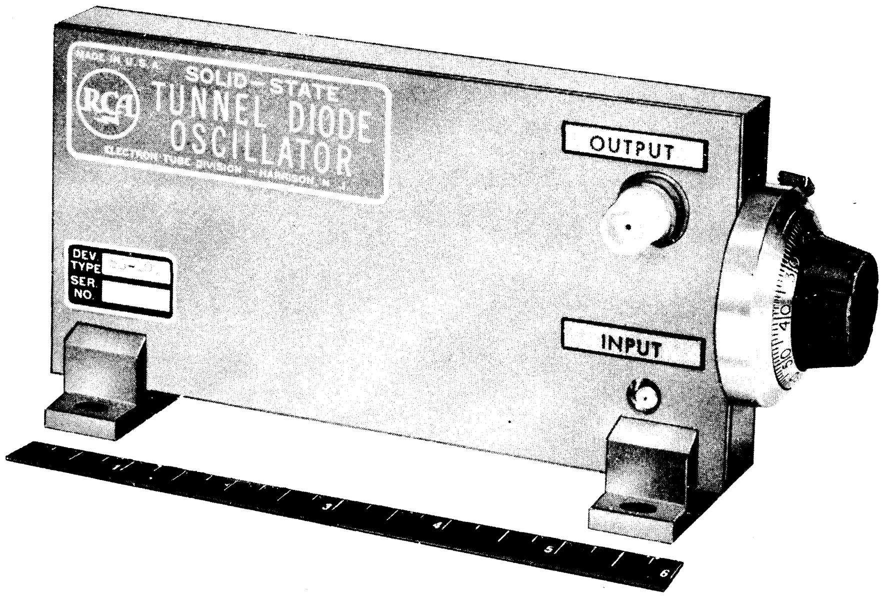


Fig. 1–1. Tunnel diode tunable (1050–1400 mc) oscillator (Courtesy Radio Corporation of America)

Chapter 1

INTRODUCTION

1–1. General

a. Fundamentally, the tunnel diode is a two-terminal semiconductor device that displays an ac (alternating current) *negative* resistance over a portion of its current-voltage curve. This phenomenon occurs if the tunnel diode is biased in the forward direction. It is this ac negative-resistance property of the tunnel diode that is exploited in a number of circuit arrangements to provide amplification, oscillation, switching, and memory functions extensively used in electronics.

b. Devices displaying ac negative resistance are not new to the field of electronics. This characteristic is found also in thyratrons, tetrode vacuum tubes, four-layer diodes, and others. However, because the tunnel diode is such a simple device, and a potentially inexpensive item, much engineering research has been directed toward increasing the number of applications. As a result, types of equipment using tunnel diodes are becoming more numerous and are being offered on the commercial market.

1–2. History of Semiconductor Devices

a. Crystal Rectifier. The first use of a crystal semiconductor as a rectifier (detector) was in the early days of radio. A crystal was clamped in a small cup or receptacle and a flexible wire (cat whisker) made light contact with the crystal. Tuning of the receiver was accomplished by operating the adjusting arm until the cat whisker was positioned on a spot of the crystal that resulted in a sound in the headset. Tuning the variable capacitor provided maximum signal in the headset.

b. Point-Contact Diode (Fig. 1–2). Point-contact diodes (germanium rectifiers) were used during World War II for radar and other high-frequency applications to replace electron-tube diodes. The point-contact diode has a very low shunt capacitance and does not require heater power; these properties provide a definite advantage over the electron-tube diode in high-frequency mixing and detecting applications. The point-contact

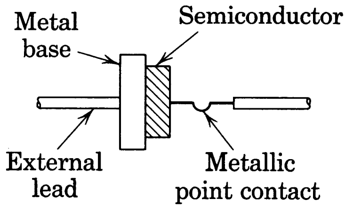


FIG. 1–2. Physical construction of point-contact diode

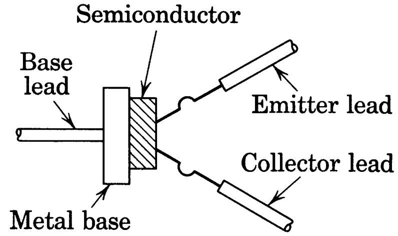


FIG. 1–3. Physical construction of point-contact transistor

diode is identical to the crystal rectifier (*a* above). The point-contact diode consists of a semiconductor, a metal base, and a metallic point contact. The connections to the point-contact diode are an external lead welded to the metallic point contact, and an external lead welded to the metal base.

c. Point-Contact Transistor (Fig.1–3). The development of the point-contact transistor was announced by the Bell Telephone Laboratories in 1948. The physicists credited with the invention were John Bardeen and Walter H. Brattain. The physical construction of the point-contact transistor is similar to that of the point-contact diode except that a third lead with a metallic point contact is placed near the other metallic point contact on the semiconductor. One lead is called an *emitter* lead; the other, a *collector* lead. When the two metallic points are properly biased with respect to the metal base, the point-contact transistor is capable of producing a power gain.

d. Junction Diode (Fig. 1–4). Rectification by semiconductor-to-semiconductor contact (a junction diode) was described in 1946 by W. H. Brattain of the Bell Telephone Laboratories. The rectifying junction diode is

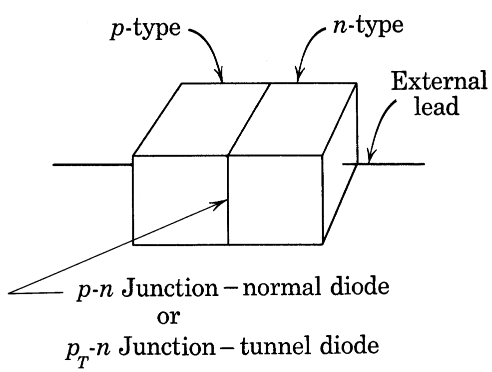
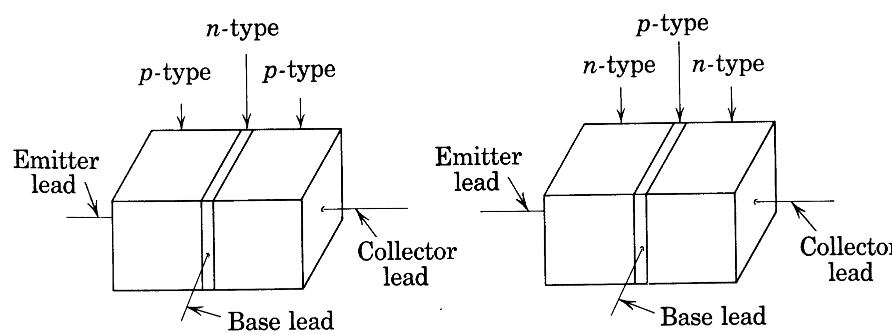


FIG. 1–4. Physical construction of rectifying junction diode, or tunnel diode

a single bar, slab, or wafer containing two dissimilar sections of semiconductor material. One section, because of its characteristics, is called a *p-type semiconductor;* the other, an *n-type* (par. 2–9 and 2–10). The connections to the rectifying junction diode consist of a lead to the *p*-type semiconductor and a lead to the *n*-type semiconductor. The rectifying junction diode can handle larger power than the point-contact diode but the rectifying junction diode has a larger shunt capacitance.

e. Junction Transistor. In 1949 W. Shockley of the Bell Telephone Laboratories published a theoretical analysis in which he predicted that a junction transistor consisting of area contacts (junctions) rather than point contacts, would be a practical device. Soon after, development of the junction transistor was announced. The junction transistor consists of two *p-n* junctions (Fig. 1–5A and B). Operation of the junction transistor is similar to that of the point-contact transistor. The junction transistor permits more accurate prediction of circuit performance, is less noisy, and is capable of handling more power than the point-contact transistor.

f. Tunnel Diode. The tunnel diode was invented in 1958 by a Japanese scientist, Leo Esaki. Thus, the tunnel diode is often referred to as the Esaki diode. The external appearance and structure of the tunnel diode is identical to the normal rectifying junction diode (Fig. 1–4). The difference between the two is in the amount of donor and acceptor impurities (Ch. 2) added to the semiconductor materials. The impurities added to the tunnel diode semiconductor materials are approximately 1000 times greater than those used in semiconductor materials for the rectifying junction diode. The junction formed in the rectifying diode is referred to as a *p-n* junction; in the tunnel diode the junction is referred to as a p_T-*n* junc-



A. *p-n-p* (semiconductor) junction transistor

B. *n-p-n* (semiconductor) junction transistor

FIG. 1–5. Physical construction of *p-n-p* and *n-p-n* junction transistors tion. It is the heavier doping (adding of impurities) of the tunnel diode semiconductor materials which causes the phenomenon of ac negative resistance (par. 3–7) that permits the use of the tunnel diode as an amplifier, oscillator, switch, etc.

1–3. Tunnel Diode Functions

a. Amplification. The tunnel diode may be used as a current, voltage, or power amplifier. For instance, a stronger signal current may be obtained

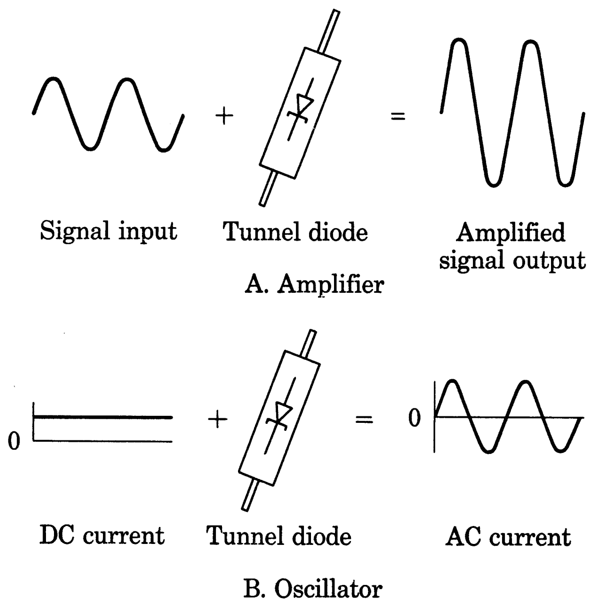


FIG. 1–6. Tunnel diode used as amplifier or oscillator

from a tunnel diode (Fig. 1–6A) than is delivered by the signal source. Various circuit arrangements provide for various amounts of signal amplification, depending on the power-handling capacity of the particular tunnel diode.

b. Oscillation. The tunnel diode may be used to convert direct-current energy into alternating-current energy; i.e., it may be used as an oscillator. When functioning in this manner, the tunnel diode shifts energy from a dc source and, in conjunction with a suitable timing (tank) circuit arrangement, generates an ac signal (Fig. 1–6B).

c. Modulation and Demodulation. The tunnel diode used in various circuit arrangements can provide amplitude modulation (variation in amplitude of an RF signal) (Fig. 1–7A) or frequency modulation (variation in frequency of an RF signal) (Fig. 1–8A). Demodulation (detection) of amplitude-modulated signals (Fig. 1–7B) or frequency-modulated signals (Fig. 1–8B) may be accomplished with tunnel diodes. These circuits are well-suited for miniature transmitters intended for short-range applications.

