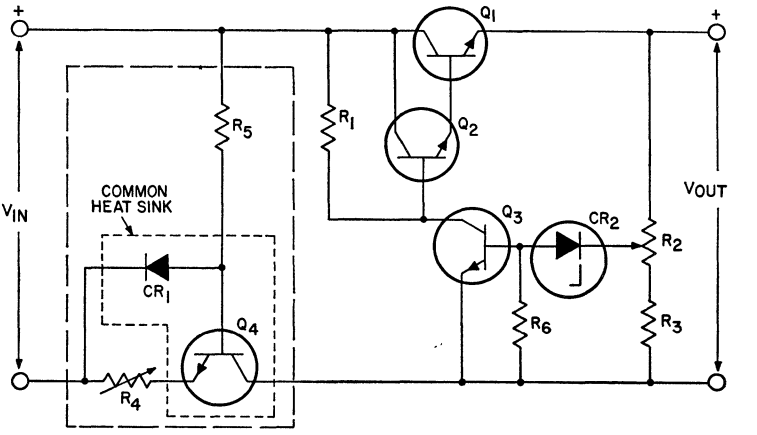
The current-limiting section (dashed line) of the regulator circuit shown in Fig. 227(a) is designed to appear as a large series resistance during current overload and as a negligible resistance during normal operating conditions. The value of resistance R_5 is designed so that, during normal regulator operation, transistor Q_4 operates in the saturated condition. For the overload condition, 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

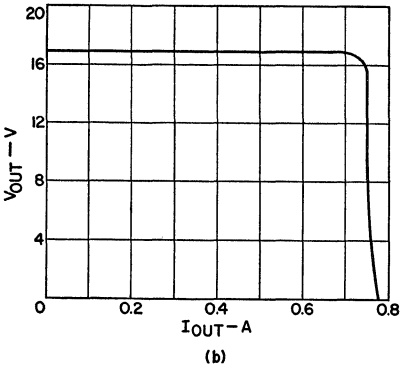
V_{in} appears across Q_4 simultaneously with the limiting value of current. Transistor Q_4 must be capable of withstanding the resulting dissipation. When the current limit is reached, the junction temperature of Q_4 rises to a value considerably above the ambient temperature. This increase in junction 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. 227(a) are shown in Fig. 227(b).

Although the series-regulator circuit shown in Fig. 227(a) provides adjustable current limiting with simple circuitry and minimum power loss during normal operation, it has the disadvantage



(a)



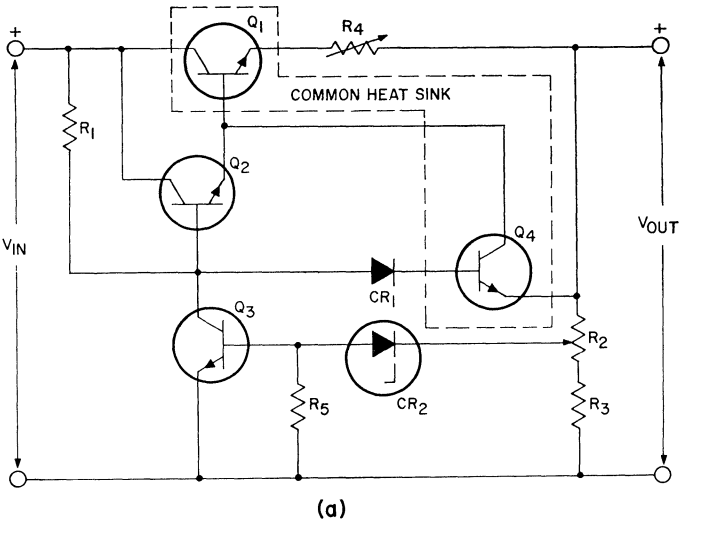
(b)

Figure 227. Series voltage regulator with transistor current-limiting circuit (inside dashed lines) added: (a) schematic diagram; (b) response characteristics.

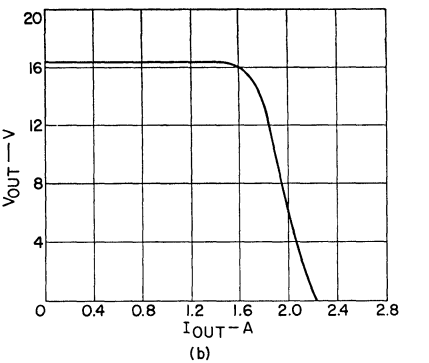
of requiring a second series transistor capable of withstanding short-circuit output current and total input voltage simultaneously.

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. 227(a) may not be practical in this application because of the additional series transistor required. The circuit shown in Fig. 228(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.

In the circuit shown in Fig. 228(a), current limiting is achieved by the combined action of the components shown inside



(a)



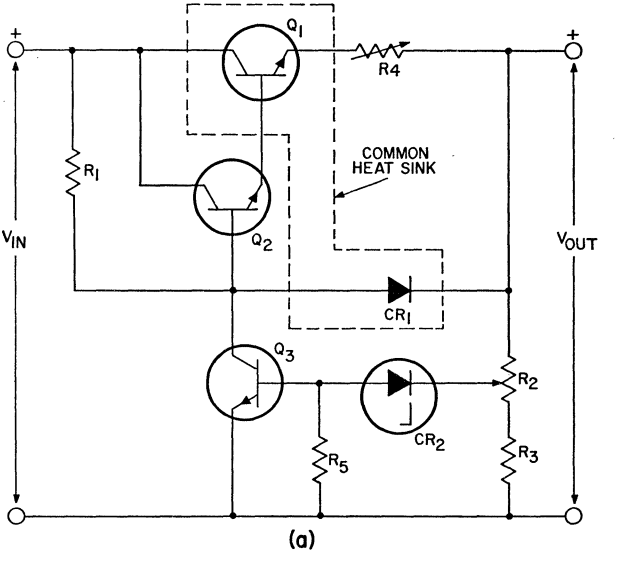
(b)

Figure 228. Series voltage regulator using pass transistor as part of current-limiting circuit: (a) schematic diagram; (b) response characteristics (for $R_4 = 0$).

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. 228(b).

The circuit shown in Fig. 229(a) is a variation of that



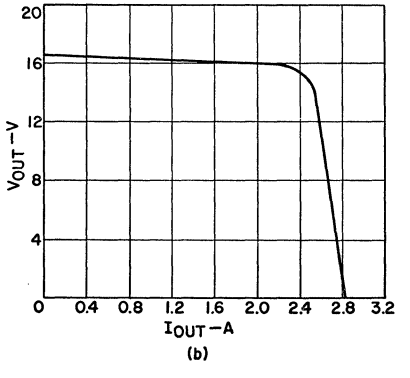


Figure 229. Series voltage regulator which uses additional transistor-diode network and series pass transistor to accomplish current-limiting function: (a) schematic diagram; (b) response characteristics.

shown in Fig. 228(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. 229(b). The circuits shown in Figs. 228 and 229 are both applicable to high-current high-voltage regulators because additional series power transistors are not required.

Fig. 230(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. 230(b). The circuit functions equally well with resistor R_4 located in the positive output lead.

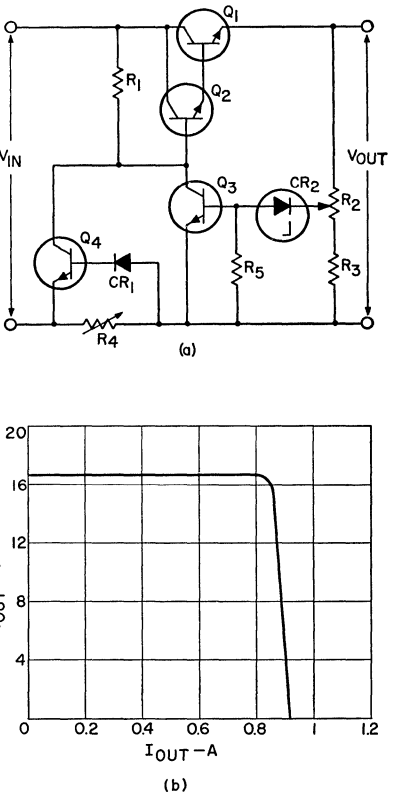


Figure 230. Current-limiting series voltage regulator in which series pass transistor must be capable of withstanding input voltage and short-circuit current simultaneously: (a) schematic diagram; (b) response characteristics.

Design of a Practical Series Regulated Power Supply

Fig. 231 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

$$V_{REF}/V_o = R_{CAL}/R_{ADJ} \qquad (254)$$

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:

$$V_o = I_o R_{ADJ} \qquad (255)$$

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. (254) 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. 232, 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 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

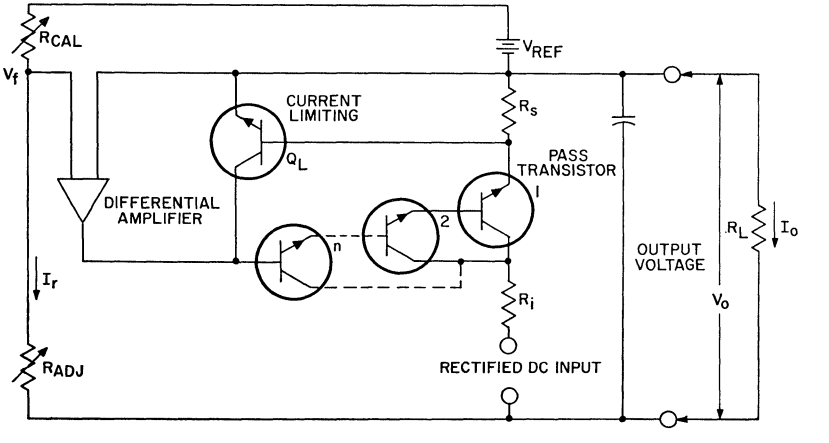


Figure 231. Voltage-regulated current-limited power supply.

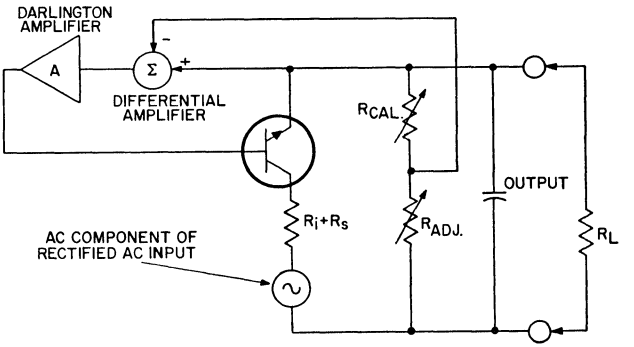


Figure 232. AC feedback circuit for a voltage-regulated power supply.

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. 231:

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, ΔV_i
Possible negative change in input voltage caused by line variation and ripple, ΔV_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:

$$V_o/20 < V_{REF} < V_o/10 \quad (256)$$

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.)

$$R_{ADJ} = 100\ V_o, \text{ or } R_{ADJ} = 1000\ V_o \quad (257)$$

4. The resistance value for R_{CAL} is calculated by use of the following equation, which is obtained by rearrangement of terms in Eq. (254):

$$R_{CAL} = (V_{REF} \times R_{ADJ})/V_o \quad (258)$$

5(a). Initially, the transistor should satisfy the following requirements:

$$V_{CEO}(sus) \geq V_i + \Delta V_i \qquad (259)$$
$$I_C(max) \geq I_o(max)$$

where V_i is the input supply voltage, ΔV_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

$$A_I = h_{FE}(min)_1 \times h_{FE}(min)_2 \times \ldots \times h_{FE}(min)_n \qquad (260)$$

Therefore, the appropriate number of driver stages is selected so that the total current gain will satisfy the following requirement:

$$A_I \geq I_o(max)/I_d \qquad (261)$$

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:

$$R_S = V_{BE}/I_o(max) \qquad (262)$$

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:

$$V_{BE1} + V_{BE2} + \ldots + V_{BE_{n-1}} + V_{CE}(sat)_n = V_x \qquad (263)$$

$$V_{CE}(sat)_1 = V_y \qquad (264)$$

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 drive 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.,

$$V_x \text{ or } V_y < [V_i - \Delta V_n - I_{max}(R_i + R_s)] - V_o \qquad (265)$$

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:

$$V_o = V_i - I_o(R_s + R_i) - V_{pass} + \Delta V_i \qquad (266)$$

This equation is rewritten to obtain an expression for the voltage dropped across the pass transistor, as follows:

$$V_{pass} = V_i + \Delta V_i - V_o - I_o (R_i + R_S) \quad (267)$$

The power dissipated by the pass transistor is given by

$$P_{pass} = I_o V_{pass} \quad (268)$$

Substitution of Eq. (267) into Eq. (268) results in the following expression for the power dissipation in the pass transistor:

$$P_{pass} = I_o (V_i + \Delta V_i) - I_o V_o - I_o^2 (R_i + R_S) \quad (269)$$

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:

$$dP_{pass}/dI = (V_i + \Delta V_i) - V_o - 2I_o (R_i + R_S) = 0 \quad (270)$$

The maximum power in the circuit occurs, therefore, when

$$I_o = \frac{V_o + \Delta V_i - V_o}{2(R_i + R_S)} \quad (271)$$

Eq. (271) 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

$$I_o = \frac{V_i + \Delta V_i}{2(R_i + R_S)} \quad (272)$$

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:

$$P_{pass(max)} = I_o (V_i + \Delta V_i) - I_o^2 (R_i + R_S) \quad (273)$$

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. (257), the value of R_{ADJ} is calculated as follows:

$$R_{ADJ} = 1000\, V_o = 100\,(120) = 120{,}000 \text{ ohms}$$

4. From Eq. (258), 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. (260), the total current gain of the driver stage is determined as follows:

$$A_I \geq I_o(max)/I_d = 300\text{ mA}/0.4\text{ mA} = 750$$

(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:

$V_{BE1} = V_{BE}$ (of TA2765 at $I_o = 0.3$ A) $= 1$ V
$V_{CE}(sat)_2 = V_{CE}(sat)$ (of 2N3440 at $I_C = 15$ mA) $= 0.5$ V
$V_{BE1} + V_{CE}(sat)_2 = 0.64 + 0.75 = 1.39\ V = V_x$
$V_{CE}(sat)_1$ (for TA2765) $= 2.5V = V_y$
$V_y > V_x$

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. (273) 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. 233. Fig. 234 shows the complete circuit diagram for the voltage-regulated current-limited power supply; Fig. 235 shows the schematic for the reference-voltage power supply.

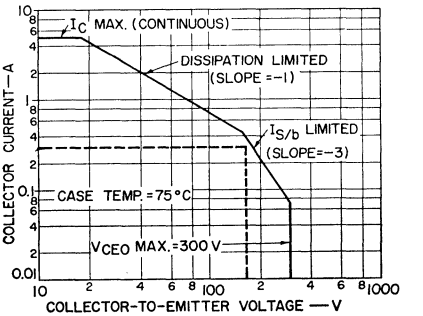


Figure 233. Safe-area curve for the RCA 2N5240 transistor. Dashed-line rectangle indicates that transistor operates within ratings in the 100-volt current-limiting (0-3 ampere) series-regulated power supply.

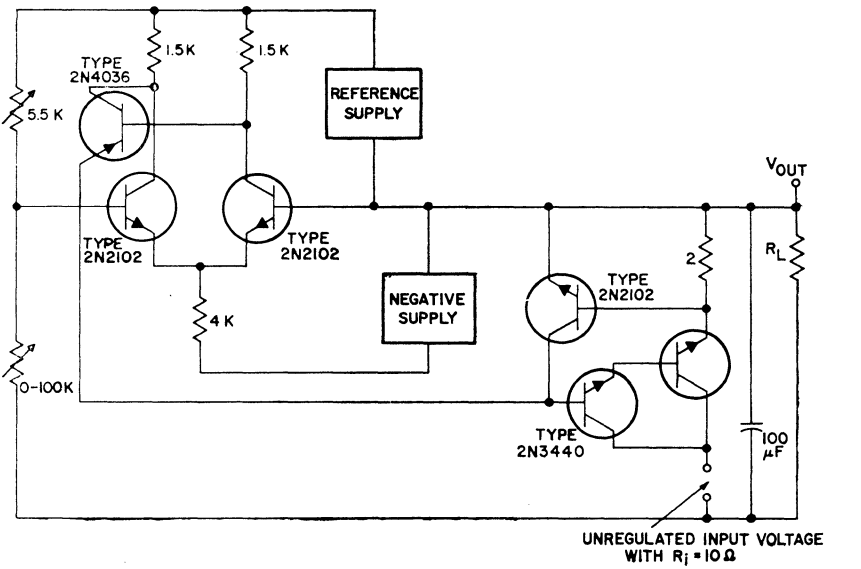


Figure 234. Schematic diagram of 100-volt series-regulated dc power supply in which output current is limited to a maximum value of 0.3 ampere.

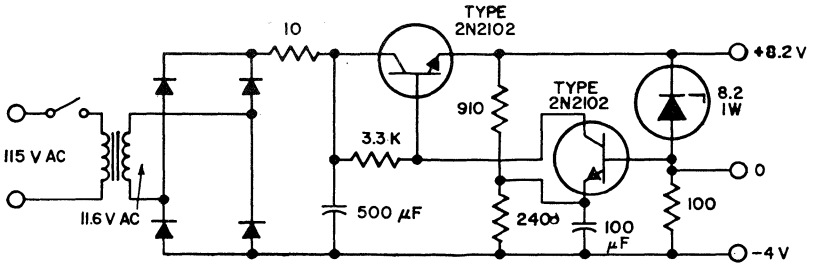


Figure 235. Schematic diagram of reference voltage supply for regulated power supply shown in Fig. 234.

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. 236.

The shunt element contains one or more transistors connected in the common-emitter configuration in parallel with the load, as shown in Fig. 237.

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 ΔV_S

Source resistance R_S

Output load resistance R_{Lo}

Output voltage V_o

Output-voltage variation ΔV_o

The terms V_S and R_{Lo} are design-center values; ΔV_S and ΔR_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:

$$I_1(\text{max}) = I_L(\text{max}) = V_o(R_{L_o} - \Delta R_L) \quad (274)$$

$$V_1(\text{max}) = V_o \text{ (under forward-bias conditions)} \quad (275)$$

$$P_1(\text{max}) = I_1(\text{max})\, V_1(\text{max}) \quad (276)$$

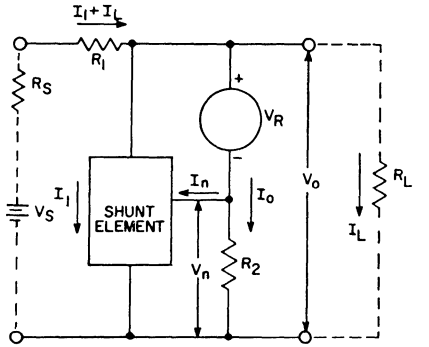


Figure 236. Basic configuration for a typical shunt regulator.

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:

$$R_2 = n/I_o \qquad (277)$$

where n is the number of stages in the shunt element.

4. The output resistance R_o of the regulator is given by

$$R_o = (2\Delta V_o/V_o)/R_{L_o} \qquad (278)$$

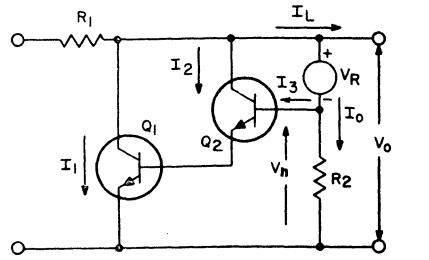


Figure 237. Shunt regulator circuit using two transistors as the shunt pass element.

If a value of series resistance is assumed for the reference R_f, this equation can be solved for h_{fe}. It can be shown that

$$R_o = \frac{R_f + \dfrac{h_{ie}\,R_2}{h_{ie} + R_2}}{1 + h_{fe}\dfrac{R_2}{R_2 + h_{ie}}} \qquad (279)$$

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:

$$I_n = I_1/(h_{FE1}\, h_{FE2} \ldots h_{FEn-1}) \qquad (280)$$

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}.

$$V_n = V_1 + V_2 + \ldots + V_{n-1} \qquad (281)$$

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 ΔV_S; R may be expressed in terms of these quantities as follows:

$$V_S+\Delta V_S = V_o+R\,[I_L(max)+I_1(max)] \quad (282)$$

$$V_S-\Delta V_S = V_o+R\,I_1(max) \quad (283)$$

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. 238. When the step-by-step procedure refers to components by reference designations, Fig. 237 indicates the component being considered.

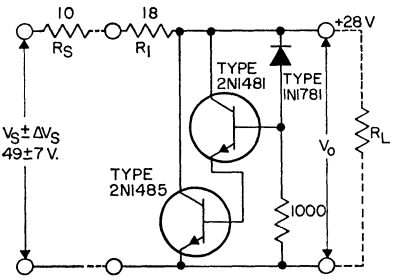


Figure 238. Schematic diagram of 28-volt shunt regulator circuit.

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:

$$I_2 = I_1/(h_{FE1}\, h_{FE2}) = 250/(50 \times 25)$$
$$= 0.20 \text{ mA}$$

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 $I_1(max) = I_L(max)$ as follows:

$$V_S - \Delta V_S = V_o + [R\ I_1(max)]$$
$$49 - 7 = 28 + (0.5\ R)$$
$$R = 28 \text{ ohms}$$

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. 238.

SWITCHING REGULATOR

Fig. 239 shows the basic configuration for a switching type of 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.

There is another method of pulse-width modulation in which 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. Transistor circuits, although not usually as simple, can also be used. It should be noted that a switching regulator operating in this mode requires that devices be used in front of any filtering. This requirement does not exist for most transistor switching regulators.

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.

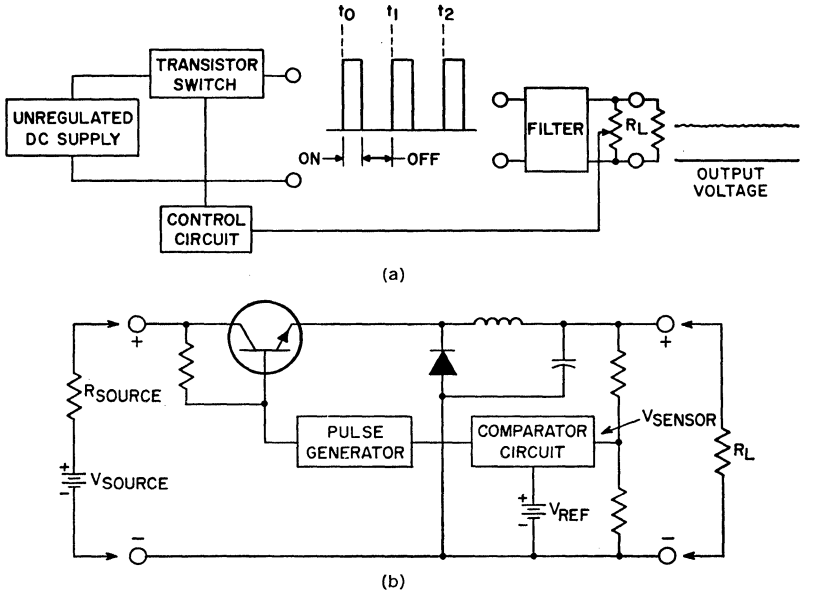


Figure 239. Basic configuration of switching type of transistor voltage regulator: (a) block diagram; (b) schematic diagram.

Filter Considerations

A fundamental part of every switching regulator is the filter. Fig. 240 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. 240(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. 240(b) has three disadvantages: (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. 241. This diode

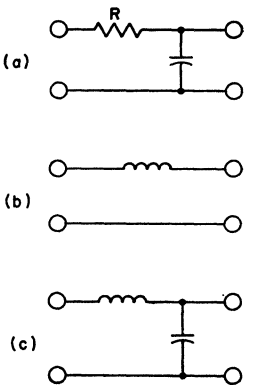


Figure 240. Typical filter circuits for use between pass element and load in a switching regulator: (a) capacitive filter; (b) inductive filter; (c) inductive-capacitive filter.

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.

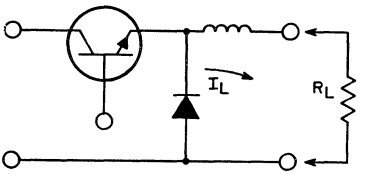


Figure 241. Use of inductance and commutating diode as filter network between pass transistor and load in switching voltage regulator.

(3) A third disadvantage of the inductive filter becomes evident during light loads. The energy stored in an inductor is given by

$$E = \tfrac{1}{2} LI^2 \qquad (284)$$

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. 240(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. 242. 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.

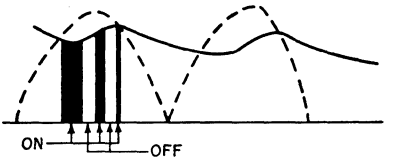


Figure 242. Effect of high-frequency switching of the switching regulator on power-supply ripple component.

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:

$$dI/dt = (1/L)\,(V_{in} - V_{out}) \qquad (285)$$

The peak current is, therefore, given by

$$I_p = [(V_{in} - V_{out})/L]\,t_{on} \qquad (286)$$

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:

$$t_r \leq 1/100f;\ t_f \leq 1/100f \qquad (287)$$

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. 243. 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. 243(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. 243(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.

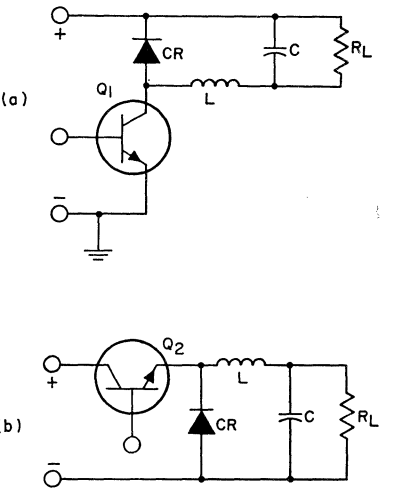


Figure 243. Basic transistor switching arrangements: (a) filter elements and load impedance in collector circuit of switching transistor; (b) filter elements and load impedance in emitter circuit of switching transistor.

The circuit of Fig. 243(b) is usually preferred when the power that results from a high V_{CE}(sat) can be tolerated.

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. 244 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. 245 shows the performance characteristics for this circuit.

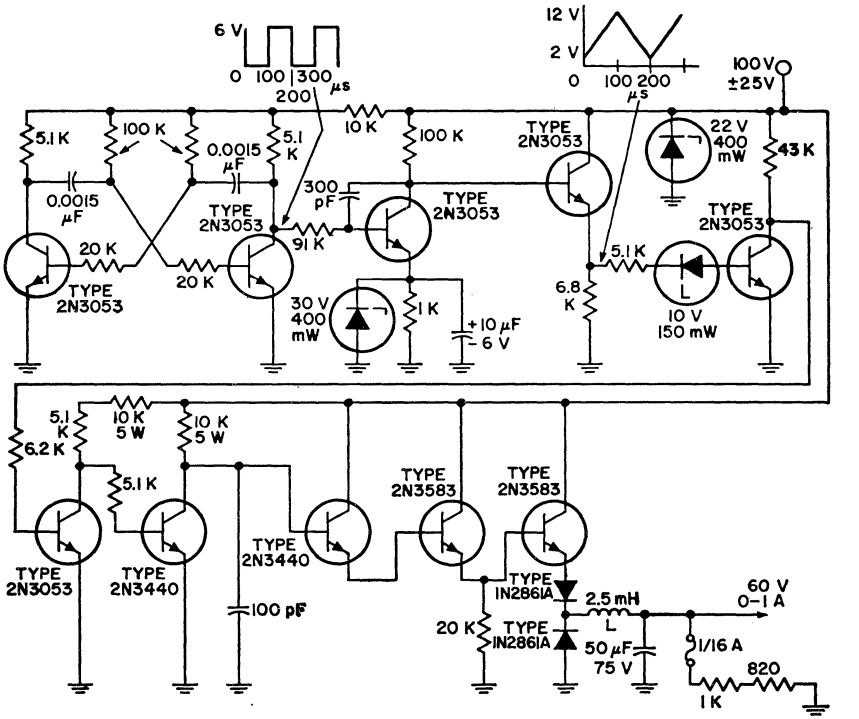


Figure 244. Typical step-down transistor switching regulator.

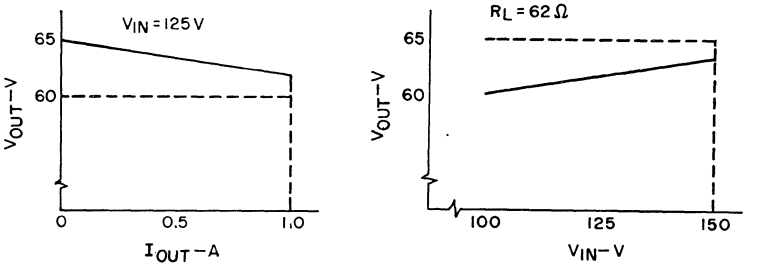


Figure 245. Performance characteristics of step-down switching regulator shown in Fig. 244.

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. In addition, SCR's are generally used whenever the desired function can be accomplished adequately by this type of device because of the economics involved.

A triac provides symmetrical bidirectional electrical characteristics. Triacs have been developed specifically for control of ac power, and are used primarily for control of power to a load from ac power lines.

GENERAL CONSIDERATIONS

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 supply voltage is reduced to zero for a short time or reversed for an even shorter time. Either operation returns the thyristor to its blocking state. Because both of these operations are accomplished 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. When only complete half-cycles or integral numbers of half-cycles are desired for a given load while control over the average power is maintained, the control is usually referred to as an **integral-cycle** or **zero-switching control.** This type of control is illustrated in Fig. 246.

Thyristor Phase Control

In most power-control applications of thyristors, partial cycles of the applied ac voltage are switched to the load. Because the

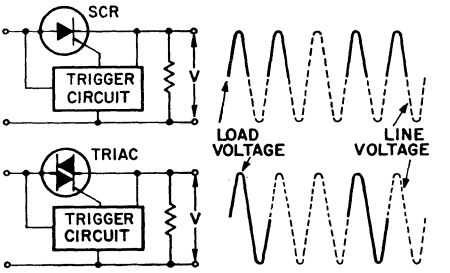


Figure 246. Integral-cycle thyristor power control circuits.

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** (θ_F). It is usually more important, however, to know and to refer to the **conduction angle** (θ_C), which 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. 247.

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. 248. 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.

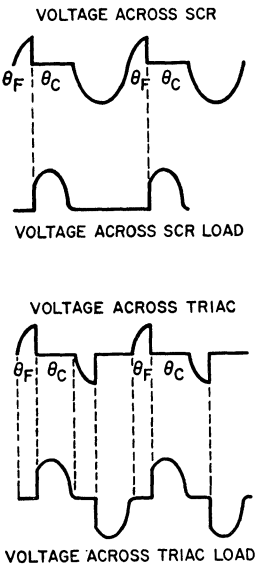


Figure 247. Voltage waveforms showing conduction angle for half-wave operation (SCR) and full-wave operation (triac) of thyristor phase-control circuits.

Current Relationships in Phase-Controlled Thyristor Circuits

In the design of thyristor power-control circuits of the types shown in Fig. 248, it is

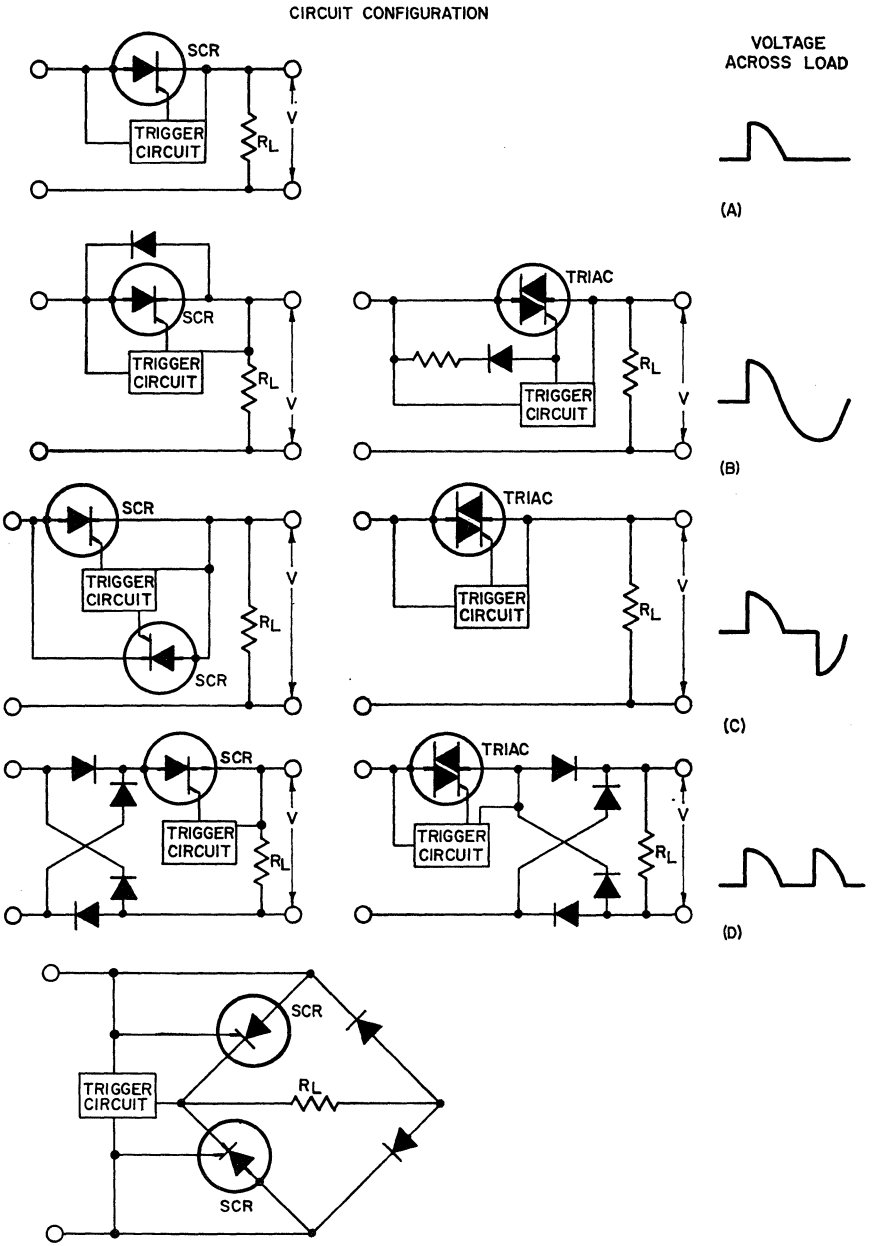


Figure 248. Basic circuit configurations for thyristor power controls and voltage waveform across the load for two complete cycles of operation.

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 I, 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. 249, 250, and 251 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. 249, 250, and 251 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.

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,

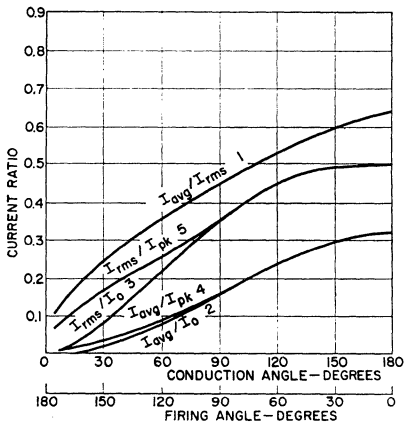


Figure 249. Ratio of SCR current as a function of conduction and firing angles for single-phase half-wave conduction into a resistive load.

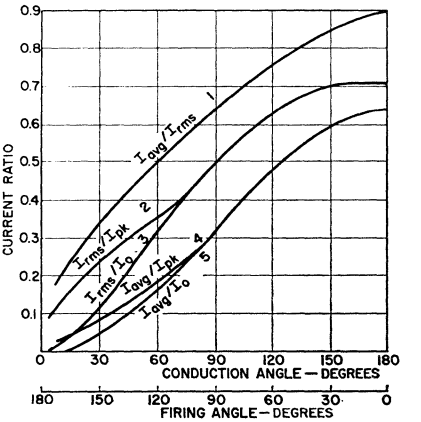


Figure 250. Ratio of thyristor current as a function of conduction and firing angles for single-phase full-wave conduction into a resistive load.

I_{pk} is equal to I_o; for conduction angles smaller than 90 degrees, I_{pk} is smaller than I_o.

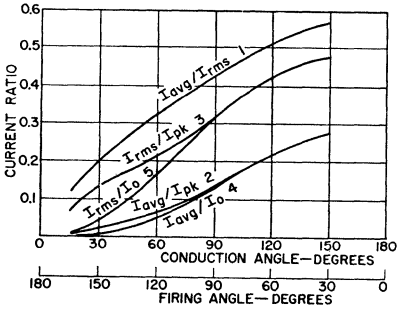


Figure 251. Ratio of SCR current as a function of conduction and firing angles for three-phase half-wave circuit having a resistive load.

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.

Example No. 1: In the single-phase half-wave circuit shown in Fig. 252, 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:

$$I_o = \frac{V_{pk}}{R_L} = \frac{170}{2.8} = 61 \text{ amperes}$$

The ratio of I_{rms} to I_o for the minimum and maximum load-current values is then calculated as follows:

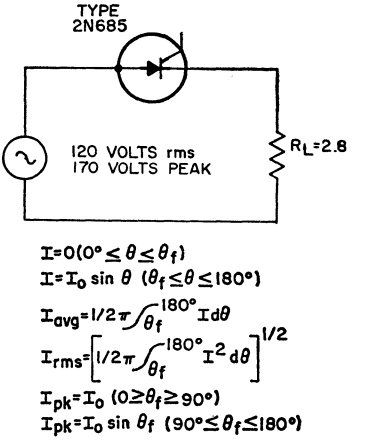


Figure 252. Single-phase half-wave circuit that operates into a resistive load and the respective equation for SCR current.

$$(I_{rms}/I_o)_{min} = 2/61 = 0.033$$

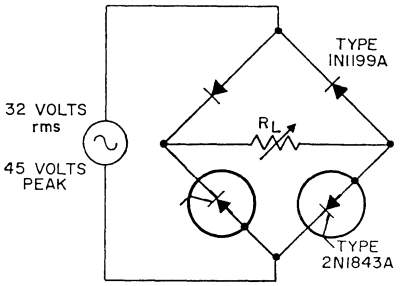
$$(I_{rms}/I_o)_{max} = 25/61 = 0.41$$

These current-ratio values are then applied to curve 3 of Fig. 249 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. 253, 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. 249 (the individual de-



$I = 0 \quad (0° \le \theta \le \theta_f)$

$I = I_o \sin \theta \quad (\theta_f \le \theta \le 180°)$

$$I_{avg} = 1/\pi \int_{\theta_f}^{180°} I \, d\theta$$

$$I_{rms} = \left[1/\pi \int_{\theta_f}^{180°} I^2 \, d\theta \right]^{1/2}$$

$I_{pk} = I_o \quad (0° \le \theta_f \le 90°)$

$I_{pk} = I_o \sin \theta_f \quad (90° \le \theta_f \le 180°)$

Figure 253. Single-phase full-wave bridge circuit that operates into a resistive load and the respective equations for SCR current.

vice 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:

$$(I_{avg}/I_o)_{min} = 3.5/225 = 0.015$$

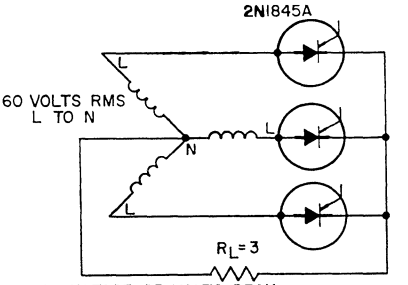
$$(I_{avg}/I_o)_{max} = 3.5/11.2 = 0.312$$

Finally, these ratios are applied to curve 2 of Fig. 249 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. 254, 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 254. Three-phase half-wave circuit that uses a resistive load and the respective equation 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. 251 for the extremes of the firing range, as follows:

$\Theta_F = 30$ degrees; $I_{rms}/I_o = 0.49$

$\Theta_F = 155$ 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:

$I_{rms(max)} = 0.49 \times 28 = 13.7$ amperes

$I_{rms(min)} = 0.06 \times 28 = 1.7$ 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:

$$P_{max} = [I_{rms(max)} \sqrt{3}]^2 R_L = 1700 \text{ watts}$$

$$P_{min} = [I_{rms(min)} \sqrt{3}]^2 R_L = 26 \text{ watts}$$

Triggering Circuits

When thyristors are triggered, the primary requirement to assure sustained forward conduction is that the gate current is of sufficient strength 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 ensure 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 always less than 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.

Many applications require that a thyristor be switched full ON or full OFF 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. 255.

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

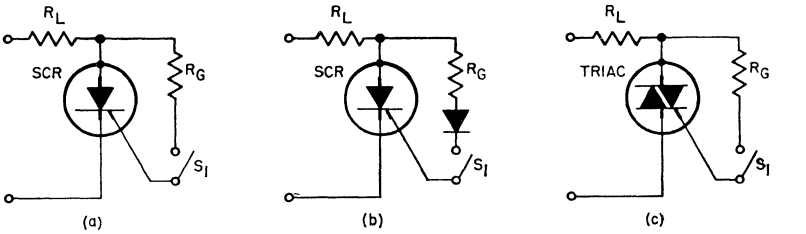


Figure 255. Simple thyristor triggering methods: (a) resistance triggering of SCR; (b) resistance-diode triggering of SCR; (c) resistance triggering of triac.

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. 255(b) illustrates the use of the SCR-and-diode combination.

When an ac resistive trigger network is used, only a certain 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. 256. 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 θ_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.

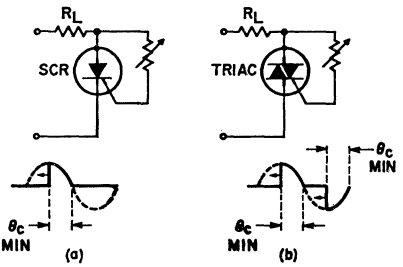


Figure 256. Degree of control over conduction angles when ac resistive network is used to trigger SCR's and triacs.

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. 257(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 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.

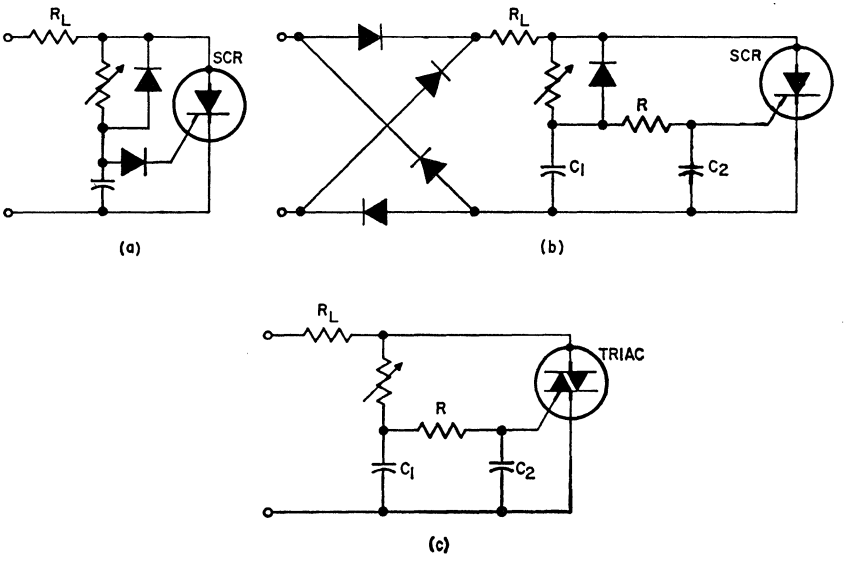


Figure 257. RC triggering networks used for phase-control triggering of thyristors: (a) series RC network and diode (for SCR); (b) RC lattice network and diode (for SCR); (c) RC lattice network without diode (for triac).

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. 257(b) and 257(c), or by use of a negative voltage across the capacitor, as shown in Fig. 257(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. 258. The voltage-current characteristic for this circuit is shown in Fig. 259.

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

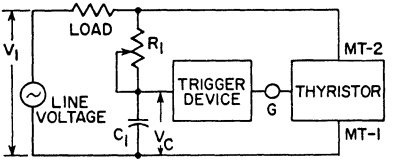


Figure 258. Thyristor power control in which switching is controlled by basic triggering-device circuit.

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

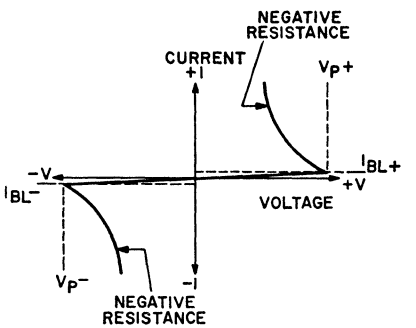


Figure 259. Voltage-current characteristic for triggering-device shown in Fig. 258.

blocking state. The blocking action of the thyristor prevents the line voltage from appearing across the load. This condition is shown in Fig. 260(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 breaks down. This breakdown is accompanied by a sudden discharge of capacitor C_1, which results in a pulse of gate

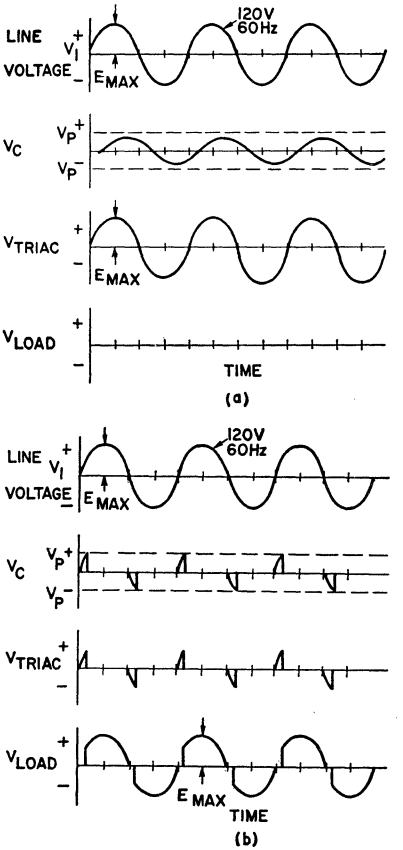


Figure 260. Voltage waveforms for circuit shown in Fig. 258: (a) thyristor in blocking state; (b) when thyristor is triggered at controlled intervals by triggering-device circuit.

current into 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 is 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. 261.

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.

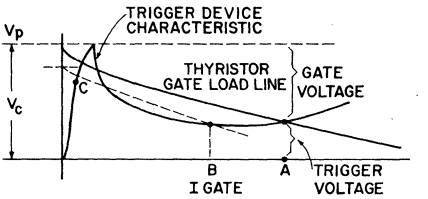


Figure 261. Load line for the circuit shown in Fig. 258.

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. 261, 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. 262 shows the typical shape of the gate-current pulse that is produced. (More specific magnitudes are shown in later diagrams for particular

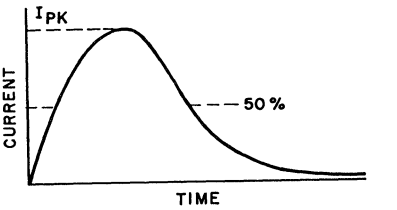


Figure 262. Typical gate-current waveform for circuit shown in Fig. 258.

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. 261.

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. 260(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 recommended types are the GE-5AH and Signalite AO57 or equivalents. 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. 263.

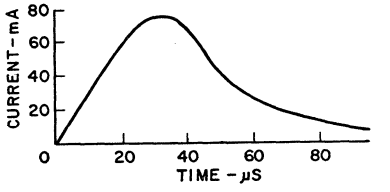


Figure 263. Typical current pulse that results from the discharge of a 0.1-microfarad capacitor through a neon bulb and a thyristor gate.

Fig. 264 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 reliable and relatively long-lived triggers can be obtained for a low price.

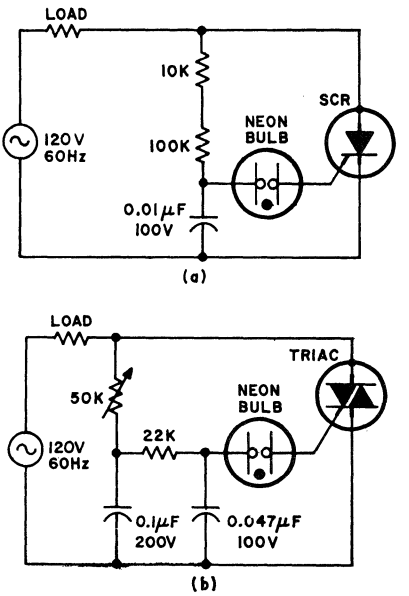


Figure 264. Circuits showing application of neon bulb as thyristor triggering device: (a) SCR power control circuit; (b) triac power control circuit.

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 silicon bidirectional 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 of 3 volts is usually imposed on the symmetry between positive and negative breakover voltages (voltage symmetry). A typical voltage-current characteristic of a bidirectional trigger diac is shown in Fig. 265.

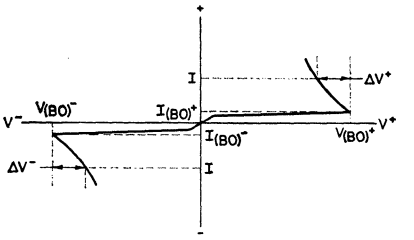


Figure 265. Typical voltage-current characteristic of a bidirectional trigger diac.

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. 266.

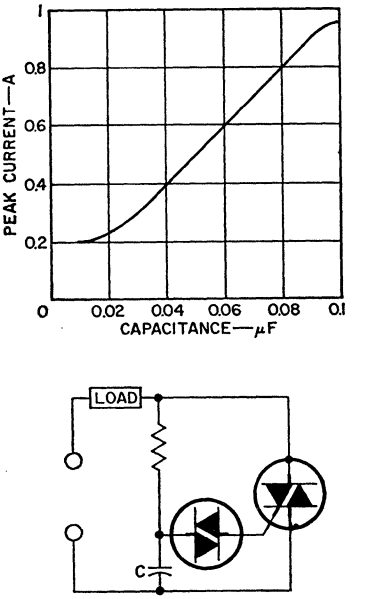


Figure 266. Peak pulse trigger currents as a function of the phase-shift capacitance.

Unijunction Transistor—The unijunction transistor is a three-terminal two-layer device formed by an emitter and a base, as illustrated in Fig. 267. One lead is connected to the emitter and the other two leads are connected to the base. Between the two base

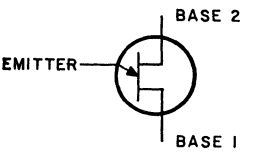


Figure 267. Schematic symbol for a unijunction transistor.

connections there is an "interbase resistance." Fig. 268 illustrates the use of this device in a pulse-triggering circuit.

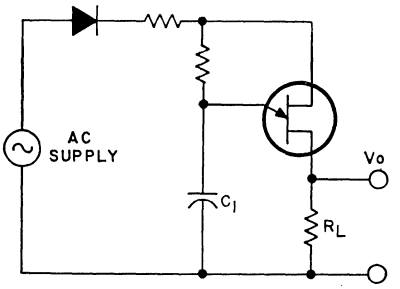


Figure 268. Pulse-triggering circuit that uses a unijunction transistor.

Proper choice of external biasing resistors, coupled with the interbase resistance, normally serves to reverse-bias the "emitter-to-base-1" diode. This condition holds the emitter-to-base-1 diode in a high-impedance state until the emitter voltage is raised to a value high enough to forward-bias this junction. As the forward-bias point is reached, the same junction switches to a low-impedance state and causes capacitor C_1 to discharge into the load resistance R_L, which may be a thyristor gate. This voltage-sensitive switching characteristic makes unijunction transistors ideal for triggering thyristors. When the biasing resistors and the emitter are connected to the same voltage source, as shown in Fig. 268, there is a degree of self-regulation of supply-voltage variation.

This regulation results because the interbase forward-bias point tracks the variation in emitter voltage V_e as V_e varies with supply voltage. Another advantage of the use of the unijunction transistor is the way in which it automatically synchronizes to an ac supply. At the end of every cycle, any charge left on the capacitor after firing is discharged when the supply goes to zero. This action occurs because the point or firing voltage is also reduced toward zero. A third advantage is the inherent ability of the device to switch relatively high currents and thus to assure positive triggering of high-gate-current thyristors.

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 ensure 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. 269.

Two-Transistor Trigger Circuit—A two-transistor trigger circuit that has characteristics similar to those of a trigger diode is shown in Fig. 270. The regenerative action of this type of circuit when either transistor begins 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. 271 can deliver trigger currents as high as 1 ampere and is more than capable of triggering all RCA thyristors.

