ELECTRONIC TECHNOLOGY SERIES

SEMICONDUCTORS and TRANSISTORS

a RIDER publication

SEMICONDUCTORS AND TRANSISTORS

Edited by

Alexander Schure, Ph.D., Ed.D.

JOHN F. RIDER PUBLISHER, INC.,

NEW YORK

Library of Congress Catalog Card No. 60-150 24

Printed in the United States of America

PREFACE

Semiconductors and the transistors derived from these materials had a limited use in the 1940's. They are now prominently used in components manufactured for electronic communication equipment.

The circuits and design considerations pertaining to semiconductors and transistors are a vital part of electronic theories and practice. It is necessary that those concerned with electronics possess a working knowledge of their essential relationships to this important, comparatively new, aspect.

The intent of this book is to present, discuss, and evaluate the important ideas relating to an understanding of the theory and characteristics of these devices. Thus, the book begins with fundamental presentations and concepts of semiconductors. It then leads into the assembly of a series of semiconductors into various types of transistors, presents basic concepts, and gives an understanding of transistor characteristics, as well as a development of circuit operations. The mathematical analyses are simple; but, the treatments are sufficiently extensive to permit the reader to develop full comprehension of the pertinent theory. To insure this aim, adequate information is given relating to broad concepts and designed for ready use. Detailed descriptions of selected major topics are presented. And, through presentation of practical situations and problems, the reader is given an opportunity to apply the principles he has learned.

Specific attention is given to atomic structure, the quantum theory, conductors, insulators, and semiconductors. Conduction by holes, semiconductors with impurities, and the semiconductor rectifier are also detailed. The p-n junction under equilibrium conditions, the p-n junction in reverse and forward bias, diode types (including photodiodes), and an extension of p-n junction theory into the transistor receive specific attention. Also discussed is a comparison of the basic transistor and vacuum tube circuits, transistor symbols, characteristics, analysis and design, drawing the load line, establishing the Q point for the common base and common collector amplier, regions of operation, static and dynamic quantities, circuit theorems, and hybrid parameters. Finally, h parameters and elements of design relating to voltage, current, and power gain are analyzed. A foundation is thereby provided upon which more advanced concepts can be built.

Grateful acknowledgment is made to the staff of the New York Institute of Technology for its assistance in the preparation of the manuscript of this book.

February 1961
New York, N. Y.

A. S.

CONTENTS

Chapter 1

INTRODUCTION TO THE SEMICONDUCTORS

Up to World War II, a material was considered either a conductor or an insulator of electricity, or it varied in values of either. As an outgrowth of crystal rectifier research during the 1940s, we began to understand the processes of a new type of conductive medium, *the semiconductor.* A semiconductor has a conductivity value somewhere between that of an insulator and a metallic conductor. The identifying characteristic of a semiconductor is that elements called **holes** (as well as electrons) are responsible for electronic conduction. (In a normal conductor, electrons are the sole carriers of electric current.) Three or more pieces of semiconductor material fastened together in a special way form a TRANSISTOR. Thus, the transistor is born after the process of semiconduction.

1. The Atom

The atomic theory of matter was one of the most important contributions to physical science. In 1805, John Dalton advanced the idea that all matter was made up of smaller particles called *atoms* (from the Greek word *atomos,* indivisible). Today, his theory is an accepted fact. *An atom is the smallest particle of an element that still possesses the properties of that element.*

What does an atom look like? Where did it come from? Why does it behave as it does? Through the years, scientists have explored the mystery surrounding the atom. Today, our atomic theory not only explains many of the atom's physical characteristics, but can predict what will happen when two or more atoms come into contact with one another.

Every atom consists of a nucleus and one or more electrons. Electrons spin around the nucleus of an atom in much the same manner as the planets of our solar system orbit the sun. Figure 1 illustrates the simplest of atoms—the hydrogen atom. Here, the nucleus is not an entity in itself, but is a term used only to designate the central part of an atom. Essentially, the nucleus contains one or more subparticles called *protons.* Just as the sun is larger than any of the planets, the proton is larger than the electron (about 1800 times larger). In the simple hydrogen atom, one electron orbits one proton. In larger atoms, many electrons orbit about a nucleus holding many protons.

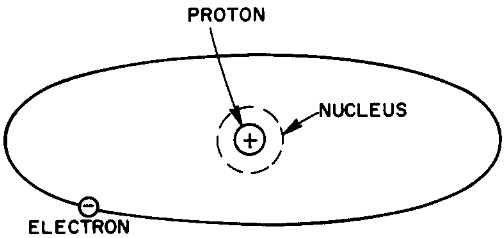


Fig. 1. The hydrogen atom.

Actually, a comparison of an atom with our solar system is quite weak, since there are several glaring differences. Whereas our solar system is millions of miles in diameter, the atom has a diameter in the order of an *Angstrom unit.*[1] For instance, if an ordinary marble were expanded to equal the size of the earth, its constituent atoms would be about the size of ping pong balls; or, a marble contains about a billion billion (10^{24}) atoms.

The motion of our solar system is governed by the attractive force of gravity. This force exists between all particles of matter. Gravitational forces also exist between all particles of an atom, but are eclipsed by a much stronger force—*electrostatic force.* Governing this force is the familiar Coulomb's Law: like electrically charged bodies repel one another, and unlike charged bodies attract one another. The electron is a negatively charged body; the proton is a positively charged body. The electron and proton have the same magnitude, but opposite types of charges. Two electrons, or two protons, tend to repel each other. A proton and an

[1] The Angstrom unit, symbolized by Å, denotes a length of 0.0000000001 meter (or, 1×10^{-10} meter).

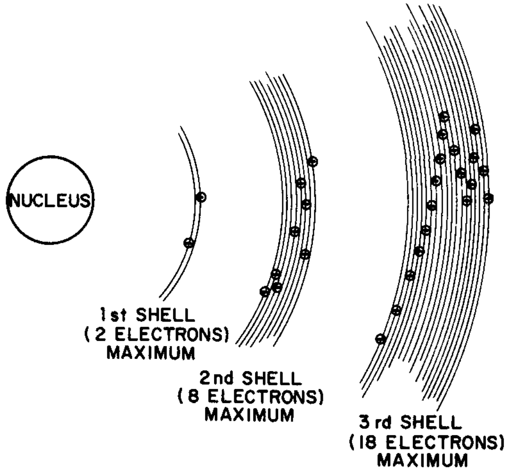


Fig. 2. Shells of an atom.

electron attract one another. For convenience, we designate both the electron and proton as having one unit of electronic charge.

An atom having the same number of protons in its nucleus as it has electrons outside is electrically neutral, with a net electronic charge of zero. This is the natural case with all atoms, and is true for the hydrogen atom of Fig. 1. If an atom gains or loses electrons, it is then called an *ion*. Ions are atoms having an excess or deficiency of electrons orbiting about the nucleus. For instance, if an atom loses two electrons, it is designated an ion with a +2 charge (or a *valence* of +2). If an atom has two extra electrons, then it becomes a negative ion with a valence of −2.

Do electrons ever hit one another when so many of them spin around the nucleus at once? NO, because *no two electrons have exactly the same orbit.* If we think of the orbit of two electrons as being different from one another, when their respective orbital radii are different, then it may be said that no two electrons travel at exactly the same distance from the nucleus. The orbital radius may also be expressed in terms of energy. Since an electron far from the nucleus has a larger potential energy than an electron close to the nucleus, we may also say that no two electrons ever have the same energy. Although all of the electrons of an atom have different energies, or orbits, they tend to crowd together in energy bands called *shells.* There is a much greater distance between the electrons of any two shells than between the electrons of a given shell (see Fig. 2).

Figure 3 schematically illustrates some common neutral atoms. This is a two-dimensional representation, assuming all the electrons travel around the nucleus in a flat plain. The dotted lines indicate different shells. (Within each shell, the electrons indicated have minute energy differences and, therefore, slightly different orbital radii.) Notice that the heavier atoms have more electrons and more protons. The number of protons in a neutral atom (or the number of electrons) is called the *atomic number*. Starting with Hydrogen, atomic number 1, the numbers extend to 92 (excluding the artificially created elements of the atomic age). If there were room to draw all 92 elements, one would notice that,

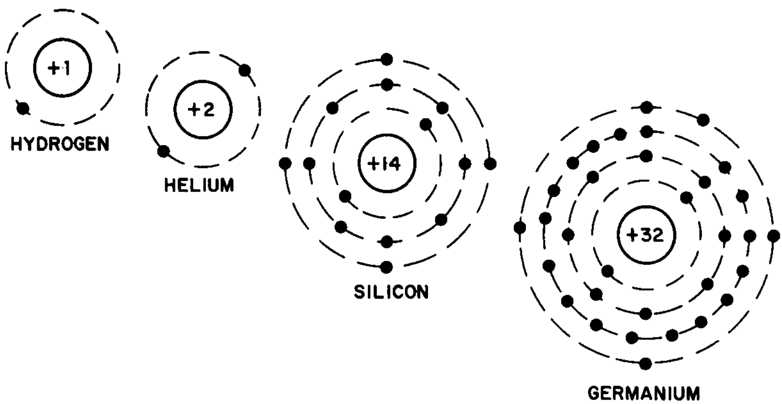


Fig. 3. Schematic representation of some neutral atoms.

as the atomic number increases, the number of electrons in any given shell do not increase continually. The first shell allows only two electrons. With atomic number 3 (Lithium), the extra electron begins a new shell, which has a maximum number of eight. After that, the third shell must be filled with eighteen electrons, and so on. It must be emphasized that, although each shell may contain only a certain maximum number of electrons, there are very many unused permissible orbits within each shell.

It was mentioned previously that the forces governing the action of an atom are electrostatic in nature. This is only partly true. An electron stays in an orbit because the electrostatic force of attraction to the nucleus *exactly balances the centrifugal force* that tends to throw the electron out of the atom. This is a very delicate balance. The velocity of the electron must be just right, because the centrifugal force varies as the square of the velocity. If the electron gains or loses any energy, this balance cannot be maintained; and the electron will move to a different orbit. The electron may not, however, jump to an occupied orbit or to a filled shell. Only discrete energy levels are allowed.

2. The Quantum Theory

The amount of energy an electron can absorb is proportional to how far it jumps from one orbit to another and, since only certain jumps are allowed, the electron may absorb only discrete amounts of energy. The reverse is also true. If an electron jumps from an outer orbit to one slightly nearer the nucleus, it will give up a discrete amount of energy. To reverse this action, *i.e.*, to make the electron jump back to an outer orbit, would necessitate the electron absorbing *exactly* the same amount of energy. These bundles of energy are called *quanta of energy*. Energy may be supplied to an atom by several means.

Energy is usually given off by an atom in the form of *light*. All visible light, in fact, all electromagnetic radiation (radio waves, heat waves, X rays, etc.) is the result of electrons in atoms jumping from higher to lower energy levels. This, in essence, is the **Quantum Theory of Light**. This theory plays a vital role in the action of semiconducting devices.

These bundles, or quanta of energy being radiated, are often called *photons*. (A photon is not energy, pure and simple.) Photons exist at the boundary line between mass and energy, possessing some of the features of each. The argument as to exactly what a photon is has raged for fifty years. However, we do know that if an electron at a higher orbit corresponding to an energy level W_1 jumps down to a lower orbit corresponding to an energy level W_2, the emitted photon will have a wavelength (λ) in Angstrom units of:

$$\lambda = \frac{12{,}400}{W_1 - W_2} \text{ Å} \qquad (1)$$

where W_1 and W_2 are in electron volts[2]

This establishes a direct relationship between the wavelength of the emitted photon and the difference in energy between the two orbits in question. Einstein's equation $E = MC^2$ (where E = energy, M = mass, and C = speed of light) told us that there is an equivalence between mass and energy. Consequently, the photon has a mass in addition to a specific energy. The shorter the wavelength the higher the energy, and the greater the mass of the photon.

---

[2] The electron volt is a unit of energy, just as the kilowatt hour is a unit of energy. It represents the amount of energy gained by an electron accelerated by a potential difference of one volt.

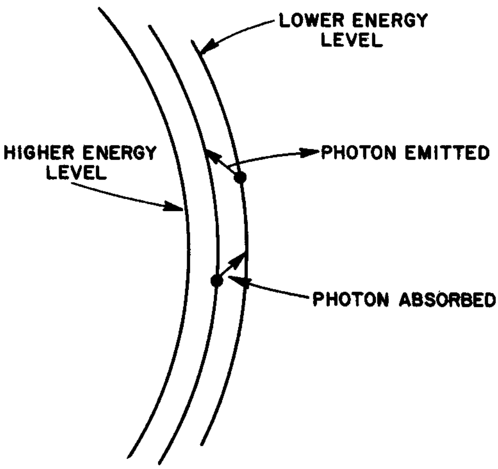


Fig. 4. Absorption and emission of photons.

When a photon has precisely the correct amount of energy (corresponding to a specific wavelength) and hits an orbiting electron of an atom, it may be absorbed by the electron—and, in the process, impart all of its energy to that electron. The electron, in turn, will jump to a higher permissible energy level, where the difference in energy between the old and the new orbit equals the energy of the photon, in accordance with Equation (1).

A completely reversible action may take place within an atom. If an electron falls to a lower permissible energy level, a photon is emitted. If a photon of the correct energy (*i.e.*, wavelength) hits an electron, it will knock the electron up to a higher energy state. It is even possible that this electron might fall back to its original orbit and re-emit a photon of the same wavelength, as shown in Fig. 4. However, if a photon of other than the correct or permissible energy hits an electron of an atom, it will pass right through—without affecting the atom and without being itself affected.

Is it possible to impart enough energy to an electron to knock it out of the atom? YES. If, by photon bombardment or some other means, an electron is given enough energy, it can jump right out of the atom. This process is called *ionization,* because an atom—minus the electron—is called an ion.

It takes energy to ionize an atom by removing an electron that is orbiting near the nucleus; whereas it takes relatively little energy to ionize an atom by removing the loosely-held outer electrons (called *valence electrons*). In this instance, the discrete energy level rule does not apply; *i.e.,* a photon need contain only enough energy to make the electron leave the atom. If it has more than this minimum energy, the departing electron will retain the additional energy in the form of velocity.

Photon bombardment is not the only way to ionize, or excite an atom. Ordinary heat is sufficient. Or, if another high speed electron from somewhere happens to hit an orbiting electron, the orbiting electron may be raised to a higher energy level—or even be knocked out of the atom. The only restriction is that the impinging electron (or any other source of energy) must possess just the right amount of energy to excite an orbiting electron to a higher energy state, or must have the minimum amount of energy to ionize the atom.

3. Atoms in Metals

Thus far in our study of the atom, we have concentrated on its behavior in an isolated state. We will now discuss how a number of similar and dissimilar atoms behave together. As previously mentioned, all matter is composed of one or more types of atoms. A material composed of only one type of atom is called an *element*. Iron, lead, hydrogen, etc., are common examples of elements.

When atoms of different elements are combined, the resulting material is called a *compound*. For instance, the compound water is made up of two atoms of the element hydrogen and one atom of the element oxygen. (The familiar symbol H_2O designates its atomic ingredients.) Yet the compound water bears little resemblance to its ingredients of oxygen and hydrogen. Another example is common table salt. It is composed of one atom each of the elements sodium and chlorine. Alone, each of these elements is a deadly poison; when combined to form salt, they become essential to life. Although there are only 92 elements, the number of compounds in existence is literally in the millions.

It is often said that all matter has three states: gaseous, liquid, and solid. When the different atoms of the material are separated by a large distance, compared to the radii of their outer orbits, then the attractive forces tending to keep them together (more will be said later) are rather weak, and the atoms drift apart to form a gas. When atoms are separated by a distance comparable to their atomic radii, about 1 Å, the various atoms slip and slide about, and the substance is in the liquid state. However, when atoms are so close together that the outer orbits of adjacent atoms overlap, the attractive forces between the atoms become large. The material is then called a solid, wherein each atom is tightly bound in a particular position. It is this solid state which interests us.

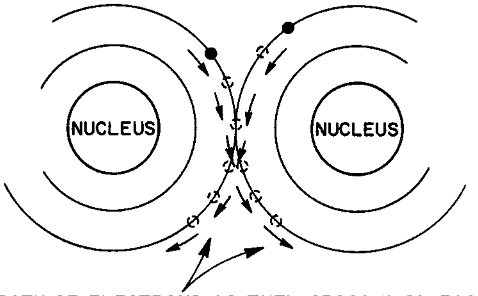
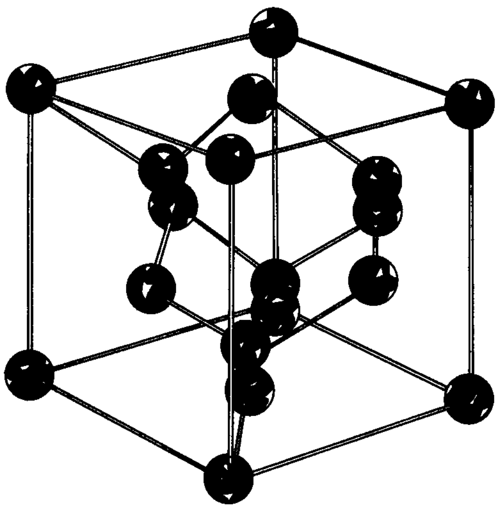


Fig. 5. The sharing of electrons.



CRYSTAL STRUCTURE OF GERMANIUM

Fig. 6. Geometric representation of a crystal structure.

A strange thing happens when atoms are so close together. The encircling electron of one atom has difficulty distinguishing the orbit of its own atom from the overlapping orbit of the adjacent atom. And, the electron spends as much time circling about its atom as it does the other. This is termed *sharing; i.e.,* the two atoms "share" an electron from each other's outer orbit. (See Fig. 5.) Sharing of electrons results in the greater strength of the solid structure.

The force of attraction between two atoms sharing one of their outer, or valence, electrons is called a *covalent force.* A covalent, or electron pair bond, exists between the two atoms. This type of bond is so important, so nearly always present in substances, that Professor G. N. Lewis of the University of California (1875-1946), who first explained its electronic structure, called it THE chemical bond. The process of sharing of electrons is *not the same* as that of ionization. The valence electrons are neither lost nor gained but shared.

Crystals are the formation of numbers of atoms joined in covalent bonds. A crystal is a body that is composed of an often-repeated geometric figure. The crystal structure of germanium is illustrated in Fig. 6. This is a simple cubic crystal structure. There are many, more complicated structures. All crystals are merely repetitions of a basic structure, called *the crystal lattice*. Figure 6 is a simple cubic lattice. Externally, a crystal takes on the appearance of its lattice. Iron Pyrite (fool's gold) has a cubic lattice and, therefore, the entire crystal is cubic. In Fig. 6, each circle symbolizes an atom, while the bars joining them indicate the covalent forces, or bonds of attraction.

Figure 7 illustrates a two-dimensional view of a crystal composed of what is called the *diamond lattice* structure. Imagine that you are close to the crystal and can see the outer valence electrons of each atom. (The electrons belonging to the inner shells of each atom are tightly bound to their parent nuclei and are of no interest to us.) We can now examine many interesting properties of metals.

Each large circle in Fig. 7 represents the nucleus and bound electrons (*i.e.*, all but the outer valence electrons of each atom). In this particular crystal, each atom has four valence electrons. Each valence electron has joined in a covalent bond with its neighboring atom. The dotted lines indicate this covalent bond and show that the valence electrons are "shared" by each neighboring atom. Note the geometrical symmetry that exists throughout the entire structure. Every valence electron of every atom is accounted for and is tied with an adjacent atom in a covalent bond. What would happen if each of the atoms of our hypothetical metal had five, instead of four, valence electrons? The extra electron associated with each atom obviously could not fit in the existing system of covalent bonds; there would be no room in the scheme. Exactly what does happen depends upon the temperature.

As was said, heat supplies energy to atoms. Absolute zero (−460°F) is the lowest possible temperature. At this point, there is no heat whatsoever. According to classical theory, the electrons of any atom possess *no* energy at this temperature. The Quantum Theory, disproved this belief. Electrons at absolute zero *do* produce energy. These *extra valence* electrons for each atom are loosely bound, because they are not tied in covalent bonds. Consequently, even at absolute zero, some of these extra electrons may have just enough energy to leave their parent atoms and wander

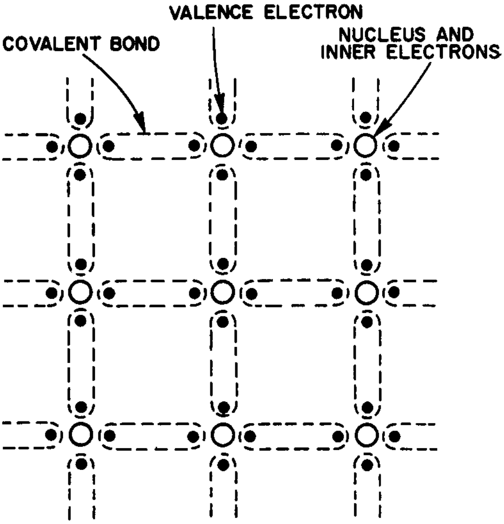


Fig. 7. A two-dimensional view of a crystal.

aimlessly throughout the metal, and not be captured by any other nuclei. These wandering electrons are called *free electrons* and the atoms that are left are called ions, because they lack a valence electron. This happens at absolute zero. However, it is not a very practical picture, because a temperature at absolute zero has never been created. Remember—at any temperature, a proportion of the electrons will have all ranges of energies. And, at absolute zero, although most of them will have very low energies, there will always be some electrons with enough energy to escape.

At higher temperatures, a greater proportion of these extra electrons will have enough energy to leave their parent atom and wander. The effect is not unlike making popcorn. When heat is applied, all of the corn is about the same temperature; but only some of the corn pops. When more heat is applied, a greater proportion of the corn pops. It is futile, however, to talk about just when a given kernel will pop. For the same reason, we cannot talk about the energy of an individual electron, but only of the energy that most of the electrons probably have. So we say that if the temperature is high enough, nearly all of these extra electrons will probably have enough energy to become free electrons.[3]

[3] It is worthwhile to note that if the metal gets hot enough (about 2000°C) the extra electrons will have so much energy that some of them will not only become free electrons but will leave the surface of the metal. This is how thermionic emission from a hot cathode is accomplished in a vacuum tube.

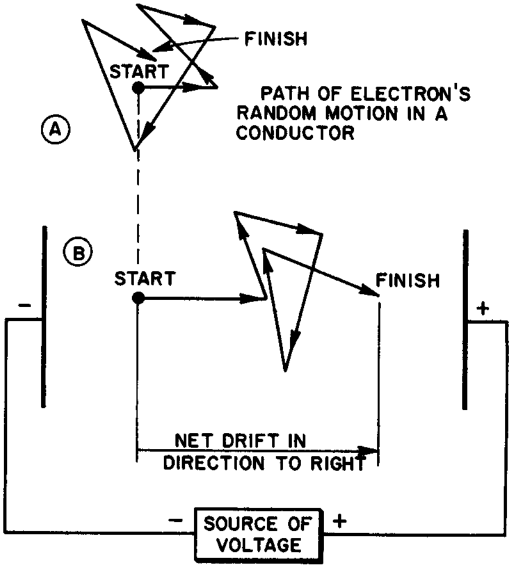


Fig. 8. Motion of an electron: (A) With no applied field; (B) With an applied field.

It is these extra electrons that are present in all metallic conductors. They are called free electrons because they are free to drift in any direction. Of course, in the absence of an externally applied force, each free electron will travel just so far before it bumps into another atom or ion. Since the direction of travel is changed during each collision, the net travel in any given direction is zero, as shown in Fig. 8 (A). Each point where the direction of travel is abruptly changed represents a collision. The distance the electron travels between collisions is called the *mean free path.*

4. A Conductor

If our piece of metal were connected across the terminals of a battery [Fig. 8 (B)], an electric field would then be set up across this intervening space. Under these conditions, the free electrons would slowly drift toward the positive terminal, as indicated in Fig. 9. These moving electrons constitute an electric current. In short, this is the process of electronic conduction in a metal. If the battery voltage is increased, the electric field is increased, and the electrons are attracted to the positive terminal with a greater force. The current will then increase. Of course, the free electrons can only reach the positive terminal by an indirect path, since they are continually colliding with other atoms.

Note that superimposed upon the random movement is a steady drift in the direction indicated. Because numerous collisions impede the flow of free electrons, it is said that the metal has resistance. The ability of free electrons to move through this maze of impeding atoms is called the *mobility of the electrons.* There is a direct relationship between the mobility, the net drift velocity in a given direction, and the applied electric field. Thus if V_d is the net drift velocity of the free electron in meters per second, ξ is the

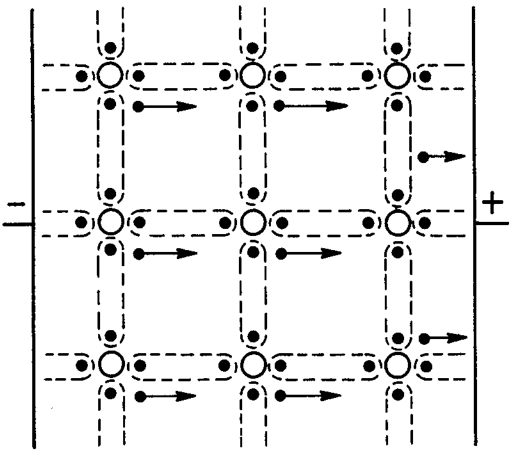


Fig. 9. Motion of free electrons towards positive terminals.

applied electric field in volts per meter, and μ is the mobility, we have the relationship:

$$V_d = -\mu\xi \qquad (2)$$

The minus sign appears because the electrons always travel in the opposite direction of the applied field.

Is atomic theory justified by Ohm's Law? We know that a metal has resistance and that current through the metal is proportional to the voltage across it. How does the atomic theory justify the fact that different metals have different resistances? Remember, our hypothetical metal had but *one* extra electron per atom. Another material could have been used that had more than one free electron to contribute to the flow of current. One, two, and sometimes, three, electrons (depending on the metal) per atom fall into this "extra" classification (*i.e.,* they are not tied in covalent bonds with other atoms) and hence can become free electrons. With all these free electrons, the material is a better conductor. Therefore, various metals have different resistances, if they have different numbers of unbonded valence electrons.

What would happen if the metal is now heated? Empirical observations prove that the resistances of *all* conductors increase with temperature. At first glance atomic theory might suggest that the resistance of a metal should decrease as the temperature is raised, because the added heat would probably liberate more of these "extra" electrons. This is true; but this effect is completely overshadowed by another. When heat is applied, the relatively massive and stable nuclei begin to vibrate back and forth in a random manner (although they never drift out of their assigned positions in the crystal structure). When this vibration occurs, the free electrons moving toward the positive electrode collide more easily with the nuclei than would if the nuclei were stable. Of course, free electrons collide with nuclei (or ions) almost all of the time—with or without heat.[4] But when the temperature is raised and the nuclei vibrate, the incidence of collision becomes so much greater that the net velocity drift towards one electrode is lessened. The conductive current decreases, equivalent to an increase in resistance. This *positive temperature coefficient* (a rise of resistance with a rise in temperature) is an identifying characteristic of all conductors. In more technical terms when the metal is heated, the mobility of the free electrons decreases at a much faster rate than the increase of thermally-generated free electrons.

5. An Insulator

We now know that a material's resistance is governed, to a large extent, by the number of free electrons that are available for conduction at a given temperature in a given volume of the material. If a material has all of its valence electrons tide in covalent bonds, as was shown in Fig. 7, there will be no free electrons for conduction and it will be a perfect insulator. There is, of course, no such thing as a perfect insulator. Some impurities in a material are always present that are not tied in interatomic bonds and, therefore, become a source of a limited number of free electrons. A good practical insulator has about 10^7 free electrons per cubic meter; a good conductor has 10^{28} free electrons per cubic meter. Thus, a good insulator has only one thousand billion billionth ($1/10^{21}$) the number of free electrons that a good conductor has.

In order to present a simplified explanation of some processes of solid state physics, a highly stylized situation was depicted. Actually, metals are rarely a single crystal, but many crystal lattices bound in an arbitrary or fixed way. Nevertheless, the remarks previously made regarding electronic conduction are true. Yet, the neat crystal of Fig. 7 is not truly representative of a conductor.

---

[4] It is well known that a conductor becomes hot when sufficient current passes through it. This is a direct result of its many electron-ion collisions.

Usually there are so many free electrons present that no heed is given to individual ions. But, in the modern concept, a metallic conductor is visualized only as a periodic three-dimensional array of tightly-bound ions, with a "sea" or swarm of free electrons moving at random throughout the structure.

6. The Semiconductor

A semiconductor falls into a special classification all its own. It is neither an insulator nor a conductor, but can be made to appear as either.

Let us again visualize a metallic crystalline solid (see Fig. 7). We know that when heat is applied to the crystal, the atoms are energized, and vibrate back and forth. This is called *thermal agitation.* If the atoms of the material have the special feature of forming very loose covalent bonds with one another, then the energy acquired by the atoms through heat might be sufficient to break some of these bonds and create free electrons for conduction. Many materials have this property; they are called *semiconductors.* Some of the commercially important semiconductors are germanium, silicon, and copper oxide. (Unless otherwise specified, germanium will be used to illustrate all aspects of semiconductor theory.) Semiconductors may be thought of as existing in two broad categories: those with, and those without, impurities. Semiconductors having no impurities are called *intrinsic semiconductors.*

Figure 7 is an intrinsic semiconductor, because the entire structure is a homogeneous mass of interlocking atoms; there are no impurities present. If the temperature of this crystal is low enough, the material will not conduct current (*i.e.,* it will be a perfect insulator). It will not conduct because the valence electrons of all the atoms are bonded to adjacent atoms. The atomic picture is exactly the same as that of a physical insulator. The difference, however, between a physical insulator and an intrinsic semiconductor is that at moderate temperature (room temperatures), some of the weakly-bonded semiconductor electrons break away and become free electrons. With an insulator, the covalent bonds are much stronger, and a moderate amount of heat would not supply enough energy to break the covalent bonds. This is probably the only atomic distinction between the two materials.

It is even possible to heat an insulator to such a high temperature that some of the covalent bonds break, resulting in an accompanying liberation of free electrons. Under these conditions, an insulator will become an intrinsic semiconductor. This is why some insulators cannot be used at high temperatures.

Heat need not be the only source of energy used to break covalent bonds in a semiconductor. If a photon of sufficient energy strikes a valence electron, it will break the bond. What will happen? Let us assume a photon has just broken one covalent bond and liberated one electron. A vacant place is now left in the location from which the electron was just ejected. Perhaps the most important single concept in semiconductor electronics is that **this vacancy, produced by the ejected electron, acts as if it were itself a positive electron.**

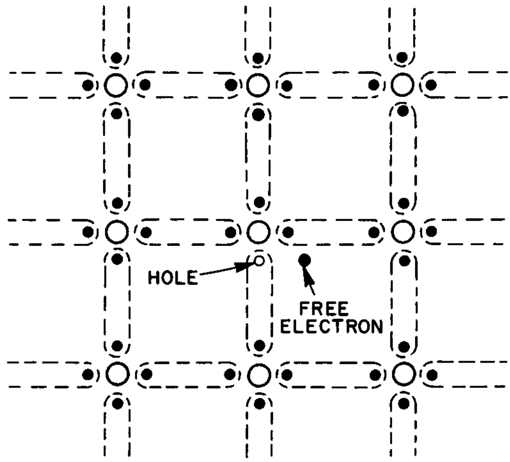


Fig. 10. Holes and electrons in an intrinsic semiconductor.

This vacancy is called a *hole* and it (along with the free electron simultaneously created) is free to move about in the crystal structure.

When an electric field is set up across a section of the crystal (see Fig. 10), the electron will drift to the positive side of the field, and the hole to the negative side. Conduction by electrons is called *electron conduction*. Conduction by holes is called *deficit conduction. That the two types of conduction may simultaneously —and even separately—take place within a semiconductor is the essence of transistor theory.*

7. Conduction by Holes

Holes may serve as conductors of electric current. If (as in Fig. 10) a covalent bond of an intrinsic semiconductor has just been broken and a hole created, it is relatively easy for a bonded electron in one of the surrounding atoms to break away and fall into the vacant place represented by the hole. Remember, *a hole is equivalent to a positively charged body* and, therefore, exerts an attractive force on other negative particles in the immediate vicin-

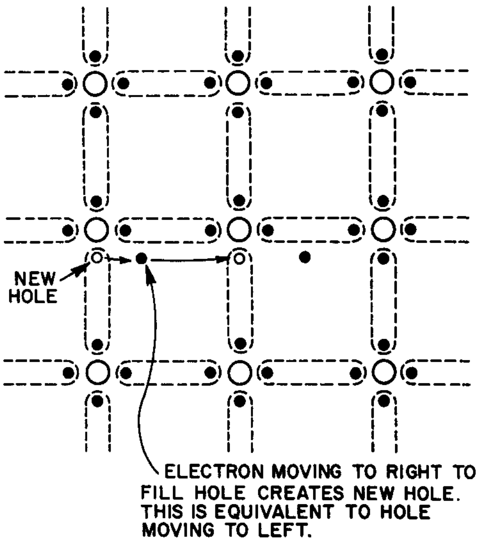


Fig. 11. Motion of electrons in an intrinsic semiconductor.

ity. In Fig. 11, it can be seen that an electron from an adjacent atom has broken its bond and filled the original hole. In the process, however, a new hole is created in the location from which the electron came.

In this manner, a hole can jump from atom to atom. A hole traveling in one direction is created by an electron traveling in the opposite direction. It is said that a *hole current in one direction is equivalent to an electron current in the opposite direction.* This is a very important concept. **Holes do not exist; they only represent the absence of electrons.**

In Fig. 10, an electron and hole (called an electron-hole pair) have simultaneously been created by the action of photon bombardment. If given sufficient time, we will see the drifting electron undoubtedly fall back into the hole. The free electron and hole have effectively disappeared and ceased to exist. The time it takes for this to happen is called the *average lifetime* of the electron-hole pair. The average lifetime of germanium and silicon is approximately 100 microseconds. After this time, annihilation of the electron-hole pair takes place due to this *recombination.*

As previously stated, under the influence of an electric field, the holes will drift to the negative side of the field, the electrons to the positive side. When a hole approaches the negative plate (see Fig. 12), a free electron jumps out of this plate and fills the hole. Remember that the plate can be considered a wire from a battery connected to one end of the semiconductor material. Since it is a conductor, this plate has a ready supply of free electrons. Thus,

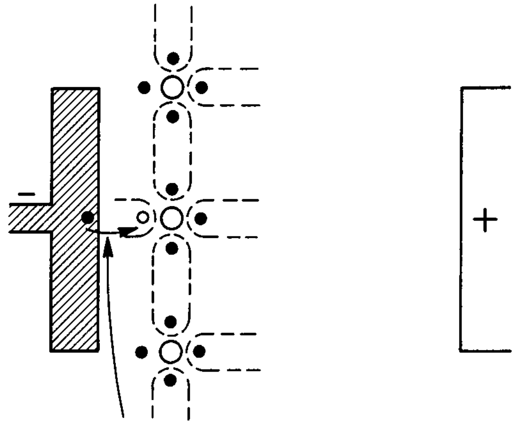


Fig. 12. Movement of holes in an intrinsic semiconductor.

for every hole attracted to the negative plate, an electron enters the semiconductor. Holes can never travel in the conducting plate or wire. Although positive holes and negative electrons move in opposite directions within the semiconductor crystal, the electric current flows in one direction—the direction of electron flow.