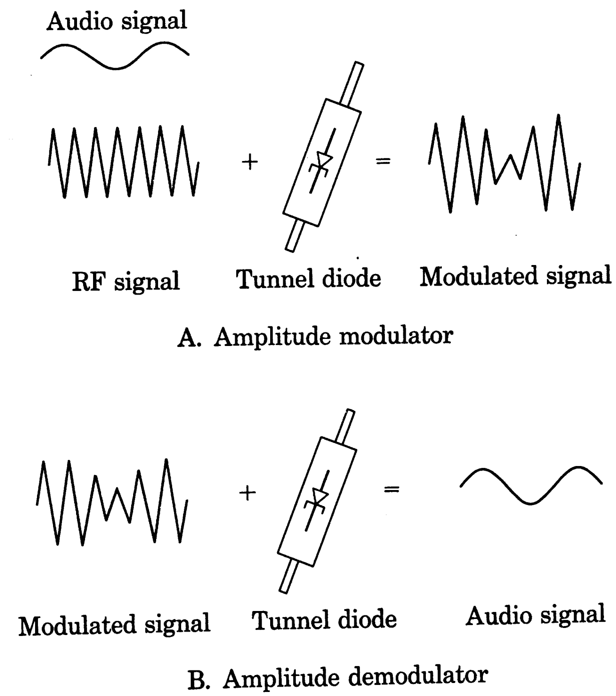


FIG. 1–7. Tunnel diode used as AM modulator or AM demodulator

d. Miscellaneous. The tunnel diode may also be used to modify the shape of signal waveforms. Waveform shaping is vital in various types of radar, teletypewriter, computer, and television circuits. Figure 1–9 indicates the use of the tunnel diode in transforming a sinewave into a square wave.

1–4. Comparison of Transistors and Tunnel Diodes

a. The transistor and the tunnel diode have common advantages over electron tubes. These are:

1. Greater power efficiency because no heater element is involved.
2. No warm-up time required.

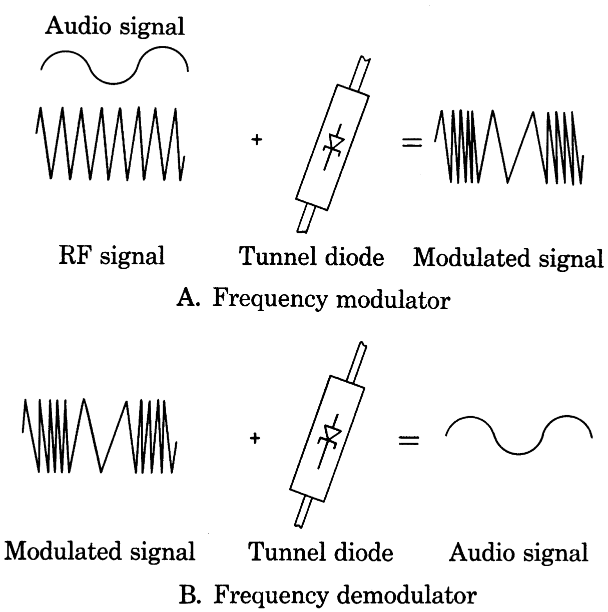


FIG. 1–8. Tunnel diode used as FM modulator or FM demodulator

3. Long life expectancy because of their physical ruggedness. Each will withstand centrifugal force, gravity, and impact tests that would completely shatter an electron tube.

4. Operation at low voltages that permit the use of smaller circuit components.

5. Extremely small in size compared to even miniature electron tubes. This permits ease of construction in microminiature modules especially desirable for use in portable equipment, aircraft, missiles, drones (unmanned aircraft), and satellites.

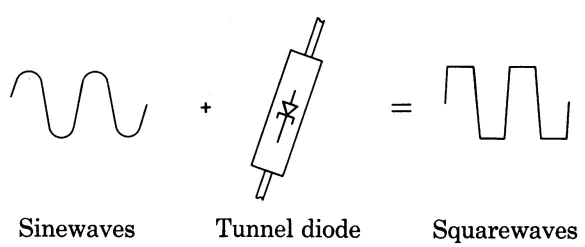


FIG. 1–9. Tunnel diode used to square waveform

b. The tunnel diode has the following advantages over transistors:

1. Much higher switching speeds. Present tunnel diodes can switch in one nanosecond (10^{-9} second or one billionth of a second). The tunneling action that occurs in a tunnel diode takes place theoretically at nearly the speed of light. Switching speed is of particular use in high-speed computers. Because of this fact it is expected that tunnel diodes will be extensively applied in this field. The transistor is limited in switching speed because of the long transit time of current carriers through its base.

2. The tunnel diode can operate through wide temperature ranges without appreciable changes in its tunneling characteristics. This property is due to the very heavy doping of the semiconductor material. Extremes of heat from absolute zero to 400° C (for gallium arsenide tunnel diodes) do not generate enough electron-hole pairs and excess electrons (par. 2–18) to increase markedly in percentage the original number of these carriers contributed by the doping materials. In the transistor, however, light doping is used and an operating temperature of 100° C materially affects its characteristics.

3. The tunnel diode can withstand relatively much larger doses of nuclear radiation compared to the transistor before its characteristics are changed. Again the reason is the heavy doping of semiconductor material (2 above).

4. Because the tunnel diode is a two-terminal device, there is no necessity to connect a third lead to an extremely small area as in the case of the transistor. This, plus the fact that only two regions of semiconductor material are required instead of three regions, as in the transistor, should result in a relatively inexpensive device.

5. Tunnel diode amplifiers can operate effectively even at microwave frequencies, producing wide bandwidth and high gain at low noise figures. Transistors have not competed with masers and parametric amplifiers at microwave frequencies. The high junction capacitances in the transistor and the long transit time of current carriers through the base region preclude use of the transistor at these frequencies. Tunnel diodes will be much lower in cost than masers and parametric amplifiers.

6. Heavy overload current of short duration will not permanently damage the tunnel diode. Actually, it is difficult to produce sufficient power in its junction to damage it.

7. Tunnel diode properties and characteristics are relatively unchanged by moisture and atmospheric gases. This fact eliminates the need for hermetic sealing that is required for transistors. The tunnel diode can be produced in very small, lightweight, encapsulated units requiring less space than transistors.

c. The main disadvantages of the tunnel diode are:

1. Gain for amplifiers and oscillators can be obtained only by operating over the ac negative-resistance characteristic of the device.
2. The tunnel diode has only two terminals that must be used for both input and output. Obviously the tunnel diode is not a unidirectional device. In simple circuit arrangements the output directly affects the input. Cascading amplifier stages becomes a major design difficulty. However, suitable unidirectional circuit arrangements can overcome this difficulty (Ch. 5).
3. The tunnel diode is a very low-voltage signal device. The available voltage swing of germanium units is less than ½ volt; gallium arsenide units provide less than 1 volt. Peak currents, however, can range from 10 μa to more than 10 amp.

1–5. Tunnel Diode Material

Materials, such as copper, silver, gold, and iron, which provide a good path for electron flow with little opposition (resistance), are referred to as *conductors* (par. 2–5*a*). Materials such as carbon in diamond form, germanium, and silicon, which provide a path for electron flow but offer moderate opposition, are referred to as semiconductors (par. 2–5*c*). Materials such as rubber, porcelain, and glass, which offer great opposition and do not provide a path for electron flow, are referred to as insulators (par. 2–5*b*). Tunnel diodes are composed of semiconductor materials such as germanium or silicon.

1–6. Summary

a. The ac negative-resistance characteristic of the tunnel diode may be employed in various circuit arrangements to perform functions normally performed by transistors and electron tubes.

b. The tunnel diode is made of a heavily doped semiconductor material, such as germanium or silicon.

c. The first crystal semiconductor was used as a rectifier (detector) in the early days of radio.

d. The point-contact transistor was invented in 1948.

e. The junction transistor was invented in 1949.

f. The tunnel diode was invented in 1958.

g. The tunnel diode may be used in circuits, such as amplifiers, oscillators, modulators, and demodulators.

h. The tunnel diode is smaller and more rugged than the electron tube. In addition, the power efficiency of the tunnel diode is greater than that of the electron tube.

i. Tunnel diodes are especially adaptable to the miniaturization of electronic equipment. Even the associated circuit components of tunnel diodes, such as capacitors, resistors, batteries, and transformers, may be made smaller than the corresponding components used in electron-tube circuits.

j. The advantages of the tunnel diode over the transistor are: faster switching; higher tolerance to high temperatures, nuclear radiation, and momentary current overloads; amplification and oscillation in the kilomegacycle region; high tolerance to moisture and atmospheric gases; lighter and less expensive to manufacture.

k. The main disadvantages of the tunnel diode are that it offers a small usable signal-voltage swing and it is not a unidirectional device.

Chapter 2

FUNDAMENTAL THEORY OF TUNNEL DIODES

SECTION I. STRUCTURE OF MATTER

2–1. General

a. A knowledge of the theory of the structure of matter is required for an understanding of the theory of the internal conduction mechanism that occurs in the tunnel diode.

b. Tunnel diodes are constructed from semiconductor materials. A comparison of the properties of conductors, semiconductors, and insulators is given in paragraph 2–5. Detailed properties and characteristics of semiconductor materials used in transistors and tunnel diodes are covered in paragraphs 2–6 through 2–12.