Fig. 272 shows an SCR circuit that uses the two-transistor regenerative-trigger network. 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

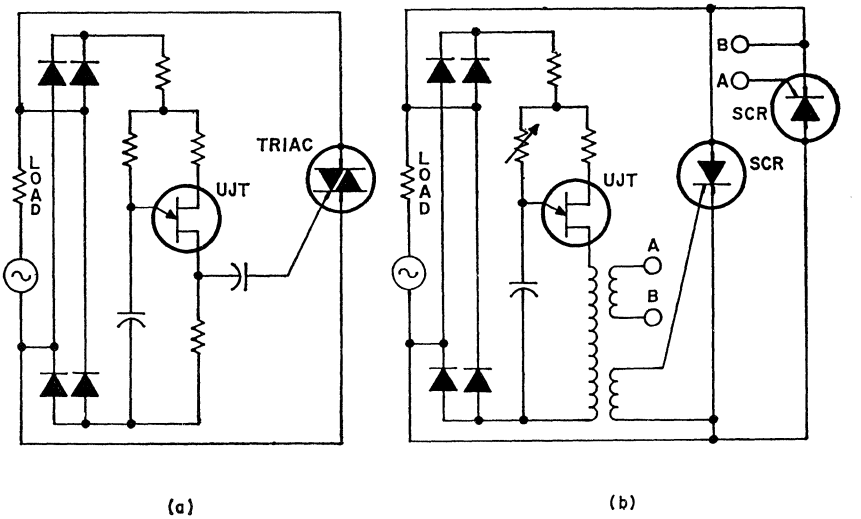


Figure 269. Circuit showing application of unijunction-transistor circuit for pulse triggering of triacs and inverse parallel SCR's: (a) triggering pulse is capacitively coupled to gate of triac; (b) triggering pulse is transformer-coupled to gate of SCR.

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 motor. Speed control is accomplished by adjustment of potentiometer R_1. For the component values shown on the schematic diagram in Fig. 272, the threshold voltage for firing the circuit is approximately 8 volts, and the maximum conduction angle is approximately 170 degrees. Table XVII shows values for operation of the circuit with various RCA SCR's.

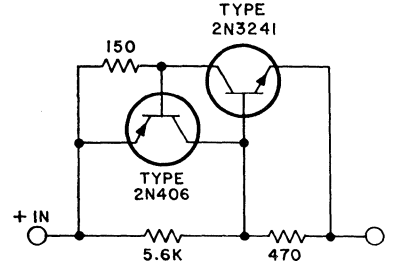


Figure 270. Two-transistor switch. Characteristics of this circuit are similar to those of a trigger diode.

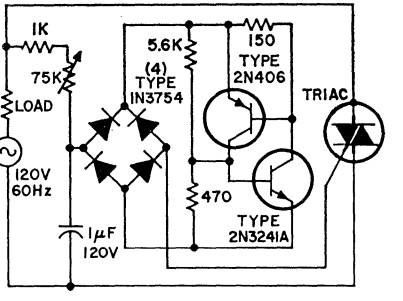


Figure 271. Circuit showing application of two-transistor switch as thyristor triggering device.

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

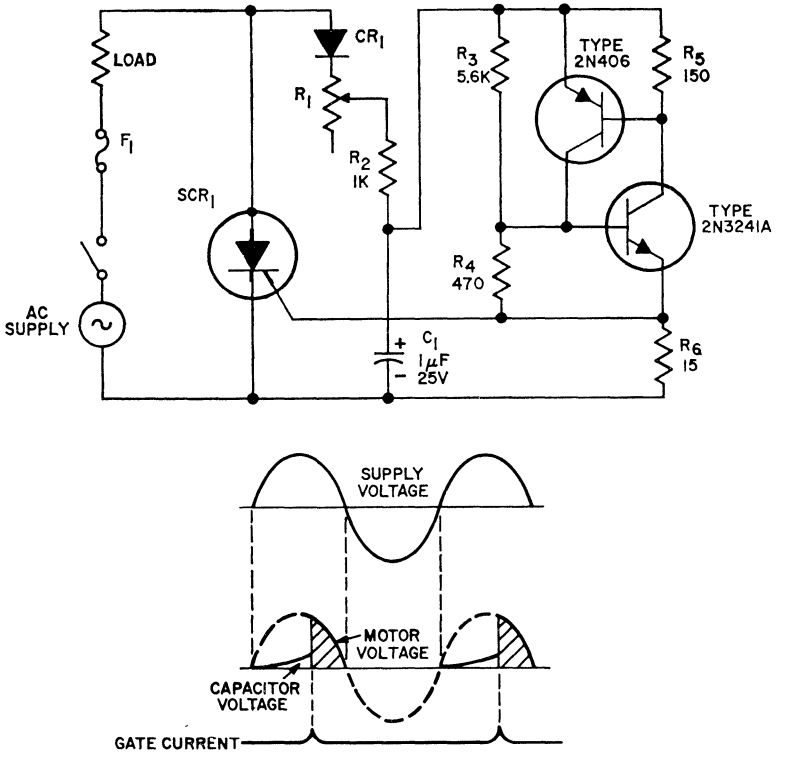


Figure 272. Half-wave SCR motor control circuit, without regulation.

Table XVII—Components For Circuit Shown in Fig. 272

AC SUPPLY	AC CURRENT	F_1	CR_1	R_1	SCR_1
120V	1A	3AG, 1.5A, Quick Act	RCA-1N3755	75K, ½W	RCA-2N3528
120V	3A	3AB, 3A	RCA-1N3755	75K, ½W	RCA-2N3228
120V	7A	3AB, 7A	RCA-1N3755	75K, ½W	RCA-2N3669
120V	25A	3AB, 25A	RCA-1N3755	75K, ½W	RCA-2N3897
240V	1A	3AG, 1.5A, Quick Act	RCA-1N3756	150K, ½W	RCA-2N3529
240V	3A	3AB, 3A	RCA-1N3756	150K, ½W	RCA-2N3525
240V	7A	3AB, 7A	RCA-1N3756	150K, ½W	RCA-2N3670
240V	25A	3AB, 25A	RCA-1N3756	150K, ½W	RCA-2N3898

triggered in either direction with positive-polarity gate signals. The only requirement is that isolation be maintained between the dc and ac current. Application of the use of the two-transistor switch to trigger thyristor control circuits is discussed further in the section on **Heater Controls.**

Application Guide for Triggering Devices—Table XVIII provides a quick reference to the prevalent types of applications for various triggering devices.

Table XVIII—Firing-Circuit Families.

Function	LOW COST Manual or Simple On-Off Power Control	MEDIUM COST Automatically Controlled or Regulated power	HIGH COST Power-Output Stage in Large Electronic or Electro-Mechanical System
Typical Applications	Light Dimmers Tool Speed Controls Appliance Speed Controls Gas Ignition Photoelectric Controls Static On-Off Switches	Regulated Power Supplies Temperature Controls Commercial DC Motor Drives Flashers Time Delays	Bulk Power Conversion for Metal Refining and Electrochemical Processes Large Industrial Motor Drives Variable-Frequency Drives Pulse Modulators Precise Process-Temperature Controls Logic-Arrays Power Output eg.—Vending Machines —Signs and Scoreboards —Computer Printer Driver
Common Characteristics of the Application	Frequently Bidirectional Small physical size an asset Low performance demands	Frugal but not poor Both dc and ac loads Thyristor fired is higher cost Technically oriented customers and applications Electrical feedback or sensor input in addition to manual control Long firing pulses often required	Firing circuit small percentage of system cost Rigid and extensive performance requirements Firing circuit often merged into other system circuits Custom engineered Primarily electrical inputs from regulators or sensors Often built up from standard logic and waveshaping circuits
Trigger Devices in Approximate Order of Preference or Use	1. Neon bulb 2. Three-layer diode 3. Five-layer diode 4. Four-layer diode 5. Unijunction transistor 6. Two-transistor regenerative circuit NOTE: For on-off control, a switch contact or single transistor may form the firing circuit	1. Unijunction transistor 2. Transistors 3. Magnetic amplifier NOTE: Firing circuit often includes several diodes, a zener, pulse transformers, control power transformers, and numerous passive components	1. Transistors

MOTOR CONTROLS

Silicon controlled rectifiers 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. 273 is a schematic of this type of motor operated from an ac supply. In reality, the universal motor is a special form of series

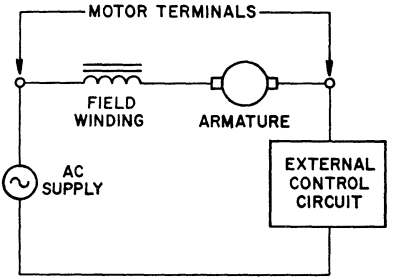


Figure 273. Schematic diagram for a series-wound universal motor.

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. 273. 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.

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. 274.

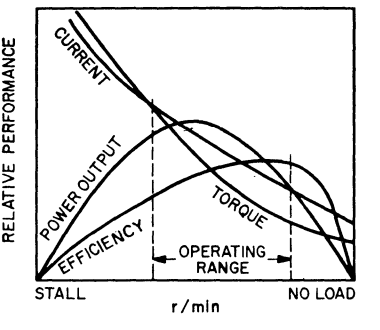


Figure 274. Typical performance curves for a universal motor.

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. 275.

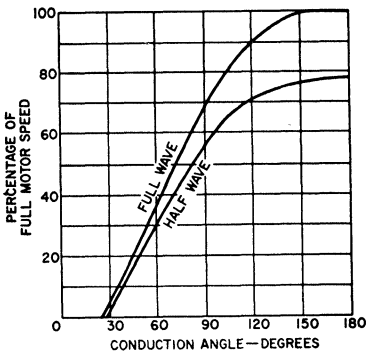


Figure 275. Typical performance curves for a universal motor with phase-angle control.

Half-Wave Control—There are many circuits available for half-wave control of universal motors; their attributes and limitations are described in detail below. Such circuits may be selected in preference to full-wave circuits for economy reasons. Motor operation with half-wave control circuits, however, in general is not as good as that obtained with full-wave control circuits. Half-wave power-control circuits are divided into two classes: regulating and non-regulating. Regulation in this instance implies load sensing and compensation of the system to prevent changes in motor speed. The half-wave proportional controls shown in Figs. 264(a) and 272 (described previously) are examples of nonregulating circuits that may be used for motor speed control.

Fig. 276 shows a fundamental circuit of a direct-coupled SCR control with voltage feedback that is highly effective for speed control of universal motors. This

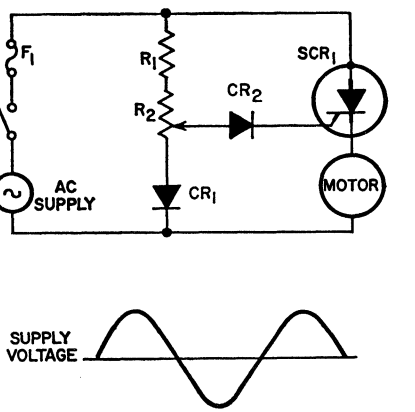


Figure 276. Half-wave SCR motor control circuit, with regulation.

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 increased 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 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

Table XIX—Components For Circuit Shown in Fig. 276

AC SUPPLY	AC CURRENT	F_1	CR_1, CR_2	R_1	R_2	SCR_1
120V	1A	3AG, 1.5A, Quick Act	RCA-1N3755	2.7K, 4W	1K, 2W	RCA-2N3528
120V	3A	3AB, 3A	RCA-1N3755	5.6K, 2W	1K, 2W	RCA-2N3228
120V	7A	3AB, 7A	RCA-1N3755	5.6K, 2W	500, 2W	RCA-2N3669
120V	25A	3AB, 25A	RCA-1N3755	2.7K, 4W	500, 2W	RCA-2N3897
240V	1A	3AG, 1.5A, Quick Act	RCA-1N3756	10K, 5W	1K, 2W	RCA-2N3529
240V	3A	3AB, 3A	RCA-1N3756	10K, 5W	1K, 2W	RCA-2N3525
240V	7A	3AB, 7A	RCA-1N3756	5.6K, 7.5W	500, 2W	RCA-2N3670
240V	25A	3AB, 25A	RCA-1N3756	5.6K, 7.5W	500, 2W	RCA-2N3898

to the load. Fig. 276 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. Circuit values for use with the various RCA SCR's are shown in Table XIX.

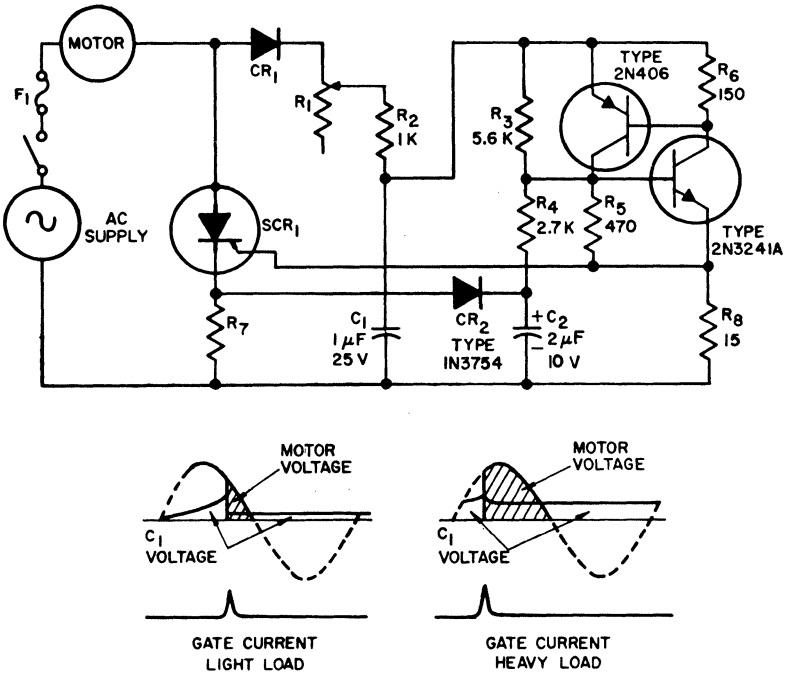


Figure 277. Half-wave SCR motor control circuit using two-transistor regenerative triggering, with regulation.

Table XX—Components For Circuit Shown in Fig. 277

AC SUPPLY	AC CURRENT	F_1	CR_1	R_1	SCR_1
120V	1A	3AG, 1.5A, Quick Act	RCA-1N3755	75K, ½W	RCA-2N3528
120V	3A	3AB, 3A	RCA-1N3755	75K, ½W	RCA-2N3228
120V	7A	3AB, 7A	RCA-1N3755	75K, ½W	RCA-2N3669
120V	25A	3AB, 25A	RCA-1N3755	75K, ½W	RCA-2N3897
240V	1A	3AG, 1.5A, Quick Act	RCA-1N3756	150K, ½W	RCA-2N3529
240V	3A	3AB, 3A	RCA-1N3756	150K, ½W	RCA-2N3525
240V	7A	3AB, 7A	RCA-1N3756	150K, ½W	RCA-2N3670
240V	25A	3AB, 7A	RCA-1N3756	150K, ½W	RCA-2N3998

Fig. 277 shows a variation of the circuit in Fig. 272. The basic difference between the two circuits is that the circuit in Fig. 277 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 XX.

Full-Wave Control—This section discusses the application of thyristors (SCR's and triacs) to provide full-wave motor control. Fig. 278 shows three thyristor full-wave controls.

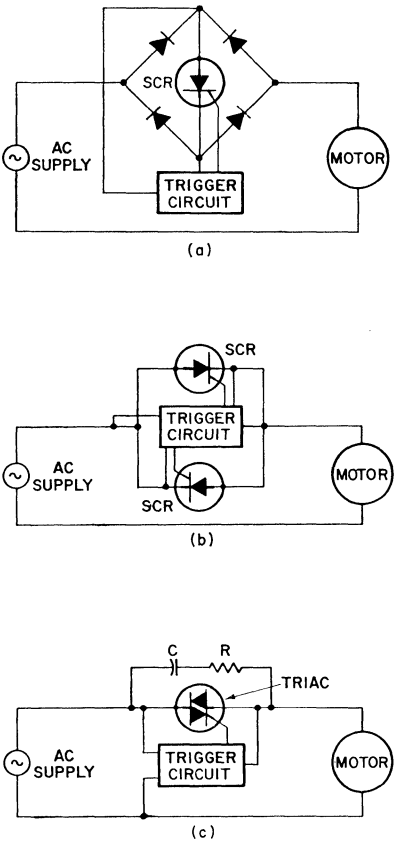


Figure 278. Full-wave thyristor motor control circuits: (a) Use of full-wave bridge rectifier to enable full-wave motor control by a single SCR; (b) Use of inverse parallel SCR's to provide full-wave motor control; (c) Use of triac to provide full-wave motor control.

The speed of induction motors can also be controlled by use of full-wave thyristor power con-

trol 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. Cooling provisions and bearings must be suitable for reduced-speed operation.

Figs. 258, 264, 266, and 278(c) show full-wave thyristor circuits suitable for use as motor-speed controls.

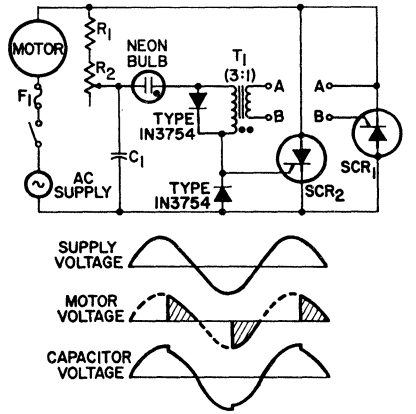


Figure 279. Full-wave SCR motor control circuit, without regulation.

A very simple SCR full-wave proportional control is shown in Fig. 279. Again, ac phase shifting and neon triggering are used to provide gate phase-angle control with a small pulse transformer utilized for isolation. The load capabilities of the circuit for various SCR's are shown in Table XXI. Conduction angles obtained with this circuit vary from 30 to 150 degrees; at the maximum conduction angle, the voltage impressed upon the load (universal motor) is approximately 95 per cent of the input rms voltage.

Fig. 280 shows a full-wave control circuit that has increased conduction angle capability. Table XXII shows the component chart for use of the circuit with various SCR's. In this circuit, the conduction angles may be varied from 5 to 170 degrees; this larger range is more desirable when higher power is to be controlled.

An SCR full-wave circuit designed for applications requiring feedback for compensation of load changes is shown in Fig. 281. Operation is similar to that of the circuits discussed previously except that this circuit has full-wave conduction with proportional control. Table XXIII gives a component list for use of this circuit with various SCR's.

Table XXI—Components For Circuit Shown in Fig. 279

AC SUPPLY	AC CURRENT	F_1	R_1	R_2	C_1	SCR_1, SCR_2
120V	1.5A	3AG, 2A, Quick Act	1K, ½W	50K, ½W	0.22 μF, 100V	RCA-2N3528
120V	5A	3AB, 5A	1K, ½W	50K, ½W	0.22 μF, 100V	RCA-2N3228
120V	10A	3AB, 10A	1K, ½W	25K, 2W	0.47 μF, 100V	RCA-2N3669
120V	25A	3AB, 25A	1K, 1W	25K, 2W	0.47μF, 100V	RCA-2N3897
240V	1.5A	3AG, 2A, Quick Act	1K, 1W	50K, 2W	0.22 μF, 100V	RCA-2N3529
240V	5A	3AB, 5A	1K, 1W	50K, 2W	0.22 μF, 100V	RAC-2N3525
240V	10A	3AB, 10A	1K, 1W	25K, 4W	0.47 μF, 100V	RCA-2N3670
240V	25A	3AB, 25A	1K, 1W	25K, 4W	0.47μF, 100V	RCA-2N3998

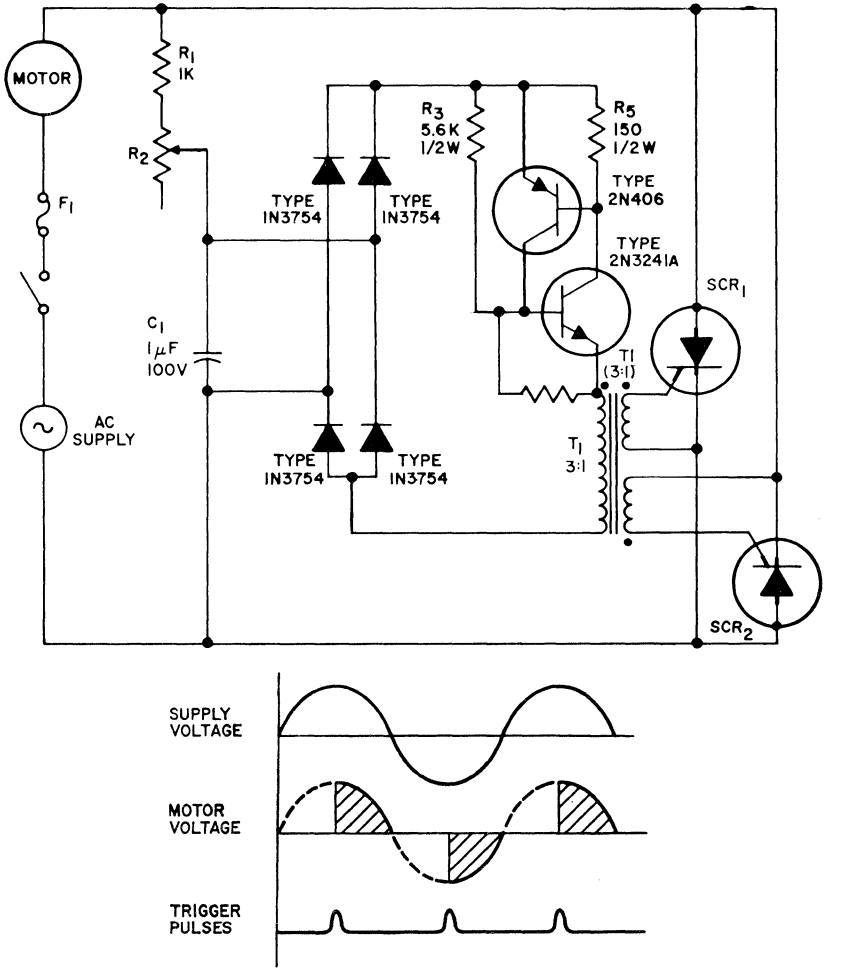


Figure 280. Full-wave SCR motor control circuit, without regulation, in which the conduction angle can be varied from 5 to 170 degrees.

Table XXII—Components For Circuit Shown in Fig. 280

AC SUPPLY	AC CURRENT	F_1	R_2	SCR_1, SCR_2
120V	1.5A	3AG, 2A, Quick Act	75K, ½W	RCA-2N3528
120V	5A	3AB, 5A	75K, ½W	RCA-2N3228
120V	10A	3AB, 10A	75K, ½W	RCA-2N3669
120V	25A	3AB, 25A	75K, ½W	RCA-2N3897
240V	1.5A	3AG, 2A, Quick Act	150K, ½W	RCA-2N3529
240V	5A	3AB, 5A	150K, ½W	RCA-2N3525
240V	10A	3AB, 10A	150K, ½W	RCA-2N3670
240V	25A	3AB, 25A	150K, ½W	RCA-2N3898

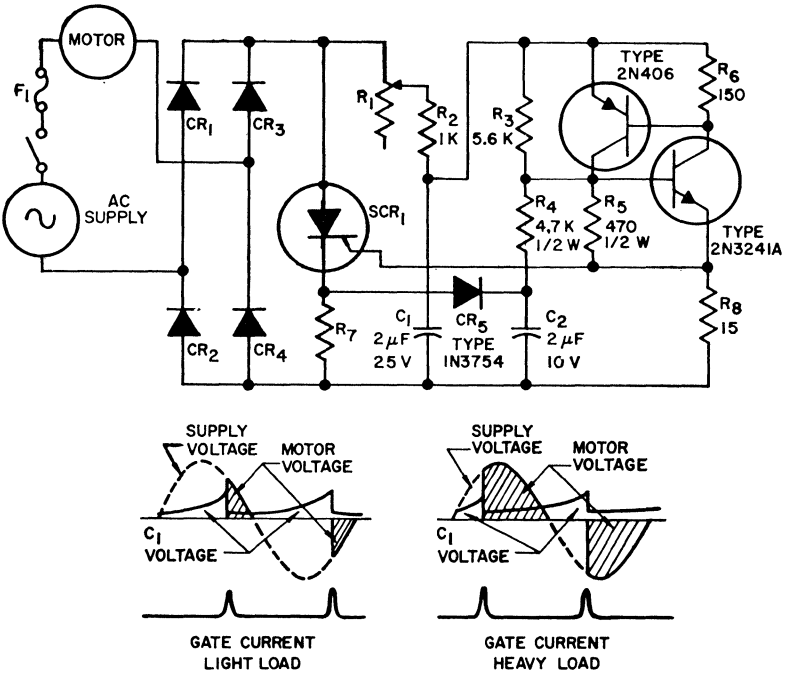


Figure 281. Full-wave SCR motor control circuit, with regulation.

Table XXIII—Components For Circuit Shown in Fig. 281

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-1N2860	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-1N2862	100K, ½W	RCA-2N3529
240V	3A	3AB, 3A	RCA-40112	100K, ½W	RCA-2N3525
240V	7A	3AB, 7A	RCA-40112	100K, ½W	RCA-2N3670

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. 282 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 re-

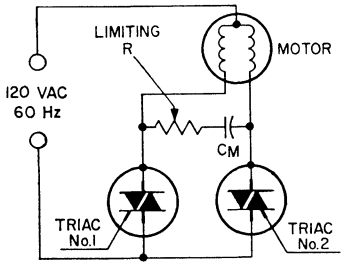


Figure 282. Motor-reversing circuit using two triacs.

versed. 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 control for door opening and closing. The block diagram in Fig. 283 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 switch disables the gate drive to the UP triac. A second keying of the transmitter provides the DOWN triac with gate drive and

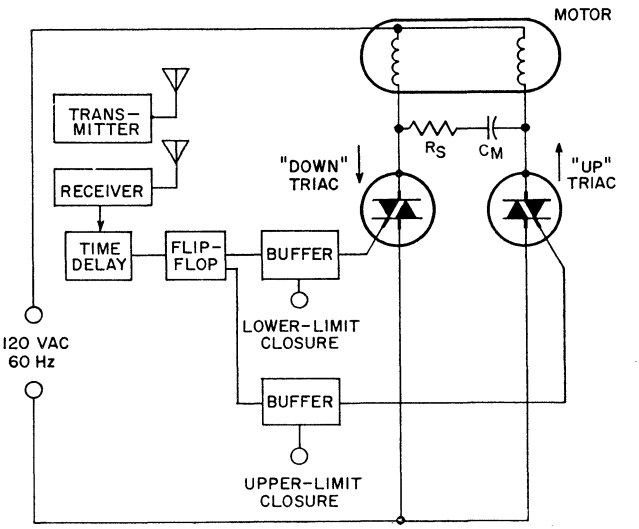


Figure 283. Block diagram of solid-state garage-door system.

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 erroneous firing. Fig. 284 shows a circuit diagram which performs the required logic for motor direction, and Fig. 285 is a timing diagram that indicates the time the time-delay monostable multivibrator must be active in order

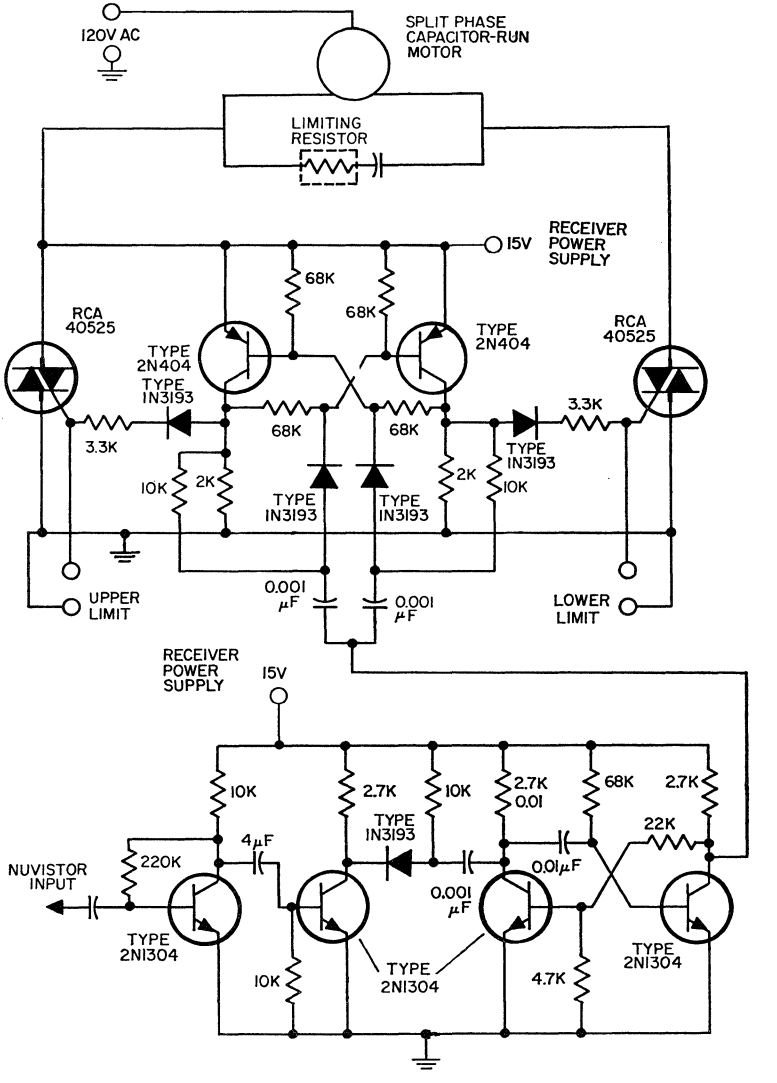


Figure 284. Garage-door control.

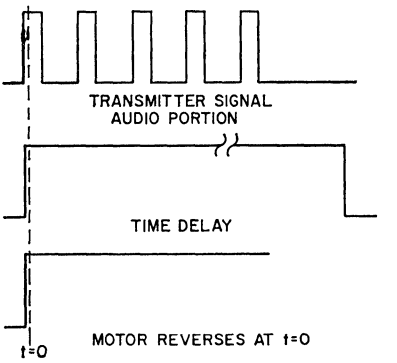


Figure 285. Timing diagram for reversing system.

to override transmitter keying. 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.

HEATER CONTROLS

Thyristors are easily adapted to the control of power to electrical heaters. Manual adjustment of the heater output may be made over an infinite range by use of phase-control circuits as described in the section on **General Considerations.** Automatic temperature regulation may be incorporated in phase-control systems through the use of a firing circuit which adjusts the conduction angle in response to feedback from a temperature sensor, such as a thermistor.

The circuits included in this section accomplish temperature regulation in a different manner. Because the heater thermal-response time is generally much slower than the period of the ac line frequency, modulation of the heater output may be accomplished by switching the thyristor fully ON for a short period and then fully OFF for a period. Through the sensitivity of the firing circuit, the temperature differential between ON and OFF states may be made extremely small. Precise temperature control is then achieved by rapid and frequent switching from ON to OFF. An advantage of this type of control is the reduction or elimination of radio-frequency interference generated by the thyristor switching, which results because the switching always occurs at or near the zero crossover point on the ac voltage wave.

The heater control shown in Fig. 286 may be used in any application in which the power requirements range from 25 watts to 600 watts maximum and temperature regulation of $\pm 0.2°C$ is adequate. For full-wave power control, the SCR is used in a rectifier bridge. The temperature at which the power is interrupted is determined by the setting of the 500-ohm potentiometer. The potentiometer sets the threshold level for the Q_1—Q_2 regenerative-switch circuit, which is connected in parallel with the Q_3 — Q_4 regenerative-switch circuit. The complete circuit is supplied with full-wave rectified voltage from the ac line through the 3000-ohm, 10-watt resistor. The SCR is turned on by the current through the Q_1—Q_2 regenerative-switch circuit after the threshold is exceeded and Q_1 and Q_2 are in conduction. The threshold level of the Q_3—Q_4 regenerative-switch cir-

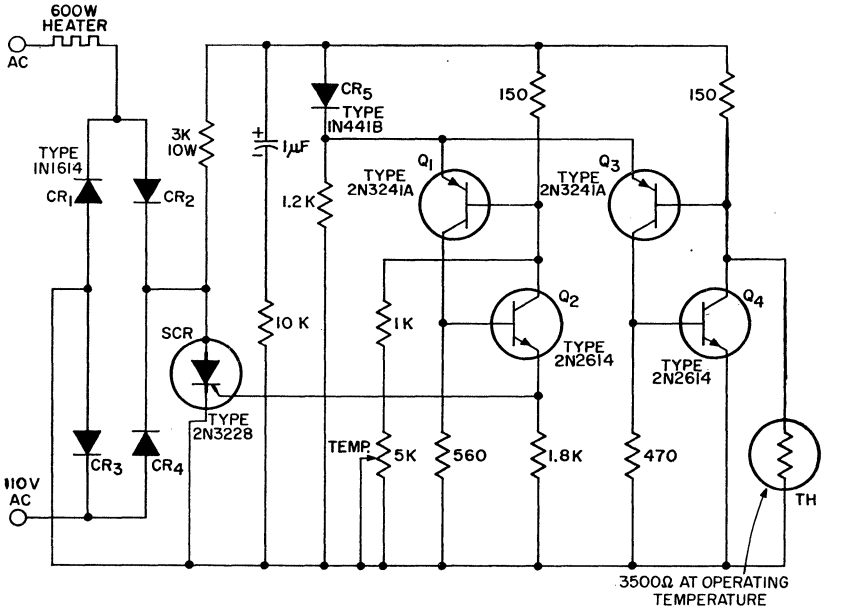


Figure 286. Half-wave SCR heater control circuit.

cuit is temperature-dependent and controlled by the thermistor. If the temperature is less than the reference temperature, the threshold level of the Q_3—Q_4 regenerative-switch circuit is higher than the threshold level of the Q_1—Q_2 circuit, and transistors Q_1 and Q_2 turn on and thus prevent the conduction threshold level required for the Q_3—Q_4 circuit from being reached. The current through the Q_1—Q_2 circuit into the gate of the SCR turns on the SCR for the rest of the half-cycle. If the temperature is higher than the set reference, the Q_3—Q_4 regenerative circuit turns on before the Q_1—Q_2 circuit, and firing of the SCR is prevented.

Fig. 287 shows a heater-control circuit that features zero-volt switching and full-wave thyristor control for loads up to 850 watts. The control circuit is the "anticipation" type that uses two thermistors as temperature sensors. For convenience, it is assumed that the resistance of the thermistor TH_2 is constant. As the voltage of the negative half-cycle (applied through a diode CR_4 and a 5600-ohm 2-watt resistor) increases to a certain magnitude (V_{th2}), it turns on transistors Q_1 and Q_2 which form a regenerative-switch circuit. Another regenerative-switch circuit composed of transistors Q_4 and Q_5 is in parallel with the Q_1—Q_2 regenerative-switch circuit. The threshold level of the $Q_4 - Q_5$ circuit, V_{th2}, is a function of the resistance of thermistor TH_1, which senses the temperature to be regulated. If the temperature on thermistor TH_1 increases, the resistance of the thermistor decreases and the threshold level V_{th1} becomes higher. If the threshold level V_{th1} is higher than the threshold level V_{th2}, the regenerative-

circuit transistors Q_1 and Q_2 turn on before the Q_4–Q_5 regenerative circuit. The current through diode CR_3 and the 4700-ohm, 2-watt resistor during the negative cycle charges the 20-microfarad capacitor. During the positive cycle, or slightly earlier, the 10-microfarad capacitor starts to discharge through the 470-ohm resistor and the gate of the SCR_2 thyristor. The discharge current into the gate of SCR_2 triggers the device. As SCR_2 turns on, it closes the circuit for the 30-microfarad capacitor connected to the anode of SCR_2. The current through the 4700-ohm, 2-watt resistor and diode CR_1 charges the capacitor. The capacitor starts to discharge through the 270-ohm resistor and the gate of SCR_1 prior to the negative half-cycle. SCR_1 turns on for the negative half-cycle to complete a full cycle of operation.

If the regenerative circuit composed of transistors Q_1 and Q_2 turns on and a portion of the emitter current of Q_2 flows into the base of transistor Q_3, transistor Q_3 also turns on. As transistor Q_3 turns on, it prevents the 20-microfarad capacitor from being charged; thus, SCR_2 is not fired for the positive half-cycle. The resistance of thermistor TH_2 (mounted in proximity to the heater element) is not constant (as had been assumed previously). The threshold level V_{th2}, which serves as the reference level to which the V_{th1} threshold level is compared, is a function of the resistance of thermistor TH_2. As the temperature increases on thermistor TH_2, the threshold level V_{th2} decreases. When V_{th2} becomes smaller than V_{th1}, the power from

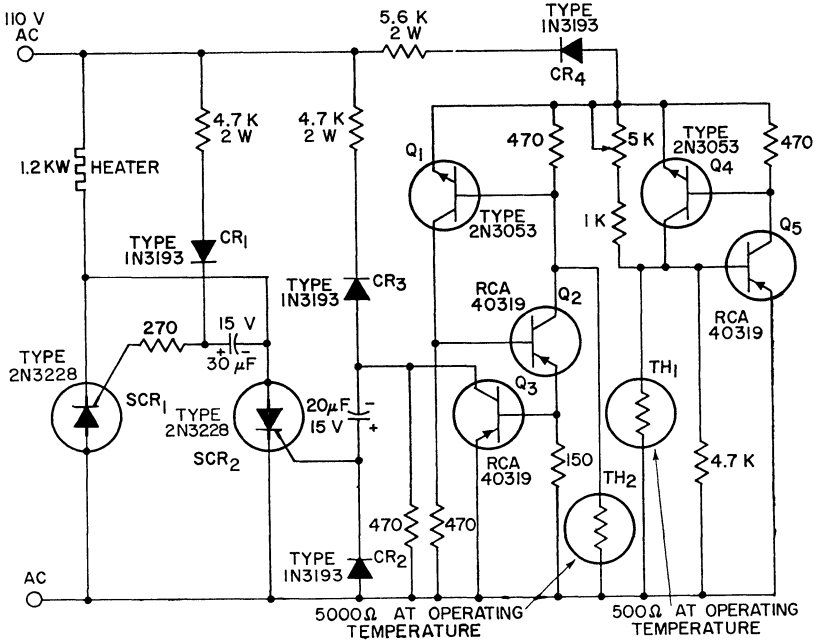


Figure 287. Full-wave zero-switching SCR heater control circuit.

the heater is turned off. This condition occurs somewhat earlier in time than if the resistance of thermistor TH_2 were constant.

INCANDESCENT LIGHTING CONTROLS

A popular application of thyristors is in controls for incandescent lighting systems. Lamp-dimmer controls make use of the basic phase-control techniques described earlier in the section **General Considerations** to vary the effective voltage applied across the lamp load.

Firing Controls for Light Dimmers

The basic circuits used in dimmer controls can be grouped into two categories: single-time-constant and double-time-constant circuits. A single-time-constant circuit is shown in Fig. 288. The operation of the circuit was described previously in the section on **Triggering Devices.**

An interesting aspect of the operation of this circuit for light dimming is the minimum level of illumination that can be achieved as the control is adjusted from its full OFF position. 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 ΔV. 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 resistive 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

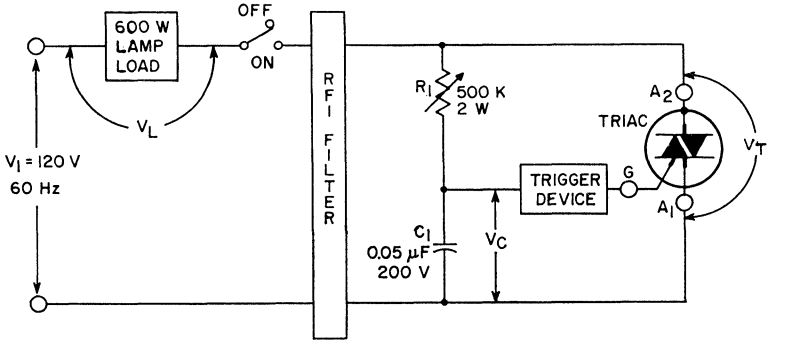


Figure 288. Single-time-constant phase-control circuit for brightness control of incandescent lamps.

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 needed, R_{TO}, to turn off the triac completely is greater than the value R_{IC} required to the triac 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." The hysteresis effect can be reduced by use of a double-time-constant circuit configuration. It should be noted that hysteresis and "quick-turn-on" are characteristics of the single-time-constant circuit.

The two basic forms of the double-time-constant circuit are shown in Figs. 289 and 290. Both these circuits produce less hysteresis and lower 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

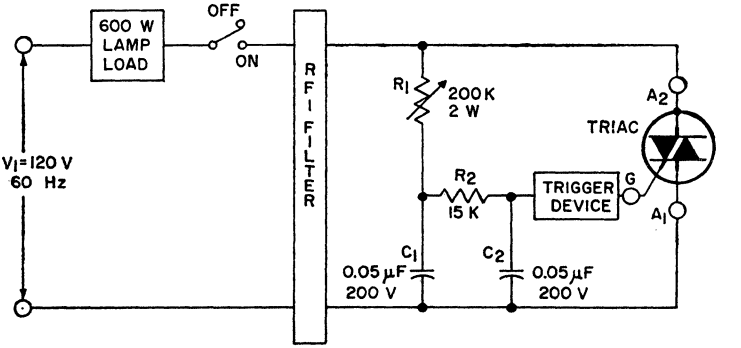


Figure 289. Double-time-constant circuit for brightness control of incandescent lamps in which triac is cut off when potentiometer is set for maximum resistance.

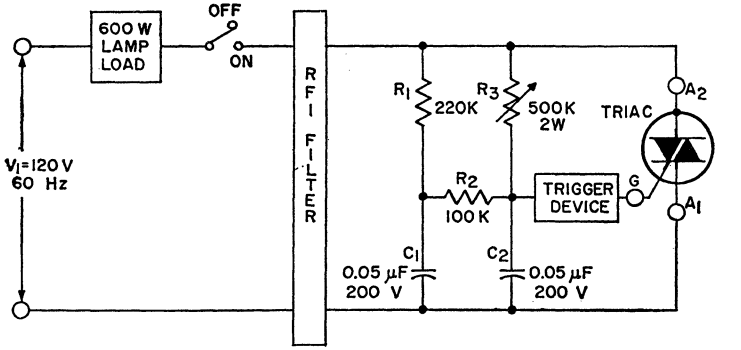


Figure 290. Double-time-constant phase-control circuit for brightness control of incandescent lamps in which a small amount of triac current flows when potentiometer is set for maximum resistance.

in Fig. 291. 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

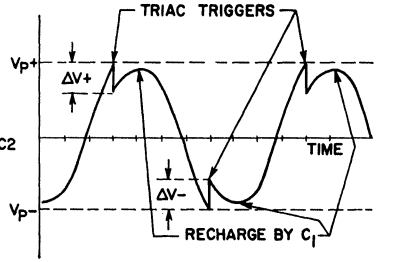


Figure 291. Voltage waveform across triggering capacitor C_2 in double-time-constant circuits shown in Figs. 289 and 290.

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. 290 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.

In the lamp-dimmer circuit shown in Fig. 292, a resistor R_3 is connected in series with the triggering diode 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.

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 de-

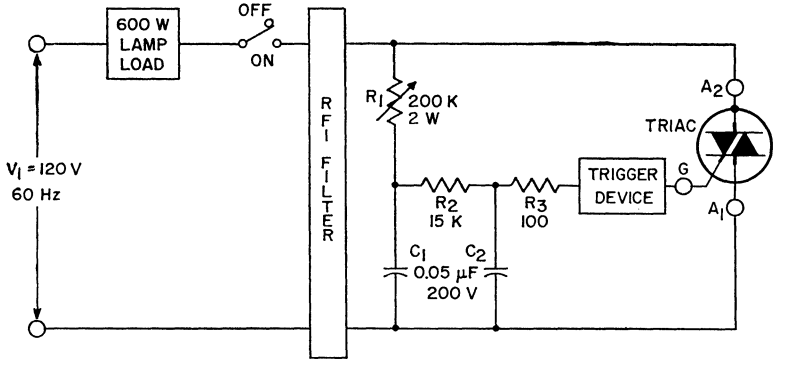


Figure 292. Double-time-constant circuit that uses a series gate resistor.

signed 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. 293 and 294.

The circuit shown in Fig. 293 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. 294 causes the lamp load to turn off as the ambient light level increases. 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

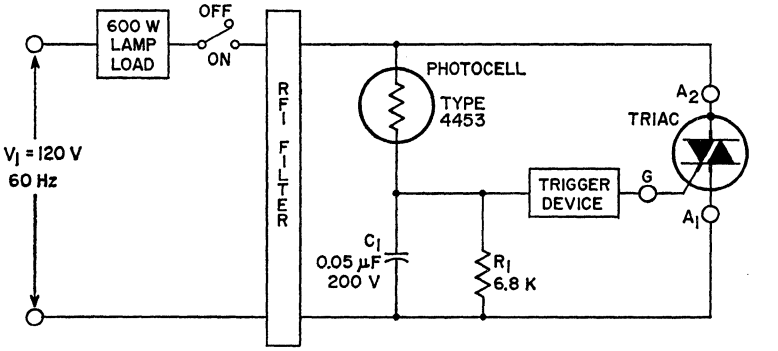


Figure 293. Photocell-operated on-off lamp control that energizes lamp load when photocell is illuminated.

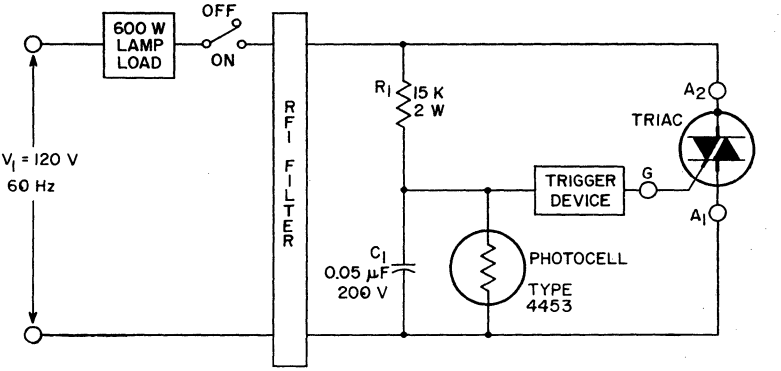


Figure 294. Photocell-operated lamp control circuit that energizes lamp load when photocell is not illuminated.

turn on at dusk and turn off at dawn automatically.

Performance Characteristics of Light Dimmers

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 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. 295. 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.

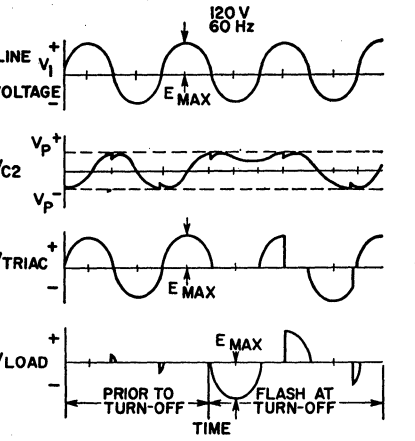


Figure 295. Waveforms during flash-at-turn-off condition in a double-time-constant phase-control circuit.

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}.

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. 259, 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. The relationship of these factors was described in the previous discussion, and the addition of a fixed-value series gate resistor was shown to produce less hysteresis than was obtained from the same circuit with series gate resistor.

Load Considerations

An important consideration in lamp-dimmer circuits is the load and its effect on the requirements for the triac. Obviously, the triac must be capable of handling the steady-state load current. In addition, the triac should be capable of handling the transient load requirements. One of the transient requirements of incandescent lamp loads is the initial or cold-lamp **inrush current,** which may include some large peak values. The magnitudes of the current peaks are a function of the instantaneous value of the line voltage and the setting of the dimmer. Fig. 295 illustrates the current waveforms for various initial values of line voltage and dimmer settings. These waveforms show that the largest peak currents are encountered when the combination of line voltage and dimmer settings is such that load current commences at the peak of the line-voltage waveform. For this case, the ratio of initial peak current to the steady-state peak value is approximately 10 to 1, and can rise as high as 15 to 1 for large-wattage lamps. This relationship implies that the triac chosen for a particular lamp load should have a subcycle surge capability sufficient to allow repeated passage of the cold inrush current without degradation of the device.

Another phenomenon associated with the incandescent lamp load is **flashover.** Flashover refers to the internal arc which is developed in the lamp when it burns out. Essentially, during burn-out, the lamp filament opens and an arc is initiated across the broken ends of the filament. Subsequently, the arc can transfer to the internal lead-in wires and remain there until it self-extinguishes or the circuit is opened. Because the currents associated with flashover can be large, the lamps have internal fuses which are designed to act quickly and open the circuit to remove the current and extinguish the arc. Damage or degradation of the triac because of an incandescent-lamp flashover can usually be avoided by selection of a triac that has a subcycle current capability which is compatible with, or in excess of, the I^2t rating of the lamp fuse. If this condition cannot be achieved in a particular application, it may be necessary to use external resistors or other devices to prevent the flashover current from exceeding the current-handling capacity of the triac.

High-Frequency Power Amplifiers

The advent of silicon rf power transistors capable of supplying up to 20 watts of output power at frequencies as high as 400 MHz has resulted in the increasingly widespread use of transistors in many commercial, industrial, and military transmitting applications. During the past few years, silicon power transistors have demonstrated the ability to provide good operating efficiency, high power gain, and excellent temperature and frequency stability when used in class A, B, or C rf power amplifiers. This capability, together with the small size and compact structure of the silicon transistors, has led to the use of these devices in many power-amplifier applications that previously were considered the sole province of electron tubes.

DESIGN OF 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 modulation requirements of transistor rf power amplifiers differ slightly from those for tube amplifiers. 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 either transistor or vacuum-tube power amplifiers. In some applications, multiple connections of the silicon power transistors may be required to develop the desired amount of output power. 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 dissi-

pation of a class A stage are high, and care must be exercised to assure thermal stability.

In applications, such as single-sideband transmitters, that require good linearity, 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. 296 shows four methods that may be used to reverse-bias a transistor stage.

Fig. 296(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

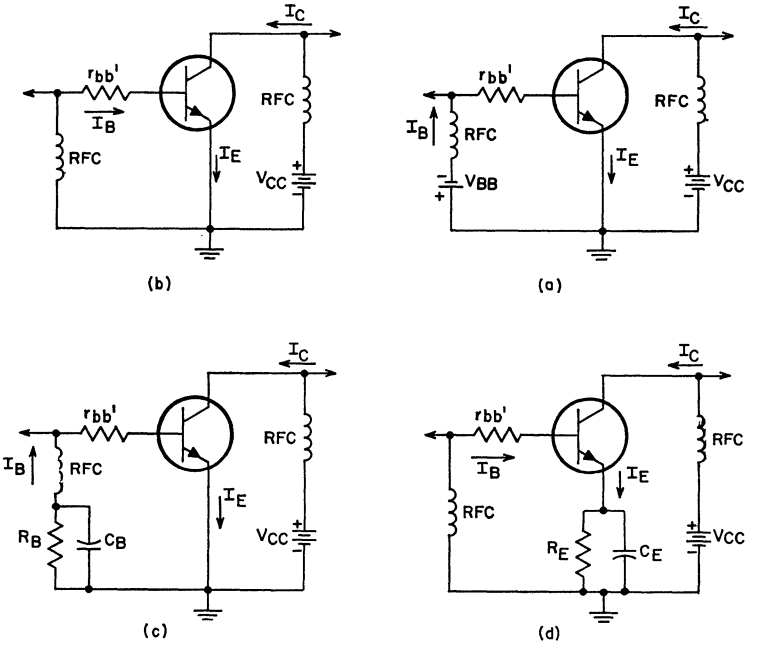


Figure 296. Methods for obtaining class C reverse bias: (a) by use of fixed dc supply V_{BB}; (b) by use of dc base current through the base spreading resistance $r_{bb'}$; (c) by use of dc base current through an external base resistance R_B; (d) by use of self bias developed across an emitter resistor R_E.

be difficult to obtain in many applications. In addition, the bypass elements required for the separate supply increase the circuit complexity.

Figs. 296(b) and 296(c) show methods in which reverse bias is developed by the flow of dc base current through a resistance. In the case shown in Fig. 296(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. 296(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. 296(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.

Matching Requirements

A simplified high-frequency equivalent circuit of an "overlay" type of transistor is shown in Fig. 297. This circuit is similar to the hybrid-pi equivalent circuit

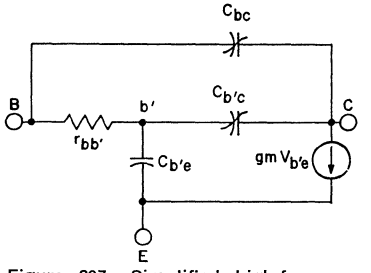


Figure 297. Simplified high-frequency equivalent circuit for an "overlay" transistor.

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 Parameters

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 determina-

tion 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

In many instances, components and constructional techniques used for rf vacuum-tube equipment applications are not suitable for rf transistor circuits. Primarily, this incompatibility between tube and transistor requirements results because of the substantially lower circuit impedances encountered in transistor circuits.

Probably the most troublesome area in high-frequency transistor circuits is frequency stability. 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. 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 tran-

sistor 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 resistance R_s and 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:

$$R_L = (V_{CC})^2/2P_o \quad (288)$$

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. 298.

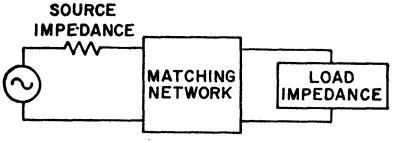


Figure 298. Basic components considered in the design of a matching network.

For the input matching network, the source is assumed to be a generator that has a 50-ohm impedance. For the output matching network, the source is the output of the transistor, which can be approximated as shown in Fig. 299.

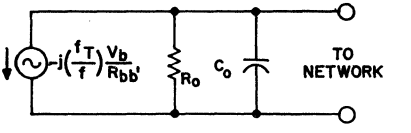


Figure 299. Equivalent circuit for the output of a transistor.