Let us summarize some salient information about intrinsic, or pure, semiconductors.

In an intrinsic semiconductor, a hole can only be created when a free electron is ejected from a covalent bond. Thus, the concentration of holes always equals the concentration of free electrons. At room temperatures, thermal agitation always creates new electron-hole pairs while, at the same time, other electron-hole pairs disappear because of recombination. The generation and recombination of electron-hole pairs results in an equilibrium between the two processes, and a definite electron-hole concentration will exist for a given temperature.

Holes have a lower mobility than free electrons. Holes must travel across a semiconductor in jumps from one atom to another, but an electron may travel a more direct path, even considering the many electron-ion collisions.

When a conductive metal is heated, the mobility of the free electrons decreases and the production of free electrons increases. The decrease in mobility, however, overshadows the latter effect, resulting in an overall *increase* in resistance, as the temperature is raised. As the temperature of a semiconductor is raised, the mobility decreases as well, and the rate of production of electron-hole pairs increases. *With a semiconductor, however, the increased production of electron-hole pairs is much greater than the decrease in mobility.* Consequently, an intrinsic semiconductor's resistance *decreases* with an increase in temperature. In other words, a conductor has a positive temperature coefficient and a semiconductor has a negative temperature coefficient. For reasons described previously, an insulator, like a semiconductor, has a negative temperature coefficient.

8. Semiconductors with Impurities

Intrinsic semiconductors are not very useful, because electrons and holes are always produced in equal numbers. Practical use of semiconductor devices is possible only because of the addition of minute quantities of impurities to the crystal structure. By virtue of these impurities, we are able to make either electrons or holes the predominant carrier of electric current.

A basic requirement in adding these impurities is that the atoms of the impurity material must be almost like the intrinsic semiconductor atoms. When this condition is met, it is possible to fit some of the impurity atoms in the semiconductor crystal lattice replacing some of the intrinsic semiconductor atoms in the process. This is easily done by mixing a minute quantity of the impurity with a batch of molten semiconductor material. When the crystal starts to form, as the mass is cooled, the impurity enters into a covalent bond with the surrounding semiconductor atoms.

What happens with the addition of one type of impurity is shown in Fig. 13. Here, an impurity atom (having 5 valence electrons) is surrounded by 5 germanium atoms, each having 4 valence electrons. Note that *one* of the impurity's valence electrons is extra; that is, it is not bonded to any other atom. For this one extra electron, we may say that it most probably will (even at room temperatures) possess enough energy to break its bond with the nucleus and become a free electron. However, note that in this case, unlike the intrinsic semiconductor, a hole *was not* created with the liberation of the electron. In addition, it must be remembered that the overall crystal structure has not been altered by the liberation of this extra electron. The ion that is left after the liberation, "looks" to the crystal almost like another germanium atom, insofar as its ability to help hold the crystal together is concerned.

Since, in this instance, a free electron has been donated to the crystal, it is said that the impurity atom is a *donor* atom. Typical donor elements that have five valence electrons are phosphorus, arsenic, and antimony. When an electric field is set up across a crystal, most of the current will be carried by

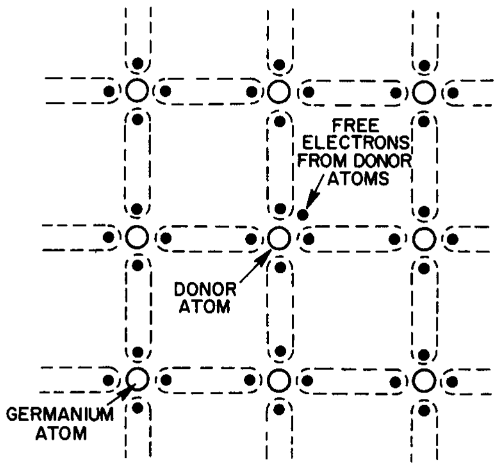


Fig. 13. Addition of impurity with atoms of 5 valence electrons to an intrinsic semiconductor with atoms of 4 valence electrons.

negative free electrons. So, the crystal is called an *n-type* semiconductor. Of course, not all the current is carried this way. There are still thermally-generated electron-hole pairs from the millions of other germanium atoms. The free electrons contributed by the donor atoms, however, far exceed the number of holes. Hence the *electrons* in an n-type semiconductor are called the *majority carriers* (of electric current), while the *holes* are called the *minority-carriers.* Electrons and holes are called *mobile* charges, as opposed to the immobile ions.

If an impurity material were added to an intrinsic germanium crystal whose atoms had one less valence electron than the germanium atom, namely three, a situation would result as illustrated in Fig. 14. The fact that this impurity has a missing place—a hole—where an electron should be is analogous to the situation that resulted when an electron-hole pair was created in an intrinsic semiconductor. An electron from a neighboring germanium atom may easily break its covalent bond and jump into this hole. We say the hole accepts an electron; thus this impurity atom is called an *acceptor* atom. Note that the material will now have a predominance of positive hole current carriers. So, the crystal is

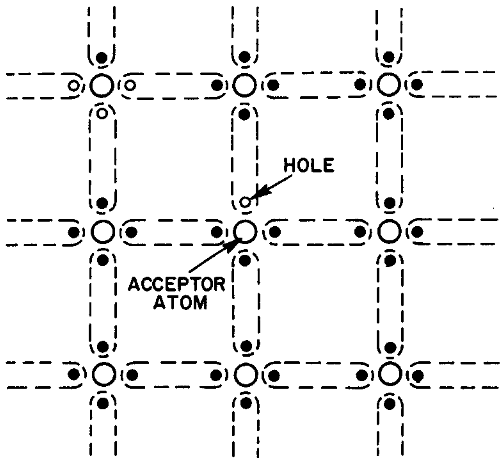


Fig. 14. Addition of impurity with atoms of 3 valence electrons to an intrinsic semiconductor.

called a *p-type* semiconductor. Again, because of thermal agitation, some electron-hole pairs will be present. Of course, the holes will far outnumber the free electrons, and so the *holes* are called *majority carriers,* and the *electrons, minority carriers.* Typical acceptor impurities are gallium, indium, and boron.

The process of adding impurity atoms to a semiconductor is called *doping.* It is a very critical process, calling for the utmost control over the amount of impurity added. It is not hard to visualize what would happen if an overabundance of impurities were added to a semiconductor. For example, if intrinsic germanium were to be overdoped with n-type material, so many free electrons from the donor atoms would permeate the crystal, that the device would undoubtedly act as a conductor.

It is interesting to note that, when an intrinsic semiconductor is doped with donor atoms, not only do the number of free electrons increase (at a given temperature), but the number of holes, which normally would be present because of thermal agitation, decreases. This happens because, with so many free electrons available, the recombination rate is higher. In other words as soon as a hole is created from the thermal agitation of a germanium atom, there is always nearby a free electron to fill in the hole. This recombination results in the death of the hole and the free electron, and the re-establishment of a covalent bond. But, since there are so many more free electrons than holes, the decrease in the number of electrons is hardly noticeable. Of course, the opposite reasoning applies when intrinsic germanium is doped with p-type impurities: the number of holes increases, and the number of electrons decreases.

Even a minute quantity of impurity added to an intrinsic semiconductor exerts an influence over its conductivity. In an intrinsic germanium crystal there are 4.52 x 10^{22} atoms per cubic centimeter. If to this are added 2.2 x 10^{14} impurity atoms per cubic centimeter, the chemical difference caused by the addition of the impurity could hardly be detected, since only one impurity atom is added for every 200 million germanium atoms. However, if each of these impurity atoms contributes a free electron, the conductivity of the crystal would increase by a factor of, approximately, seven.

At room temperatures, it can be assumed that nearly all of the impurity atoms of a doped semiconductor will ionize and, therefore, liberate holes or electrons for conduction. At these temperatures, the number of holes or electrons created because of the addition of the impurity, far exceed the number of thermally-generated electron-hole pairs. If, however, the temperature is raised high enough, and the number of electron-hole pairs created far exceeds the number of impurity holes or electrons, the total number of electrons will approximately equal the total number of holes. For example, if the temperature of the doped germanium crystal referred to above is raised enough, every germanium atom would be ionized. In fact, every covalent bond would probably be broken, liberating four electron-hole pairs per atom, or about 20 x 10^{22} electrons and holes per cubic centimeter. Whereas, at room temperature, the electrons contributed by the donor atoms far exceeded the number of holes thermally generated, now essentially the opposite is happening. In fact, there are so many more electron-hole pairs thermally that the additional electrons contributed by the donor atoms hardly affect the ratio of the number of holes to the number of electrons. In the example cited, the new ratio of holes to electrons is approximately 1/1.000001. Consequently, at high enough temperatures, the number of holes will roughly equal the number of electrons—regardless of the type or amount of doping. This equality of holes and free electrons was previously attributed to an intrinsic semiconductor. Consequently, a doped semiconductor is said to have "gone intrinsic" at high temperatures. This effect places a definite upper temperature limit on a semiconductor's operation. Practically, germanium semiconductor devices are usable up to about 85°C, whereas silicon devices will operate up to, and sometimes over, 200°C. The reason for this difference (in the upper operating temperature between the two materials) is that it takes about 1.25 electron volts of energy to break a silicon covalent bond, but only 0.75 electron volt to break a germanium covalent bond. Therefore, at high temperatures—where a large amount of thermal energy is supplied to the crystal—more germanium covalent bonds are likely to be broken than silicon covalent bonds.

In addition to a high temperature limitation, n-and p-type materials have a low operational temperature limit. When the temperature is so low that the thermal energy supplied to the impurity atoms is insufficient for their ionization, the material effectively becomes an intrinsic semiconductor. This effect is so important that for low temperature operation, heaters are often supplied to semiconductor devices.

9. Review Questions

1. If oxygen has an atomic number 16, how many protons does it have in its nucleus? How many electrons are there in the neutral oxygen atom?
2. What is a photon? How is our daily life dependent upon it?
3. If a photon has 12 electron volts (ev) of energy, and it strikes an electron in an atom whose ionization potential (energy needed for ionization) is only 10 electron volts, does the photon ionize the atom? If so, where do the extra 2 electron volts go?
4. Do electrons have any energy at absolute zero? How do the classical and modern theories differ on this point?
5. Why do conductors have positive temperature coefficients, and semiconductors negative temperature coefficients?
6. Although holes and electrons move in opposite directions under the influence of an electric field, why does this result in a current in one direction?
7. What are the conditions under which a doped semiconductor appears to be an intrinsic semiconductor? Does this effect vary among different semiconductor materials?
8. What happens to a hole if it collides with a free electron?
9. Is it possible to know the exact energy of any given electron at a particular time?
10. What would be the outcome if an intrinsic semiconductor crystal were doped simultaneously with donor and acceptor atoms?

Chapter 2

THE P-N JUNCTION DIODE

10. The Semiconductor Rectifier

Alone, p- or n-type semiconductor crystals conduct current in either direction. However, when a p- *and* n-type crystal are placed together, the combination exhibits the important property of rectification; that is, the property of conducting current in one direction only. Because the property of rectification is dependent upon the *junction,* or interface between the p- and n-crystals, semiconductor rectifiers are often called *p-n junction diodes.*

11. The P-n Junction Diode Under Equilibrium Conditions

Let us place a piece of hypothetical n- and p-type germanium material together. The doped semiconductor materials are schematically illustrated before contact by Fig. 15 (A). (For the present, we will not consider thermally generated electron-hole pairs from germanium atoms. Hence, only holes, electrons, and impurity ions are shown.) Before ionization, each neutral donor atom in the n material still possessed its extra electron. Each atom may also be represented as a donor ion with a + 1 charge, its extra electron possessing a − 1 charge. The net charge on each atom is still zero, even though we show the ion and electron separately.

Similarly, in the p material, each acceptor atom can be shown as an ion with a − 1 charge, its hole with a + 1 charge. This method of illustration is advantageous because now a neutral and an ionized atom may be treated with the same symbols. In Fig. 15A, *some* of the donor atoms are ionized (due to thermal energy)

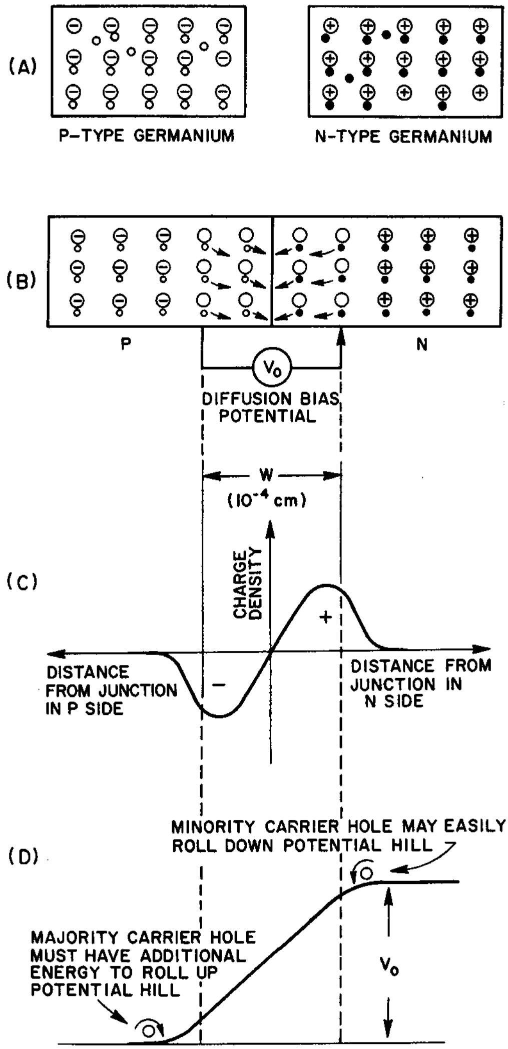


Fig. 15. A. P-type germanium and n-type germanium; B. The p-n junction diode. C. Charge density distribution. D. Electrostatic potential variations.

and the **electrons** that have been released are freely moving about. In the p-type material, the reverse is happening: some **holes** have been liberated and are roaming around the crystal.

Figure 15A is a highly stylized illustration and does not portray the exact atomic picture of either crystal. However, this analysis allows a meaningful explanation of the p-n junction action. For instance, almost all the impurity atoms are ionized, even at room temperatures. Yet in the illustration, only a few electrons and holes are shown separated from their respective parent atoms (ions). This is a seemingly incorrect representation. However, although almost all the impurity atoms are ionized, both crystals are homogeneous (or even) in charge density. That is, in the n-type crystal, the number of free electrons in any given area of the crystal is about the same. Assuming the donor ions are evenly distributed throughout the crystal (they are), then in any given region of the crystal the net electrical charge is zero; *i.e.,* the total negative charge of the free electrons in this region effectively cancels the total positive charge of the donor ions in the same region. Therefore, the crystal has a neutral charge throughout its structure, and we illustrated it as essentially containing neutralized donor ions, or just plain neutral atoms. Physically, this assumption may be questioned; electrically, it is true.

Because the thermally agitated free electrons wander about, there will always be some variation in charge density. At some points in the crystal there will be momentary concentrations of free electrons. To account for this slight concentration of negative charge, some of the electrons in Fig. 15A are shown separated from their parent donor atoms.

In the p-type crystal (Fig. 15A), we have also simplified an involved process. Here, however, holes were shown to move in the space between donor atoms. This, too, is justified. Holes may skip from germanium atom to germanium atom, and these atoms are not shown. A hole does not roam around the crystal exactly as a free electron does, because a hole is not a physical entity. A hole is only the absence of an electron in a covalent bond. But, when we view large sections of the crystal at once, there are so many holes, and the atoms are so close together, that little error is introduced when we assume that holes can move freely. It may be said (using the same logic we applied to the n-type crystal) that the charge density everywhere is approximately zero. Or, as with free electrons in the n-material, the concentration of positive holes in any given region is about equal to the concentration of negative acceptor ions. The small variations in this net zero charge density are indicated by the presence of a few holes, separated from their parent atoms.

What would happen if the n- and p-type crystals were placed together, without an external electrical connection? On the right side of the junction between the two materials (see Fig. 15B) there is a concentration of free electrons; on the left side, a concentration of holes. If the device has been fabricated correctly, the junction between the two materials does not represent a discontinuity of the atomic crystal structure and the two materials together represent one homogeneous figure. Consequently, since there is a high concentration of free electrons on one side of this junction, and an extremely small concentration of free electrons on the other side of the junction, the free electrons will tend to drift across the junction. This type of thermal drifting is called *diffusion* and is not ordinarily encountered in conductors.

Electrons tend to diffuse, or drift from points of high concentration to points of lower concentration, in the same way as the odor from an uncapped bottle of ammonia diffuses throughout a room free of moving air current. Electron diffusion across the junction occurs because this is the only direction that the thermally agitated electrons may take without almost immediately colliding with other free electrons. Much as water seeks its own level, these electrons (if left to themselves) will spill across the junction and create an even concentration throughout the crystal. The holes, too, will diffuse across the junction from left to right.

When an electron diffuses across the junction, it leaves its positive ion and deposits its own negative charge to the p material. Similarly, a hole diffusing across the junction adds its positive charge to the n material and leaves a negative ion. The electrons and holes that scatter across the junction do not diffuse very deeply into their respective materials. An electron crossing the junction almost immediately combines with a hole that (up to that point) was neutralizing a negative acceptor ion, assuming the hole had not already diffused across the junction. If it had already diffused, the electron would probably combine with some other hole in the immediate vicinity. A similar argument applies to hole diffusing to the n material. A hole may meet an electron while crossing the junction. The collision results in the inevitable annihilation of both carriers.

Before this electron-hole recombination, the charge density was zero in each side of the crystal. Now the ions near the junction in both materials have been stripped of their neutralizing holes or electrons and, therefore, possess a charge. These ions are called *uncovered ions.* The charge density distribution across the diode (see Fig. 15C) is no longer zero. Note that in the immediate vicinity of the rows of uncovered ions, the charge density is at a maximum. It is negative in the p material and positive in the n material. There is a smooth transition from negative to positive charge density across the junction region. Of course, the rest of the crystal has a neutral, or zero charge density.

As soon as ions are uncovered in the vicinity of the junction, an electric field is created between these oppositely charged ions. An electric field always exists between any two points of charge density variation. Visualize a fictitious generator across the junction region, with a potential V_o representing the net potential drop across the field. Since the electric field is created because of the diffusion of holes and electrons across the junction, the voltage existing across the field is called the *diffusion bias potential.* On most diodes, the diffusion bias potential is about 0.2 to 0.3 volt. The electric field, although created by electron-hole diffusion, has a repelling effect on the further diffusion of these electrons or holes across the junction. Obviously, the positive ions in the n material will repel any positive holes attempting to cross the junction. Also, the negative ions in the p material constitute an effective barrier against further migration of free electrons across the junction. Therefore, the space between the donor and acceptor uncovered ions is called the barrier, or *space charge region,* and its width is denoted by the symbol W. This space charge region is about 10^{-4} cm wide.

Figure 15D illustrates the electrostatic potential variations across the device. Any hole attempting to cross the barrier region must travel up an electrostatic hill of a height equal to the diffusion bias, V_0. In simpler terms, the positive hole must travel toward the uncovered positive donor ion IN SPITE of the repellent force existing between these two positively charged bodies. The "hill" terminology is imaginary and helps get across the idea of a hole climbing a hill against this symbolic gravity, which is caused by the repellent force. It cannot do this unless an additional force is applied.

Similarly, an electron must travel up a potential hill of the same magnitude when crossing to the p side. Since the electron is negative, Fig. 15D more correctly illustrates the picture of a "hill," when it is visualized upside down.

Thermally generated electron-hole pairs from the germanium atoms in the diode will not produce a net current flow across the junction. In the n material, electrons that are thermally generated from the breaking of germanium covalent bonds act just as other free electrons in this material. Most of the holes that are thermally generated will immediately disappear by recombination with the free electrons that are present. The few holes that escape annihilation diffuse over to the junction and fall down the potential hill, as shown in Fig. 15D. This hole current is precisely balanced by another hole current going in the opposite direction, which results when a few holes (majority carriers in the p side) have enough thermal energy to climb the potential hill. The height of the potential barrier automatically adjusts itself to insure that this happens. Consequently, since both hole currents have the same magnitude and travel in the opposite direction, the net hole current is zero.

In the p side, those thermally generated electron minority carriers that escape recombination fall down the potential hill to the n side. Similarly, this electron current is exactly balanced by another (resulting from some electrons in the n side having enough thermal energy to climb the potential hill and go to the p side) and, therefore, the net electron current is also zero.

The above could have been deduced by simple logic. Without an external electrical connection to the crystal, there could never be a continuous current flow across the junction. Electric current must flow in a loop from the generator to the load, and back again. Therefore, it is obvious that, for whatever the reason, the net electron and hole currents across the junction must be zero.

12. P-n Junction Diode in Reverse Bias

To have the property of rectification, the p-n junction diode must conduct current in one direction only. Let us first consider how the diode becomes practically an insulator current flow.

In Fig. 16A, our p-n diode has been connected to a battery, with the positive battery terminal connected to the end of the n material and the negative battery terminal connected to the end of the p material. (We assume that the ends of the diode are solidly connected to a metallic end plate of negligible resistance.) This method of connection—positive battery terminal to n material and negative battery terminal to p material—is called *reverse,* or *blocking bias.* We shall show that when the diode is reverse biased, it has a maximum resistance to the flow of current.

With the application of reverse bias, a field is created between the end plates. This field is in a direction such that the positive plate attracts the negative free electrons in the n material, and the negative plate attracts the positive holes in the p material. With this crowding of the majority carriers at either end of the diode (as shown in Fig. 16A), the region of uncovered ions increases in width. Figure 16B shows how, by broadening the area of high charge density, this broadening of the space charge, or barrier region, is accomplished.

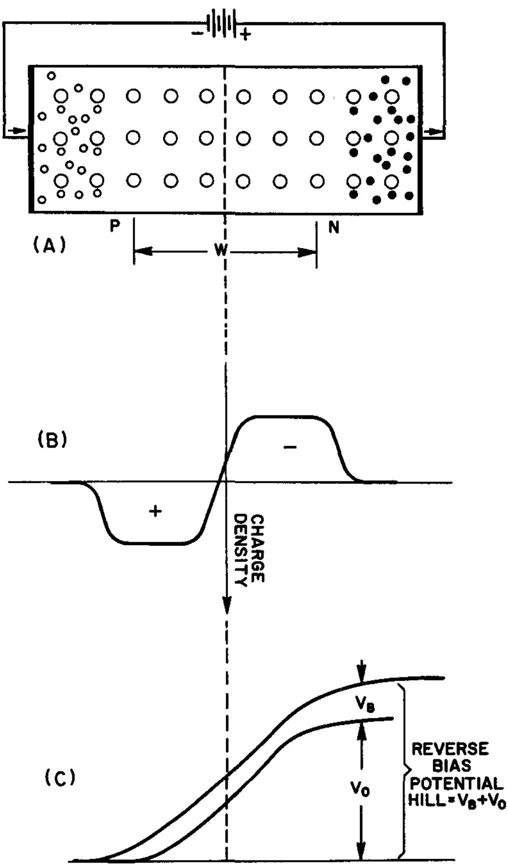


Fig. 16. A. P-n junction diode reverse biased. B. Charge density distribution. C. Electrostatic potential variations.

A comparison of the polarity of this electric field (between the end plates of the diode) with the polarity of the natural diffusion bias field of Fig. 15B shows that they are in the same direction and aid each other. In other words, the natural diffusion bias, V_0, across the junction has been effectively increased by the amount of the reverse bias battery potential E. The new potential hill created is shown in Fig. 16C. The hill is now much higher than before the application of reverse bias. Because of this high hill, the majority carrier holes in the p side will probably never have enough thermal energy to climb, or surmount, the barrier. Moreover, the electrons in the n side will not be able to climb this hill and diffuse to the p side. Consequently, even with the application of a slight amount of reverse bias, conduction by majority carriers cannot take place across the junction. Only 0.1 or 0.2

volt of reverse bias is needed to almost completely stop this majority carrier migration across the junction.

If the action of a diode were based entirely on the impurities contained in the material, then a reverse-biased diode would be a perfect insulator. However, at anything but very low temperatures, electron-hole pairs are still generated. The majority carriers created in this way (*i.e.*, the electrons thermally generated in the n side, and the holes similarly produced in the p side) will behave just as the other majority carriers do. They do not contribute to a conductive current.

The thermally generated minority carriers, however, present a problem. For instance, thermally generated minority carrier holes in the n material will be repelled, instead of attracted to the positive plate. They will diffuse to the junction and fall down the potential hill. Of course, recombination will annihilate many of these holes long before they ever reach the junction. The fact that the potential is much higher than before the application of reverse bias is unimportant. A hole will roll down the hill, so to speak, regardless of its height. The holes that escape recombination and cross the junction to the p side migrate to the negative plate. The free electrons from this conductive plate will combine with the holes. Hence, thermally generated holes in the n material give rise to an electron current, through the connecting wire from the battery.

This analysis applies equally to the thermally generated electrons in the p material. The electrons left after recombination will cross the junction by falling down the potential hill. Then they will drift over to, and enter, the positive conductive plate, thus causing an electron current in the wire. Minority carrier generation, in both n and p materials, causes an electron current in the wire that moves in the same direction. (See Fig. 15A.) Essentially, for every hole that combines with a free electron from the left end plate of the p material, a free electron enters the right hand plate of the n material.

Now if, starting from zero, in Fig. 16A, the reverse bias battery voltage were increased, this small reverse bias current would advance to a certain point, and then stop. The limiting case would be reached, when all of the thermally generated minority carriers contributed to the conduction current. Since the amount of minority carrier generation is wholly dependent upon temperature, if the bias voltage were increased beyond that point, there would be no further increase in current. For this reason, the reverse bias current is called the *reverse saturation current,* and is designated by the symbol I_0. I_0 reaches its maximum value at a reverse bias voltage of 0.1 or 0.2 volt. (This effect is analogous to the saturation of plate current in a vacuum tube. At the point where the plate voltage is so high that all of the electrons thermally emitted from the cathode are attracted to the plate, a

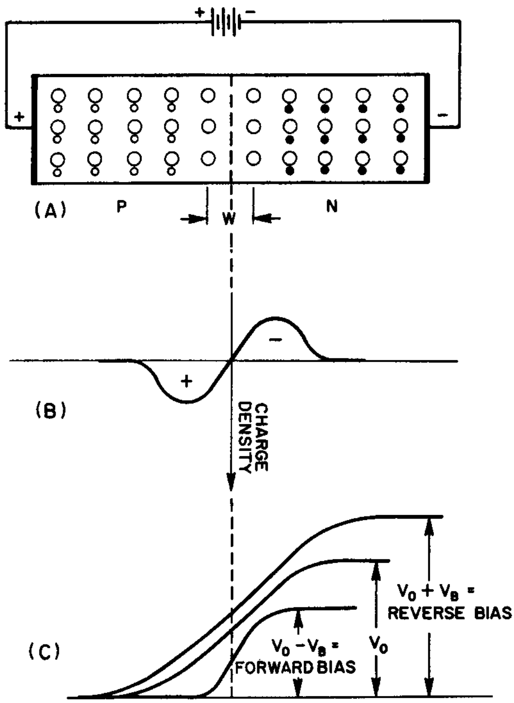


Fig. 17. A. P-n junction diode forward biased. B. Charge density distribution. C. Electrostatic potential variations.

further increase of plate voltage cannot produce more electrons. Hence, the current is the same with increasing plate voltage beyond the point of saturation.)

The performance of a rectifier depends greatly upon its ability to retard the flow of current in one direction. In a vacuum diode, this reverse current is zero, because electrons never travel from plate to cathode. In a semiconductor, however, we have reverse saturation current. At room temperatures, so few electron-hole pairs are generated from the breaking of covalent germanium bonds that the reverse saturation current is in the order of a few microamperes. For all practical purposes, the diode may be considered an insulator. Unfortunately, I_0 varies radically with temperatures. (I_0 for germanium, increases about 11% per degree centrigrade in temperature.) While at room temperatures I_0 is of little consequence, it may become serious at high temperatures. Stability problems in high temperature transistor operation are caused by the heat of this current.

13. P-n Junction Diode in Forward Bias

When the battery across our diode is reversed (the positive terminal is connected to the end of the p material and the negative terminal is connected to the n material), the diode is said to be *forward* biased. Biased in this direction (see Fig. 17A), the field across the diode is of a polarity such as to repel the majority carriers. That is, the free electrons in the n side are repelled from the negative end plate and forced toward the junction. Morever, the holes in the material will be repelled from the positive end plate and will crowd together near the junction. It is easy to see now that the width of the region of uncovered ions is much less. Essentially, the application of forward bias lowers the height of the potential barrier. When Fig. 17A is compared with Fig. 17B, it is seen that the polarity of the applied forward bias is opposing the natural diffusion bias potential, V_0. This is shown in Fig. 17C, wherein the potential barrier height is illustrated for the forward, reverse, and unbiased condition. Using our previous analysis for the reversed and unbiased diode, the lowering of the potential hill implies that many holes in the p material now have enough thermal energy to climb the low potential hill. Also, the electrons in the n material will rush over to the p side. The small potential barrier offers little opposition to this boiling mass of mobile charges.

As the electrons cross the junction and enter the p material, they meet holes going in the opposite direction, towards the n material. Some of these electrons will disappear due to recombination. Others will penetrate quite deeply into the p material, before recombining with the many holes always present on this side of the junction. Eventually, due to recombination, all of the free electrons injected into the p material will disappear. At a distance far from the junction in the p material, the current is composed entirely of holes. (The relationship between total, hole, and electron current across a p-n diode is shown in Fig. 18.)

Notice that in the n material, the holes that rush across the junction immediately start to recombine and, therefore, the hole current in the n material gradually decreases until it disappears entirely. Since doping symmetry exists in this diode (each material has been equally doped), the same situation results in the p material. We can conclude that, at great distances from the junction (a large distance compared with W), the forward current will be carried by the appropriate majority carriers. **The process whereby majority carriers from one type of material cross the junction and move into a different material, where they become minority carriers, is called *minority carrier injection.*** It is a very important principle in the action of transistors.

Let us trace the path of current around the circuit. Starting from the positive electrode of the battery, free electrons drift

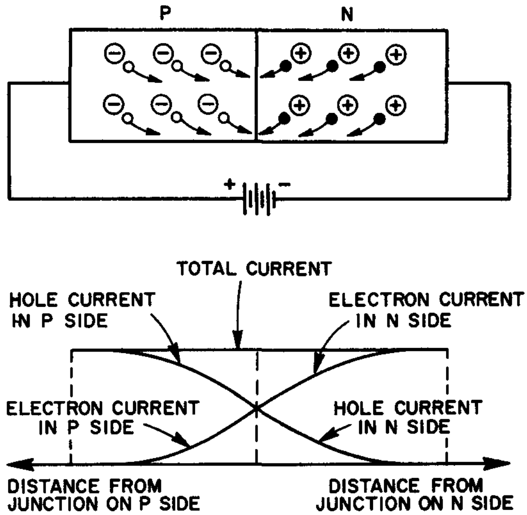


Fig. 18. Current carrier in the p region is by holes; in the n region by electrons.

through the conductive wire and enter the conductive end plate of the p material, as shown in Fig. 18. The free electrons then leave the end plate and enter the semiconductor, where, almost instantly, they combine with holes and disappear. On the right hand side of the diode, the majority carrier free electrons enter the conductive plate and proceed down the wire toward the battery. The electrons (through an electrochemical action) move through the battery. This completes the circuit. The current entering the diode must be exactly the same as the current leaving the diode. So, the total current must be the same at all points in the diode (Fig. 18) where if the hole and electron currents are linearly added at any point, they always total the same overall current.

If the forward bias potential directly opposes the natural diffusion bias and tends to reduce the barrier potential—"Would the resistance of the diode be reduced to zero if the forward bias were adjusted to equal this diffusion bias? And wouldn't this zero resistance make the forward current extremely large?"

For two reasons the answer to both questions is NO. First, the semiconductor and end plate materials themselves have a certain resistance. Second, with any appreciable forward current through the diode, a voltage drop will occur across this resistance, and the full bias potential will not be reflected across the junction. Hence, when the external forward bias potential reaches a magnitude com-

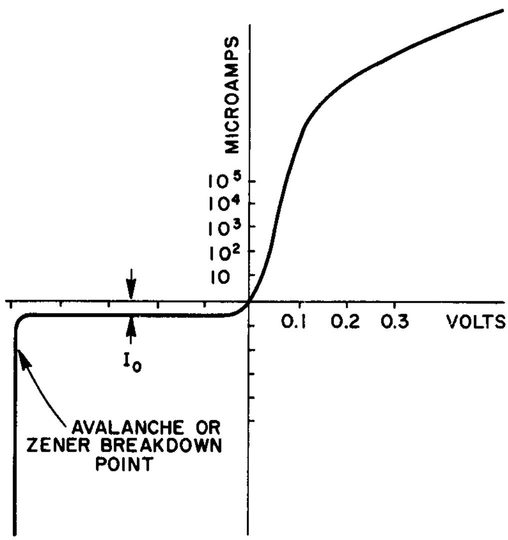


Fig. 19. Typical current voltage characteristics of p-n junction diode showing Zener breakdown.

parable to the diffusion bias potential, V_0 (about 0.1 to 0.3 volt), the forward current becomes limited by the above-mentioned resistances. This can be seen from the volt-ampere characteristics of the p-n diode of Fig. 19. Here, as the forward bias increases from zero to about 0.1 volt, the forward current increases at an exponential rate. When about 0.2 volt is reached, the increase in current becomes a linear function (straight line) of the forward bias. The straight portion is caused by the resistance of the material. It cannot be assumed that the current may become arbitrarily large. Each manufactured diode has a maximum permissible forward current, which is based on the ability of the device to dissipate the heat generated in its resistance. A good practical rectifier diode has a back-to-forward resistance of 100,000:1.