2–2. Matter, General

Matter is any substance that has weight (mass) and occupies space. Examples of matter are: air, water, plants, and metals. As these examples indicate, matter may be found in the gaseous, liquid, or solid state. Matter is found in nature as elements (*a* below), or compounds (*b* below). The elements and compounds are made up of molecules (*c* below), atoms (*d* below), and subatomic particles (*e* below).

a. Element. Matter consists of one or more basic materials which are called elements. Scientists have definite proof that at this time 102 elements exist. An element is defined as a substance that can be neither decomposed (broken up into a number of substances) by ordinary chemical changes nor made by chemical union of a number of substances. Copper, iron, aluminum, and gold are examples of metallic elements; oxygen, hydrogen, and sulfur are nonmetallic elements.

b. Compound. A substance containing more than one element and usually having properties different from those of its elements is called a *compound.* For example, water is made up of hydrogen and oxygen. Therefore, water is a compound.

c. Molecule. A *molecule* is defined as the smallest particle of matter that can exist by itself and still retain the properties of the original substance. If a drop of water, a compound, is divided until the smallest possible particle is obtained and is still water, that particle is known as a molecule. An idea of the size of a molecule may be obtained by imagining that a stone is first broken into two pieces, that the two pieces are then broken into four pieces, and that this process is continued. The smallest particle of stone which could be obtained by this process would be a molecule. Actually, it is impossible to crush a stone into its molecules; we can only crush it into dust. One small particle of dust is composed of thousands of molecules.

d. Atom. An atom is defined as the smallest part of an element that can take part in ordinary chemical changes. For simplicity, the atom may be considered to be the smallest particle that retains its identity as part of the element from which it is divided. Since there are approximately 102 known elements, there must be 102 different atoms, or a *different* atom for each element. All substances are made of one or more of these atoms.

e. Subatomic Particles. The atom itself can be subdivided into still smaller, or subatomic, particles. The nature of these subatomic particles is covered in paragraph 2–4.

2–3. Structure of Atom

Figure 2–1 shows the structure of matter. Parts A and B of the illustration are real; parts C through F are imaginary but are based on extensive laboratory and theoretical data.

a. Figure 2–1A shows the metal as it appears to the unaided eye.

b. In Fig. 2–1B the magnification is 100 diameters. Note that the metal has a crystalline structure. The crystals are small, nonuniform in shape, and irregularly arranged. This arrangement of crystals is referred to as a *polycrystalline structure.* All metals reveal a polycrystalline structure when seen through a microscope. The properties and characteristics of polycrystalline materials are quite different from the properties and characteristics of *single* crystal materials. Germanium and silicon, when processed for use in semiconductor devices, are *single crystal* materials (par. 2–7).

c. Magnification to 100,000 diameters (Fig. 2–1C) gives evidence of the presence of individual atoms or subatomic particles.

d. In Fig. 2–1D magnification is 10 million diameters. At this magnification the metal atoms appear as identical dots in straight rows.

e. At a magnification of 100 million diameters a single metal atom fills the entire area (Fig. 2–1E). This single atom resembles the solar system with a central body, called a nucleus, about which a number of smaller particles (electrons) move in outer orbits. Each of the electrons in this atom has a charge of electricity identical with the charge on any of the

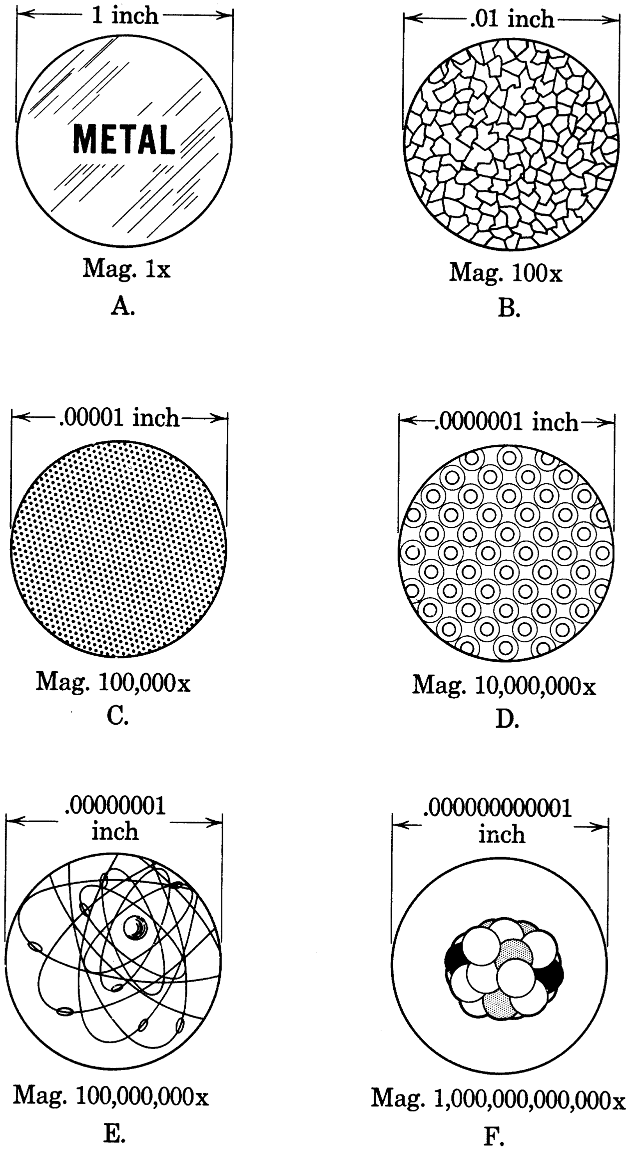


FIG. 2–1. Structure of a metal element

other electrons. The charge associated with an electron, the elemental charge, is the smallest electrical charge. The charge on the electron is arbitrarily designated a *negative charge*. For each negative electron that orbits about the nucleus, there is a positive proton in the nucleus, so that the atom is electrically neutral.

f. Increased magnification (Fig. 2–1F) shows that the enlarged nucleus contains two kinds of particles. The positively charged particles are called *protons*. The uncharged particles are called *neutrons*.

2–4. Electrons, Protons, and Neutrons

It has been established that the metal atom consists of a positively charged nucleus with negatively charged electrons that orbit around the nucleus.

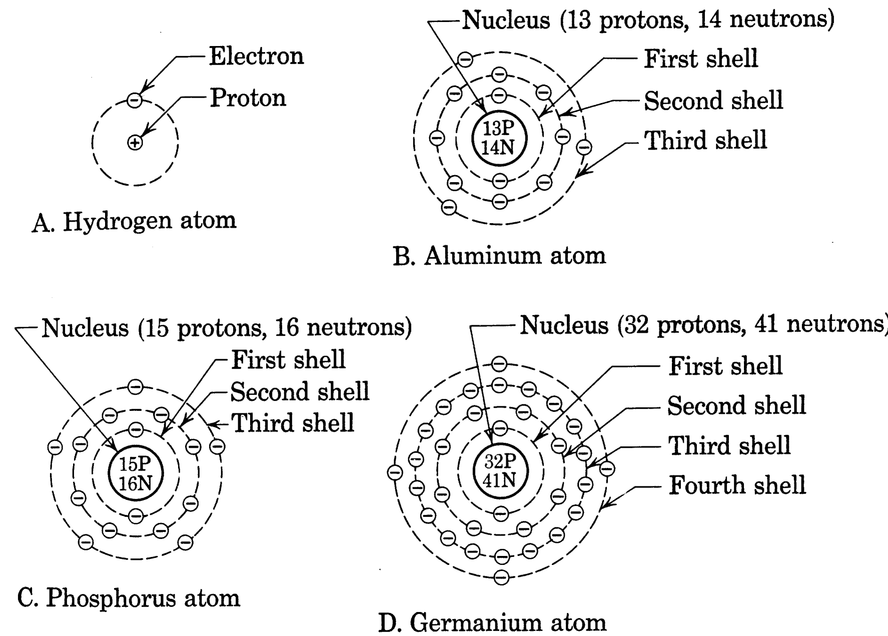


FIG. 2–2. Structure of atoms

Actually the atoms of *all* elements (oxygen, silver, hydrogen, etc.) contain a central nucleus and orbiting electrons.

a. Examples of Atomic Structure. 1. The atomic structure of the hydrogen atom (Fig. 2–2A) is the simplest of all atoms. It contains one electron orbiting around the nucleus which consists of one proton. The negative charge on the electron equals the positive charge on the proton and the atom is electrically neutral.

2. The nucleus of the aluminum atom (Fig. 2–2B) contains 14 neutrons and 13 protons. The positive charges of the 13 protons balance the negative charges of the 13 electrons and the entire atom is neutral. Note that the outermost shell has three electrons. The importance of these three electrons is explained in paragraph 2–10.

3. The phosphorus atom (Fig. 2–2C), a more complex structure, has 15 orbital electrons in three separate rings or *shells*. In this atom the outermost ring has five electrons. The importance of these five electrons is explained in paragraph 2–9.

4. The germanium atom (Fig. 2–2D), an even more complex atom, has 32 protons and 41 neutrons in the nucleus. The 32 orbital electrons revolve in four separate shells, with four electrons revolving in the outer incomplete shell. Paragraph 2–7 explains the importance of this arrangement.

b. Building Blocks and Their Characteristics. The difference in the various elements and their characteristics is in the *number* and *arrangement* of the electrons, protons, and neutrons of which each atom is composed. All electrons are basically identical regardless of the atom, and therefore the element, of which they are a part. The same can be said of all protons and all neutrons. Since matter is composed of atoms of positively charged protons, negatively charged electrons, and uncharged neutrons, these charged and uncharged particles are the fundamental building blocks of all matter. The charge on the negative electron or the positive proton is considered the elemental unit of electrical charge. Because the elemental unit is too small a quantity of electricity for practical purposes, a larger unit called the coulomb is commonly used. One coulomb of electricity represents 6.28 million, million, million (6.28×10^{18}) electrons. Although they have equal and opposite quantities of charge, a proton's mass is 1850 times greater than the mass of the electron. The neutron is equal in diameter and mass to the proton. Relatively great distances exist between the electrons and the protons of an atom. A copper one-cent piece magnified to the size of the earth's path around the sun (approximately 584,000,000 miles) would show electrons the size of baseballs spaced about 3 miles apart.