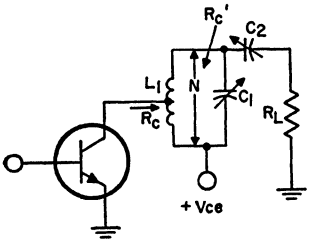
Output-Circuit Design—When the dc supply voltage and power output are specified, the circuit designer must determine the load

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. 300 and 301 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



FOR N:1 TURN RATIO

(1) $R_c = \frac{V_{ce}^2}{2P_O}$ (FOR CLASS C)

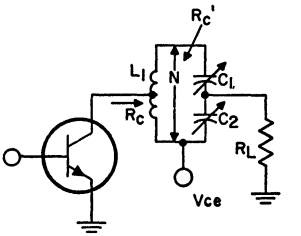
(2) $X_{L1} = \frac{R_c'}{Q_L} = \frac{N^2 R_c}{Q_L}$

(3) $X_{C2} = R_L \sqrt{\frac{N^2 R_c}{R_L} - 1}$

(4) $X_{C1} = \frac{N^2 R_c}{Q_L} \cdot \frac{1}{\left(1 - \frac{X_{C2}}{Q_L R_L}\right)}$

Figure 300. Tuned-circuit output coupling method and design equations in which output is transferred to load by a series coupling capacitor.

to the tuned circuit. The transformed R_L across the entire tuned circuit is stepped down to match the collector by the proper turns



FOR N:1 TURN RATIO

(1) $R_c = \frac{V_{ce}^2}{2P_O}$ (FOR CLASS C)

(2) $X_{L1} = \frac{N^2 R_c}{Q_L}$

(3) $X_{C1} = \frac{N^2 R_c Q_L}{(Q_L^2 + 1)} \left[1 - \frac{R_L}{Q_L X_{C2}}\right]$

(4) $X_{C2} = \frac{R_L}{\sqrt{\frac{(Q_L^2 + 1) R_L}{N^2 R_c} - 1}}$

Figure 301. Tuned-circuit output coupling method and design equations in which output to the load is obtained from a capacitive voltage divider.

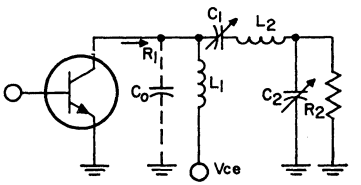
ratio of the coil L_1. If the value of the inductance L_1 is chosen properly and the portion of the output-coil inductance between the collector and ground is sufficiently high, the harmonic portion of the collector current in the tuned circuit is small. Therefore, the contribution of the harmonic current 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.

Although the circuits shown in Figs. 300 and 301 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. 302 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. 302 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. 303 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. 303(a) and is included as part of L_2 for networks of the type shown in Fig. 303(c). In Fig. 303(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. 303(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. 303(a) and 303(b) require the

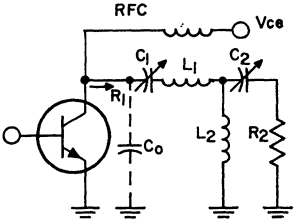


(a) FOR $R_1 < R_2$

(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) $X_{L_2} = Q_L R_1 \left[1 + \dfrac{R_2}{Q_L X_{C_2}}\right]$



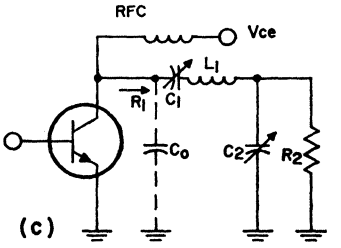
(b) FOR $\dfrac{Q_L X_{Co}}{\sqrt{R_1 R_2}} > 1$

(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) $X_{C_2} = \dfrac{R_2}{Q_L}\left[\dfrac{Q_L X_{Co}}{\sqrt{R_1 R_2}} - 1\right]$

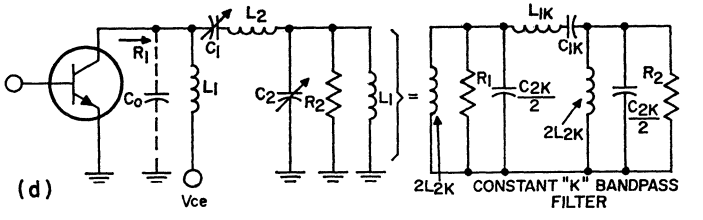


(c)

(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) $X_{L_1} = \dfrac{Q_L X_{Co}^2}{R_1}\left[1 + \dfrac{R_2}{Q_L X_{C_2}}\right]$

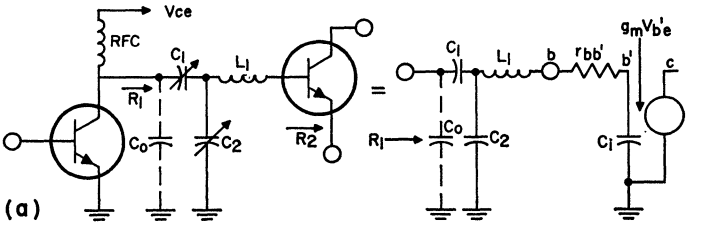


LET $C_{2K} = 2\,C_{OUT}$; $R_1 = R_2$; f_1 = LOW FREQ. CUTOFF; f_2 = HI-FREQ. CUTOFF

(1) $(f_2 - f_1) = \dfrac{1}{2\pi C_o R_L}$ (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) $C_2 = C_o = \dfrac{C_{2K}}{2}$

Figure 302. Additional transistor output-coupling networks including transistor output capacitance.

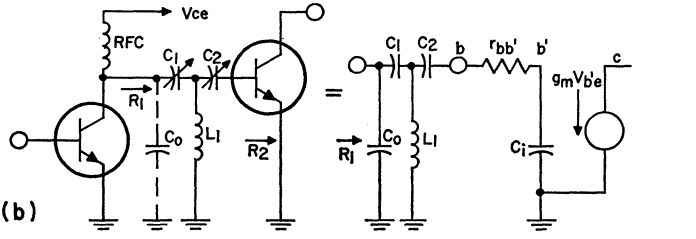


FOR $X_{L_1} >> X_{c_1}$; $R_1 > R_2 = r_{bb'}$

(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) $X_{c_2} = \frac{r_{bb'}(Q_L^2+1)}{Q_L} \cdot \frac{1}{\left[1-\sqrt{\frac{R_1 r_{bb'}(Q_L^2+1)}{X_{co}^2 Q_L^2}}\right]}$

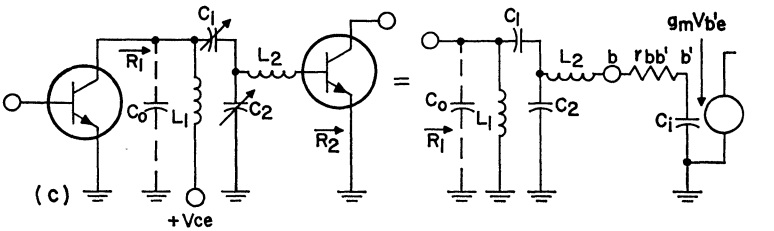


FOR $X_{c_2} >> X_{c_1}$; $R_1 > R_2 = r_{bb'}$

(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 $R_1 > R_2$; $R_2 = r_{bb'}$; $X_{L_2} \gg X_{c_1}$

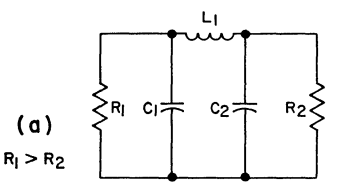
(1) $X_{L_1} = \frac{R_1}{Q_L}$

(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 303. Transistor input-circuit coupling networks.



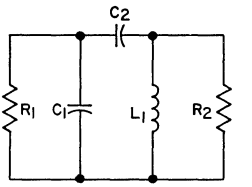
(a)

$R_1 > R_2$

(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) $X_{L_1} = \frac{Q_L R_1}{Q_L^2+1}\left[1+\frac{R_2}{Q_L X_{C_2}}\right]$



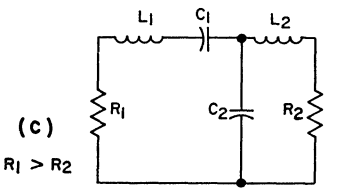
(b)

$R_1 < R_2$

(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)

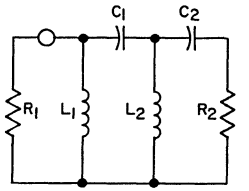
$R_1 > R_2$

(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)

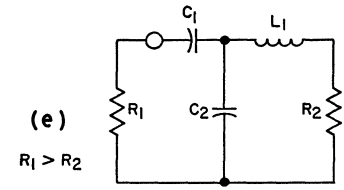
$R_1 > R_2$

(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) $X_{C_2} = Q_L R_2$



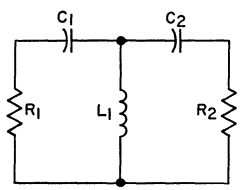
(e)

$R_1 > R_2$

(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 304. Other suitable rf-amplifier coupling networks for maximum power transfer.

collector of the driving transistor to be shunt-fed by a high-impedance rf choke. Fig. 303(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. 304 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 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. 305 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 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 XXIV, which, together with Fig. 306, indicates the direction of travel along the curve and makes it unnecessary to determine the plus or minus sign 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. 307. The node that has the highest Q dominates; this Q is then defined as the operating Q of the circuit.

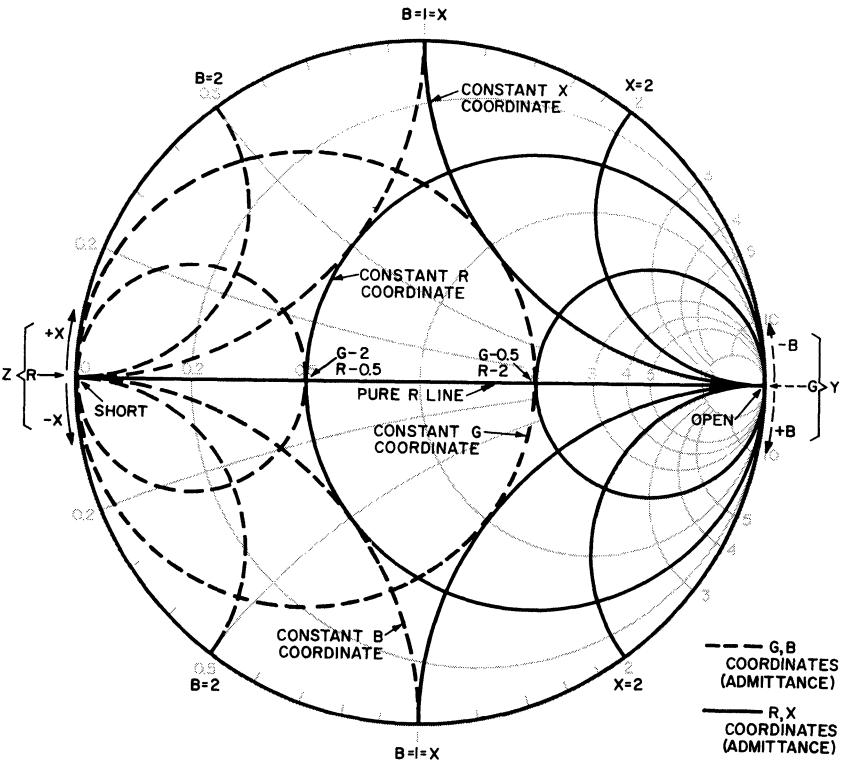


Figure 305. Impedance-admittance chart.

Table XXIV—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.

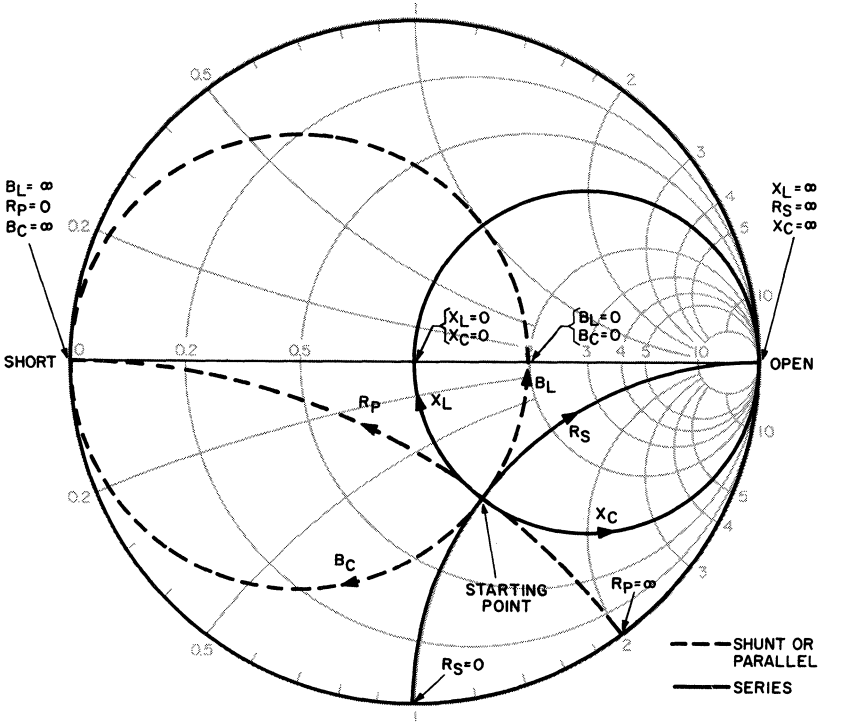


Figure 306. Method used to trace constant-parameter curves for matching-network components.

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 admittances represented by the normalized values.) The ohm (Ω) and mho (υ) symbols are retained on the chart to distinguish between impedance and susceptance. A 50-ohm impedance is used as the normalizing factor 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. 305, 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. 308 and 309 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. 308 on Fig. 309 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 standard 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.

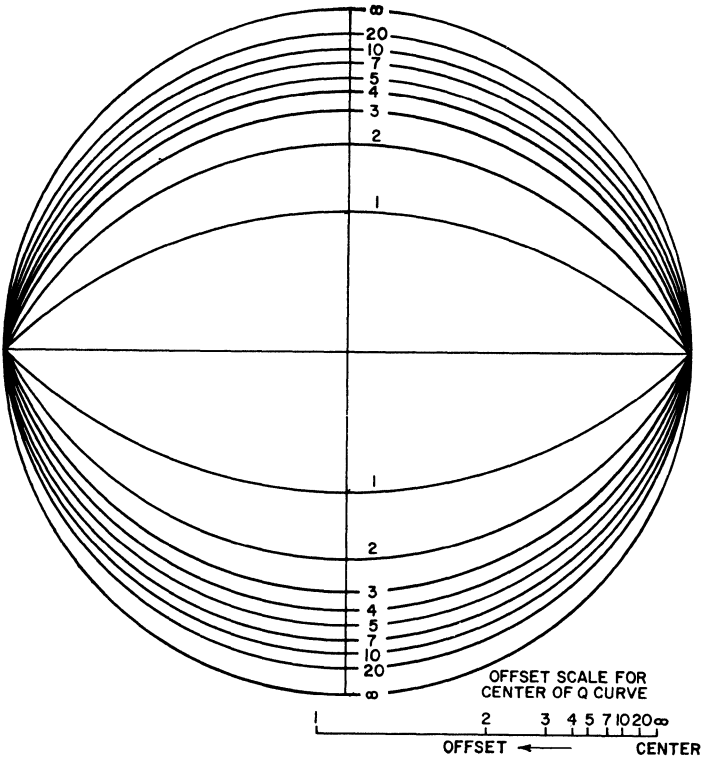


Figure 307. Chart of constant-Q curves.

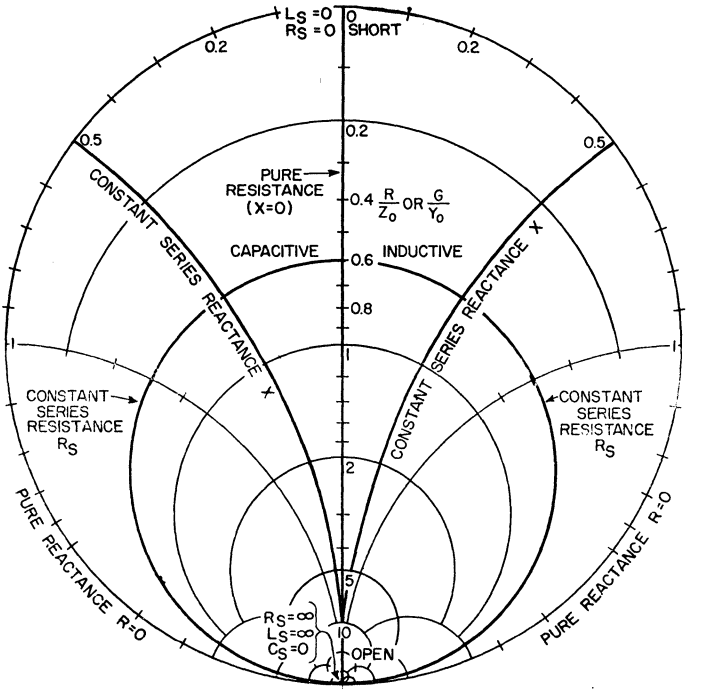


Figure 308. Impedance chart

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. 310 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 reactions 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. 310. 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 XXIV 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. 310. 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 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 XXIV.)

1. **Design of tapped-C network.** Fig. 311 shows the circuit configuration and the network map

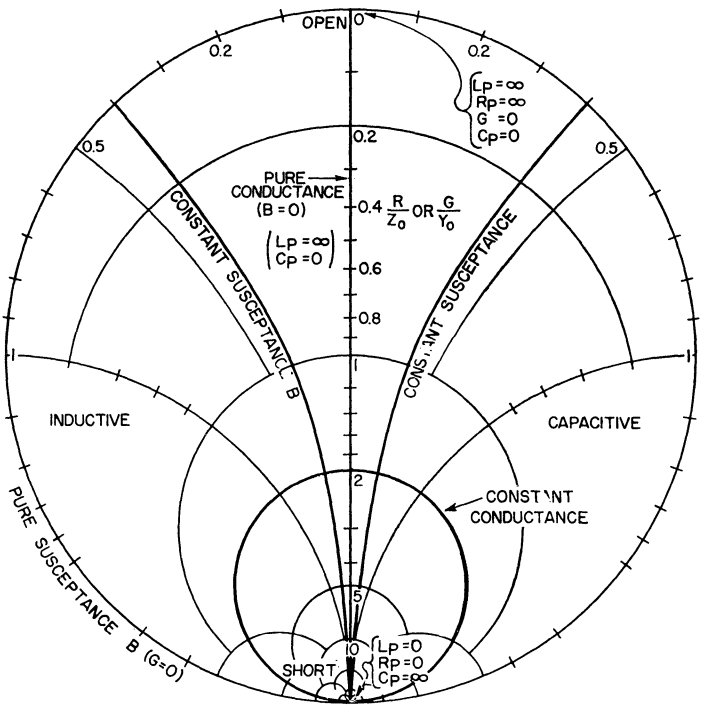


Figure 309. Admittance chart.

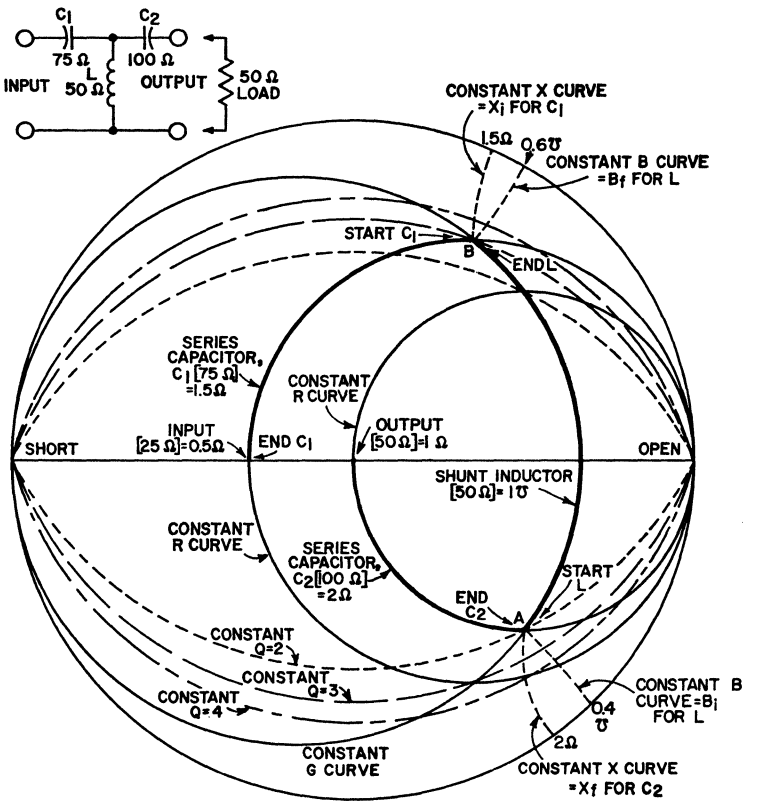


Figure 310. Typical output matching network and network map used to determine required input impedance for this network.

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 from the Q chart, Fig. 307.

(c) The curve for the shunt L is traced along the constant R_p = 1.0υ curve (from the admittance chart) from the termination point 1 + j0Ω 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 + j0Ω 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 XXIV, 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:

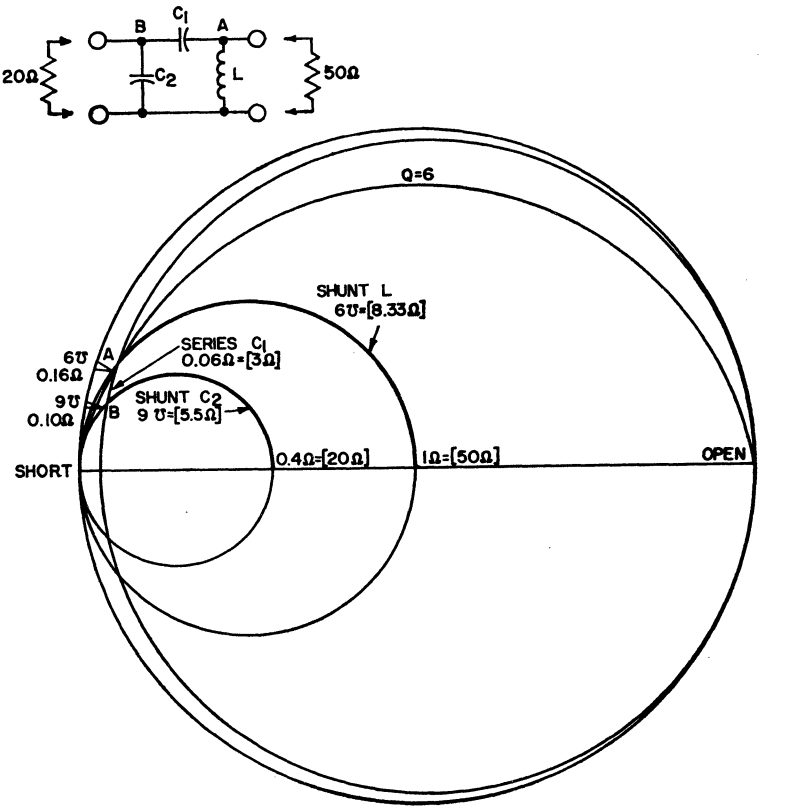


Figure 311. Circuit configuration and network map used to determine the component values for a tapped-C matching network.

$$\Delta B_L = B_{(at\ A)} - B_{(at\ 1+j0\Omega)} = 6\upsilon - 0 = 6\ \text{mhos}$$

$$\Delta X_{c_1} = X_{(at\ B)} - X_{(at\ A)} = 0.10 - 0.16 = 0.06\ \text{ohm}$$

$$\Delta B_{c_2} = B_{(at\ 0.4+j0\Omega)} - B_{(at\ B)} = 0 - 9\upsilon = 9\ \text{mhos}$$

(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 XXV. 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 determined 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. 312 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

Table XXV—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; $X_{C_1} = \Delta X(50) = 3\Omega$
9	Shunt C_2	B	(0.4 + j0)Ω; ΔB = 0 − 9 = 9υ; $X_{C_2} = 50/\Delta B = 5.5\Omega$

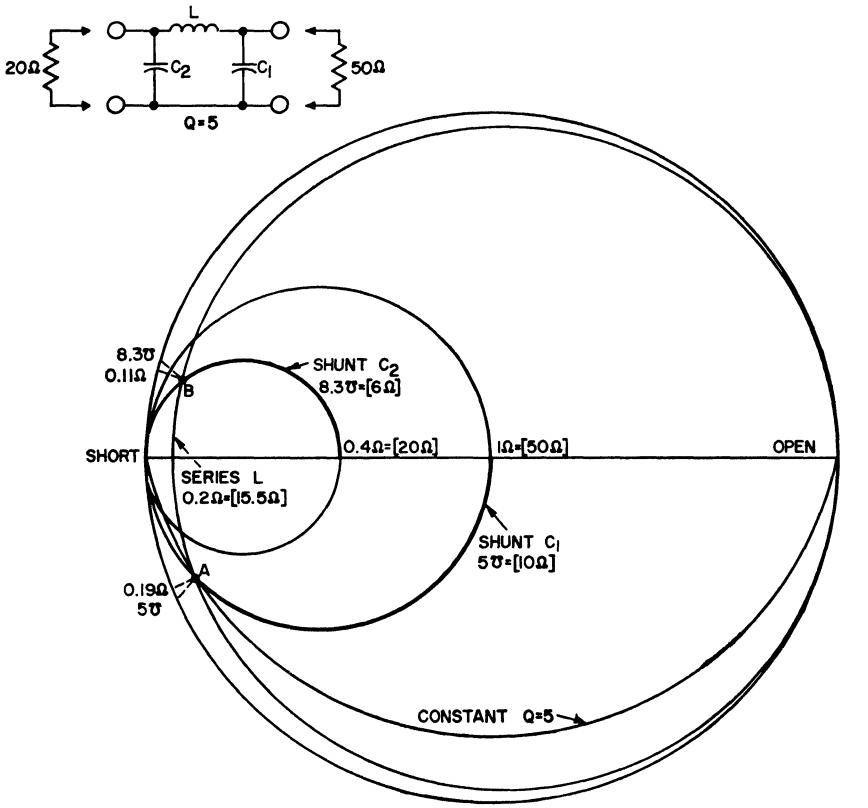


Figure 312. Circuit configuration and the network map used to determine the component value for a pi matching network.

Table XXVI—Procedure for Determining Component Values for Pi Matching Network

Step No.	Component Curves 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	5υ	0.19Ω
6	B	8.3υ	0.11Ω

	Compute Component Values COMPONENT	INITIAL POINT	FINAL POINT	PARAMETER CHANGE (NORMALIZED VALUES)	COMPONENT VALUE
7	Shunt C_1	(1 + j0)Ω	A	ΔB = 5 − 0 = 5υ	$X_{C1} = 50/\Delta B = 10\Omega$
8	Series L	A	B	ΔX = 0.11 + 0.19 = 0.3Ω	$X_L = 50(\Delta X) = 15\Omega$
9	Shunt C_2	B	(0.4 + j0)Ω	ΔB = 0 − 8.33 = 8.33υ	$X_{C2} = 50/\Delta B = 6\Omega$

determine the component values for the pi network are given in Table XXVI.

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. 313 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 XXVII 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

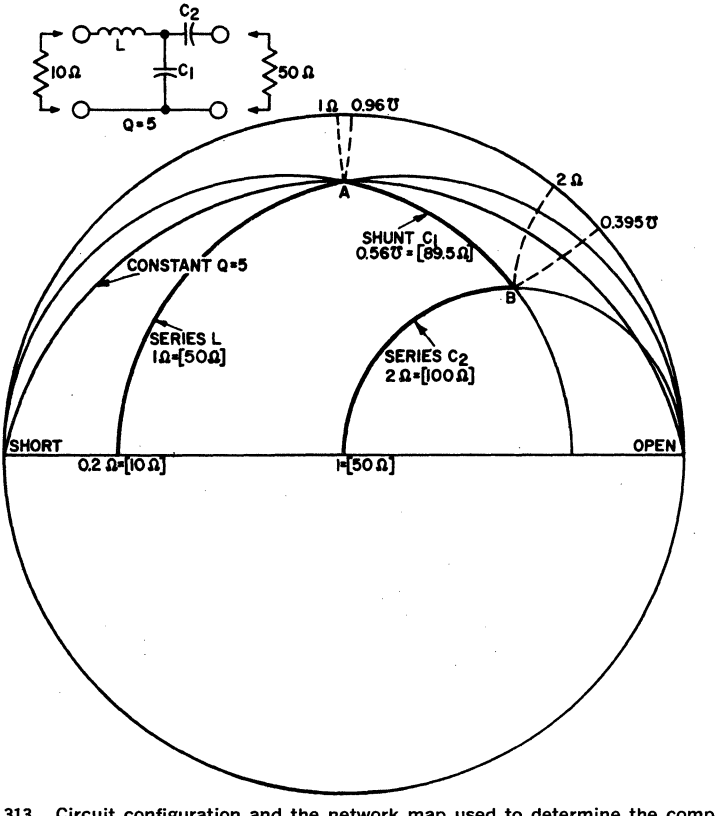


Figure 313. Circuit configuration and the network map used to determine the component values for a "lossy"-L matching network.

Table XXVII—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	ΔX = 1.0 − 0 = 1.0Ω	X_L = 50Ω
C_1	A,B = 0.96	B, B = 0.395	ΔB = 0.395 − 0.96 = 0.56υ	X_{CL} = 89.5Ω
C_2	B,X = 2.0	1.0	ΔX = 0 − 2.0 = 2.0Ω	X_{C2} = 100Ω

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.

4. **Design of network containing four unspecified components.** Fig. 314 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 purposely 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 XXVIII.

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

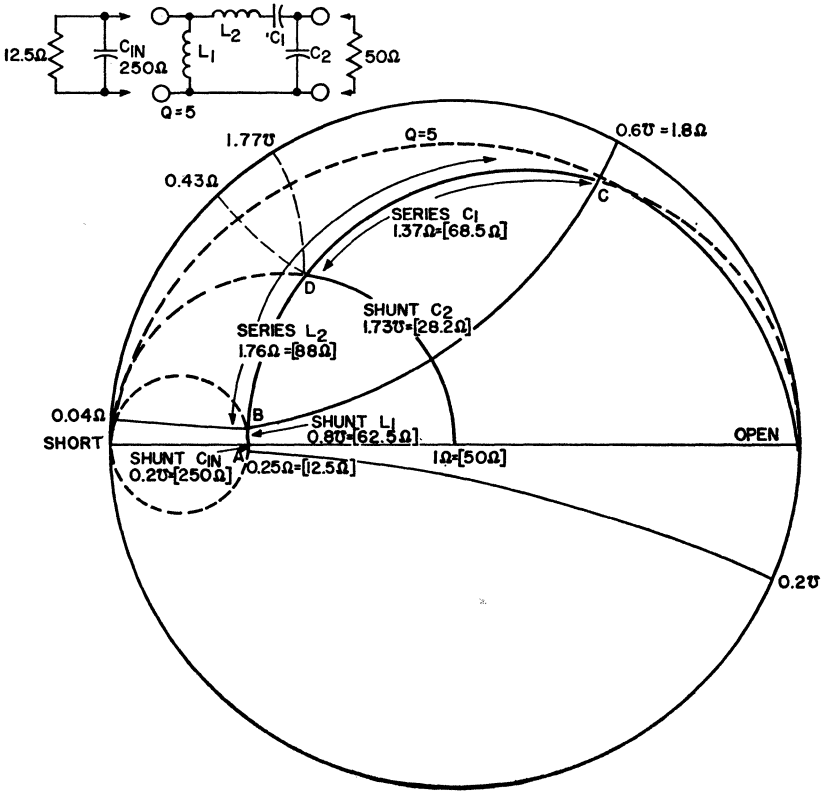


Figure 314. Circuit configuration and network map used to determine component values for matching network that includes four unspecified components.

Table XXVIII—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_{1n}	$(0.25 + j0)\Omega$	A, determined by given value for XC_{1n}
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_{1n}	0.25Ω	A,B = 0.2υ	ΔB = 0.20 = 0.2υ	X_{C1n} = 250Ω
L_1	A,B = 0.2υ	B,B = 0.6υ	ΔB = 0.6 + 0.2 = 0.8υ	X_{L1} = 62.5Ω
L_2	B,X = 0.04Ω	C,X = 1.8Ω	ΔX = 1.8 − 0.04 = 1.76Ω	X_{L2} = 88Ω
C_1	C,X = 1.8Ω	D,X = 0.43Ω	ΔX = 0.43 − 1.8 = 1.37Ω	X_{C1} = 68.5Ω
C_2	D,B = 1.77υ	1.0Ω	ΔB = 0 − 1.77 = 1.77υ	X_{C2} = 28.2Ω

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. 312) 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. 315. 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

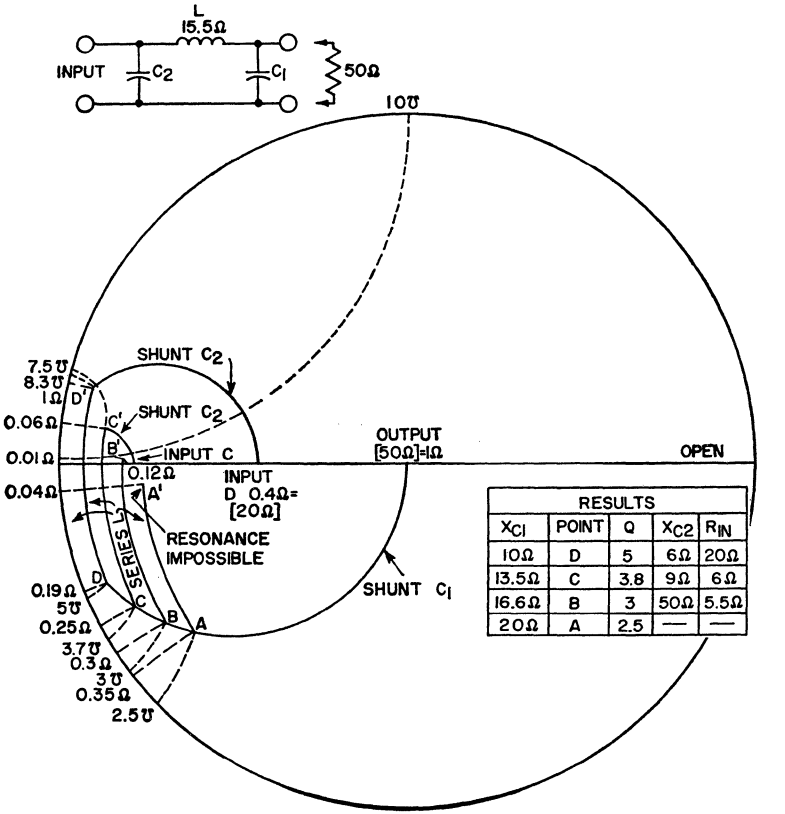


Figure 315. Circuit configuration for pi matching network and network map showing tuning range for variable components used in this type of matching network.

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.

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. In the uhf and microwave-frequency ranges it may be more desirable to use short sections of transmission lines 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:

$$Z_o = (R_S \times R_L)^{1/2} \quad (289)$$

where R_S is the source or input impedance and R_L is the load or output impedance.

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_o 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:

$$Z_o = |Z_s| = (R_S^2 + X_S^2)^{1/2} \quad (290)$$

The real impedance Z_L can be determined from the Smith Chart or by use of the following equation:

$$Z_L = \frac{R}{1 - \frac{X}{|Z_S|}} \quad (291)$$

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:

$$Z_{in} = 7.5 + j\,8.0 \text{ ohms} \quad (291)$$

$$Z_{CL} = 6.5 + j\,33 \text{ ohms} \quad (292)$$

The amplifier uses an air-line input circuit and an air-line output circuit similar to that shown in Fig. 316.

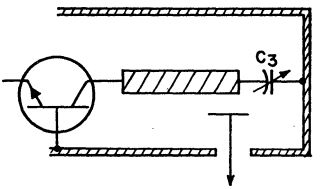


Figure 316. Capacitive-probe-coupled output cavity.

For the design of the output circuit, the optimum characteristic impedance Z_o of the output line is calculated from Eq. (290) to be 31 ohms. The collector load impedance Z_{CL} is normalized as follows:

$$Z_{CL}' = Z_{CL}/Z_o = 0.18 + j0.915 \quad (293)$$

Point Z_{CL}' is then located on the Smith Chart shown in Fig. 317 and rotated about the constant-VSWR circle toward the load. The 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

$$Z_L' = 1.39 - j\,3.3 \quad (294)$$

The load impedance Z_L is then equal to

$$Z_L = Z_o Z_L' = 36(1.39 - j\,3.3) = 50 - j\,119 \text{ ohms} \quad (295)$$

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. 317 to be equal to 0.33 λ (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. 318(a).

For the design of the input circuit, a characteristic impedance Z_o of 11 ohms is calculated from Eq. (290). The input impedance Z_{in} is normalized as follows:

$$Z_{in}' = (Z_{in}/Z_o) = 0.68 + 0.725 \quad (296)$$

Point Z_{in}' is then located on the Smith Chart shown in Fig. 319 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

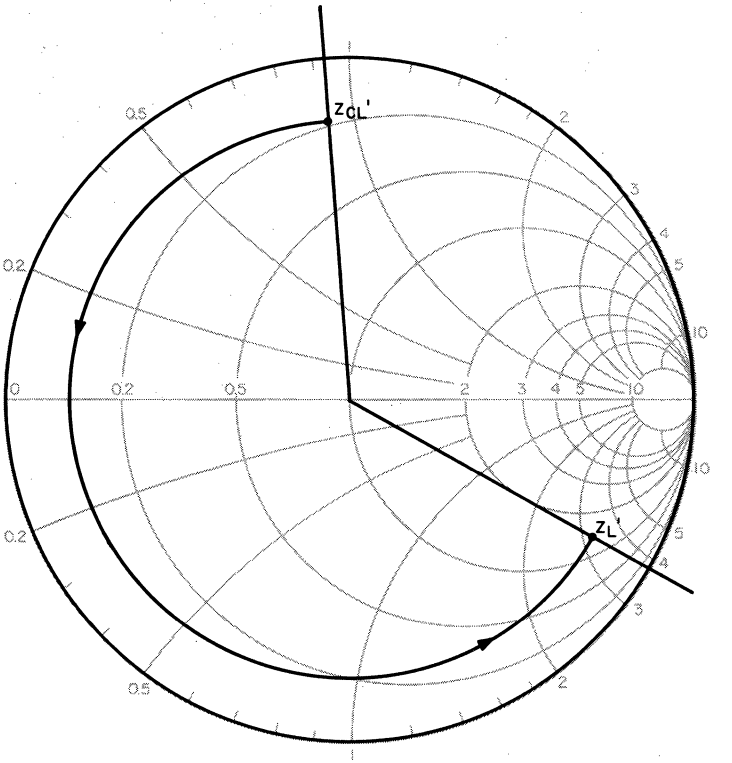
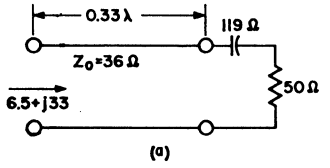
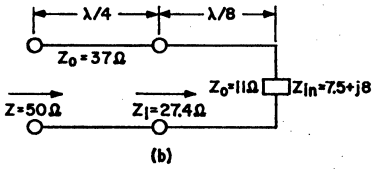


Figure 317. Smith-chart admittance plot for design of output transmission-line matching section.





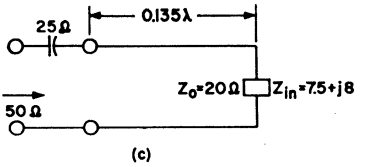


Figure 318. Input and output transmission lines with transmission-line matching sections added to provide required impedance transformation: (a) output line; (b) input line using series matching section; (c) input line using series matching section foreshortened by reactive elements.

and the normalized source impedance is, therefore, 4.55 ohms.) However, Fig. 319 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. 319, 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. (289) to be 37 ohms. Such an input circuit is shown in Fig. 318(b). Another possible input circuit uses added reactive elements, as shown in Fig. 318(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 ϵ is the dielectric constant of the insulator material of the stripline.

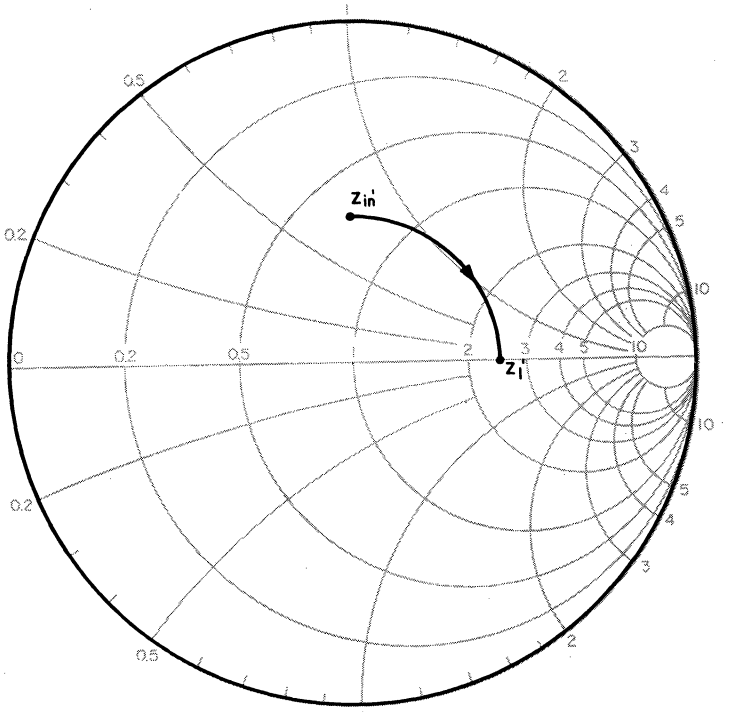


Figure 319. Smith-chart admittance plot used for design of input transmission-line matching sections.

MARINE RADIO

Marine radios are used primarily for ship-board communications involving the safety or navigation of the carrying vehicle; citizens-band radios used aboard boats are covered in a separate section. The term "radio", as used here, refers to a receiver-transmitter combination. Telegraphy, disaster and emergency radios, public ship communications, or navigation-only radios are not discussed. The type of radio considered is the type normally used on small private power boats, which features simple operation and requires no technical operator training. Radios of this type add to safety of the boat because of their ability to transmit on distress and calling frequencies (2.182 MHz, 156.3 MHz, and 156.8 MHz) constantly monitored by other boats and by the Coast Guard.

Ship-to-ship and ship-to-shore communications are permitted over certain frequencies assigned by the Federal Communications Commission and shared by all boats; all marine-radio operation is regulated by the FCC. These regulations concern the granting of licenses, assignment of frequencies, technical specifications of radio equipment, and allowable usage for each frequency. The exact regulations depend upon the geographic location of the boat, its use, and its size. A condensed listing of the FCC regulations for marine radio, as they affect transmitter design, is given in Table XXIX.

Marine radios are designed to operate in either the hf or the vhf band. The transmitter for the hf band (1.605 to 27.5 MHz) usually covers from 2 to 3 MHz in four channels with dc input of 25 watts (rf output of 12 to 15 watts). Higher-power transmitters usually cover the frequency range from 2 to 5 MHz and feature more channels. The rest of the hf band is covered by special-purpose transceivers not usually found on private boats. The vhf band is not as popular as the hf band in marine radio. Because of overcrowding on the hf band, however, the vhf band is becoming increasingly popular. VHF transmitters usually cover the entire vhf band, from 156 to 174 MHz, and feature frequency-modulated output.

Table XXIX—Marine-Radio Transmitter Regulations*

Frequency Band	Modulation	Bandwidth kHz	Deviation kHz	Frequency Accuracy %	Frequency Accuracy Hz	Power Output** Min.	Power Output** Max.	Required Frequency	Min. No. of Channels
1605 kHz-	AM	8.0	—	0.02	—	15	150	2182	3
2070 kHz	SSB	3.5	—	—	50		400		
2070 kHz-	AM	8.0	—	0.005	—				
2080 kHz	SSB	3.5	—	0.005	—				
2080 kHz-	AM	8.0	—	0.02	—				
3500 kHz	SSB	3.5	—	—	50				
4000 kHz-	AM	8.0	—	0.005	—		150		
27.5 MHz	SSB	3.5	—	—	50		1000		
156 MHz-					—		40	156.3	—
174 MHz	FM	40.0	15	0.002	—	—	100	156.8	—

* These are not complete specifications and are intended only as a general guide for transmitter design. See FCC Rules and Regulations, Part 83.

** Value depends upon location, ship weight, and ship type.

Transistorized equipment offers the advantages of small size, ruggedness, and improved reliability. In addition, the inverter power supply may be eliminated because the transmitter can operate directly from the 13.6 volts available from the boat battery.

Design Problems

Amplitude modulation in a transistorized marine-radio transmitter is an important design consideration. The transistor used must deliver the peak power required for upward modulation. The theoretical requirement is for a peak envelope power (PEP) of four times the cw power at 100-percent modulation. This requirement is seldom met, however, for the following reasons: the gain of a transistor decreases at high peak currents, the $V_{CE(sat)}$ increases at high peak current, and modulator losses rise during upward modulation. A minimum performance requirement is set by the FCC at 75 per cent upward and 100 per cent downward for an equivalent total modulation specification of $(100 + 75)/2 = 87.5$ per cent. This specification corresponds to a PEP requirement of 3.1 times the cw power.

Marine-radio owners want reliable, foolproof radio equipment, particularly because it is used in times of emergency. The owners/operators, however, generally lack technical knowledge and have inadequate concern for maintenance. Unfortunately, maintenance of the vulnerable antenna and transmission line has a direct effect on the performance of the radio. Optimum performance occurs only when the transmitter is properly loaded. A mismatch causes low power output and may cause spurious outputs or oscillation. Even more important, antenna and transmission-line faults place a stress on the transmitter output stage and could destroy the transistor. The best transistors for marine-radio transmitters, therefore, are those that can withstand such load faults without damage.

The hf-band radio must operate over a wide frequency range, as mentioned previously, usually 2 to 3 MHz. Because of this relatively low operating frequency and broad range, tuning capacitors capable of operating throughout the range must be large in value and, therefore in size. Unfortunately, capacitors (and variable inductors) of large value are not commercially available, and changes in operating frequency require the switching of entire tuned circuits or the changing of taps on coils with simultaneous switching to separate sets of capacitors. The low, wide frequency band also complicates antenna design. Because the antenna only approximates a constant impedance, the load presented to the transmitter may range from 5 to 50 ohms with a parallel capacitance in the ±200-picofarad range. The impedance varies with operating frequency, length of transmission line, and installation. Whatever the impedance, the tuned circuits in the transmitter output stages must match this value properly and transform it to the optimum load impedance for the output transistor. Because of this requirement, the tuned-circuit components must be capable of operating over very wide frequency ranges.

VHF FM Transmitter

New FCC regulations indicate that marine-radio communications will be gradually shifted from the 2-to-3-MHz AM band into the 156-to-174-MHz FM band. These new rulings have stimulated an interest in the design of all-solid-state vhf FM transmitters. Although the new FCC regulations are not entirely firm, they indicate that the maximum allowable power output for such marine-radio transmitters shall not exceed 25 watts with the provision that a maximum power output of 1 watt shall be imposed for harbor communications.

The new interest in vhf FM marine-radio communications systems has led to the design of multiplier-amplifier and power-amplifier sections of a vhf FM transmitter, such as that represented by the block diagram shown in Fig. 320. The circuits operate at any frequency in the 156-to-157.4-MHz marine band without necessity of retuning. Only a change in the crystal frequency of the exciter section is required to change the output frequency of the transmitter.

Multiplier-Amplifier—Fig. 321 shows a schematic diagram for the multiplier-amplifier section of the transmitter. This section, which consists of three frequency-multiplier stages and a power-amplifier stage, can deliver a power output of 1 watt over the frequency range of 156 to 157.4 MHz. With the output of the power-amplifier stage connected to the antenna (through

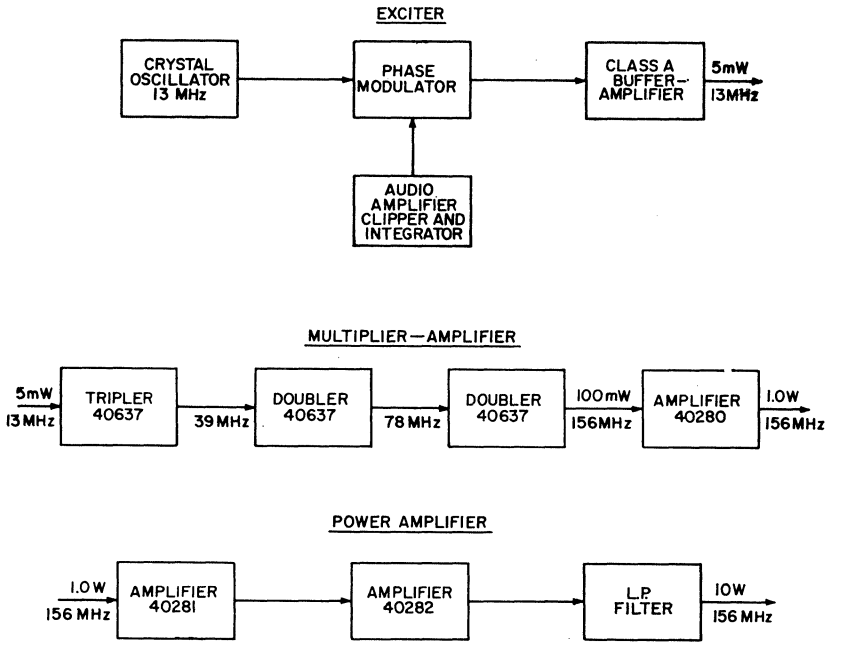
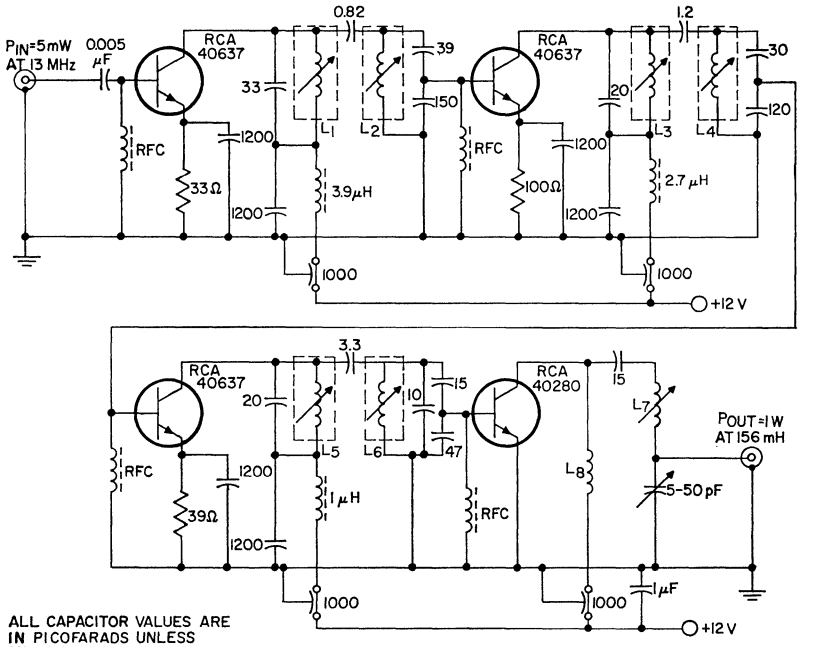


Figure 320. Block diagram of a typical vhf FM marine-band transmitter.



L_1, L_2 = 10½ turns No. 22 enamel wire, close-wound, 10/32" slug tuned coil forms 15/64" OD, shield can ½" x ½" x 1"; slug = carbonyl S.F. or equiv.

L_3, L_4 = 4½ turns No. 22 enamel wire, close-wound, 10/32" slug tuned coil forms 15/64" OD, shield can ½" x ½" x 1"; slug = carbonyl S.F. or equiv.

L_5, L_6 = 1½ turns No. 20 B.T., ¼" long, close-wound, 10/32" slug tuned coil forms 15/64" OD, shield can ½" x ½" x 1"; slug-carbonyl S.F. or equiv.

L_7 = 2½ turns No. 20 B.T., ¼" long, close-wound, 10/32" slug tuned coil forms 15/64" OD, shield can ½" x ½" x 1"; slug = carbonyl S.F. or equiv.

L_8 = 2 turns No. 20 B.T., 3/16" D, 3/16" long

RFC = 4 turns No. 30 enamel wire, ferrite bead, Ferroxcube No. 56-590-65/4B or equiv.

Figure 321. VHF marine-band multiplier-amplifier.

a low-pass filter), this section may be used to provide the maximum power output of 1 watt specified for harbor communications.

Each frequency-multiplier stage uses an RCA-40637 transistor, and the power-output stage uses an RCA-40280 transistor. The 5-milliwatt 13-MHz input signal applied to the first multiplier stage (tripler) from the exciter section of the transmitter is sufficient to saturate the multiplier stages; as a result, any amplitude modulation generated in the frequency-multiplier stages is eliminated. The second and third multiplier stages operate as frequency doublers.