In the discussion of forward bias, we have so far neglected a consideration of the reverse saturation current. Now even if this current did exist, we could not properly call it by the same name, for the diode is not reverse-biased. Regardless of the bias situation, however, the electron-hole pairs are still generated; they are dependent only upon the temperature. As with reverse bias, these thermally generated holes in the p side act like holes already present. Electrons thermally generated in the n side act the same as the other majority carriers. The minority carriers that are thermally generated will be attracted to the end plates, but will probably never get there. On the way, they will combine with the ever-present majority carriers. The forward current of a diode is usually thousands of times greater than this minute thermal current and, therefore, electron-hole pairs generated during forward conduction are of little consequence.

14. Zener Breakdown Voltage

It was shown previously that the reverse saturation current would be constant with an increase of reverse bias. This is not strictly true. As the bias is increased, the thermally-generated electrons in the p material rush to the junction faster and faster. They are accelerated by the field that exists across the region. If the reverse voltage is made large enough, these electrons may collide with other bound electrons and knock them out of their covalent bonds. These newly created electrons may gain enough speed to knock still more electrons out of covalent bonds. Each new electron created in this method may liberate many more electrons. The cumulative process is called *avalanche multiplication,* and results in a very large reverse current. In this state, the diode is said to be in *avalanche breakdown.* (See Fig. 19.)

Even if (owing to lack of sufficient velocity of the thermal electrons) avalanche breakdown did not occur, the electric field across the junction can become large enough, with a high reverse voltage, to directly rupture the germanium covalent bonds. The plethora of new free electrons created in this way amounts to the same high reverse current as with avalanche multiplication. Although this reverse voltage breakdown may occur from two distinct causes, this phenomenon is usually referred to as the *Zener breakdown,* after the man who first explained how covalent bonds could be broken with a high electron field. The voltage at which breakdown occurs is termed the Zener voltage, or the *Zener breakdown voltage.*

It is significant to note that once the reverse voltage has been reduced below the Zener value, the diode will recover; that is, the reverse current will once again consist only of the reverse saturation current. Of course, if the current during breakdown is too great, the junction may be permanently damaged. It has been found that, during manufacture, it is possible to control the magnitude of the Zener voltage. Special "Zener" diodes are manufactured as voltage regulators, having breakdown voltages varying from about 10 to over 400 volts. The diode is valuable as a voltage regulator because, during breakdown, the reverse voltage is nearly constant for a wide variation of current. This can be seen from the almost vertical aspect of the breakdown part of the diode volt-ampere characteristic in Fig. 19.

15. Diode Types

There are over 2000 different types of semiconductor diodes currently available. They vary in size, power handling capacity, Zener breakdown voltages, forward and reverse current, temperature limitations, package dimensions and style, etc. One can always consult a diode catalogue for specific needs; however, a general knowledge of some of the characteristics of the most popular diode types is advantageous.

The ***point contact diode*** is about the oldest type of semiconductor device. The earliest radio receivers were "crystal sets." The diode in this radio was a piece of galena (lead sulfide) crystal. The body of the crystal served as one connection, and a "cat's whisker," or fine wire probe, served as the other. To make the diode work, the cat's whisker was carefully placed against the surface of the crystal until a spot with the proper impurities was found. Both the diode and the receiver were so simple that home radio receivers were constructed by the thousands, as soon as the first radio programs went on the air. Since there was no way of knowing the precise location of the impurity atoms in the crystal structure (in fact, nobody even knew how it worked), exasperating hours were often spent searching the face of the crystal with the cat's whisker for the most sensitive spot (where the station seemed to be received the loudest).

Today, the construction of point contact diodes is a fine art. Figure 20(A) shows a point contact diode. The cat's whisker is still present, but is now permanently attached to the most sensitive spot. To create a junction, a surge of current is passed through the cat's whisker that has been attached to an n-type germanium slab. During the short period of current flow, a small

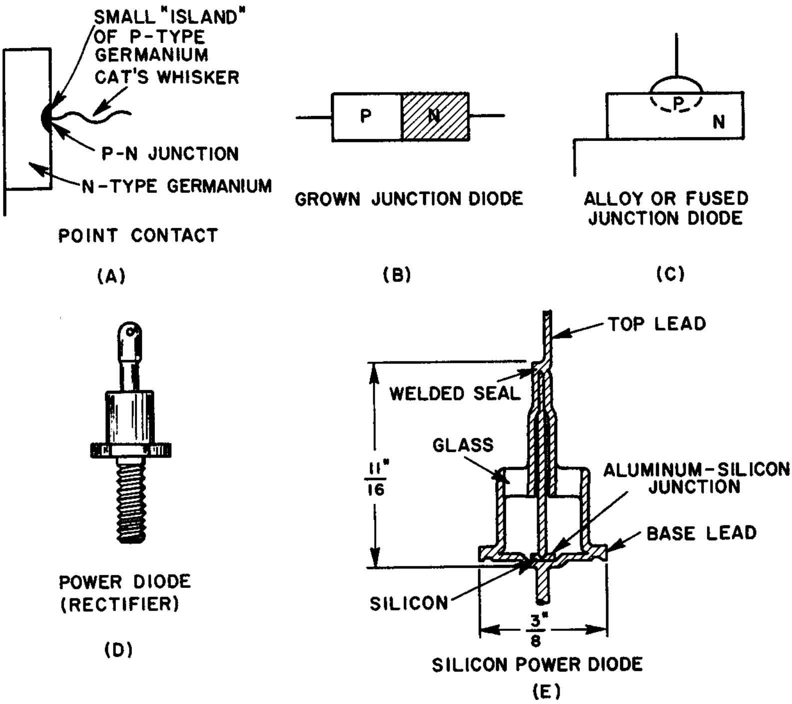


Fig. 20. Most popular diode types: (A) point contact; (B) grown junction; (C) alloy or fused junction; (D) power rectifier; (E) silicon power.

dot or "island" of p-type germanium is found where the wire makes contact with the crystal. Thus, a p-n junction is formed within an extremely small area. Typical forward currents of 4 to 11 ma with a forward bias as small as 1 volt are achievable. The diode is able to withstand reverse voltages of about 250 volts.

Germanium point contact diodes are used for relatively low power applications such as video second detectors, computer flip-flops, and low-level power supplies. Silicon point contact diodes are often used for microwave mixers. However, they are more fragile than the germanium kind and will stand only a few volts in the reverse direction. Strictly speaking, a point contact diode is a junction diode, because a junction exists. However, because the junction is so small, it is usually referred to merely as a point contact diode. (The term junction diode usually refers to other types of diodes.)

A *grown junction diode* (Fig. 20 (B)) is prepared by adding to pure molten (intrinsic) germanium an impurity of the proper kind to make it either a p or n type. With the temperature controlled precisely, a crystal starts to form, or grow about a tiny seed crystal. As the doped crystal starts to form, it is slowly withdrawn from the molten mass of germanium, called the *melt*. At the proper point, the melt is doped with a different kind of impurity, so that the crystal grown thereafter is of the opposite type.

An ***alloy*** or ***fused junction diode*** (Fig. 20 (C)) is prepared by placing a small dot or pellet of acceptor impurity, such as iridium, on one surface of a wafer of n-type germanium. When this combination is fired to the proper temperature, the iridium fuses into the germanium wafer. Thus a p-n junction is formed between the iridium pellet and the germanium wafer.

In the ***diffused junction diode,*** a p-n junction is formed by exposing a piece of n-type germanium or silicon to a gaseous p-type gas. The gas diffuses through the crystal lattice and, where the diffusion stops, a junction is formed.

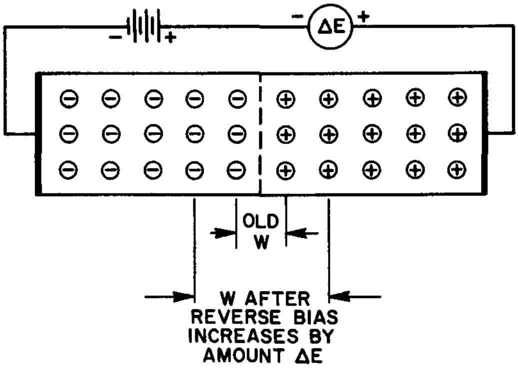
POWER DIODES refer to diodes used in rectifiers delivering an appreciable amount of power to a load. Figure 20 (D) is a picture of a typical stud-mounted power diode that can carry an average forward current of more than 5 amperes. Fused and diffused construction are used. Typical units can withstand inverse peak voltages up to one thousand volts. Much of the compactness of these diodes is due to the fact that with such a low forward resistance, little power is dissipated in the unit. Hence, large heat radiating fins are not needed.

For high temperature operation (above 85°C) ***silicon power diodes*** are used. Figure 20 (E) shows an actual size cross section of a silicon junction power diode. Typical diodes with this case size can handle 750 ma at temperatures up to 200°C.

16. Junction Capacitance

The barrier, or space charge region, W, in Figs. 15C, 16A, and 17A is important in still another respect. If a given reverse bias is applied to a diode, a certain space region will result from a given width of uncovered donor and acceptor ions. If the reverse bias is increased slightly by an amount ΔE, then the width of the uncovered ions grows slightly larger by an amount ΔW. It does this because, with an increase in reverse bias, the end conducting plates of the diode have a slightly greater attraction for the majority carriers. Therefore, they remove these mobile charges from the extremities of the space charge region. This can be seen in Fig. 21, where the outer row of holes and electrons are

Fig. 21. The effect of barrier capacitance.

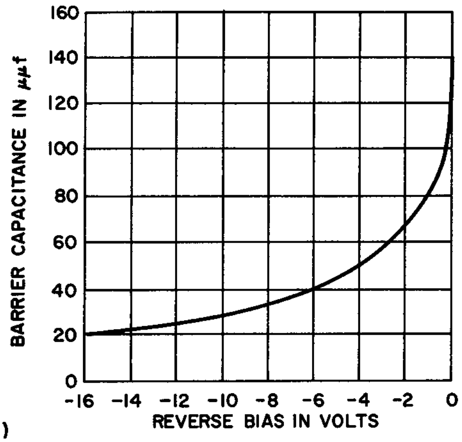


removed from their respective ions, and diffuse toward the ends of the crystal. Of course, this movement of charges is only momentary; the current stops as soon as the space charge width has adjusted to the new, higher value of reverse bias. It happens almost instantaneously. The time is limited, principally, by the resistance of the semiconductor material.

This phenomenon is precisely what happens within a capacitor, under the action of a change in applied voltage. For example, if the applied voltage across a capacitor is constant, no current flows. If, suddenly, the voltage increases a slight amount, there is a momentary surge of current, as electrons are removed from one plate and deposited on the other plate. Any device that exhibits this property must be called a capacitor. In a semiconductor diode, this capacitance of reverse bias is called the *barrier capacitance.* The barrier capacitance is proportional to the junction cross-sectional area, and inversely proportional to the space charge width W. Hence, as the reverse bias is increased, W increases, and the capacitance decreases.

All semiconductor diodes possess this barrier capacitance in some degree. Effectively, the capacitance is in parallel with the diode. Thus, during reverse bias at high radio frequencies, current is likely to flow through this capacitance, rather than being stopped by the high back resistance of the diode. This capacitance, then, places a high-frequency limitation on the use of the diode as a detector, or in some other application involving radio frequencies. The junction area of a point contact diode, however, is very small, and so its barrier capacitance is in the order of a fraction of a micromicrofarad ($\mu\mu f$). This is compared to a typical grown junction diode with a cross-sectional area of 1 square millimeter, and an accompanying barrier capacitance of 5–50 $\mu\mu f$. This low barrier capacitance of the point contact diode is the reason they are used almost exclusively in the uhf and microwave region.

Normally, barrier capacitance is a hindrance. Certain diodes, however, have been produced that take advantage of this capacitance. In fact, they are actually used as voltage variable diode capacitors. In manufacture, certain physical and chemical features affecting this capacitance are maximized, so as to create a diode having as great a change in capacitance, for a given voltage charge, as possible. Figure 22 (A) is a graph of a typical voltage variable diode capacitor. Notice that the capacitance decreases as the reverse bias increases. It is this control of capacitance that makes the diode useful as a replacement for a reactance tube in automatic



(A)

(B)

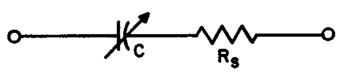


Fig. 22. (A) Graph of a typical voltage variable diode capacitor. (B) Schematic representation.

frequency control circuits, and similar devices. It has a greater change or capacitance for a given change in applied voltage than a reactance tube. And, the power necessary to make it operate is infinitesimal, compared to the heater and plate power required by a conventional reactance tube.

The voltage variable diode capacitor (marketed under the name of *Varicap,* among others) is not a perfect device. The resistance of the semiconductor material places an effective high-frequency limit on its use. Figure 22 (B) is a schematic representation of the device. The power losses in the series resistance, R_s, place an upper limit of several hundred megacycles on its use. In early 1959, Microwave Associates, Inc. began production of a much improved version of this diode capacitor, called the *Variactor.*

It had a series resistance approximately 100 times less than previously manufactured capacitor diodes. The upper frequency limit of this diode is still to be determined, but it is thought to be many thousands of megacycles.

Because of its nonlinear capacity *vs* voltage relationship and low loss, a new kind of amplifier became possible: the *parametric amplifier.* The principal advantage of this kind of amplifier (other than the fact that it is a solid state device operating in the microwave region) is that because of its low R_s, it contributes very little noise to the signal it is amplifying. This low noise parametric amplifier, which features the *Variactor* diode as the amplifying element, is now being used as an r-f amplifier in long distance radar installations, radio astronomy, and the like.

17. The Tunnel Diode

Another semiconductor diode recently introduced is the *Esaki Tunnel Diode.* The phenomenon of electrons traveling within certain energy levels in an especially prepared p-n junction was first observed by Leo Esaki of Japan. Since this tunneling effect became known, active research has been devoted towards improving the diode's constructional techniques, as well as understanding the theory of operation. At this writing, tunnel diodes no larger than a normal p-n junction diode are being used in experimental oscillators, amplifiers, detectors, etc., up to the 10,000-mc range. The startling performance of the tunnel diode on almost no power, low noise, and eventually, low cost, makes it a prime candidate for the replacement of the transistor in many fields.

18. Photoresistors and Photodiodes

When radiation falls on an intrinsic semiconductor crystal, its conductivity increases. This was emphasized earlier, when it was shown how high energy photons from bombarding light will break some of the germanium covalent bonds, and create electron-hole pairs. The resistance of the material decreases, as the number of free electrons and holes increases. Thus, an intrinsic semiconductor is really a light-sensitive resistor; more appropriately, a *photoresistor.* The dark resistance, or the resistance of the crystal in the absence of light, is in the neighborhood of four-five thousand ohms. The resistance may drop to below one thousand ohms with a strong light. Several manufacturers make photoresistors, although (owing to their low dark resistance) their popularity is not as great as the *photodiode.*

When light falls on one side of a p-n junction diode, new electron-hole pairs are generated along with the thermally generated electron-hole pairs. If the diode is under open-circuit conditions, then the light-generated minority carriers will diffuse across the junction. In this instance, light is acting as a minority-carrier injector. However, the net current in the diode under open-circuit conditions must always remain zero. Consequently, to cancel this minority current, majority carriers must flow across the junction in the opposite direction and at the same rate. The majority carriers must overcome the potential hill created by the diffusion bias, before they can diffuse across the junction. Hence, the barrier potential automatically lowers itself to allow sufficient majority carriers to cross the junction and cancel the light-induced minority current. Fundamentally, the barrier height has been reduced as a result of the light. A voltage will appear across the diode's terminals exactly equal to the amount by which the potential barrier has been reduced. If the external circuit is closed, then this external voltage will cause a current to flow. This process, which is the generation of electric power from the energy of incident light, is called the *photovoltaic effect.* The device is called a *photovoltaic cell;* more popularly, a *sun* or *solar battery.* Under short-circuit conditions, the cell will cause a current of up to 50 μa per millilumen. These sun batteries are currently being used to power uhf transmitters in America's Explorer satellites, as well as in portable radios, and many other devices.

When reverse bias is applied to a p-n junction diode exposed to light, the resulting electron-hole pairs generated increase the reverse current. Figure 23A shows the response characteristics of a typical photodiode, and Fig. 23B shows an overall and cross-sectional photodiode. Notice that a lens is provided to insure that entering light is concentrated on a small area of the crystal. So that a minimum amount of light-created electron-hole pairs recombine, as they diffuse to the junction after generation, the light must be shown on a spot very near the junction. The sensitivity of the device therefore decreases, as the light is shown on the crystal at points farther from the junction.[1]

Unfortunately, the sensitivity of a photodiode is largely dependent on the operating temperature. Both the light and dark currents

---

[1] Although photodiodes are constructed to maximize their light-sensitive characteristics, ALL semiconductor diodes are light-sensitive. The performance of a diode as a rectifier may be seriously impaired if light is allowed to enter the junction region. If the diode has a transparent encasing (many do), special precautions must be taken to insure that the unit is not exposed to light.

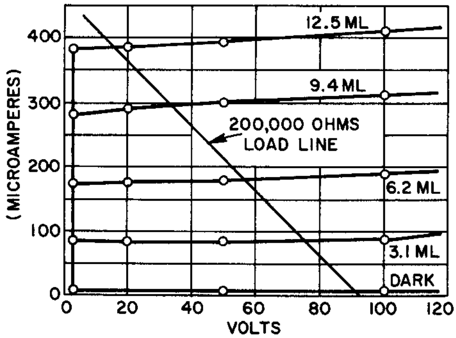


Fig. 23A. Average response characteristics of a photodiode.

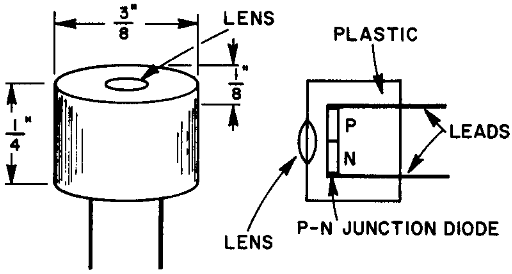


Fig. 23B. Two views of the p-n junction photodiode.

increase with a temperature increase. However, the light current increases at a greater rate than the dark current. Hence, the diode is more sensitive at higher operating temperatures. When the application requires a constant dark current over a wide temperature range, balanced bridge circuits are used to cancel out this dark current. Photodiodes are used in almost every instance where conventional photocells are used. These include headlight dimmers, motion picture sound systems, computer punch-card optical readout system, and many others.

19. Review Questions

1. What is the difference between diffusion and random drifting of thermally agitated electrons?
2. Describe the process that occurs when a piece of n and p crystals are placed together.
3. What is meant by uncovered ions? How are uncovered ions important in forming the space charge region?
4. Why is the word saturation contained in the term "reverse saturation current"?

5. If a semiconductor rectifier is to be used in a plate voltage supply for a radio, what are some of the important characteristics of the diode you must know before installing it?
6. What is the difference between avalanche and Zener breakdown? Is a diode likely to be damaged if this occurs?
7. What is meant by junction capacitance, and how does it occur?
8. Why are photodiodes more popular than photoresistors?
9. How are the characteristics of a semiconductor diode rectifier altered if light is allowed to enter the junction region?
10. What is the optimum position for light to shine on a photodiode? Why?
11. Explain the theoretical difference between a photoresistor, photodiode, and photovoltaic cell?

Chapter 3

INTRODUCTION TO THE TRANSISTOR

There are many ways to describe the action of a transistor. Here, the subject will be approached in terms of the material studied in Chapter 2. A theoretical description of a transistor can be viewed as merely an extension of p-n junction theory.

In Chapter 2, a reversed bias diode was almost an insulator to the flow of electric current. If any germanium covalent bonds were broken, due to thermal energy, the minority carriers thus liberated (electrons in the p side and holes in the n side) were instantly attracted to the junction by the strong electric field existing in this region. They fell down the potential hill, eventually entered the end plates, and contributed to a conductive current. Although this reverse saturation current is of little value (it is chiefly responsible for lowering the back resistance of the diode), the fact that a largely unused potential hill exists across the junction is intriguing. We have a hill, but nothing to roll down the hill. If we could obtain electrons and/or holes to roll down this hill, easily and inexpensively, then a useful device might be created.

The problem is somewhat analogous to one faced by Dr. Lee DeForest. He was confronted with a vacuum tube diode that conducted current quite efficiently in one direction. He was seeking a method by which he could easily control this flow of electrons from cathode to plate. His solution was to insert a wire mesh (called a grid) between the cathode and plate. When he

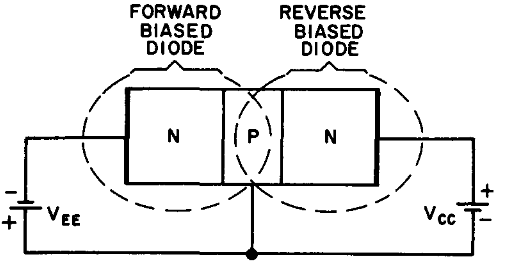


Fig. 24. Pictorial diagram of a junction transistor shown as a forward and reverse biased junction diode with common p material.

applied a small potential to this grid, he found that the intensity of the electron stream could be controlled. It took very little voltage variation on the grid to create a rather large voltage variation across a plate load resistor: an amplifier was created. To accomplish approximately the same thing with a semiconductor (*i.e.,* to create an amplifier), a method must be found to produce a supply of holes or electrons. If we could easily vary this supply, then the reverse current could be varied.

One method of creating electrons or holes is to shine a light on the junction region of a reversed bias diode. The intensity of the light could be varied in accordance with the information to be amplified. This is in essence a photodiode. A photodiode is excellent for its specific purpose. However, to require a light (with its accompanying lenses and power supply) to be an integral part of a semiconductor amplifier would not be advantageous.

Another way of controlling the reverse current is to vary the temperature of the diode. Unfortunately, due to the mass of the object, there would be a rather long delay between the time of application of heat and the time when the semiconductor reached the desired new temperature. This sluggishness is called *thermal capacity*. What is needed is a method instantaneous in action. Consequently, an all-electronic means of generating holes or electrons, and then injecting them into the junction region of a reversed bias diode.

A process called minority carrier injection takes place within a forward bias diode. (See Chapter 2.) In this process, the thermal energy given to electrons in the n material is sufficient to cause them to climb a relatively small potential hill. When they cross the junction to the p material, they become minority carriers. This is precisely the process we are looking for. The question remaining is: Can a forward bias diode be attached to a reversed bias diode in a way such as to insure that minority carriers created in this forward bias diode can be injected into the junction region of the reversed bias diode? In practice, this is done quite simply (see Fig. 24) by attaching both diodes to a common element. The resultant device is called a **junction transistor,** for it is made up of *two junction diodes having a common p material.* It is also called an **n-p-n transistor,** because the piece of p material is sandwiched between two pieces of n material.

The left-hand junction of this transistor forms a forward bias diode. When in forward bias (as evidenced by the polarity of the

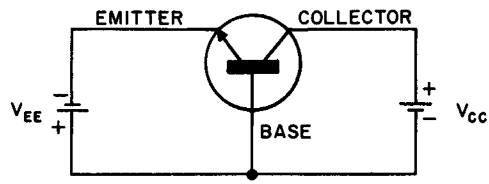


Fig. 25. Schematic representation of Fig. 24.

battery voltage, V_{EE}), minority carrier electrons are injected into the p region. For this reason, the left-hand forward bias junction is called the *emitter.* The right-hand junction is in reverse bias. If it were by itself—unattached to the rest of the transistor—the only reverse current would be the reverse saturation current. The strong electric field that exists across this junction, however, will attract or collect any minority carriers that are injected into this region by the emitter. Because the right-hand junction attracts minority carriers, it is called the *collector.* The region between the emitter and collector is called the *base.* Schematically, this transistor is illustrated in Fig. 25. The arrow on the emitter pointing away from the base indicates that this is an n-p-n transistor. It points in the direction that *positive current* flows when the emitter is in forward bias. Hence, in forward bias, electrons flow in the opposite direction of the arrow.[1]

So as to correlate the many things that are happening within an n-p-n transistor, the various electron and hole currents will be dealt with in detail.

The same n-p-n transistor is depicted in Fig. 26, where the emitter is shown in forward bias. Consistent with previous p-n junction theory, minority carriers will be injected across the emitter-base junction into the p material. Since it is this injection of minority carriers (electrons) into the base region that is the

---

[1] The direction of positive current (flowing from plus (+) to minus (−) in the external circuit) is purely an engineering convention. In this book, the direction of electron flow will be taken as the direction of the current.

sole purpose of the emitter, the efficiency of the system could be improved by making the free electrons carry the greatest proportion of the forward current in the vicinity of the emitter—base junction—instead of the 1:1 ratio between holes and electrons that normally exists in any symmetrically doped forward bias diode (see Fig. 18). This is done by doping the n material more heavily than the p material. It insures that on the right side of the emitter-base junction, almost all of the forward current will be carried by electrons. This is a very desirable situation because

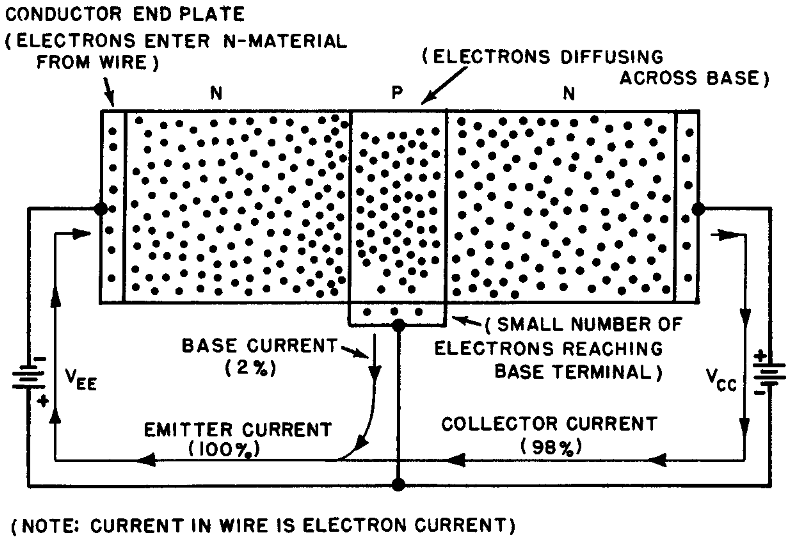


Fig. 26. The chief current carriers in an n-p-n junction transistor are electrons.

(with reference to Fig. 18) the electron current in the p material decreases exponentially as the distance from the junction increases, due to the recombination of these free electrons with the majority carrier holes. Therefore, the greater the concentration of electrons at the emitter junction, the greater will be the concentration of electrons at any nominal distance from this junction.

Now if these electrons were uninfluenced by the collector, they would proceed through the base material to the base terminal connection in normal junction diode manner. However, with a reversed bias diode in the same region, any injected minority carrier electrons that happen to drift, by diffusion to the right, and actually reach the collector will fall down a large potential hill. Consequently, two things tend to happen at once. The minority carrier electrons injected in the p material diffuse down to the base terminal connection (as would happen in any forward bias diode) while at the same time they diffuse to the right towards the collector. They diffuse toward the collector because there is a reduced concentration of electrons in this vicinity, *i.e.,* any electrons reaching the collector will immediately fall down the potential hill leaving the base side of the collector-base junction deficient in this type of mobile carriers. This diffusion of electrons to the collector is to be maximized in a successful junction transistor.

Unfortunately, although the base p material is not as heavily doped as the emitter material, millions of holes are still present. Some of these majority carrier holes are bound to combine with the injected electrons, as they diffuse across the base region to the collector. Whether there will be any electrons left to actually reach the collector is largely a function of the distance they have to travel through the base material. To reduce the number of electron-hole recombinations, and at the same time reduce the likelihood of too many electrons diffusing down to the base terminal, the width of the base wafer is made exceedingly small (about one-thousandth of an inch thick) with respect to its cross-sectional area. This practical solution assures that most of the injected minority carriers actually reach the collector (about 98%). The fraction of the emitter current transferred to the collector circuit is called the *current gain.*

The reader may ask: Can a transistor be operated backwards, with the collector and emitter interchanged? This can be done. The operation, however, would be at greatly reduced efficiency, since the transistor is designed primarily as a one-way device. For instance, the area of the collector junction is usually larger than the area of the emitter junction. The collector is made larger so as to offset the spreading effect of the free electrons as they cross the base region. If the collector were smaller in cross-sectional area, some of the free electrons might hit the surface of the base material or, more likely, diffuse to the base electrode connection. In addition, the emitter resistance is often made less than the collector resistance, to lessen the input power needed to create a given emitter current. For these reasons, transistors are usually manufactured asymmetrically. (Symmetrical transistors, with equal emitter and collector junction areas are made, but only for specialized tasks. An example of one application for a symmetrical transistor would be a transistor phase detector. Symmetrical transistors do not operate at maximum efficiency in either direction.)

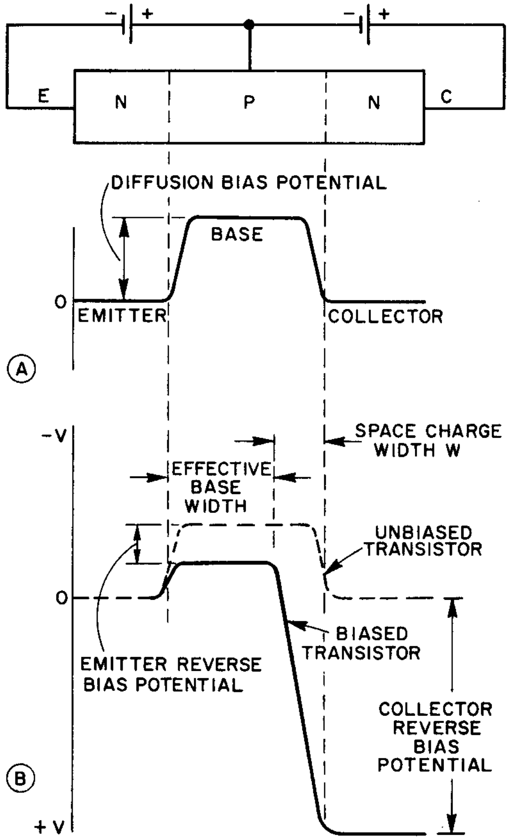


Fig. 27. Electrostatic potential variation across n-p-n junction transistor. (A) unbiased; (B) forward biased emitter.

The electrostatic potential variations across an unbiased n-p-n junction transistor are shown in Fig. 27. For the unbiased case, the up and then down potential hill is merely the potential hill of two unbiased p-n junction diodes placed back to back (this can be checked by referring back to Figs. 16A, 16B, and 16C). As was true in the discussion of diffusion bias, the potential barriers adjust themselves so that no net current will flow across either junction. If forward bias is placed across the emitter (Fig. 27), the emitter-base barrier potential hill will be lowered by the amount of the forward bias. With the application of reversed bias to the collector, the collector-base barrier potential hill will be raised by the amount of the reversed bias. From Fig. 27, it can be seen that electrons easily climb the small emitter potential hill, diffuse across the base region, and then fall down the high collector potential hill. Once the free electrons reach the collector and fall down the hill, they diffuse over to the collector terminal, and enter the connecting wire. Proceeding through the collector and emitter bias batteries, they enter the n semiconductor material, and the circuit is complete. The current flowing in the common, or base, lead in Fig. 26 is caused by several things. (1) The holes that combine with some of the injected electrons as they diffuse across the base region must be supplied from the base terminal (this is done by electrons entering the base terminal from the base p material). (2) The forward current of the emitter diode will contain some holes from the p region and these also must be created by electrons entering the base terminal from the base material. (3) The reverse saturation current (I_{CO}) of the reversed bias collector diode must also flow through this same base terminal lead. (This latter effect is so small as to be considered negligible at room temperature.)

Since the current entering the emitter is essentially equal to the current leaving the collector, one may believe that amplification cannot be achieved. This, however, is not true. The amplification property of a transistor depends on the fact that the emitter and collector circuits have vastly different resistances. In effect, current is transferred from the low-resistant emitter circuit to the high-resistant collector circuit. The word *transistor* itself is a contraction of the words *transfer-resistor.*

As a simple example of how the n-p-n transistor acts as an actual amplifier, let us examine Fig. 28. Here, a signal source, V_i, has been connected to the emitter, in series with the emitter bias source. A load resistance, R_L, has been connected in series with the collector bias supply. Connected in this manner, with the base

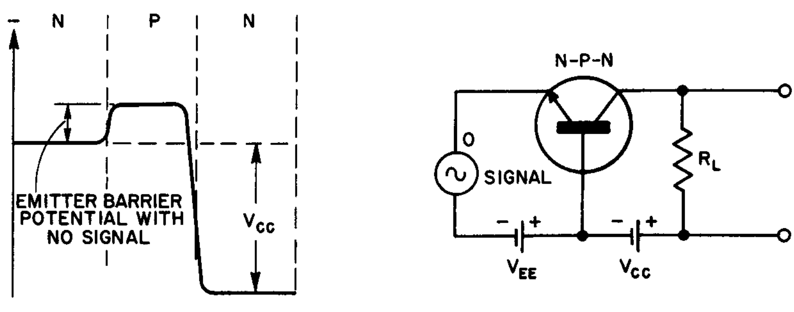


Fig. 28. The n-p-n transistor as an amplifier, with zero signal.

terminal common to both emitter and collector bias batteries, the circuit is referred to as a *common, or grounded, base amplifier.*

When the signal voltage is zero, the number of free electrons crossing the emitter junction and entering the base region is determined solely by the emitter-base bias, as depicted in Fig. 28, which, for convenience, is a repetition of Fig. 27 (B). Now, assume that the signal voltage is increased slightly in the negative direction. Since the signal source is in series with the emitter bias battery, a deflection in the negative direction will add to the emitter-base forward bias battery potential, causing the emitter

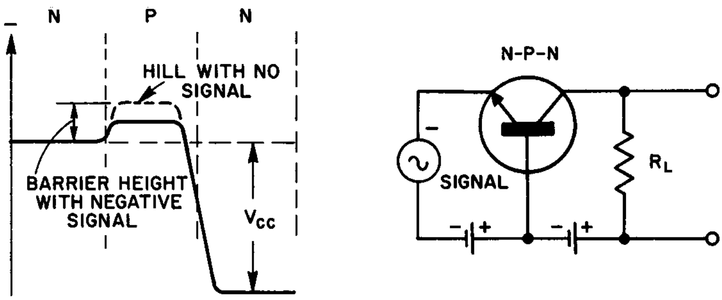


Fig. 29. The n-p-n transistor as an amplifier, with negative-going signal.

potential hill to be reduced further. This permits a greater number of free electrons to enter the base region (Fig. 29), slightly increasing the emitter current. Since most of the electrons entering the base region will diffuse to the collector circuit and through the load resistance, R_L, this increase in emitter current, due to the signal fluctuation, will produce an approximate equal current increase in the load resistance.