2–5. Conductors, Semiconductors, and Insulators (Fig. 2–3)

In the field of electricity all materials are placed in three main categories: conductors, semiconductors, or insulators. The category into which a material is placed depends on its ability to conduct electricity. This, in turn, depends on its atomic structure.

a. Conductors. A good conductor is a material that has a *large number of loosely held electrons.* All metals conduct electricity, but some are better conductors than others. A one-centimeter cube (each edge measures one centimeter) of silver, copper, or aluminum has a resistance of less than three millionths of an ohm. Silver is a better conductor than copper, but copper is used more extensively because it is less expensive. Aluminum is used as a conductor where weight is a major factor, for example, on long-span high-tension lines.

b. Insulators. A material classified as an insulator has an atomic structure that does not permit the movement of electrons from atom to atom. An insulator (or dielectric) is a material that has few loosely held electrons. Actually there are no perfect insulators. However, such materials as glass,

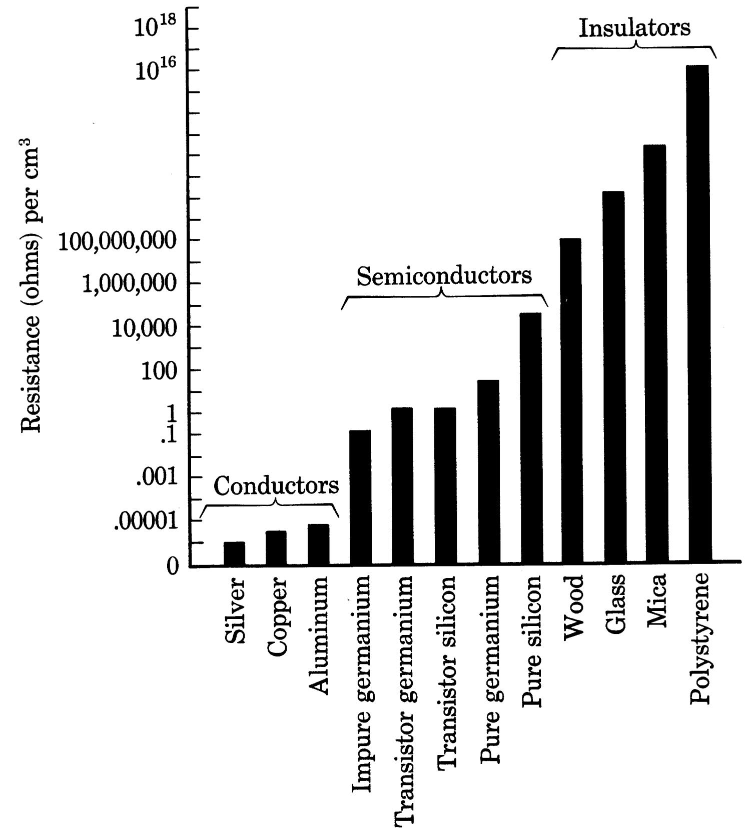


Fig. 2–3. Chart of resistance in ohms for a centimeter cube of conductors, semiconductors, and insulators

rubber, mica, and polystyrene have a resistance of several millions of ohms per centimeter cube. For most practical purposes these are considered nonconducting materials.

c. Semiconductors. Semiconductors are materials which are neither good conductors nor good insulators. Germanium and silicon fall into this category. At room temperature pure germanium has a resistance of 60 ohms per centimeter cube. Its resistance is several million times greater than that of copper and several million times less than mica. Pure silicon is a worse conductor, having a resistance of 60,000 ohms per centimeter cube at room temperature. Both impure germanium and impure silicon derived from related compounds have a resistance of 0.2 and 0.4 ohm per centimeter cube, depending on the type and amount of impurities present. Germanium and silicon, for use in transistors and rectifying diodes, have carefully controlled amounts of impurities added (par. 2–8 and 2–10), and each has a resistance of 2 ohms per centimeter cube at room temperature. This resistance decreases rapidly as the temperature rises. The resistance of semiconductor materials used for tunnel diodes is 1000 times less than the resistance of semiconductor materials used for transistors and rectifying diodes. Note that while only germanium and silicon as semiconductors are discussed, there are many other semiconductors.

SECTION II. IMPURITIES AND CURRENT CARRIERS IN CRYSTALS

2–6. Crystals, General

a. General. Most solids, except those exhibiting a biological structure of cells, such as leaves, branches, and bone, reveal a crystal structure when studied under a microscope. Many substances, such as rocks and metals which are not usually considered crystalline, reveal a specific crystal pattern when studied under a microscope.

b. External Characteristics. The most commonly known characteristics of crystals are their angles and their planes. Snow crystals, for instance, although formed in an infinite number of geometric patterns, contain only 60° angles. Some materials, such as common salt (sodium chloride), form cubes; other materials form long needles, rhomboids, or variations of hexagonal or rectangular structures. Each material has a characteristic form.

c. Internal Structure. X rays have been used to investigate the internal structure of crystals. The wavelengths of X rays approximate the distance between the atoms or molecules of crystals. When X rays are beamed through a crystal, the rays are deflected and distributed in accordance with the specific arrangement of the atoms or molecules of the crystal. When the resultant deflected rays are photographed, the photograph invariably shows a specific pattern depending upon the substance of the crystal. With the pattern indicated on the photograph, and through complex mathematical analyses, scientists have been able to construct models of the internal structure of a given crystal. These analyses have indicated that the atoms of crystals are arranged in specific patterns; that one atom is not closely related with another atom only, but rather is related equally to a number of adjacent, equidistant atoms. The specific arrangement of atoms depends on the size and number of atoms present and on the electrical forces between them. The physical, electrical, optical, and mechanical characteristics of the crystal depend on the forces between the atoms. The crystals discussed in paragraphs 2–7 through 2–8 are semiconductor crystals, such as germanium and silicon. These materials are used in making semiconductor devices, e.g., tunnel diodes and transistors. The most important common characteristics among these materials is that each of their atoms has four electrons in the outer shell. The discussion therefore applies to any of these semiconductor materials.

2–7. Pure Semiconductor Crystal

a. General. Figure 2–4 shows a pure semiconductor crystal. Each sphere represents a semiconductor atom less the four electrons (valence or outer-orbit electrons) that are in the outer (fourth) incomplete shell of the atom (Fig. 2–2D). The sphere (Fig. 2–4) contains the nucleus of the atom, and all the tightly bound electrons that orbit around the nucleus. Since the atom of any element is electrically neutral, the sphere has a net positive charge of four. Throughout this text, the sphere will be referred to as the *semiconductor core.*

b. Single Crystal Structure. The dashed lines in Fig. 2–4 form two cubes. Note that the four semiconductor cores between the two cubes are shared equally by the cubes. If the illustration were to be extended in all directions, the sharing of the corner semiconductor cores would be extended to all adjacent cubes. This repeated, uniform, *cubical structure* constitutes a *single semiconductor crystal.* The properties and characteristics of *single crystal* materials, such as germanium and silicon (as prepared for use in semiconductor devices), are quite different from the properties and characteristics of *polycrystalline materials,* e.g., copper and aluminum (par. 2–3*b*). The term crystal used throughout this text will refer to *single crystal material* only.

c. Lattice Structure. 1. It has been established that electrons rotate constantly in relatively fixed orbits about the nucleus. In a crystal, the rotation of one valence electron of a given atom is coordinated with the rotation of one valence electron of an adjacent atom. The coordinated rotation of two valence electrons (one from each of two adjacent atoms) results in the formation of an *electron-pair bond.* The electron-pair bonds shown diagrammatically in Fig. 2–4 are also referred to as *valence bonds.* The electron-pair bonds cause the cores to be attracted toward each other. The positive charges on the cores cause the cores to repel each other. When a balance of the forces of attraction and repulsion is obtained, the crystal is said to be in a state of equilibrium.

2. Each semiconductor core is equidistant from four adjacent semiconductor cores. Note that each core is interconnected with adjacent cores by four electron-pair bonds. This condition exists since each semiconductor

atom contains four valence electrons in its outer shell. This arrangement of semiconductor cores and electron-pair bonds is referred to as a *lattice.* To avoid congestion, Fig. 2–4 shows only a few of the electron-pair bonds.

d. Conductivity. The valence electrons of such good conductors as copper or aluminum are loosely bound to the nucleus of the atom, and they move quite readily through the conductor under the influence of an electric or magnetic field. Valence electrons which form part of an electron-pair bond,

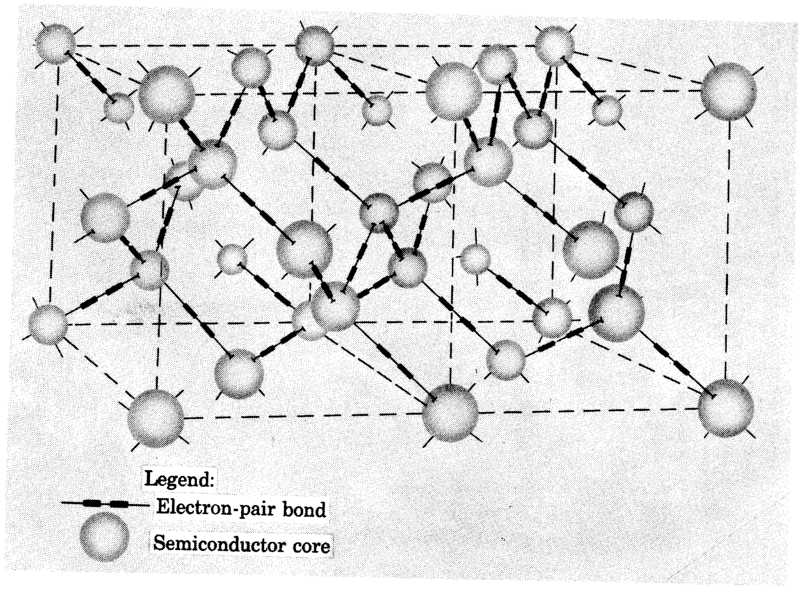


Fig. 2–4. Pure semiconductor crystal, lattice structure

however, are bound in the electron-pair bond and are not free to take part in conduction. Crystalline materials, such as germanium and silicon, the valence electrons of which are bound, are poor conductors (Fig. 2–3) under normal conditions. Only if the material is subjected to high temperatures or strong radiation will the electron-pair bonds separate and partial electrical conduction occur.