Emitter-bias resistors are used in the multiplier stages to reduce current drain and dissipation in the multiplier transistors. The coupling between stages is accomplished by double-tuned coils to achieve maximum rejection of unwanted frequencies. The impedance transformation from the

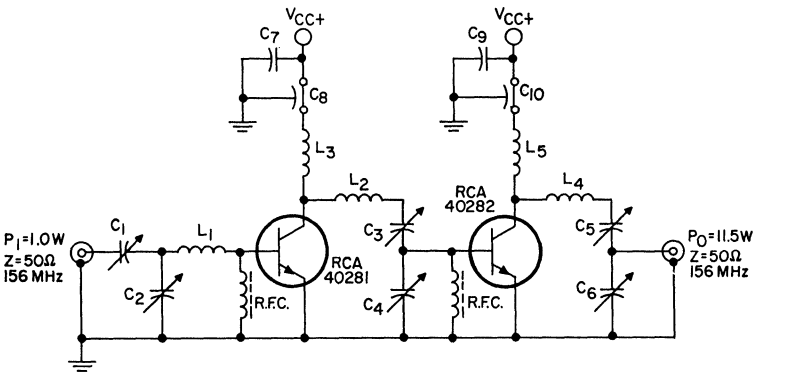
preceding collector to the following base is made possible through the use of a capacitive divider, which also forms a part of the resonant circuits. Heat sinks are not required for the first two multiplier transistors. The third multiplier and the amplifier transistors require fin radiators. The output from the second doubler is 100 milliwatts at 156 MHz. With 100 milliwatts into the power-amplifier stage, either the ouput to the antenna or the input to the power-amplifier section of the transmitter is 1 watt.

Power Amplifier—When a power output greater than 1 watt is required, the power-amplifier section of the transmitter provides two additional stages of power amplification. Fg. 322 shows a schematic diagram for the power amplifier. This circuit uses an RCA-40281 transistor in the driver stage and an RCA-40282 transistor in the output stage. With the 1-watt, 156-MHz output from the multiplier-amplifier applied to the driver stage, the power amplifier provides more than 10 watts of power output.

CITIZENS-BAND TRANSMITTERS

The Federal Communications Commission established the citizens band (CB) with 23 channels near 27 MHz and 48 channels near 465 MHz for business and personal use. Although many restrictions exist on transmitters for these bands and on the operating procedures, there are few restrictions and no technical requirements imposed upon the operators. As a result, the citizens band has become very popular.

Some of the equipment restrictions are listed in Table XXX. Most notable of the restrictions is the maximum dc input power to the final transmitter stage. Many restrictions on production qualification, on harmonic interference, and on operation are detailed by



C_1, C_2, C_3, C_4, C_5, C_6, = Arco 404 or equiv.
C_7, C_9 = 0.022 μF, 25 V
C_8, C_{10} = 1000 pF, ceramic standoff
L_1, L_5 = 1 turn No. 20 wire, ¼" ID, 3/32" long
L_2, L_3 = 1½ turns No. 20 wire, ¼" ID, 3/32" long
L_4 = 2 turns No. 18 wire, ¼" ID, 5/16" long
RFC = 4 turns No. 30 wire, ferrite bead Ferroxcube No. 56-590-65/4B or equiv.

Figure 322. 156-MHz marine-band power amplifier.

Table XXX—FCC Specifications for Citizens-Band Transmitters

Use	Classification	Frequency	Max DC Input Power	Formal License	Other Restrictions
Signaling	Class C *	5 channels 26.955-27.255 MHz	5 W	Yes [1]	No Voice
		27.255 MHz	30 W	Yes [1]	No Voice
Communications	Class D *	23 channels near 27 MHz	5 W	Yes [1]	AM or SSB Modulation, crystal-controlled
	Class A *	48 channels 462.55-466.45 MHz	60 W	Yes [1]	
Handy-Talky	Restricted ** Radiation Device	Same band as Class D, but not channels	0.1 W	No	AM or SSB modulation
Control Model Aircraft Communications	Class C *	5 channels 72.08-75.64 MHz	5 W		
	Class B *	All Class A frequencies plus 465.00 MHz	5 W	Yes [1]	

[1] Type Approval Required.
* Current FCC restrictions detailed in Vol. VI, Part 97 of FCC Regulations.
** Current FCC restrictions detailed in FCC Regulations.
(These units are not, in fact, CB sets, but share the same frequency band with CB.)

the FCC in the publications listed. Canada also established a citizens band with channels common to FCC specifications. Other Canadian specifications are different; licensing is not reciprocal.

There are three popular types of citizens-band transmitters. The most popular type is the 100-milliwatt hand-held transceiver, which usually incorporates small-signal transistors. The second type in popularity is the "full-5-watt" set designed for maximum operating range with from 5 to 23 channels. The dc voltage supply used in the full-5-watt set is 13.6 volts for both fixed and mobile locations. The third type of set is hand-held and operates on batteries, but has a transmitter power capability of 1 or 2 watts and provides more range than the 100-milliwatt sets.

Extra features are added to each type to interest hobbyists. For business use, for example, extra quality is built in. In every design, however, cost is very important because the citizens-band market is highly competitive.

Design Problems

Amplitude modulation in a transistorized citizens-band transmitter is an important design consideration. The peak upward-modulation power requirement must be met by the transistors. The theoretical requirement is for a PEP of 4 times the cw power. This requirement is seldom met because of both technical and economic reasons. Unfortunately, there is no standard for the quality of the modulation. It has been referred to as "talk power", although there is no definition for this term. The quality of modulation directly influences the range and performance of the transmitter. However, good modulation is difficult to achieve and is expensive. Because no specification exists, it is often compromised.

The citizens-band equipment appeals to non-technical persons who

generally assume the equipment to be foolproof. A particularly dangerous type of mistreatment is operation into a mismatched antenna/transmission-line combination. This type of operation may cause low power output and spurious outputs and oscillations, or may destroy transmitter transistors. A mismatched transmitter load condition must be evaluated carefully. Some transistor manufacturers recommend only certain transistor types for citizens-band AM service. The best transistors for citizens-band transmitters can withstand any antenna or transmission-line faults without damage.

The 23 channels available for citizens-band operation at 27 MHz have close frequency spacing; consequently, transmitter circuits that have a moderate loaded Q cannot operate at any channel without retuning. Harmonic-output specifications are strictly enforced by the FCC and require extra care in circuit design. Additional harmonic filtering is required at the output of the transmitter and usually takes the form of a pi-network, low-pass filter. Because fundamental-frequency oscillators are normally used, there is no output below the operating frequency and no filtering is needed below this frequency.

Design Example and Transistor Requirements

The most popular circuit for 5-watt (input) transmitters is one that employs three stages. An oscillator stage employing a third-overtone crystal rated at a dissipation of 2 milliwatts is used to establish the frequency. This stage drives a class C intermediate amplifier which, in turn, drives an output stage. The last two stages are collector-modulated. A block diagram of this arrangement is shown in Fig. 323.

This circuit is popular because it uses an economic combination of available transistors. The 12-volt supply is readily available in automobiles and easily generated in a base station and, when amplitude-modulated, rises to 24-volt peaks. The 12-volt supply is almost universally accepted in transistorized citizens-band set designs.

The modulator portion of the 5-watt transmitter can be operated in either the class A or the class B push-pull mode. The class A output type requires continuous dissipation of approximately 7 watts in either "receive" or "transmit" operating modes. The push-pull class B output type draws little power during "receive" and "transmit" operating modes, and almost no power

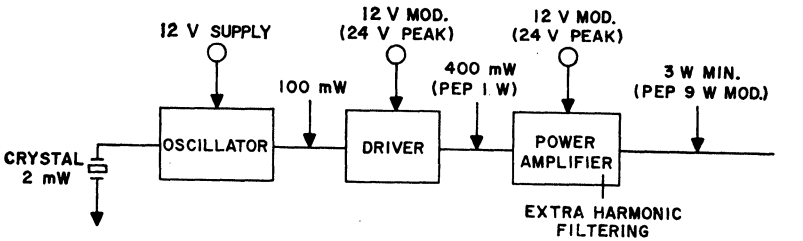


Figure 323. Block diagram of a 5-watt (dc input power to final amplifier) citizens-band transmitter.

during standby. This advantage is particularly important in battery-operated transceivers. The class B modulator requires smaller heat sinks, but larger transformers, than the class A type. Each circuit has been used in 5-watt CB transmitters.

The requirements for the transistors in the circuit of Fig. 323 were considered when the circuits were selected; the selection of the best circuit and the selection of the best transistor for that circuit are in no way independent. The best guide for selection of a circuit is a comparison of the circuits and transistors recommended by the transistor manufacturer. The cost of the transistors greatly affects the cost of the circuit; therefore, the transistors must be carefully selected.

The oscillator stage operates at the output frequency determined by the crystal (i.e., at the third harmonic of the crystal). This output frequency must be held constant to within ±0.005 per cent to meet FCC requirements. The crystals employed usually have a dissipation limit of 2 milliwatts. The dissipation limit is based on economic considerations and on the loss of frequency accuracy caused by increases in dissipation. The dissipation limit restricts the available 27-MHz power output of the oscillator stage. The power level is low, and the stage is usually designed for class A bias to assure starting without introduction of a dissipation problem in the transistor. A gain of 17 dB can be obtained from low-cost silicon transistors. This gain provides a minimum rf output of 100 milliwatts from the oscillator.

The output stage is designed for maximum cw power output within the FCC limit of 5 watts of dc input power. RF output of more than 3 watts is typical. The output stage must also withstand operation into mismatched loads. The output stage is required to deliver more than 3 watts of output when modulated; a theoretical peak envelope power (PEP) of four times the 3-watt cw output can be obtained at 100-per-cent modulation. The four-times capability is not usually achieved because of the increase in rf V_{sat} of transistors at high current, the decrease in transistor gain at high currents, and modulator losses (as explained in the general section on AM). Although no standards exist, a PEP of 9 watts, or three times the 3-watt cw power (i.e., 85-per-cent modulation) is considered good modulation. This power is obtained by modulating the collector supply voltage of the output transistor from a nominal value of 12 volts to a peak value of 24 volts. Therefore, the output transistor must deliver 9 watts at 27 MHz with a 24-volt supply. The requirements for the output transistor are listed in Table XXXI.

The intermediate amplifier (driver) must make up the difference between the rf drive required by the output stage and the rf power supplied by the oscillator. For the circuit discussed, the driver stage must increase the oscillator output of 100 milliwatts to 1 watt at the modulation peaks; this increase corresponds to a gain of 10 dB. Unmodulated, the required gain is 100 to 400 milliwatts, or 6 dB.

Again, the modulation peaks dominate transistor requirements. It is best to take advantage of the high supply voltage available from the modulated supply, rather than

Table XXXI—CB Transmitter Transistor Requirements

	Oscillator	Output Amplifier		Intermediate Amplifier	
		Modulated	Unmodulated	Modulated	Unmodulated
Operation	Crystal-Controlled Oscillator	Modulated RF Amplifier		Modulated RF Amplifier	
Designed for	Min. Crystal Diss. Max. P.O. (gain)	Maximum Efficiency (gain **)		Efficiency ** and Gain	
Operating Mode	Class A required for starting	Class C		Class C	
Supply Voltage	12 V *	24 V *	12 V *	24 V *	12 V *
Typical Performance	17 dB min. gain 2 mW crystal diss.	9 W out **	min. 3 W out max. 5 W in	1 W	0.4 W
Operating Current	35 mA Typical		415 mA max.		75 mA
Breakdown Voltage	24 V *	48 V *	24 V *	48 V *	24 V *
Mode	Open base	Reverse bias ***		Reverse bias ***	
Max. Dissipation	420 mW	See Text			
Typ. Dissipation	250 mW		2 W		0.5 W

* High line and transients must be added.
** Parameter is compromised to achieve other requirements.
*** Supplied by rf drive.

to use the straight 12-volt supply, to obtain the required 6-dB gain. Thus, modulation of the driver reduces the required performance of the driver transistor and, in addition, reduces circuit cost and improves modulation linearity. The distortion caused by power feedthrough in the output transistor on the downward-modulation peaks is severe when a constant 1-watt drive is used. Modulation of the drive to the final amplifier causes the output transistor to operate at nearly constant gain, rather than at constant drive, and reduces the feedthrough effect. The intermediate amplifier, therefore, is modulated and supplies a PEP of 1 watt with a 24-volt supply and a constant 100-milliwatt drive. The intermediate amplifier is usually operated with reverse bias to improve efficiency, primarily because the use of this bias mode reduces the dissipation requirements of the transistor and the power-output requirements of the modulator. The requirements for a transistor for the intermediate amplifier are also listed in Table XXXI.

The modulator circuit must supply audio power to the transmitter equal to one-half the dc input power. For the transmitter circuit described, the dc input power is the sum of the input power to the output stage (5 watts) and the input power to the intermediate amplifier (12 V $\times$ 75 mA = 90 mW). The audio power must be nearly 3 watts to modulate the transmitter fully. The audio input originates at a microphone. A large amount of audio amplification is necessary in the small-signal class A stage (microphone amplifier). The microphone amplifier is followed by an audio power amplifier which must deliver audio power of 3 watts. Either a class A or a class B push-pull audio power amplifier may be used.

The efficiency of a class A output stage approaches 50 per cent at full output power. If allowance is made for some change in bias current and some transformer losses, the nominal dissipation is approximately 7 watts. The tran-

sistor must be capable of withstanding this dissipation at high ambient temperatures (approaching 75°C). The results of calculations of required thermal resistance for the modulator output transistor are given in Table XXXII.

Table XXXII—Modulator Output-Transistor Requirements

	Class A Type	Class B Push-Pull Type
Designed for	Class A Amplifier 3 W Audio Output	Class B 3 W Audio Output
No. of Transistors	1	2
Audio Input Power	**	**
Supply Voltage	12 V *	12 V *
Operating Current:		
Peak	1.0 A	0.5 A
Average	0.5 A	0.2 A
Breakdown Voltage	24 V *	24 V *
Mode	Low	Reverse Bias ***
Dissipation: Continuous	7 W	0
Talk Time	7 W	1.5 W each
Max. Thermal Resistance (Transistor and Heat Sink for 75°C Ambient)	17°C/W	125°C/W

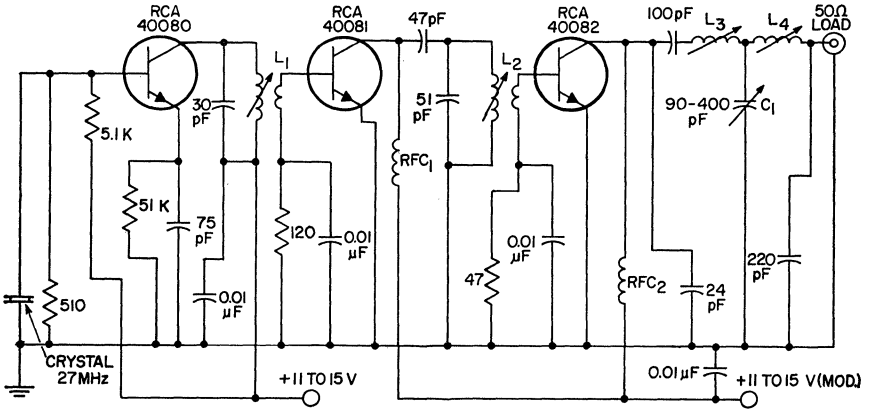
* High line voltage and transients must be added.
** This parameter is compromised to achieve other requirements.
*** Supplied by audio drive power.

The peak operating current for the output transistor can be calculated as 1 ampere. The gain of the circuit may be calculated from the characteristics of the transistor selected at a peak current of 1 ampere and the operating junction temperature.

A class B push-pull modulator requires two power transistors and two transformers, but reduces idling current. The idling current is only that required to reduce cross-over distortion, approximately 20 milliamperes in each transistor. This idling current results in a dissipation of 13.6 V × 0.020 A = 0.272 watt in each transistor. When a signal is present, the transistor dissipation rises. The peak dissipation occurs at 40 per cent of the rated output power and is equal to 20 per cent of the rated output power. For the example, the worst-case transistor dissipation is 0.272 W + (0.2 × 3) W, or less than 0.9 watt. It is possible, therefore, to use simple finned heat sinks and to wire the transistors on the printed-circuit board. Heat sinks are supplied with certain RCA transistors. The transistors must handle a peak current of 0.5 ampere; gain may be calculated on the basis of the highest junction temperature expected with the heat sinks used. This gain is usually higher than that for the class A circuit.

In selection of the output transformer for the modulator, many factors must be considered, not the least of which are cost and size. The transformer is required to match the output impedance of the modulator to the speaker used for public address and receiver service. It must have a winding to modulate the transmitter supply voltage with an impedance of V_{CC}/I_{DC}. The unbalanced direct current through this winding cannot be allowed to saturate the transformer. The winding must have a low dc resistance to reduce the voltage for the rf transistors and decrease the upward modulation. The windings driven by the modulator transistors must have a low dc resistance. An unbalanced direct current may also flow through this winding.

A complete circuit design for a CB transmitter and modulator is shown in Fig. 324. This circuit uses straight-forward design



C1 = Arco No. 429 or equiv.
L1 = 14:3 turns No. 22 wire, 1/4" CTC coil form with "green dot" core, 0.75-1.2 μH, Q = 100
L2 = 14:2¾ turns No. 22 wire, 1/4" CTC coil form with "green dott" core, 0.75-1.2 μH, Q = 100
L3 = 11 turns No. 22 wire, 1/4" CTC coil form with "green dot" core, 0.5-0.9 μH, Q = 120
L4 = 7 turns No. 22 wire, 1/4" CTC coil form with "green dot" core, 0.21-0.34 μH, Q = 140
RFC1, RFC2 = 15 μH, Miller No. 4624 or equiv.

Figure 324. 27-MHz citizens-band transmitter.

methods and has the following features:

(1) It can be economically produced.

(2) The number of tuning components is minimized. The rf transformers, L_1 and L_2, have only one adjustment, which varies both coupling and inductance. However, one adjustment cannot optimize both resonance and impedance matching. Therefore, the transformer turns ratio may have to be varied if any other characteristics are changed. Transformer L_2 may be optimized by changing the values of the 47- and 51-picofarad capacitors.

(3) The output stage is coupled to the transmission line and antenna through a double-pi network. This network reduces the harmonic output below the 50-dB reduction required by the FCC. One section of the double-pi network should always be peaked for maximum output; the other section is adjusted to a limit of 5 watts dc input to the final amplifier as required by the FCC.

MOBILE 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 practiced in 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 as high as 100 watts. Overseas, power-output requirements are much more moderate; the most common power-output levels are in the 10-watt range.

Cost Considerations

Today, all-solid-state mobile radios still cost more than vacuum-tube radios that have equivalent power-output ratings. However, all-solid-state mobile radios permit lower operating costs because they are relatively maintenance-free. In many mobile-radio applications (such as in ambulances, police patrol cars, and fire-fighting equipment) where failures cannot be evaluated in terms of money, solid-state radio equipment is the only logical choice. Even in less critical services (such as taxicab or truck dispatching), down-time on communications equipment is difficult to tolerate and is usually costly.

Transistor Requirements

The transistors in the rf power stages are the heart of every solid-state transmitter. In fact, the present frequency and power capabilities of these transistors constitute one of the major limitations to the advance of new mobile-transmitter design. For example, one of the basic requirements for a good rf power transistor is high f_T (appreciable current gains at high radio frequencies). Other requirements are high power dissipation and current-handling capability. However, these requirements are conflicting. Good current-handling capability requires a large emitter periphery, while high gains at high frequencies require that capacitances, and thus emitter area, be minimized. Some overlay transistors have emitter periphery-to-area ratios of 10 to 20 mils of emitter edge per square mil of emitter area, and 1 to 2 mils of emitter edge per square mil of base area; these transistors satisfy the requirements for rf power stages very well.

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 cur-

rent limiting are also somewhat more difficult.

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. In addition, all elements are usually isolated from the case to permit access to the emitter. The use of an emitter-biasing resistor is of great importance at these operating voltages because it contributes to good reliability under mismatched load conditions. Transistors intended for 12-volt operation have very high peak-current capability to realize substantial amounts of rf power at low voltage. The emitter electrode is usually connected to the case internally so that inductance in the emitter path and, consequently, gain degeneration are kept to a minimum. If reverse biasing is desired in grounded-emitter amplifiers, it can be obtained by use of a resistor in the base circuit.

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. In view of these factors, the well-proven vacuum-tube techniques become practically useless. Because the collector-to-base capacitance of the transistor is voltage-dependent, the neutralization of large-signal transistor amplifiers is impractical.

It has been observed that large-signal transistor amplifiers perform best when the output matching circuit presents a high impedance to the harmonic currents generated at the collector. The design of such networks was discussed previously in the section on **Network Design.**

Parallel Operation of Output Transistors

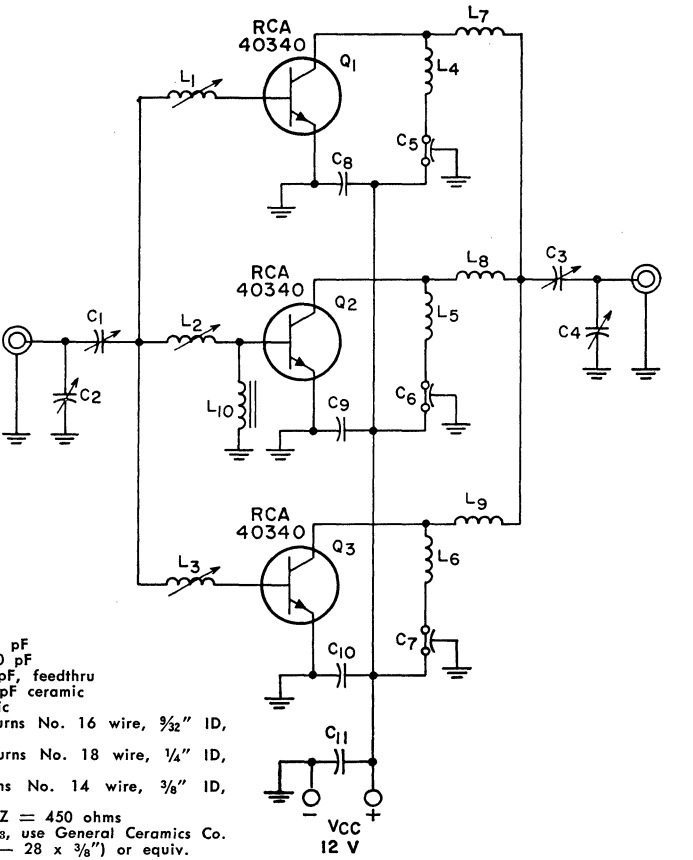
In applications that require more rf power output than one transistor can deliver, two or more transistors must be operated in parallel. Parallel operation in transistor circuits is somewhat more complex than in vacuum-tube circuits. The main problem is to assure that the load and collector currents are shared equally by all transistors in the parallel array. For this reason, the bases of the transistors should not be tied together directly, but rather should be connected through individual base-input coils which allow adjustment of drive to the individual transistors. The collectors of the paralleled transistors can be tied together directly, as is the usual practice in lower-power amplifiers operating from 24- to 28-volt supplies, in which collector load impedances are relatively high. For best parallel-operation efficiency in amplifiers operating from 12-volt supplies, where low collector load impedances are encountered, the

collectors of the transistors should be connected through individual collector coils. This technique permits parallel operation at higher impedance levels. Fig. 325 shows a circuit diagram of a 50-MHz power amplifier that delivers 60 watts with 12-volt collector supply voltage; this circuit demonstrates the techniques described.

Instabilities in VHF Transistor Amplifiers

In transistor vhf power amplifiers, the most common instabilities occur at frequencies far below operating frequency 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 circuits 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



C_1, C_3 = 65 — 340 pF
C_2, C_4 = 100 — 560 pF
C_5, C_6, C_7 = 1000 pF, feedthru
C_8, C_9, C_{10} = 1800 pF ceramic
C_{11} = 0.2 pF ceramic
L_1, L_2, L_3 = 2-1/2 turns No. 16 wire, 9/32" ID, 1/4" long
L_4, L_5, L_6 = 3-1/2 turns No. 18 wire, 1/4" ID, 3/16" long
L_7, L_8, L_9 = 4 turns No. 14 wire, 3/8" ID, 1/2" long
L_{10} = ferrite choke, Z = 450 ohms
Note: For coils L_1—L_3, use General Ceramics Co. Q_2 material (1/4" — 28 x 3/8") or equiv.

Figure 325. 50-MHz parallel-transistor power amplifier that delivers 50 watts of output power from a 12-volt dc supply.

following simple precautions, as indicated in Fig. 326:

(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.

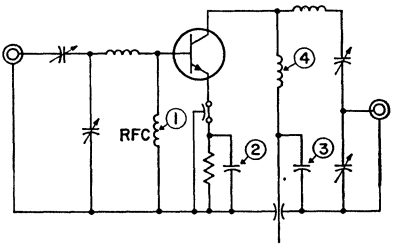


Figure 326. Circuit indicating areas where simple precautions prevent low-frequency oscillations.

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.

(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 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.

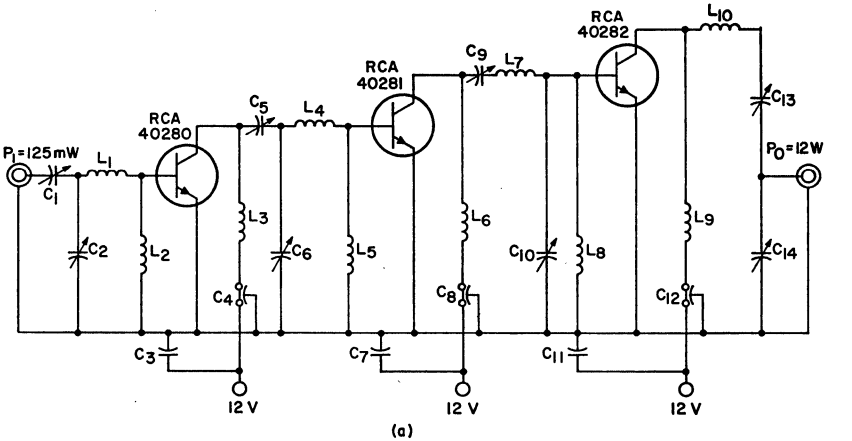
175-MHz Power Amplifier

Fig. 327 shows a 175-MHz amplifier-chain design that incorporates the techniques discussed for calculating matching networks and avoiding low-frequency oscillations. The amplifier operates from a 12-volt collector supply and delivers 35 watts at 175 MHz with an input of 125 milliwatts. The over-all efficiency of the amplifier chain is 60 per cent. The first part of the amplifier, Fig. 327(a), consists of three stages that provide output powers of 1, 4, and 12 watts from each consecutive stage. The second part of the amplifier, Fig. 327(b), has only one stage consisting of three transistors in parallel. This stage delivers 35 watts of power with a 12-watt rf input.

Reliability

Mobile-radio application places a severe requirement on transistors. The devices must withstand the no-load conditions created by objects near the transmitting antenna, or a break in the transmission line anywhere between zero and half wave length. Under these conditions, the transistors must handle not only increased dissipation, but also sudden energy surges that can destroy them in a matter of microseconds.

The development of transmitters immune to these failures is a result of a joint effort between semiconductor-device and mobile-radio manufacturers. To avoid excessive junction temperatures, the equipment manufacturer must select transistors of sufficiently low thermal resistance. If a transistor lacks enough dissipation capability, two should be used, even though one could deliver the required rf output power. To protect the devices under high-ambient-temperature operating conditions, it is necessary to use adequately



C_1 = 3.35 pF
C_2, C_6, C_{10}, C_{24} = 8.60 pF
C_3, C_7, C_{11} = 0.01 μF
C_4, C_8, C_{12} = 1500 pF
C_9, C_{10}, C_{13}, C_{14}, C_{23} = 7 — 100 pF
C_9 = 14 — 150 pF
C_{15} = 1.5 — 20 pF
C_{17}, C_{18}, C_{19} = 0.2 pF
C_{20}, C_{21}, C_{22} = 1500 pF

L_1 = 2 turns No. 16 wire, 3/16" ID, 1/4" long
L_2, L_5, L_8 = ferrite rf choke, Z = 450 ohms
L_3, L_6, L_{11} = rf choke, 1 μH
L_4, L_7 = 3 turns No. 16 wire, 3/16" ID, 1/4" long
L_9 = 1 1/2 turns No. 16 wire, 1/4" ID, 3/8" long
L_{10} = 2 turns No. 16 wire, 1/4" ID, 5/16" long
L_{12}, L_{13}, L_{14} = 3 1/2 turns No. 16 wire, 1/4" ID, 3'8" long (slug tuned)
L_{15}, L_{16}, L_{17} = 2 turns No. 18 wire, 1/8" ID, 1/8" long
L_{18}, L_{19}, L_{20} = 2 turns No. 18 wire, 1/4" ID, 1/4" long

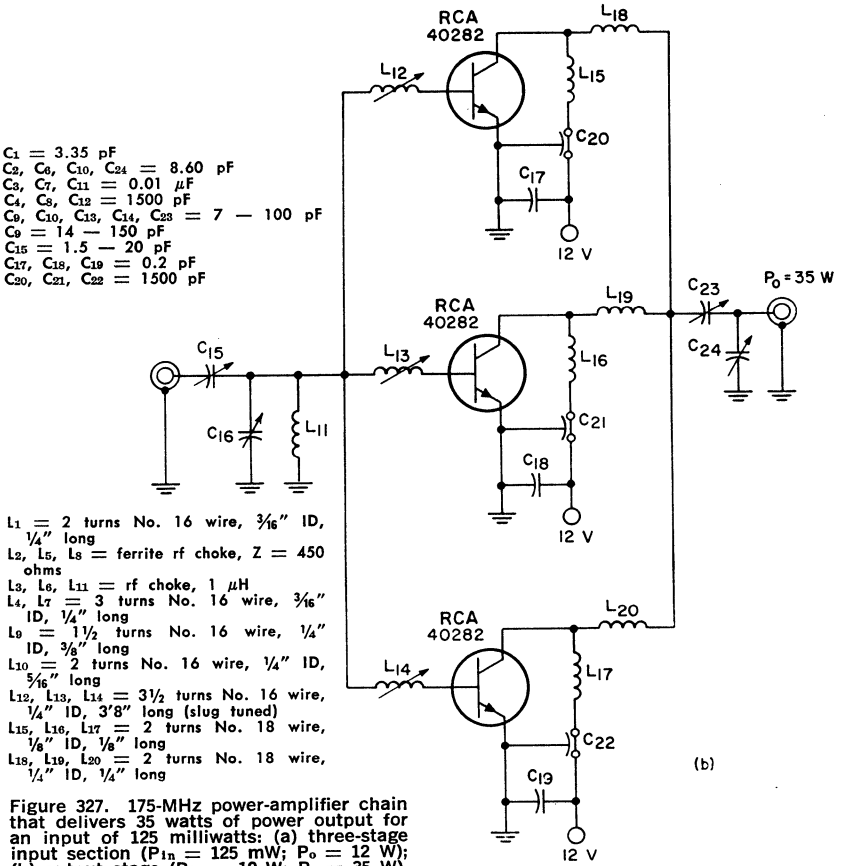


Figure 327. 175-MHz power-amplifier chain that delivers 35 watts of power output for an input of 125 milliwatts: (a) three-stage input section (P_{in} = 125 mW; P_o = 12 W); (b) output stage (P_{in} = 12 W; P_o = 35 W).

sized heat sinks, as well as current limiting to prevent excessive junction temperature rise under mismatched load conditions. As an added precaution, a thermostat can be mounted on the heat sink to reduce the transmitter power in the event that the temperature becomes excessive.

The protection of transistors from instantaneous failure is more difficult because the time response of current or voltage limiters is not fast enough. Biasing the emitters of transistors operating from 24- to 28-volt supplies helps to prevent this type of failure. Fig. 328 shows a circuit which is sufficiently fast in response time to protect the devices from instantaneous energy surges that result from mismatched load conditions. This circuit operates on the principle of reflected power. Under matched load conditions, there is no output from the VSWR detector. The control amplifier is saturated, and the gain-controlled rf amplifier operates at maximum gain. For this condition, maximum power output is obtained from the power amplifier. If a mismatch occurs, a negative voltage from the VSWR bridge brings the control amplifier out of saturation and thus reduces the gain in the gain-controlled rf amplifier. Gain is reduced because the base of the rf amplifier becomes more negative with respect to the emitter, and because the unsaturated control amplifier has a degenerative effect on the rf amplifier. Because the gain of the gain-controlled rf amplifier is reduced, the drive to the power amplifier is decreased to safe levels. Once the load mismatch is removed, the system returns instantaneously to normal operating conditions.

SINGLE-SIDEBAND TRANSMITTERS

The increase in communication traffic, especially in the hf and vhf ranges, necessitates more effective use of the frequency spectrum so that more channels can be assigned to a given spectrum. It has been shown that one of the more efficient methods of communication is through the use of single-sideband (SSB) techniques. In the

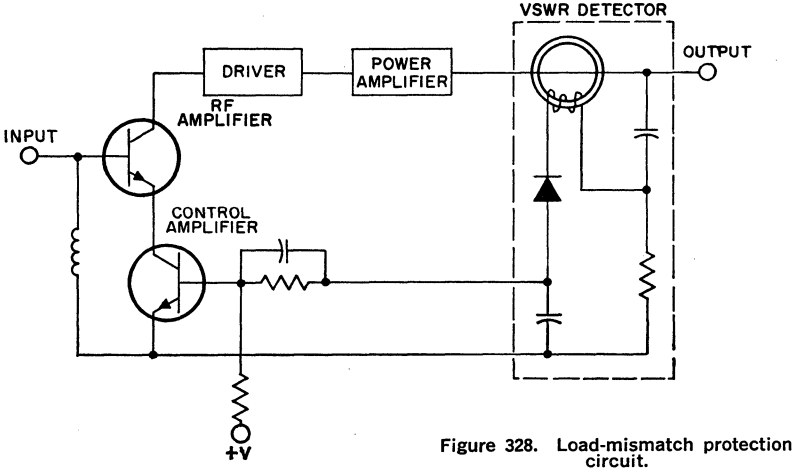


Figure 328. Load-mismatch protection circuit.

past, the power-amplifier stages of an SSB transmitter invariably employed tubes because of the lack of suitable high-frequency power transistors. Recent transistor developments, however, have made it feasible and practical to design and construct all-solid-state single-sideband equipment for both portable and vehicular applications.

Unlike most commercially available rf power transistors, which are normally designed primarily for class C operation, an SSB transistor is designed for linear applications and should have a flat beta curve for low distortion, and emitter ballast resistance for stability and degeneration. In high-power amplifiers, transistor junctions experience wide excursions in temperature and a means must be provided to sense the collector-junction temperature so that an external circuit can be used to provide bias compensation to prevent an excessive shift in operating point and to avoid catastrophic device failure as a result of thermal runaway.

Advantages of SSB Transmission

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:

$$e = E_o (1+m \cos 2\pi f_m t) \sin 2\pi f_c t$$

$$= E_o \sin 2\pi f_c t + m E_o \cos 2\pi f_m t \sin 2\pi f_c t \quad (297)$$

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:

$$e = E_o \sin 2\pi f_c t + (m E_o/2) \sin 2\pi (f_c+f_m) t + (m E_o/2) \sin 2\pi (f_c-f_m) t \quad (298)$$

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 unmodu-

lated 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. 329. The resultant signal envelope

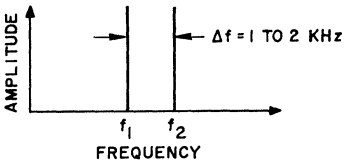


Figure 329. Frequency spectrum for a typical two-tone signal without distortion.

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 PEP rating of a single-sideband system is two times the average power rating.

Intermodulation Distortion

Nonlinearities in an amplifier generate intermodulation (IM) distortion. The important IM 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. 330 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 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

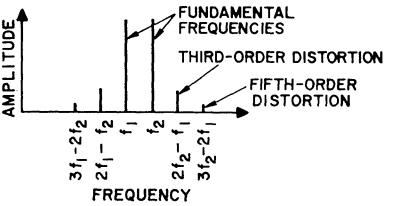


Figure 330. Frequency spectrum showing the frequency relationship of some distortion products to two test signals f_1 and f_2.

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 signal-to-distortion 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 IM distortion for an RCA developmental transistor at various output-power levels is shown in Fig. 331.

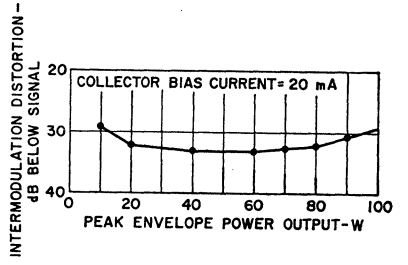


Figure 331. Typical intermodulation distortion in an RCA developmental transistor at various output power levels.

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 developmental 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 developmental transistor has a high emitter-periphery-to-collector-area and -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. 332 is a schematic representation of the transistor showing the separate emitters with resistors in series with

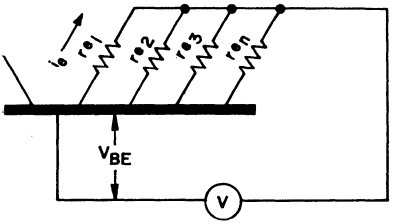


Figure 332. Schematic representation of the separate emitter sites in an "overlay" transistor with a resistor connected in series with each site.

each site. The voltage drop across each site is expressed by the following equation:

$$V = V_{BE} + i_e r_e \qquad (299)$$

Changes in V_{BE} have an exponential effect on the emitter current i_e, as follows:

$$\begin{aligned} i_e &= i_s \left[e^{(q/kT)\, V_{BE}} - 1\right] \\ &= i_s \left[e^{(q/kT)\, (V - i_e r_e)} - 1\right] \end{aligned} \qquad (300)$$

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:

$$(1/R_t) = (1/r_{e1}) + (1/r_{e2}) + (1/r_{e3}) + \ldots + (1/r_{en}) \qquad (301)$$

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 IM distortion to power output reveals that third-order distortion increases at both high and low output levels, as shown in Fig. 331. 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. 333.

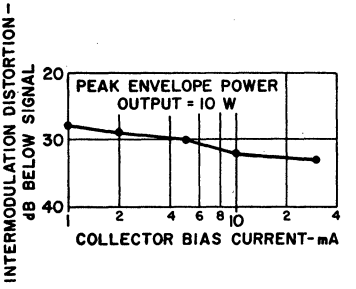


Figure 333. Intermodulation distortion as a function of collector bias current for the RCA SSB developmental transistor.

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 collector bias current of from 2 to 20 milliamperes for the RCA SSB developmental transistor is adequate to deliver 100 watts PEP. Fig. 334 shows a curve of power output as a function of supply voltage with distortion maintained at −30 dB.

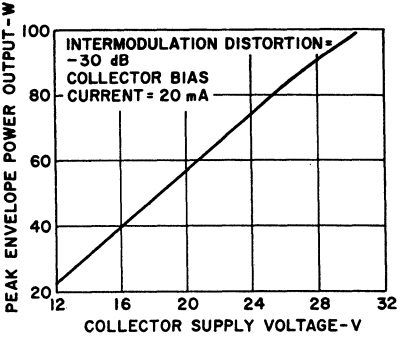


Figure 334. Peak envelope power as a function of collector supply voltage for the RCA SSB developmental transistor.

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 condi-

tion 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.

It is particularly difficult to maintain the bias current of a transistor high-power class AB amplifier at a constant level. As the drive increases, the dissipation increases and the junction temperature rises. If the conventional biasing technique is employed (an ac-bypassed emitter resistor and a constant voltage supply to the base), the varying emitter current that results from the varying drive changes the voltage drop across the emitter resistor and causes the bias to shift with drive. If a constant-current base-bias supply is used, the drive power is rectified and the bias point is changed.

The problem of maintaining a stable quiescent current is caused by a reduction in the V_{BE} of the transistor when the temperature rises. The base-to-emitter voltage decreases at a rate of approximately 2 millivolts per °C rise in temperature. Unless this condition is compensated for (i.e., bias voltage made to vary according to the V_{BE} decrease), the transistor is destroyed by the thermal effects.

Bias-point control for the SSB developmental transistor is accomplished by use of a diode placed next to the transistor pellet in the same package. The cathode of the diode is connected internally to the emitter lead. The anode of the diode is connected to a fourth terminal, as shown in Fig. 335. The diode is forward-biased between 1 to 5 milliamperes to provide a forward-voltage drop

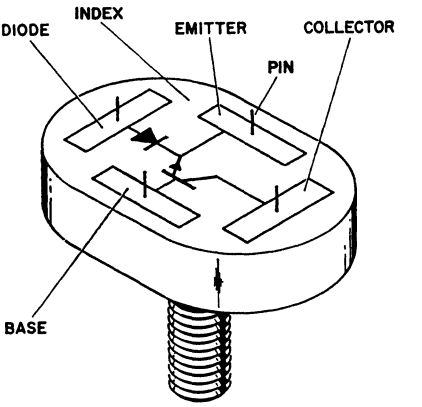


Figure 335. Package outline for the RCA SSB developmental transistor showing internal-package diode used for transistor bias-point control.

that is temperature-sensitive. At such a low current the diode operates in the low-conductance region where it does not provide the stiff voltage necessary for the transistor bias. In this case, the diode acts merely as a thermometer; an external amplifier must be used for current amplification. Compensation is achieved because the diode has approximately the same temperature coefficient for its forward-voltage drop as does the base-emitter junction of the transistor. Good tracking is obtained by mounting the diode and transistor pellets in the same case in very close proximity to minimize any thermal time lag. Temperature coefficient depends, to a large extent, upon the operating current. If the diode current can be adjusted so that it is approximately equal to the base current, good compensation can be achieved. The block diagram of a current amplifier that uses a low-conductance diode is shown in Fig. 336.

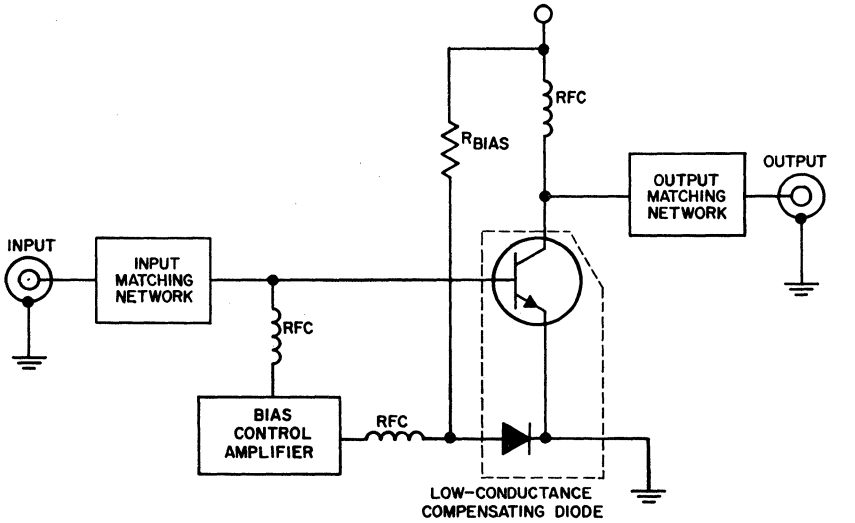


Figure 336. Block diagram of 30-MHz amplifier that uses a low-conductance diode for temperature compensation.

The schematic diagram of the current amplifier is shown in Fig. 337. The current amplifier employs a dc differential amplifier. The output voltage is the bias source for the power transistor. The use of a differential amplifier makes the entire amplifier relatively insensitive to temperature variations. Two additional stages are used for current amplification with negative feedback for stability.

Transistor collector-bias current can be adjusted by varying the 100-ohm potentiometer connected

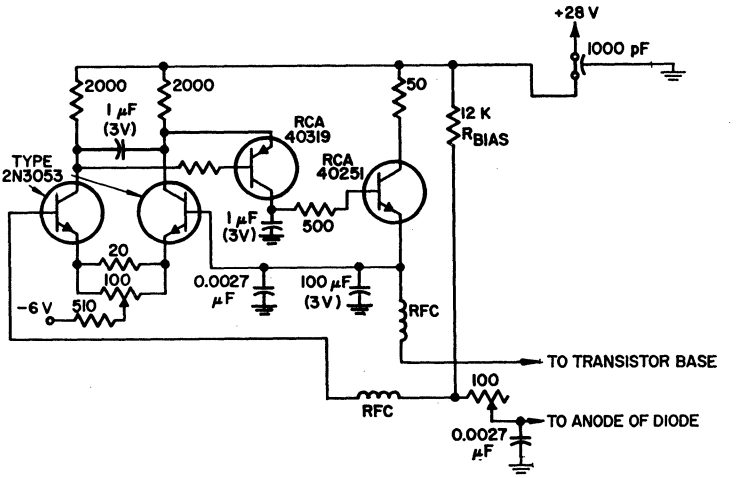


Figure 337. Linear 30-MHz amplifier with temperature-compensating circuit.

in series with the temperature-compensating diode. The diode current established by R_{bias} determines the degree of compensation. Overcompensation occurs when diode current is greater than base current. Fig. 338(a) shows collector quiescent current, initially biased at 10 milliamperes, as a function of case temperature. With compensation, the transistor is thermally stable even for case temperature as high as 150°C. Without compensation, however, the transistor tends toward thermal runaway at a case temperature of approximately 75°C.

Because both input and output are isolated through rf chokes, the external circuit provides compensation without degrading the rf performance of the power amplifier. Fig. 338(b) shows that no appreciable decrease in output power nor much increase in the third-order IM distortion occurs with increasing case temperatures up to $T_C = 120°C$. The slight decrease in output power and the increase in distortion, together with a decrease in collector efficiency, can be attributed to a rise in rf saturation voltage and a decrease in transistor beta at high temperature.

Despite the extra circuit needed to achieve temperature stabilization, the approach provides a practical solution for achievement of reliable operation of a class AB amplifier over a wide temperature range. The use of a small diode as a temperature-sensing element offers the following advantages:

(a) Diode and transistor pellets need not be matched for forward-voltage drop.
(b) Transistor quiescent current can be either overcompensated or undercompensated against changes in temperature by variation of the diode current.
(c) A diode idling current as low as 1 to 5 milliamperes can be used.
(d) Current of less than 50 milliamperes at 28 volts is needed to operate the external compensating circuit.

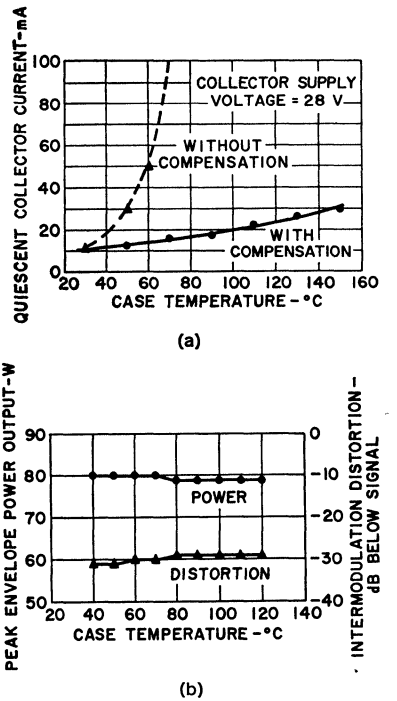


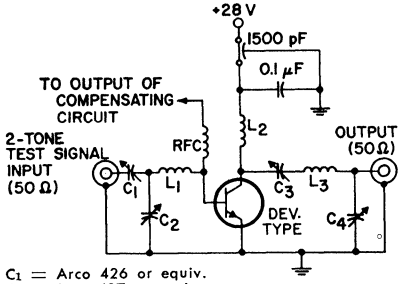
Figure 338. Performance characteristics for the 30-MHz amplifier: (a) collector current as a function of case temperature with and without temperature compensation; (b) output power and intermodulation distortion as a function of case temperature.

Typical Linear Amplifier

The common-emitter configuration should be used for the 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. 339 shows a schematic diagram of a narrow-band, high-power, 30-MHz amplifier. The amplifier provides an output power in excess of 100 watts PEP from a 28-volt power supply. The impedance of the base-to-emitter junction of the RCA SSB developmental 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



C_1 = Arco 426 or equiv.
C_2 = Arco 427 or equiv.
C_3 = 80-480 pF, Arco 469 or equiv.
C_4 = 140-680 pF, Arco 466 or equiv.
L_1 = 3 turns No. 14 wire, ¼" ID, ½" long
L_2 = 3 turns No. 10 wire, ½" ID, ⅜" long
L_3 = 3-½ turns No. 10 wire, ⅝" ID, ½" long

Figure 339. Narrow-band, high-power 30-MHz amplifier.

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. 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 and 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.

In low-power linear amplifiers, the use of temperature-compensating circuits is sometimes not necessary provided that the transistor output power is less than 50 per cent of its maximum cw power rating. The RCA-2N5070 transistor is useful in such application. 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.)

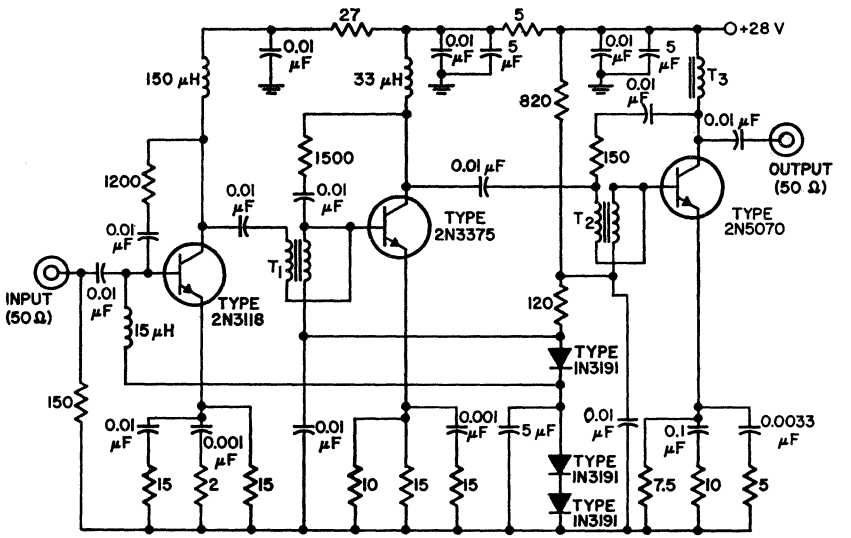
Fig. 340 shows a 2-to-30-MHz wideband linear amplifier that uses other types of RCA rf transistors. At 5 watts (PEP) output, IM distortion products are more than 40 dB below one tone of a two-tone signal. Power gain is greater than 40 dB.

The SSB Developmental Transistor

The RCA SSB developmental transistor is mounted in a plastic stud package. A diagram of the electrical connections for this transistor is shown in Fig. 336. Typical SSB performance of the transistor for −30 dB distortion is tabulated below for 30-MHz operation with the transistor biased at a quiescent collector current of 10 milliamperes:

P_o (PEP) at 28V = 90 W
Power Gain = 13 dB
Collector Efficiency = 50%

The effect of forward bias on the performance of the transistor is illustrated in Fig. 333. The figure shows an improvement in linearity when the transistor bias is increased from the threshold of collector conduction (class B) to a collector quiescent current of 25 milliamperes. Biasing the transistor further into class AB under high output power improves distortion performance



T_1, T_2 = 18 turns twisted pair. No. 28 enamel wire on Q_1 CF 102 form

T_3 = 50 turns No. 30 enamel wire on CF 102 Q_1 form

Figure 340. 2-to-30-MHz linear power amplifier.

very slightly. Even though the transistor has ballast resistance and a temperature-compensating diode, care must be used to prevent transistor damage by selecting the minimum quiescent collector current to obtain the desired low level of distortion at the maximum output-power level.

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 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.

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 XXXIII.

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

Table XXXIII—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.

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 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 broadband transmitter.

A narrow-band transmitter is mechanically tuned in each stage in 18 steps (or more if necessary), each covering 1 MHz. A particular problem arises in assuring that each component is varied in accordance with its own tuning-versus-frequency curve. This technique requires special cams and mechanical linkage. Remote operation of the transmitter requires expensive and complex

servo mechanisms. The inherently high Q of vacuum-tube circuits has forced the development and refinement of these techniques to produce good aircraft radios. Transistors may be used in place of tubes in narrow-band transmitter designs that use conventional circuits.

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.

Design Example

The design considerations for an aircraft-radio transmitter are illustrated by an example. The example is a popular transmitter which meets the requirements for privately owned single-engine aircraft. This transmitter is designed to provide a minimum output of 6 watts at a modulation of 85 per cent and to operate from the 12-volt airplane battery without converters. It is small enough to be mounted on the instrument panel.

Transmitters smaller than the one in the example are readily designed by removal of the high-power stage. Transmitters larger than the example have a much higher cost factor (rf-watt-output per dollar), and have smaller markets. However, they can be designed by use of the same techniques as those set forth in the example.

The design begins with the output stage and moves toward the crystal oscillators. The oscillators are not discussed in this section because they are low-power types and are not limited by the performance of available transistors. A block diagram of the transmitter is shown in Fig. 341.