If the signal voltage moves in the positive direction the same amount that it moved in the negative direction, the emitter potential hill will be increased by the amount of this deviation, since the signal voltage now subtracts from, instead of adds to, the emitter bias potential. This increase in the height of the emitter potential hill decreases the diffusion of free electrons to the base region, and there will be less current flow in the load resistance (Fig. 30). It can be seen now that the fluctuations of emitter current due to the signal voltage changes are essentially reproduced in the collector circuit. In the low-resistant emitter circuit, these signal induced current changes produce a very small voltage change. However, in the collector circuit, where the load resistance is relatively high, the same current changes produce a large voltage variation. Hence, voltage amplification is achieved in the device.

For a more quantitative treatment of the same phenomenon, let us assume that the signal voltage increases in the negative

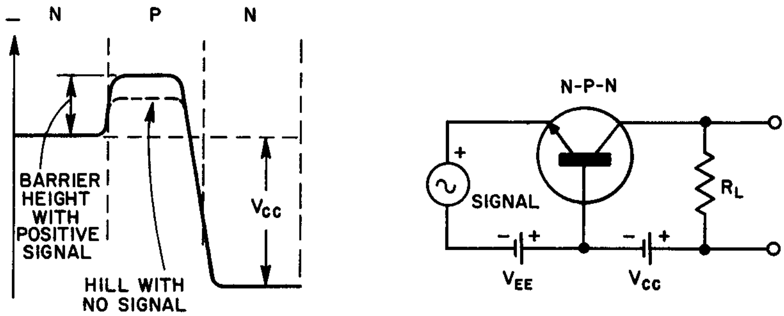


Fig. 30. The n-p-n transistor as an amplifier, with positive-going signal.

direction by an amount ΔV_i (pronounced delta V_i). The forward-biased emitter represents a low resistance, r_e, so that this slight increase in emitter voltage will produce a correspondingly large increase in emitter current of:

$$\Delta I_E = \frac{\Delta V_i}{r_e} \tag{1}$$

The amount of this increase of emitter current that reaches the collector is dependent upon the *current gain* of the transistor. If, for instance, 98% of the free electrons injected into the base region always reach the collector, then 98% of this *increase* in emitter current would also reach the collector. Hence, the amount of ΔI_E reaching the collector (and producing a change in collector current, ΔI_C) is equal to the current gain $\times$ ΔI_E. The current gain for a common base circuit is called alpha (α). Then:

$$\Delta I_C = \alpha \Delta I_E \tag{2}$$

This change in collector current will produce a corresponding change in the voltage across the load resistance (called the output voltage, V_o) equal to:

$$V_o = \alpha \Delta I_E R_L \tag{3}$$

The voltage amplification of the amplifier (A_v) is defined as the ratio of the change in output voltage to the change in input voltage. Therefore:

$$A_v = \frac{\Delta E_o}{\Delta E_i} \qquad (4)$$

Expressed in terms of the load and emitter resistances:

$$A_v = \frac{\Delta E_o}{\Delta E_i} = \frac{\alpha \Delta I_E R_L}{\Delta I_E r_e} = \frac{\alpha R_L}{r_e} \qquad (5)$$

Since α is approximately 1, the voltage amplification is essentially equal to the ratio of R_L/r_e. With a load resistance of, say, 3000 ohms, and a typical emitter resistance of about 30 ohms, the voltage amplification is 3000/30 = 100.

A power gain is also realized. The change in power dissipated in the load resistance (power output, P_o) due to the change in collector current is:

$$P_o = \Delta I_C{}^2 R_L = [\Delta I_E \alpha]^2 R_L = \Delta_E{}^2 \alpha^2 R_L \qquad (6)$$

The change in input power dissipated in the emitter forward resistance (input power, P_i) is:

$$P_i = \Delta I_E{}^2 r_e \qquad (7)$$

The power gain (A_p) is the ratio of the two:

$$A_p = \frac{P_o}{P_i} = \frac{\Delta I_E{}^2 \alpha^2 R_L}{\Delta I_E{}^2 r_e} = \frac{\alpha^2 R_L}{r_e} \qquad (8)$$

The details of the grounded base amplifier will be dealt with in Chapter 4. This example was given only to demonstrate the principle of power and voltage amplification occurring when a current is transferred from a low-resistance circuit to a higher resistance circuit. In practice, other factors must be considered before a transistor amplifier may be designed.

20. Comparison of Transistors to Vacuum Tubes

Let us examine some of the similarities and dissimilarities between the transistor and the vacuum tube. Knowing their relationship, we can often call upon our previous knowledge of vacuum tubes, when designing transistor circuits.

Both transistors and vacuum tubes are designed to amplify a signal. Both the transistor and the triode vacuum tube have three basic elements. In fact, the transistor is often referred to as a *semiconductor triode*. This external similarity of the two devices

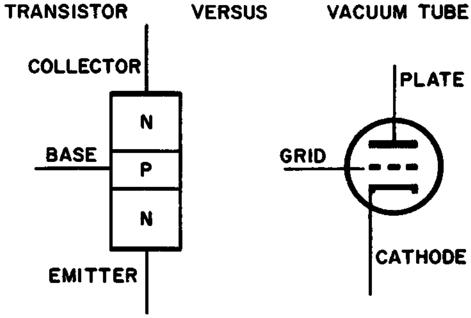


Fig. 31. Comparing transistor and vacuum tube elements.

suggests that the emitter, collector, and base of a transistor have an analogous relationship to the cathode, plate, and grid, respectively, of a triode vacuum tube (Fig. 31).

The current carriers in a vacuum tube (electrons) are liberated by the cathode. The equivalent job in a transistor is performed by the emitter; it emits current carriers. Hence, the cathode in a vacuum tube is analogous to the emitter in a transistor.

Current carriers in a vacuum tube travel through the vacuum tube to the plate, and then to the external circuit. In a transistor, current carriers travel across the base region, and to the collector. The plate of a vacuum tube is analogous to the collector in a transistor.

The electrons of a vacuum tube must flow through the grid before reaching the plate. In a transistor, the current carriers must flow through the base section before reaching the collector. The grid is analogous to the base. This relationship is further appreciated when we consider the effect of an input signal. In a vacuum tube, the signal varies the voltage between the cathode and grid. In a transistor, the signal varies the voltage between the emitter and base.

The transistor is essentially a CURRENT-OPERATED DEVICE; that is, differences in current flow control its operation. The common base current amplification factor, α, is the ratio of a change in collector current produced by a change in emitter current. The vacuum tube is a VOLTAGE-OPERATED DEVICE; electrostatic potential variations between its elements control its operation. The voltage amplification factor μ is the electron tube equivalent to α. It is the ratio of a change in plate voltage produced by a change in grid voltage.

Along with the above analogous relationships, in a vacuum tube, the current carriers move through a vacuum; in a transistor, they move through a solid.

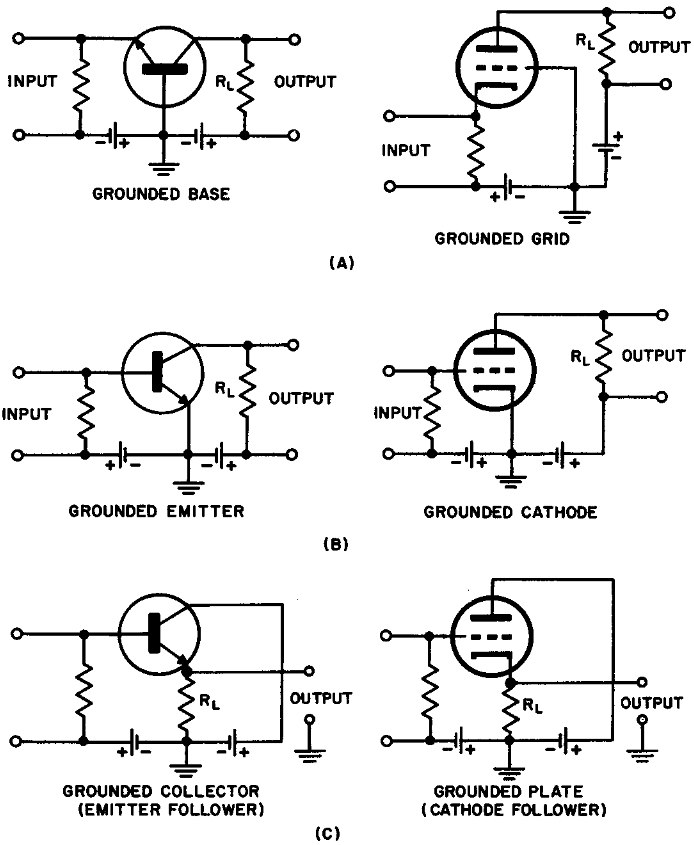


Fig. 32. Comparison between the transistor and accompanying tube configurations. (A) Grounded Base vs Grounded Grid; (B) Grounded Emitter vs Grounded Cathode; (C) Grounded Collector vs Grounded Plate.

21. Comparison of Basic Transistor and Vacuum Tube Circuits

This analogy of a transistor's elements to those of a vacuum tube suggests that a grounded base amplifier is similar to the grounded grid vacuum tube amplifier, as shown in Fig. 32 (A).[2]

[2] To further illustrate the analogy, in the illustration the input signal to the transistor amplifier is applied across a resistor. This is done to allow the signal to be fed through a capacitor and hence prevent bias current from flowing through the signal source, which is not shown. The signal-induced voltage variations across this resistor will add and subtract from the emitter bias in the same way as a signal placed directly in series with the emitter bias battery.

In the grounded, or common, grid circuit, the cathode current equals the plate current; in the grounded, or common, base circuit, the emitter current approximately equals the collector current. (Voltage and power gains for the common base amplifier are in the order of 100 and 400.) The input resistance of both grounded base and grounded grid amplifiers ranges from 30 to about 300 ohms. The output resistance is about 100,000 to 500,000 ohms, for both circuits. There is no signal polarity inversion in either circuit; *i.e.*, when the input signal goes positive, the output signal also goes positive.

The grounded base configuration is not the only method of connecting a transistor as an amplifier. Figure 32 (B) shows a *grounded,* or *common emitter,* amplifier. (It derives its name from the fact that the emitter is common to both the input and output.) Similar to the common base amplifier, the signal in this circuit is also applied between the emitter and base, except that in this instance the collector current does not flow through the input circuit. Both the tube and transistor are capable of current gains, and there is a signal polarity inversion between input and output. The grounded emitter circuit has a higher input resistance (300-1,000 ohms) and a lower output resistance (5–50,000 ohms) than the grounded base amplifier. For these reasons, the grounded emitter is the most popular general transistor amplifier circuit. (Current gains much greater than one are achievable. The reason for this will be dealt with in Chapter 4.)

The *grounded collector* circuit of Fig. 32 (C) is similar to the grounded plate, or cathode follower, vacuum tube circuit. The input resistance is much higher than the other two configurations. The output resistance is very low. This circuit is primarily used for impedance matching between two circuits. There is no signal polarity inversion between input and output. For this reason, the grounded collector circuit is often called an *emitter follower,* signifying its similarity to the cathode follower. Both circuits have a very high current and power gain, with a voltage gain of less than one.

We have concerned ourselves with the n-p-n transistor. If holes may act as positive electrons, is not a p-n-p transistor possible? It is not only possible, but it is more popular than the n-p-n transistor. The two types of transistors are practically identical in their outward performance, except that holes are now the major current carriers, and have a slightly lower mobility than free electrons. Thus, a separate theoretical discussion of p-n-p transistors is unnecessary.

A p-n-p common base transistor amplifier is shown in Fig. 33.

Notice that the symbol has the emitter arrow pointing *towards* the base. (The emitter arrow points in the direction that positive current will flow when the emitter is forward biased. Of course the arrow points in the opposite direction of electron flow.) In addition, the polarity of the bias batteries is opposite to that in the n-p-n transistor. It is important to immediately recognize the schematic differences between n-p-n and p-n-p transistors. Failure to do this will probably result in connecting the bias batteries backwards. A transistor is a physically rugged device; however,

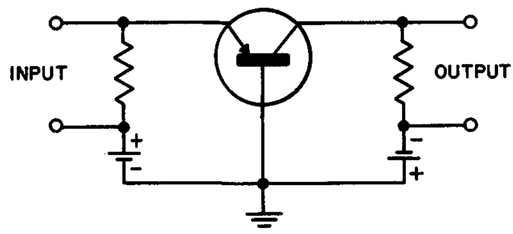


Fig. 33. A p-n-p common base transistor amplifier.

it is very sensitive to electrical overloads, etc. Connecting the collector batteries backwards will result in ruining the device. Become familiar with the n-p-n and p-n-p versions of the three basic amplifier configurations. These circuits are presented in Fig. 34, along with their accompanying bias battery polarities.

Transistors and diodes are classified according to the same rules. Transistors may be called *grown junction* or *diffused junction* or *diffused base* or *surface barrier* or *point contact,* etc. Junction formation is the governing rule.

The grown junction transistor is produced by the technique described in connection with the grown p-n diode. Grown junction transistors are characterized by a rather low noise figure and relatively stable characteristics.

The construction of the diffused or alloy junction transistor is analogous to the alloy diode. Two small dots of indium are placed on either side of a thin wafer of n-type germanium. The entire structure is heated to about 500°C, for a short time (Fig. 35). Under the influence of heat, the indium dots dissolve into the germanium wafer (this action is similar to soldering action), changing the n-type germanium to p-type germanium. Notice that the collector dot is larger than the emitter dot. This is to counteract the spreading effect of the holes as they cross the base region on their way to the collector. If it is larger than the emitter, it will collect a greater percentage of this divergent flow of holes.

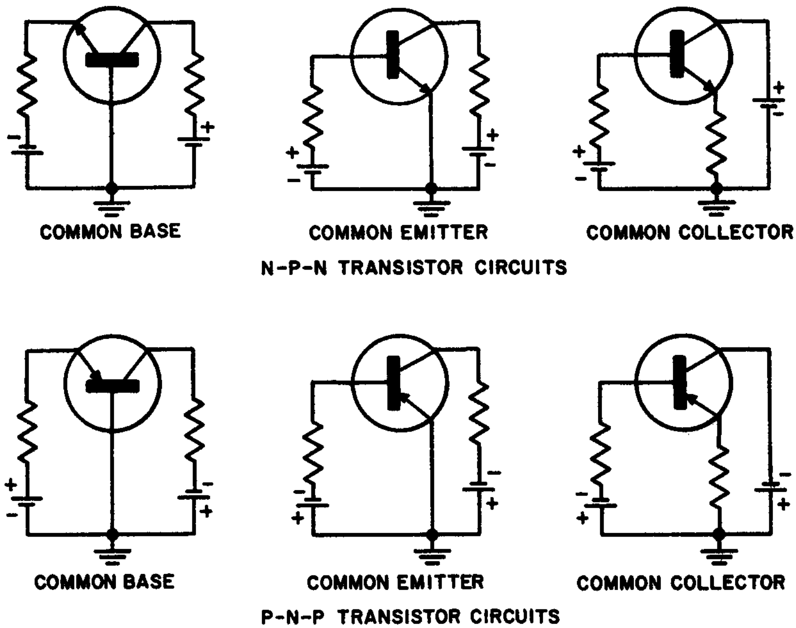


Fig. 34. Basic transistor configuration; (A) n-p-n transistor circuits; (B) p-n-p transistor circuits.

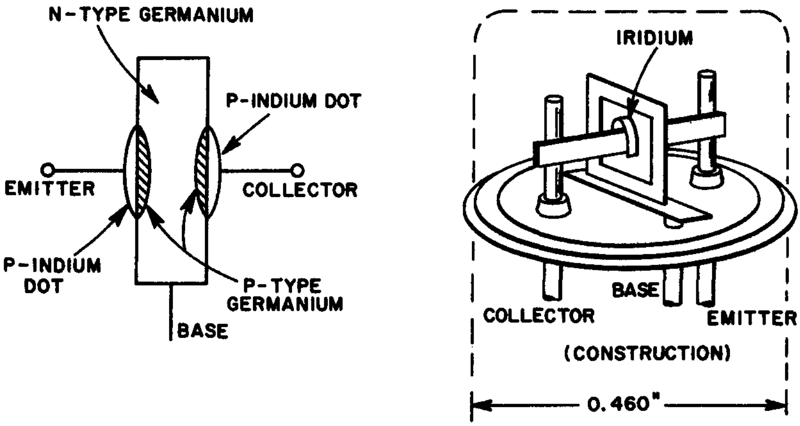


Fig. 35. Diffused (alloyed) junction transistor.

A diffused base transistor does not have an analogous diode type. To manufacture the transistor, a wafer of p-type germanium 0.003 inch thick is heated to the proper temperature. Under the action of heat, a layer of n-type impurity is caused to diffuse a short distance into the germanium slab (Fig. 36), thus forming the base region. A thin strip of p-type germanium, forming the emitter,

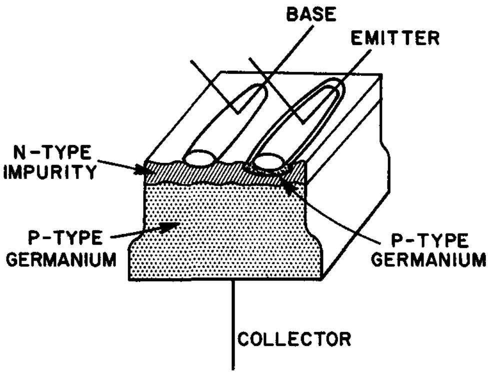


Fig. 36. Diffused base junction transistor.

is evaporated and alloyed atop this n-impurity slab. The collector becomes the original p-type germanium wafer. Double gold leads are welded to the emitter and base elements. This type of construction possesses several inherent advantages. It allows operation at an extremely high frequency. The *2N509* and *2N537* today are in commercial production, operating in the 1000 mc region—something considered fantastic a short time ago. Although fabricated of germanium, this transistor can withstand a temperature up to 100°C. It has a very low collector barrier capacitance and thus helps to provide good high-frequency characteristics. The diffused base transistor was used in the telemetering transmitters of both the Vanguard and Explorer satellites, on a frequency of about 106 mc. This is but one striking example of how the transistor has surpassed the vacuum tube in certain specialized tasks requiring reliability, ruggedness, and efficiency.

The *Philco* surface barrier transistor is another type of high frequency transistor. It utilizes the junction principle for its operation, but does not use p- *and* n-type germanium. It utilizes only n-type germanium. The emitter and collector are electroplated on opposite sides of an n-type germanium wafer. The emitter and collector materials do not diffuse into the n material (as in the diffused junction transistor), but remain on the surface only.

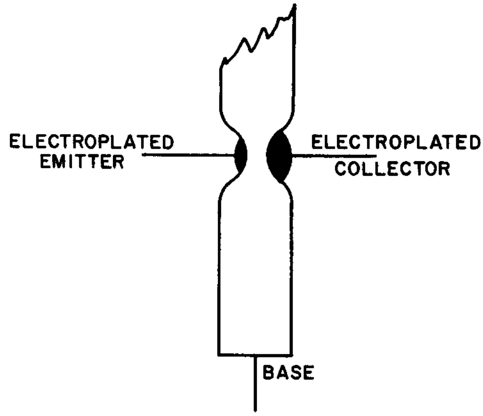


Fig. 37. Surface barrier junction transistor.

Figure 37 illustrates the details of the surface barrier transistor.

Because of a rather complex energy phenomenon that exists at the surface of the germanium, a high negative electric field is built up here, repelling free electrons attempting to enter the surface region. It is this surface of the germanium crystal that is referred to as the "barrier." One metal contact to the crystal is made positive; it reduces this negative barrier; the contact serves as the emitter. The other metal contact is made negative with respect to the main crystal; this serves as the collector. The main wafer of n-type germanium serves as the base. The entire structure is an n-p-n transistor configuration. Its high frequency performance is due to the extremely thin layer of germanium separating the collector from the emitter. This transistor is used as a video amplifier, an oscillator, etc., in the vhf region.

The point contact transistor was the first type of transistor available. The method of manufacture is quite similar to that of making a single point-contact diode. Two phosphor-bronze wires are placed against a wafer of n-type germanium. A surge of current is passed through the two wires, thus forming two islands of p-type germanium directly beneath the point where the wires touch the germanium. Although a point contact transistor may serve in certain circuits where two other types of transistors would otherwise be needed (such as a multivibrator), because of its poor reliability and unstable characteristics, it is rarely used today. Most transistor manufacturers have discontinued production of point contact transistors. For this reason, and because point contact transistor theory is not completely understood, we will deal exclusively with the various types of junction transistors.

22. Review Questions

1. Why is the base region of a grown junction transistor rather narrow with respect to the width of the emitter or collector sections?
2. Why is the effective base region narrower with a reversed bias collector junction than with an unbiased collector junction?
3. Why does a common base transistor have a voltage and power gain when the current gain is less than unity?
4. Where does the word transistor originate?
5. What three vacuum tube circuits are analogous to the common base, common emitter, and common collector transistor circuits respectively?
6. Why could not temperature be used as a minority carrier injector into the base region?
7. What is a symmetrical transistor? Are normal transistors asymmetrical? If so, why?

Chapter 4

THE TRANSISTOR AS A CIRCUIT ELEMENT

The analysis and design of a transistor circuit is a logical step-by-step procedure. We will develop a simple method by which a transistor amplifier may be designed—knowing the basic characteristics of the transistor to be used and the specifications of the proposed amplifier. Here, we will deal mostly with transistor characteristics and the problems involved in establishing the quiescent point of operation for the different configurations. Since we will be dealing with many types of voltages and currents, an explanation of the various symbols to be used and the rules governing their use is in order.

Basically, any voltage or current, etc., may be denoted by a letter and a subscript. If the letter is of the upper case variety, it indicates a dc or an rms value. If it is a lower case letter, it represents an instantaneous value. Thus I, V, and P represent rms or dc values of current, voltage, and power, respectively. Instantaneous values of the same quantities: i, v, and p.

The subscript has a variety of meanings. If voltage is being measured, the subscript usually consists of two letters. The first letter denotes the electrode at which the voltage is measured, and the second subscript letter denotes the reference electrode with respect to which the measurement is made. Dc and instantaneous *total* values are represented by upper case subscripts, and instantaneous variations from the no-signal value (called the quiescent value) are represented by lower case letters. Bias supply voltages are indicated by a repeated upper case subscript.

TABLE 1

COMMON TRANSISTOR SYMBOLS

V_{EE} = The emitter bias supply voltage.
V_{CC} = The collector bias supply voltage.
V_{CE} = The quiescent voltage between collector and emitter.
V_{CB} = The quiescent voltage between collector and base.
V_{cb} = The rms signal voltage, or the rms variation from the quiescent value.
v_{CB} = The instantaneous total value of collector to base voltage.
v_{cb} = The instantaneous variation from the quiescent value of the collector to base voltage.

NOTE: The last four symbols are shown in Fig. 38, which represents the voltage variations from collector to base of a hypothetical transistor.

Figure 39 shows the polarity convention for the n-p-n and p-n-p common base configuration. Here, whether the transistor is an n-p-n or p-n-p device, the emitter and collector voltages are always DEFINED as being positive with respect to the base. Hence, if (as actually happens) the emitter is negative with respect to the base, as in an n-p-n transistor, the negative sign will accompany this voltage, because it is opposite to the convention. Remember to use the base as a reference. Then the actual voltages on the other elements will have the correct and logical polarity.

It is now known that current in a conductor is actually the flow of electrons. This was not known to early circuit theorists. Hence, a convention was adopted that electric *current* flows from positive to negative. It was visualized that there was a rise in potential within a battery (from minus to plus) and a fall of potential outside the battery (from plus to minus). Early circuit theory was developed with this convention in mind. When the electron theory was developed, it was too late to change the old convention. Hence, the old concept that current flows from + to − is still in use. It has often been argued that the convention should be changed. The argument centers about what we mean by current flow.

Basically, current flow is the flow of electrical charges—be they minus electrons, or positive ions or holes. In a vacuum tube,

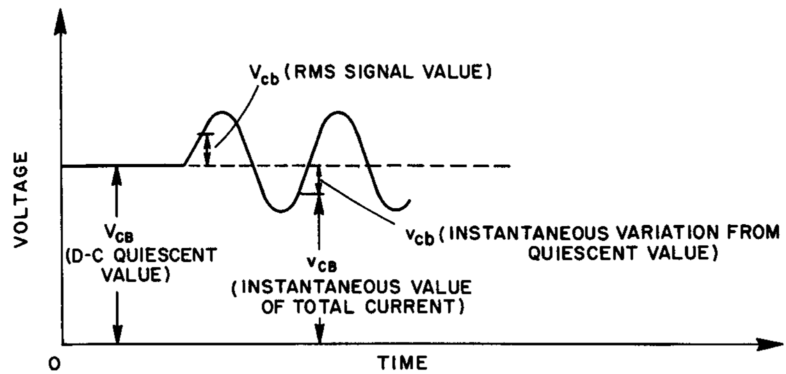


Fig. 38. Voltage variations from collector to base of a hypothetical transistor.

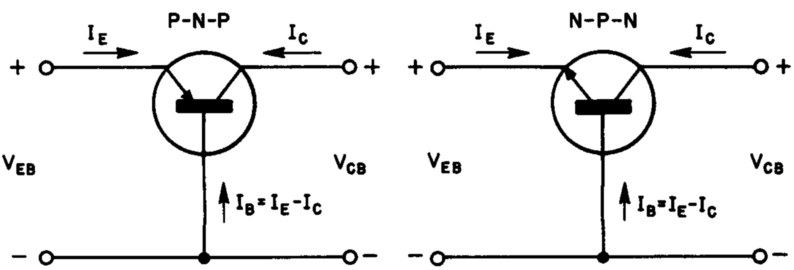


Fig. 39. Polarity convention for n-p-n and p-n-p common base configurations.

electrons are the main carriers of current and so the old convention of current flowing from the plate to cathode might be considered incorrect. However, there are gas tubes where positive ions *do* travel from plate to cathode. In transistors, neither the electron nor the positive current theory is completely correct. Both positive holes and negative electrons are the carriers of electricity, depending on whether the transistor is of the n-p-n or p-n-p variety.

Because almost all transistor literature (and even the schematic symbol of a transistor) has been developed with positive current theory in mind, this convention will be applied here. Even if the reader is aware of the electron theory, confusion will not arise. No matter which convention is followed, the SAME results are arrived at. In Fig. 39, this current convention is applied to the two types of transistors. The emitter arrow in each case always points in the direction of positive current flow, when the junction is in forward bias. Notice also that the currents are defined as being positive for both configurations, when they flow into the transistor. Now, under actual bias conditions (shown in Fig. 40), the emitter and collector are not always positive with respect to the base, and the currents do not always flow into the transistor.

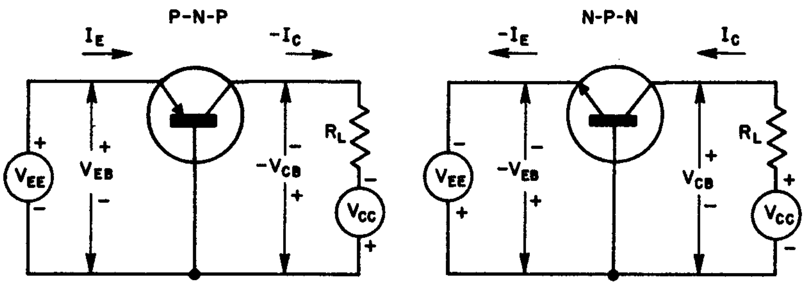


Fig. 40. N-p-n and p-n-p common base configurations under actual bias conditions.

When the direction of current is opposite to the defined direction into the transistor, a negative sign is assigned to it. Thus the p-n-p transistor has a positive emitter current and bias, and a negative collector current and bias. The n-p-n transistor (Fig. 40) has a negative emitter current and bias, and a positive collector current and bias.

23. Transistor Characteristics

In Chapter 3, it was noted that even when no signal was applied to the common base amplifier, an emitter and collector current flowed. The presence of this constant bias current is necessary for the proper operation of a transistor, just as the presence of plate and grid bias voltages are necessary for the successful operation of a vacuum tube amplifier. This no signal, or quiescent bias current in a transistor presents an average current, which the signal current either adds to or subtracts from. It is important that this quiescent current is of the proper magnitude, so that equal positive and negative excursions of the input signal produce symmetrical variations of the output current about an average value. An incorrect quiescent bias current may cause one or more forms of distortion in the output signal.

A constant quiescent bias current flowing in the collector will

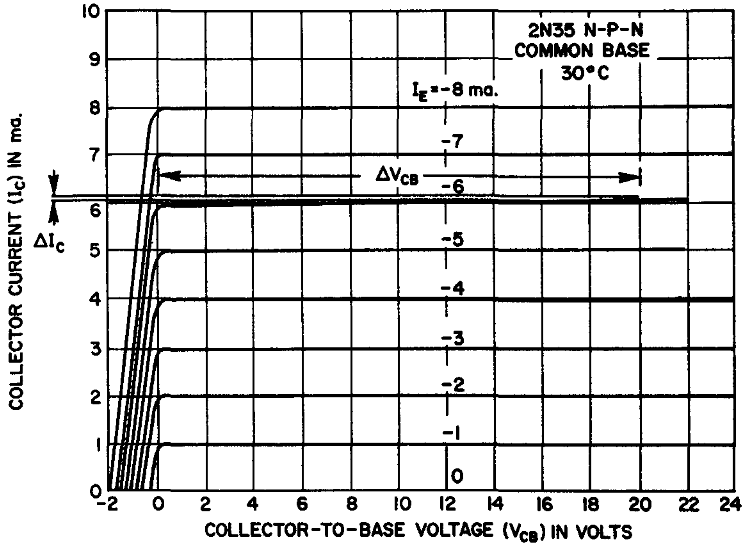


Fig. 41. Static collector characteristics for the common base configuration of a 2N35 n-p-n transistor.

cause a collector-to-base voltage drop. If the signal current in the emitter circuit causes the collector current to vary above and below its quiescent value, then the collector-to-base voltage will also vary above and below its quiescent value, as shown in Figs. 28, 29, and 30.

Unfortunately, the resistance of the emitter and collector junctions are not always constant. Hence, Ohm's Law cannot always be used to express the relationship between the various currents and voltages in a transistor. For this and other reasons, this information is usually supplied by the manufacturer in graphical form.

The two most commonly used graphs of transistor characteristics deal with the collector and emitter respectively. Of these, the collector graph is the most popular and useful. Figure 41 shows a graph of the static collector characteristics for the common base configuration of a *2N35* n-p-n junction transistor. The graph consists of a series of curves. Each curve shows, with the emitter current held constant, the collector current variation as the collector-to-base voltage is changed. The curves are obtained quite simply, using the test apparatus shown with the circuit of Fig. 42. After the collector current is observed and plotted for variations in the collector-to-base voltage, and for a given value of emitter

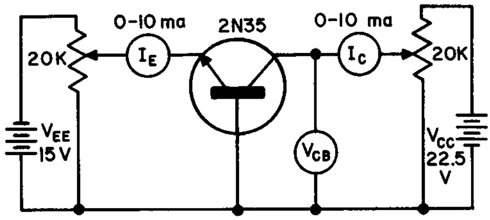


Fig. 42. Test circuit required to obtain characteristics of Fig. 41.

current, the emitter current is changed, and the procedure is repeated. Different curves are shown for increments of 2 ma of emitter current. The entire collection is called a family of collector curves, with the emitter current as a parameter (a value of emitter current set for each curve). The family of curves shows what value of emitter current must exist to produce any combination of V_{CB} and I_C.

The minus sign in front of each emitter current line in the graph exists because in an n-p-n transistor, I_E flows out of the transistor.

Many interesting facts come to light with the aid of this graph. For instance, a rough estimate of the *dynamic* output resistance of the transistor may be obtained. This is the resistance that the load "sees" when looking back into the transistor between the emitter and the base leads. It is analogous to the dynamic plate resistance may be obtained by dividing a certain change in the resistance may be obtained by dividing a certain change in thc collector-to-base voltage by the accompanying change of collector current, while the emitter current is held constant. This is the slope of each $I_C = V_{CB}$ line. For example, in Fig. 41, the 6 ma emitter line moves upward about 0.2 ma as the collector-to-emitter voltage varies from 0 to 20 volts. The output resistance (R_o) is then:

$$R_o = \frac{\Delta V_{CB}}{\Delta I_C} = \frac{20 \text{ volts}}{0.2 \text{ ma}} = 100{,}000 \text{ ohms} \qquad (9)$$

As previously stated, the output resistance of the common base configuration is quite high. This permits a high load resistance. The voltage gain is roughly proportional to the load resistance divided by the emitter resistance. Notice that the output impedance or resistance is quite constant over a wide variation of V_{CB} and I_C, as evidenced by the fact that the slope or tilt of each base current line is about the same. (It is important to state here that the dynamic R_o is not the resistance that would be measured by connecting a voltmeter across the collector base junction and an ammeter in series with the plate. This volt-ammeter resistance is the dc *static* resistance. The slope of the line is the ac *dynamic* resistance. Chapter 5 contains a more detailed discussion of the differences between these two resistances.)

Equation (2) gave us a relationship relating the change in collector current, the change in emitter current, and the current gain of the common base.

$$\Delta I_C = \alpha \Delta I_E \tag{10}$$

Solving this equation for α, yields a definition for this quantity:

$$\alpha = \left.\frac{\Delta I_C}{\Delta I_E}\right|_{V_{CB}\ \text{constant}} \tag{11}$$

α is taken with a constant collector-to-base voltage. Its official name is the *forward short-circuit current gain for the common base configuration.*[1] To obtain this information from the common base collector characteristics, it is noted that for a constant collector-to-base voltage of 2 volts, the collector current varies from 0 to 6 ma, while the emitter current only varies from 0 to 5.95 ma. Thus:

$$\alpha = \frac{\Delta I_C}{\Delta I_E} = \frac{5.95\ \text{ma}}{6.00\ \text{ma}} \cong .995 \tag{12}$$

(The symbol $\cong$ means approximately equal to)

α is not related to the external circuit. It applies only to the common base configuration. The value obtained from this graph (as obtained from any graph) is not exact; the manufacturer specifies a more reliable value. α for the *2N35* is given as 0.98; typical values for α vary between .90 and .99. (Again, this is the dynamic α, of concern when the amplification of ac signals are involved. When dealing with dc switching circuits, the dc, or static α is useful.)