2–8. Impurities

a. General. It is possible for the atoms of substances having more or less than four valence electrons to join the crystal lattice structure of the semiconductor material. These substances whether found in the semiconductor material in its natural state, or added intentionally during the processing of the semiconductor material for use in tunnel diodes or transistors, are referred to as *impurities.*

b. Donor and Acceptor Impurities. Two groups of substances exhibit the important characteristic of joining the lattice structure of semiconductor materials. The substances in one group are called *donors;* in the second group, they are called *acceptors.*

1. The atoms of substances classified as donors have five valence electrons in the outer incomplete shell. Some of the substances that have been used as donors are phosphorus (Fig. 2–2C), arsenic, and antimony. The characteristics and properties of semiconductor materials containing donor atoms are discussed in paragraph 2–9.
2. The atoms of substances classified as acceptors have three valence electrons in the outer incomplete shell. Some of the substances that have been used as acceptors are aluminum (Fig. 2–2B), gallium, boron, and indium. The characteristics and properties of semiconductor materials containing acceptor atoms are covered in paragraph 2–10.

2–9. *n*-Type Semiconductor

a. Figure 2–5 shows a semiconductor crystal in which one of the semiconductor atoms has been replaced by a donor impurity (par. 2–8*b*1). The dark sphere in the illustration represents the nucleus of the donor atom and all the tightly bound electrons that orbit around the nucleus. The valence electrons are not included in the sphere. The donor impurity contains five valence electrons. Note that four of the valence electrons of the

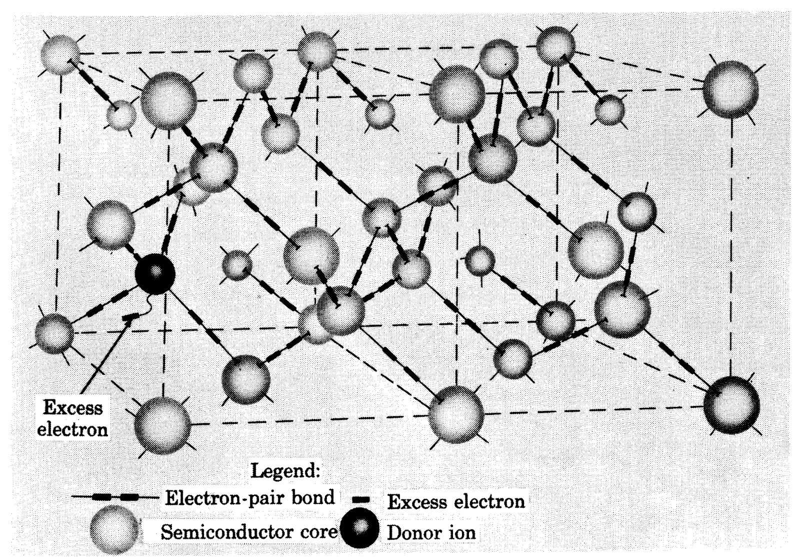


Fig. 2–5. Semiconductor crystal with donor atom

donor form electron-pair bonds with electrons of four neighboring semiconductor atoms. The electrons of both the semiconductor and the donor atoms that enter into electron-pair bonds form a very stable structure and are not readily removed from the bonds.

b. The fifth valence electron of the donor cannot form an electron-pair bond since there are no adjacent electrons available. This electron becomes an excess electron. The donor nucleus has a very weak influence over the

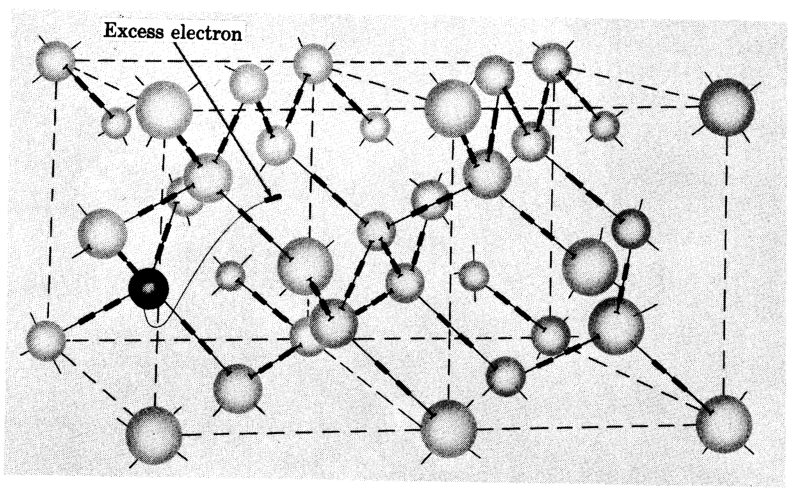


FIG. 2–6. Semiconductor crystal with donor atom showing movement of excess electron

excess electron. Actually, only one one-hundred-seventieth ($1/170$) of the energy required to remove an electron from an electron-pair bond is required to remove the excess electron from the donor. At normal room temperature (70° F), enough thermal (heat) energy is present to cause the excess electron to break away from the donor and wander through the space between the crystal lattices (Fig. 2–6).

c. When the excess electron leaves (ionizes from, or is donated by) the donor atom, the donor atom then possesses a positive charge equivalent to the negative charge of one electron. An atom that loses or gains an electron is called an *ion*. For that reason, the spheres that represent the donor (Figs. 2–5 and 2–6) are called donor ions.

d. Note that the semiconductor crystal that contains a donor ion (positive) also contains an excess electron (negative). The crystal taken as a whole therefore is electrically neutral; i.e., the crystal possesses a net charge of zero.

e. Semiconductors containing donor impurities are referred to as *n-type* semiconductors. The letter *n* refers to the *negative* charge of the excess electron.

2–10. *p*-Type Semiconductor

a. Figure 2–7 shows a semiconductor crystal in which one of the semiconductor atoms has been replaced by an acceptor impurity (par. 2–8*b*2).

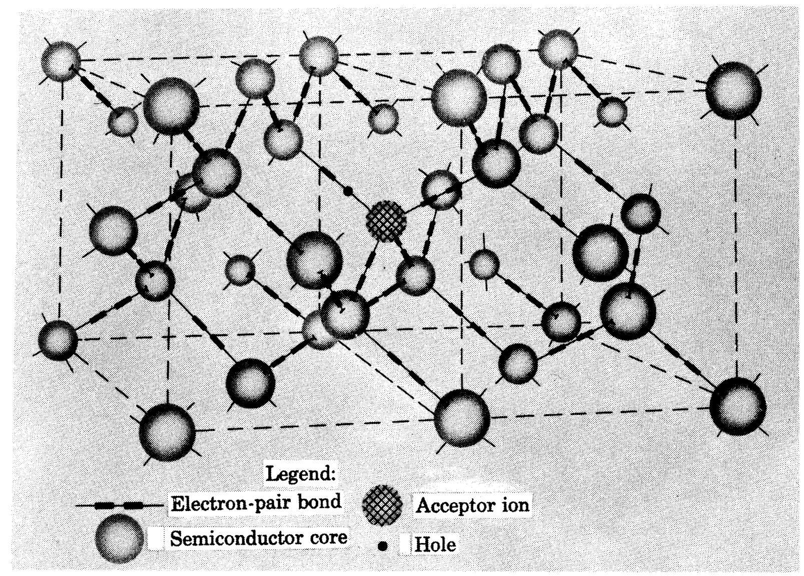


FIG. 2–7. Semiconductor crystal with acceptor atom

The cross-hatched sphere contains the nucleus of the acceptor atom and all the tightly bound electrons that orbit around the nucleus; the valence electrons are not included in the sphere. The acceptor impurity contains three valence electrons. Note that the three valence electrons of the acceptor form electron-pair bonds with electrons of the neighboring semiconductor atoms.

b. One valence electron of the fourth neighboring semiconductor atom cannot form an electron-pair bond because the acceptor has only three valence electrons. In this condition, an electron-hole arrangement exists. The position that would normally be filled with an electron is designated as a *hole*.

c. It is possible for an electron from an adjacent electron-pair bond to absorb enough thermal or electrical energy to break its bond (Figs. 2–8 and 2–9) and fill in the hole in the original electron-hole arrangement (Fig. 2–7).

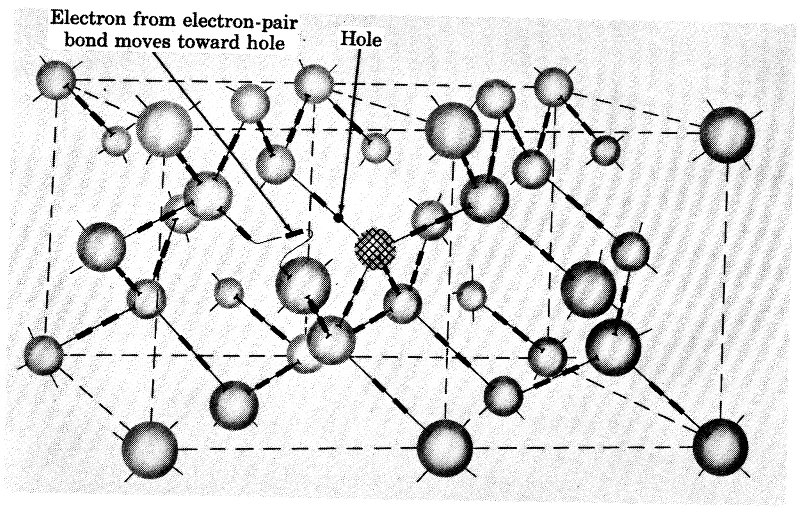


FIG. 2-8. Semiconductor crystal with acceptor atom, showing electron from electron-pair bond moving toward hole

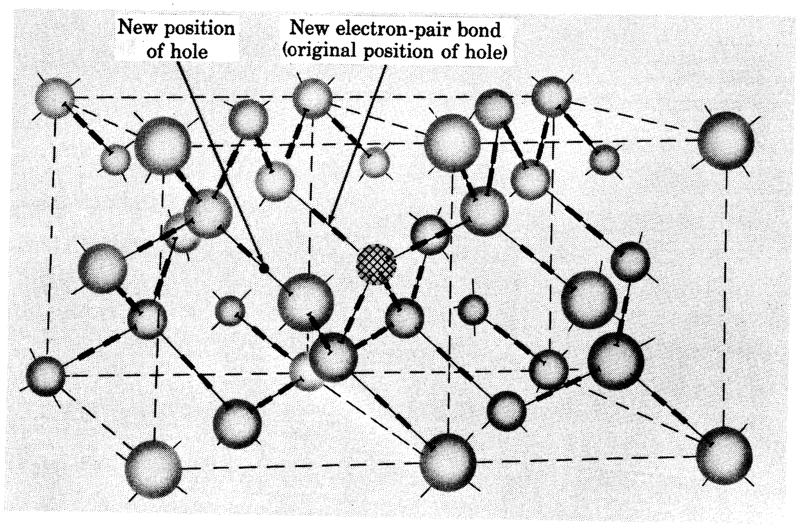


FIG. 2-9. Semiconductor crystal with acceptor atom showing new position of hole

Note that the hole has moved to a new position in Fig. 2-9. When the hole moves to the new position, two important changes take place.