The output stage of the transmitter is critical. It handles the most power, costs the most, and must be capable of accepting a load mismatch. The output power

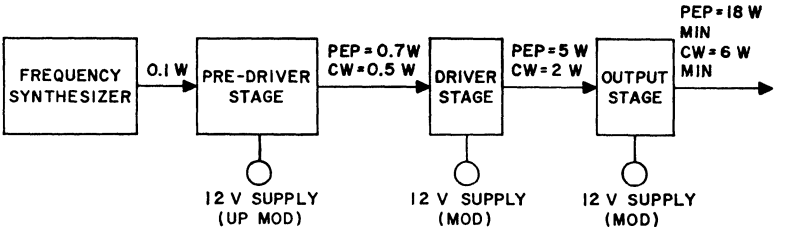


Figure 341. Block diagram of transmitter designed to deliver a minimum of 6 watts of rf power at 85-per-cent modulation for operation from a 12-volt airplane battery without converters.

for the transmitter is a minimum of 6 watts cw with a 12-volt supply (13.6 volts minus modulator and other voltage drops). The most important requirement is for upward modulation to a theoretical limit of 4 times the carrier power output, which corresponds to 24 watts PEP. FAA specifies 85 per cent minimum modulation, which is equivalent to 18 watts PEP. The limit of 24 watts is not achieved, however, because of the reduction in transistor gain at high currents, transistor rf $V_{CE}(sat)$, and modulator losses.

The output transistor must deliver the required PEP; it must also withstand a mismatched load (transmission lines and antenna). The severity of the load mismatch is increased with any increases in PEP, supply voltage, and/or ambient temperature. The severity of the mismatch is limited by any transmission-line and harmonic-filter losses (at high VSWR), by power-supply limiting, and by the energy stored in the modulation transformer. Therefore, a true test of the ability of the transistor to withstand load mismatch can be made only in a completed transmitter. The completed transmitter must also be tested for stability and interference output with a mismatched load. Before the transmitter is completed, the best guide to load-mismatch capability of transistors is the transistor manufacturer.

The requirements for the output transistor used in the transmitter are given in Table XXXIV. It should be noted that gain is

Table XXXIV—RF Transistor Requirements

	OUTPUT STAGE		DRIVER STAGE		PREDRIVER STAGE	
Operation	Modulated RF Amp.		Modulated RF Amp.		Modulated RF Amp.	
Drive	Antenna		Output Stage		Driver Stage	
Designed for	Max P.O. From Transistor At 24 V (Gain**) Load-Mismatch Capability		P.O. Gain Linearity		Gain Pre-Distortion	
Supply Voltage	12 V*	24 V*	12 V*	24 V*	12 V*	24 V*
Typical Performance	6 W Min. 8 W Typ.	18 W Min. PEP 24 W Typ.	2 W	7 W	0.5 W	1 W PEP
Operating Current	0.7 A		0.25 A		0.1 A	
Breakdown Voltage	24 V*	48 V*	24 V*	48 V*	24 V*	48 V*
Mode	Reverse Bias***		Reverse Bias***		Reverse Bias***	
Dissipation	See Text		1 W	1.5 W		

* High supply-voltage limit and transients must be added.
** This parameter is sacrificed to achieve other requirements.
*** Supplied by rf drive

of secondary importance as compared to power output and load-mismatch capability. The gain achieved in this example is only about 2.6 times the carrier power output, or 4.1 dB. As a result, drive power of 7 watts peak must be delivered by a modulated driver transistor.

The driver transistor must deliver a PEP of 7 watts with a peak supply voltage of 24 volts, and a cw power of 2 watts at 12 volts. These values yield a PEP-to-cw power ratio for the driver transistor of 7:2, which is greater than that for the output transistor (8:6) and requires the use of smaller peak currents relative to the maximum rating. Modulation of this stage has three important advantages: (1) it reduces the over-all modulation distortion; (2) it reduces the dissipation requirement of the driver transistor; and (3) it reduces rf drive to the output transistor on the downward modulation and thus reduces the feedthrough power which tends to distort downward-modulation peaks.

The predriver stage is a class C power stage which must generate cw power of 0.5 watt and a PEP of 1 watt. This stage is also modulated and has a constant drive power. Modulation of this stage reduces the dissipation requirement, but can cause distortion problems. Modulation of more than two stages distorts downward modulation because transistors exhibit a threshold effect. At low supply voltages, the gain and power output decrease very rapidly. This decrease causes a "notching" in the audio waveform and zero rf power output if more than two stages are modulated. If only two stages are modulated, the feedthrough power prevents complete cutoff of the transistors.

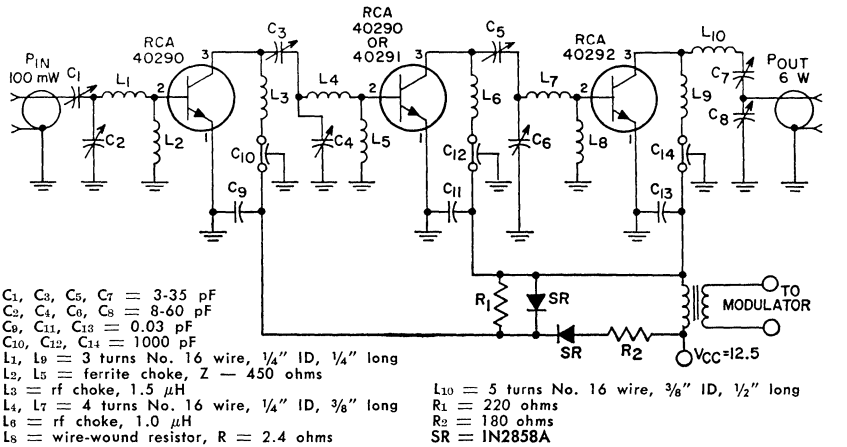
Both improved upward modulation and undegraded downward modulation may be achieved simultaneously if only up-modulation is used in the pre-driver stages. This purposeful distortion partially compensates for the nonlinear (gain-current) characteristic of the transistor, and is swamped out and undetectable if two fully modulated stages follow the distorted stages.

The gain of the pre-driver stage is usually high because the operating currents are small. A constant drive power of 100 milliwatts is needed. This power level is available from the frequency synthesizer (channel selector).

The actual circuit design for a narrow-band transmitter follows conventional techniques, as described above. The completed design for the example is shown in Fig. 342. The transmitter requires a good-size modulator capable of delivering an audio power equal to half the total dc input power to the transmitter. For a 6-watt transmitter, the dc power input is approximately 12 watts, and a full 6 watts of audio power is required.

Four combinations of modulators and power supplies are listed in Table XXXV, together with the advantages and disadvantages of each. For aircraft radios, each combination should be considered along with the other features planned for the radio (e.g., multiple voltages, remote operation, size, and weight). Only two modulator designs are included in the table.

The series modulator uses one transistor to reduce an input sup-



C_1, C_3, C_5, C_7 = 3-35 pF
C_2, C_4, C_6, C_8 = 8-60 pF
C_9, C_{11}, C_{13} = 0.03 pF
C_{10}, C_{12}, C_{14} = 1000 pF
L_1, L_9 = 3 turns No. 16 wire, 1/4" ID, 1/4" long
L_2, L_5 = ferrite choke, Z — 450 ohms
L_3 = rf choke, 1.5 μH
L_4, L_7 = 4 turns No. 16 wire, 1/4" ID, 3/8" long
L_6 = rf choke, 1.0 μH
L_8 = wire-wound resistor, R = 2.4 ohms
L_{10} = 5 turns No. 16 wire, 3/8" ID, 1/2" long
R_1 = 220 ohms
R_2 = 180 ohms
SR = IN2858A

Figure 342. 35-MHz, 6-watt, narrow-band amplitude-modulated transmitter.

ply voltage of 28 volts to 12 volts, and to modulate this voltage. The input voltage must exceed the peak modulated supply voltage for the transmitter. The dissipation of the modulator transistor is maximum when there is no modulation. The dissipation then is equal to the transmitter operating current times the difference between the input voltage and the unmodulated output voltage. This dissipation must exceed the unmodulated dc power input to the transmitter. For the example, the peak dissipation would be (1 ampere) (28-12) = 16 watts. The worst-case dissipation occurs with maximum input voltage.

The size and weight of the heat sink for the modulator transistor is balanced by the elimination of the modulator transformer. Elimination of this transformer allows the use of simple feedback to reduce distortion.

The elimination of supply-voltage transients from the 13-volt

Table XXXV—Modulator and Power-Supply Combinations for Aircraft Radio

Type of Power Supply	Inverter	None	Active Limiter	Passive Transient Limiter
Min. Input Voltage	12 V	25 V	13.6 V	13.6 V
Output Voltage	30 V + Transients	25 V + Transients	13 V	13.5 V
Type of Modulator	Series-Type Class A		Transformer-Type Push-Pull Class B	
Main Advantage	Excellent Transient Limit	Excellent Transient Limit, Low Cost	Low Loss	Simplicity
Other Advantages	Light-Weight Transformers (High Freq.) Convertible To Any Supply Voltage Low Distortion	Light Weight No Transformers Low Distortion	No Inverter No Inverter Noise or Hash Small Heat Sinks	No Inverter No Inverter Noise or Hash Small Heat Sinks
Disadvantages	Expensive, Large Heat Sinks	Large Heat Sinks	Large Transformers (Audio)	Large Transformers (Audio)

internal supply results directly from the feedback used to reduce modulator distortion. It is necessary only that the series transistor have a breakdown voltage greater than any expected transient, and a pulsed-power capability to absorb the transient.

The transformer type of modulator usually employs a push-pull class B circuit. This type of operation requires two transistors, but produces an efficient modulator that has low transistor dissipation. The total transistor dissipation is a maximum of 40 per cent of the designed output power, or 2.4 watts for the example. This dissipation occurs when the modulator output is 40 per cent of the maximum.

The modulation transformer must be rather large to meet all its requirements. The resistance of the windings must be low, or the modulated performance of the transmitter is degraded. The output winding matches an impedance of $V_{CC}/I_{DC} = 12/1 = 12$ ohms for this example. An unbalanced dc current equal to I_{DC} (the transmitter operating current) flows through this winding. This current may rise if there is a mismatch in the output of the transmitter; there should be a 25-per cent power-output margin designed into the transformer.

Fig. 343 shows a block diagram and Fig. 344 shows the circuit schematic of another aircraft transmitter. This transmitter operates over the frequency range from 118 to 136 MHz and requires no retuning to cover the entire band. The push-pull output stage in this broadband transmitter delivers 40 watts of peak envelope power. The first stage of the four-stage chain is operated as a class A amplifier to produce a maximum amount of amplification of the relatively low output from the frequency synthesizer and to provide isolation between synthesizer and modulated stages. The remaining three stages of the transmitter operate class C, and all are amplitude-modulated. The nominal supply voltage for the transmitter is 12.5 volts.

COMMUNITY-ANTENNA TELEVISION

Community-antenna television systems (CATV) 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 off-air signals, and the 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.

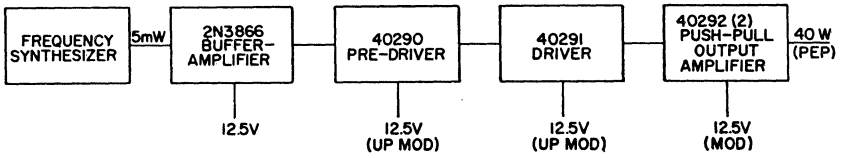
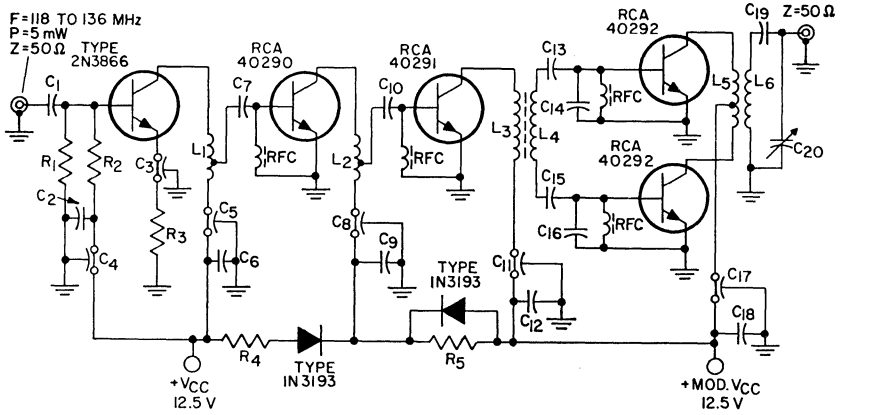


Figure 343. Block diagram of broadband 118-to-136-MHz Aircraft Transmitter.



C_1 = 330 pF, Arco S.M. or equiv.
C_2 = 0.005 μF, ceramic
C_3, C_4, C_5, C_8, C_{11}, C_{17} = 1000 pF, feedthru
C_6, C_9, C_{12}, C_{18} = 0.05 μF, ceramic
C_7 = 50 pF, 5%, Arco S.M., or equiv.
C_{10}, C_{13}, C_{15} = 82 pF, 5%, Arco S.M., or equiv.
C_{14}, C_{16}, C_{19} = 150 pF, 5%, Arco S.M., or equiv.
C_{20} = Variable capacitor, 8-to-60 pF, Arco 404 or equiv.
L_1 = 7 turns No. 22 wire, 13/64" ID, 9/16" long, tapped at 1.5 turns
L_2 = 5½ turns No. 22 wire, 13/64" ID, close wound, tapped at 2 turns
L_3 = 6 turns No. 22 wire, 13/64" ID, interwind with L_4 on IRN-9 core material
L_4 = 4 turns No. 22 wire, 13/64" ID, interwind with L_3 on IRN-9 core material
L_5 = 5 turns No. 22 wire, 13/64" ID, center-tapped, interwind with L_6
L_6 = 5 turns No. 22 wire, 13/64" ID, interwind with L_5
RFC = 1 turn No. 28 wire, ferrite bead, Ferroxcube No. 56-590-65/4B or equiv.

Figure 344. 118-to-136-MHz broadband amplifier for aircraft radio.

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.

System Operation

Fig. 345 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 of the order of 2500 feet. At various points along the trunk line, 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

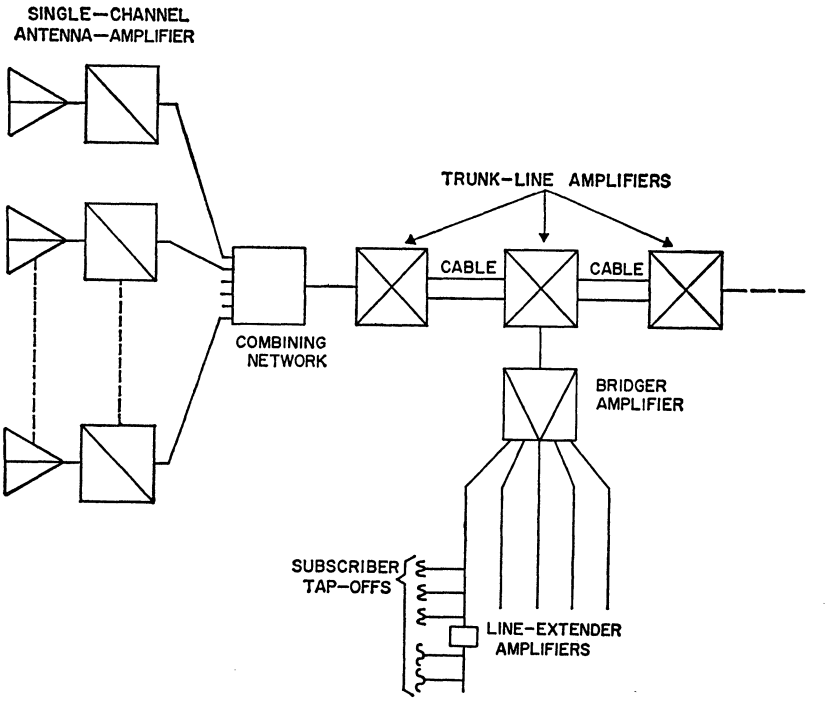


Figure 345. Simplified community-antenna Television (CATV) system.

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 the entire television band, from 54 to 216 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. 346. Typical loss is 0.4 dB per 100 feet at channel 2, and 1 dB per 100 feet at channel 13. The spacing presently used between trunk amplifiers is 25-dB cable length at channel 13. The gain of the trunk

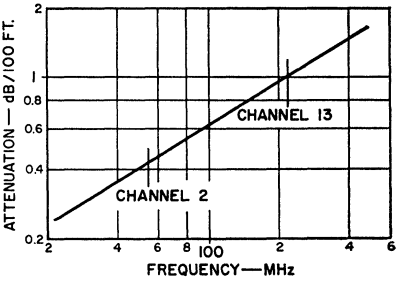


Figure 346. Attenuation characteristics of a coaxial cable as a function of frequency.

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 "slop" or "tilt." The amplifier gain must be higher at the high end of the band than at the low end. The gain of a bridger amplifier is somewhat lower than that of the wide-band amplifier (about 18 dB). These amplifiers usually contain 2 to 4 stages of individual amplification, depending on the gain requirement.

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 is also referred to as "windshield-wiper" effect. Cross-modulation results when several channels are passing through a wide-band 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.

"Snowy" pictures can be avoided if the signal at any point in a system is made strong 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.

Some current, typical requirements for trunk-line amplifiers to be used in a CATV cascaded system are as follows:

Bandwidth
= 54 to 216 MHz
Input Operating Level
= 10 dBmV
Output Operating Level
= 35 dBmV
Maximum Output Capability
= 50 to 55 dBmV
Gain
= 25 dB
Response
± 1 dB over the band
Noise Figure
= 12 dB at channel 13
= 8 dB at channel 2
Tilt
= 12 dB over the frequency range
Voltage Available
= 30 V (+ 15V and −15V)

These performance specifications must be met in outdoor temperatures ranging from −40 to 140°F.

Transistor Wide-band Amplifier

The gain-bandwidth product of a transistor connected in a common-emitter configuration is equal to f_T. Thus, the bandwidths of an

uncompensated common-emitter amplifier stage may be expressed as follows:

$$BW = f_T/h_{fe} = f_T\,(1-\alpha_o)/\alpha_o \quad (302)$$

where h_{fe} is the low-frequency common-emitter current gain of the transistor and α_o is the low-frequency common-base current gain. Eq. (302) 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. (302). Fig. 347 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. 348(a), in which parasitics are not included. One difficulty with the circuit of Fig. 348(a) is the capacitance C_c, which prevents

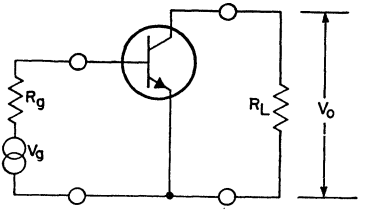


Figure 347. Common-emitter transistor amplifier.

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. 348(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

$$C_{eq} \simeq (1/W_t\, r_e')\,(1+W_t\, C_C\, R_L) \quad (303)$$

With the aid of the simplified circuit of Fig. 348(b), the following equations for the gain and band-

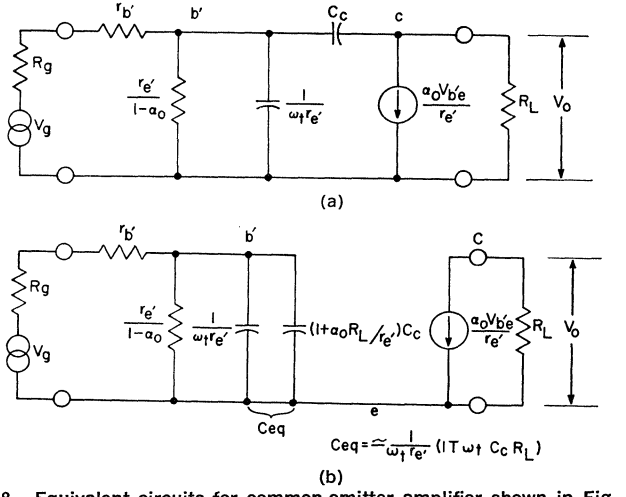


Figure 348. Equivalent circuits for common-emitter amplifier shown in Fig. 347: (a) without parasitic elements; (b) simplified equivalent circuit.

width of the amplifier are derived:

$$\text{Gain} = \frac{V_o}{V_g} = \left[\frac{\alpha_o R_L}{(R_g+r_b')(1-\alpha_o)+r_e'}\right]\left[\frac{1}{1+\frac{p(R_g+r_b')r_e' C_{eq}}{(1-\alpha_o)(R_g+r_b')+r_e'}}\right] \quad (304)$$

$$BW = \frac{(1-\alpha_o)(R_g+r_b')+r_e'}{(R_g+r_b')r_e' C_{eq}} = \frac{W_t}{(1+W_t C_c R_L)}\left[1-\alpha_o+\frac{r_e'}{(R_g+r_b')}\right] \quad (305)$$

Eq. (305) shows that the bandwidth is decreased by an increase 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 ω_T and by a decrease in r_b' and C_c. Thus, a transistor suitable for wideband operation should have high f_T (or ω_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. (304) 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 determnied from Eqs. (304) and (305), is not inversely proportional to the bandwidth. One of the important criteria of a wideband 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. 349. 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.

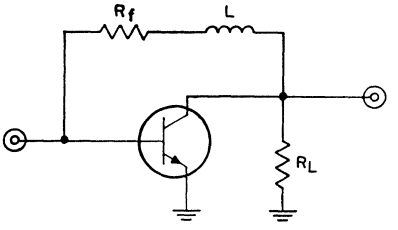


Figure 349. Common-emitter amplifier that uses shunt RL feedback to increase circuit bandwidth.

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:

$$L = (R_f + r_b' + R_L)\, R_f/2\pi\, f_t\, R_L. \quad (306)$$

Emitter Degeneration—Another common method of trading gain for bandwidth is to use emitter regeneration or emitter peaking, as shown in Fig. 350. Simple resistance-capacitance feedback is

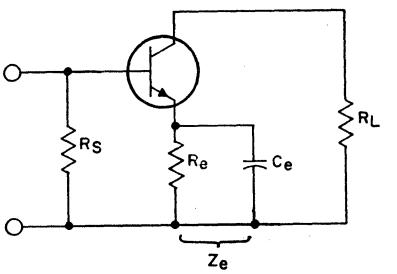


Figure 350. Transistor amplifier that uses emitter degeneration to increase circuit bandwidth.

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 of this amplifier, when $Z_e = R_e$, are as follows:

$$\text{Gain} = \frac{V_o}{V_g} = \left[\frac{\alpha_o R_L}{\alpha_o R_e + r_e' + (1-\alpha_o)(R_s + r_b' + R_e)}\right] \left[\frac{1}{1 + \frac{pC_{eq}\, r_e' (R_s + r_b' + R_e)}{\alpha_o R_e + r_e' + (1-\alpha_o)(R_s + r_b' + R_e)}}\right] \quad (307)$$

$$BW = \frac{\alpha_o R_e + r_e' + (1-\alpha_o)(R_s + r_b' + R_e)}{r_e' (R_s + r_b' + R_e)\, C_{eq}} \quad (308)$$

Comparison of Eqs. (305) and (308) 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 $1/(15 f_T R_e)$.

Two-Stage Wide-band Amplifier

Degenerative feedback around two stages offers large bandwidth increase with minimum loss of gain-bandwidth product. Two such feedback methods are shown in Fig. 351. The circuit shown in Fig. 351(a) uses a combination of voltage and current feedback. The resistance R_{e2} provides emitter degeneration in the second stage, while the resistance R_{f2} provides a form of current feedback to the first stage. The current gain is approximately equal to the ratio R_{f2}/R_{e2}. Because the feedback is obtained from a low-impedance point and is returned to a point of somewhat higher impedance around the current amplification of both transistors, substantial increases in bandwidth (up to 0.5 f_T) can be obtained for a constant gain-bandwidth product.

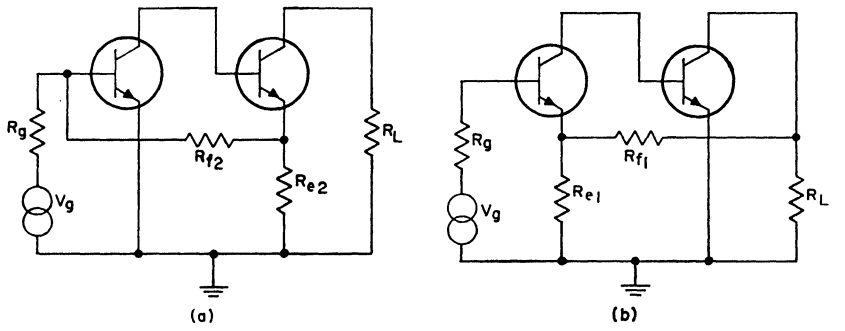


Figure 351. Two-stage wide-band rf amplifiers that use a combination of voltage and current feedback to increase circuit bandwidth: (a) Feedback is coupled from emitter of output transistor (low-impedance point) to base of input transistor (higher impedance point); (b) feedback is coupled from collector of output transistor (high-impedance point) to emitter of input transistor (low-impedance point).

The circuit shown in Fig. 351(b) is satisfactory when moderate bandwidth (up to 0.1 f_T) is desired. This circuit differs from that of Fig. 351(a) in that the feedback is obtained from a high-impedance point and is applied only around the current amplification of the second stage.

Series and Shunt Peaking

Series peaking or shunt peaking schemes can also be used to extend the bandwidth of an amplifier. Fig. 352(a) shows a series interstage inductance that connects the base of the second stage and the collector of the preceding transistor. This inductance increases the bandwidth by forming a resonant impedance-matching network with transistor input and output capacitances at frequencies near the amplifier cutoff frequency. The value of L can be determined from the following equation:

$$L = \frac{r_e' (r_b' + R_L)^2 C_e}{4 [r_e' + (1-\alpha_o)(R_L + r_b')]} \quad (309)$$

The shunt peaking method shown in Fig. 352(b) employs an inductance in series with the input shunting resistance. The inductance is adjusted so that the load is reduced in the frequency range in which the transistor gain begins to decrease. The required inductance L is equal to $R_L/15\ f_T$.

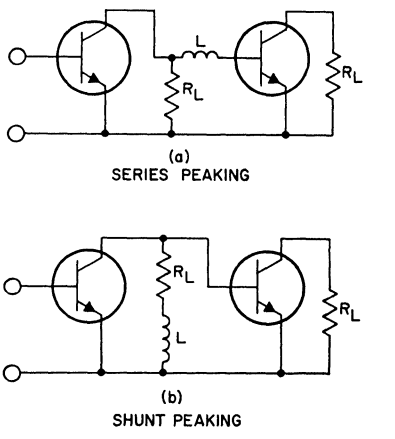


Figure 352. Circuits showing use of peaking coils to increase bandwidth: (a) series peaking; (b) shunt peaking.

Wideband Transformer

The choice of dc bias for a wideband transistor amplifier to be used in CATV depends on such factors as power-output requirements, power dissipation, cross-modulation, noise figure, gain, and bandwidth. In general, the transistor should be biased at a point that provides the maximum gain-bandwidth product f_T. In addition, the stage that requires high-level output should be biased at a point at which good linearity or minimum distortion can be obtained. The first stage should be biased at a point for which a minimum noise figure can be achieved. The operating conditions at a given voltage of 15 to 30 volts (available in a CATV system) are determined by the collector load resistance R_L in the circuits shown in Figs. 348 through 352. Because the input-source and output-load resistances in a CATV wideband amplifier are 75 ohms, wideband transformers must be used to provide the proper operating conditions of the transistor.

The wideband transformer shown in Fig. 353 consists of a ferrite toroid around which a twisted pair of wires are wound. This transformer is of the transmission-line type and has excellent bandwidth ratios. The transmission lines take the form of twisted pairs. In this transformer, the coils are so arranged 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).

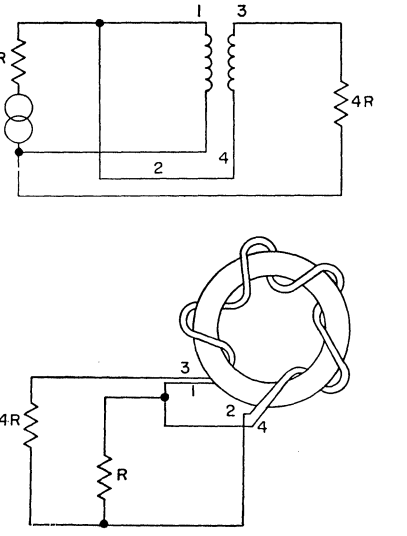


Figure 353. Wideband transformer. This transformer may be used to provide a 4:1 impedance ratio, as indicated in top diagram. The transformer is basically a twisted-pair transmission line wound about a ferrite torroid, as shown in lower diagram.

The 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 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 important that coupling be high at all frequencies, or the transformer action fails.

The transformer shown in Fig. 353 has an impedance ratio of 4:1. The high-frequency response of this transformer may be calculated by use of the following equation:

$$\frac{\text{Power Available}}{\text{Power Output}} = \frac{(1+3\cos \mathrm{B}\,l)^2 + 4\sin^2 \mathrm{B}\,l}{4\,(1+\cos \mathrm{B}\,l)^2} \quad (310)$$

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 Noise-Figure Considerations

The noise figure, defined as the ratio of the input signal-to-noise ratio to the output signal-to-noise ratio, of a single-stage transistor amplifier is a function of frequency and transistor parameters, as shown by the following equation:

$$\mathrm{NF} = 1 + \frac{r_b'}{r_g} + \frac{r_e'}{2\,R_g} + \frac{(r_b'+r_e'+R_g)^2}{2\,\alpha_o\,R_g\,r_e'}\left[\frac{I_{co}}{I_E} + \frac{1-\alpha_o}{\alpha_o} + \left(\frac{f}{f_T}\right)\right] \quad (311)$$

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 α_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. (311) 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. (311) is differentiated and the result is set equal to zero and solved for R_g. The following equation for R_g is then obtained:

$$R_g\,(\text{optimum}) = \left[(r_e'+r_b')^2 + \frac{\alpha_o\,r_e'\,(2\,r_b'+r_e')}{\frac{1-\alpha_o}{\alpha_o} + \left(\frac{f}{f_T}\right)^2 + \frac{I_{co}}{I_E}}\right]^{\frac{1}{2}} \quad (312)$$

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. (312) becomes small, and the optimum source resistance approaches $(r_b' + r_e')$. The typical noise figure for a high-frequency silicon transistor at 216 MHz is in the order of 2.5 to 3 dB. Such a noise figure cannot be obtained in a CATV amplifier because the source resistance for such an amplifier (75 ohms) is not the optimum source resistance required for minimum noise figure.

Cross-Modulation Distortion

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. Cross-modulation in a transistor amplifier is produced mainly by nonlinear characteristics of the emitter-to-base diode and by transfer characteristics that are not linear.

The voltage-current relations for the emitter-to-base diode of a transistor can be expressed as follows:

$$I_e = I_o [e^{(q-V)/kT} - 1] \qquad (313)$$

Expansion of Eq. (313) shows that the emitter current contains fundamental as well as harmonic components. The relative magnitude of the harmonics is proportional to the applied voltage. Therefore, when two or more signal voltages are applied, cross-modulation products are generated.

Transfer characteristics that are not linear, such as the current gain of a common-emitter stage, are another source of cross-modulation. If the output collector current is proportional to the input base current, no cross-modulation exists. However, because the common-emitter current gain (h_{fe}) is not constant with current levels, the output current is not proportional to the input current. When two or more signals are present, the amplitude of the first signal is dependent on the amplitudes of the other signals, and cross-modulation results.

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. 354. 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.

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 view-

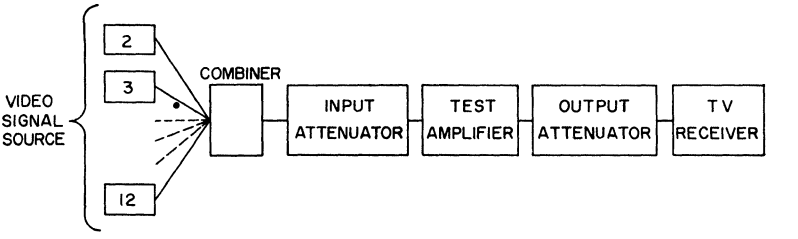


Figure 354. Block diagram of basic test setup used for cross-modulation measurements.

ing channel. This technique provides a white screen which does not change during the test, and allows more consistent and critical observations.

A test for cross-modulation can also be made by use of a field-strength meter in an arrangement such as that shown in Fig. 355. Two channel frequencies are used in the test. The measurement procedure is as follows:

1. The equipment is set up as shown in Fig. 355.
2. Generator No. 1 is set to 150 MHz modulated 30 per cent by 1000 hertz and the field-strength meter is tuned to 150 MHz.
3. The output of generator No. 1 is adjusted to the rated output of the amplifier.
4. The potentiometer is adjusted and the voltmeter is calibrated for a convenient level. This level then corresponds to 100-per-cent cross-modulation.
5. The modulation is then removed from the input signal.
6. Generator No. 2 is set to 210 MHz modulated 30 per cent by 1000 hertz and the field-strength meter is tuned to 210 MHz.
7. The output of generator No. 2 is adjusted to provide the rated output of the amplifier. (If the amplifier has a flat response, the output of the two signal generators will be equal.)
8. The field-strength meter is tuned to 150 MHz.
9. The voltmeter is set to proper scale for reading, and is read. The percentage of cross-modulation is then calculated based upon the 100-per cent level set in step 4.

This method is good enough to determine the cross-modulation of an amplifier on a relative basis. For accurate cross-modulation measurement, multi-channel frequencies must be used.

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

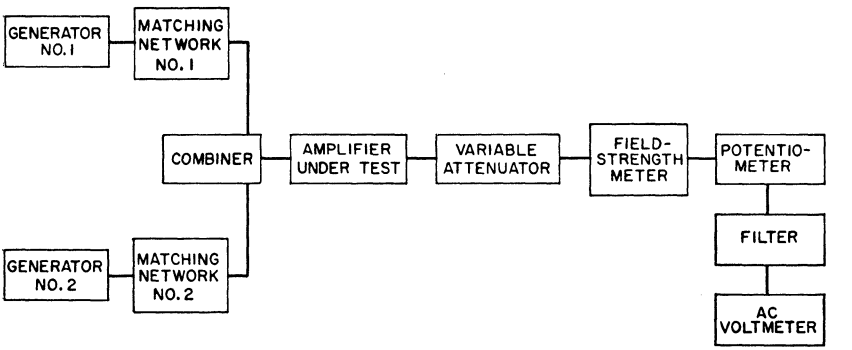


Figure 355. Block diagram of cross-modulation test setup in which a field-strength meter is used.

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. 356 shows the f_T of a typical 2N5109 as a function of collector current at a

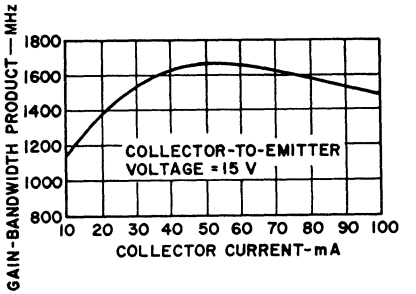


Figure 356. Gain-bandwidth product as a function of collector current for a typical RCA-2N5109 Transistor.

V_{CE} of 15 volts. The f_T measured at a collector current of 50 milliamperes is 1.5GHz; f_T is within 20 per cent of its maximum value from 25 to 100 milliamperes. Fig. 357 shows the f_T of a typical 2N5109 as a function of V_{CE} at

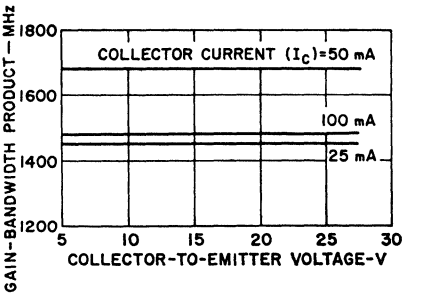


Figure 357. Gain-bandwidth product as a function of collector voltage for a typical RCA-2N5109 Transistor.

collector currents of 25, 50, and 100 milliamperes, respectively. The electrical characteristics of this transistor are summarized in Table XXXVI.

Table XXXVI—Electrical Characteristics of RCA-2N5109 Overlay Transistor (Case Temperature = 25°C)

CHARACTERISTIC	SYMBOL	TEST CONDITIONS DC COLLECTOR VOLTS		DC CURRENT (mA)		LIMITS		UNITS
		V_{CB}	V_{CE}	I_B	I_C	Min.	Max.	
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	BV_{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.

Fig. 358 shows the noise figure of a typical 2N5109 as a function of collector current. The noise figure is measured with the

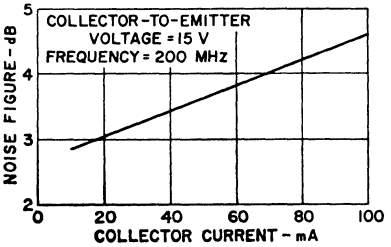


Figure 358. Noise figure as a function of collecter current for a typical RCA-2N5109 transistor.

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.

Choice of Operating Conditions for 2N5109

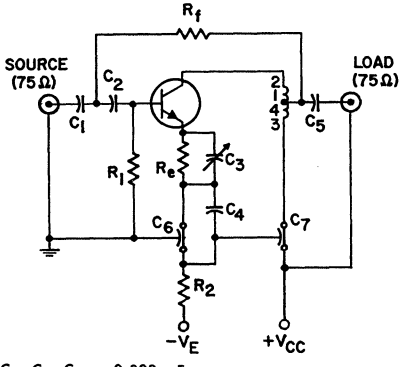
The most important parameter in the input stage of a CATV system is the noise figure. Distortion is not usually important in the input stage because the voltage and current swings of the transistor are small. The dc bias of the transistor should be chosen for mimimum noise figure. An RCA-2N5109 used 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. 356 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. 356 and 357 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. 356, 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. 357 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. 359. This common-emitter class A amplifier uses the RCA-2N5109 transistor and is designed for



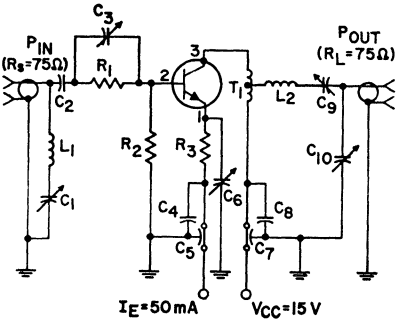
C_1, C_2, C_5 = 0.002 μF
C_3 = 8-60 pF, Arco 404 or equiv.
C_4 = 0.03 μF
C_6, C_7 = 1500 pF
R_1 = 390 ohms, ½ watt
R_2 = 330 ohms, 1 watt
R_e = 6.8 ohms, ½ watt
R_f = 200 ohms, ½ watt
T = 4-turn bifilar winding, 3/16" ID, No. 30 wire; Core: G.I. material Q_1 or equiv.

Figure 359. 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.

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. The dual-dc-power-supply arrangement is consistent with present CATV systems, which have both positive and negative 15-volt supplies.

The amplifier of Fig. 359 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.

Fig. 360 shows the use of the 2N5109 in a more sophisticated wideband amplifier. This amplifier also employs a 4:1-impedance wideband transformer and emitter peaking and degeneration. Output impedance matching at the upper cutoff frequency is provided by L_2, C_9 and C_{10}. This method results in a small increase in gain at the upper cutoff frequency. The same



C_1 = 1.5 — 20 pF
C_2 = 0.002 pF
C_3 = 55 — 300 pF
C_4, C_8 = 0.02 pF
C_5, C_7 = 1000 pF
C_6 = 32 — 250 pF
C_9 = 7 — 100 pF
C_{10} = 8 — 60 pF
L_1 = 7 turns, 3/16" diameter, No. 26 wire
L_2 = 3 turns, 3/16" diameter, No. 26 wire
R_1 = 30 ohms
R_2 = 390 ohms
R_3 = 220 ohms
T = 4-turn bifilar winding, 3/16" ID; Core: G.I. material Q_1 or equiv.

Figure 360. Single-stage wideband amplifier using the RCA-2N5109 transistor which employs various matching, peaking, and feedback techniques to increase the gain-bandwidth product.

technique is used in the input through L_1 and C_1. The combination of C_3 and R_1 tends to reduce the gain of the amplifier at low frequencies through the effect of R_1. At high frequency, full gain

is obtained through the shunting effect of C_3. As compared to the amplifier circuit of Fig. 359, this amplifier can provide 1 to 2 dB more gain under the same operating conditions. However, the amplifier circuit of Fig. 360 is more complicated, particularly its tuning procedure.

Trunk line and bridger amplifiers usually contain 2 to 4 stages. The single-stage circuits of Figs. 359 and 360 provide the basic building blocks for such an amplifier. The first stage must have a low noise figure and, for that reason, "tilt" control is usually not used in this stage. Because it increases the low-frequency noise figure, tilt control is usually incorporated only in the second stage.

VHF AND UHF MILITARY RADIO

Military radios, which operate in the vhf and uhf ranges, vary greatly in requirements. Telemetering devices may operate with as little output as 0.25 watt, while communication systems may require outputs of 50 watts and more. Modulation may be AM, FM, PM (pulse modulation), or PCM (pulse-code modulation). Equipment may be designed for fixed, mobile, airborne, or even space applications. Although the circuits described in this section apply only to specific military applications, they are representative of the general design techniques used in all military vhf and uhf radio equipment.

Sonobuoy Transmitters

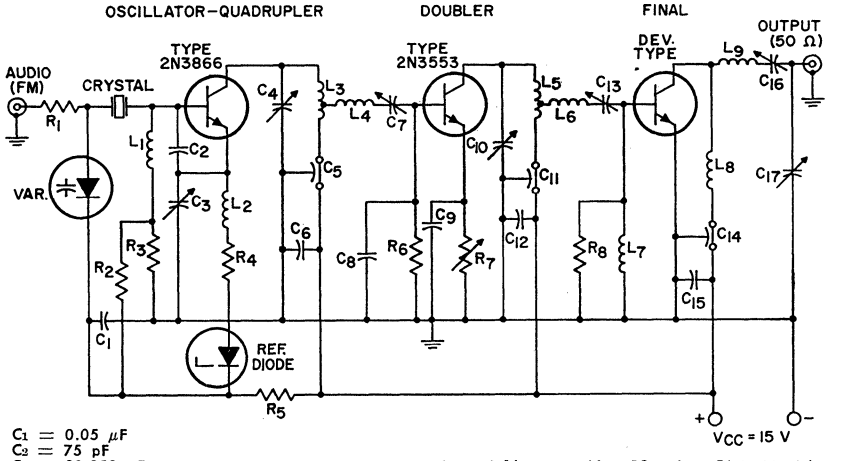
A sonobuoy is a floating submarine-detecting device that incorporates an underwater sound detector (hydrophone). The audio signals received are converted to a frequency-modulated rf signal which is transmitted to patrolling aircraft or surface vessels. The buoy is battery-operated and is designed to have a very limited active life.

Typical requirements for the rf-transmitter section of the sonobuoy are as follows:

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

Figure 361 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. 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.



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, 1/4" ID x 1/2" long
L_4 = 5 turns No. 22 wire, 3/16" ID (close wound)
L_5 = 2-1/4 turns No. 16 wire, 1/4" ID x 3/8" long
L_6 = 2-1/4 turns No. 18 wire, 7/32" ID (close wound)
L_7, L_8 = 1.0 μH
L_9 = 5 turns No. 16 wire, 1/4" ID x 3/8" 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 361. 1.5-watt (rf power output) sonobuoy transmitter.

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.

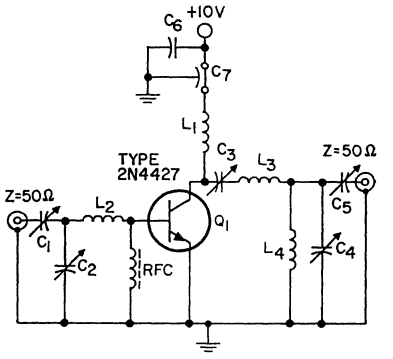
The final power output is developed by an RCA developmental transistor which operates as a straight-through class C amplifier at 160 MHz. A pi network matches this output to the 50-ohm line. The spurious output (measured directly at the output port) is more than 35 dB down from the carrier. This suppression 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. 362 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.

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 these lower supply voltages.

Sonobuoy circuits, in general, must be reliable, simple, and low in cost. The three-stage transmitter circuit shown in Fig. 361 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 0.5 watt.



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 μF, 50 V
C_7 = 1000 pF, feedthru
L_1 = 0.75 μH
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 362. 0.5-watt-175-MHz sonobuoy rf power output stage.

Air-Rescue Beacon

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
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. 363. Collector modulation, as well as some driver modulation, is used to achieve good down-modulation 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.

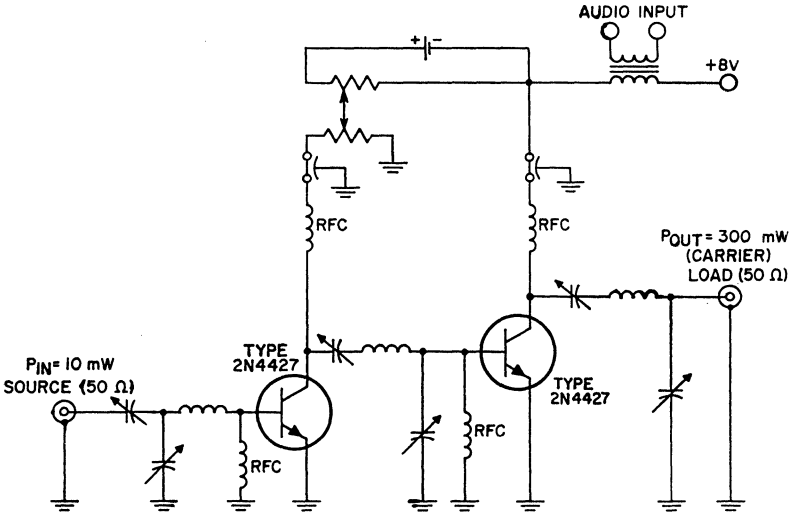


Figure 363. Driver and output stage for a 243-MHz beacon transmitter.

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. 364 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.

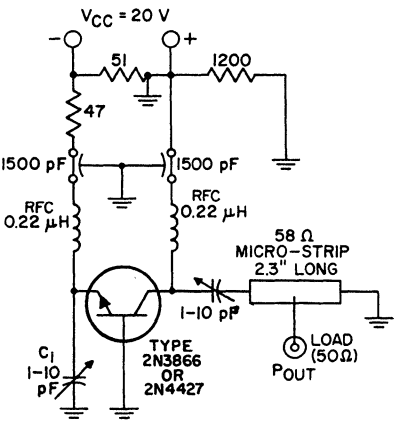


Figure 364. 1-watt, 500-MHz microstripline oscillator using the RCA 2N3866 or 2N4427 transistor.

Fig. 365 shows a very simple lumped-constant oscillator circuit for operation in the 700-to-1000-MHz frequency range.

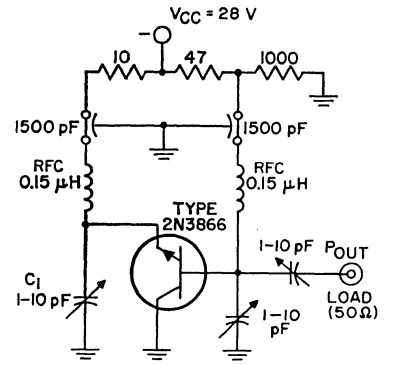


Figure 365. 0.5-watt, 1000-MHz lumped-constant oscillator using the RCA 2N3866 transistor.

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 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.

Military 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 limited to class B and class C amplifiers for use in driver or final output stages that provide power outputs in the range from 5 to 40 watts.

Fig. 366 shows the use of the 2N3733 transistor in a 5-watt power-amplifier circuit of microstripline construction. A common-emitter configuration is used. The transmission-line elements are constructed from 1/32-inch Teflon fiberglass microstrip board. The 750-ohm ferrite choke in the base return provides low-frequency stability for this circuit.

Fig. 367 shows the schematic diagram and components required for 225-MHz and 400-MHz lumped-constant circuits using the RCA-2N5016 transistor. This transistor can provide a power output

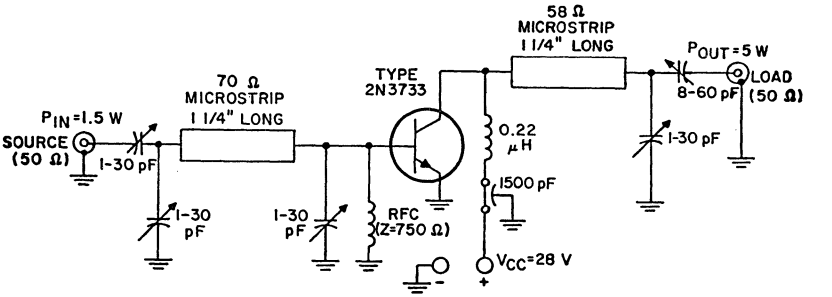
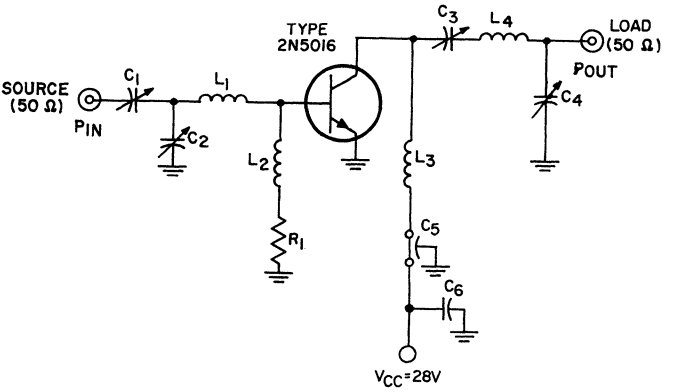


Figure 366. 5-watt, 375-MHz stripline amplifier.



FOR 225-MHz OPERATION
C_1 = 4-40 pF
C_2 = 7-100 pF
C_3 = 3-35 pF
C_4 = 8-60 pF
C_5 = 1500 pF, feedthru
C_6 = 0.01 μf, disc ceramic
L_1 = 1.3 turns No. 16 wire, ¼" ID, 3/16" long
L_2 = ferrite choke, Z = 750 ohms
L_3 = rf choke, 0.44 μH
L_4 = 4 turns No. 16 wire, ¼" ID, 0.3" long
R_1 = 0.68 ohms, wire wound, 1 watt

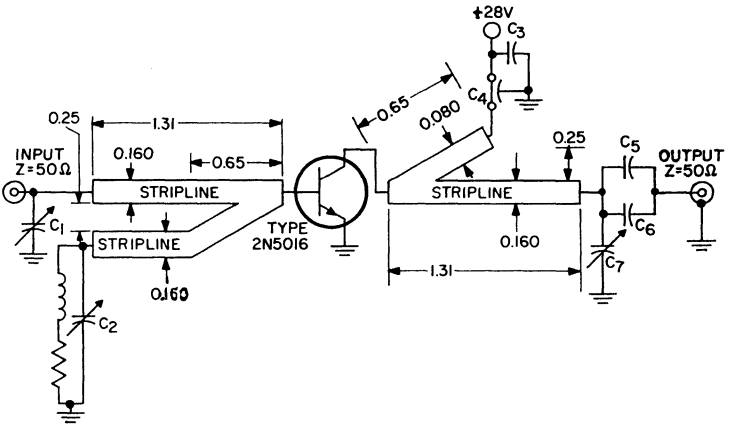
FOR 400-MHz OPERATION
C_1, C_3 = 1.5-20 pF
C_2, C_4 = 3-35 pF
C_5 = 1000 pF, feedthru
C_6 = 0.01 μF, disc ceramic
L_1 = 1.3 turns No. 16 wire, ¼" ID, ¼" long
L_2 = ferrite choke, 2 = 750 ohms (or 0.12-pH choke)
L_3 = rf choke, 0.13 μH
L_4 = 3 turns 1/32" x 1/8" copper ribbon, 3/16" ID, ½" long
R_1 = 0.68 ohms, wire wound, 1 watt

Figure 367. Lumped-constant power amplifier circuit for operation at 225 MHz or 400 MHz.

of more than 20 watts at 225 MHz and more than 15 watts at 400 MHz when operated from a 28-volt supply. A T-network is used to match the relatively low impedance of the input (1 to 3 ohms) of this device to a 50-ohm source impedance. A pi network is used to match the output to a 50-ohm load. Self-bias is provided by the base resistance R_1.

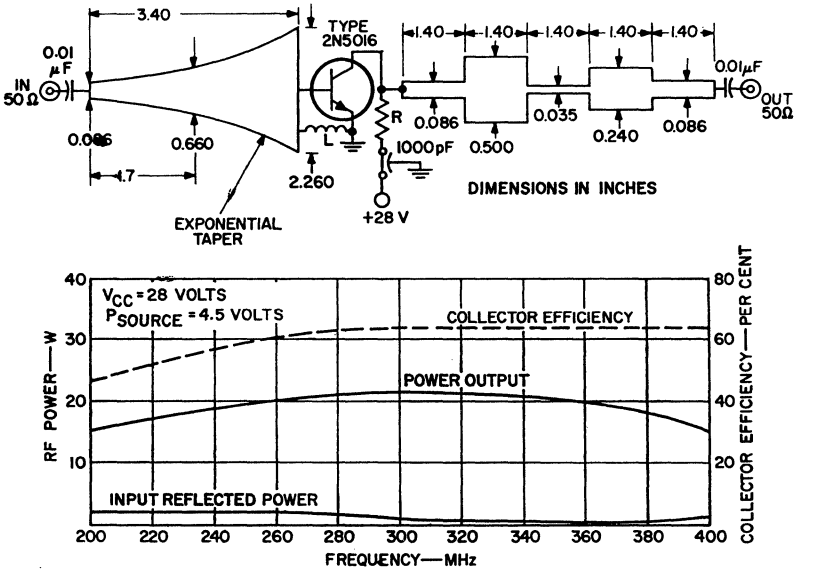
Figure 368 shows the use of a 2N5016 transistor in a microstrip amplifier circuit that provides an ouput of 15 watts at 400 MHz when operated from a 28-volt collector supply. The transmission lines are constructed on 1/16-inch Teflon fiberglass circuit board with the foil on one side acting as a ground plane. The side lines are used to tune out the transistor input and output reactances. The main lines, together with capacitors C_1 and C_7 are used to provide real-to-real impedance transformations. Base self-biasing is accomplished through resistor R_1. Even though no dc blocking is employed on the 50-ohm input, proper biasing is obtained, provided that the 50-ohm source impedance has a dc impedance of 50 ohms or more.