Although it cannot be detected from the graph, there is a minute collector current flowing even when the emitter current is zero. Since, when the emitter current is zero, no minority carriers are injected into the base region, this current can only come from thermally-generated minority carriers. It is the familiar reverse saturation current; but, this time residing between the collector-base junction. It is called I_{CO} to denote that it is the *collector*

---

[1] The word "short circuit" only refers to the ac value; dc voltages and currents are present.

reverse saturation current. At room temperature (30°C), I_{CO} is only a few microamperes, and so can be neglected with respect to the much larger collector current. However, it is highly dependent on the temperature—approximately doubling in value for every 10°C rise in temperature. Where I_{CO} is typically only 5 μa at room temperature, it will have increased to about 650 μa at 100°C. Since I_{CO} is often of a magnitude that cannot be ignored, the expression relating the collector current with the emitter current must be modified slightly to include this term. Of course, thermally-generated minority carriers in the base region "look" to the collector, just as injected minority carriers do. Hence, these two currents add in the collector circuit. Equation (2) must now be modified to read:

$$I_C = \alpha I_E + I_{CO} \tag{13}$$

It is apparent that large heat-induced increase in I_{CO} might disrupt the circuit operation, although the common base is not as sensitive to these changes in I_{CO} as other configurations. The

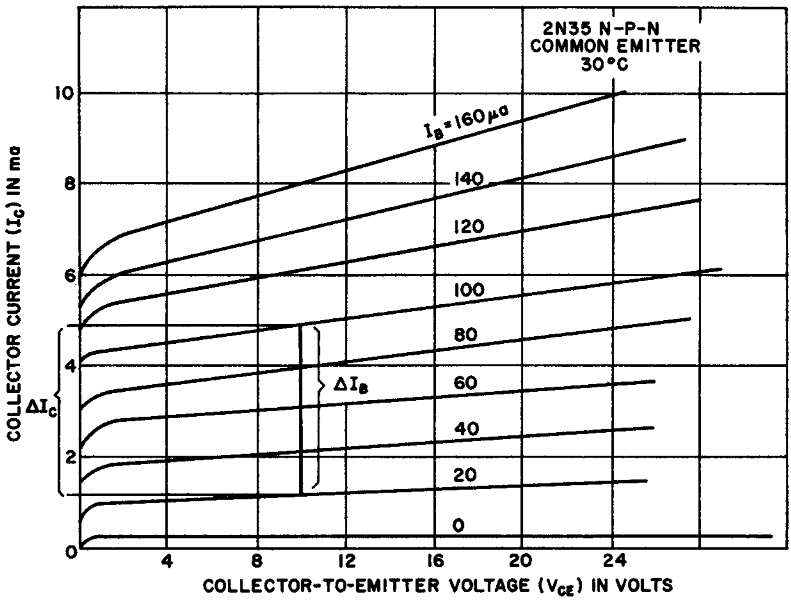


Fig. 43. Static collector or output characteristics for the common emitter configuration of a 2N35 n-p-n transistor.

common base static collector characteristics must be taken at a given temperature, and this information must be presented along with the graph. Transistor characteristic graphs are usually shown for a temperature of 30°C.

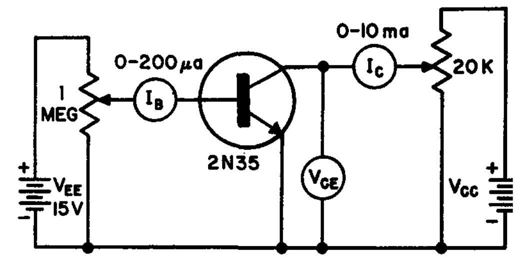


Fig. 44. Test circuit required to obtain characteristics of Fig. 43.

24. Common Emitter Characteristics

Figure 43 shows the static collector or output characteristics of the common emitter configuration. The curves were obtained in the same manner as they were for the common base, using only the common emitter test circuit of Fig. 44. In this instance, collector current is plotted for variations in collector-to-emitter voltage, with the base current as a parameter. Notice that all of the base current curves merge into the same vertical line near $V_{CE} = 0$. This graph is similar to a pentode vacuum tube's static plate characteristics.

It is immediately noticed from Fig. 43 that the slope of each line is greater than the common base. This signifies that the dynamic output resistance is less. With the common emitter, the output resistance is approximately the resistance betwen the collector and the emitter. It can be seen that along the 120 μa base current line, the collector current rises 0.6 ma for a variation of collector-to-emitter of 2 to 8 volts. Hence the output resistance is:

$$R_O = \frac{\Delta V_{CE}}{\Delta I_C} = \frac{7 \text{ volts}}{0.6 \text{ ma}} = 11{,}600 \text{ ohms} \quad (14)$$

As with the common base configuration, the output resistance is relatively constant with different values of emitter and collector currents. (Later we will make use of this very important fact.)

The base current in a transistor is equal to the difference between the collector and emitter currents. If this base current could be varied, it would result in a change in the collector current. Since the base current represents the difference between the collector and emitter currents, a change in this *difference* current *must* result in a change in either the collector or emitter current, or both. But since the emitter current is the source of the collector current, both the emitter and collector change. In the common base configuration, this fact held little significance, since the emitter and not the base current controlled the collector current. In the common emitter, however, this is not the case. The signal is applied between the base and emitter in such a way as to produce changes in the base current. Since the majority of transistor circuits in practice employ this common emitter configuration, a further study of common emitter current gain is in order.

Neglecting I_{CO} for the moment, the relationship between the change in emitter current and the change in collector current in the common base amplifier was found to be:

$$\Delta I_C = \alpha \Delta I_E \tag{15}$$

The base current is *always* equal to the difference between the emitter and collector current. Hence, the change in base current must be equal to the difference between the change in the emitter and collector currents.

$$\Delta I_B = \Delta I_E - \Delta I_C \tag{16}$$

Substituting $\alpha \Delta I_E$ from Equation (15) for ΔI_C in Equation (16), we have:

$$\Delta I_B = \Delta I_E - \alpha \Delta I_E \tag{17}$$

$$\Delta I_B = \Delta I_E (1 - \alpha)$$

Solving this for ΔI_E, we find that:

$$\Delta I_E = \frac{\Delta I_B}{1 - \alpha} \tag{18}$$

But Equation (15) still holds for any configuration. Hence, we substitute $(\Delta I_B/1 - \alpha)$ in Equation (18) for ΔI_E in Equation (15):

$$\Delta I_C = \alpha \Delta I_E = \alpha \left(\frac{\Delta I_B}{1 - \alpha}\right) = \left(\frac{\alpha}{1 - \alpha}\right) \Delta I_B \tag{19}$$

Therefore, the change in collector current is equal to the change in base current × the quantity $(\alpha/1 - \alpha)$. The term $(\alpha/1 - \alpha)$ is the current gain for the common emitter, corresponding to the α found in the common base circuit. It is denoted by the symbols α_{CB}, α_E and β. The last symbol (pronounced beta) is by far the most popular term and will be used in this book. In terms of α, the common base current gain, β is defined as:

$$\beta = \frac{\alpha}{1 - \alpha} \qquad (20)$$

The values of α and/or β are given by the manufacturer. With the aid of Equation (20) a conversion of either value may be made. As opposed to α, whose value is never greater than 1 in a junction transistor, β may assume values ranging from 10 to 100, depending upon the transistor. As can be seen from Equation (20), the closer that α approaches 1, the greater β becomes. The *2N35* has an α of 0.98, and so β equals $(0.98/1 - 0.98) = 49$.

When α is near 1, a very slight variation in α has a greatly magnified effect on β. Unfortunately, the state of the transistor art has not progressed to the point where α may be held constant during manufacture. One can expect a certain variability in α between different transistors of the same type, even ones made by the same manufacturer. This slight α variability gives rise to much greater changes in β, and so the common emitter characteristic curves are subject to a rather wide variation in practice.

From the common base relationship $\Delta I_C = \alpha \Delta I_E$, we previously obtained an expression for α:

$$\alpha = \frac{\Delta I_C}{\Delta I_E}$$

Similarly, for the common base:

$$\Delta I_C = \beta \Delta I_B \qquad \text{(a)}$$

hence:

$$\beta = \frac{\Delta I_C}{\Delta I_B} \qquad \text{(b)} \qquad (21)$$

With the aid of Equation (21), β may be obtained directly from the common emitter collector static characteristics of Fig. 43. For a collector-to-emitter voltage of 10 volts, a base current variation from 60 to 160 μa produces a collector current variation from 3 to 7.5 ma. Hence β is found to be:

$$\beta = \frac{\Delta I_C}{\Delta I_B} = \frac{4.95\,\text{ma} - 1.2\,\text{ma}}{100\,\mu\text{a} - 20\,\mu\text{a}} = \frac{3.25 \times 10^{-3}\,\text{amperes}}{80 \times 10^{-6}\,\text{amperes}} = 47 \qquad (22)$$

This value agrees rather nicely with the calculated value of 49 for the *2N35* using Equation (20).

The affect of I_{CO} upon the collector current is much greater than for the common base. Its effect may be appreciated from the following consideration: I_{CO} for any configuration flows across the collector-base junction. The electron current resulting from I_{CO} must enter or leave the base terminal, depending on whether the device is a p-n-p or an n-p-n transistor. In either case, the base current is varied by this reverse saturation current. This change in the base current is tantamount to a minute signal and, therefore, gives rise to a collector current of $I_C = I_{CO}\ (\beta) = I_{CO}\ [\alpha/1-\alpha]$ in accordance with part (a) of Equation (21). However, the reverse saturation current flows not only in the base but also in the collector. Hence, the total collector current due to I_{CO} is the sum of two terms: the magnified I_{CO} of the base plus the I_{CO} actually flowing in the collector.

$$I_C = I_{CO} + I_{CO}\left(\frac{\alpha}{1-\alpha}\right)$$

$$I_C = I_{CO}\left(1 - \frac{\alpha}{1-\alpha}\right)$$

$$I_C = I_{CO}\left(\frac{1-\alpha}{1-\alpha} - \frac{\alpha}{1-\alpha}\right)$$

$$= I_{CO}\left(\frac{1}{1-\alpha}\right) = \frac{I_{CO}}{1-\alpha} \qquad (23)$$

We know that the total collector current is the sum of the normal or quiescent collector current plus the collector reverse saturation current. Adding equations (a) of Equations (21) and (23), we obtain an expression for the total collector current for the common emitter.

$$I_C = \beta I_B + \frac{I_{CO}}{1-\alpha} \qquad \text{(a)}$$

$$I_C = I_B\left(\frac{\alpha}{1-\alpha}\right) + \frac{I_{CO}}{1-\alpha} \qquad \text{(b)} \qquad (24)$$

Since I_{CO} is multiplied by the factor $(1/1-\alpha)$ when it reaches the collector, the circuit is much more sensitive to temperature variations than the common base, where I_{CO} is not magnified at all when it reaches the collector. For instance, with $\alpha = 0.98$ as with the *2N35*, the quantity $(1/1-\alpha)$ is equal to 50. Thus the effect on the collector current due to I_{CO} is about fifty times as great for the common emitter as for the common base.

The collector current for the curve of $I_B = 0\ \mu a$ represents this magnification of I_{CO} in the collector circuit. Here, I_{CO} is about

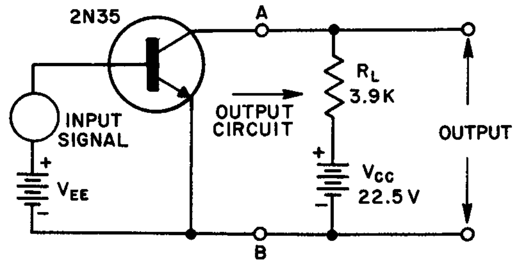


Fig. 45. A practical common-emitter voltage amplifier employing the 2N35 transistor.

0.3 ma, whereas it was only a few microamperes in the common base configuration of the same transistor. This magnification of I_{CO} in the common emitter circuit gives rise to special problems of circuit stability when operating at high temperatures. These problems will be dealt with in Chapter 5.

25. Drawing the Load Line

The static collector characteristics of the common emitter circuit of Fig. 43 remain the same regardless of the external circuit. By itself, this graph only refers to the voltage directly across, and the currents flowing through, the transistor terminals. With an actual amplifier, we must combine this data with data regarding the external circuit, so that the operation of the amplifier may be graphically portrayed.

In Fig. 45, a more practical common emitter voltage amplifier employing the *2N35* transistor is shown. The point A and B separate the transistor proper from the external output circuit. It is known from the common base characteristics how the internal circuit of the transistor amplifier (*viz.*, the transistor proper) reacts to different voltages and currents. We wish to determine how it will react when the external circuit is connected at points A and B. This can be done graphically by determining the appropriate current and voltage existing across the collector-base terminals (A and B) when these terminals are shorted and open-circuited. These conditions will give us two points of interest on the collector characteristics graph. Between these two points a line called a *load line* will be drawn. Many operational properties of the amplifier may be explained with the aid of this load line.

To do this coherently and logically, so that we may gain as much information as possible, the problem will be approached in a simple mathematical manner. In Fig. 46, the same transistor circuit is shown. For the moment, we are interested only in the collector output circuit and will not concern ourselves with the

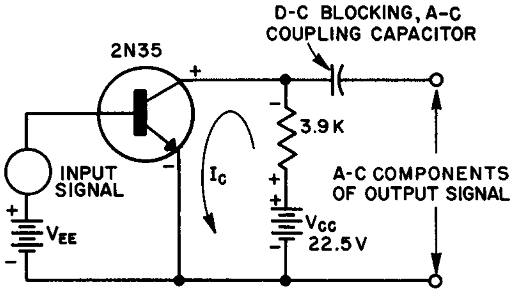


Fig. 46. Circuit of Fig. 45, with addition of d-c blocking coupling capacitor.

base circuit. The collector current flows in the loop indicated in the figure. It will cause a voltage drop across the collector-emitter terminals and the load resistance. The sum of these two voltage drops *must* add up to the collector bias supply voltage. This is Kirchhoff's famous voltage law: *the sum of all of the voltage drops around any closed loop is zero.* Taken in the direction of positive current travel, and assuming that all voltage drops external of the source are positive and the voltage rise in the source is negative (negative drop), we proceed as follows around the loop:

$$\begin{aligned} &- \text{(rise in battery)} \\ &+ \text{(drop across load resistance)} \\ &+ \text{(drop across emitter-base terminals)} = 0 \end{aligned} \tag{25}$$

or

$$-V_{CC} + V_{CE} + V\ (\text{across } R_L) = 0 \tag{26}$$

or, when the resistive voltage drops are expressed in terms of I × R:

$$-V_{CC} + V_{CE} + I_C R_L = 0. \tag{27}$$

$$I_C = \frac{V_{CC} - V_{CE}}{R_L} \tag{28}$$

Equation (27) is merely an expression equating all of the voltage drops in the collector loop to zero. Now, if we solve it for I_C, with the stipulation that $V_{CE} = 0$, this will correspond to a short-circuit condition where the current is limited purely by R_L. Thus:

$$I_C \Big|_{V_{CE} = 0} = \frac{V_{CC}}{R_L} \tag{29}$$

(The vertical line in Equation (29) serves to separate I_C from the conditions under which it is being measured.)

This short-circuit collector current is located on the common

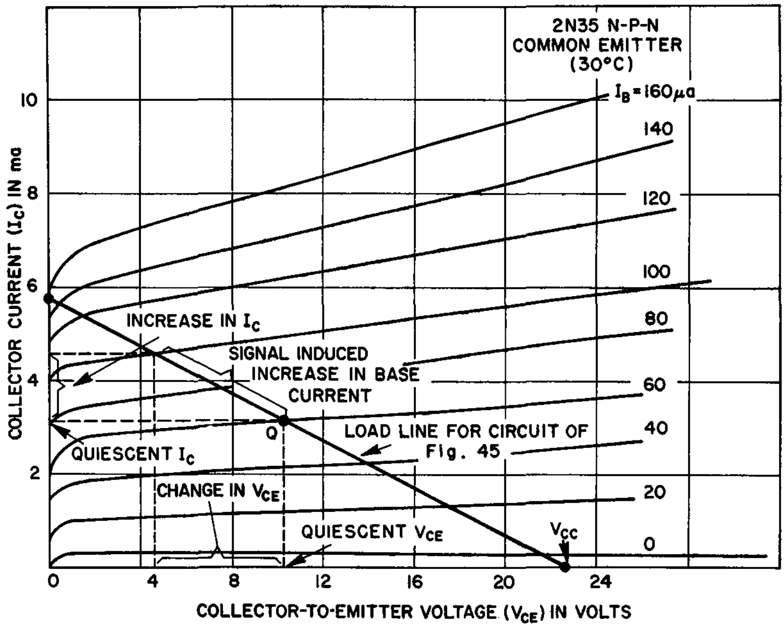


Fig. 47. Common emitter characteristics.

emitter characteristics of Fig. 47 above the zero collector-to-emitter voltage point.

If the transistor were open-circuited in such a way as to stop current flowing from emitter to collector, then no current would flow in the circuit. If the collector current is reduced to zero, this corresponds to an open circuit of the transistor terminals. It may be expressed mathematically by solving Equation (27) for V_{CE}.

$$V_{CE} = V_{CC} - I_C R_L \tag{30}$$

and letting $I_C = 0$

$$V_{CE} = V_{CC} - (0) R_L = V_{CC} \tag{31}$$

This point ($V_{CE} = V_{CC}$) is also located on the base line of the graph opposite the zero collector current point. Now connect a line between the two points just located. This line is the load line.

Here we have two graphs, dealing with two entirely different subjects, superimposed on one another. One is a graph of the collector characteristics, which has nothing to do with the external circuit. The other straight line graph is governed *entirely* by the external circuit and would be the same irrespective of the kind of transistor used. The *combination* of these two graphs gives us the solution of our problem. We may now specify the action of the entire amplifier. The important conclusion to be gained from the above is that the operation of the amplifier **must always be on the load line. For any given value of base current, only those combinations of V_{CE} and I_C that represent points on the load line are attainable.** In effect, this only means that between the open and short circuit conditions on the load line, there lie intermediate values of resistance which must conform to Ohm's Law and, therefore, must lie in a line between the two end points. The graphical construction of a load line is a very important operation in transistor work.

Since the operation of the transistor must always be on the load line, the job of the circuit designer is to pick a quiescent point of operation. The signal will make the collector current vary above and below this operating point, hence it is usually advantageous to place the quiescent point (Point *Q* in Fig. 47) in the middle of the load line—or at least in the middle of the base current curves which the load line crosses. In our example, this point is taken where the 60 μa base current line crosses the load line. Using the *Q* point and the load line, we can graphically illustrate the operation of the amplifier.

The signal is applied to the base. The quiescent base current is 60 μa. We will assume that the signal causes an increase in base current of 40μa. Since we know that the transistor must always operate along the load line, we know that the resultant change in collector current from its quiescent value of 3.1 ma will be found at the intersection of the load line with the new base current value of 100 μa. This increase is 1.5 ma. The amplitude of the voltage change across the transistor represents the output signal, and this is also read directly off the graph. (NOTE: Since the V_{CC} remains constant, the change in V_{CE} must equal the change of voltage across R_L. This change in V_{CE}, representing the difference between the quiescent V_{CE} and the new V_{CE}, is found on the V_{CE} axis directly below the intersection of the load line with the 100 μa base current line.

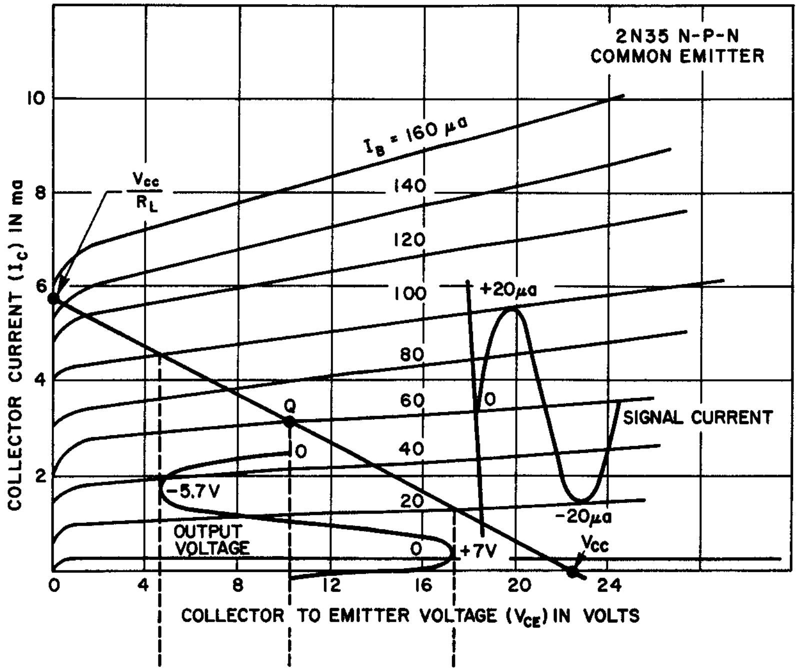


Fig. 48. Common emitter characteristics with a-c signal applied.

It is clear that the transistor is acting as a variable resistor. When the signal increases the base current, the collector-to-emitter resistance decreases and V_{CE} decreases. The opposite happens when the signal causes the base current to decrease. In Fig. 48, the same amplifier is depicted with a sine wave signal applied to the base. The 60 μa quiescent base current line is used as the zero reference point for the input signal. This corresponds to a quiescent output voltage (V_{CE}) of 10.2 volts. The plus and minus 20 μa base current variations due to the signal cause the collector-to-emitter voltage to vary − 5.7 volts and + 7.0 volts, respectively, from the quiescent value of 10.2 volts. Notice that although the input signal was perfectly sinusoidal, the output voltage is slightly distorted. There is some variation of the spacing of the base current lines although it is somewhat exaggerated in this example. If the output were taken directly across the collector-emitter terminals, there would always be the quiescent voltage present. This is not desirable. It is the job of the amplifier to amplify a weak varying signal, not to contribute a large dc component to it. For this reason, the ac component—the fluctuations in V_{CE}—are passed through the coupling capacitor of Fig. 46 and on to the next circuit. The capacitor blocks the dc component, and so is often called a blocking capacitor. Our primary interest is not in the quiescent voltage or current, but only in the signal variations.

The current amplification of this common emitter amplifier may be determined graphically as the ratio of the total swing of the collector current over the total swing of the base current.

$$A_i = \frac{\Delta I_C}{\Delta I_B} = \frac{4.60\,\text{ma} - 1.35\,\text{ma}}{100\,\mu\text{a} - 20\,\mu\text{a}}$$

$$= \frac{3.35 \times 10^{-3}\,\text{amperes}}{80 \times 10^{-6}\,\text{amperes}} = 42 \qquad (32)$$

Here A_i refers to the operational current gain; *i.e.,* the current gain with the external circuit connected. β refers to the short circuit current gain. β for the *2N35* common emitter is 47, as previously measured. Comparing A_i *with* β, confirms the interesting fact that the greatest current gain achievable by an amplifier is the short-circuit gain. In an actual amplifier, the short-circuit current gain is only approached, not reached.

Normally, if the voltage gain of the amplifier in Fig. 46 were to be computed, we would take the ratio of the change of the voltage across the load resistance to the change in the signal voltage. Remember that the collector characteristic curves (Fig. 43) relate *base current* (and not base-to-emitter voltage) to other transistor voltages and currents. A mere knowledge of the change in base-to-emitter voltage (ΔV_{BE}) would be sufficient to calculate the change in collector current and the base-to-emitter voltage. lationship between the base current and the base-to-emitter voltage. Unfortunately, the emitter-base (unlike the collector-base) junction resistance varies substantially with large changes of voltage across the junction. Conseqently, voltage and current do not have a linear relationship. Since in our example of Fig. 46, the signal source is connected directly to the base, we can expect that any signal voltage change will be reflected across the base-emitter junction. If the signal is sinusoidal, the base-to-emitter voltage must also change sinusoidally, BUT THE BASE CURRENT WILL NOT CHANGE SINUSOIDALLY. Because this is a non-linear V-I relationship, it will undoubtedly be somewhat distorted, the exact amount being dependent upon the signal amplitude. Since the usual function of the device is to amplify a signal voltage faithfully, without adding any components of its own, this distortion may hinder its usefullness as an amplifier.

With the aid of a graph of the input characteristics relating V_{BE} to I_B (not shown), we can set the bias level to produce the correct quiescent base current of 60 μa. This can be done because the quiescent bias is a static condition and regardless of the junc-

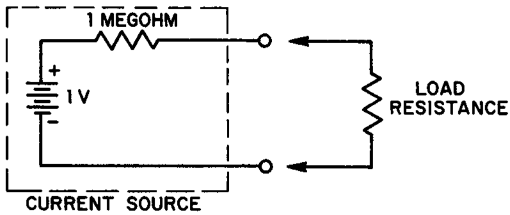


Fig. 49. A constant current source.

tion resistance, the correct operating point can always be obtained, using the load line method described in connection with the collector circuit. However, the *signal* current passing through the base must be sinusoidal. This cannot be done if the signal *voltage* is used as the controlling medium in Fig. 46. Irrespective of the operating point selected, the resistance of the junction will vary when the signal causes the voltage across the junction to change above or below the quiescent condition. This change in emitter junction resistance causes a change in the slope of the static input characteristics (Fig. 47). Since we have stated that a transistor is a current-operated device, it would seem desirable to *make* the sinusoidal signal voltage produce sinusoidal current variations across the emitter-base junction. To do this, we need to make the signal a current, rather than a voltage, source.

Everyone is familiar with voltage generators. A perfect voltage generator is a voltage source that never changes its voltage with a varying load. In practice, large generators or batteries approach this state. A current generator, however, is a device that delivers the same current to the load without regard to the load resistance. This device is not usually encountered in electronics. A current generator may be simulated by placing a voltage source in series with a very large resistance. as shown in Fig. 49. In Fig. 49, the voltage source has a potential of 1 volt and the series resistance is 1 megohm. If the external terminals were shorted—corresponding to a load of 0 ohms—the current flowing through the short circuit would be I = E/R = 1/1,000,000 = 1 μa. Now, if the short circuit were replaced by a 1-ohm resistor, the current would be

1/1,000,001 (approximately 0.999999 μa). This is still 1 μa, for all practical purposes. Even if the load resistance were raised to 100 ohms, the current would still be 1/1,000,100 = 0.9999 μa. A practical current source is created from a voltage generator in series with a high resistance.

If this current source represents the signal source, the input current (the base current) will be independent of the resistance of the emitter-base junction. This is shown in Fig. 50, where the signal has been coupled to the base through a capacitor and a high resistance, R_g. The capacitor in this case only serves the

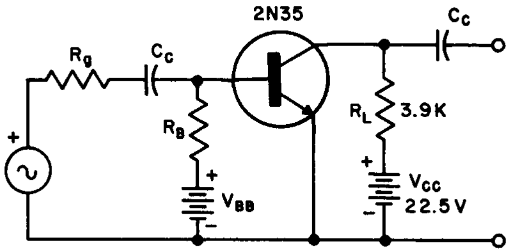


Fig. 50. A common emitter amplifier.

purpose of keeping the bias voltage out of the signal source. It is considered a short circuit to the signal voltage alternating component. R_g makes the base signal current independent of the emitter-base junction resistance. The bias current is also fed through a high resistance R_B. R_B serves the same purpose to the bias voltage source, as R_g serves to the signal voltage source—namely, to make the quiescent bias current independent of the emitter-base junction resistance. Note that the signal is fed to the emitter-base junction resistance *and* R_B in parallel. This in no way alters the previous discussion. Since R_B is so much larger than the emitter-base resistance, it may be neglected insofar as the signal is concerned.

It is not to be construed from the above that R_g is always a physical resistance. R_g usually represents the effective output resistance of the previous stage. In Chapter 3, it was mentioned that the output resistance or impedance of the common emitter and common base transistor configurations was much larger than their input resistances. Insofar as distortion is concerned, this is a most desirable situation, for it means that the output of each amplification stage will act as a current source for the subsequent stage. Where the input signal is extremely small (such as from a dynamic microphone, etc.), the signal excursions to either side of the *Q* point are reduced to such an extent that the variations in input resistance are insignificant. Hence, the necessity of a high input resistance of the signal source is not as stringent, and one transistor may be transformer-matched to another to achieve maximum power transfer. Chapter 5 deals with this type of *small signal* amplifier.

The common emitter amplifier of Fig. 50 is a more realistic circuit than the simplified circuits previously illustrated. Actually, there is a further improvement that seems rather obvious. The base and collector bias batteries have a common connection with the same polarity. Hence, there is no reason to be content with two bias

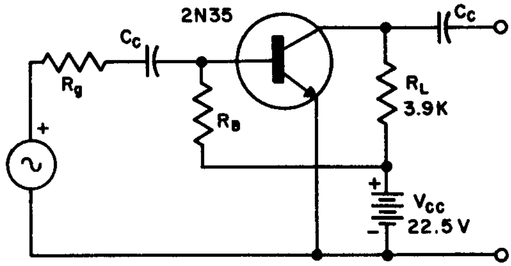


Fig. 51. A practical version of a common emitter amplifier.

batteries. The circuit is readily modified to that of Fig. 51. Calculation of the voltage gain of this amplifier is now much easier. The voltage gain may be practically stated as the ratio of the change in voltage across R_L, produced by a change in the signal voltage. It is realized that with the large series resistance R_g, the voltage gain will decrease; but, with the choice of a lower voltage gain, or excessive distortion, the former is the more desirable.

26. Establishing the Proper Emitter Bias

In the common emitter circuit of Fig. 52 (a redrawn version of Fig. 51), the quiescent base current is independent of the quiescent collector current. The method of base biasing in Fig. 52 is, therefore, called **fixed bias.** The collector load line is the same as in the circuit of Fig. 46 (as shown in the graph of Fig. 48) because R_L and V_{CC} are the same. The problem at hand is to select the proper bias resistor, R_B, to yield the required 60 μa.

Applying Kirchhoff's voltage law to the base loop, we proceed by equating the voltage drops around this loop to zero, as shown in Fig. 52.

$$-V_{CC} + V\ (\text{across}\ R_B) + V_{BE} = O \qquad (33)$$

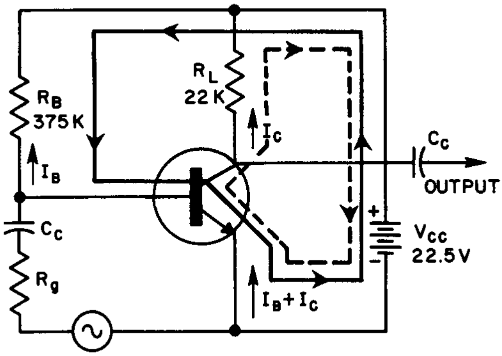


Fig. 52. A common emitter circuit.

or

$$-V_{CC} + I_B R_B + V_{BE} = 0 \qquad (34)$$

The previous discussion involving a current source, stated that the quiescent base current will be independent of the base-emitter voltage drop, providing R_B is large. With this in mind, we can logically neglect V_{BE} in the above equation. (In all future calculations, this will be done.) Considering that V_{BE} is only a few millivolts, as compared with V_{CC}, 22.5 volts, little error is made.

$$-V_{CC} + I_B R_B = 0 \qquad (35)$$

$$R_B = \frac{V_{CC}}{I_B} \qquad (36)$$

In this example, the quiescent I_B has been selected as 60 μa. Hence:

$$R_B = \frac{22.5 \text{ volts}}{60 \times 10^{-6} \text{ amps}} = 3.75 \times 10^{+5} \text{ ohms}$$

$$= 375{,}000 \text{ ohms} \qquad (37)$$

This establishes the quiescent point for the fixed bias common emitter amplifier.

Summary of Procedure for Determining the Quiescent Point for the Common Emitter Fixed-Bias Amplifier

1. For a given load resistance, write the collector voltage loop equation. Solve for V_{CE} with $I_C = 0$; solve for I_C with $V_{CE} = 0$. This yields the open and short-circuit voltage and current respectively. Locate these points on the collector characteristics. Draw the load line between them.

2. Select the operating points as approximately the middle of the region of equally spaced base current lines. This *Q* point establishes the base current needed.
3. To select the required base resistor, R_B, to yield this base current, write Kirchhoff's voltage law equation around the base loop, neglecting the base-emitter junction voltage drop. Solve for R_B.

27. Establishing the Q Point for the Common Base Amplifier

The procedure for establishing the *Q* point for the common base circuit is approximately the same as for the common emitter amplifier. The general method evolved for the common emitter will be applied to the *2N35* common base amplifier of Fig. 53.

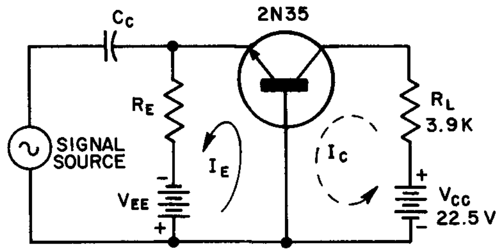


Fig. 53. 2N35 common base amplifier circuit.

Assuming that the same value of load resistance (3900 ohms) will be used, we determine the equation for the load line by writing Kirchhoff's voltage law equation around the collector loop, as indicated by the dotted line. It may be argued that the collector current does not actually travel from the collector to the base and back through V_{CC} and, hence, Kirchhoff's law does not apply. This is not true! A knowledge of the actual path of collector current is not necessary to write these equations, as long as a *loop* is formed where current *could* flow, if conditions were right. Kirchhoff's voltage law states only that the sum of the voltage drops around this closed loop equals zero.

In Fig. 53, the voltage drops around the collector loop are:

$$-V_{CC} + I_C R_L + V_{CB} = 0 \qquad (38)$$

Solving this for V_{CB} with the condition $I_C = 0$, we determine the open circuit voltage between the collector-base terminals. Thus:

$$V_{CB} = V_{CC} \qquad (39)$$

Now solving Equation (38) for I_C, with $V_{CB} = 0$, establishes the collector current with the collector-base terminals shorted.

$$I_C = \frac{V_{CC}}{R_L} \tag{40}$$

Note that the open and short circuit equations are identical for the common base and common emitter circuits. In practice, for either circuit, it is not necessary to write down the complete

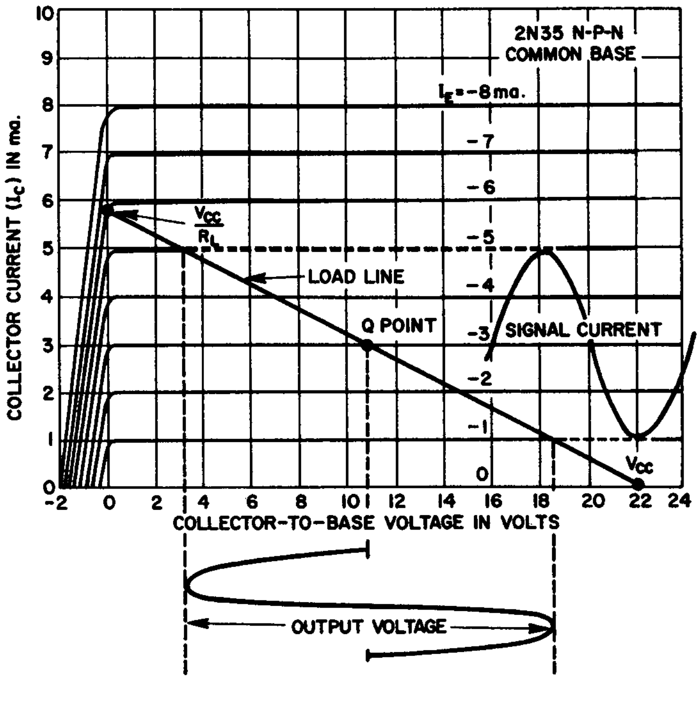


Fig. 54. Common base collector characteristics.

voltage loop expressions—just apply Equations (39) and (40). These points are located on the appropriate collector characteristic graph as indicated in Fig. 54. Since the values of V_{CC} and R_L are the same in this and the previous example, the points are the same, resulting in identical load lines.