1. The first change is that the acceptor atom has been ionized; i.e., the acceptor has acquired (or *accepted*) an electron and is now an ion. A negative charge exists in the immediate vicinity of the acceptor.

2. The second change is that the semiconductor atom, which requires four valence electrons, is left with only three valence electrons. The semiconductor atom, lacking an electron, has a net positive charge equivalent to the negative charge of the electron. Because of the crystal structure of the semiconductor, the positive charge of the semiconductor atom is not diffused or scattered, but is concentrated in the hole in the electron-hole arrangement. Furthermore, laboratory experiments have shown that the positive hole moves within the crystal in the same manner that an excess electron moves within the crystal. The concept of holes is very important in understanding the operation of semiconductor devices; the properties and characteristics of holes are discussed more thoroughly in paragraph 2-12.

d. Note that the semiconductor crystal which contains an acceptor ion (negative) also contains a hole (positive). The crystal taken as a whole therefore possesses a net charge of zero.

e. Semiconductor materials containing acceptor impurities are referred to as *p-type semiconductors.* The letter *p* refers to the *positive* charge of the hole.

2-11. Movement of Hole

Figure 2-10 is a two-dimensional representation of the mechanism involved in the movement of a hole through a crystal. In Fig. 2-10A, the hole is in the upper left-hand corner. An electron from an adjacent electron-pair bond moves to the position of the hole. The hole (Fig. 2-10B) is now midway between top and bottom of the crystal and slightly to the right of its original position. This process is repeated in Figs. 2-10C and D, until the hole is at the right-hand side of the crystal (Fig. 2-10E). The complete path of the hole through the crystal is shown in Fig. 2-10F.

2-12. Holes, Properties and Characteristics

For an understanding of the theory of semiconductor devices, it is convenient for the reader to think of the hole as a specific particle. Holes in motion, like electrons in motion, constitute an electrical current (par. 2-11). There are differences, however, which must be kept in mind:

a. The hole can exist only in a semiconductor material, such as germanium or silicon. This is because the hole depends for its existence on a specific arrangement of electrons (electron-pair bonds) and atoms as is found in crystal substances (par. 2–10). Holes *do not* exist in such conductors as copper and aluminum.

b. The hole is deflected by electric and magnetic fields in the same manner that electrons are deflected. Because the hole possesses a charge equal and opposite to that of the electron, the direction of deflection of the hole is

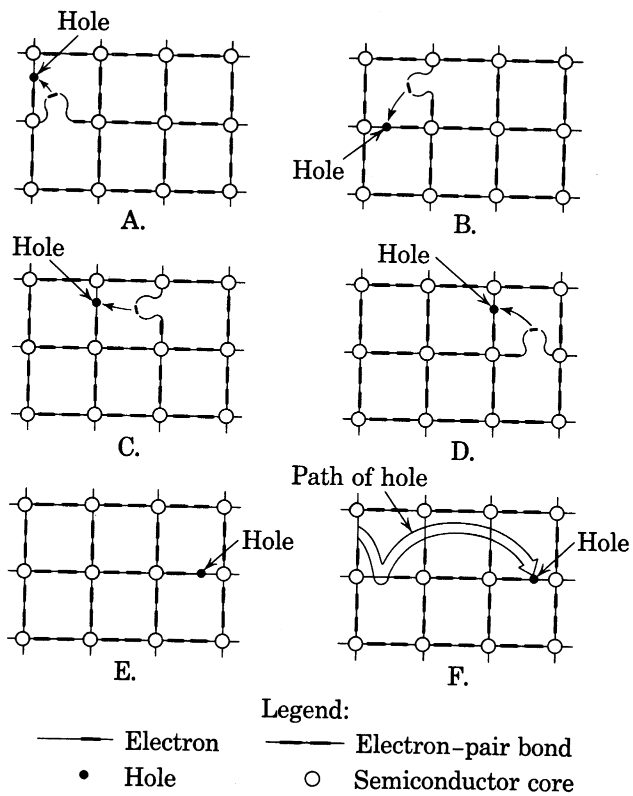


FIG. 2–10. Movement of hole through crystal

opposite to that of the electron. In an electric field, for instance, the electron moves toward the positive potential; the hole moves toward the negative potential.

c. In the field of electronics, the electron is considered indestructable. When a hole is filled by an electron from an adjacent electron-pair bond, the hole is considered as having moved from one position to another (Figs. 2–8 and 2–9). When a hole is filled by a free or excess electron, the hole no longer exists. This statement is supported by the fact that a semiconductor material containing an equal number of donor and acceptor atoms has none of the properties of the *p*-type or *n*-type semiconductor.

SECTION III. *p-n* JUNCTIONS

2–13. General

When an *n*-type semiconductor (par. 2–9) and a *p*-type semiconductor (par. 2–10) are joined in the same crystal, an unusual but important phenomenon occurs at the surface where contact is made between the two types of semiconductor. The contact surface is referred to as a *p-n junction.*

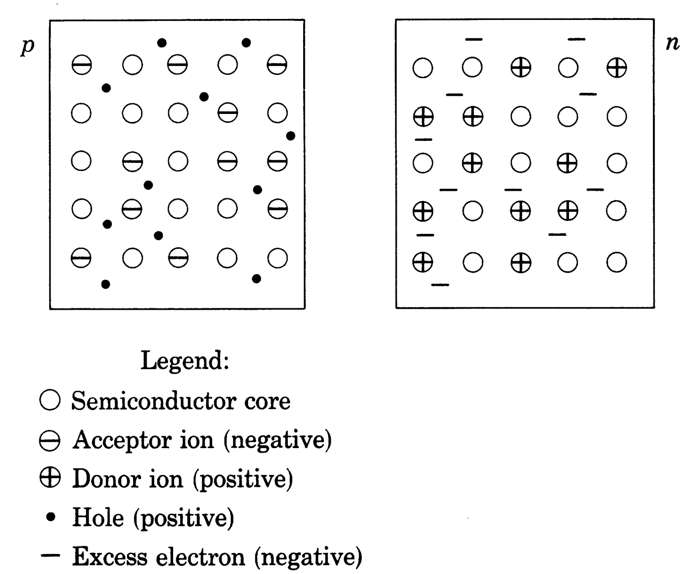


FIG. 2–11. Separated sections of *p*-type and *n*-type semiconductor

The phenomenon that occurs at *p-n* junctions permits the use of semiconductors, such as germanium and silicon, in circuits normally employing electron tubes. The detailed theory of *p-n* junctions is covered in paragraphs 2–14 through 2–18.

2–14. *p-n* Junction, General

a. Figure 2–11 shows a section of *p*-type semiconductor and a section of *n*-type semiconductor.

1. For clarity, the electron-pair bonds are not shown; only the holes, the excess electrons, the semiconductor cores, and the donor and acceptor ions are represented.

2. For discussion purposes, Fig. 2–11 shows a large number of acceptor ions in the *p*-type and a large number of donor ions in the *n*-type germanium.

In practice, however, semiconductor materials, such as those used for transistors and rectifying diodes, contain approximately one impurity atom per 10 million germanium atoms.

3. If one could actually look inside the bulk semiconductor material, one would see the semiconductor cores and the impurity ions vibrating within their lattice positions because of thermal energy. However, the cores and

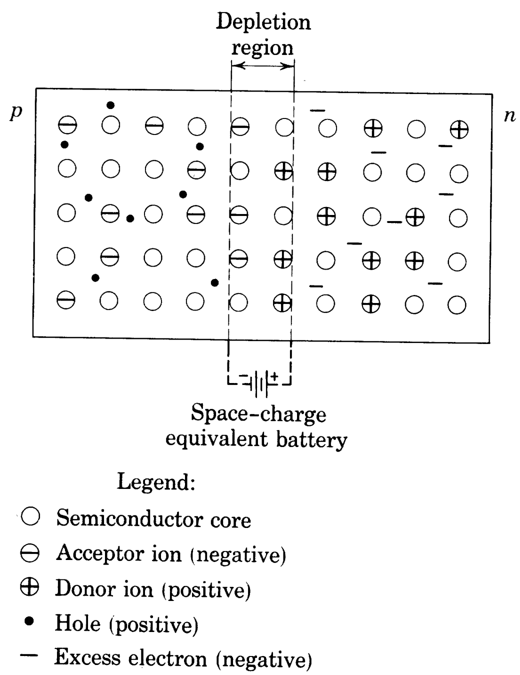


FIG. 2–12. Joined sections of p-type and n-type semiconductor

the ions *do not* leave their lattice positions and therefore do not constitute a current. The cores and the ions may be considered to be stationary. The holes and the excess electrons would be seen to move at random within the material. The movement of the holes and the electrons is due to thermal energy; this movement of charges in the absence of an applied field is called *diffusion.* (Diffusion is the movement of carriers from a region of high concentration to a region of lower concentration.) Even though the holes are in motion they are evenly distributed throughout the p-type semiconductor; the excess electrons are evenly distributed throughout the n-type semiconductor.

b. Figure 2–12 shows the same two sections of semiconductor (*a* above) joined to form a p-n junction.

1. Note that no external circuits or voltages have been connected to the material; nor is the material exposed to external electric or magnetic fields.