Fig. 369 shows a microstripline amplifier circuit designed for broadband operation over the frequency range of 225 to 400 MHz. This circuit is constructed on a 1/32-inch Teflon fiberglass circuit board with the foil on one side acting as a ground plane for the microstrip elements. This amplifier employs a broadband step-transformer output network which transforms the 50-ohm load down to about 20 ohms for the collector load and provides the best match in the center of



C_1, C_2, C_7 = trimmer capacitor, 2 to 18 pF, Amperex HTIOA/218 or equiv.
C_3 = 0.03 μF, ceramic disc
C_4 = 470 pF, feedthru, Allen Bradley FA5C or equiv.
C_5, C_6 = 0.005 μF, ceramic disc
R_1 = 5.1 ohms, 0.5 watt, carbon
Notes:
1. Broad = 1/16" Teflon board (ε = 2.6), Budd Co. Polychem Div., Grade 108T, 1 oz, double-clad copper, or equiv.
2. Dimensions in inches.

Figure 368. 400-MHz microstripline amplifier using the RCA-2N5016 transistor.



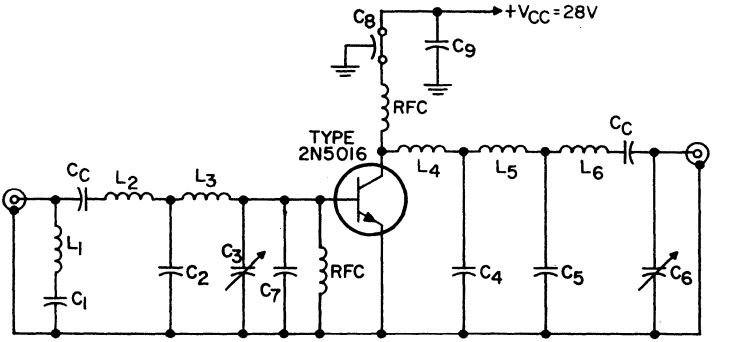
R = 0.27-ohm 1-watt wire wound
L = 0.1-μH rf choke

Notes:
Board = 1/32" Teflon (ε = 2.6), Budd Co. Polychem Div. grade 108T, 1 oz, double clad copper, or equiv.
Dimensions in inches.

Figure 369. Schematic and performance curves of a 200-to-400-MHz microstripline broadband rf amplifier.

the band. The fact that only an approximate match of the transistor output capacitance is provided at midband does not result in serious performance degradation. The input network consists of an exponentially tapered impedance transformer which matches the complex input impedance of the transistor at 400 MHz. At frequencies below 400 MHz, the input network matches the real part of the base impedance, but not the reactive part. The resultant mismatch offsets the increasing gain of the 2N5016 transistor as lower operating frequencies are approached and results in a relatively constant power output over the band. With 5 watts of drive, this circuit develops about 15 watts of output power across the 225-to-400-MHz band with a total output variation of 1.5 dB.

The lumped-constant circuit shown in Fig. 370 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



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 370. Lumped-constant 225-to-400-MHz power amplifier.

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 provides the last series inductor of the input network, and capacitor C_c is again used for dc blocking. As in the microstripline circuit, shown in Fig. 369, 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. 371.

The new RCA hermetic stripline package makes possible the design of broadband power amplifiers without compromise of hermetic reliability. This new radial lead package, shown in Fig. 372, uses ceramic-to-metal seals and is superior to uhf plastic packages in two respects: it has lower parasitic inductances and is hermetically sealed. For example, the first transistor in a series of hermetic radial-lead packages, had a dynamic input impedance of 1.5 + j1.2 ohms at 400 MHz as compared to 2 + j2 ohms for a plastic 2N5017. Fig. 373 shows the power output and efficiency of this developmental transistor as a function of input power, measured at a frequency of 400 MHz and a collector-to-emitter voltage of 28 volts. This transistor can deliver 19 to 20 watts of power output with a gain of 6.5 dB and a collector

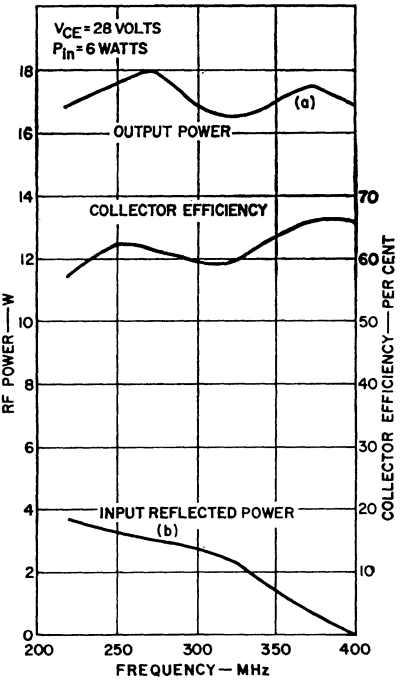


Figure 371. RF power output, input reflected power, and collector efficiency of the RCA 2N5016 transistor as functions of frequency.

efficiency of approximately 70 per cent at 400 MHz. An important feature of this transistor

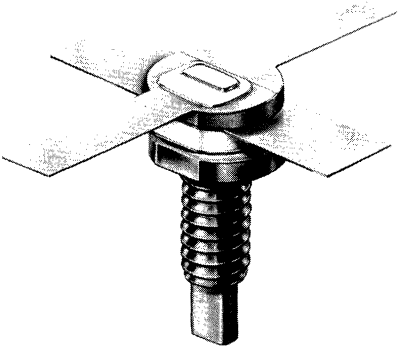


Figure 372. Hermetic Strip-Line Type Ceramic-to-Metal transistor Package (Isolated Electrodes)

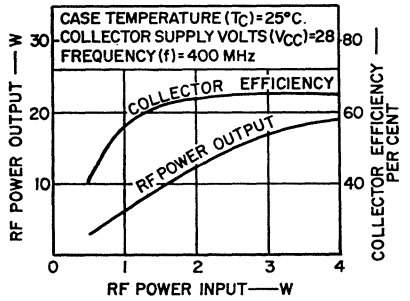


Figure 373. Typical power output or collector efficiency as a function of power input at 400 MHz.

is that the power gain is linear within 0.6 dB at power outputs in the range from 3 to 18 watts. The transistor can also supply 20 watts of output with a gain of more than 10 dB at 225 MHz. Fig. 374 shows the power output as a function of frequency for the developmental radial-lead transistor. Fig. 375 shows the circuit configuration for a 400-MHz amplifier that uses this transistor.

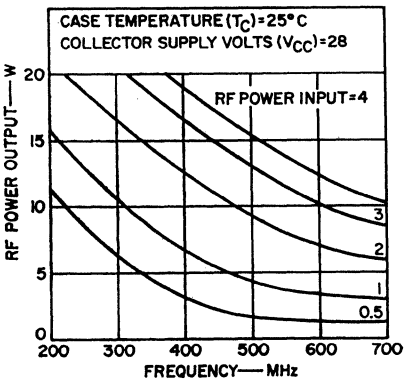


Figure 374. Typical power output as a function of frequency.

MICROWAVE AMPLIFIERS AND OSCILLATORS

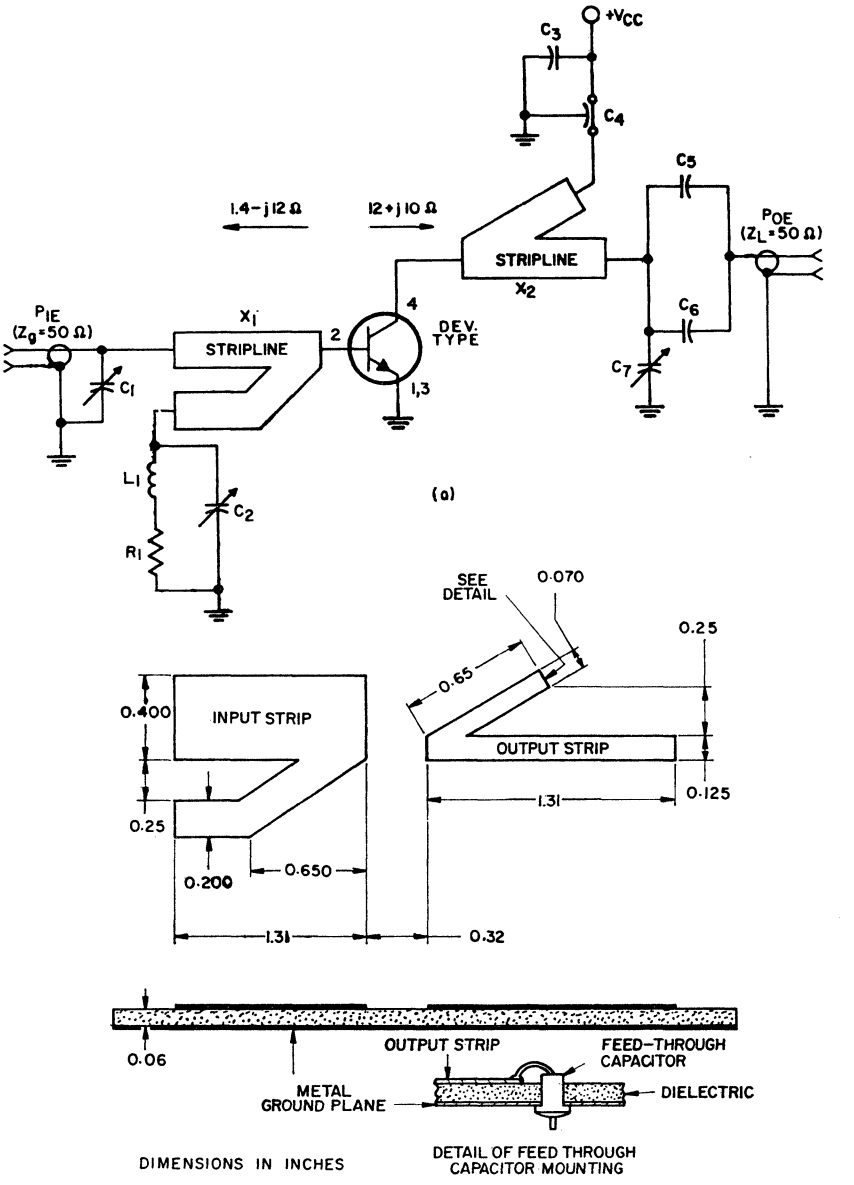
Continued development of rf power transistors has extended their high-power capability into the microwave region. Commercial microwave transistors capable of cw power outputs up to 1 watt are now available for use at frequencies up to 2.3 GHz. CW power outputs of 10 watts at 1 GHz and 5 watts at 2 GHz have been obtained with laboratory developmental transistors, and future prospects are even more promising.

This section describes the use and capabilities of overlay transistors in such applications as microwave straight-through amplifiers and fundamental-frequency oscillators. The use of overlay transistors as amplifier-multipliers and oscillator-multipliers also provides important building blocks in microwave equipment are discussed separately in the section on **Frequency Multipliers.**

Circuit-Design Considerations

The important performance criteria in microwave power-amplifier circuits are power output, power gain, efficiency, and bandwidth. Transistors suitable for power amplification must be capable of delivering power efficiently with sufficient gain in the frequency range of interest.

The power output that can be obtained from a transistor is determined by the current- and voltage-handling capabilities of the device in the frequency range of interest. Because high-voltage operation of high-frequency transistors is usually not prac-



C_1, C_2, C_7 = 2 to 18 pF, Amperex HTIOMA/218, or equiv.
C_3 = 0.03 μF, disc type
C_4 = 470 pF, feedthru, Allen-Bradley FA5C, or equiv.
C_5, C_6 = 0.005 μF, disc type
L_1 = 0.22 μH, rf choke
R_1 = 5.1 ohms, ½ W, carbon
X_1, X_2 = Details given in (b)

Figure 375. 400-MHz stripline rf amplifier test circuit for measurement of power output: (a) circuit diagram; (b) details of striplines.

tical, high power output requires a transistor which has a high current capability. The overlay transistor is suitable for high-power operation at ultrahigh and microwave frequencies because its construction provides a substantial increase in the ratio of emitter periphery to emitter area which results in a high current-handling capability.

A proper collector load impedance that provides the necessary voltage and current swings over the entire frequency range must be maintained to obtain maximum power output from the transistor. The real part of this load impedance should be a function of the collector supply voltage V_{CC} and the power output P_o. For class B operation, in which the collector rf voltage is sinusoidal, the real part of the load impedance can be expressed as follows:

$$R_L = \frac{[V_{CC} - V_{CE(sat)}]^2}{2P_o} \quad (313)$$

The dynamic output capacitance C_o of the transistor is very nearly equal to the open-circuit output capacitance C_{ob} at the supply voltage used. For class C operation, in which the collector rf voltage waveform resembles the waveform of a half-wave rectifier, the real part of the load impedance is modified as follows:

$$R_L = K\frac{[V_{CC} - V_{CE(sat)}]^2}{2P_o} \quad (314)$$

where the constant K is less than unity and is equal to 0.8 when the collector voltage waveform approaches that of a half-wave rectifier. In class C operation, the dynamic collector output capacitance C_o depends upon the conduction angle and is generally substantially higher than the capacitance C_{ob}. The imaginary parts of the load and transistor impedances are made part of the output circuit (where they are generally tuned out).

The power gain PG of a transistor power amplifier may be expressed in many forms; one of the simplest is as follows:

$$PG = (|h_{21}|^2 R_L)/[4 R_e (Z_{in})] \quad (315)$$

where h_{21} is the dynamic current-transfer ratio (current gain) of the transistor, R_L is the real part of the collector load impedance determined from the required power output, and $R_e(Z_{in})$ is the real part dynamic input impedance (when the collector is loaded by a complex load impedance Z_L).

Eq. (315) shows that, for high-gain operation, large-signal power transistors should have (a) high current gain, which is also required for small-signal operation; (b) constant current gain with current-level variations for large-signal operation; and (c) low dynamic input impedance. This type of performance can usually be obtained with overlay transistors.

For good efficiency, transistors are usually operated under the signal-bias condition at which the collector-to-base junction is reverse-biased and the emitter-to-base junction is partially forward-biased by the input-drive signal. The collector efficiency of a transistor amplifier is defined as the ratio of the rf power output at the frequency of interest to the dc power input. Therefore, high efficiency implies that circuit losses should be minimal, that the ratio of the real part of the dynamic output admittance of the transistor to its collector load

conductance should be maximum at the frequency of interest, and that the circulation of harmonic current in the circuit should be minimized. For efficient operation, a transistor that has a small dynamic output conductance should be used in a low-loss circuit that can block the flow of all harmonic currents. Because transistor output conductance increases very rapidly with frequency, a high ratio of output conductance to collector-load conductance is essential for efficient operation.

The bandwidth of a transistor power amplifier is limited by three factors: (1) the intrinsic transistor structure, (2) transistor parasitic elements, and (3) the external circuitry, i.e., the input and output matching networks. For a given collector load resistance and frequency, the inherent bandwidth of a transistor power amplifier increases with the gain-bandwidth figure of merit f_T and is inversely proportional to the collector-to-base capacitance C_{ob}. Transistors suitable for wideband operation, therefore, should have a high f_T and a low C_{ob}.

Parasitic elements contributed by the transistor package impose limitations on bandwidth. The most critical parasitic elements are base and emitter inductances. Either type of inductance can be reduced greatly by grounding the corresponding terminal directly to the package. These inductances, in series with the intrinsic transistor structure, increase the input Q and thus reduce the bandwidth. Additional parasitic inductances also increase the variation in transistor input and output impedances with frequency.

A suitable package for transistors required to provide stable operation over large bandwidths in microwave-frequency applications should have low common-lead inductance, low shunt and feedthrough capacitances, and good thermal-resistance properties. Such requirements are readily obtainable by use of coaxial or stripline transistor circuit structures designed for use at microwave frequencies. Best high-frequency performance is provided by common-base configurations in coaxial packages of the type shown in Fig. 376. The lower section of

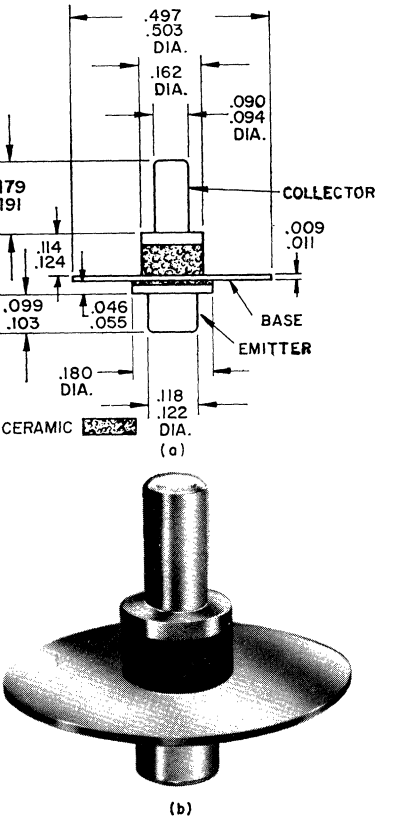


Figure 376. RCA-2N5470 coaxial transistor (a) outline; (b) external view.

the coaxial package, which contains the emitter coaxial lead, is insulated from the flange, which serves as the base lead, by an aluminum oxide disc. Another disc separates the base flange from the copper collector lead. The use of this package in a coaxial circuit requires a proper heat sink. Fig. 377 shows a coaxial transistor package that uses a standard beryllium oxide ring to facilitate heat conduction from the center conductor to the outside conductor of an air-dielectric line section. This type of arrangement is useful for 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.

The coaxial line package is useful well into the S-band frequency range. The JEDEC TO-60 and TO-39 packages are useful in amplifier circuits to 1 GHz and in oscillator circuits to 2 GHz. The hermetic construction of these packages leads to increased device reliability under stringent environmental operating conditions.

The choice of transistor configurations at uhf and microwave frequencies is dependent upon both performance and stability requirements. Common-base amplifiers provide higher gains than common-emitter amplifiers at frequencies above the f_T of the transistor. Collector efficiency for common-base and common-emitter circuits is about the same. It is generally acknowledged, however, that a common-emitter configuration provides a more stable circuit at frequencies below the f_T of the transistor. This assumption arises from the linear analysis of transistors in which parasitic elements are not included. For high-power operation at uhf and microwave frequencies, transistor parasitics contributed by the package must be treated with the intrinsic

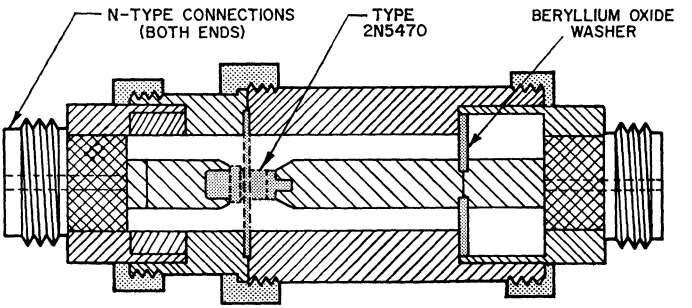


Figure 377. Heat sink for use with coaxial transistor package.

transistor. Stable operation can also be obtained in common-base configurations provided parasitic elements can be controlled.

Because transistor parameters change with the power level, certain forms of instabilities (e.g., hysterisis, parametric oscillations, and low- and high-frequency oscillations) can be incurred in both common-emitter and common-base circuits. Usually, most of these instabilities can be eliminated or minimized by careful design of the bias circuit, by proper location of the transistor ground connections, and by use of packages in which parasitic inductance and capacitance are held to a minimum. Stable operation has been obtained at 2 GHz with both common-base and common-emitter configurations. However, common-base coaxial packages have been empirically found to be more stable at the higher frequencies.

Circuit-Design Approach

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, such as h_{21} in Eq. (315), 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 with the large-signal values at 28 volts. Because 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. 378. The 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

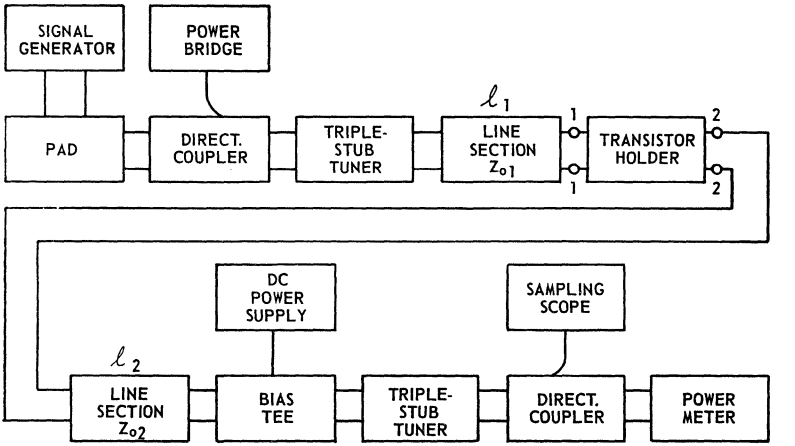


Figure 378. Block diagram of test setup used to determine input and output impedances of transistors.

in a slotted-line set-up or with a vector voltmeter. The conjugate of this impedance equals the dynamic 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.

In addition to determining dynamic input impedance and load impedance, the system shown in Fig. 378 is also 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 sections (l_1 and l_2, respectively). Eighth-wavelength ($\lambda/8$) line sections are preferred for l_1 and l_2 because, as pointed out in the discussion of **Matching Networks**, such sections exhibit the lowest VSWR and the smallest line losses.

An alternative method of determining the dynamic input impedance is shown in Fig. 379. This method uses a well-padded, high-power signal generator connected in series with a slotted-line setup.

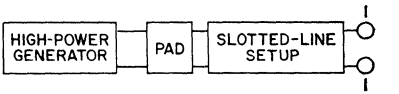


Figure 379. Block diagram of dynamic-impedance test setup that may also be used to test transistor performance capability.

The setup beyond terminals 1-1 is identical to that of Fig. 378. 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.

Circuit-Design Techniques

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 the manufacturers published data, the input and output circuits can be designed. The network design methods described in the section **Matching Networks** may be used. For most microwave-circuit applications, however, either air-line or stripline arrangements are generally used.

This section discusses only some simple designs of the types shown in Fig. 380. 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. 380(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-wave-

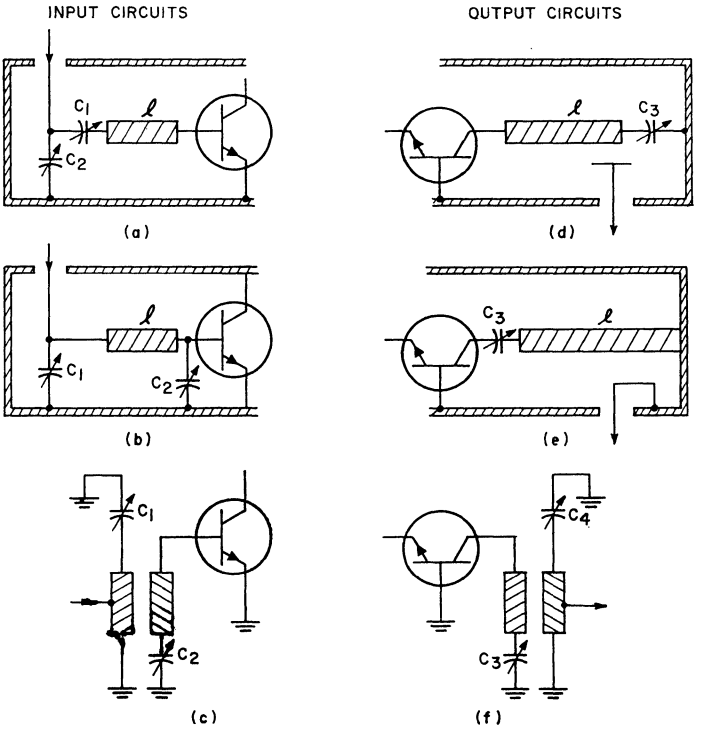


Figure 380. Transistor input and output coupling circuits suitable for use at microwave frequencies: (a) direct-coupled input network using series tuning capacitor; (b) direct-coupled input network using shunt tuning capacitor; (c) resonant-line input circuit; (d) capacitive-probe-coupled output cavity; (e) inductive-probe-coupled coaxial output cavity; (f) resonant-line output circuit.

length long and its characteristic impedance Z_0 is determined by use of Eq. (290), 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 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. 380(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. The characteristic impedance of this line can be determined from Eq. (289). Capacitor C_1 may be used to adjust for minor differences between transistors.

The output circuit shown in Fig. 380(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. The location of this probe is best determined empirically because a mathematical relation for such a case has not been developed. 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.

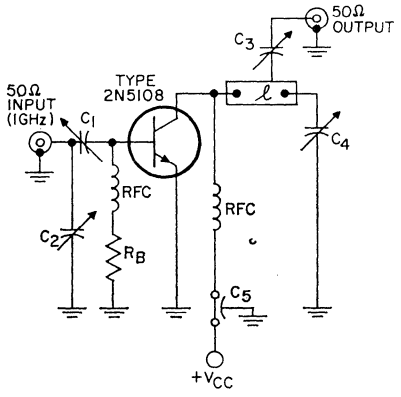
The ouput circuit shown in Fig. 380(e) is similar to that shown in Fig. 380(d) except that inductive loop coupling is used. Again, the design of the coupling loop is empirical. In general, the inductive loop is placed near the ground (high-current) end of the line; in fact, it may be tapped directly to the center conductor. Conversely, capacitive probes are generally located near the high-voltage end of the line.

The coupling networks shown in Fig. 380(a) through Fig. 380(e) can apply to either input or output circuits, and the specific illustrations are used for discussion only. The circuits shown in Figs. 380(c) and 380 (d) make use of inductive coupling and are particularly suitable for stripline circuits. This technique enables additional circuit isolation as well as additional filtering action. The technique can also be extended to provide pass-band filtering.

Microwave Amplifier Circuits

The RCA 2N5108 transistor can be used in the common-emitter amplifier mode at L-band frequencies. A typical circuit con-

figuration for operation at the 1.0-to-1.5-GHz range is shown in Fig. 381. This circuit can provide an output power of 1 watt at 1 GHz when operated from a



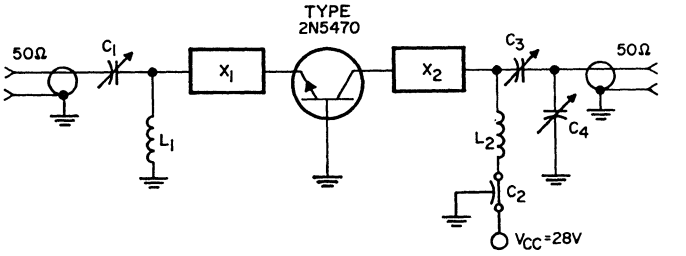
C_1, C_2, C_3, C_4 = Variable capacitor, 0.3 to 3.5 pF, Johanson piston type or equiv.
C_5 = 420 pF, feedthru
l = As described in text
R_B = 2.7 ohms, ¼ W
RFC = 0.1 μH

Figure 381. 1-GHz, 1-watt amplifier using the RCA-2N5108 transistor.

28-volt power supply. The emitter of the transistor is directly connected to the ground plane of the stripline circuit board. The input circuit consists of the capacitors C_1 and C_2 and the parasitic lead inductance of the 2N5108 transistor. The output circuit uses a capacitively loaded 50-ohm section of stripline which is resonant at the operating frequency. Power output is taken from this line at the proper impedance level. Amplifier power gain is in the order of 6 dB and collector efficiency is about 35 per cent.

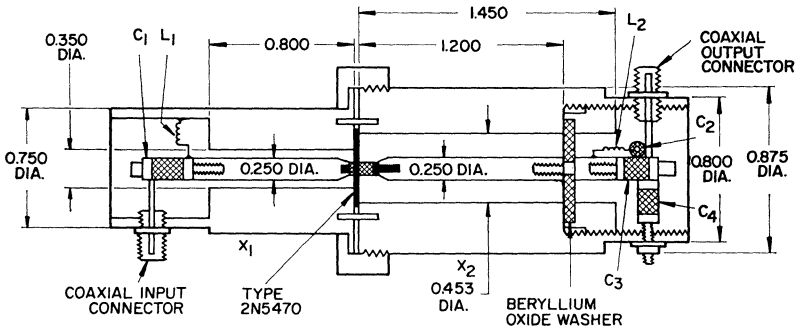
For operation at high L-band or low S-band frequencies, the RCA-2N5470 coaxial transistor can provide greater stable power outputs. Fig. 382 shows a coaxial-line amplifier circuit which can provide 1.2 watts of output power at 2 GHz with a 28-volt power supply. The coaxial transistor is placed in series with the center conductors of the coaxial air lines, and the base is properly grounded to separate the input and output cavities, as shown in Fig. 380. The input line L_1 has a characteristic impedance Z_0 of 20 ohms and a length of about 0.80 inch. This line directly transforms the input impedance of the 2N5470 (at 2 GHz) to an impedance of about 50 ohms at the input. The output line L_2 has a characteristic impedance of 36 ohms and a length of about 1.8 inch. Power output is taken from the capacitor network loading this output line. The two rf chokes isolate the rf lines from the bias supply.

The 2N5470 coaxial-line amplifier can supply a cw power output of 1.2 watts at a gain of 5 dB and has a collector efficiency greater than 35 per cent. Fig. 383 shows the variation in power output of the 2N5470 transistor as a function of frequency at the 0.2-watt and 0.3-watt drive levels, and Fig. 384 shows the transistor power output and collector current as a function of drive power at the 2-GHz operating point. Because of the excellent input- and output-circuit isolation (within the 2N5470 transistor as well as in this coaxial-circuit design), the common-base circuit configuration shown in Fig. 382 is extremely stable. Improved efficiencies and power outputs can be expected by use of boron nitride dielectric-loaded "heat sink" lines, as discussed previously.



C_1 = 0.8 to 10 pF, Johnason 4355, or equiv.
C_2 = 1000 pF, feedthru, Allen-Bradley FB2B, or equiv.
C_3 = 0.3 to 3.5 pF, Johnason 4701, or equiv.
C_4 = 0.35 to 3.5 pF, Johnason 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)

(b)



DIMENSIONS IN INCHES

(b)

Figure 382. 2-GHz power amplifier using the RCA-2N5470 coaxial transistor: (a) circuit schematic; (b) construction details.

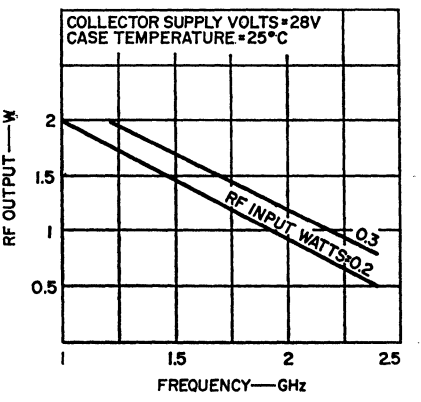


Figure 383. Power output as a function of frequency for the RCA-2N5470 transistor.

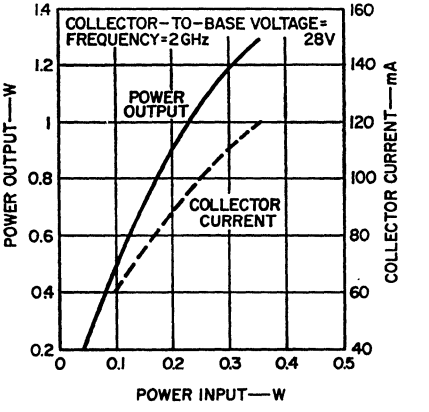


Figure 384. Power output as a function of input for the 2-GHz amplifier shown in Figure 382.

FUNDAMENTAL-FREQUENCY OSCILLATORS

Transistors capable of power amplification are also suitable for power oscillation. The most important 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.

Basic Microwave Oscillator Circuits

The parasitic elements of a transistor package can sometimes be used to form an economical microwave oscillator circuit. The high-frequency oscillations discussed previously usually occur at a frequency close to the output frequency of the amplifier when the input power is removed. This form of instability can be attributed to the parasitic elements of the package which set up the frequency of oscillation with the intrinsic transistor. Such parasitic elements can be used to form a transistor oscillator operated at microwave frequencies, provided the frequency of oscillation can be controlled.

Fig. 385(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. Although the transistor configuration is not too well defined in this oscillator circuit, the device can be grounded in high-frequency operation 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. 385(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. 385(c) shows another method of coupling power output from the oscillator.

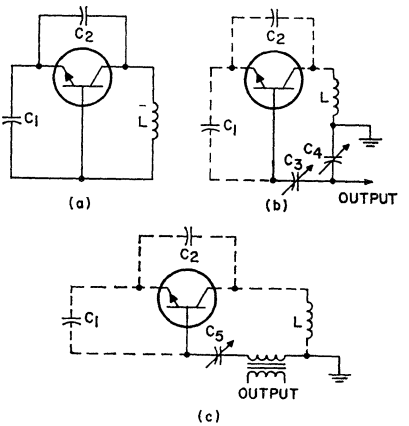
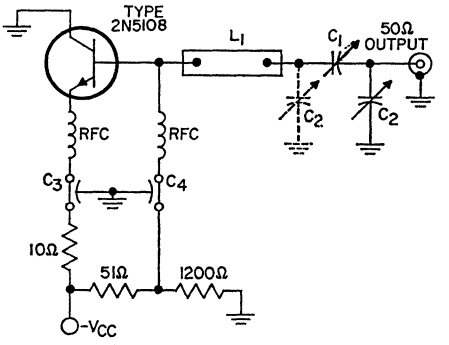


Figure 385. Colpitts oscillator for use at microwave frequencies: (a) basic ac circuit configuration; (b) basic ac circuit with collector returned to ground through parasitic inductance L and the output taken from base through capacitive voltage divider; (c) basic ac circuit with transformer coupled output.

L-Band Oscillators

Fig. 386 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 386. 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. 387 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

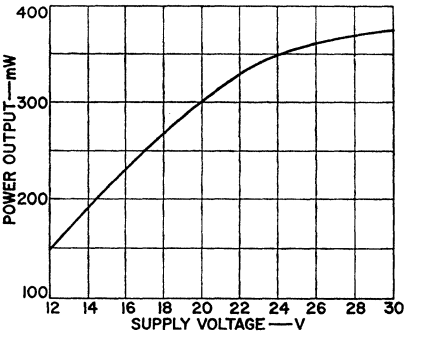
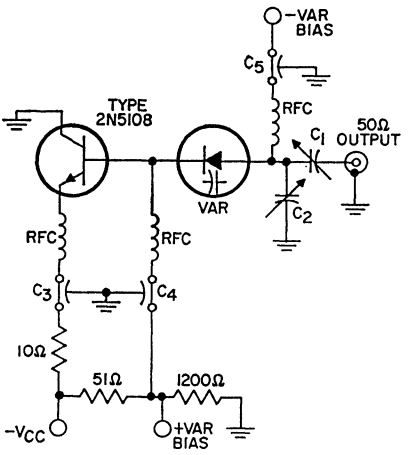


Figure 387. Power output as a function of supply voltage for the 1.68-GHz oscillator shown in Fig. 386.

range. For example, the line length is increased to 0.8 inch 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. 386) can be supplied by a high-Q varactor diode, as shown in Fig. 388. 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



RFC = 0.1 μH
C_1, C_2 = 1 to 7 pF, piston capacitors
C_3, C_4, C_5 = 470 pF, feedthru
Var. = Described in text

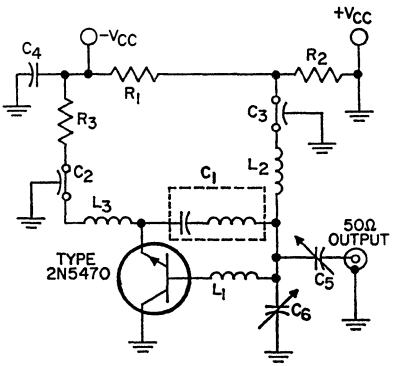
Figure 388. Wideband Varactor-tuned L-Band Oscillator using the RCA-2N5108 transistor.

with an inductive-reactive divider-network, such as a tapped 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.

S-Band Oscillators

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 in 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. 389 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.



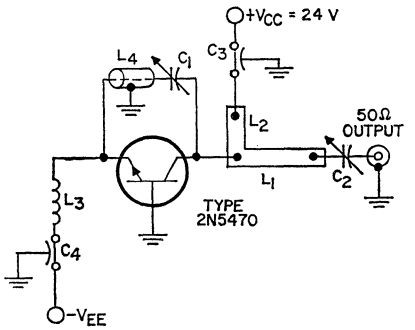
C_1 = 0.82 pF, "gimmick" capacitor (manufactured by Quality Components, Inc. St. Mary's, Pa.)
C_2, C_3 = 100 pF, Allen-Bradley FA5C or equiv.
C_4 = 0.01 pF, disc ceramic
C_5, C_6 = Trimmer capacitor 0.35 to 3.5 pF, Johanson Type 4702 or equiv.
L_1 = 0.05" length of No. 22 wire
L_2, L_3 = 4 turns 7-mil wire, .062" ID x 3/16" long
R_1 = 51 ohms, 1/2 watt
R_2 = 1200 ohms, 1/2 watt
R_3 = 5 to 10 ohms, 1/2 watt

Figure 389. Lumped-constant 2-GHz oscillator circuit using the RCA-2N5470 transistor.

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. 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 390 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 390. 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.

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 permits 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 operates 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:

$$f_{max} = (PG)^{\frac{1}{2}} f = [(1/8\pi)(1/r_{bb}' C_c \tau_{ec})] \quad (316)$$

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 τ_{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:

$$f_{VCB} = 1/[2\pi C_{min} (r_b'+r_s)] \quad (317)$$

where C_{min} is the minimum collector-to-base capacitance and r_s is the collector series resistance.

Fig. 391 shows a cross-sectional view of an overlay transistor that indicates the capacitance and loss distributions. Fig. 392 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

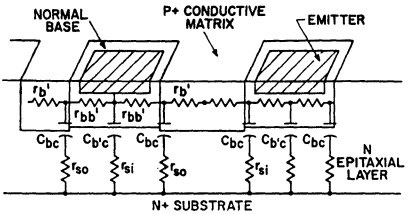


Figure 391. Cross-sectional view of an overlay transistor indicating the capacitance and loss distributions.

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 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 extrin-

sic 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. 392 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

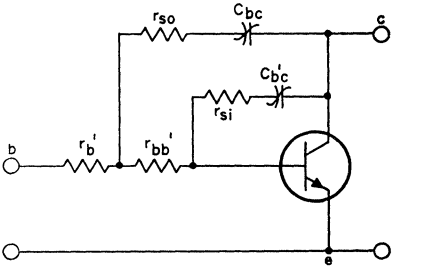


Figure 392. Circuit showing the nonlinear impedance factors that make possible frequency multiplication with overlay transistors.

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:

$$C_{bc} = K\,(\phi - V)^{-n} \qquad (318)$$

where K is a constant determined by the area and doping of the junction, θ 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 393 shows the variation in the collector-to-base capacitance 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

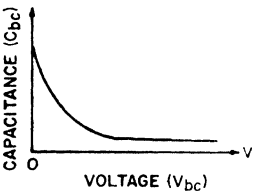


Figure 393. Collector-to-base capacitance C_{bc} as a function of collector bias in overlay transistors.

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:

$$Q = \int C_{bc}\, dV \qquad (319)$$

If the capacitance C_{bc} is defined as in Eq. (317), the integration indicated in Eq. (318) can be performed with respect to the voltage V to obtain the charge. The result of this integration, shown in Fig. 394, shows the variation in the charge Q as a function of the voltage V_{BC}.

If a sinusoidal voltage such as that shown in Fig. 395(a) is developed by the amplifier section of the overlay transistor to drive the

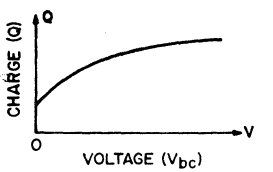


Figure 394. The charge Q in the collector-to-base junction as a function of collector-to-base voltage in an overlay transistor.

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. 395(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.

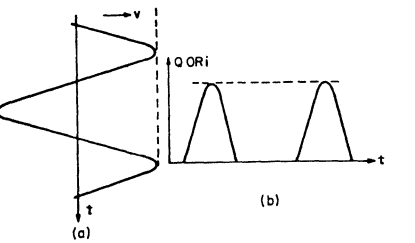


Figure 395. (a) Sinusoidal voltage developed by the amplifier section of an overlay transistor to drive the nonlinear collector-to-base capacitance; (b) distorted charge, or current, waveform produced by the nonlinear collector-to-base capacitance of an overlay transistor in the generation of harmonic 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 simi-

lar to that of the series-type varactor frequency multiplier.

Fig. 396 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

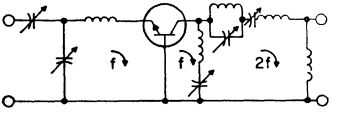


Figure 396. Basic configuration for use of an overlay transistor in a common-base frequency doubler.

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. 397 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 addi-

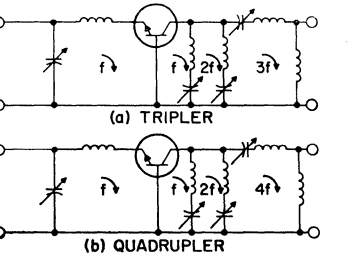


Figure 397. Basic configurations for use of an overlay transistor in a common-base frequency tripler (a) and frequency quadrupler (b).

tional 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.

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). These dynamic impedances can be determined experimentally as described in the section on **Microwave Amplifiers and Oscillators.** 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 hysterisis, 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 de-

creased. 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 minimized, 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. 398.

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. 398 to provide a low impedance to the spurious component in addition to that provided by the rf bypass capacitance C_1.

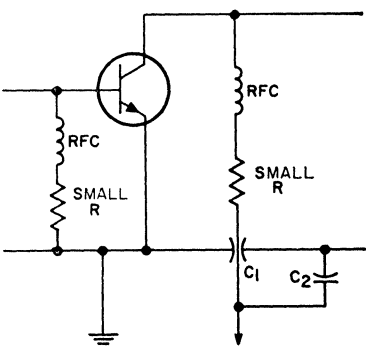


Figure 398. Circuit showing biasing techniques and bypassing capacitances used to eliminate instabilities in common-emitter frequency multipliers.

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. 399 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 3.4 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

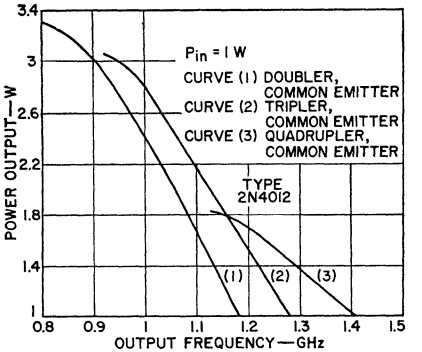


Figure 399. Power output of the RCA-2N4012 overlay transistor as a function of frequency when operated in common-emitter doubler, tripler and quadrupler circuits.

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. 399 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 this section 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. 400 shows the complete circuit diagram of a 400-to-800-MHz doubler that uses the 2N4012 tran-

sistor. This circuit uses lumped-element input and idler circuits and a coaxial-cavity output circuit.

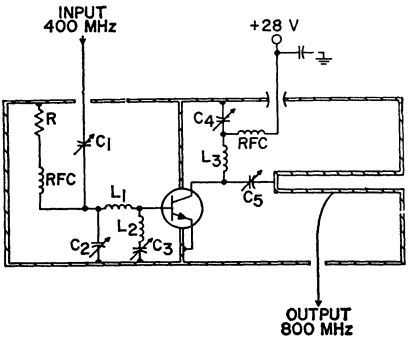


Figure 400. 400-to-800-MHz common-emitter transistor frequency multiplier.

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 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. 400.

Fig. 401 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

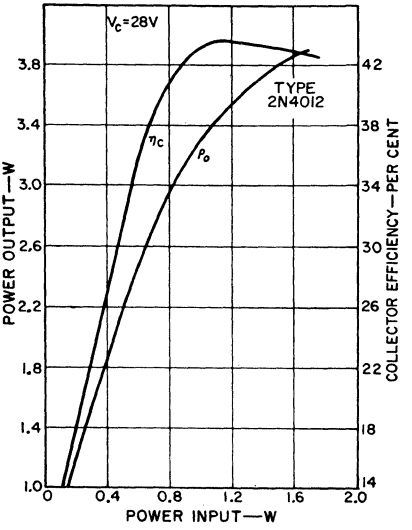


Figure 401. Output power and collector efficiency as a function of input power for the 400-to-800-MHz frequency doubler.

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. 401. The efficiency is 43 per cent measured at an input power of 1 watt. The 3-dB bandwidth of this circuit measured at 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

shown in Fig. 402. This curve is obtained with the circuit tuned at 28 volts. The curves of Figs. 401 and 402 indicate that the transistor amplifier-multiplier circuit is capable of amplitude modulation.

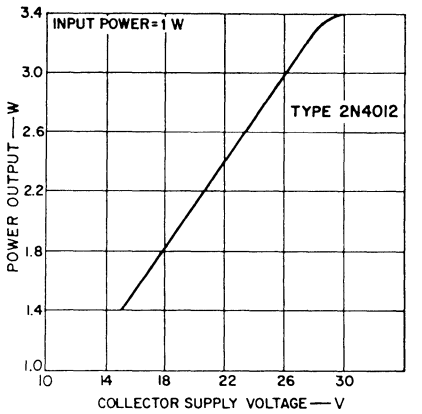


Figure 402. Power output as a function of supply voltage for the 400-to-800-MHz frequency doubler.

367-To-1100-MHz Tripler

The 367-to-1100-MHz tripler shown in Fig. 403 is essentially the same as the doubler shown in Fig. 400 except that an additional idler loop, (L_4, 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. 404 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

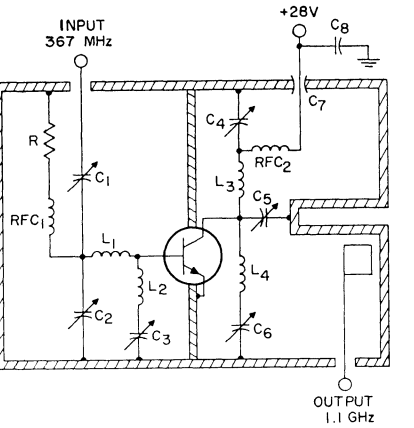


Figure 403. 367-MHz-to-1.1-GHz common-emitter transistor frequency tripler.

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 2.3 per cent. The spurious-frequency components measured at

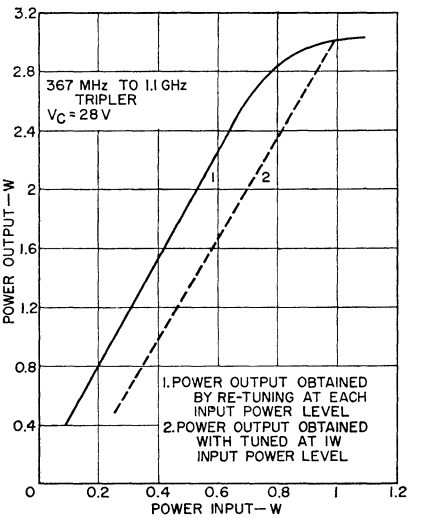


Figure 404. Power output as a function of power input for the 367-MHz-to-1.1-GHz frequency tripler.

the output are as follows: −22dB at 340 MHz, −30dB at 680 MHz, −35dB at 1360 MHz.

The variation of power output with collector supply voltage at an input drive level of 1 watt is shown in Fig. 405. The variation of collector efficiency is also shown. These curves were obtained with the circuit tuned at 28 volts.

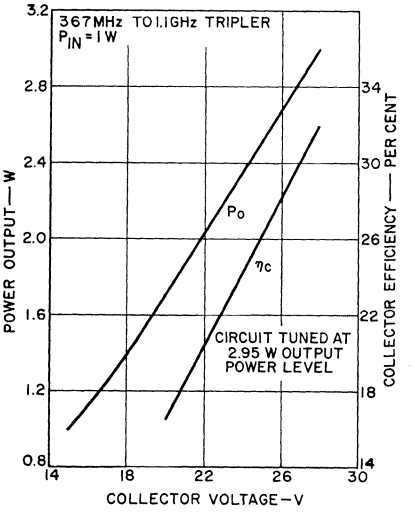


Figure 405. Power output as a function of collector supply voltage for the 367-MHz-to-1.1-GHz frequency tripler.

A 367-MHz amplifier that used the same circuit configuration and components as those of the tripler circuit shown in Fig. 403 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 to 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. 406 illustrates the use of the 2N4012 transistor in a 150-to-450-MHz frequency tripler. The input coupling network is designed

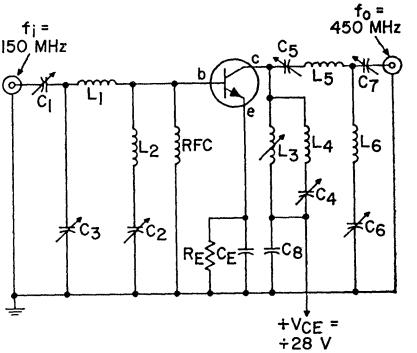


Figure 406. 150-to-450-MHz common-emitter transistor frequency tripler.

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

formed by C_5, C_6, C_7, L_5, and L_6 provides the required collector loading for 450-MHz power output. Fig. 407 shows the 450-MHz power output of the tripler as a function of the 150-MHz power

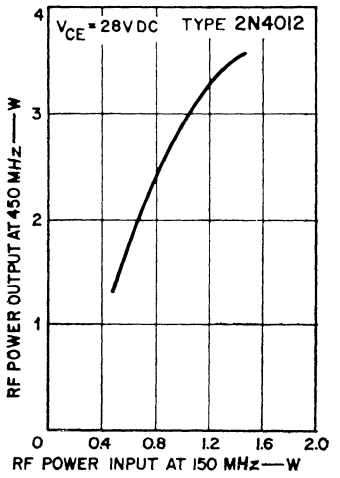


Figure 407. Power output as a function of power input for the 150-to-450-MHz frequency tripler.

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. 408.

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

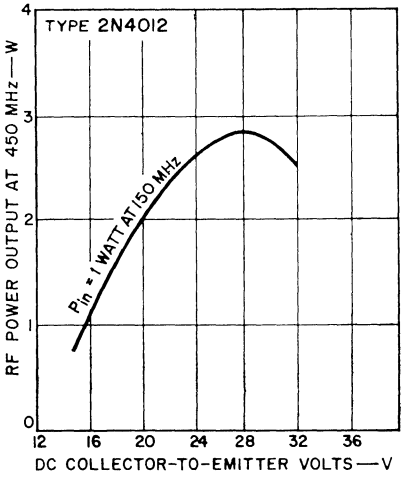


Figure 408. Power output as a function of collector supply voltage for the 150-to-450-MHz frequency tripler.

are obtained. Fig. 409 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

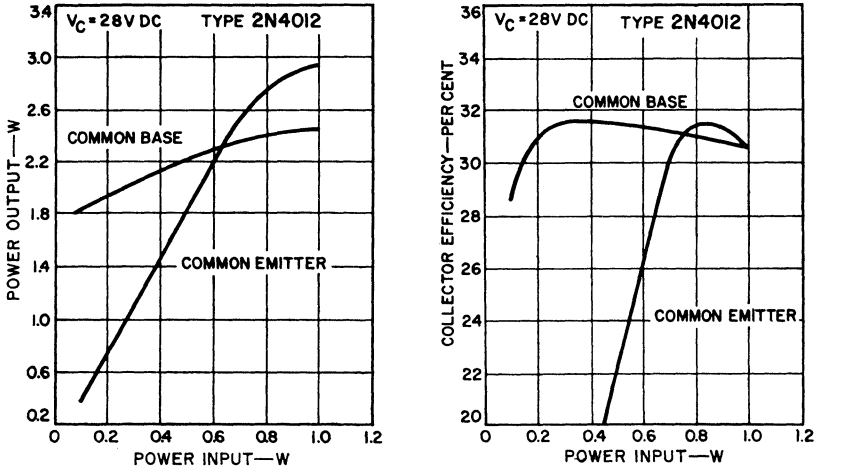


Figure 409. Comparison of performance characteristics of common-base and common-emitter tripler circuits using the RCA-2N4012 transistor: (a) Power output as a function of power input; (b) collector efficiency as a function of power input.

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.

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. 410 shows an oscillator-quadrupler circuit that uses a selected 2N3866 transistor. This circuit can deliver a power output 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.

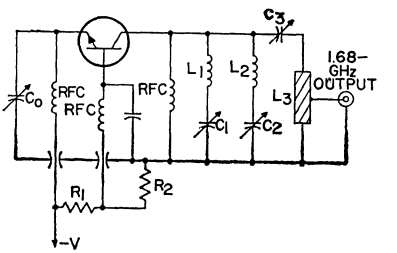


Figure 410. Oscillator-quadrupler circuit.