Once the collector load line has been drawn, the quiescent point may be chosen. We will choose the Q point at $I_E = 3$ ma. This emitter current may be established in a manner similar to that previously used for the base current of the common emitter amplifier. With V_{EE} = 22.5 volts, we again write the expression for the voltage drops around the emitter loop as shown by the solid line in Fig. 53. Again, we assume that the emitter current flows from the base to emitter, through the emitter bias battery, and back to the base again. This reasoning is the same as that for the path for collector current. As long as we account for *all* of the voltage drops around the closed emitter-base loop, our assumptions are justified. The voltage loop equation is:

$$-V_{EE} + V_{EB} + I_E R_E = 0 \tag{41}$$

Applying the same method to the common emitter circuit, we neglect the few millivolts of drop across the emitter-base junction and the equation becomes:

$$-V_{EE} + I_E R_E = 0 \tag{42}$$

or

$$R_E = \frac{V_{EE}}{I_E} \tag{43}$$

Substituting the value for the quiescent I_E and V_{EE} of the circuit, R_E becomes:

$$R_E = \frac{22.5}{3 \times 10^{-3}} = 7.5 \times 10^3 = 7500 \text{ ohms} \tag{44}$$

This establishes the quiescent point for the common base amplifier.

28. Establishing the Q Point for the Common Collector Amplifier

As with the vacuum tube cathode follower, establishing the quiescent point for the common collector (emitter follower), is rather involved, compared to the other two configurations. The emitter follower has an emitter resistor that is common to both the input and output circuits. Hence, the collector and base circuits are now dependent upon one another; in the other configurations, they were independent. With the common base and common emitter circuits, the input bias current could be specified without regard to the collector circuit—as evidenced by Equations (43) and (46). Now, however, because of this common resistor, the collector current will be present in the analogous equation for the common collector circuit. Consequently, there are two equations and two unknowns which must be solved simultaneously, before the *Q* point may be determined. A rather straightforward graphical method has been devised.

In Fig. 55 (A), a typical emitter follower circuit is shown. This circuit has the primary function of impedance matching, as does the cathode follower. Figure 55 (B) shows a somewhat more practical version of the same circuit, utilizing a common bias battery. The collector load line may be obtained in the normal

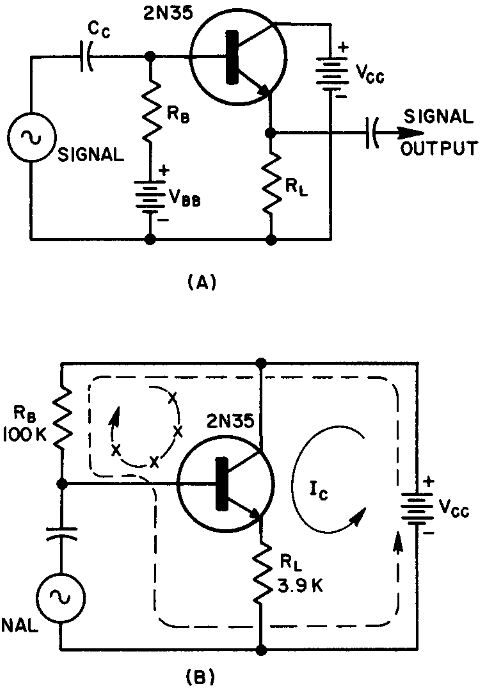


Fig. 55. Common emitter follower circuits: (A) typical; (B) practical.

way, by writing the expression for the voltage drops around the collector loop. The path of the collector current for this equation is shown by the solid line. Using the same rules developed for the other two configurations, we may write:

$$-V_{CC} + I_C R_L + I_B R_L + V_{CE} = 0 \qquad (45)$$

Here, we are presented with the fact that both the base current and the collector current flow through the same load resistance. Each current contributes to the voltage drop across this resistance, as evidenced by the $I_C R_L$ and $I_B R_L$ terms in Equation (45). This might cause trouble were it not for the fact that the magnitude of the base and collector currents are so different. Whereas the base current is in microamperes, the collector current is in milliamperes. For all practical purposes, then, this minute base current can be neglected with respect to the much larger collector current. The base current can little add or detract from the voltage drop already established by the much larger collector current. Consequently, as an engineering approximation, we are justified in neglecting the $I_B R_L$ term in Equation (45). Hence:

$$-V_{CC} + I_C R_L + V_{CE} = 0 \tag{46}$$

This is the same equation obtained for the collector load line in the common base and common emitter configuration. Skipping the intermediate steps and solving directly for the short circuit current and open circuit voltage, we have:

$$V_{CE}\Big|_{I_C = 0} = V_{CC} = 22.5 \text{ volts} \tag{47}$$

$$I_C\Big|_{V_{CE} = 0} = \frac{V_{CC}}{R_L} = \frac{22.5 \text{ volts}}{3900 \text{ ohms}} = 5.77 \text{ ma} \tag{48}$$

The two points are located on the common emitter collector characteristics of Fig. 56, and the load line is drawn in the usual manner. (The emitter follower is but a modified common emitter amplifier. Hence, the common emitter characteristics are used for both configurations.)

The next step is to write down the voltage drops around the base loop (the path of the base current is indicated by the dotted line in Fig. 55 (B)).

$$-V_{CC} + I_B R_L + V_{BE} + I_B R_B + I_C R_L = 0 \tag{49}$$

Again, it is seen that both the collector and base currents flow through the load resistance, R_L. Now, however, we are very much interested in the small base current, and so its presence cannot be neglected with respect to the collector current. Solving the equation for I_B and neglecting (as usual) the minute V_{BE}, we obtain:

$$-V_{CC} + I_B (R_L + R_B) + I_C R_L = 0 \tag{50}$$

$$I_B = \frac{V_{CC} - I_C R_L}{R_L + R_B} \tag{51}$$

In this last expression, we know V_{CC}, R_L, and R_B because they were given. I_B and I_C are the only unknowns in the expression. Picking a value for I_C will yield a corresponding value for I_B. However, this is the function of the load line equation. If this is

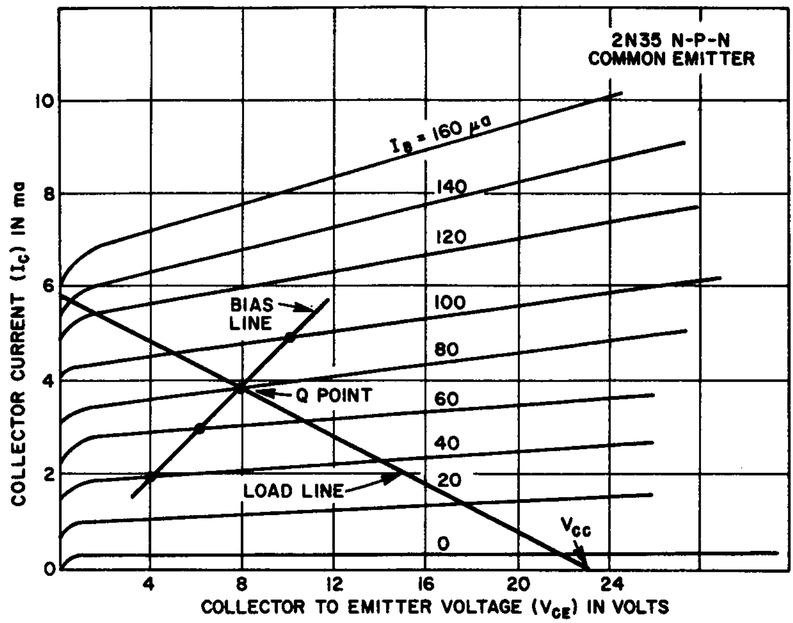


Fig. 56. Common emitter collector characteristics.

true, then no new information is gained by writing the base loop expression. It will be presently shown that Equation (51) is but a re-arranged version of Equation (45). To do this, we will have to perform some mathematical gymnastics.

First, V_{CE} is *always* the algebraic sum of $V_{CB} + V_{BE}$,

or

$$V_{CE} = V_{CB} + V_{BE} \qquad (52)$$

When (as usual) we neglect the minute base-to-emitter voltage drop, Equation (52) becomes:

$$V_{CE} = V_{CB} \qquad (53)$$

Second, write the voltage drop equation around the loop in Fig. 55 (B) containing R_B and defined by the line containing X's.

$$-V_{CB} + I_B R_B = 0 \qquad (54)$$

or

$$V_{CB} = I_B R_B \qquad (55)$$

But, we now know that $V_{CB} = V_{CE}$ from Equation (53). Hence, V_{CE} may be substituted for V_{CB} in Equation (55), yielding:

$$V_{CE} = I_B R_B \qquad (56)$$

This is a very important substitution and will be referred to again. Returning to Equation (45) and substituting $I_B R_B$ for V_{CB},

$$-V_{CC} + I_C R_L + I_B R_L + I_B R_B = 0 \qquad (57)$$

which, when solved for I_B, is identical to the load line Equation (51). Hence, no new information has been gained by writing the base loop expression in the form of Equation (51).

This problem may be eliminated by substituting in Equation (51) for I_C, its value as defined by the load line Equation (46). Therefore, Equation (51) now becomes:

$$I_B = \frac{V_{CC} - R_L \overbrace{\left(\frac{V_{CC} - V_{CE}}{R_L}\right)}^{= I_C \text{ from Equation (46)}}}{R_L + R_B} \qquad (58)$$

In this expression, only I_B and V_{CE} are unknown. When simplified, this expression becomes:

$$I_B = \frac{V_{CC} - R_L \left(\frac{V_{CC} - V_{CE}}{R_L}\right)}{R_L + R_B}$$

$$I_B = \frac{V_{CC} - V_{CC} + V_{CE}}{R_L + R_B} = \frac{V_{CE}}{R_L + R_B}$$

or

$$V_{CE} = I_B (R_L + R_B) \qquad (59)$$

When values are substituted for our example, $R_B = 100{,}000$ ohms and $R_L = 3900$ ohms, in this latest expression, we have:

$$V_{CE} = I_B \text{ (103,900 ohms)} \qquad (60)$$

Equations (59) and (60) define what is called a *bias line*. The equations indicate that V_{CE} and I_B are related in a certain way. If values are picked for I_B and the resulting V_{CE} is evaluated in each case, several points may be located on the collector characteristics, and the bias line drawn between the points. The point where the bias line intersects the load line, represents the simultaneous solution of Equations (46) and (51), and indicates the quiescent point of operation. In the example at hand, three points will be selected by choosing three convenient values of I_B (values that base current lines have been drawn for). The results are:

V_{CE} (in volts)	I_B (in μa)	
4.18	40	
6.19	60	(61)
10.30	100	

The bias line resulting from these points intersects the load line at $I_B \cong 78\ \mu$a, $I_C \cong 3.75$ ma, and $V_{CE} \cong 7.9$ volts. This is the quiescent point of operation for this circuit. The steps in obtaining the Q point may be summarized as follows:

1. Write the collector voltage loop equation and draw the load line equation in the usual way, using Equations (47) and (48).
2. Write the base voltage loop equation and solve for I_B. Substitute in this equation the value for I_C obtained from the load line equation. This always reduces to Equation (59).
3. Insert circuit constants in Equation (59).
4. Select several convenient base currents and solve for the corresponding V_{CE}. Plot these points so that the bias line crosses the load line. The intersection of the load and bias lines defines the Q point.

29. Finding R_B

The previous discussion of the common collector circuit covers the procedure for finding the Q point, when the circuit constants (R_L, R_B, and V_{CC}) are given. The practical problem, from a design standpoint, is how to find the correct R_B once the load line has been drawn. (Factors governing the selection of R_L and V_{CC} will shortly be discussed.)

Once the load line is drawn, to find the appropriate R_B, it must first be decided where the Q point is to be located. For clarity, it will be assumed that the Q point has been chosen at the same point as in the solution of the previous problem (illustrated in Fig. 56). If we *assume* that this is to be the Q point, we may also *assume* the values of I_B and V_{CE} corresponding to that point on the collector characteristics. In this instance they are 78 μa and 7.9 volts, respectively. With this information, Equation (59) may be solved directly for R_B, since R_B is the only unknown in the equation.

$$V_{CE} = I_B\,(R_L + R_B)$$

$$R_L + R_B = \frac{V_{CE}}{I_B}$$

$$R_B = \frac{V_{CE}}{I_B} - R_L \qquad (62)$$

$$R_B = \frac{7.9 \text{ volts}}{78\,\mu\text{a}} - 3.9 \times 10^3 \text{ ohms}$$

$$R_B = 101 \times 10^3 - 3.9 \times 10^3 = 97{,}100 \text{ ohms} \qquad (63)$$

This is within 3% of the value originally selected for R_B, and is excellent for a graphical solution. In fact, with a required R_B of 97,100 ohms, a 100,000 ohm-resistor would probably be used as the closest standard value available.

This method of selecting R_B from a specified Q point is general and may even be used when there is a separate base bias battery, as shown in Fig. 55 (A). Of course, in this case, V_{BB} would be present in Equation (51), instead of V_{CC}. Hence, it will not cancel with the V_{CC} term introduced in Equation (58), and the simplified expression of Equation (59) will contain this V_{BB} term. Other than this, the procedure is exactly the same as outlined for the single bias battery case. (See problem 7 at the end of this chapter.)

30. Distortion and Regions of Operation

The object of a transistor amplifier is to convert variations of the input signal current into variations of the output collector current. The input and output current in any successful amplifier must be, as nearly as possible, exact replicas of one another—other than the fact that they may not be of the same amplitude. In the common emitter circuit, the output current is larger than the input current. In the common base circuit, they are approximately equal. In either case, the essential job of the amplifier is to transfer current from a low resistant input circuit to a high resistant output circuit and, in the process, not contribute any signal of its own.

The ability of the collector current to faithfully reproduce the shape of the input signal current (for any configuration) depends upon the extent that the base or emitter current lines (depending

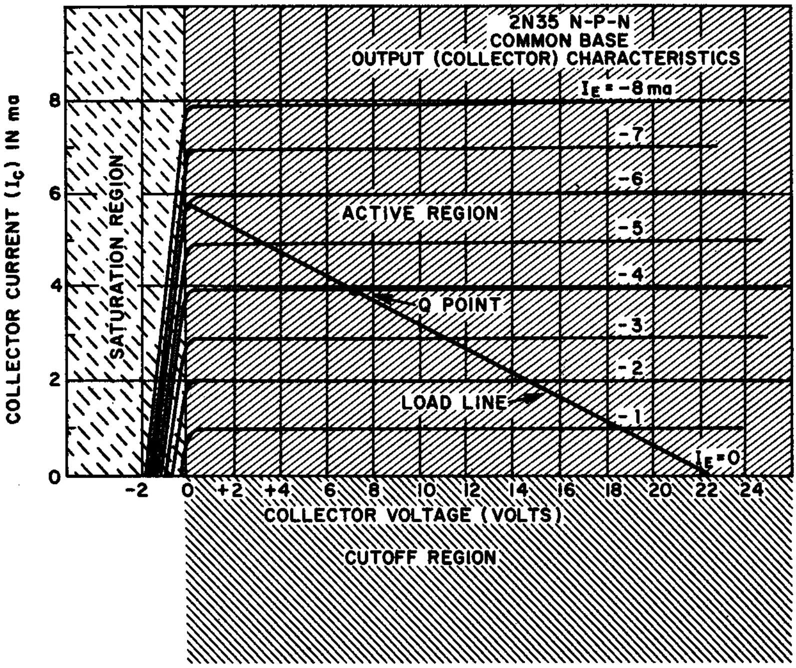


Fig. 57. Saturation region of a common base amplifier.

upon the type of amplifier being considered) are evenly spaced. This is a very important consideration and is never completely realized in practice.

31. Active Region

Examination of Figs. 48 and 54 shows that the base current and emitter current lines, respectively, are almost evenly spaced. This is the region in which the load line is drawn and (in accordance with the above discussion) is the only region where successful amplification with the least distortion can take place. It is called the *active region.* This is the region that has been used for all previous discussions of the common base and common emitter amplifiers. *The active region is characterized by the fact that the emitter is forward biased and the collector is reversed biased.* In this region the collector and emitter current are related by Equations (13) and (24). Normally, the quiescent point will be taken as right in the middle of this region as shown in Fig. 57. From the figure, it can be seen that, as the emitter current increases in increments of 1 ma from the Q point, the collector increases in approximate increments of 1 ma. As long as this happens, the amplifier is operating in the active region and distortion is at a minimum. Theoretically, then, β or α (depending on the configuration) should remain constant throughout the active region.

32. Saturation Region

If the emitter current in the common base amplifier of the above discussion increases beyond a point of about 5.9 ma, the emitter current lines tend to crowd together in the saturation region of Fig. 57. Examination of the figure shows that this crowding takes place to the left of $V_{CB} = 0$. Here, the collector voltage is of the opposite polarity and produces a positive instead of a negative bias. This *region in which both the emitter and collector are forward biased is called the saturation region.* It is indicated on the graph to the left of the $V_{CB} = 0$ point. Assuming that the Q point remains at the point indicated on the graph, this saturation condition will exist for part of the input signal cycle, if the input signal current is large enough to cause the emitter current to rise above a critical value. The fact that V_{CB} will reverse and cause a forward collector bias even when V_{CC} remains unchanged in magnitude and polarity becomes apparent from examination of the collector voltage loop equation (Equation 38). (This equation is for the common base amplifier of Fig. 53, for which the load line of Fig. 57 has been drawn.)

$$-V_{CC} + V_{CB} + I_C R_L = 0 \tag{64}$$

Solving this for V_{CB}:

$$V_{CB} = V_{CC} - I_C R_L \tag{65}$$

From Equation (65), it is noted that V_{CC} and I_C have an opposite effect on V_{CB}. Now V_{CC} and R_L of the relation are always constant, but if I_C increases sufficiently, the right hand side of the equation will become zero. This is the point that separates the saturation region from the active region. If the value of I_C increases still further, the effective value of V_{CB} becomes negative, and the transistor goes into the saturation region. The region of negative V_{CB} is shown in Fig. 57.

Solving Equation (65) for I_C with the condition that $V_{CB} = 0$, we find the critical collector current where saturation begins:

$$I_C \Big|_{Sat} = \frac{V_{CC}}{R_L} \qquad (66)$$

As long as the collector current exceeds this critical value, the transistor will remain in saturation. Of course, the collector current does not vary spontaneously by itself, but is controlled by the emitter current. The critical value of emitter current at which saturation begins may be estimated from Equation (65) by substituting αI_E [from Equation (15)][1] for I_C, hence:

$$V_{CB} = V_{CC} - (\alpha I_E)\, R_L \qquad (67)$$

Setting $V_{CB} = 0$ as the point where saturation starts:

$$0 = V_{CC} - \alpha\, I_E R_L \qquad (68)$$

or

$$I_E \Big|_{Sat} = \frac{V_{CC}}{\alpha\, R_L} \qquad (69)$$

When the transistor enters the saturation region, *bottoming* is said to have taken place, since the emitter current lines abruptly drop to the *bottom* of the characteristics in this region. The most obvious evidence that the transistor has bottomed, or entered the saturation region, is that large changes in emitter current produce almost no change in collector current. In Fig. 21, for example, a change of emitter current from -6 to -8 ma (along the load line) only produces a change of about 0.3 ma collector current.

If a signal is applied to the transistor that drives it to saturation, that portion of the input signal that exceeds the critical saturation current will be clipped and will not appear in the output signal. This is demonstrated in the graph of Fig. 22 by an input signal that causes the emitter current to exceed the critical value [as established by Equation (69)]. Here $R = 3900$ ohms, $V_{CC} = 22.5$ volts, and $\alpha = 0.98$, for the *2N35* common base amplifier of Fig. 53. Applying Equation (69), we find that analytically, the critical emitter current is:

$$I_E\ (Sat) = \frac{V_{CC}}{\alpha\, R_L} = \frac{22.5}{(0.98)\ (3900\,\text{ohms})} = 5.9\,\text{ma} \qquad (70)$$

---

[1] I_{CO} is considered negligible and for the common base, $I_C \cong a I_E$.

From Fig. 57, it can be seen that saturation, or bottoming, begins at slightly less than 6 ma, about 5.9 ma. At this point the 5.9 ma emitter current line (if one were shown) crosses both the load line and where $V_{CB} = 0$. As can be seen, the entire part of the signal that exceeds this critical emitter current—namely that from

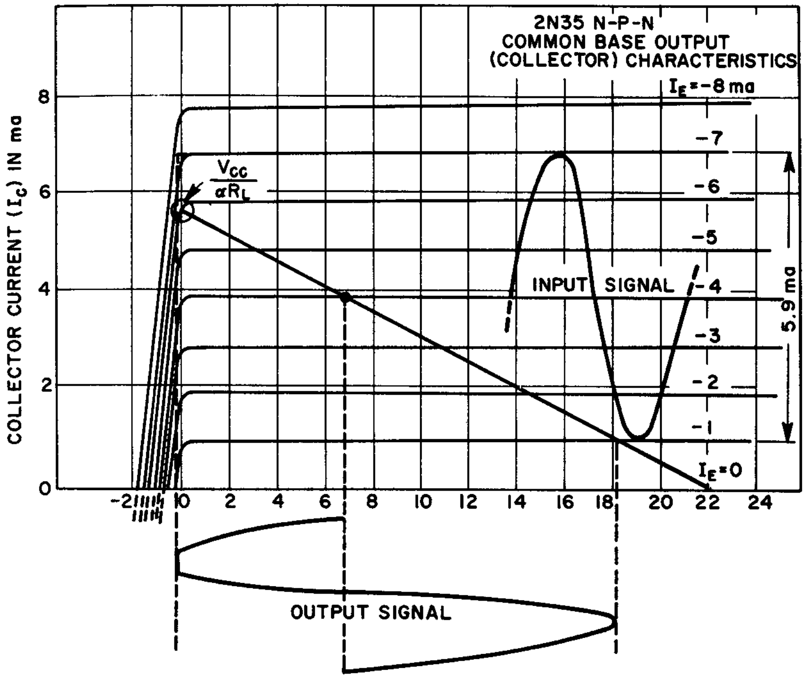


Fig. 58. Common base collector characteristics with applied a-c signal driving transistor to saturation.

approximately 6 to 7 ma—is clipped off and is not present in the output current. The clipped voltage swing across the load resistance is shown on a vertical axis at the bottom of the graph.

If the transistor is to be used as a linear amplifier stage, extreme care must be exercised to insure that the input signal does not rise to a value such as to drive the transistor to saturation, or into the nonlinear region, where the output is not an amplified version of the input.

33. Cutoff Region

The *cutoff region* of a transistor is defined as the region where both the collector and emitter are in reverse bias. It is shown in Fig. 57. The existence of this region is obvious when it is remembered that only when the emitter is in forward bias can minority carriers be injected into the base material. With no

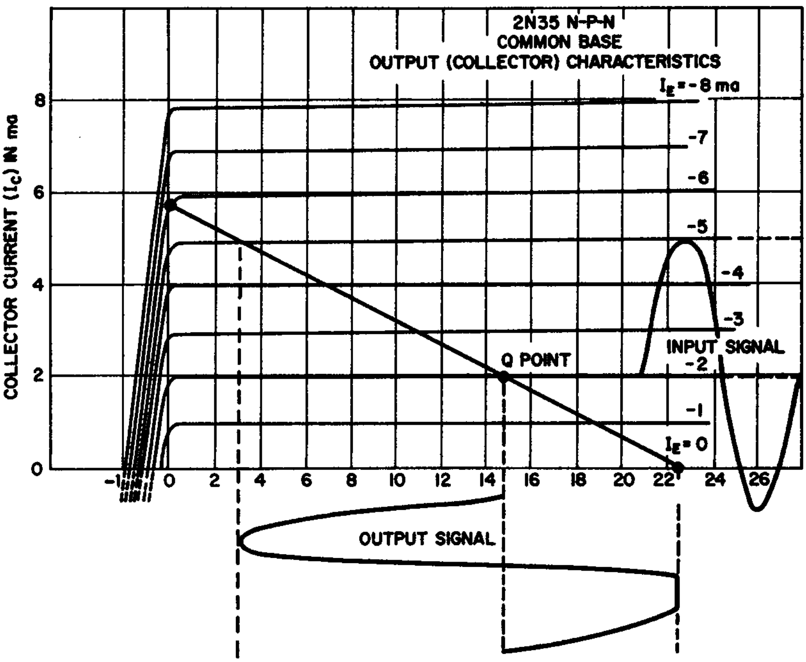


Fig. 59. Common base collector characteristics—different quiescent point.

minority carriers being injected, the only collector current is the reverse saturation current. If the input signal current causes the emitter voltage to go to zero, or into reverse bias, that portion of the signal present in the cutoff region will be clipped from the output signal, as it is in the saturation region.

In Fig. 59, an input signal is applied to the same common base amplifier, except that this time the stage is operating at a lower

Q point (referring to the load line starting at 22.5 volts). The part of the signal that drives the emitter into cutoff is clipped from the output signal.

34. Interrelation of Q Point, V_{CC}, and Regions of Operation

From the previous discussion, it should be evident that there are several reasons why the Q point is usually established in the middle of the active region. In both Figs. 58 and 59, the input signal is of the same peak-to-peak (p-p) amplitude, 5.9 ma. In the first example, the Q point was established at $I_E = 4.0$ ma, and the top part of the input signal is seen to drive the transistor into saturation. In the second example, the bottom part of the same 5.9 ma peak-to-peak signal causes the transistor to enter the cutoff region when the Q point is established at $I_E = 2.0$ ma. Obviously then, picking a Q point too near either the cutoff or saturation end of the load line will result in a clipped and distorted output signal. Of course, the input signal may be kept quite small, thus allowing a greater margin in the selection of the Q point, without danger of clipping.

If in Fig. 58, the Q point had been selected at a point slightly lower than 4.0 ma, the 5.9 ma peak-to-peak signal would just fit in the active region—neither driving the circuit to cutoff nor saturation. This is the extreme, or limiting case, where the peak-to-peak emitter signal just equals the value $V_{CC}/\alpha R_L$. To avoid clipping of either the negative or positive peaks, the operating point must be located at a point of $I_E = V_{CC}/2\alpha R_L$ or, exactly in the middle of the active region. As has been mentioned, it is not a good practice to use an input signal this large, as temperature variations, battery drain, etc., will cause the Q point to shift somewhat. In this limiting case, any drift of Q point will cause clipping, due to saturation or cutoff.

Then, it is advantageous to use as large a V_{CC} as possible, if large signals are to be handled without distortion. This so-called dynamic range of the amplifier is a direct function of the value of V_{CC}. To illustrate the concept, the same load line has been drawn for three values of V_{CC} in Fig. 60. Notice that for the left-hand load line (corresponding to $V_{CC} = 2$ volts), the input signal of 1 ma peak-to-peak nearly approaches saturation and cutoff. This is because a small V_{CC} was chosen for the value of load line resistance. For the middle load line with a V_{CC} of 8 volts, the amplifier will handle a peak-to-peak input signal of 4 ma without overloading (*i.e.*, reaching saturation or cutoff). For the load line using $V_{CC} = 14$ volts, an input peak-to-peak signal of 7 ma can be applied. In each case, the load line corresponds to an R_L of 2000 ohms. Thus, although the input signal handling capacity, the dynamic range, of the amplifier has been increased by increasing the value of V_{CC}, the amplification has not increased. In each case, the ratio of input current to output voltage is 2.

In Fig. 61, the same amplifier is operated with a constant V_{CC}, but with three different load line resistances. As can be seen, as the resistance of the load line *increases,* the amplification *increases,* and the dynamic range *decreases.* As the resistance of the load line *decreases,* the amplification *decreases,* and the dynamic range *increases.* Thus, although the slope of the load line governs the amplification, it also affects the dynamic range.

The larger the V_{CC} used, the larger the input signal that can be handled without distortion. There is, however, a limit on the tolerable value of V_{CC}. If it is too large and the transistor happens to go into cutoff momentarily, then $V_{CB} = V_{CC}$ [from Equation

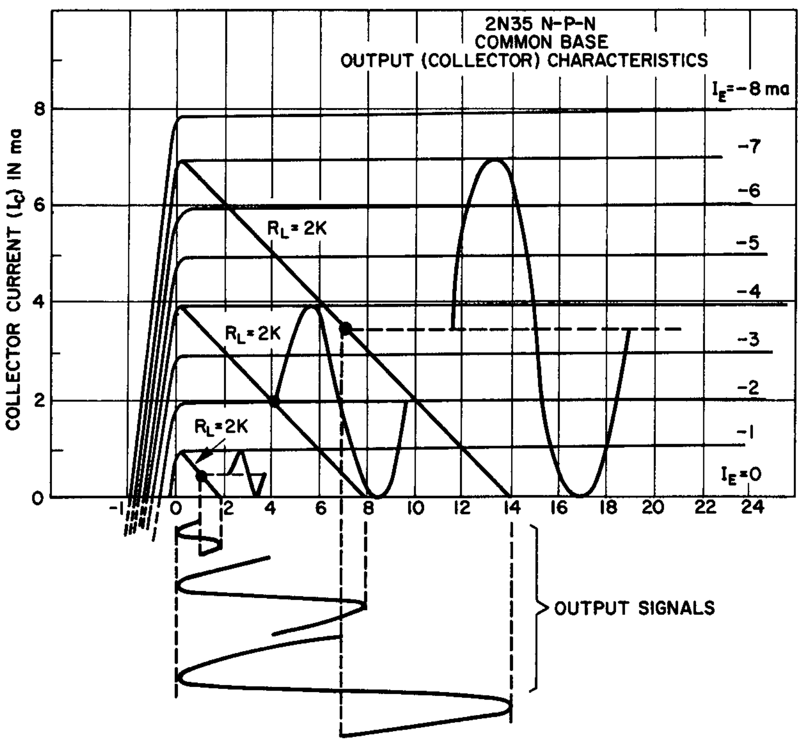


Fig. 60. The dynamic range of a common base amplifier.

(38)]. For this reason, V_{CC} is not usually chosen greater than the maximum tolerable reverse voltage across the collector-base junction, unless provisions in the signal amplitude and/or bias circuitry are made, to prevent the stage from ever entering the cutoff region. If this should occur, the ensuing Zener breakdown of the collector junction will possibly ruin the transistor. Of course, if a very large R_L is used, the current will not damage the device.

Up to this point in our study of transistor circuits, we have not considered the effect of the load into which the signal is fed. We have only considered signal voltage variations existing across the collector load resistance. The existence of a load connected to R_L through the coupling capacitor or transformer often has the effect of reducing the effective load presented to the transistor. Note that an excessive increase in the magnitude of R_L is not the answer to the problem of securing greater amplification.

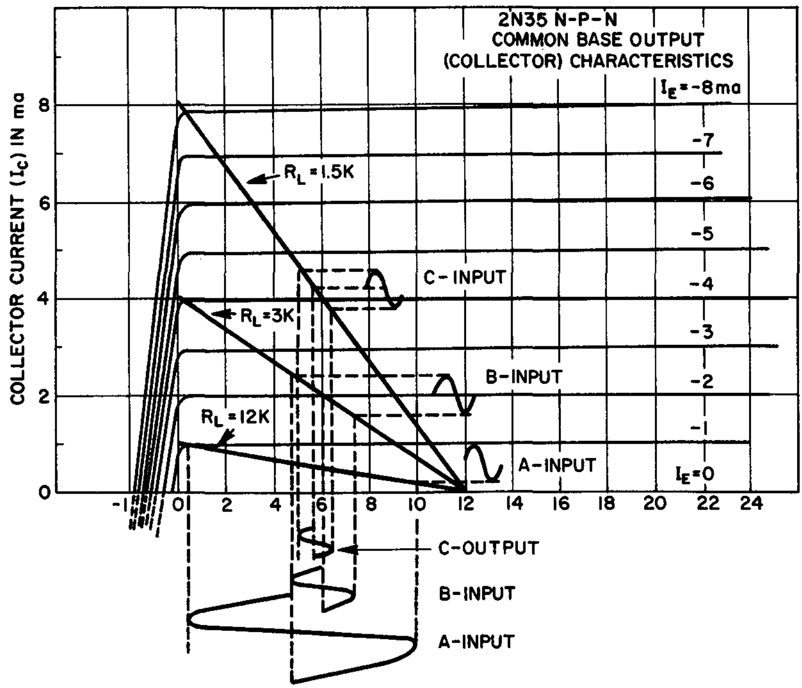


Fig. 61. Three different load resistances applied to the common base amplifier of Fig. 60.

35. Review Questions

1. Why is the reverse saturation current more of a problem with the common emitter configuration than with the common base configuration?
2. How is it possible to observe the difference in the output resistance between the common base and common emitter amplifiers by examination of their static collector characteristics?
3. What restrictions does the load line place on the operation of an amplifier?
4. What is the advantage of having a high output resistance of a signal source connected to the input of a transistor amplifier?
5. Explain the difference between V_{BE}, V_{be}, v_{BE}, and v_{be}.
6. For what purpose is the bias line used?
7. With reference to Fig. 55 (A), assume that the Q point is to be established at $I_B = 40$ ma. Calculate the required R_B if $V_{BB} = 5$ volts, $V_{CC} = 22.5$ volts and $R_L = 3900$ ohms. Draw the load and bias lines on one of the common emitter output characteristics.
8. For the common base amplifier of Fig. 53, $R_L = 3300$ ohms. The Q point is to be at $I_E = 4$ ma.
   (a) Find the proper R_E, assuming V_{EE} and V_{CC} are 10 and 20 volts, respectively. (b) What peak amplitude of a sinusoidal input current is allowed before driving the device into saturation? (c) Cutoff? (d) Which is reached first as the amplitude of the input signal is slowly increased from zero?
9. Why is it often advisable to limit V_{CC} to a value not greater than V_{CB}? Under what conditions is it permissible?
10. If it was known that the input signal would always cause operation in the active region, could V_{CC} be safely made greater than V_{CB}?

Chapter 5

SMALL SIGNAL ANALYSIS

In Chapter 4, the graphical characteristics of a transistor were described. It was shown that as the input signal varied the bias point, the collector current varied, and amplification was achieved. This method of analysis is usually used where large signals—causing large shifts from the quiescent point—are involved, such as in a power amplifier. When the input signal is very small, it is difficult, if not impossible, to graphically measure the shift of the Q point.