2. One would normally expect the holes in the p-type semiconductor and the electrons in the n-type semiconductors to flow toward each other, combine, and eliminate all holes and excess electrons. When the two types of semiconductor are joined, however, after a few combinations of holes and electrons result, a restraining force is set up automatically to preclude total combination. This restraining force is called a barrier. The cause and nature of the barrier are discussed in paragraph 2–15.

2–15. Junction Barrier

a. When the p-type semiconductor and the n-type semiconductor are joined (Fig. 2–12) some of the holes in the p-region and some of the excess electrons in the n-region diffuse toward each other and combine. Each combination eliminates a hole and an *excess* electron; the excess electron is now part of an electron-pair bond. This action occurs for a short time in the immediate vicinity of the junction. Negative acceptor ions in the p-region and positive donor ions in the n-region and near the junction are left uncompensated. Additional holes that would diffuse into the n-region are repelled by the uncompensated positive charge of the donor ions. Electrons that would diffuse into the p-region are repelled by the uncompensated negative charge on the acceptor ions. As a result, *total* recombination of holes and electrons cannot occur.

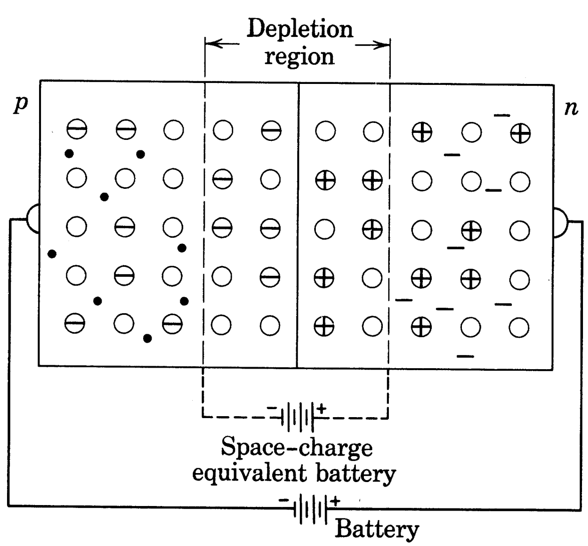
b. The region containing the uncompensated acceptor and donor ions is referred to as the *depletion region,* i.e., there is a depletion of holes and a depletion of excess electrons in this region. Since the acceptor and the donor ions are immobile (fixed) and are charged electrically, the depletion region is also referred to as the *space-charge region.* The electric field between the acceptor and the donor ions is called a *barrier.* The effect of the barrier is represented by the imaginary space-charge equivalent battery. The physical distance from one side of the barrier to the other is referred to as the *width* of the barrier, the width of the barrier depending on the density of holes and excess electrons in the germanium crystal. The difference of potential from one side of the barrier to the other is referred to as the *height* of the barrier. The height of the barrier is the intensity of the electric field (voltage of space-charge equivalent battery) and is measured in volts. With no *external* batteries connected, the barrier height is on the order of tenths of a volt.

c. It is stated in *a* above that total recombination of electrons and holes cannot occur. Inspection of the polarity of the space-charge equivalent battery confirms this statement. Note that the electrons in the n-type germanium are already at the *highest* positive potential (positive terminal of space-charge equivalent battery) with the crystal. The holes in the p-type germanium also are at the highest negative potential (negative terminal of space-charge equivalent battery) within the crystal. This condition precludes the movement of holes or electrons across the p-n junction.

2–16. p-n Junction, Reverse Bias

a. Figure 2–13 shows what happens when an external battery with the indicated polarity is connected to a p-n junction. Note that the negative terminal of the battery is connected to the p-type semiconductor and that the positive terminal of the battery is connected to the n-type semiconductor. The holes are attracted toward the negative terminal and away from the junction. The electrons are attracted toward the positive terminal and away from the junction. This action widens the depletion region and increases the barrier height (potential). Compare the width of the depletion region of Fig. 2–12 with that of Fig. 2–13.

b. Since the depletion region widens until the barrier height (potential of space-charge equivalent battery) equals the potential of the external



Legend:
- ○ Semiconductor core
- ⊖ Acceptor ion (negative)
- ⊕ Donor ion (positive)
- • Hole (positive)
- – Excess electron (negative)

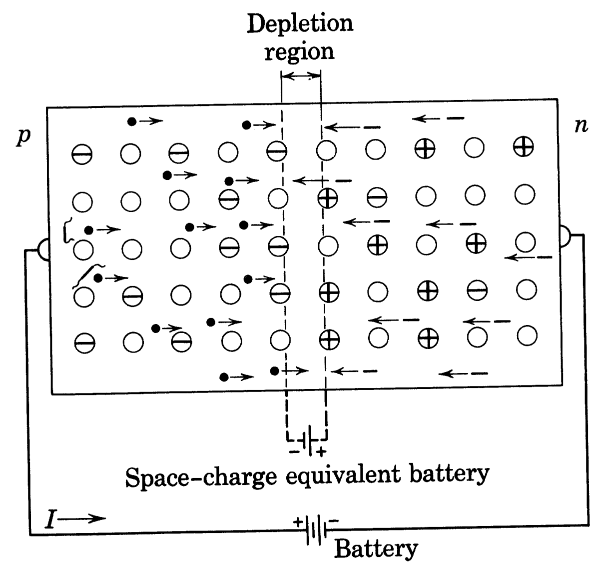
FIG. 2–13. p-n junction showing reverse bias

battery, no current flow of holes or electrons occurs because the battery voltages are in opposition. In this condition, the p-n junction is biased in the *reverse* direction; or simply, a *reverse bias* is placed across the p-n junction.

c. It is possible to apply a reverse bias greater than the largest possible barrier height. However, if this is done, the crystal structure will break down. In normal applications, this condition is avoided. The crystal structure will return to normal when the excess reverse bias is removed, provided that overheating does not permanently damage the crystal.

2–17. p-n Junction, Forward Bias

a. Figure 2–14 shows what happens when an external battery with the indicated polarity is connected to a p-n junction. Note that the positive



Legend:
- ○ Semiconductor core
- ⊖ Acceptor ion (negative)
- ⊕ Donor ion (positive)
- • Hole (positive)
- – Excess electron (negative)
- ⌒ Electron from electron-pair bond

FIG. 2–14. p-n junction showing forward bias terminal of the battery is connected to the *p*-type semiconductor and the negative terminal of the battery is connected to the *n*-type semiconductor. The holes are repelled from the positive terminal of the battery and drift toward the junction. The electrons are repelled from the negative terminal of the battery and drift toward the junction. Because of their acquired energy, some of the holes and the excess electrons penetrate the depletion (space-charge) region and combine.

b. For each combination of an excess electron and a hole that occurs, an electron from the negative terminal of the external battery enters the *n*-type semiconductor and drifts toward the junction. Similarly, an electron from an electron-pair bond in the crystal, and near the positive terminal of the external battery, breaks its bond and enters the positive terminal of the external battery. For each electron that breaks its bond, a hole is created which drifts toward the junction. Recombination in and about the space-charge region continues as long as the external battery is applied.

c. Note that there is a continuous electron current (I) in the external circuit as indicated by the arrow. The current in the *p*-type semiconductor consists of holes; the current in the *n*-type semiconductor consists of electrons. In this condition, the *p-n* junction is said to be biased in the forward direction. If the forward bias is increased, the current increases.

Note: Throughout this text *current flow* refers to *electron current* rather than *conventional current.*

d. In paragraph 2–15*b* it was stated that the barrier potential with no external battery connected is on the order of tenths of a volt. It would appear, therefore, that an external battery of very low voltage (about 1 volt) would eliminate the barrier completely. However, the larger the voltage of the external battery, the greater the current flow through the crystal. Since the crystal has a relatively high resistivity, increased current causes increased voltage drop on both sides of the barrier. The remaining voltage of the external battery does not overcome the barrier completely. Normally, 1 to 1½ volts is used to bias the *p-n* junction in the forward direction. If excessive forward bias is used, excessive current will cause excessive thermal agitation and breakdown of the crystal structure.

2–18. Diode Action

a. Paragraphs 2–14 through 2–17 cover the mechanism of rectification through a *p-n* semiconductor diode. Figure 2–15 is a plot of current flow versus voltage applied to a practical *p-n* junction. Note that current flow in the forward bias direction is quite high (measured in ma). However, current flow in the reverse bias direction, although low (measured in μa), is not zero as might have been expected (par. 2–16).

1. In the *p*-type semiconductor some electrons in the electron-pair bonds gain enough energy to break and move out of the electron-pair bond structure. This action produces some additional holes which add to the existing holes caused by the acceptor impurities. The electrons that have broken their bonds become excess (free to move) electrons; these electrons are called *minority* carriers in the *p*-type semiconductor because they are always outnumbered by the holes which are referred to as *majority* carriers. The energy required to release an electron from an electron-pair bond may be in the form of heat, light, or other radiation. The opposite action can also occur; i.e., the excess electron can lose energy and fall into a hole, thus eliminating the excess electron and the hole. This action of generation and elimination of excess electrons and holes goes on continuously within the semiconductor material even at average room temperature.

2. In the *n*-type semiconductor the electrons that gain enough energy (1 above) to break their bonds, create excess electrons that add to the *majority* carriers (electrons provided by the donor impurities). The resultant holes, always outnumbered by the excess electrons, are referred to as the *minority* carriers.

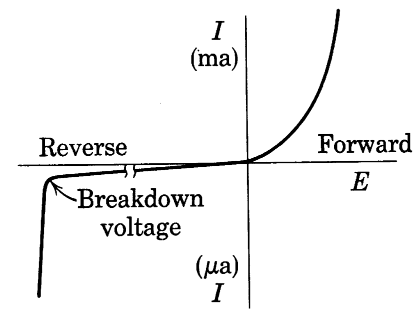


FIG. 2–15. Chart of current through, and the voltage across, a *p-n* junction

b. When the *p-n* junction is biased in the reverse direction for the majority carriers (par. 2–16), the *p-n* junction is biased in the *forward* direction (par. 2–17) for the minority carriers, electrons in the *p*-type semiconductor and holes in the *n*-type semiconductor. The internal mechanism of conduction for the minority carriers when *forward* biased (majority carriers reverse biased) is identical with that for forward-biased majority carriers.

c. Note that when a very high reverse bias is applied (Fig. 2–15), a high reverse current flows. This high current is not due to the minority carriers. A breakdown of the single crystal structure occurs (par. 2–16*c*).