Control and Low-Frequency Power Amplifiers

SILICON power transistors offer many advantages when used in the power-output and driver stages of high-power audio amplifiers and in applications such as ultrasonic generators, servomechanism control systems, inverters, and automobile ignition systems. In these applications, silicon transistors can be used, over a wide range of ambient temperatures, to develop power output of tens of watts to drive loudspeakers, ultrasonic transducers, or servo motors. Alternatively, silicon-transistor amplifiers may be required to increase the output of some type of transducer to a level at which it may be used to control some process or indicator.

GENERAL CONSIDERATIONS

Transistor power amplifiers may be designed for operation in either linear or pulsed (switching) service. In pulsed service, the transistors are switched, usually in response to a control signal, between cutoff and saturation to develop a rectangular-wave output. This switching may be symmetrical to provide equal ON and OFF times, or may be asymmetrical for increased or decreased ratios of ON time to OFF time as determined by the control function desired. In linear service, the circuit designer can select any one of three basic classes of operation for the transistors. This selection is dictated by a combination of such factors as required power output, dissipation capability, efficiency, gain, and distortion characteristics.

Classes of Operation for Linear Amplifiers

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. 411. At low power levels, the class A amplifier can also be coupled to the load by resistor, capacitor, or direct coupling techniques.

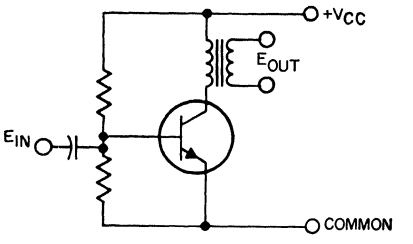


Figure 411. Basic class A, transformer-coupled amplifier.

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. 412.

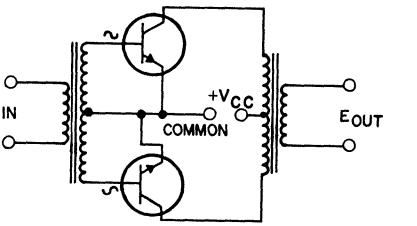


Figure 412. Basic class B, push-pull transformer-coupled amplifier.

If conduction 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 for Linear Amplifiers

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. 413. 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

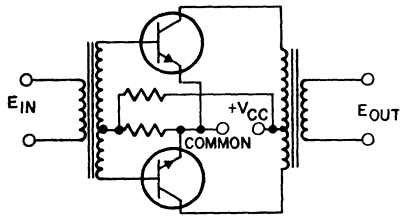


Figure 413. Basic class AB, push-pull, transformer-coupled amplifier.

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.

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. 414. 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

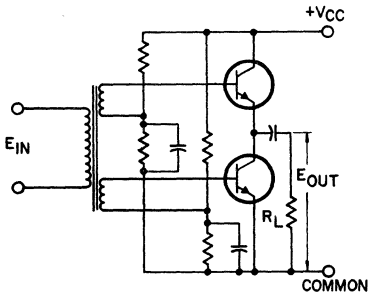


Figure 414. Class AB, push-pull amplifier with series output connection.

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. 415. 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. The driving transistors are connected directly to the bases of the output transistors, as illustrated in Fig. 416.

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 ad-

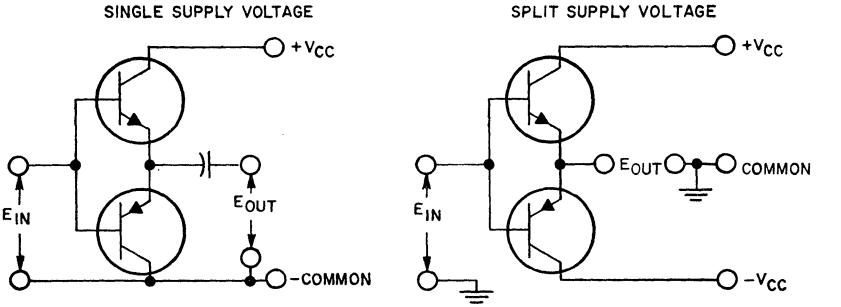


Figure 415. Circuit arrangements for operation of complementary output stages (a) from single dc supply; (b) from symmetrical dual (positive and negative) supplies.

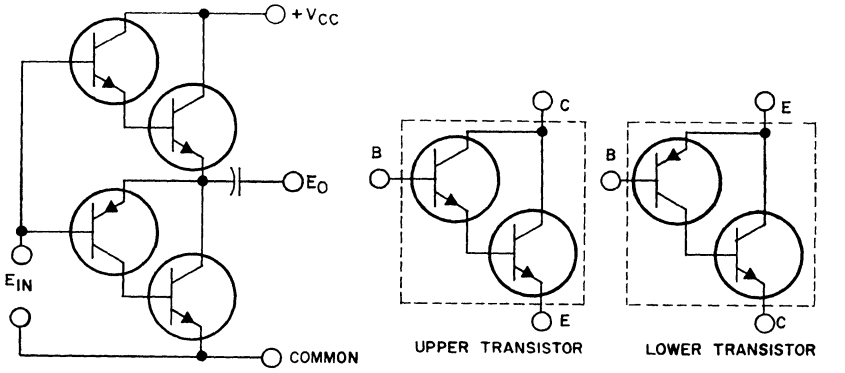


Figure 416. Compound output stage in which output transistors are driven by complementary driver transistors: (a) over-all circuit; (b) upper transistor pair; (c) lower transistor pair.

vantage of a split supply is that it makes possible direct connection to the load and thus improves low-frequency response.

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 be required to supply the drive needed. There are several complex phase-splitting circuits; a few of them are shown in Fig. 417.

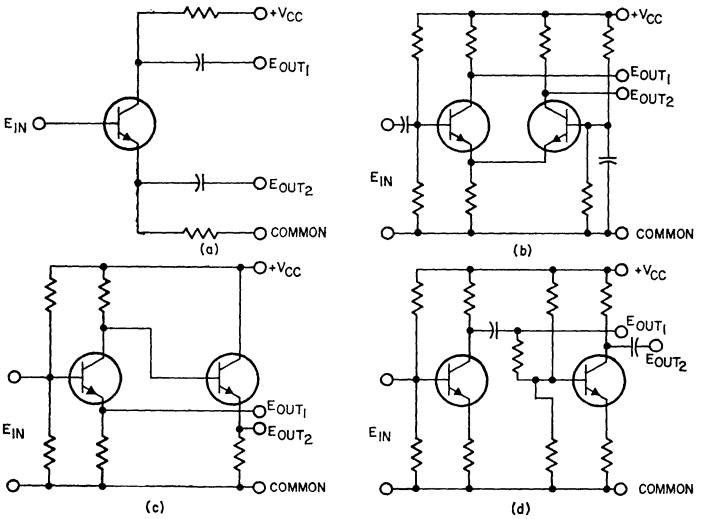


Figure 417. Basic phase-inverter circuits: (a) single stage phase-splitter type (b) two-stage emitter coupled type; (c) two-stage low impedance type; (d) two-stage similar-amplifier type.

Other Design Considerations

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 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 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 trans-

former 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.

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.

Linear single-phase servo amplifiers are usually audio power amplifiers that operate at a single frequency. The servo amplifier is usually designed and biased for class AB or class B operation. Greater distortion can be tolerated than would be acceptable in a high-fidelity audio amplifier. Hum and noise may still be problems if they are near the operating frequency used in the servo application. Servo amplifiers often use transformer-coupled driver and output stages because the operation is generally at a single frequency.

The amplifiers described can be used with any ultrasonic equipment requiring a sine-wave drive. The previous discussion of low-frequency amplifiers is applicable

to ultrasonic equipment except that consideration must be given to the frequency involved; ultrasonic equipment usually operates in the frequency range between 20 kHz and 100 kHz. The oscillator power amplifier already described is best used at low frequencies; however, special attention must be given to the selection of transistors to assure that they will operate efficiently at the frequency chosen.

The power oscillator may also be used as an ultrasonic generator; it supplies sufficient energy for the load without the need for any additional amplification. This amplifier type is particularly useful at lower power levels.

A dc-to-ac inverter may be used as a source of power for an ultrasonic transducer. The design of such circuits is covered in the section on **Power Conversion.**

AUDIO-FREQUENCY 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 sources.

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. 418 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.

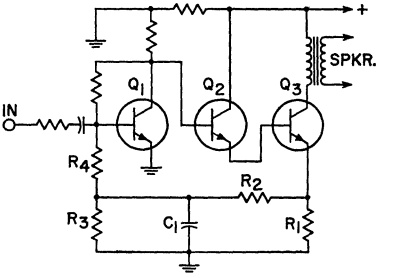


Figure 418. Three-stage, transformer-coupled, class A amplifier.

An output capability of 5 watts with 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. 418 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. 419 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 frequency 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.

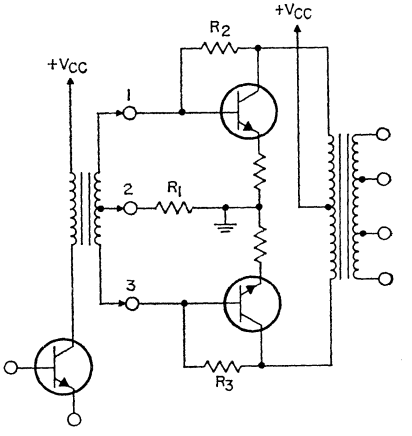


Figure 419. Class AB, push-pull, transformer-coupled audio output stage.

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. 419, 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. 420. 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.

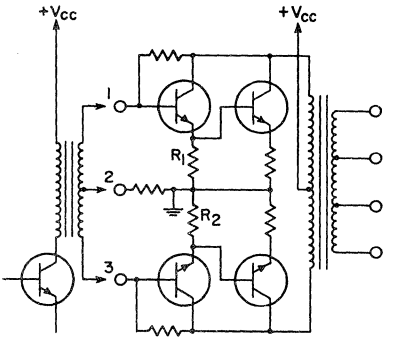


Figure 420. Class AB, push-pull, transformer-coupled audio output stage in which Darlington pairs are used to reduce drive requirements of output transistors.

An alternative method of phase inversion is to use a transistor in a phase-splitter circuit, such as those shown in Fig. 417. 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. 421 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

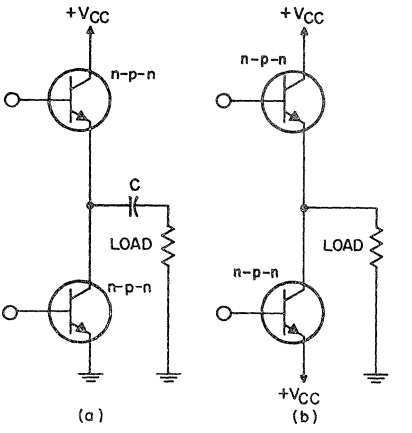


Figure 421. Circuit arrangements for operation of series output circuit from (a) a single dc supply and (b) symmetrical dual supplies.

secondary windings, as shown in Fig. 422. The split secondary windings are required because of

the mode in which each of the series output transistors operates.

If ground were used as the drive reference for both secondary windings of the circuit shown in Fig. 422, 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.

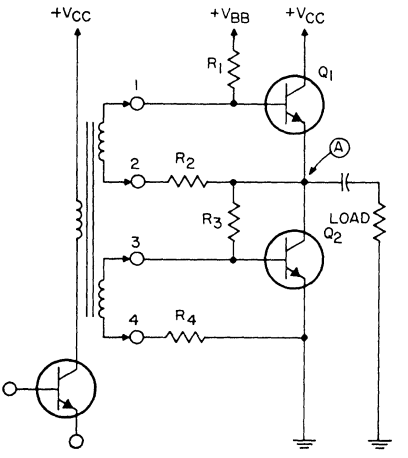


Figure 422. Circuit using a driver transformer that has split secondary windings to provide phase inversion for push-pull operation of a series-output circuit.

The disadvantages of a driver transformer discussed previously also apply to the circuit shown in Fig. 422. 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. 423 shows a circuit in which the transistor phase inverter is used, together with a Darlington output stage to minimize loading

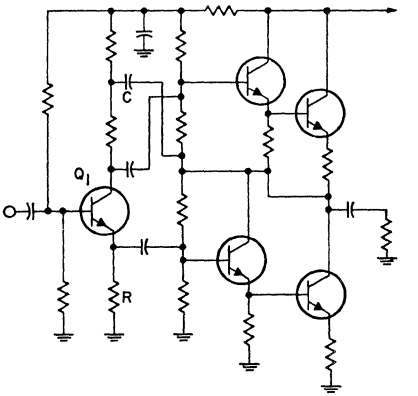


Figure 423. Push-pull series-output amplifier in which driver and output transistors are connected as Darlington pairs and drive-signal phase inversion is provided by phase-splitter stage Q_1.

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.

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. 424 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

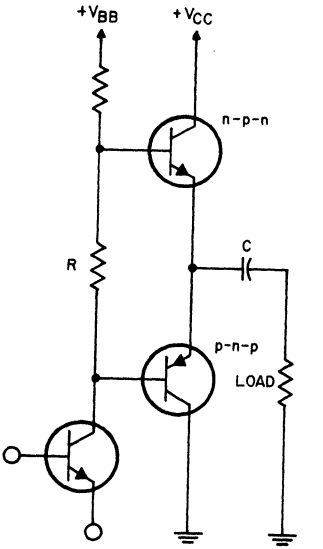


Figure 424. Basic complementary type of series-output circuit.

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. 425, the driver transistors provide the necessary phase inversion. A simple but descriptive way

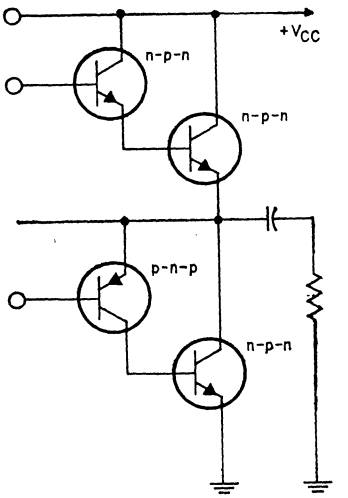


Figure 425. Basic quasi-complementary type of series-output circuit.

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. 426. 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

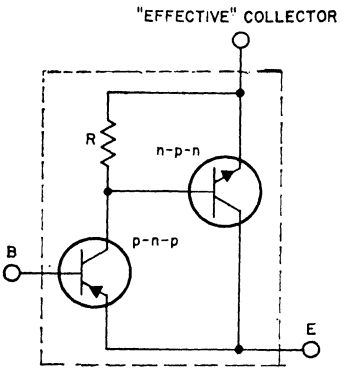


Figure 426. Connection of p-n-p driver transistor to n-p-n output transistor.

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. 426 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. 427, are compared in Fig. 428.

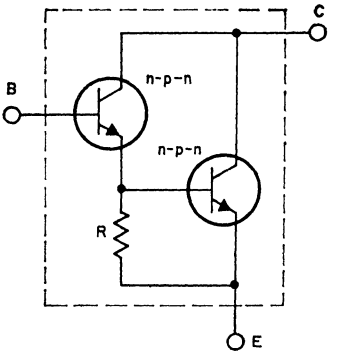


Figure 427. Darlington connection of n-p-n driver transistor to n-p-n output transistor.

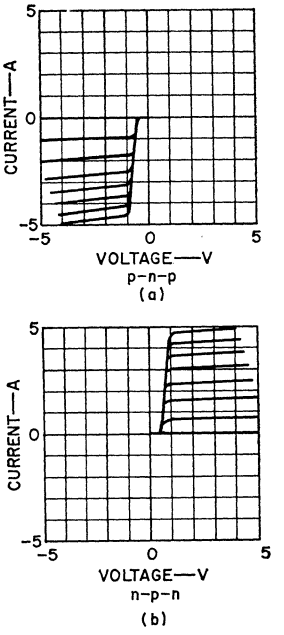


Figure 428. Output characteristics for (a) p-n-p/n-p-n driver-output transistor pair shown in Fig. 426 and for (b) Darlington pair of n-p-n transistors shown in Fig. 427.

The saturation characteristics of the over-all circuit in both cases is 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. 424, 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 as stable as that of the complementary amplifier, but presents no problem for silicon transistors.

A typical quasi-complementary amplifier is shown in Fig. 429. 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

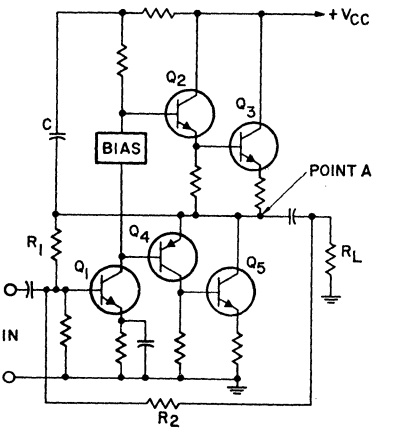


Figure 429. Quasi-complementary audio power amplifier that operates from a single dc supply.

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. 421(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. 429, 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. 430. 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

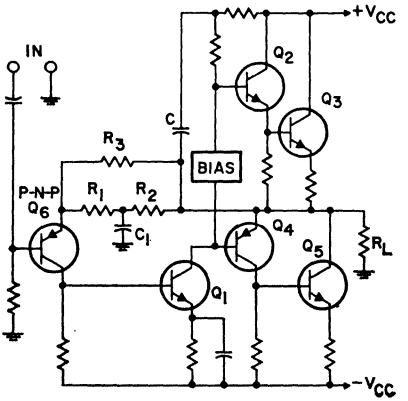


Figure 430. Quasi-complementary audio power amplifier that operates from symmetrical dual dc power supplies. The p-n-p transistor input stage is required to prevent ripple component from driving amplifier.

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:

$$P_o = I(rms) \times E(rms) = (I_p E_p)/2 = (I_p^2 R_L)/2 = E_p^2/2R_L$$

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. 431 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.

Audio-Power-Amplifier Rating Methods—The **Institute of High Fidelity** (IHF) and the **Electronic Industries Association** (EIA) have attempted to standardize

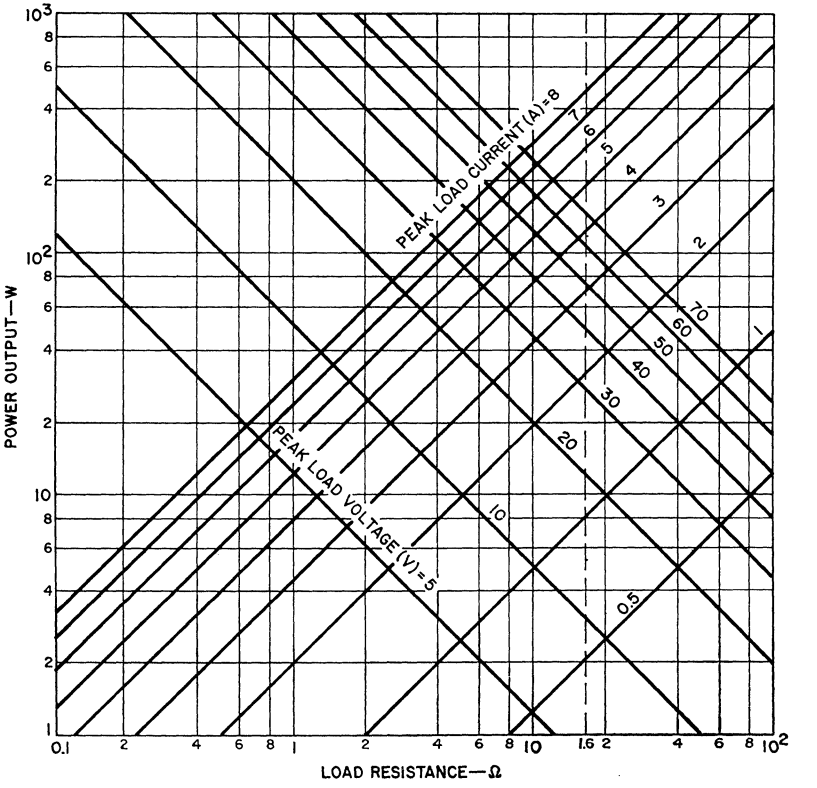


Figure 431. Peak transistor currents and load voltages for various output powers and load resistances.

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 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 METHODS: 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 av-

erage 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. 432 shows a typical regulation curve for a rectifier power supply that has a capacitive input filter. The voltage is a linear

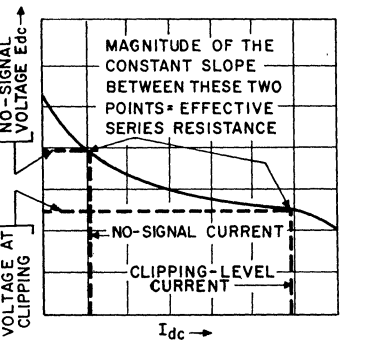


Figure 432. Regulation curve for capacitive rectifier power supply.

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. 433 and 434 show equivalent circuits for capacitive-input

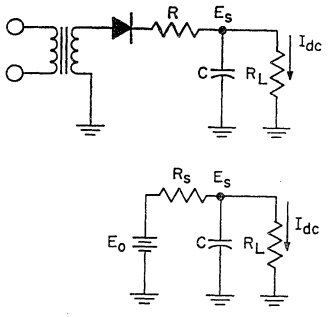


Figure 433. Equivalent circuit for single-ended capacitive-input rectifier supply.

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_o 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_o, as follows:

$$E_s = E_o - R_s I_{dc} \qquad (320)$$

Eq. (320) shows that the supply voltage, E_s, is equal to the no-signal supply voltage, E_o, 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

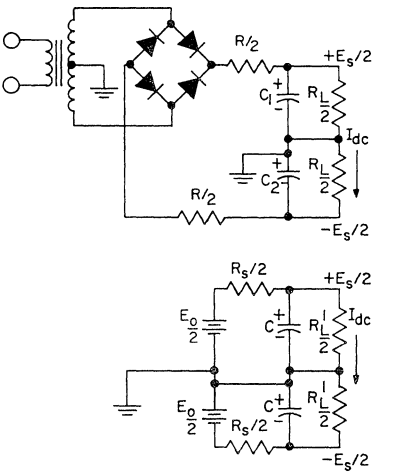


Figure 434. Equivalent circuit for split capacitive-input rectifier supply.

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:

$$I_{dc} = \frac{I_{pk}}{\pi} \qquad (321)$$

The power output is related to the peak output current, as follows:

$$P_o = \frac{I_{pk}^2}{2} R_L \qquad (322)$$

where R_L is the speaker load resistance. Consequently, the supply current is related to the power output by

$$I_{dc} = \left(\frac{2 P_o}{\pi^2 R_L}\right)^{1/2} \qquad (323)$$

Combining Eqs. (320) and (323),

$$E_s = E_o - R_s \left(\frac{2 P_o}{\pi^2 R_L} \right)^{1/2} \quad (324)$$

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 π^2 is 10. Eq. (324) then becomes

$$E_s = E_o - 0.158 R_s (P_o)^{1/2} \quad (325)$$

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. 432. Using the values for voltage and current from Fig. 432, the music power rating, based on the no-signal voltage (44.8) and an 8-ohm load, is given by

$$P_{music} = \frac{E_o^2}{8 R_L} = 31.5 \text{ watts/channel} \quad (326)$$

If no parasitic losses are assumed, therefore, the stereo music power would be 63 watts.

[The factor, 8, in Eq. (326) is derived in converting peak-to-peak volts to rms volts.]

The effective series resistance, according to Fig. 432, is approximately 6 ohms. 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. 435. The

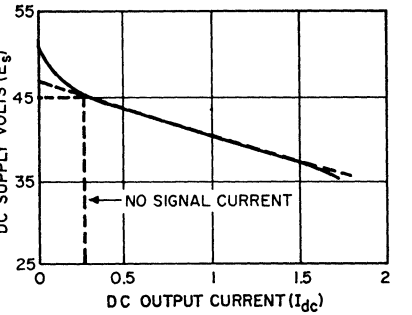


Figure 435. Typical power-supply regulation curve.

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 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. 436. The effect of a nonregulated power supply on transistor dissipation is then examined.

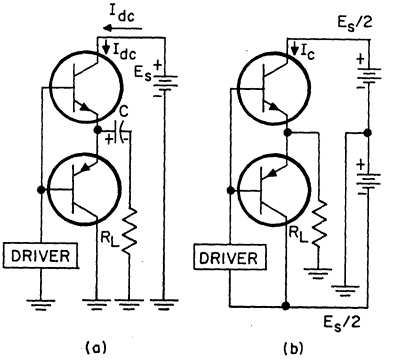


Figure 436. Typical complementary-symmetry circuits.

REGULATED SUPPLY: When the amplifier circuit shown in Fig. 436(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

$$I_{pk}\,(max) = E_s/2\,R_L \qquad (327)$$

Because the supply delivers current on alternate half-cycles, the average supply current I_{dc} is given by

$$I_{dc} = I_{pk}/\pi \qquad (328)$$

The power P_o delivered by the supply can then be expressed as follows:

$$P_s = (I_{pk}\,E_s)/\pi \qquad (329)$$

The power delivered to the load P_o is given by

$$P_o = (I_{pk}^2\,R_L)/2 \qquad (330)$$

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:

$$P_T = (P_s - P_o)/2$$

$$P_T = \frac{I_{pk}\,E_s}{2\pi} - \frac{I_{pk}^2\,R_L}{4} \qquad (331)$$

If Eq. (331) is differentiated and solved for the peak load current I_{pk} at maximum average transistor dissipation, the following expression is obtained:

$$I_{pk} = E_s/(\pi\,R_L) \qquad (332)$$

When this value is substituted in Eq. (331), 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:

$$\frac{P_T\,(max)}{P_o\,(max)} = \frac{2}{\pi 2} \qquad (333)$$

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

$$P_o\,(max\ diss.) = \frac{E_s^2}{2\pi^2\,R_L} \qquad (334)$$

The ratio of the power output at maximum dissipation P_o (max diss.) to maximum power output $P_o(max)$ is then given by

$$\frac{P_o\,(max\ diss.)}{P_o\,(max)} = \frac{4}{\pi^2} \qquad (335)$$

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. (320). If this value for E_s is substituted in Eq. (331), the following result is obtained:

$$P_T = \frac{I_{pk} E_o}{2\pi} - \frac{R_s (I_{pk})^2}{2\pi^2} - \frac{(I_{pk})^2 R_L}{4} \quad (336)$$

The partial derivative of Eq. (336) 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. (336) to determine the maximum transistor dissipation P_T(max), as follows:

$$\frac{dP_T}{dI_{pk}} = \frac{E_o}{2\pi} - I_{pk} \frac{2R_S + \pi^2 R_L}{2\pi^2} \quad (337)$$

$$I_{pk} = \frac{E_o \pi}{2R_S + \pi^2 R_L} \quad (338)$$

$$P_T(\text{max}) = \frac{E_o^2}{8R_S + 4\pi^2 R_L} \quad (339)$$

Clipping begins at the point where the peak collector current I_{pk} is given by

$$I_{pk} = \frac{E_o \pi}{R_S + 2\pi R_L} \quad (340)$$

Power output at clipping can then be expressed as follows:

$$P_o(\text{clipping}) = \frac{E_o^2 \pi^2 R_L}{2(R_S + 2\pi R_L)^2} \quad (341)$$

If $R_s = 0$ is substituted in Eq. (341), the power output may be expressed as follows:

$$P_o = E_o^2 / 8R_L \quad (342)$$

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:

$$\frac{P_T(\text{max})}{P_o(\text{music})} = \left[\frac{\pi^2}{2} + \frac{R_S}{R_L}\right]^{-1} \quad (343)$$

The power output at which maximum average transistor dissipation occurs P_o (max diss) is related to the music power output as follows:

$$\frac{P_o(\text{max diss.})}{P_o(\text{music})} = \left[\frac{\pi^2}{4} + \frac{R_S}{R_L} + \frac{R_S^2}{\pi^2 R_L^2}\right]^{-1} \quad (344)$$

The continuous power output at the clipping level, P_o (clipping) is related to the music power output by the following expression:

$$\frac{P_o(\text{clipping})}{P_o(\text{music})} = \left[1 + \frac{R_S}{\pi R_L} + \frac{R_S^2}{4\pi^2 R_L^2}\right]^{-1} \quad (345)$$

Eq. (343), (344), and (345) are plotted in Fig. 437. 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. 437 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 example, a 100-watt amplifier

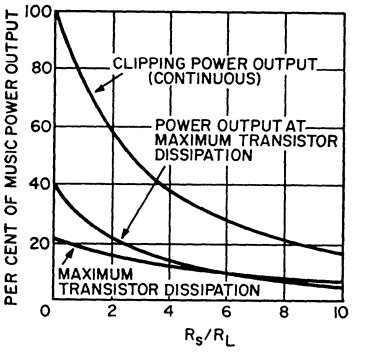


Figure 437. Power output and dissipation as functions of R_S and R_L.

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. 438. 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 initiated 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.

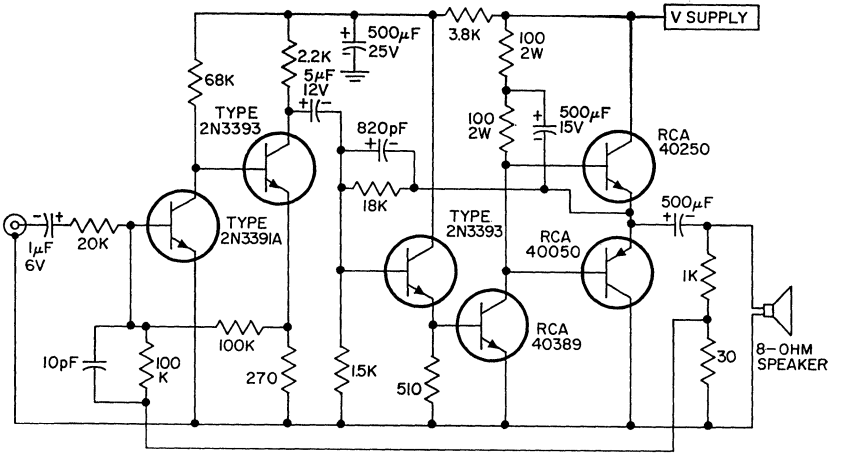


Figure 438. Class B complementary-symmetry power amplifier.

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. 437.

Maximum Ratio of Music Power to Continuous Power Output—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:

$$R_S = \left[\frac{(8R_L)\ P_o\ (\text{music})}{\overline{I}^2}\right]^{1/2} - \frac{E_S\ (\text{min})}{\overline{I}} \qquad (346)$$

where E_s(min) is the minimum voltage required for 100 dB of acoustical power output and I is the current required for 87 dB of acoustical power output, less the idle current. [E_s(min) should, in practice, be increased by peak output-circuit voltage losses.]

Eq. (346) is plotted in Fig. 439. The value of R_s is the absolute maximum value of effective 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.

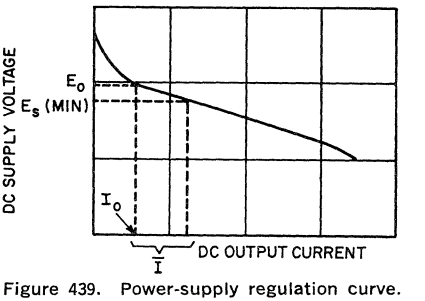


Figure 439. Power-supply regulation curve.

Use of Fig. 440 in conjunction with Fig. 437 shows that very high ratios of music power output to continuous power output may be employed without sacrifice of the subjective ability of 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.

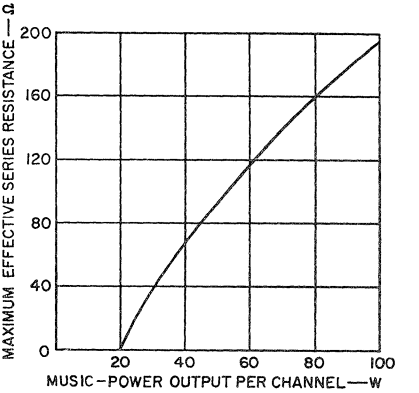


Figure 440. Maximum effective series resistance as a function of music power output.

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. 441 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

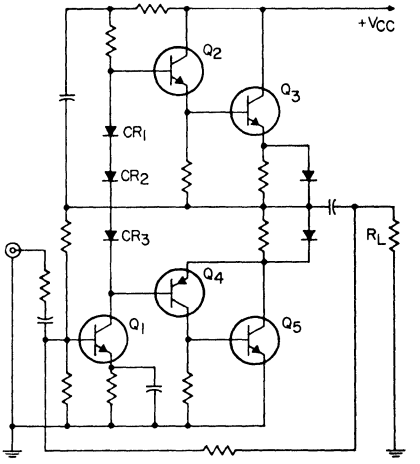


Figure 441. Quasi-complementary amplifier that incorporates two stabilization networks.

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. 429, do not improve the operating-point stability of the over-all circuit.

The circuit shown in Fig. 441 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. 442, which shows the load-line characteristics of a transistor in a class AB push-pull circuit

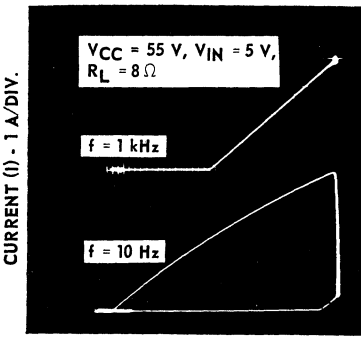


Figure 442. Effect of large signal phase shift on the load-line characteristics of a transistor at low frequencies.

similar to that shown in Fig. 401 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.

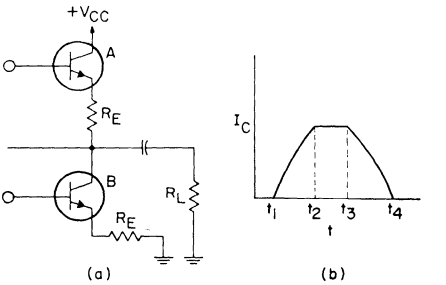


Figure 443. (a) Class B series-output stage, (b) collector-current waveform under overdrive (clipping) conditions.

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. 443(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. 443(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. 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. 443(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. 444 shows the 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-

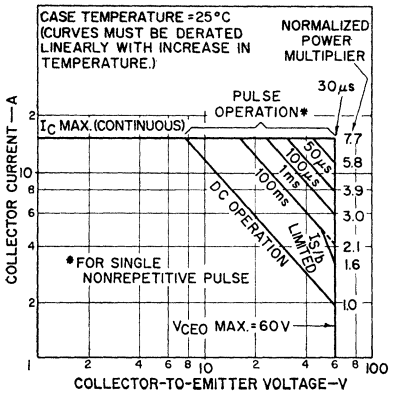


Figure 444. Safe-area-of-operation rating chart for the RCA-2N3055 homotaxial-base transistor.

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. 445. 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 margin of the amplifier. This technique in no way affects the normal operation of the amplifier.

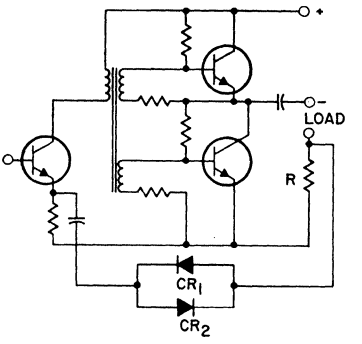


Figure 445. Push-pull power amplifier with short-circuit protection.

A second approach to current limiting is illustrated by the circuit shown in Fig. 446. 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. 447 illustrates a dissipation-limiting technique that provides positive protection under all loading conditions. The limiting action of this circuit is shown in Fig. 448. 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.

Universal-Amplifier Design Approach

A novel approach to the design of high-quality audio power

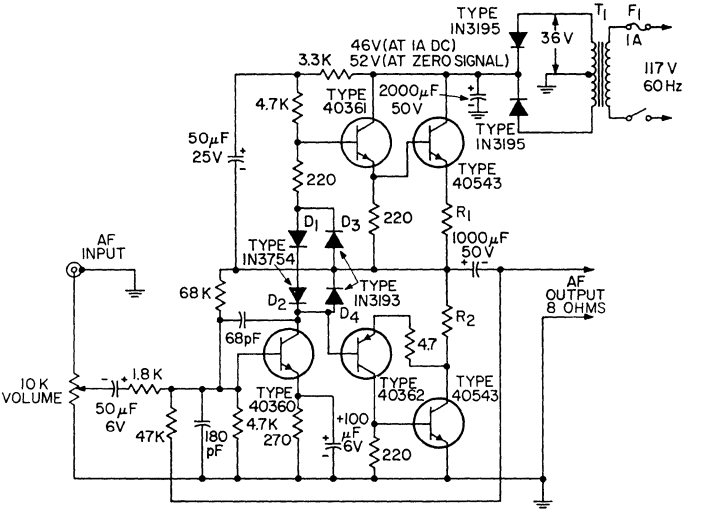


Figure 446. 25-watt (rms) quasi-complementary audio amplifier using current-limiting diodes (D_3 and D_4).

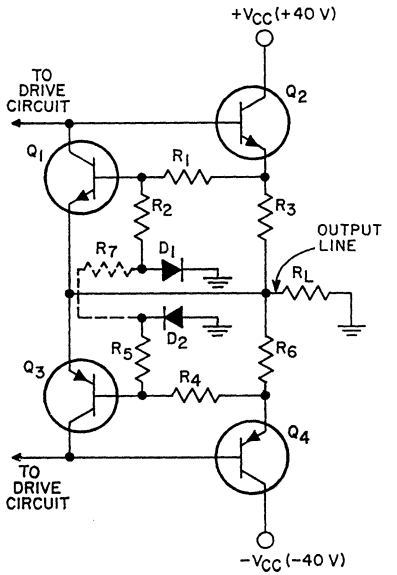


Figure 447. Quasi-complementary audio output in which diode-resistor biasing network is used to prevent complementary transistors Q_1 and Q_2 from being forward-biased by the output voltage swing.

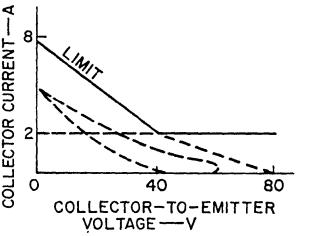


Figure 448. Load lines for the circuit of Fig. 447. Load lines showing effect of the inclusion of high-resistance diode-resistor network in the forward-biasing path of Q_1 are shown dotted.

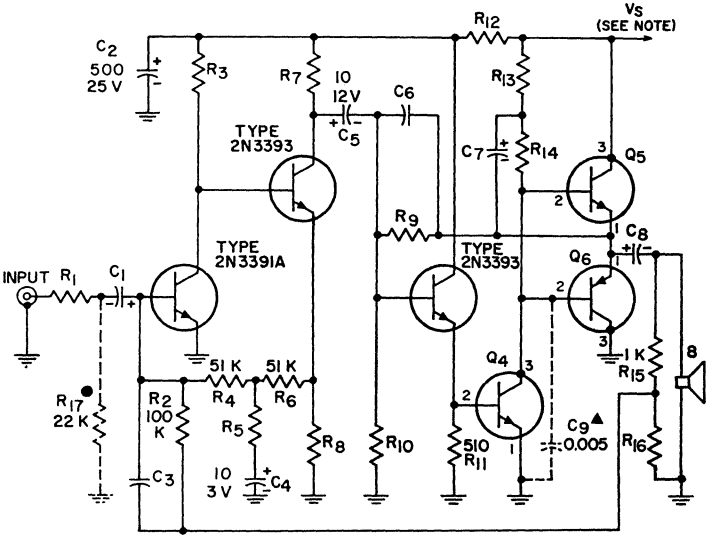
amplifiers is demonstrated by the RCA "Universal Audio-Amplifier Program." In this program, the objective is to devise basic circuit configurations that can be used for a broad family of practical audio amplifiers to cover a wide range of output-power levels. Thus far, the program has provided two basic 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 Complementary-Symmetry Circuit—Fig. 449 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 output-stage transistors, and the supply voltage must be changed to alter the output-power level. Table XXXVII 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. 449.



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 449. 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).

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

Table XXXVII—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	40618	40615	40621	40620	40622	40624	40627
Q_6 (p-n-p output)	40612	40609	40022	40614	40022	40619	40050	40623	40626
Capacitance values:									
C_1 (μF/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 (μF/V)	500/12	500/12	500/12	500/12	500/12	500/12	500/15	500/15	500/15
C_8 (μF/V)	500/25	500/25	500/25	500/25	500/25	500/25	500/50	500/50	500/50
C_9 (μF)	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	—	—	—

of current, as established by the output voltage-divider network. 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, and 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 transistor 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 XXXVIII. Fig. 450 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. 451 shows the basic configuration of a quasi-complementary-symmetry circuit that can provide outstanding performance for the most stringent requirements of

Table XXXVIII—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

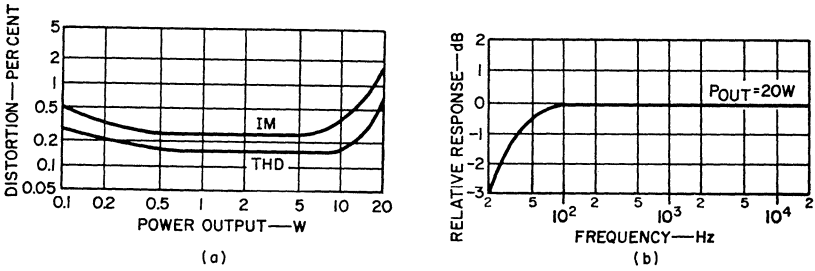
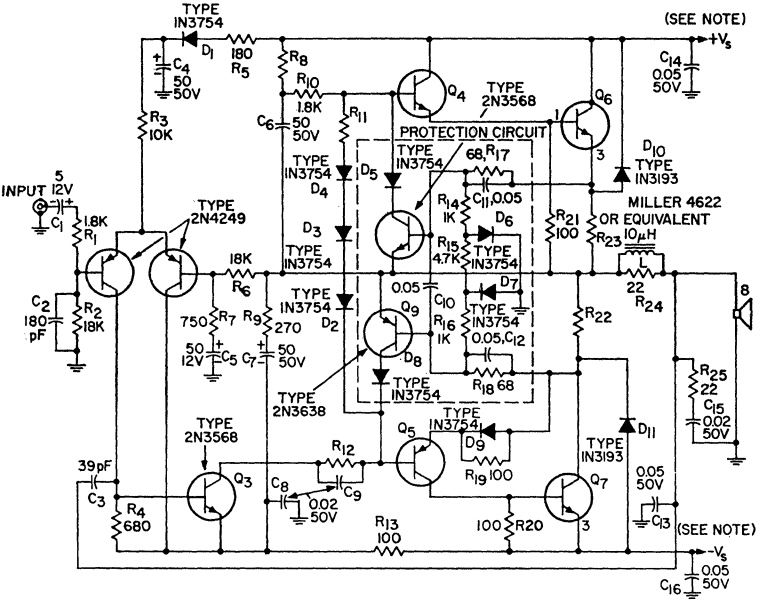


Figure 450. Performance characteristics for the 20-watt complementary-symmetry amplifier: (a) distortion as a function of power output; (b) relative response as a function universal complementary-symmetry design.)



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 451. 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).

manufacturers of high-fidelity amplifiers. This basic circuit can be used for four separate audio amplifier circuits that provide continuous sine-wave power outputs 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 XXXIX lists the other component requirements for the four circuits. The voltage chart 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 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 out-

Table XXXIX—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)	4[illegible]	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

put 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-frequency 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} ensure a transconductance match between the upper and lower Darlington pairs to minimize low-level distortion.

Table XXXX 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

Table XXXX—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

of power output and of output response as a function of frequency for the 60-watt amplifier, shown in Fig. 452, are indicative of the performance capabilities of these amplifiers.

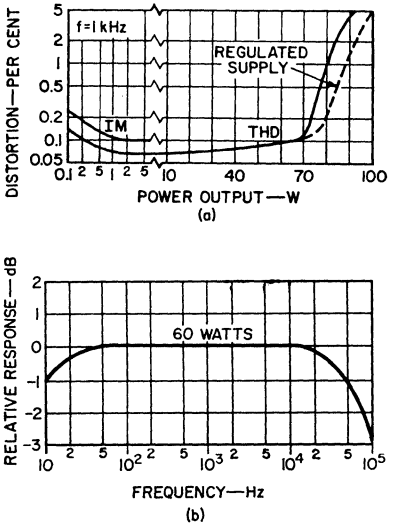


Figure 452. Performance curves for the 70-watt quasi-complementary-symmetry audio amplifier: (a) distortion as a function of power output; (b) relative response as a function of frequency. (These curves are representative of the performance provided by the universal quasi-complementary-symmetry design.)

Bridge-Amplifier Design Approach

Fig. 453 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 supply 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.

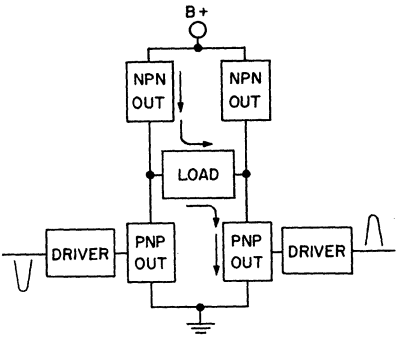


Figure 453. Block diagram of bridge type of audio-amplifier circuit.

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. 454 shows the basic configuration of an experimental breadboard circuit designed to

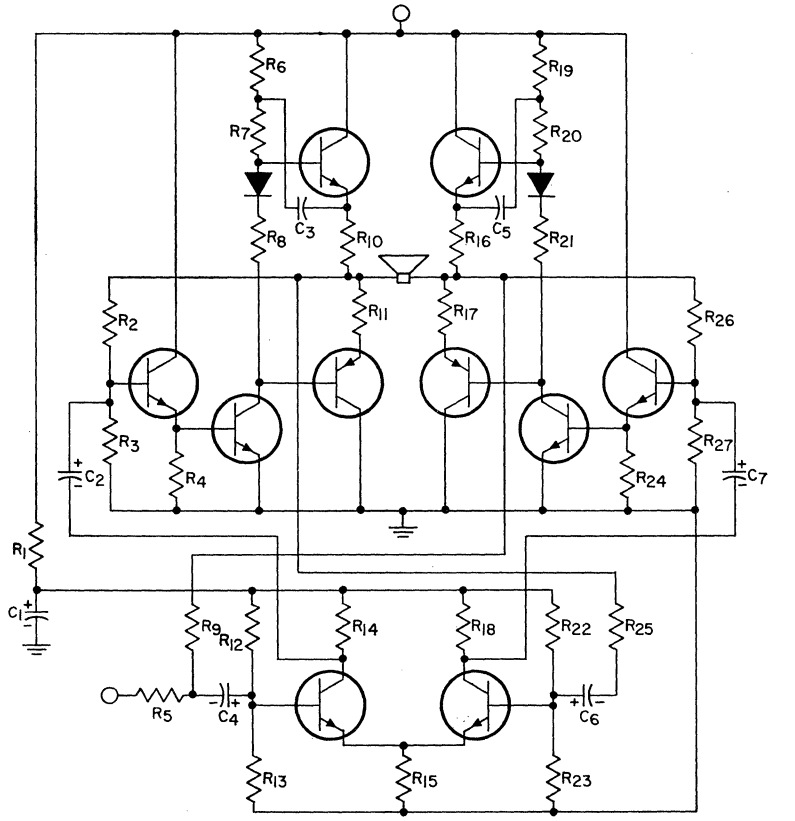


Figure 454. Basic circuit configuration for a bridge type of audio amplifier.

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 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 ensure 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. 455 shows curves of total harmonic distor-

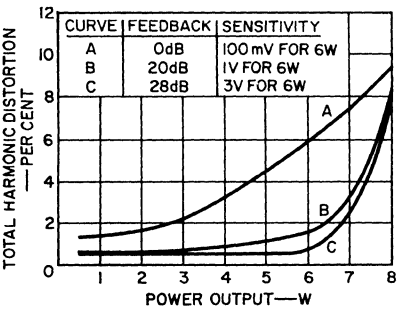


Figure 455. Total harmonic distortion (at 1 kHz) of the bridge audio amplifier as a function of power output for different values of balanced loop feedback. (Distortion performance is comparable to that of a single-ended amplifier that provides one-quarter of the power output for the same dc supply voltage.)

tion as a function of power output for operation of the bridge amplifier with 0 dB, 20 dB, and 28 dB of balanced feedback. Figs. 456 and 457 show total harmonic distortion and relative response as functions of frequency for the bridge amplifier operated with 20 dB of balanced feedback.

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.

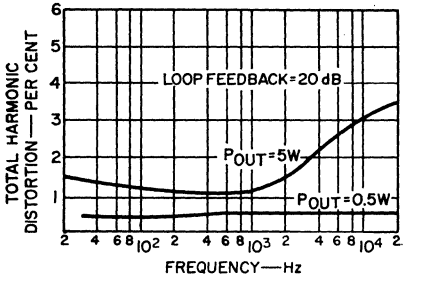


Figure 456. Total harmonic distortion of the bridge audio amplifier as a function of frequency.

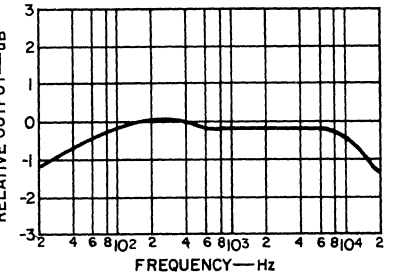


Figure 457. Relative response of the bridge audio amplifier.

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 XXXXI 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.

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 XXXXI—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. 458 shows the strain (change in length per unit length) as a function of magnetic field

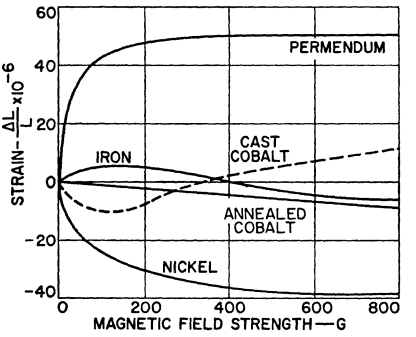


Figure 458. Strain as a function of magnetic field strength for several magnetostrictive materials.

strength for several magnetostrictive materials. It can be seen that nickel gets shorter as the magnetic field is increased, while Permendum gets longer. Fig. 459 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 ΔL peak to peak.

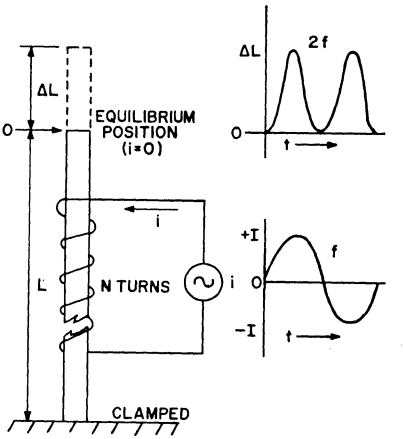


Figure 459. Reaction of a bar of material that has a positive strain coefficient to an alternating magnetic field when no static biasing field is used. Waveforms show change in length of bar (top) and alternating current (bottom) used to produce the magnetic field.

Fig. 460 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 ΔL. Under these condi-

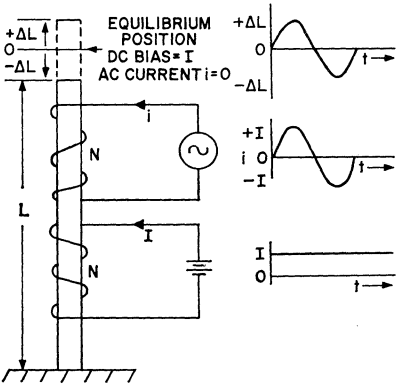


Figure 460. Reaction of bar of material that has a positive strain coefficient to an alternating magnetic field when static biasing is employed. Waveforms show change in length of bar (top), alternating current used to produce alternating field (center), and direct current (bottom) used to produce the bias field.

tions, 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 (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. 461 shows an example of this method.

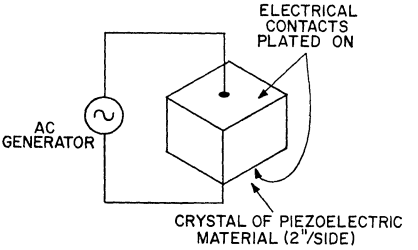


Figure 461. Application of an alternating voltage to a piezoelectric crystal to produce high-frequency sound waves.

In the design of equipment that uses electromechanical transducers, a useful equivalent circuit for the transducer must be available. Fig. 462(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. 462(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-

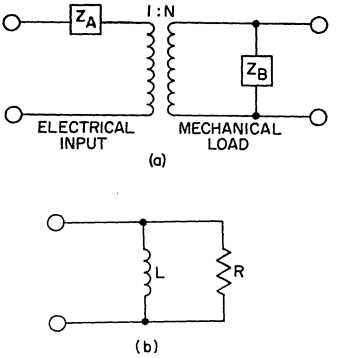


Figure 462. (a) Actual equivalent circuit and (b) simplified approximation of a magnetostrictive transducer.

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. 463 shows a typical circuit.

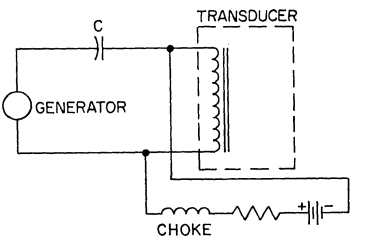


Figure 463. Circuit showing application of electrical power to a magnetostrictive transducer.

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 inductive reactance of the transducer is cancelled.