A new method of analysis is presented in this chapter. It makes use of a technique of substituting an *equivalent* electrical circuit in place of the transistor. The equivalent circuit, once developed, can be used to simulate the transistor in nearly all aspects of performance. This equivalent transistor circuit is developed with the assumption in mind that the characteristics of the transistor (α input and output impedance, etc.) do not change as the input signal moves the Q point back and forth on the load line. The assumption is accurate to a good approximation, when the input signal is *small*. Hence, the name *small signal analysis.* (Unless otherwise stated, only small signals will be dealt with here.)

36. Static and Dynamic Quantities

Consider the static input characteristics of a common emitter junction transistor, shown in Fig. 62. The input resistance of the transistor's emitter-base junction may be measured at any point on the curve, by merely observing the ratio of emitter-to-base voltage to the base current. This is called the *static* or dc input resistance. It exists only at the point of measurement. For a different base current, a different static input resistance is observed.

Resistance on any graph possessing voltage and current coordinates is the slope of the curve, or line. The slope is equal to the ratio between the vertical and horizontal travel of the line. Thus, in Fig. 62, the slope of the dotted line extending from zero to point A is R = 0.3 volts/190 μa = 1500 ohms. This is approximately the dc resistance (as measured by the volt-ammeter method) across the emitter-base junction. The slope, or tilt of the dotted line would become greater or less, as the junction resistance correspondingly became greater or less.

Assume that in Fig. 62, a given quiescent base current has been established at point *A*. What is the resistance to *a small increase* in base current? If the input signal swings the point on the curve from *A* to *B*, the increase in emitter-base voltage is only 0.01 volt for an increase in base current of 35 μa. The resistance for this *increase* is only 285 ohms. Clearly, this is different from the dc resistance of 1500 ohms. It is this 285-ohm dynamic resistance that will be presented to an ac signal applied to the junction.

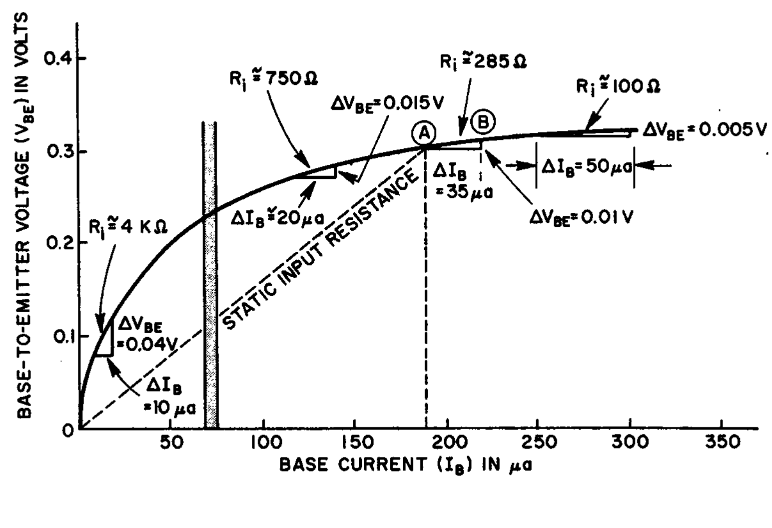


Fig. 62. Typical transistor static input characteristics.

Now establishing the fact that the dynamic resistance changes for different values of quiescent base current, how is the dynamic resistance measured accurately, if it is continually changing? The problem is solved practically by drawing a small right triangle under the curve, as indicated in the illustration. Only when the hypotenuse of that triangle (always a straight line) lies on the

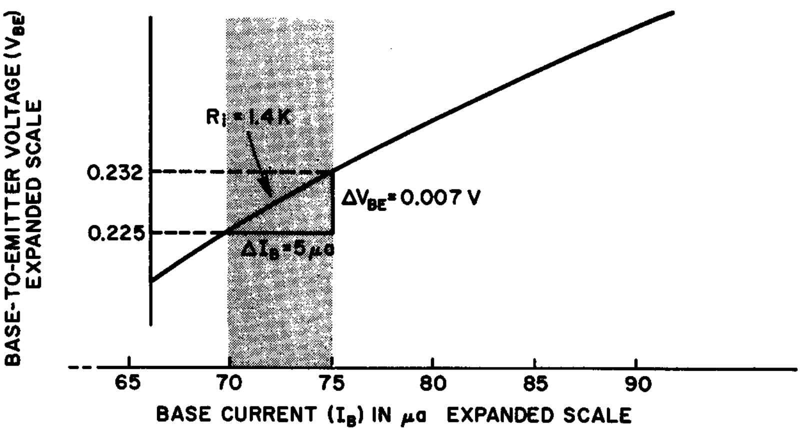


Fig. 63. Expanded portion of Fig. 62.

curve are we able to state the dynamic resistance. This required that the triangle be exceedingly small, so that the curve "looks" like a straight line at that point. It can be seen that this is not strictly true for the triangles drawn. In Fig. 63, the small shaded section has been expanded. The portion of the total curve represented in this shaded portion is so small that it is nearly straight. The hypotenuse of the right triangle drawn under this portion of the curve lies on the curve nearly perfectly, and the dynamic resistance can be stated with some confidence. Too, if the input signal does not drive the operating point outside the shaded section, it can be considered that the input resistance is constant over this section, for the slope is constant. What has been said about the input resistance remaining constant for small changes in input signal applies to the other parameters of a transistor.

Under the condition that a transistor's characteristics remain constant during a small signal swing about the operating point, these characteristics may be measured and a simple electrical circuit, an equivalent circuit may be constructed, possessing the same *constant* characteristics that were measured. This equivalent transistor circuit may then be substituted for the transistor in the external circuit. Once this has been done, an *analytical,* rather than *graphical* analysis of the system is possible. Small signal

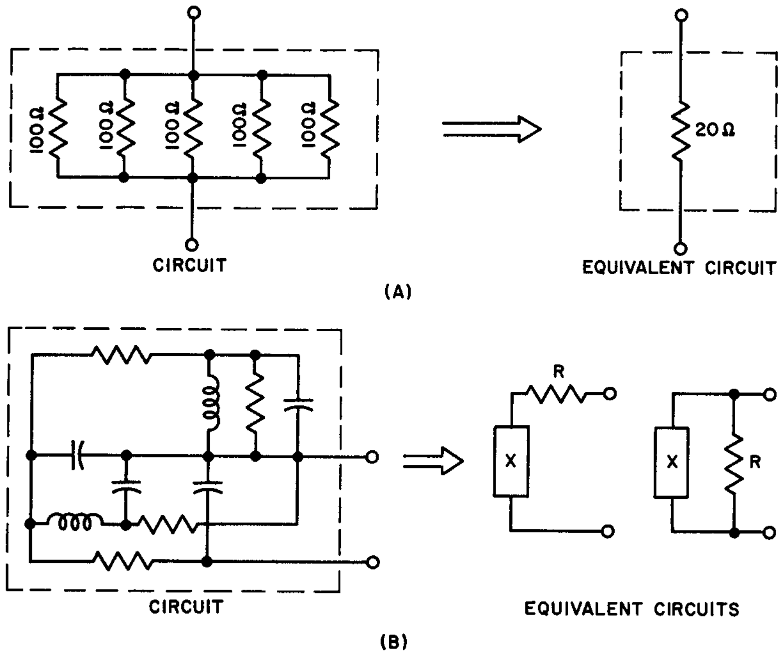


Fig. 64. Equivalent circuits: (A) simple example; (B) complex example.

analysis of a transistor circuit is quicker and more direct, and brings to light many facets of circuit performance not evident from a graphical approach.

37. Equivalent Circuits—"Black Boxes"

An equivalent circuit is a circuit that *electrically* represents the performance of some device. About the simplest example of an equivalent circuit is illustrated in Fig. 64 (A). Here, a group of five 100-ohm resistors are connected in parallel. The total resistance measured would be 20 ohms. Thus, an equivalent electrical circuit for these resistors would be *one* 20-ohm resistor. A dotted line has been drawn around the resistors in each case to denote that, to the world outside THE BOXES ARE EQUAL. Indeed, if one excludes the difference in power dissipation between the boxes, it is impossible to tell from an external measurement how many resistors are present. Only a presence of 20 ohms resistance is sensed. Thus, from only an external measurement of what is in the box, one may reproduce the electrical equivalence. If a capacitor or inductor or resistance, or any combination or number of these were in the box, they could only present *one* impedance to the output terminals. Once this terminal impedance has been determined by an external measurement, then a single resistor in series, or parallel with, the proper reactance could replace a very complicated circuit. This is symbolized in Fig. 64 (B). When there are sources of voltage or current involved, several electrical theorems must be employed to develop an equivalent circuit.

38. Circuit Theorems

Two fundamental theorems of modern circuit theory are *Thévenin's* and *Norton's theorems.* The use of Thévenin's theorem may be illustrated by Fig. 65 (A). Here, a dry cell has been connected to a 9-ohm external resistance. The cell has an open circuit voltage of 10.0 volts and an external resistance of 1 ohm. Thévenin's theorem says that if the imperfect battery is represented by a perfect voltage source[1] of 10 volts in series with the battery's internal resistance, then viewed from the terminals, an exact equivalence exists. The example may be so trivial that the significance of the theorem may be overlooked. What is important is that at no time is it necessary to open the battery to determine its characteristics. All that is necessary is to measure the resistance from the external terminals, along with the open-circuit terminal voltage.

In Fig. 65 (B), the same battery is connected to a somewhat more complicated circuit. What is desired is to replace the battery and the two resistors with a single voltage source and a single resistor. To begin to simplify this circuit, we could represent the

[1] A perfect or constant voltage source is one that maintains a constant voltage across its terminals regardless of the current flowing—even under short circuit conditions, where an infinite current flows. This theoretical representation is only realized in practice with a battery of infinite capacitance, but it is necessary to use with Thévenin's theorem. A voltage source is considered to have no resistance, only voltage.

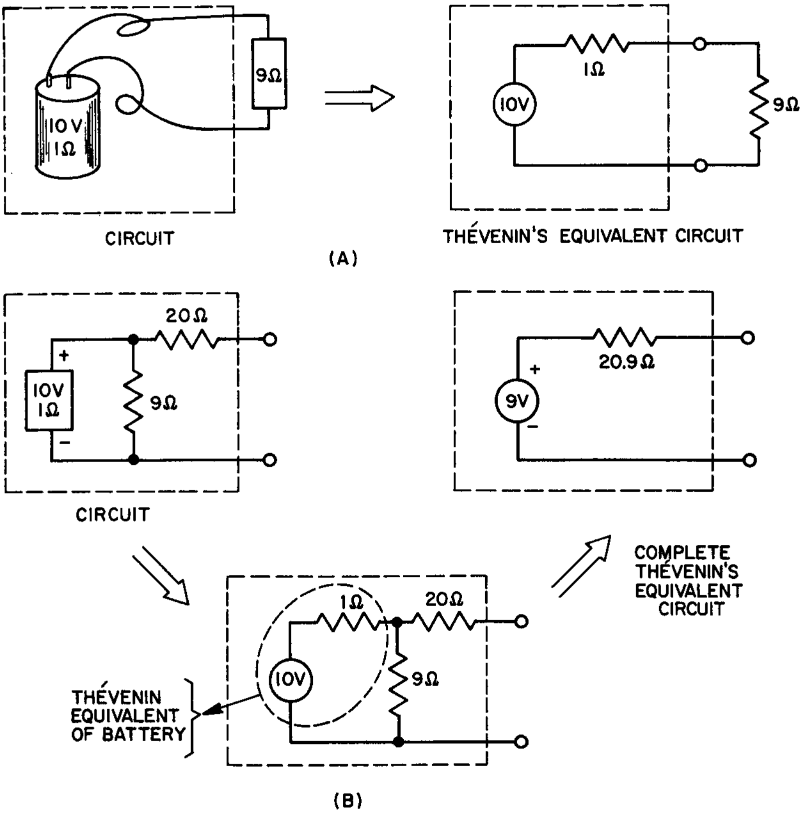


Fig. 65. Thévenin's equivalent of active circuits: (A) simple; (B) complex.

battery by its Thévenin equivalent. This is done inside the circle, as indicated. Next, we measure the open-circuit voltage existing at the output terminals of the network. Since with open circuit conditions at the output, no current will flow through the 20-ohm resistor, the open-circuit voltage exists across the 9-ohm resistor. The battery current is then: $I = E\,(R_1+R_2)$ $10\,(1+9) = 1$ ampere. The voltage drop across the 9-ohm resistor is then: $E = IR = 1 \times 9 =$ 9 volts. Therefore, 9 volts also exists at the output terminals. This is the voltage source for the Thévenin's equivalent circuit. The external resistance of the network (the resistance presented to the output terminals) may be measured in practice, by a number of methods. It may be calculated by noting that the 20-ohm resistor is in series with the parallel combination of 1 and 9 ohms. The parallel combination reduces to 0.9 ohm; thus, the total output resistance is $0.9 + 20 = 20.9$ ohms. Then, the Thévenin's equivalent may be represented as a perfect voltage source of 9 volts in series with a resistance of 20.9 ohms. It is impossible to tell from an external measurement that the two circuits are different. This is why the box made by a dotted line is often called a **black box.** It denotes that we are not interested in whether the equivalent circuit approximates the original circuit layout, but only whether it reacts on the output terminals in the same way.

Norton's Theorem may also be used to construct an equivalent circuit, for the same networks as illustrated for Thévenin's Theorem. If the internal resistance (as measured across the output terminals) is connected in parallel with a perfect, or constant, current source[2] equal to the current flowing across the shorted output terminals of the network, an exact equivalent circuit is formed. In Fig. 66 (A), the same battery used for Thévenin's Theorem is shown. The current source shown in the equivalent circuit equals the current flowing when the 9-ohm resistor is shorted. The arrow denotes the direction of current flow. It should be in the direction to produce the same polarity voltage across the terminals as with the original circuit. In the second example, the short circuit (Fig. 66 (B)) current may be calculated from the original circuit (Fig. 55 (B)), but assuming that Thévenin's Theorem is correct, we may short the terminals of the Thévenin's equivalent circuit (Fig. 55 (B)) and obtain: I = 9 volts/20.9 ohms = 0.43 ampere. In both the Norton and Thévenin equivalent circuits, the open-circuit voltage at the terminals is 9 volts. If any external load is connected to these terminals, the resulting voltage—current flow, power dissipation, etc.—will be the same for both equivalent circuits, as well as the original circuit.

To sum up the construction for the Thévenin's and Norton's equivalent circuits, assume that an unknown active network within a black box is given. Assume that the open circuit terminal voltage and short circuit current are measurable, and denoted as V_o and I_s

---

[2] A perfect current source is a device that has a constant current flowing through it at all times, regardless of the external circuit conditions. If the load is open-circuited across a current source, an infinite voltage exists. This theoretical representation is only partially realized in practice with a high voltage source in series with a very large resistance. (See Chapter 4 for a more detailed discussion.) A constant current source is considered to have an infinite resistance, and a voltage determined by the I x R of the external circuit.

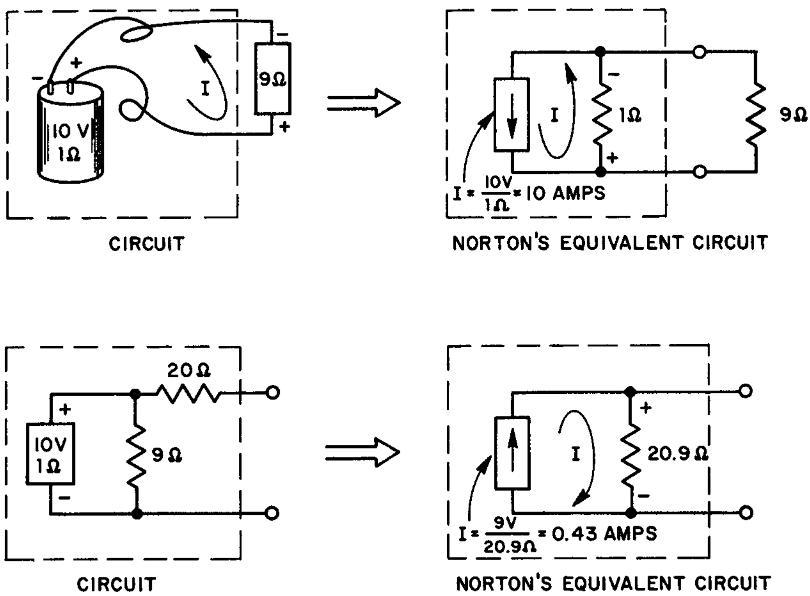


Fig. 66. Norton's equivalent of active circuits: (A) simple; (B) complex.

respectively. Then, the Thévenin's equivalent is a voltage source equal to V_0 volts in series with a resistance equal to V_0/I_s. The Norton's equivalent is a current source equal to I_s in parallel with a resistance equal to V_0I_s. This relationship is shown in Fig. 67.

Kirchhoff's voltage law is again needed: The algebraic sum of the voltage drops around any closed loop is zero. Figure 68 gives some examples of circuits and their related voltage law equations.

These theorems are absolutely essential in representing any active electronic circuit, for in many cases, the internal complexity of the device is so great (*i.e.,* as in a transistor) that specifying a precise circuit is impossible. What is possible, however, is to construct an electronic equivalent circuit that will duplicate the device's action at its output terminals.

A transistor has two input and two output terminals. It may be represented by a black box having an input and output set of terminals. This is the standard representation of a two *terminal-pair network,* and is shown in Fig. 69. *All* two terminal-pair networks may be analyzed, knowing only 4 quantities—the voltage across the input and output terminals, and the currents flowing in or out of these terminals. The complexity of the internal cir-

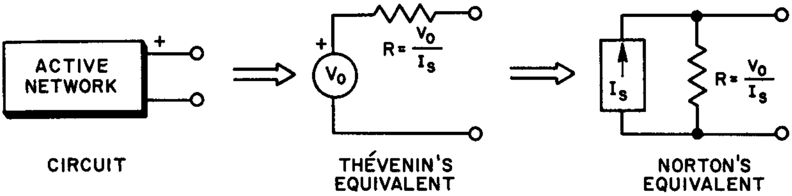


Fig. 67. Norton's vs Thévenin's equivalent of an active circuit.

cuitry of the box is only evident when the voltage and current relationships at its terminals are known. Once these relationships are known, it is possible to construct an equivalent circuit that will duplicate these terminal volt-ampere relationships.

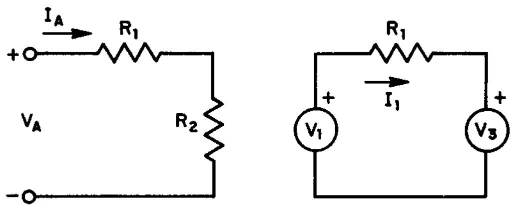


Fig. 68. Applications of Kirchhoff's Law.

$-V_A + V_{R1} + V_{R2} = 0$

$V_A = I_A R_1 + I_A R_2$

$V_A = I_A (R_1 + R_2)$

$-V_1 + V_{R1} + V_3 = 0$

$V_1 = V_3 + I_1 R_1$

CONVENTION: TREAT VOLTAGE RISES (FROM − TO +) AS NEGATIVE, AND VOLTAGE FALLS (+ TO −) AS POSITIVE.

It is known that a transistor has a current amplification, and an input and output impedance. It is also known that putting a voltage across the output terminals of a transistor will cause a voltage to exist across the input terminals. This latter fact was demonstrated in the introduction to the transistor, where it was shown that a transistor is not strictly a one-way device—there exists some reverse current amplification. In fact, bilateral, or sym-

Fig. 69. 2 terminal-pair network.

metrical, transistors are available where the current gain is the same in both directions. This is mentioned to point out that (unlike vacuum tubes) the input terminals are not completely isolated from the output terminals. This reverse current amplification will create a voltage drop across the input terminals when a voltage is placed across the output terminals. This effect of output voltage on input voltage is called the *reverse voltage amplification,* or just *voltage feedback ratio.*

39. Hybrid Parameters

All of the above parameters manifest themselves in the voltage-current relationship of the transistor, when represented as a two terminal-pair network. Many mathematical equations may be written expressing this relationship. Since any equations written should be capable of laboratory verification, only a few have reached any degree of popularity, it being rather difficult to perform certain tests on a transistor, such as open-circuiting the output terminals. A set of equations that has become universally accepted combines the best features of both open-and short-circuit tests of a transistor's input and output terminals. These equations are called *hybrid equations,* for they are formed by combining the best feaetures of two other sets of equations. The defining hybrid equations are:

$$v_1 = h_{11}i_1 + h_{12}v_2 \tag{71}$$

$$i_2 = h_{21}i_1 + h_{22}v_2 \tag{72}$$

Equations (71) and (72) are written in terms of the input and output terminal voltage-current relationships. The unknown quantities, the *h's* are called *hybrid parameters.*

Solving Equation (71) for h_{11}, and letting $v_2 = 0$:

$$h_{11} = \left.\frac{v_1}{i_1}\right|_{v_2 = 0} \tag{73}$$

= the input resistance with the output terminals shorted for ac $(v_2 = 0)$.

Solving Equation (71) for h_{12} with $i_1 = 0$:

$$h_{12} = \left.\frac{v_1}{v_2}\right|_{i_1 = 0} \tag{74}$$

= ratio of the output voltage that is fed back to the input, with the input terminals open circuited $(i_1 = 0)$

Solving Equation (72) for h_{21} and letting $v_2 = 0$:

$$h_{21} = \left.\frac{i_2}{i_1}\right|_{v_2 = 0} \quad = \text{current gain, } \alpha\text{, with output terminals shorted for ac } (v_2 = 0) \tag{75}$$

Solving Equation (72) for h_{22} and letting $i_1 = 0$:

$$h_{22} = \left.\frac{i_2}{v_2}\right|_{i_1 = 0} \quad = \text{output conductance with input terminals open circuited for ac } (i_1 = 0) \tag{76}$$

These four hybrid parameters define the input and output impedance, and the current or voltage gain (under short- or open-circuit conditions) in both directions. It is important to reiterate that these parameters refer only to small signal *ac values,* as evidenced by the use of *small* letters. Also the terms "open and shorted" apply only to the ac values, and it is assumed that the dc conditions necessary to establish the correct bias conditions are present.

It has become standard practice to use the terminology:

11 = i = input

12 = r = reverse transfer

21 = f = forward transfer

22 = o = output

In addition, to differentiate between the three configurations:

b = common base

e = common emitter

c = common collector

Adding these two subscripts to each *h* parameter, instead of the numbers, the *h* parameters for the three configurations may be written as shown in Table 5-1.

Small ac, or instantaneous, signals are denoted by small case letters. They are equivalent to small changes in the quiescent value, which are denoted by the symbol Δ. Hence the *h* parameters for, say, the common emitter configuration may be defined as:

$$h_{ie} = \left.\frac{v_{be}}{i_b}\right|_{v_{ce} = 0} = \left.\frac{\Delta V_{BE}}{\Delta I_B}\right|_{V_{CE} = \text{Constant}} = \text{input impedance} \tag{77}$$

$$h_{re} = \frac{v_{be}}{v_{ce}}\Bigg|_{i_c = 0} = \frac{\Delta V_{BE}}{\Delta V_{CE}}\Bigg|_{I_B = \text{Constant}} = \text{voltage feedback ratio or reverse open circuit voltage gain} \quad (78)$$

$$h_{fe} = \frac{i_c}{i_b}\Bigg|_{v_{ce} = 0} = \frac{\Delta I_C}{\Delta I_B}\Bigg|_{V_{CE} = \text{Constant}} = \text{Forward small signal short-circuit current gain, } \beta \quad (79)$$

$$h_{oe} = \frac{i_c}{v_{ce}}\Bigg|_{i_b = 0} = \frac{\Delta I_C}{\Delta V_{CE}}\Bigg|_{I_B = \text{Constant}} = \text{Output conductance} \quad (80)$$

The *h* parameters for the other two configurations may be defined in a similar way.

TABLE 2

h PARAMETER CONFIGURATIONS

Common Base	$h_{ib} = h_{11}$ = *input* resistance, common *base*
	$h_{rb} = h_{12}$ = *reverse* voltage gain, common *base*
	$h_{fb} = h_{21}$ = *forward* current gain, common *base*
	$h_{ob} = h_{22}$ = *output* conductance, common *base*
Common Emitter	$h_{ie} = h_{11}$ = *input* resistance, common *emitter*
	$h_{re} = h_{12}$ = *reverse* voltage gain, common *emitter*
	$h_{fe} = h_{21}$ = *forward* current gain, common *emitter*
	$h_{oe} = h_{22}$ = *output* conductance, common *emitter*
Common Collector	$h_{ic} = h_{11}$ = *input* resistance, common *collector*
	$h_{rc} = h_{12}$ = *reverse* voltage gain, common *collector*
	$h_{fc} = h_{21}$ = *forward* current gain, common *collector*
	$h_{oc} = h_{22}$ = *output* conductance, common *collector*

40. Graphical Determination of Hybrid Parameters

Defining the *h* parameters in terms of small changes of average value of voltages or currents, as in Equations (77) to (80), suggests that a graphical measurement of these quantities is possible—for the graphical static input and output characteristics should contain every combination of tests performable on the transistor, when treated as a two-terminal-pair network.

Figure 70 is a graph of the static input characteristics (V_{BE} *vs* I_B). The procedure for obtaining the input *h* parameters (h_{ie} and h_{re}) from this graph is essentially that of measuring the quantities in Equations (77) and (78), under the restriction that the appropriate quantity, or parameter, is held constant. In Fig. 70, the small-signal input resistance of the transistor is (as before) merely the slope of the V_{BE}–I_B curve. As was mentioned, the slope, or tilt, of this line ($V_{CE} = 7$ *volts*) is a function of the dependent on

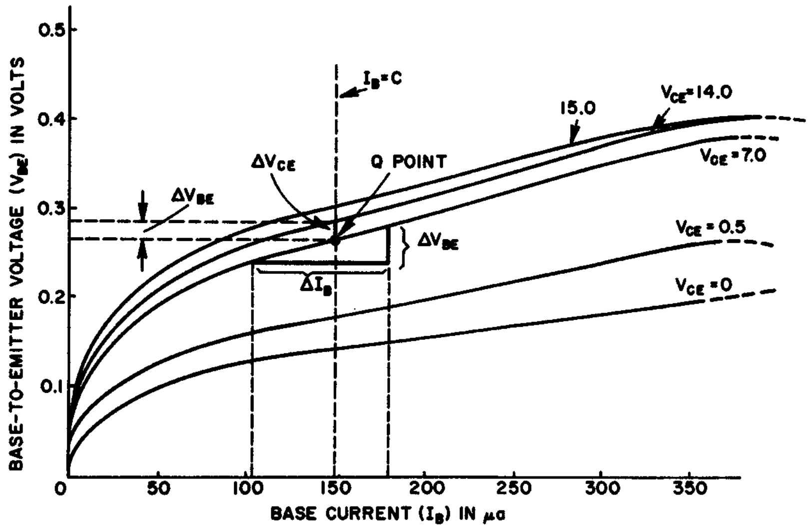


Fig. 70. Typical static input characteristics, for measurement of h_{ie} and h_{re}, of junction transistor.

the base current. Consequently, it is necessary to establish a dc operating point, before the slope is measured. For this example, a point of 150 μa base current has been picked. As before, the slope is measured by drawing a small triangle under the V_{CE} line, thus measuring the vertical and horizontal travel of the line. If the V_{CE} line is rather curved at the point where the triangle is drawn, a small triangle is necessary, for it is *assumed* that the hypotenuse of this triangle—formed by the V_{CE} line—is straight. Thus, the error in this approximation is less as the triangle becomes smaller. The compromise between drawing a small triangle and the ability to measure it is within the student's discretion.

$$h_{ie} = \frac{\Delta V_{BE}}{\Delta I_B}\Big|_{V_{CE} = 7 \text{ volts}} = \frac{(0.28\text{-}0.24) \text{ volts}}{(108\text{-}105)\ \mu a}$$

$$= \frac{0.04}{75 \times 10^{-6}} = 530 \text{ ohms} \tag{81}$$

Notice that the input resistance h_{ie} becomes greater (greater slope), as the base current decreases. Of course, under small-signal conditions, where the variation of base current during the signal swing is minute, this variation is not a problem. Only with this in mind, can any figure be given for the input impedance.

h_{re}, the reverse voltage amplification with the input terminals open circuited for ac, is measured on the same input characteristic graph. At the same value of V_{CE} (or as close as is possible), the ratio of a change in V_{CE} to an accompanying change in V_{BE} is measured at a constant I_B of 150 μa, corresponding to the same Q point. Here, the small change in V_{CE} is measured between the 14.0 and 7.0 V_{CE} lines on most input characteristic graphs vary widely in their spacing, tending to crowd together for high values of collector to emitter voltages. Therefore, adjacent V_{CE} lines, or lines as close together without incurring measurement difficulties, should be chosen.

$$h_{re} = \frac{\Delta V_{BE}}{\Delta V_{CE}}\Big|_{I_B = 100 \text{ ma}} = \frac{(0.288 - 0.266) \text{ volts}}{(14 - 7) \text{ volts}} = \frac{0.022}{7}$$

$$\cong 3.15 \times 10^{-3} \tag{82}$$

Notice that h_{re} is a dimensionless quantity. Due to the large variation of V_{CE} line spacing, h_{re} may vary considerably with the operating point. h_{re} may vary in a typical transistor from 6×10^{-2} to 1×10^{-3}.

To obtain the remaining two h parameters, we must consult the static output characteristics for the same transistor (Fig. 71). Carefully keeping the same operating point as in the graph of the input characteristics, we perform the same measurements, but with different parameters.

$$h_{fe} = \frac{\Delta I_C}{\Delta I_B}\Big|_{V_{CE} = 7.0 \text{ volts}} = \frac{(4.5\text{-}2.9) \text{ ma}}{(150\text{-}100)\ \mu a} = \frac{1.6 \times 10^{-3}}{50 \times 10^{-6}}$$

$$= \frac{160 \times 10^{+1}}{50} \cong 32 \tag{83}$$

Typical values for this small signal current gain (β) are 20 to over 100. It is important to stress that this is the small signal ac current gain, as opposed to the dc current gain. The two β's are denoted as h_{fe} and H_{FE}, respectively.

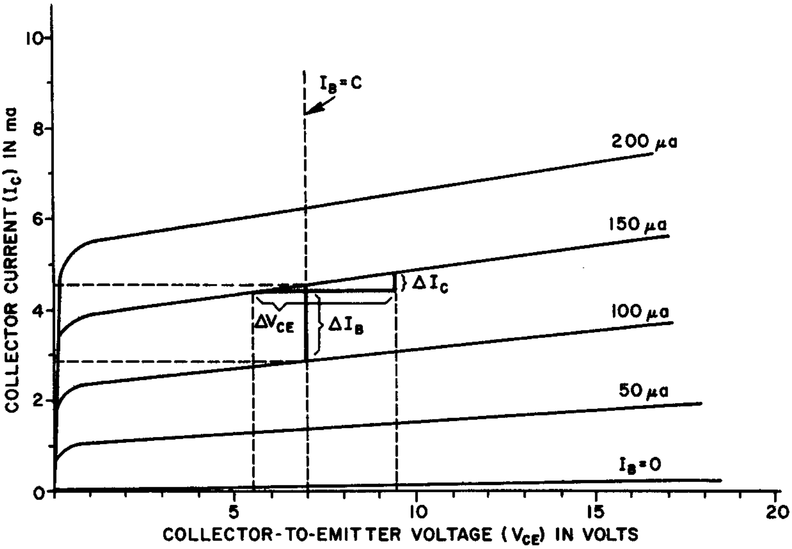


Fig. 71. Typical static output characteristics, for measurement of h_{fe} and h_{oe}, of junction transistor.

h_{oe}, the output conductance of the transistor with the input terminals open circuited for ac, is measured on the output characteristic graph in the same way as the input resistance (h_{ie}) is measured on the input characteristic graph. It is the slope of the base current line. Keeping the same dc operating point of 150 ma of base current and a collector to emitter voltage of 7 volts, we construct a small triangle under the base current line, to measure the horizontal and vertical change. Thus:

$$h_{oe} = \left.\frac{\Delta I_C}{\Delta V_{CE}}\right|_{I_B = 150\,\text{ma}} = \frac{(4.8 - 4.4)\ \text{ma}}{(9.5 - 5.5)\ \text{volts}} = \frac{0.4}{4.0} \times 10^{-3} = 1 \times 10^{-4}\,\text{mhos} = 100\,\mu\text{mhos} \tag{85}$$

This is actually a rather low output resistance for the common emitter configuration. Typical values of output resistances vary from 20,000 ohms to 100,000 ohms, corresponding to an output conductance, h_{oe}, of 10-50 μmhos.

This graphical determination of the small signal h parameters has been illustrated with the common emitter circuit. Since the output curves of this configuration are more tilted, determination of h_{fe} and h_{oe} is more accurate. h parameters obtained in this way are sufficiently accurate for most design work. The method of obtaining the common base h parameters from the input and output characteristics is carried on in exactly the same way.

41. h Parameters for the Common Base and Common Collector Circuits

Although the h parameters may be obtained graphically, often the input characteristics (sometimes the output characteristics also) are not given on the manufacturer's data sheet for the transistor. In such cases, the h parameters will often be given along with the conditions under which they were measured (*i.e.*, operating point, frequency, etc.). Today, with the advent of high-frequency transistors, several sets of h parameters may be given, one for each frequency. The low-frequency measurement is usually made at 1 kc.