Fig. 464(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. 464(b) shows the approximate equivalent circuit used to represent a piezoelectric transducer for the purpose of

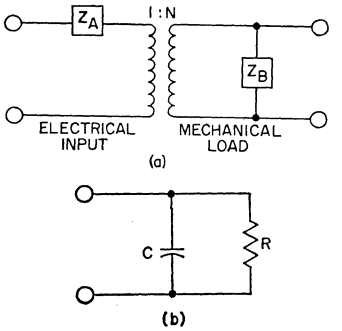


Figure 464. (a) Actual equivalent circuit and (b) simplified approximation of a piezoelectric crystal.

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.

The majority of ultrasonic applications employ a continuously oscillating power source. In fact, the only application listed in Table XXXXI 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 XXXXI 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 manual.

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.

Fig. 465 shows one method of obtaining a voltage sine wave

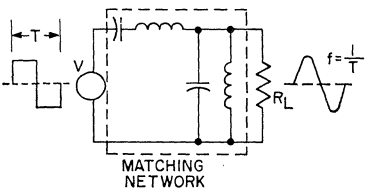


Figure 465. Use of a transducer and resonant matching network to convert a square-wave input to a sinusoidal output. Reactive component of transducer is used as the shunt inductor or capacitor of the matching network depending upon whether a magnetostrictive or electrostrictive type of transducer is used.

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. 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. 466. 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. Theoretically, 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.**

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. 467. With this arrangement, the operating frequency is always the resonant frequency of the load.

Fig. 468 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

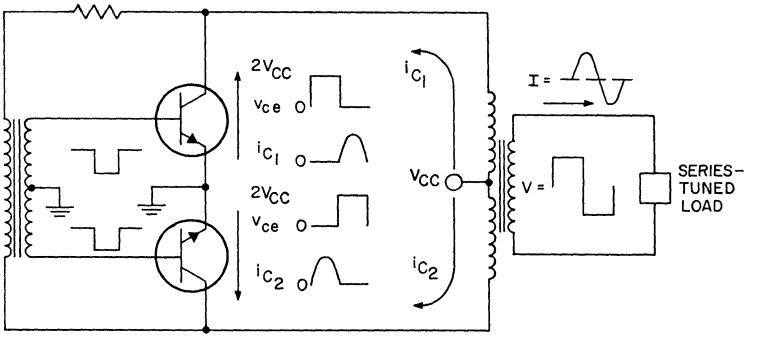


Figure 466. Use of a push-pull switching inverter to drive a transducer that forms part of a series-tuned load circuit.

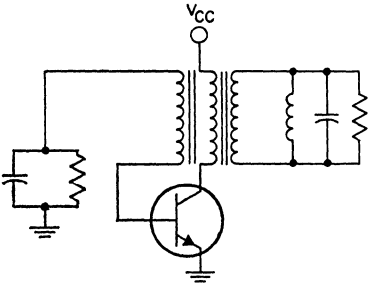


Figure 467. Class C oscillator that operates into a tuned load circuit.

equal to the reciprocal of the resonant frequency f_r of the load. Therefore, if f_r changes, there is

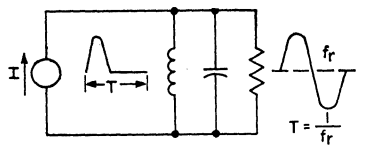


Figure 468. Simplified equivalent circuit for the class C oscillator shown in Fig. 467.

a corresponding change in T. Fig. 469 shows the collector voltage and collector current for the class C oscillator.

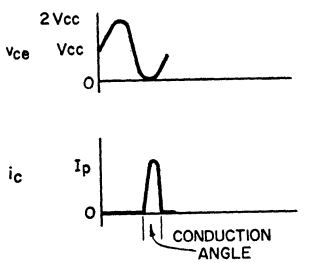


Figure 469. Collector voltage and current waveforms for the Class C oscillator shown in Fig. 467.

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. 469, 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. 470. This condition occurs when the

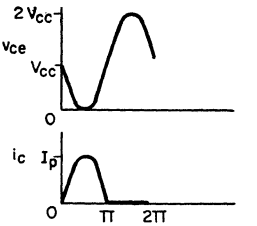


Figure 470. Collector voltage and current waveforms for an oscillator circuit that has a conduction angle of 180 degrees.

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 η can be calculated for a typical transistor operated in a circuit of this type. 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 conduction angle $\theta = \pi$ (maximum power output), the quantities P_s, P_L, P_d, and η 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 η 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. 471 shows the collector current and voltage waveforms for this condition. If 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:

$$P_s = \frac{1}{2\pi}\int_{\pi/6}^{5\pi/6} V_{CC}\, I_p \sin\frac{3}{2}\theta\, d\theta$$

$$= \left[\frac{V_{CC}\, I_p}{2\pi}\left(-\frac{2}{3}\cos\frac{3}{2}\theta\right)\right]_{\pi/6}^{5\pi/6}$$

$$= \frac{-V_{CC}\, I_p}{3\pi}\left[\cos\frac{3}{2}\left(\frac{5\pi}{6}\right) - \cos\frac{3}{2}\left(\frac{\pi}{6}\right)\right]$$

$$= \frac{V_{CC}\, I_p}{3\pi}(2)\left(-\frac{1}{\sqrt{2}}\right)$$

$$= 0.15\, V_{CC}\, I_p = 150 \text{ watts}$$

$$P_L = \frac{1}{2\pi}\int_{\pi/6}^{5\pi/6} V_{CC}\sin\theta\, I_p \sin\frac{3}{2}\theta\, d\theta$$

$$= \frac{V_{CC}\, I_p}{2\pi}\int_{\pi/6}^{5\pi/6}\sin\theta\sin\frac{3}{2}\theta\, d\theta$$

$$= \frac{V_{CC}\, I_p}{2\pi}\left[\frac{\sin\left(\frac{3}{2}-1\right)}{2\left(\frac{3}{2}-1\right)} - \frac{\sin\left(\frac{3}{2}+1\right)}{2\left(\frac{3}{2}+1\right)}\right]_{\pi/6}^{5\pi/6} = \frac{V_{CC}\, I_p}{2\pi}\left[\sin\frac{1}{2}\theta - \frac{\sin\frac{5}{2}\theta}{5}\right]_{9/\mu}^{5\pi/6}$$

$$= \frac{V_{CC}\, I_p}{2\pi}(0.966 - 0.05 - 0.26 + 0.193)$$

$$= \frac{0.85}{2\pi} V_{CC}\, I_p = 0.135\, V_{CC} I_p = 35 \text{ watts}$$

$P_d = P_s - P_L = 0.015\, V_{CC}\, I_p = 15$ watts

$\eta = P_L/P_s = 90$ 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.

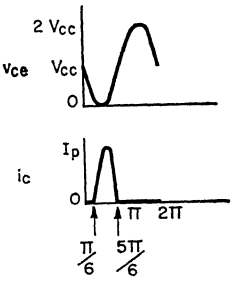


Figure 471. Collector voltage and current waveforms for an oscillator circuit that has a conduction angle of 120 degrees.

Example No. 4: The design of a practical class C oscillator which has a conduction angle θ of 120° and an over-all circuit efficiency η 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 induc-

tance 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:

$$2\pi f = 1/(LC)^{\frac{1}{2}}$$
$$C = 0.01 \text{ microfarad}$$

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 $TR_{J\text{-}C} \leq 7°/$ 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 **High-Frequency Power Amplifiers.** Fig. 472 shows the schematic diagram of the completed circuit, and Fig. 473 shows the circuit waveforms.

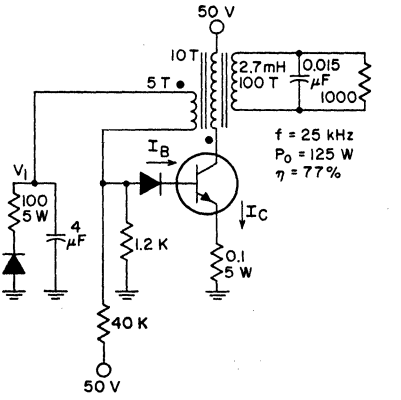


Figure 472. 125-watt, 25-kHz, class C oscillator.

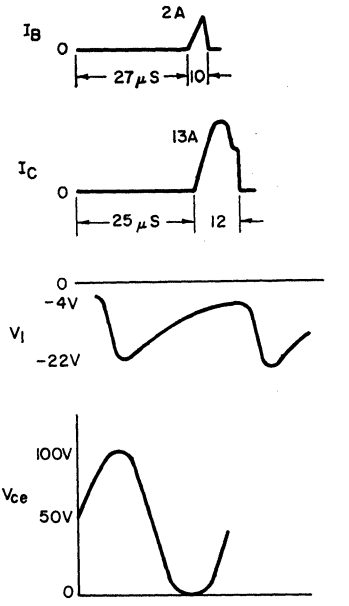


Figure 473. Current and voltage waveforms for the class C oscillator shown in Fig. 472.

SERVO AMPLIFIERS

A servomechanism system is a feedback control system in which the difference between a reference input and some function of a controlled variable is used to supply an error signal to the control elements. The error-signal amplifier is used to drive the control element to reduce the difference between the two functions to zero. The amplifiers used for such functions may be operated in either linear or switching service depending upon the type of control system desired.

Linear Servo Amplifiers

The design of a linear electronic servo amplifier depends on the nature of the input error signal and the output requirements. Selection of transistors and detailed circuit design are accomplished by normal design procedures as discussed in the section on **Audio-Frequency Power Amplifiers.** The types of amplifiers that are used may be classified according to input and output requirements as follows:

(1) dc-input, dc-output
(2) dc-input, ac-output
(3) ac-input, dc-output
(4) ac-input, ac-output

Special circuit design is usually necessary in all classes to control phase shift over the range of frequencies involved. Normally, the ac signal is of constant frequency (60 or 400 Hz); the design problems, therefore, are not too difficult.

Servo amplifiers in classes (2), (3), and (4) are essentially normal amplifiers except that class (2) types may include a "chopper" and in class (3) types the output stage includes rectification and, possibly, filtering.

In class (1), the usual stability and drift problems common to dc amplifiers must be solved. If necessary, a direct-coupled amplifier may be used; when possible, however, the more desirable "chopper-stabilized" dc amplifiers should be used. A chopper-stabilized dc amplifier consists of two switches, sometimes built into the same package and driven synchronously. The input switch chops the dc signal and thus converts it into a series of pulses that can be amplified by a normal ac amplifier. The second switch is connected to the amplifier output, and synchronously rectifies the ac output signal. With this technique, the dc drift of the amplifier is not a factor in the dc output signal.

For the case of the direct-coupled amplifier, the output stages follow conventional class A or class B designs except for special techniques used to assure thermal stability and drift-free operation when the amplifier is operated over the entire expected temperature range.

For applications in which the circuit arrangement must drive an amplifier with a dc input and produce a reversible-phase ac output (class 2), special techniques must be employed in the output stage. This type of servo amplifier, which is similar to the balanced modulator with suppressed carrier, is the most common, and is found in many X-Y plotters.

Fig. 474 shows a simple control system with direct feedback. In this circuit the reference input and the feedback are passed to a comparator in which an error signal is produced and amplified to produce the controlled variable.

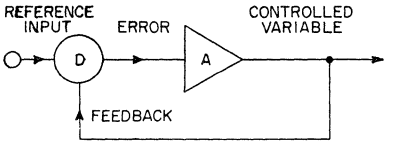


Figure 474. Block diagram of a simple control system that uses direct feedback.

Fig. 475 shows a phase-sensitive amplifier which takes a dc input and produces an ac output. The collectors of both transistors are fed from the ac supply, which must be in phase with the servo-system supply so that a phase reference is provided. Both collectors go positive at the same half-cycle, and when there is no input both carry the same amount of current. Under these conditions, the output transformer produces zero voltage. If terminal A becomes more positive, Q_1 conducts more heavily than Q_2; thus, there is a net field, and a resultant net voltage, coupled through the transformer windings. If terminal B becomes positive with respect to terminal A, the phase of the ac output voltage reverses. This type of amplifier can easily be adapted to the case of ac input to ac output with the addition of an input transformer. An application of an amplifier similar to that of Fig. 412 has been previously identified as an X-Y plotter. In this application, a servo motor is supplied with a voltage from the servo supply system to one winding; the output voltage of this amplifier is applied to another winding. This motor drives the mechanical pen in one of the axes. A position voltage is obtained from a potentiometer setup connected to the mechanical assembly. This voltage is compared to the input voltage of the instrument, and the resulting error voltage is applied to the amplifier. The servo motor drives the assembly in the direction that tends to cancel the error voltage. A similar arrangement is used for the other axis.

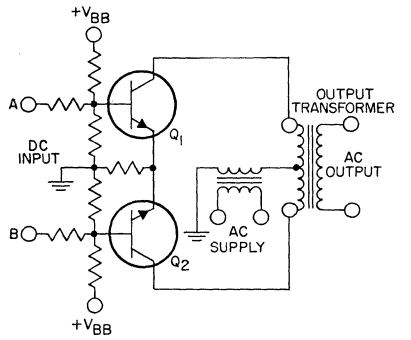


Figure 475. Phase-sensitive linear servo amplifier that develops an ac output for a dc input.

The basic principle of feedback control tends to produce accurate performance because the control system endeavors continually to correct for any error. This corrective action, however, can give rise to a condition of unstable action when the control elements exhibit large amounts of amplification and phase shift. An unstable system may produce large sustained oscillation or erratic control, either of which makes the entire system useless.

In an effort to increase the accuracy of a system, the usual approach is to increase the gain of the servo amplifier. This approach always reduces the stability of the system because undesirable effects of the amplifier become more pronounced. Obviously, the requirements for stability are incompatible with those of accuracy; there is always a trade-off between these factors.

The requirements of the servo amplifier are easily described, but are usually difficult to obtain. As discussed in the section on **Audio-Frequency Power Amplifiers,** negative feedback is usually employed to maintain the over-all gain of the amplifier. In an audio system, if the gain varies with aging components, or if the output dc drifts with changes in ambient temperature, nothing disastrous occurs to the over-all operation of the system. In a servo system, however, steps must be taken to assure a constant gain under all operating conditions. Therefore, large amounts of feedback are employed in applications of this nature. Because of the large variation of parameters that can be expected from transistors, designs tend to be ultra-conservative, with more devices employed to provide a given amount of gain than in a conventional audio amplifier. In practically every stage, some form of temperature compensation is employed to maintain the operating point and the gain of the stage. It is not uncommon to employ 60 dB of over-all feedback in such an amplifier. In such applications, steps must be taken to assure that the amplifier itself is stable. Careful analysis using the Nyquist stability criterion is the only way to assure not only good amplifier performance but good over-all system stability. The Nyquist stability criterion places on a firm mathematical basis the physical fact that instability results when the feedback signal is equal in magnitude and in phase with the actuating signal. Thus, this criterion indicates the necessary conditions for stability in terms of the ratio between the input signal and the feedback signal.

For applications in which a high-power dc amplifier is required, a circuit similar to the 70-watt quasi-complementary-symmetry amplifier shown in Fig. 116 or to the universal quasi-complementary amplifier shown in Fig. 451 may be employed. The input circuit must be modified to provide a dc capability.

Switching-Mode Servo Control

Switching-mode servo controls afford an efficient means for amplification of directional information. As an alternative to the use of cascaded linear stages to drive a class B push-pull output stage, this switching mode of control allows the active elements of the amplifier to operate in either saturation or cutoff. Because a relatively small length of time is spent in the active region of the devices, where power dissipation is high, the average power dissipation is lower. The efficiency of the over-all system, therefore, is higher. Switching servos are used in stable platforms for guidance and navigational systems, control of memory access devices in computer and data-processing systems, and other applications in which efficiency is a prime factor. Fig. 476 shows the circuit diagram of a pulse-width-modulated output stage for a servo-motor-drive system.

General Circuit Description—Fig. 477 shows the block diagram of a pulse-duration-modulated switching-servo control system. The directional error signal (commonly a 400-Hz sine wave) is used to modulate the pulse width of a square-wave pulse having a con-

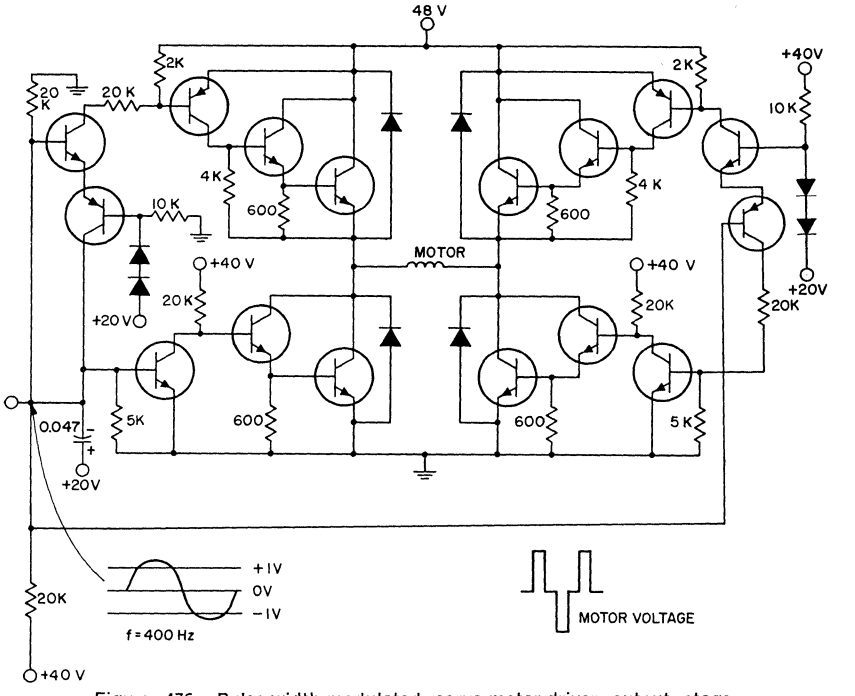


Figure 476. Pulse-width-modulated servo-motor-driver output stage.

stant repetition rate. The waveforms at the input and output of the pulse-duration modulator are shown in Fig. 478.

The error signal has a frequency f_0, and the ramp-generator waveform has a fundamental frequency f_1 which is much greater than f_0. The resulting pulse-duration-modulated (PDM) signal has three components. One is a dc or average component, the second is the error component having a frequency spectrum equal to the error signal (f_0), and the third is a summation of components of the upper harmonics of f_0.

The pulse is amplified in the switching driver and power switch amplifier. It is then sent through a band-pass filter that has a cut-off frequency equal to that of the error signal. The band-pass filter appears as a large impedance to the collectors of the power switching transistors at harmonics of

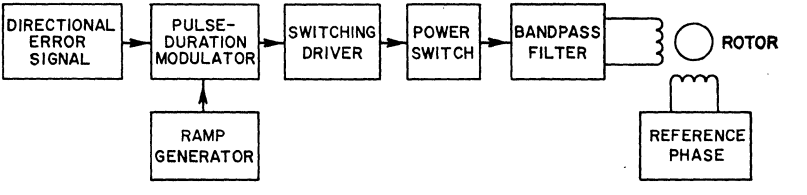


Figure 477. Block diagram of a pulse-duration-modulated switching-servo control system.

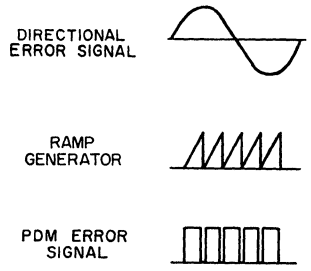


Figure 478. Input and output waveforms for the pulse-duration-modulated servo control system.

f_o. Thus, only the fundamental component of collector current flows in the servo-control winding. The original sinusoidal error signal is then almost completely recovered. The quality of the recovered signal is an inverse function of harmonics introduced into the system by the switching characteristics of the amplifiers and a direct function of the sharpness of cutoff of the band-pass filter.

The error-phase flux and the recovered-phase flux form a resulting control flux that directs the servo rotor.

Practical Circuit—Fig. 479 is the circuit diagram of a switching-servo-control amplifier that uses a pulse-width modulator. The pulse width is modulated by adjustment of the trigger level of the OR gate. A constant-amplitude ramp function appears at point (a). The linearity of the ramp is enhanced by the constant-current source which charges capacitor C_1.

Whenever the voltage at point (a) is greater than the threshold voltage V_T, Q_1 turns on. At this time, Q_2 turns off because its base-emitter junction is reverse-biased. The threshold trigger voltage is continuously varying as a function of the modulating input voltage at the input of the transformer.

When the input signal is at its greatest value, the threshold is greatest. As a result, the OR gate switches later in the ramp period. The pulse duration then varies with the instantaneous magnitude of the modulating signal. Fig. 480 shows the waveforms and the tim-

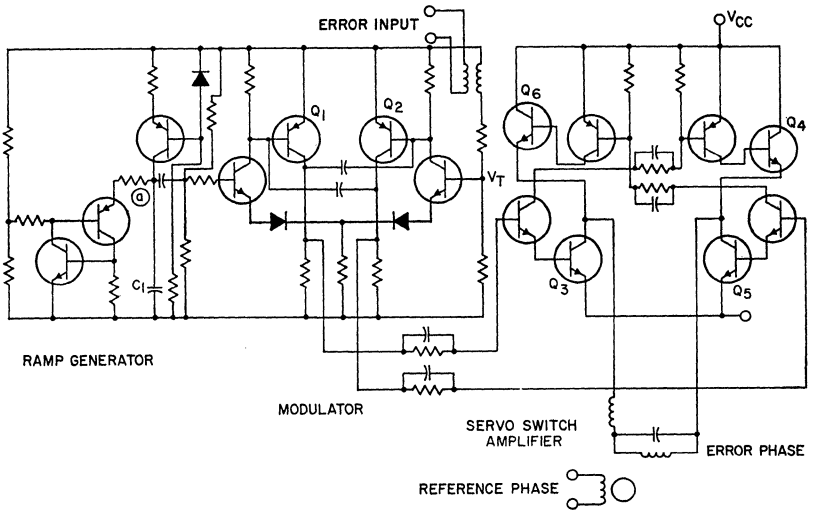


Figure 479. Switching-servo control amplifier with a pulse-width modulator.

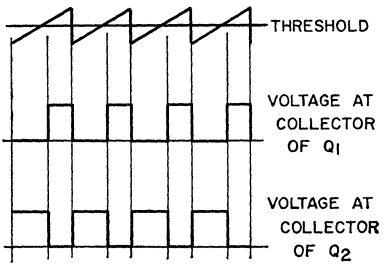


Figure 480. Waveforms showing the threshold switching control for the control amplifier shown in Fig. 479.

ing relationships between the threshold voltage and the collector voltages of Q_1 and Q_2.

The pulses at the collectors of Q_1 and Q_2 provide an alternating current through the load which consists of the error field winding and the filter. This filter should be designed as a low-pass filter with a cutoff at 400 Hz. The flux in this winding then becomes a function of the fundamental frequency of the error signal only.

Efficiency Consideration—Fig. 481 shows the approximate voltage and current switching waveforms of the power output transistors of the switching servo amplifier shown in Fig. 479. It can be shown that the switching configuration is much more efficient than a class B push-pull system. The latter system has a maximum ideal efficiency of 78 per cent. An approximate expression for ideal efficiency for the switching servo amplifier can be derived if the following simplifying assumptions are made:

1. A perfectly loss-free filter with a cutoff frequency of 400 Hz is used.
2. The current waveform through the load is purely sinusoidal, and a perfect reproduction of the input signal.
3. The current has a peak value of $V_{CC}/R_L = I_p$, where R_L represents the dissipation of real power in the servo.
4. The current and V_{CE} switching waveforms are linear, as shown in Fig. 481.
5. The leakage current I_{CO} is negligible.

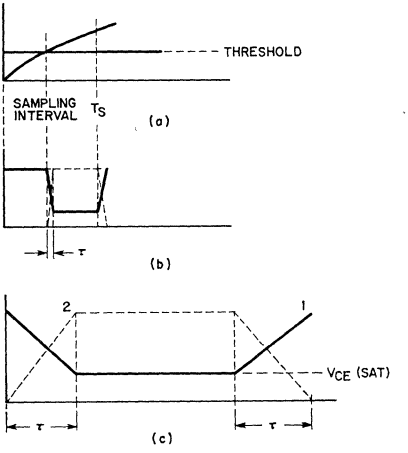


Figure 481. Approximate switching waveforms of the output transistors in the control amplifier shows in Fig. 479.

First, an expression for power dissipated in transistor pair Q_3-Q_4 or Q_5-Q_6 during the turn-off (or turn-on) is defined.

$$\text{Total } V_{CE}(t) = 2\,C_{CE}(sat) + [V_{CC} - 2\,V_{CE}(sat)]\,(t/\tau) \quad (347)$$

$$I_C(t) = I_p - I_p\,(t/\tau) \quad (348)$$

The average power dissipated can be approximated as follows:

$$P_d > (1/\tau) \int_0^{\tau} [2\ V_{CE}(sat) + \{V_{CE} - 2\ V_{CE}(sat)\}\ (t/\tau)]\ [I_p - I_p\ (t/\tau)]\ dt$$
$$= (I_p/6)\ [4\ V_{CE}(sat) + V_{CC}]\ (\tau/T) = P_s \qquad (349)$$

During each sampling interval, both transistor pairs are switching; therefore, $2P_S$ is dissipated. During this interval, switching takes place twice; therefore, a total of $4P_S$ is dissipated during switching. If there are N sampling intervals in an error-signal cycle that lasts for T seconds, then $4NP_S$ is dissipated where $T = NT_s$.

Next, the approximate power dissipated during all time other than the switching time is calculated. The current is sinusoidal and throughout the cycle one set of transistors is on.

$$P_{avg} \cong (1/T) \int_0^{T} 2\ V_{CE}(sat)\ I_p \sin(2\pi t/T)\ dt = [2\ V_{CE}(sat)\ I_p/\pi]$$

Then, the power dissipated in the devices during this time is less than a value P_1, expressed by the following equation:

$$4\ N\ P_s + [2\ V_{CE}(sat)\ I_p/\pi] = P_1$$

The power supplied by the V_{CC} source is given by

$$(2\ V_{CC}\ I_p)/\pi = P_{in}$$

Then, the efficiency η is approximately equal to $(P_{in} - P_1)/P_{in}$, that is,

$$\eta = 1 - (P_1/P_{in}) = 1 - (\pi P_1/2\ V_{CC} I_p)$$

The following typical values are used in an example to show how the ideal efficiency is calculated:

$N = 20$
$V_{CE}(sat) = 2.5$ volts
$I_p = 2$ amperes
$V_{CC} = 200$ volts
$\tau = 2 \times 10^{-6}$ second
$t = 1/400 = 2.5 \times 10^{-3}$ second

With these values, the following results are obtained:

$$P_s = (I_p/6)\ [4\ V_{CE}(sat) + V_{CC}]\ (\tau/t) = (2/6)\ [4 \times 2.5 + 200]\ [(2 \times 10^{-6})/(2.5 \times 10^{-3})] = 0.56$$

$$P_1 = 0.056 + [2\ V_{CE}(sat)\ I_p/\pi] = [(2 \times 2.5 \times 2)/3.14] + 0.056 = 3.25$$

$$P_{in} = (2\ V_{CC} I_p)/\pi = (2 \times 200 \times 2)/3.14 = 255$$

$$\eta = (P_{in} - P_1)/P_{in}) = 1 - (3.25/255) = 1 - 0.012 = 99 \text{ per cent}$$

Bibliography

Aronson, H. L., "Applying Power Transistors to Control", *Control Engineering*, October 1956.

Beaufoy, R., and Sparkes, J. J., "The Junction Transistor as a Charge Control Device", *ATE Journal*, Vol. B, pp. 310–327, October, 1957.

Birman, P., "An Analysis of Cooling Methods for Reliable Operation of Components", *Electronics*, June 22, 1962.

Breece, H. T., "High Speed Inverters Using Silicon Power Transistors", RCA Application Note, SMA–37.

Brown, Morgan, and Stephenson, "The Design of Transistor Push-Pull DC Converters", *Electronic Engineering*, October, 1959.

Carley, D. R., P. L. McGeough, and J. F. O'Brien; "The Overlay Transistor, Part 1: New Geometry Boosts Power", *Electronics*, August 23, 1965.

Caulton, M., H. Sobol, and R. Ernst; "Generation of Microwave Power by Parametric Frequency Multiplication in a Single Transistor", *RCA Review*, June, 1965.

Chang, Z. F., and Turner, C. R., "Characterization of Second Breakdown in Silicon Power Transistors", RCA Application Note, SMA-21.

Chipp, R. D., "Community Antenna Television Systems", *XEEE Spectrum*, July, 1966.

Dewitt and Rossoff, *Transistor Electronics*, McGraw-Hill, New York, 1957.

Donahue, D. J., and A. Jacoby, "Part II: Putting the Overlay to Work at High Frequencies", *Electronics*, August 23, 1965.

Ebers, J. J., and J. L. Moll, "Large-Signal Behavior of Junction Transistors", *Proc. IRE*, Vol. 42, Dec. 1954.

Eimbinder, J., and C. R. Turner, "Determining Permissible Dissipation for Silicon Power Transistors", *Electronic Design*, August 16, 1961.

Elementary Circuit Properties of Transistor, SEEC Vol. 3, J. Wiley & Sons, 1964.

Frederick, J. R., *Ultrasonic Engineering*, J. Wiley & Sons, 1965.

Gartner, W., *Transistors: Principles, Design and Applications*, D. Van Nostrand Co., 1960.

Gold, R. D., and Sondermeyer, J. C., "Designing Silicon Transistor High-Fidelity Amplifiers", *Electronics World*, Sept., Oct., and Nov., 1966.

Goldman, W. E., "9 Ways to Improve Heat Sink Performance", *Electronic Products*, Oct. 1966.

Greiner, R. A., *Semiconductor Devices and Applications*, McGraw-Hill Book Company Inc., New York, 1961. Pages 274-299.

Grinich and Noyce, "Switching Time Calculations for Diffused Base Transistors", *IRE WESCON Convention Record*, Part 3., Electron Devices, August, 1958.

Gutzwiller, F. W., and T. D. Sylvan, "Power Semiconductor Ratings Under Transient and Intermittent Loads", *Communication and Electronics*, Jan. 1961.

Hartz, R. S., and F. S. Kamp, "Power Output and Power Dissipation in Class B Transistor Amplifiers," preprint No. 545, Audio Engineering Society 33rd Convention, Oct. 16 - Oct. 19, 1967.

Hartz, R. S., and F. S. Kamp, "Stereo Amplifier Power Ratings," *Audio*, April 1968.

Hunter, L. P., *Handbook of Semiconductor Electronics*, McGraw Hill Book Co., Inc., 1958.

Jensen, J. L., "An Improved Square-Wave Oscillator Circuit", *IRE*

Transactions on Circuit Theory, September 1957.

Kolody, O. A., and R. R. Langendorfer, "Measure Transistor Y-Parameters", *Electronic Design*, August 30, 1966.

Kuhn, N., "Simplified Flow-Graph Analysis", *The Microwave Journal*, November 1963, pp. 59-66.

Lee, H. C., "Microwave Power Generator Using Overlay Transistors", *RCA Review*, June 1966.

Lee, H. C., "Microwave Power Transistors", *WESCON Record*, 1966.

Lee, H. C., and G. J. Gilbert, "Overlay Transistors Move into Microwave Region", *Electronics*, March 21, 1966.

Lee, H. C., and R. Minton, "Designing Transistor Multipliers", *Microwaves*, November 1965.

Light, L. H., and M. Hooker, "Transistor DC Converters", The Institute of Electrical Engineers, Paper #1862R, April 1955.

Light, L. H., "The Design and Operation of Transistor DC Converters", *Mullard Technical Communications*, Vol. 2, 17, Feb. 1956.

Matyckas, S. J., "RF Power Transistors In Vehicular Radio Communications Equipment", *WESCON Record*, 1966.

Millman, Taub: *Pulse, Digital, and Switching Waveforms*, McGraw-Hill Book Co., Inc., New York, N.Y., 1965.

Minton, R., "Semiconductor High-Frequency Power Amplifier Design", *WESCON Record*, 1966.

Minton, R., "Design of Large Signal VHF Transistor Power Amplifiers", *RCA Application Note*, SMA-36.

Moll, J. L., "Large Signal Transient Response of Junction Transistors", *Proc. IRE*, Vol. 42, pp. 1773-1784, December, 1954.

Morin, F. J., and J. P. Maita, *Physical Review*, Volume 96, 1954.

Nielson, E. G., "Behavior of Noise Figure in Junction Transistors", *Proc. IRE*, June 1951.

Nowicki, J. R., "New High-Power LC Converter Circuits", *Mullard Technical Communications*, April, 1960.

Nyquist, H., "Regeneration Theory", *Bell System Technical Journal*, Volume II, January 1932.

Pappenjus, E. W., W. B. Bruene, and E. O. Schoenike, *Single-Sideband Principles and Circuits*, McGraw-Hill Book Co., Inc., N.Y.

Pettit and McWhorter, *Electronic Amplifier Circuits*, McGraw-Hill Book Co., Inc., N.Y., 1961.

Phillips, A. B., *Transistor Engineer*, McGraw-Hill Book Co., Inc., New York, N.Y., 1962.

Physical Electronics and Circuit Models of Transistors, SEEC Vol. 2, Chap. 8.

Roddam, T., *Transistor Inverters and Converters*, 1963, Van Nostrand, N. 4, p. 166.

Ruthroff, C. L., *Some Broad-Band Transformers, Proc. IRE*, Aug. 1959.

Schiff, P., "Second Breakdown in Transistors Under Conditions of Cutoff", RCA Application Note, SMA-30.

Schlicke, H. E., *Essentials of Dielectromagnetic Engineering*, 1961, John Wiley, New York.

Searle, C. L., Boothroyd, A. R., Angelo, Jr., E. J., Gray, P. E., and Pederson, D. O., *Elementary Circuit Properties of Transistors, SEEC/* Vol. 3, John Wiley & Sons, Inc., New York, New York, 1964.

Sears, F. W., and M. W. Zemansky, *College Physics*, Addison-Wesley, Cambridge, Chapter 17, 1956.

Simmons, C. D., "Transistor Heat Sinks and Their Evaluation", *Semiconductor Products*, Volume 1, No. 1, Jan./Feb. 1958.

Sorensen, A. A., "Semiconductor Electronics: Part 4, Transistor Network Analysis", *Electro-Technology*, January 1961.

Steinberg, "Ultrasonics in Industry", *Proc. IEEE*, Vol. 53, No. 10, Oct. 1965.

Tatum, J., "RF Large Signal Transistor Amplifiers, Part I—Theoretical Considerations", *Electrical Design News*, May, June, and July, 1965.

Turner, C. R., "Design of Transistorized DC-to-DC Converters", *Electronic Design*, Sept. 16 and 30, 1959.

Uchrin, G. C., and Taylor, W., "A New Self-Excited Square-Wave Transistor Power Oscillator", *Proceedings of the IRE*, Jan. 1955.

Wolf, H., "Recent Advances in High-Frequency Power Transistors", *Solid State Design*, June 1963, pp. 21-28.

Index

RCA Technical Publications

on Electron Tubes, Semiconductor Products, and Batteries

COPIES of the publications listed below may be obtained from your RCA distributor or from Commercial Engineering, Radio Corporation of America, Harrison, N. J.

Electron Tubes

● RCA INTERCHANGEABILITY DIRECTORY OF INDUSTRIAL-TYPE ELECTRON TUBES—ID-1020-H (10⅞" x 8⅜")—12 pages. Lists more than 2300 basic type designations for 22 classes of industrial tube types; shows the RCA Direct Replacement Type or the RCA Similar Type, when available. Single copy free on request.

● RCA INDUSTRIAL RECEIVING-TYPE TUBES—RIT 104F (10⅞" x 8⅜")—24 pages. Concise technical data on over 200 types used in military, industrial, and commercial equipment. Includes application guide, chart of prototype versus similar RCA industrial types, interchangeability list of domestic versus RCA replacements, terminal diagrams, and socket and connector information. Price 25 cents.*

● RCA RECEIVING TUBES AND PICTURE TUBES—ERT-1275M (10⅞" x 8⅜")—56 pages. Contains classification chart, application guide, characteristics chart, and base and envelope connection diagrams on more than 1300 entertainment receiving tubes and picture tubes. Price 40 cents.*†

● RCA INTERCHANGEABILITY DIRECTORY OF FOREIGN vs. U.S.A. RECEIVING-TYPE ELECTRON TUBES—ERT-197E (8⅜" x 10⅞")—8 pages. Covers approximately 800 foreign tube types used principally in AM and FM radios, TV receivers, and audio amplifiers. Indicates U.S.A. direct replacement type or similar type if available. Price 10 cents.*

● RCA NUVISTORS—INDUSTRIAL AND MILITARY—NIT-140—28 pages. A guide for communication equipment designers, researchers and experimenters. Describes design features and performance characteristics. Contains characteristic charts, curves, socket information, dimensional outlines, and terminal diagrams. Price 35 cents.*

● RCA PERIODICALLY FOCUSED TRAVELING-WAVE TUBES—ICE-204—56 pages. Contains theory of operation, design features, and performance characteristics of RCA periodically focused traveling-wave tubes. Prince 50 cents.*

● RCA RECEIVING TUBE AND PICTURE TUBE SUBSTITUTION GUIDE—ERT-198—Price 25 cents.*

● RCA PHOTOMULTIPLIER AND IMAGE TUBES—PIT-700 (10⅞" x 8⅜")—36 pages. Includes concise data on RCA photomultiplier tubes, gas and vacuum photodiodes, sockets and shields for phototubes, and dimensional outlines for photo and image tubes. Price 60 cents.*

● RCA PHOTOMULTIPLIER TUBES FOR NEW-EQUIPMENT DESIGN—PIT-703—16 pages. Reviews some of the applications of photomultiplier tubes. RCA's wide selection is demonstrated by a composite graph of spectral responses; a matrix of spectral response designations versus configuration further assists in preliminary selection of tube types. Additional characteristics are tabulated to help narrow the choice. Price 35 cents.*

● RCA PICTURE TUBE PRODUCT GUIDE—COLOR AND BLACK & WHITE—PIX-300B—24 pages. Includes inter-

changeability chart and characteristics chart on all industry types where RCA has a replacement for both black-and-white and color picture tubes. Basing diagrams and illustrations depicting safety features are also included. Price 30 cents.*

● PRODUCT GUIDE FOR RCA POWER TUBES—PWR-506B—40 pages. Contains tabulated data on all RCA power tubes in order of type designation within each general class of service. Includes maximum ratings, temperature ratings, heater or filament requirements, outline drawings, and basing diagrams. Price 30 cents.*

● RCA INDUSTRIAL TUBES PRODUCT GUIDE—TPG-200C (10⅞" x 8⅜")—28 pages. Covers all RCA industrial-tube product lines. Gives a brief description of each product line together with quick-selection data. Single copy free on request.

● RCA STORAGE TUBES AND CATHODE-RAY TUBES—STC-900B—16 pages. Contains technical information on RCA storage tubes, special-purpose kinescopes and oscillograph-type cathode-ray tubes including display-storage tubes, radechons, scan-conversion tubes, flying-spot tubes, monitor, projection, transcriber, and view-finder kinescopes; as well as data on fluorescent screens. Price 20 cents.*

● RCA TRAVELING-WAVE TUBE CLASSIFICATION CHARTS—MWD-101C—4 pages. Contains catalog-type data. Single copy free on request.

● RCA PENCIL TUBE CLASSIFICATION CHARTS—MWD-102B—4 pages. Contains catalog-type data. Single copy free on request.

● RCA CAMERA TUBES—CAM-600A—26 pages. Contains classification charts, defining data and typical characteristic curves for RCA image orthicons and vidicons. Camera tubes recommended for new equipment design are highlighted. Price 50 cents.*

● VIDICONS—CAM-700—16 pages. Supplies tube selection guidance and data on RCA vidicons for commercial, educational, industrial, and military service. Also included are tube replacement information and typical vidicon characteristic curves. The information contained in this publication supersedes the vidicon section of the booklet CAM-600A. Price 30 cents.*

● TECHNICAL BULLETINS—Authorized information on RCA receiving tubes, transmitting tubes, and other tubes for communications and industry. Be sure to mention tube-type bulletin desired. Single-copy on any type free on request.

Semiconductor Products

● RCA SEMICONDUCTOR PRODUCTS DATABOOK—SPD-100. Two loose-leaf binders for standard 8½" x 11" data booklets with more than 900 pages of data and curves on RCA semiconductor devices such as transistors, silicon rectifiers, and semiconductor diodes. Available on a subscription basis. Price $15.00* including service for first year. Also available with RCA Electron Tube Handbook HB-3 at special combination price of $30.00.*

● RCA SEMICONDUCTOR PRODUCTS GUIDE—SPG-201D (10⅞" x 8⅜")—44 pages. Contains classification chart, index, and ratings and characteristics on RCA's line of transistors, silicon rectifiers, semiconductor diodes, and photocells. Price 75 cents.*

● RCA DIFFUSED-JUNCTION SILICON RECTIFIER STACKS AND BRIDGES—SRS-300—10 pages. Contains technical data on RCA's diffused-junction silicon rectifier stacks and bridges. Characteristics of basic rectifier circuits are also given to assist in selection of proper RCA rectifier device. Price 20 cents.*

● RCA SMALL-SIGNAL SILICON N-P-N TRANSISTORS—SST-210—8 pages. Contains technical data on 2N2102 family of silicon transistors including high-voltage types, very-high voltage types, linear-beta types, and general types. Also includes quick-reference guide. Price 20 cents.*

● DESIGN OF TRANSISTOR SWITCHING CIRCUITS FOR DATA-PROCESSING EQUIPMENT—CTG-161—42 pages. Gives design considerations for a variety of transistor switching circuits for data-processing equipment such as logic gates, flip-flops, and memory drivers. It includes a review of switching theory, design procedures, methods of specifying characteristics and ratings for computer switching transistors; examples of design procedures; typical circuits using RCA transistors; and a complete listing of RCA Computer Transistors with ratings, characteristics, and performance data. Price 75 cents.*

● RCA MOS FIELD-EFFECT TRANSISTORS PRODUCT GUIDE—MOS-160—20 pages. Includes comprehensive data on RCA dual insulated-gate and single insulated-gate MOS FET's in easy-to-find format plus background information on MOS construction and application. Price 20 cents.*

● HEAT-SINK GUIDANCE FOR RCA THYRISTORS USING TO-5 AND "MODIFIED TO-5" PACKAGES—SCR-501—6 pages. Application guide on heat-sink methods for RCA thyristors. Single copy free on request.

● RCA HOMETAXIAL BASE SILICON POWER TRANSISTORS—HBT-400A—18 pages. Contains data, dimensional outlines and theoretical information on hometaxial-base silicon power transistors. Price 30 cents.*

● RCA LOW-NOISE COMMUNICATION-TYPE TRANSISTORS—CTG-165—Contains quick-selection graphs and charts and capsule data for RCA Bipolar Transistors and MOS Field-Effect Transistors for Low-Noise VHF and UHF Communication and Industrial Instrumentation Applications. Includes special characteristics curves showing quick-selection chart containing curves, G_p (dB) and NF (db) vs. f (30 to 1000 MHz) for each listed transistor type. Single copy free on request.

● MOUNTING HARDWARE FOR RCA INDUSTRIAL SEMICONDUCTOR DEVICES—MHI-300—4 pages. Contains mounting information for RCA industrial transistors, thyristors, and rectifiers. Single copy free on request.

● RCA RF POWER TRANSISTORS—RFT-700B—6 pages. Contains data, selection guide, and a quick-selection graph on RCA "overlay" transistors. Single copy free on request.

● RCA PHOTOCELLS—SOLID-STATE PHOTOSENSITIVE DEVICES—CSS-800A—32 pages. Contains detailed and updated information on RCA cadmium-sulfide and cadmium-sulfo-selenide photoconductive-cell characteristics, an extended section on photoelectric measurements, a new section describing design, new circuits, and an extension replacement guide. Price 35 cents.*

● RCA PHOTOCONDUCTIVE CELLS—File No. 312—8 pages. Contains descriptive material, characteristic curves, and classification charts on RCA cadmium-sulfide and cadmium-sulfo-selenide brood-area photoconductive cells. Single copy free on request.

● RCA SILICON POWER TRANSISTOR APPLICATION GUIDE—1CE-215—28 pages. For designers of industrial and military equipment. Discusses ratings, stability conditions, parameters and equivalent circuits. Includes design procedures and specific design equations for several transistor circuits. Price 50 cents.*

● SILICON VHF TRANSISTORS APPLICATION GUIDE—1CE-228—20 pages. For designers of industrial and military equipment. This guide describes the capabilities of RCA silicon vhf transistors for application at frequencies up to 300 MHz. Includes typical circuits for the 2NF1491 family of silicon vhf transistors. Maximum ratings and characteristics are included. Price 50 cents.*

● RCA THYRISTORS (SCR's AND TRIACS)—SCR-500A—22 pages. Contains tabulated data, classification charts and dimensional outlines for all-diffused silicon thyristors. Price 40 cents.*

● RCA TOP-OF-THE-LINE SOLID-STATE REPLACEMENT GUIDE—SPG-202-E—48

pages. Lists 31 RCA "Top-of-the-Line" SK-Series replacement semiconductor devices which can replace more than 9600 types of transistors, integrated circuits, and rectifiers used in entertainment electronic equipment, including U.S.A. industry-standard (EIA) types, foreign types, and types identified only by device-manufacturers' part numbers. Price 15 cents.*

● TRANSISTORIZED VOLTAGE REGULATOR APPLICATION GUIDE—1CE-254—12 pages. Discusses transistorized voltage regulators of the series and shunt types. Includes design considerations, step-by-step design procedures, and the solutions to sample design problems. Price 20 cents.*

Batteries

● RCA BATTERY MANUAL—BDG-111 (10⅞" x 8¾")—68 pages. Contains information on dry cells and batteries carbon zinc, mercury, and alkaline types. Includes battery theory and applications, detailed electrical and mechanical characteristics, a classification chart, dimensional outlines, and terminal connections on each battery type. Price 50 cents.*†

● RCA BATTERIES—BAT-134H (10⅞" x 8⅜")—36 pages. Technical data on 146 carbon-zinc, alkaline, and mercury batteries for consumer and industrial applications. Includes replacement information for 4000 portable radios, and cross-references 860 domestic battery types to their RCA replacements. Price 35 cents.*†

Test and Measuring Equipment

● INSTRUCTION BOOKLETS — Illustrated instruction booklets are available for all RCA test instruments at the prices indicated below.

WA-44A (Audio Signal Generator) $0.50*
WA-44C (Audio Signal Generator) 1.00*
WO-33A (Super Portable Oscilloscope) 1.00*
WO-88A (5-in. Oscilloscope) ...0.75*
WO-91A (5-in. Oscilloscope) ...1.00*
WO-91B (5-in. Oscilloscope) ...1.00*
WR-36A (Dot-Bar Generator) .0.50*
WR-46A (Video Dot/Crosshatch Generator) 1.00*
WR-49A (RF Signal Generator) 0.50*
WR-49B (RF Signal Generator) 1.00*
WR-50A (RF Signal Generator) 1.00*
WR-51A (Stereo FM Signal Simulator) 1.00*
WR-52A (Stereo FM Signal Simulator) 1.00*
WR-61B (Color-Bar Generator) 1.00*
WR-64A (Color Bar/Dot/Crosshatch Generator) 1.00*
WR-64B (Color/Bar/Dot/Crosshatch Generator) 1.00*
WR-67A (Test-Oscillator) 0.25*
WR-69A (Television/FM Sweep Generator) 1.00*
WR-70A (RF-IF-VF Marker Adder) 0.75*
WR-86A (UHF Sweep Generator) 0.50*
WR-99A (Marker Calibrator) ..1.00*
WT-100A (Electron-Tube Micro Mho Meter) 1.75*
WT-100A (Electron-Tube Micro Mho Meter, Ser. No. 1001 and over) 2.00*
WT-100A (Tube Chart 1CE-163) 3.00*
WT-110A (Automatic Electron-Tube Tester) 0.75*
WT-110A (1CE-174 Card Punch Data) 0.25*
WT-110A (1CE-234 Card Punch Data) 1.00*
WT-115A (Color Picture Tube Tester) 0.50*
WV-37A (Radio Battery Tester) 0.25*
WV-37B (Radio Battery Tester) 0.25*
WV-38A (Volt-Ohm-Milliammeter) 0.50*

WV-65A (VoltOhmyst†) 0.25*
WV-74A (High Sensitivity AC VTVM) 0.75*
WV-75A (VoltOhmyst†) 0.25*
WV-76A (High Sensitivity AC VTVM) 0.75*
WV-77A (VoltOhmyst†) 0.25*
WV-77B (VoltOhmyst†) 0.25*
WV-77E (VoltOhmyst†) 1.00*
WV-84C (Ultra-Sensitive DC Microammeter) 0.75*
WV-95A (Master VoltOhmyst†) 0.25*
WV-97A (Senior VoltOhmyst†) 0.75*
WV-98A (Senior VoltOhmyst†) 1.00*
WV-98B (Senior VoltOhmyst†) 1.00*
WV-98C (Senior VoltOhmyst†) 0.50*
195-A (VoltOhmyst†) 0.25*

° Trade Mark Reg. U.S. Pat. Off.

* Prices shown apply in U.S.A. and are subject to change without notice.

† Suggested price.

Other RCA Technical Manuals

● **RCA TRANSISTOR MANUAL**—SC-13 (8⅜" x 5⅜")—544 pages. Contains up-to-date definitive data on over 770 semiconductor devices including tunnel diodes, silicon controlled rectifiers, varactor diodes, conventional rectifiers, and many classes of transistors. Features easy-to-understand text chapters, as well as tabular data on RCA discontinued transistors. Contains over 40 practical circuits, complete with parts lists, highlighting semiconductor-device applications. Price $2.00.*†

● **RCA LINEAR INTEGRATED CIRCUITS**—IC-41 (8¼" x 5⅜")—352 pages. Contains basic principals involved in design and application of linear integrated circuits—includes description of silicon monolithic fabrication process—derivation of design equations and performance criteria—schematic diagrams, operating characteristics, and performance data for RCA multiple-function silicon integrated circuits for a variety of linear applications. Price $2.00.*†

● **RCA SOLID-STATE HOBBY CIRCUITS MANUAL**—HM-90 (8⅜" x 5⅜")—224 pages. Contains complete construction information on 35 circuits of general interest to all experimenters. Circuits use diodes, transistors, SCR's, triacs, MOS transistors, integrated circuits, and light and heat detectors. Circuit operation is described in detail; construction layouts, photographs, schematic diagrams, and parts lists are given; and full-size drilling templates are included for most circuits to simplify construction. Price $1.75.*†

● **RCA TUNNEL DIODE MANUAL**—TD-30 (8⅜" x 5⅜")—160 pages. Describes the microwave and switching capabilities of tunnel diodes. Contains information on theory and characteristics, and on tunnel-diode applications in switching circuits and in microwave oscillator, converter, and amplifier circuits. Includes data for over 40 RCA germanium and gallium arsenide tunnel diodes and tunnel rectifiers. Price $1.50.*†

● **RCA SILICON CONTROLLED RECTIFIER EXPERIMENTER'S MANUAL**—KM-71 (8⅜" x 5⅜")—136 pages. Contains 24 practical and interesting control circuits that can be built with a complement of active devices available in kit form. Includes photographs, schematic diagrams, and descriptive writeups. Also includes brief descriptions of solid-state components used (rectifiers, transistors, SCR's) and short section on trouble-shooting. Price 95 cents.*†

● **RCA RECEIVING TUBE MANUAL**—RC-26 (8" x 5¼")—656 pages. Contains technical data on more than 1400 receiving tubes for home-entertainment use. Includes six easy-to-read text chapters that provide basic information on electron-tube operation, ratings and characteristics, and applications. Also features a detailed application guide for receiving tubes; quick-reference charts for replacement and discontinued RCA receiving tubes, black-and-white and color picture tubes, and voltage-regulator and voltage-reference tubes; and a Circuits section that includes schematic diagrams, descriptive writeups, and complete parts lists for 36 practical electron-tube circuits for a wide variety of applications. Price $1.75.*†

● **RCA TRANSMITTING TUBES**—TT-5 (8¼" x 5⅜")—320 pages. Gives data on over 180 power tubes having plate-input ratings up to 4 kw and on associated rectifier tubes. Provides basic information on generic types, parts and materials, installation and application, and interpretation of data. Contains circuit diagrams for transmitting and industrial applications. Features lie-flat binding. Price $1.00.*†

● **RCA PHOTOTUBE AND PHOTOCELL MANUAL**—PT-60 (8¼" x 5⅜")—192 pages. Well-illustrated informative manual covering fundamentals and operating considerations for vacuum and gas phototubes, multiplier phototubes, and photocells. Also describes basic applications for these devices. Features easy-to-use selection chart for multiplier phototubes. Data and performance curves given for over 90 photo-sensitive devices. Price $1.50.*†

● **RADIOTRON° DESIGNER'S HANDBOOK**—4th Edition (8¾" x 5½")—1500 pages. Comprehensive reference covering the design of radio and audio circuits and equipment. Written for the design engineer, student, and experimenters. Contains 1000 illustrations, 2500 references, and cross-referenced index of 7000 entries. Edited by F. Langford-Smith. Price $7.00.*†

° Trade Mark Reg. U.S. Pat. Off.

* Prices shown apply in U.S.A. and are subject to change without notice.

† Suggested price.

RCA | Electronic Components | Harrison, N.J. 07029