Manufacturers usually give the h parameter values for one configuration, usually the common base or common emitter. If this is not the configuration of interest, a conversion must be made from the h parameter values of one configuration to ahother. Table 4 gives the hybrid parameter conversion formulae from one configuration to another.

h parameter values for a typical transistor in the three configurations are shown in Table 3. Notice that in the table, the forward short circuit current gain is *negative* for the common base and common collector. This will be explained in the subsequent section on h parameter equivalent circuits. Assuming that these values were originally measured with the common emitter circuit, the table may be checked with the aid of the conversion formulae of Table 3. Calculating the common base h parameters with the aid of this table, we have:

$$\text{Given: } h_{ie} = 2000 \text{ ohms}$$
$$h_{re} = 600 \times 10^{-6}$$
$$h_{fe} = 49$$
$$h_{oe} = 20\ \mu\text{mhos}$$

$$h_{ib} = \frac{h_{ie}}{1 + h_{fe}} = \frac{2 \times 10^3}{1 + 49} = \frac{2 \times 10^3}{5 \times 10^1} = 0.4 \times 10^2 \text{ ohms}$$

$$= 40 \text{ ohms} \qquad (85)$$

$$h_{rb} = \frac{h_{ie}h_{oe}}{1 + h_{fe}} - h_{re} = \frac{(2 \times 10^3)\ (20 \times 10^{-6})}{1 + 49} - 600 \times 10^{-6}$$

$$= \frac{(400 \times 10^{-4})}{50} - 6 \times 10^{-4}$$

$$= 8 \times 10^{-4} - 6 \times 10^{-4}$$

$$= 2 \times 10^{-4} \qquad (86)$$

$$h_{fb} = \frac{-h_{fe}}{1 + h_{fe}} = \frac{-49}{1+49} = \frac{-49}{50} = -0.98 \qquad (87)$$

$$h_{ob} = \frac{h_{oe}}{1 + h_{fe}} = \frac{20 \times 10^{-6}}{1+49} = \frac{20 \times 10^{-6}}{50} = 0.4 \times 10^{-6}\,\text{mhos}$$

$$= 0.4\,\mu\text{mhos} \qquad (88)$$

TABLE 4

h PARAMETER VALUES FOR A TYPICAL TRANSISTOR IN THE THREE CONFIGURATIONS

Common Emitter	*Common Base*	*Common Collector*
$h_{ie} = 2000$ ohms	$h_{ib} = 40$ ohms	$h_{ic} = 2000$ ohms
$h_{re} = 6 \times 10^{-4}$	$h_{rb} = 2 \times 10^{-4}$	$h_{rc} = 1$
$h_{fe} = 49 = \beta$	$h_{fb} = -0.98 = -\alpha$	$h_{fc} = -50$
$h_{oe} = 30\,\mu$mhos	$h_{ob} = 0.4\,\mu$mhos	$h_{oc} = 30\,\mu$mhos

Manufacturers have had a field day with *h* parameter notations. Although the notation convention used in this book is by far the most popular, a knowledge of the other notations used is necessary to quickly comprehend manufacturer's specifications. Table 5 equates the notations used in this book with others commonly used.

42. Variation of h Parameters with Operating Conditions

It is important to reiterate that a set of given or measured *h* parameters are only valid for one set of operating conditions. The extent to which they vary with different operating conditions, varies with the parameter and the actual transistor. Many manufacturers give graph and tables to indicate the extent of *h* parameter variation with operating conditions. Figure 72 is a page from a typical manufacturer's data sheet. It shows the variations of *h* parameters for temperature, emitter current, collector voltage, and junction temperature variations. Notice that the *h* parameter variations are *normalized* to the values found at a

TABLE 3

HYBRID PARAMETER CONVERSION FORMULAS

To convert from:

CB to CE

$$h_{ie} = \frac{h_{ib}}{1 + h_{fb}}$$

$$h_{re} = \frac{h_{ib}\, h_{ob}}{1 + h_{fe}} - h_{rb}$$

$$h_{fe} = \frac{-\, h_{fb}}{1 + h_{fb}}$$

$$h_{oe} = \frac{h_{ob}}{1 + h_{fb}}$$

CE to CB

$$h_{ib} = \frac{h_{ie}}{1 + h_{fe}}$$

$$h_{rb} = \frac{h_{ie}\, h_{oe}}{1 + h_{fe}} - h_{re}$$

$$h_{fb} = \frac{-\, h_{fe}}{1 + h_{fe}}$$

$$h_{ob} = \frac{h_{oe}}{1 + h_{fe}}$$

CB to CC

$$h_{ic} = \frac{h_{ib}}{1 + h_{fb}}$$

$$h_{rc} = \frac{1 - h_{ib}\, h_{ob}}{1 + h_{fb}}$$
$$\cong 1$$

$$h_{fc} = \frac{-\, 1}{1 + h_{fb}}$$

$$h_{oc} = \frac{h_{ob}}{1 + h_{fb}}$$

CE to CC

$$h_{ic} = h_{ie}$$

$$h_{rc} = 1 - h_{re}$$
$$\cong 1$$

$$h_{fc} = -\,(1 + h_{fe})$$

$$h_{oc} = h_{oe}$$

TABLE 3 (continued)

HYBRID PARAMETER CONVERSION FORMULAS

To convert from:

CC to CE	CC to CB
$h_{ie} = h_{ic}$	$h_{ib} = \frac{h_{oc}}{h_{fc}}$
$h_{re} = 1 - h_{rc}$	$h_{rb} = h_{rc} - \frac{h_{ic}\,h_{oc}}{h_{fc}} - 1$
$h_{fe} = -(1 + h_{fc})$	$h_{fb} = -\frac{h_{fc} + 1}{h_{fc}}$
$h_{oe} = h_{oc}$	$h_{ob} = -\frac{h_{oc}}{h_{fc}}$

standard operating condition. The technique of normalization involves dividing all the readings by the standard value (*i.e.,* the value at the standard operating conditions). To find the correct set of *h* parameters for a given transistor and its operating conditions, it is first necessary to find the published values for the standard conditions. Then, consult the graph (if any), to find the percentage variation at the new operating conditions.

43. h Parameter Equivalent Circuits

Now that we have defined a set of *h* parameters conforming to open- and short-circuit tests on the input and output terminals of our black box, we must *construct* an equivalent circuit that is capable of reproducing the results of these tests. This will be accomplished if the circuit is designed to conform to the input and output defining *h* parameter equations.

Earlier the common base circuit was usually used by manufacturers as a reference configuration, with all information given in terms of it. Today, the trend is towards using the common emitter circuit since it is the most often used configuration.

To construct the C.E. (common emitter) equivalent circuit, we start with the input defining *h* parameter equation.

TABLE 5

HYBRID NOTATIONS IN CURRENT USAGE

Common Emitter	$h_{ie} = h_{bb} = h_{11e} = 1/r_{11e}$
	$h_{re} = h_{bc} = h_{12e} = u_{bc} = u_{re}$
	$h_{fe} = h_{bc} = h_{21e} = \alpha_{cb}$
	$h_{oe} = h_{cb} = h_{22e} = 1/r_{22e}$
Common Base	$h_{ib} = h_{ee} = h_{11b} = 1/r_{11b}$
	$h_{rb} = h_{ec} = h_{12b} = u_{ec} = u_{rb}$
	$h_{fb} = h_{ce} = h_{21b} = \alpha_{ce}$
	$h_{ob} = h_{cc} = h_{22b} = 1/r_{22b}$
Common Collector	$h_{ic} = h_{bb} = h_{11c} = 1/r_{11c}$
	$h_{rc} = h_{be} = h_{12c} = u_{be} = u_{rc}$
	$h_{fc} = h_{eb} = h_{21c} = \alpha_{eb}$
	$h_{oc} = h_{ee} = h_{22c} = 1/r_{22c}$

$$v_1 = h_{ie}i_1 + h_{re}v_2$$

Where $v_1 = v_{be}$, $v_2 = v_{ce}$, and $i_1 = i_b$ (89)

This equation relates input voltage to the input current and output voltage. Certain conclusions may be drawn, when it is rewritten in more definitive form:

$$v_1 \text{ (input signal voltage)} = h_{ie} \text{ (ohms)} \times i_1 \text{ (input current)} + h_{re} \text{ (pure number)} \times v_2 \text{ (output voltage)}.$$

This equation must obviously conform to Kirchhoff's voltage law.[3] Since we know that v_1 is the signal input voltage (a voltage rise), the other side of the equation must be in a voltage fall. The first term on the right side of this equation is the voltage drop caused by i_1 flowing through the input resistance, h_{ie}. The second term is a voltage source which is dependent upon the output voltage. From this, an input circuit may be drawn, as shown in Fig. 73 (A).

3 See Chapter 4, for a detailed explanation of this law.

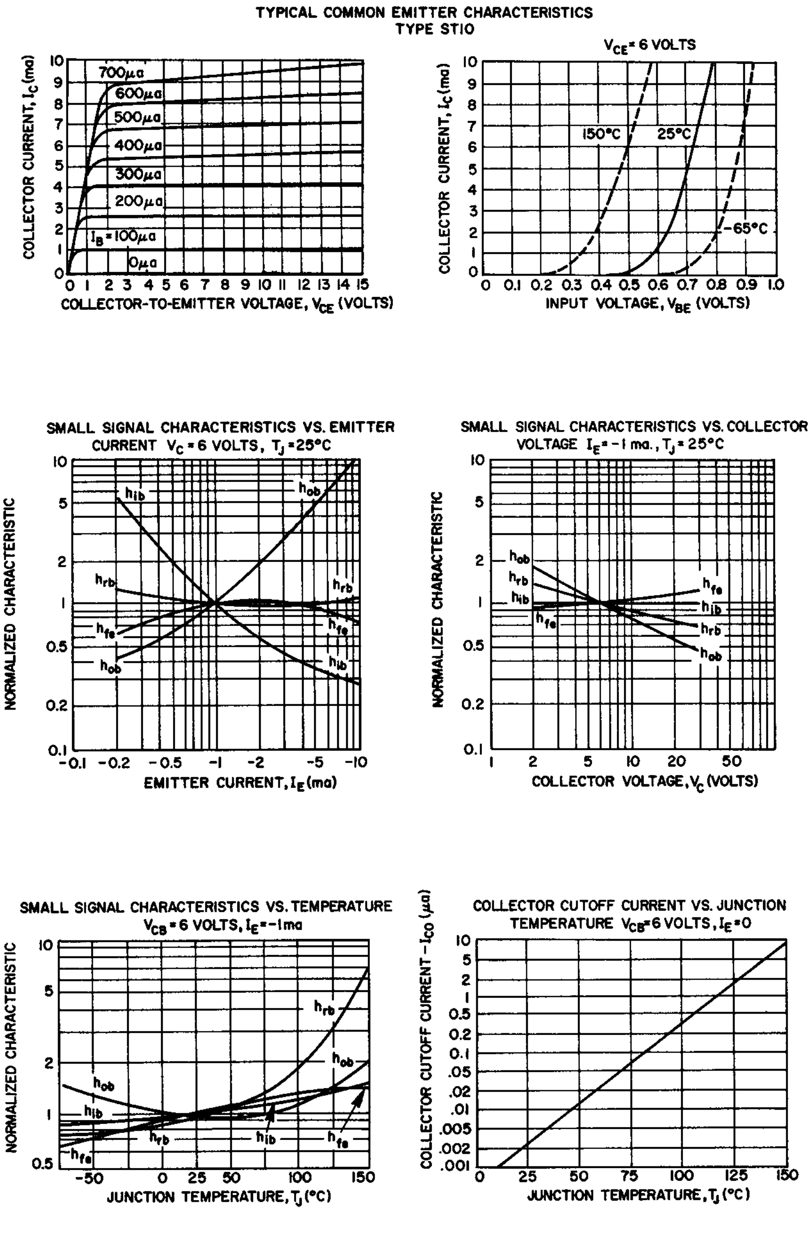


Fig. 72. Typical data sheet.

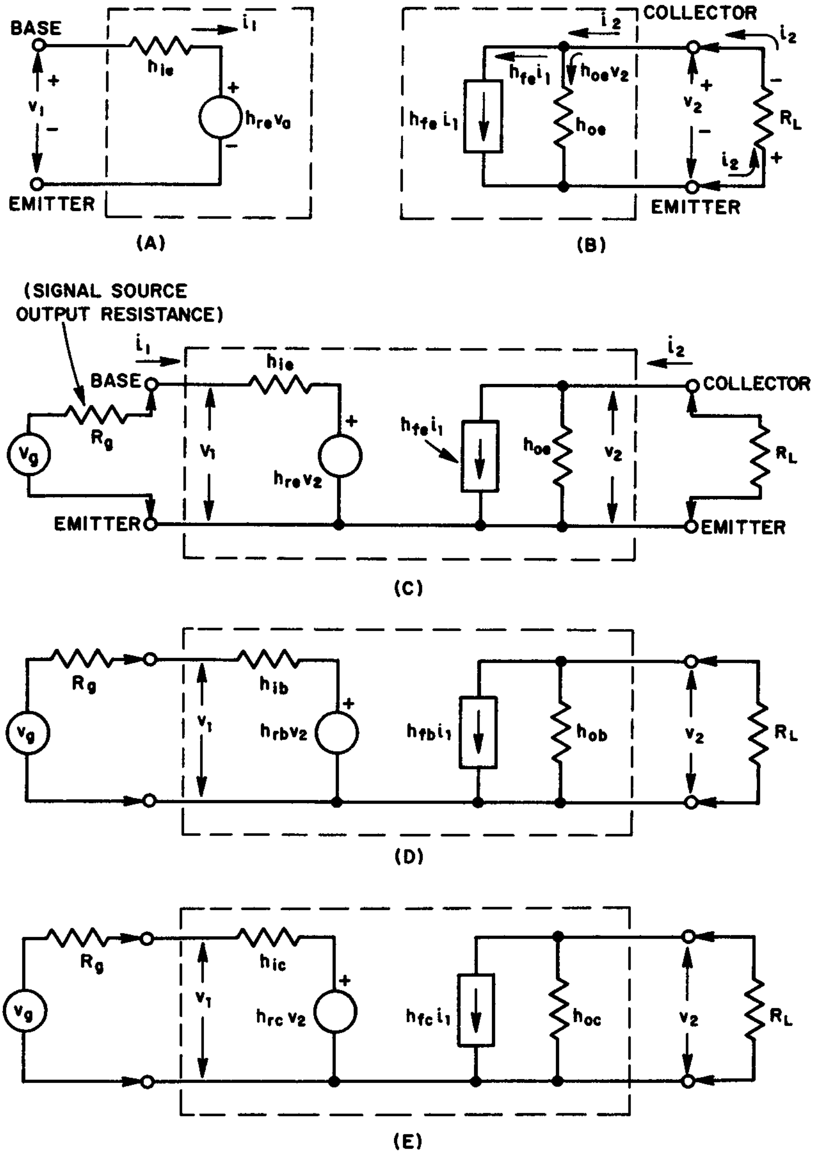


Fig. 73. (A) Common emitter input equivalent circuit; (B) Common emitter equivalent circuit; (C) Complete common h parameter emitter equivalent circuit; (D) Complete h parameter common base equivalent circuit; (E) Complete h parameter common collector equivalent circuit.

The dc bias circuitry has been omitted, as this is a small signal circuit only. If Kirchhoff's voltage law equation is applied to this input circuit, Equation (89) is evolved.

Notice that the feedback voltage source is of the same polarity as the signal input voltage and thus opposes the flow of signal current, i_1. This feedback voltage process is called *negative feedback,* since the output signal is fed back in opposition to the input voltage. This can be demonstrated by rewriting the input equation as:

$$v_1 - h_{re}v_2 = h_{ie}i_1 \qquad (90)$$

The *h* parameter equation for the output is:

$$i_2 = h_{fe}i_1 + h_{oe}v_2 \qquad (91)$$

This may also be written in a more explanatory form:

$$i_2 \text{ (output current)} = h_{fe} \text{ (pure number)} \times i_1 \text{ (input current)}$$
$$+ h_{oe} \text{ (output conductance)} \times v_2 \text{ (output voltage)}$$

The output equivalent circuit is shown in Fig. 73 (B). i_2 is the output (load) current. The first term on the right hand side, $h_{fe}i_1$, is the amount of current generated in the current source. It is dependent upon the input current and the forward short circuit current gain. The second term, $h_{oe}v_2$, is also a current (from Ohm's law). Then Equation (91) conforms to Kirchhoff's current law and says that i_2 splits at the node indicated, part going down h_{oe} and having a value of h_{oe2}, and part going through the current generator.

Putting the input and output equivalent circuits together, as in Fig. 73 (C), we now have the complete common emitter equivalent circuit. Notice that the output current generator "forces" current up from the bottom to the top of the load resistance. Then the load voltage (V_{RL}) is opposite in phase to the input voltage. This 180° phase shift is a property of the common emitter circuit, just as with the common cathode tube circuit. The C.E. is the only one of the three configurations that has this phase reversing property. This phase reversal can also be deduced from the fact that the actual load voltage is opposite in polarity to the assumed output voltage (v_2) polarity. The output current, i_2, which flows through the load, has the assumed direction. Consequently, there is no current phase reversal between the input and output.

One of the most useful aspects of the *h* parameter equivalent circuit is that each configuration has the same equivalent circuit, if the correct subscript is inserted for each parameter. In Figs.

73 (D) and (E), the complete equivalent circuit is shown for the common base and common collector circuits, respectively. The signal source and load resistance are shown with each circuit. The Thévenin's equivalent of the signal source is also shown.

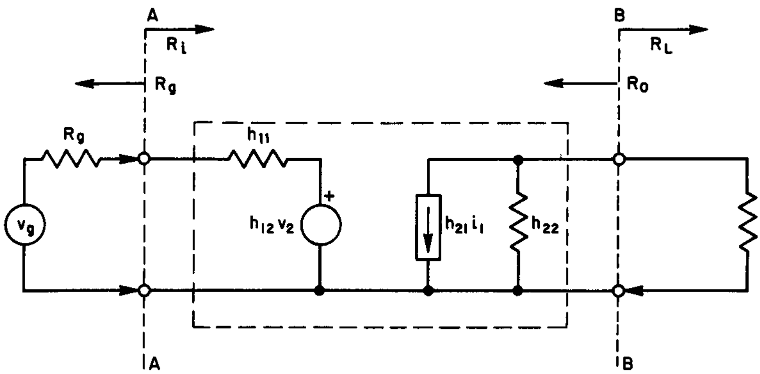


Fig. 74. A generalized amplifier equivalent circuit.

It is evident that in all three equivalent circuits, the output current generator is pointed in the same direction, even though there is no voltage phase reversal with the common base and common collector. To reconcile this inconsistency, the current generators for the common base and common collector circuits must be turned around. This is done quite effectively by defining h_{fb} and h_{fc} as being negative. Therefore, although the current source is drawn downward (to be consistent with the C.E. circuit), the current really flows upward, around, and down through the load resistance, in the *opposite* direction assumed for i_2. Consequently, the *current* undergoes a 180° phase shift, and there is no voltage phase reversal.

The h parameters are defined and measured under short- or open-circuit conditions, whereas in an operational amplifier stage, the input and output terminals are never short-circuited. The h_{11} parameter, for instance, does not represent the operating input resistance, R_i. The same may be said for the other h parameters. Consequently, when operating (in circuit) parameters are to be determined, they must be calculated from the equivalent circuit under *operating* conditions. Since, but for the proper subscript, all three h parameter equivalent circuits are identical, the formulae for R_i (input resistance), R_o (output resistance), A_v (voltage gain), A_i (current gain), and A_p (power gain) will be the same for all three configurations. (The derivation of these formulae by the simultaneous solution of several sets of equations will not materially add to the present study of transistors and, therefore, they are given here without proof.)

In Fig. 74, a generalized amplifier equivalent circuit is shown The *h* parameter formulae to follow are in reference to this generalized circuit. To obtain the specific formulas for the C.E., C.B., or C.C. configurations, insert the proper *h* parameter subscript in the formula, as indicated in Table 2.

44. Source Resistance

In the generalized amplifier of Fig. 74, the source output resistance (R_g) is the resistance "seen" when "looking" to the left of the line A-A. This will be the same whether the Thévenin's or Norton's equivalent of the signal source is shown.

45. Input Resistance

The input resistance of the amplifier (R_i) is the resistance "seen" by the source "looking" to the right of the line A-A *with the amplifier load connected.*

$$R_i = \frac{v_1}{i_1} = \frac{h_{11} + \Delta R_L}{1 + h_{22}R_L} \qquad (92)$$

Where $\Delta = h_{11}h_{22} - h_{12}h_{21}$

46. Output Resistance

The output resistance (R_o) of the amplifier stage is defined as the resistance seen by the load looking to the left of the line B-B *with the signal source connected.*

$$R_o = \frac{v_2}{i_2} = \frac{h_{11} + R_g}{h_{22}R_g + \Delta} \qquad (93)$$

47. Voltage Gain

The voltage gain (A_v) is the ratio of the output voltage to the input voltage, when the transistor is in the circuit with the source and load connected.

$$A_v = \frac{v_2}{v_1} = \frac{-h_{21}R_L}{h_{11} + \Delta R_L} \tag{94}$$

NOTE: A negative A_v denotes a voltage phase reversal.

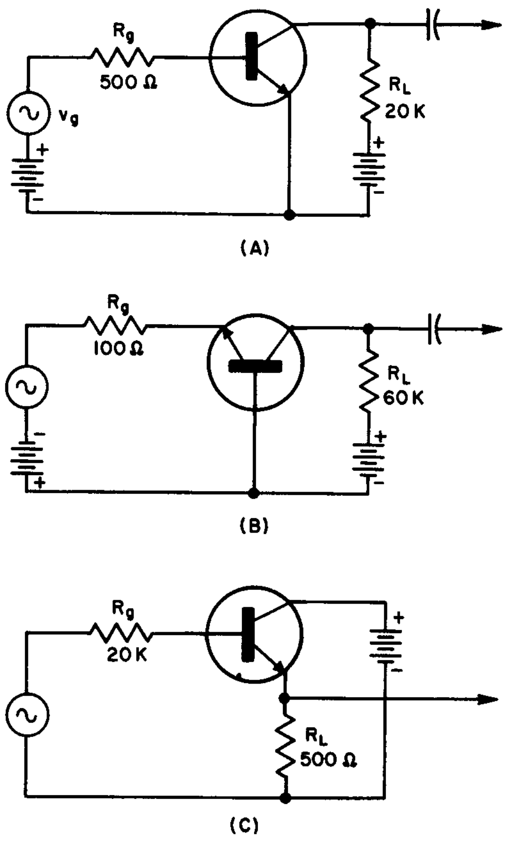


Fig. 75. (A) Common emitter circuit; (B) Common base circuit; (C) Common collector circuit.

48. Current Gain

The current gain (A_i) is the ratio of the output (load) current to the input current with the source and load connected.

$$A_i = \frac{i_2}{i_1} = \frac{+h_{21}}{1 + h_{22}R_L} \tag{95}$$

NOTE: A negative A_i denotes a current phase reversal.

49. Power Gain

The power gain (A_p) is the ratio of the power delivered to load, to the power delivered to the input of the amplifier. Since the input power, $P_i = v_1 i_1$, the output power, $P_o = v_2 i_2$. Then:

$$A_p = \frac{P_o}{P_i} = \frac{v_2 i_2}{v_1 i_1} = \frac{v_2}{v_1} \times \frac{i_2}{i_1} = |A_v|\,|A_i| \quad (96)$$

The vertical bars denote that A_i and A_v should be treated as + regardless of signs, so that A_p is always positive. A_p may be obtained directly by:

$$A_p = \frac{h^2_{21} R_L}{(1 + h_{22} R_L)\ (h_{11} + \Delta R_L)} \quad (97)$$

As an example of how these formulae are applied, let ùs calculate R_i, R_o, A_v, A_i, and A_p for a one-stage amplifier, in each of the three configurations. We will use the transistor whose h parameters are listed in Table 4. Again, bias voltages, etc. are not considered here, as it is assumed that they are correct. For the common emitter circuit of Fig. 75 (A)., we have:

$$h_{ie} = 2000 \text{ ohms}$$

$$h_{re} = 6 \times 10^{-4}$$

$$h_{fe} = 49$$

$$h_{oe} = 3 \times 10^{-5}$$

The manufacturer recommends $R_g = 500$ ohms, $R_L = 20{,}000$ ohms. Then:

$$\Delta = h_{ie} h_{oe} - h_{fe} h_{re}$$

$$= (2 \times 10^3)\ (3 \times 10^{-5}) - (6 \times 10^{-4})\ (49)$$

$$= 6 \times 10^{-2} - 2.94 \times 10^{-2}$$

$$= 3.06 \times 10^{-2}$$

$$R_i = \frac{h_{ie} + \Delta R_L}{1 + h_{oe} R_L}$$

$$= \frac{2 \times 10^3 + (3.06 \times 10^{-2})\ (2 \times 10^4)}{1 + (30 \times 10^{-6})\ (2 \times 10^4)}$$

$$= \frac{20 \times 10^2 + 6.12 \times 10^2}{1 + 60 \times 10^{-2}}$$

$$= \frac{26.12 \times 10^2}{1.6} = 16.3 \times 10^2$$

$$= 1630 \text{ ohms} \qquad (98)$$

$$R_o = \frac{h_{ie} + R_g}{h_{oe} R_g + \Delta}$$

$$= \frac{2 \times 10^3 + 5 \times 10^2}{(30 \times 10^{-6})\ (5 \times 10^2) + (3.06 \times 10^{-2})}$$

$$= \frac{2.5 \times 10^3}{1.5 \times 10^{-2} + 3.06 \times 10^{-2}}$$

$$= \frac{2.5 \times 10^3}{4.56 \times 10^{-2}} = 0.55 \times 10^5$$

$$= 55{,}000 \text{ ohms} \qquad (99)$$

$$A_v = \frac{-h_{fe} R_L}{h_{ie} + \Delta R_L}$$

(From Equation (98), $h_{ie} + \Delta R_L = 26.12 \times 10^2$)

$$= \frac{-(49)\ (2 \times 10^4)}{26.12 \times 10^2}$$

$$= -3.75 \times 10^2 = -375 \qquad (100)$$

$$A_i = \frac{+h_{fe}}{1 + h_{oe} R_L}$$

(from Equation (98), $1 + h_{oe} R_L = 1.6$)

$$= \frac{+49}{1.6} = +30.5 \qquad (101)$$

$$A_p = |A_v|\,|A_i| = (375)\ (30.5) = 11{,}400 \qquad (102)$$

Using the same procedure for the common base circuit of Fig. 75 (B), the manufacturer recommends $R_g = 100$ ohms, $R_L = 60{,}000$ ohms. The *h* parameters for this configuration from Table 4 are:

$$h_{ib} = 40 \text{ ohms}$$

$$h_{rb} = 2 \times 10^{-4}$$

$$h_{fb} = -0.98$$

$$h_{ob} = 0.4\ \mu\text{mhos}$$

$$\Delta = h_{ib}h_{ob} - h_{fb}h_{rb}$$

$$= (40)\ (0.4 \times 10^{-6}) - (-0.98)\ (2 \times 10^{-4})$$

$$= 16 \times 10^{-6} + 1.96 \times 10^{-4}$$

$$= 0.16 \times 10^{-4} + 1.96 \times 10^{-4}$$

$$= 2.12 \times 10^{-4}$$

Then: $$R_i = \frac{h_{ib} + \Delta R_L}{1 + h_{ob} R_L}$$

$$= \frac{40 + (2.12 \times 10^{-4})\ (6 \times 10^4)}{1 + (0.4 \times 10^{-6})\ (6 \times 10^4)}$$

$$= \frac{40 + 12.72}{1 + 2.4 \times 10^{-2}} = \frac{52.72}{1 + 0.024}$$

$$= \frac{52.72}{1.024} = 51 \text{ ohms} \tag{103}$$

$$R_o = \frac{h_{ib} + R_g}{h_{ob} R_g + \Delta}$$

$$= \frac{40 + 100}{(0.4 \times 10^{-6})\ (1 \times 10^2) + (2.12 \times 10^{-4})}$$

$$= \frac{140}{0.4 \times 10^{-4} + 2.12 \times 10^{-4}}$$

$$= \frac{140}{2.42 \times 10^{-4}} = 57.5 \times 10^4$$

$$= 575{,}000 \text{ ohms} \tag{104}$$

$$A_v = \frac{-h_{fb} R_L}{h_{ib} + \Delta R_L}$$

[from Equation (103), $h_{ib} + \Delta R_L = 52.72$]

$$= \frac{-(-0.98)\ (6 \times 10^4)}{52.72}$$

$$= \frac{+5.9 \times 10^4}{5.272 \times 10^1} = 1.14 \times 10^3 = 1{,}140 \tag{105}$$

$$A_i = \frac{h_{fb}}{1 + h_{ob} R_L}$$

[from Equation (103), $1 + h_{ob} R_L = 1.024$]

$$= \frac{-0.98}{1.024} = -0.95 \tag{106}$$

$$A_p = |A_v|\,|A_i| = (1140)\,(0.95) = 1080 \qquad (107)$$

For the CC circuit of Fig. 75 (C), using recommended values of R_g = 20,000 ohms and R_L = 500 ohms, we again calculate the operating parameters using the C.C. *h* parameters from Table 4.

$$h_{ic} = 2000 \text{ ohms}$$

$$h_{rc} = 1$$

$$h_{fc} = -50$$

$$h_{oc} = 30\,\mu\text{mhos}$$

$$\begin{aligned}\Delta &= h_{ic}h_{oc} - h_{fc}h_{rc}\\ &= (2 \times 10^3)\,(30 \times 10^{-6}) - (-50)\,(1)\\ &= 60 \times 10^{-3} + 50\\ &= 0.06 + 50 = 50.06 \cong 50.1\end{aligned}$$

$$\begin{aligned}R_i &= \frac{h_{ic} + \Delta R_L}{1 + h_{oc} R_L}\\ &= \frac{2 \times 10^3 + (50.1)\,(5 \times 10^2)}{1 + (30 \times 10^{-6})\,(5 \times 10^2)}\\ &= \frac{2 \times 10^3 + 250.5 \times 10^2}{1 + 150 \times 10^{-4}}\\ &= \frac{2 \times 10^3 + 25.05 \times 10^3}{1 + 0.015}\\ &= \frac{25.25 \times 10^3}{1.015} \cong 25.25\text{K ohm} \qquad (108)\end{aligned}$$

$$\begin{aligned}R_o &= \frac{h_{ic} + R_g}{h_{oc} R_g + \Delta} = \frac{2 \times 10^3 + 20 \times 10^3}{(30 \times 10^{-6})\,(20 \times 10^3) + 50.1}\\ &= \frac{22 \times 10^3}{600 \times 10^{-3} + 50.1} = \frac{22 \times 10^3}{0.6 + 50.1}\\ &= \frac{22 \times 10^3}{50.7} = 0.435 \times 10^3 = 435 \text{ ohm} \qquad (109)\end{aligned}$$

$$A_v = \frac{-h_{fc} R_L}{h_{ic} + \Delta R_L}$$

[from Equation (108), $h_{ic} + \Delta R_L = 25.25 \times 10^3$]

$$= \frac{-(-50)\,(500)}{25.25 \times 10^3} = \frac{+(5 \times 10^1)\,(5 \times 10^2)}{25.25 \times 10^3}$$

$$= \frac{25 \times 10^3}{25.25 \times 10^3} = 0.98 \tag{110}$$

$$A_i = \frac{+h_{fc}}{1 + h_{oc} R_L}$$

[from Equation (108), $1 + h_{oc} R_L = 1.015$]

$$= \frac{-50}{1.015} = -49.2 \tag{111}$$

$$A_p = |A_v|\,|A_i| = (0.98)\,(49.2)$$

$$= 48.2 \tag{112}$$

The results of these calculations are given in Table 6. From this table, certain generalized comments may be made with regard to the three configurations.

The common emitter circuit has a high (−) voltage and a high (+) current gain. It is the only configuration with a voltage *and* current gain greater than unity. It has the highest power gain of the three circuits. It is the only configuration that can achieve a power gain in a multistage r-c coupled amplifier. For these reasons, together with the fact that only one bias battery is required, it is the most popular configuration.

The common base circuit has a (+) voltage and (−) current gain. It has the highest voltage gain of the three circuits, while having a current gain less than unity. The common base circuit has the lowest input and highest output resistance of the three configurations. Hence, it is often used to match a low to high impedance.

The common collector circuit (emitter follower circuit) has a (+) voltage and (−) current gain. It has the highest current gain of the three circuits, while having a voltage gain less than unity. It has the highest input and lowest output resistance of the three configurations. Therefore, it is often used to match a high to low impedance.

50. Isolation Between Input and Output

Examination of Equations (92) and (93) shows that the input resistance of the transistor is affected by the value of the load resistance, and the output resistance is affected by the value of the source resistance. These phenomena have been mentioned previously, but now we have a quantitative expression dealing with it. It is very important, for unlike a vacuum tube, the input and output are not isolated. The load must be specified before the

TABLE 6

TABULATION OF RESULTS OF EQUATIONS (98) TO (112)

Item	*Common Emitter*	*Common Base*	*Common Collector*
R_i	1630 ohms	51 ohms	25,250 ohms
R_o	55,000 ohms	575,000 ohms	435 ohms
A_v	−375	+ 1,140	+ 0.98
A_i	+ 30	− 0.95	− 49.2
A_p	11,400	1080	48.2

TABLE 7

INPUT AND OUTPUT RESISTANCE OF TYPICAL TRANSISTOR WHEN SOURCE AND LOAD RESISTANCE EQUAL ZERO AND INFINITY OHMS.

Quantity & Condition	$\dfrac{R_i}{R_L = 0}$	$\dfrac{R_i}{R_L = \infty}$	$\dfrac{R_o}{R_g = 0}$	$\dfrac{R_o}{R_g = \infty}$
Reduced Formula	$= h_{11}$	$= \dfrac{\Delta}{h_{22}}$	$= \dfrac{h_{11}}{\Delta}$	$= \dfrac{1}{h_{22}}$
Common Emitter	2000 ohms	1000 ohms	65,000 ohms	33,000 ohms
Common Base	40 ohms	550 ohms	190,000 ohms	2.5 megs
Common Collector	2000 ohms	1.6 megs	40 ohms	33,000 ohms

input resistance is known, and the source resistance must be specified before the output resistance is known. Equations (92) and (93) have been solved for the condition that the source and load resistance is both zero and infinity. The results are given in

Table 7. This table may be used as a quick rule-of-thumb guide for evaluating the input and output resistance under the conditions mentioned. It is noted that the common collector configuration has the least isolation between input and output.

51. Review Questions

1. Define the essential differences between large and small signal analysis.
2. Why is it that h_{21} is $+\beta$ but $-\alpha$?
3. Verify the common collector h parameters listed in Table 4, using the formulae of Table 3.
4. Which transistor configuration would you use for the greatest voltage gain, the greatest current gain, the greatest power gain? Why?
5. Why is the common emitter configuration the most popular? Is it always the best circuit to use? Under what conditions would you use the other circuits?
6. With the common emitter h parameters developed in Equations (81) to (84), calculate the voltage, current, power gain, and the input and output resistance in the C.E. configuration shown in Fig. 75 (A). Use $R_g = 1000$ ohms, $R_L = 15{,}000$ ohms.
7. Convert the h parameters referred to in question 6 to common base h parameters. Calculate the quantities asked for in question 6 for an $R_g = 200$ ohms, $R_L = 50{,}000$ ohms.
8. Convert the h parameters referred to in question 6 to common collector h parameters. Calculate the quantities asked for in question 6 for an $R_G = 50{,}000$ ohms, $R_L = 100$ ohms.
9. Short circuit the output and calculate the input resistance for the configurations referred to in questions 6, 7, and 8. Use the abbreviated formulas of Table 7.
10. Is it possible to use the h parameter equivalent circuit at high frequencies? At other operating points than the parameters were measured at? How does one convert from the parameters of one operating point to the parameters of another operating point?