SECTION IV. *p-n* JUNCTION, ENERGY LEVEL CONSIDERATIONS

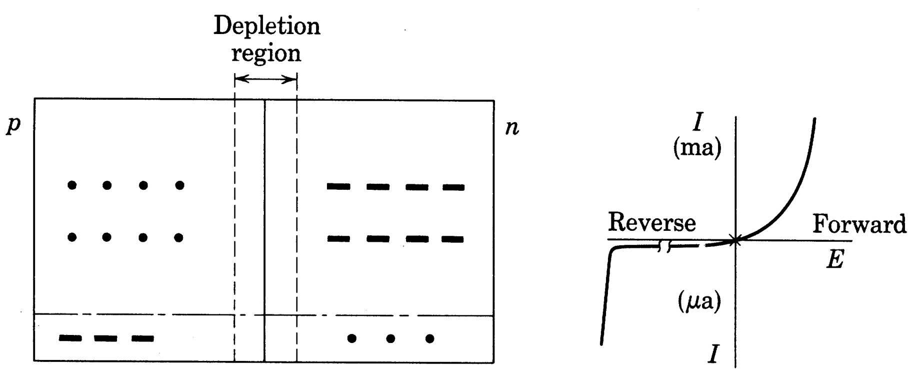
2–19. General

a. The *p-n* junction conduction and nonconduction mechanisms as discussed in paragraphs 2–13 through 2–18 are based purely on the attraction and repulsion of unlike and like charges, respectively. To that extent the discussions are valid. However, to understand the difference in the conduction mechanisms between the normal diode and the tunnel diode, it is necessary to reconsider the normal diode conduction mechanism to take into account relative energy levels of holes and excess electrons. These energy levels during normal diode action are discussed in paragraphs 2–20, 2–21, and 2–22.

b. Figures 2–16, 2–17, and 2–18, showing no bias, reverse bias, and forward bias correspond, respectively, to Figs. 2–12, 2–13, and 2–14. Figures 2–12, 2–13, and 2–14 are more realistic in that they show more accurately the spacial distribution of holes and excess electrons throughout the semiconductor. Figures 2–16, 2–17, and 2–18 line up the holes and excess electrons according to their relative energy levels. This line-up of current carriers helps visualize the effect of energy levels on *p-n* junction action. For greater clarity these figures do not show semiconductor atoms or impurity ions. Additional simplification is obtained by separating the minority and majority carriers by a long-short dash line.

2–20. Energy Levels, Unbiased *p-n* Junction

Figure 2–16 shows essentially the same conditions as Fig. 2–12 when a region of *p*-type semiconductor and *n*-type semiconductor are joined. In the usual manner a depletion region (par. 2–15) is formed which acts as a barrier to the net flow of holes or excess electrons across the junction.



Legend:
- • Hole (positive)
- ▬ Excess electron (negative)

Majority carriers above long-short dash line
Minority carriers below long-short dash line

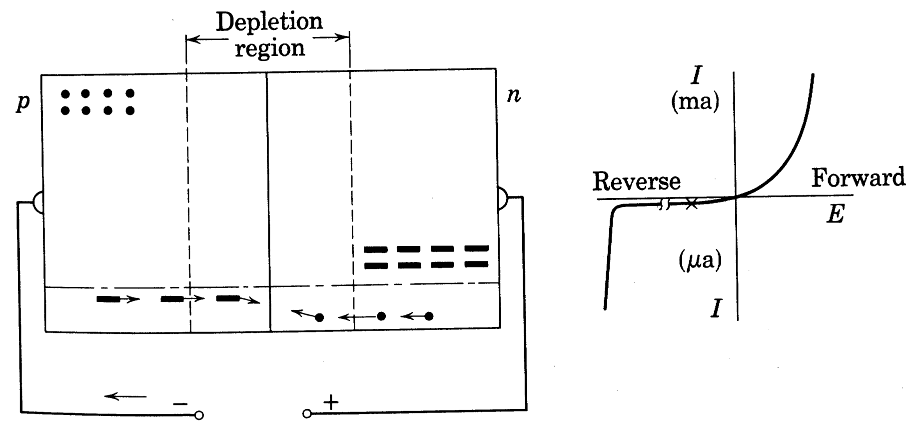
FIG. 2–16. *p-n* junction, unbiased, showing relative electron-energy levels of majority carriers and relative electron-energy levels of minority carriers

a. Assume that the long-short dash line is the lowest electron energy level for the majority carriers and that electron energy increases vertically up. Then the condition of equilibrium shown here is a small but equal number of majority holes and majority excess electrons at a medium energy level. The × on the current-voltage chart shows zero current flow caused by majority carriers.

b. The same condition (*a* above) prevails for the minority carriers. The lowest electron-energy level for the minority carriers is considered to be the bottom edge of the device. Note that relative energy levels between minority and majority carriers are not demonstrated in this diagram because they are not important to the discussion. As in the case of the majority carriers there is no net current flow across the barrier due to the minority carriers.

2–21. Energy Levels, Reverse-biased *p-n* Junction

Figure 2–17 shows what happens when a reverse bias (negative to *p*-type material and positive to *n*-type material) is applied. As previously explained (par. 2–16), the depletion width increases because the majority electrons move toward the positive terminal and the majority holes move toward the negative terminal. There is no current flow due to majority



Legend:
- • Hole (positive)
- ▬ Excess electron (negative)

Majority carriers above long-short dash line
Minority carriers below long-short dash line

FIG. 2–17. *p-n* junction, reverse biased, showing relative electron-energy levels of majority carriers, and relative electron-energy levels of minority carriers

carriers. There is a small current flow due to opposite movements of the minority carriers. This reverse-biased current is marked × on the current-voltage chart.

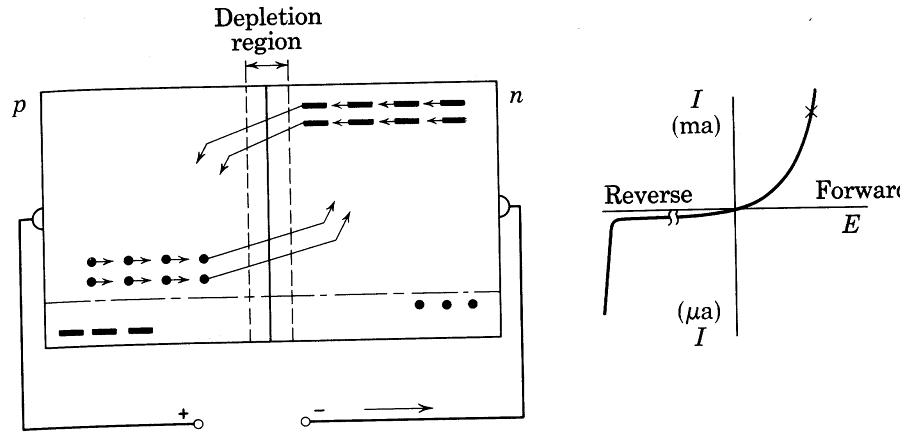
a. The same phenomenon is explained in terms of the energy levels of the carriers. If a battery were placed across an ordinary conductor, electrons would flow from the negative terminal of the battery, through the conductor, to the positive terminal of the battery. The reason for the current flow is that the *negative* potential excites the electrons to a high energy level proportional to the magnitude of the negative potential. With the greater energy, the loosely held electrons of the conductor atoms leave their atoms. Electrons will always move toward the lowest energy level; the positive terminal of the battery represents the lowest electron-energy level. This phenomenon is comparable to water seeking its lowest level. Water at a high level, if free to flow, will flow downhill. In its movement downhill, the water gives up energy. This energy is often used to drive hydroelectric plants.

b. Note what happens to the electron-energy levels of the majority carriers (Fig. 2–17) when a reverse bias is applied. The majority electrons, at the highest positive battery potential, are also at the *lowest* electron-energy level. *The electron-energy level of the hole indicates the energy level that electrons would assume if they occupied the position of the hole.* Since electrons seek the *lowest* energy level, the majority electrons do not move into the holes. No current flow due to majority carriers can be expected any more than water can be expected to flow uphill.

c. In the case of the minority carriers, however, the situation is reversed. The minority electrons are raised to a high energy level by the negative potential of the battery; the holes are reduced to a low electron-energy level by the positive terminal of the battery. The minority electrons therefore will readily flow from their high energy level to the low energy level of the holes. A small external current in the direction shown can be detected in the external circuit.

2–22. Energy Levels, Forward-biased *p-n* Junction

a. Figure 2–18 shows a forward-biased *p-n* junction (positive terminal of battery to *p*-type semiconductor, negative terminal to *n*-type semiconductor). Because of the negative potential of the battery, the majority electrons are raised to a high electron-energy level. The positive potential of the battery places the majority holes at a low electron-energy level. The majority electrons and the majority holes will readily flow toward the barrier. Within and about the barrier, numerous combinations of holes and electrons take place. This current flow due to a forward bias *high enough* (approximately 0.5 to 1 volt) to excite the majority electrons to an energy



Legend:

- • Hole (positive)
- ▬ Excess electron (negative)

Majority carriers above long-short dash line

Minority carriers below long-short dash line

FIG. 2–18. *p-n* junction, forward biased, showing relative energy levels of majority carriers and relative energy levels of minority carriers

level sufficient to overcome the barrier in large numbers is referred to as an *injection* current. A large external current flows in the direction indicated by the arrow. This current is marked by a × on the current-voltage chart.

b. The applied forward bias causes the minority electrons to assume a low energy level and the minority holes to assume a high energy level. Minority current flow therefore does not occur.

SECTION V. p_T-n JUNCTION, TUNNEL DIODES

2–23. General

a. The normal rectifying diode discussed in paragraphs 2–19 through 2–22 uses semiconductor materials lightly doped with one impurity atom for 10^7 (10 million) semiconductor atoms. If the semiconductor materials forming a junction are doped to the extent of *1000 impurity atoms* for 10^7 semiconductor atoms, the resultant current-voltage characteristic appears as shown in Fig. 2–19. Compare Figs. 2–15 and 2–19. The most important aspect of this characteristic (Fig. 2–19) is the peak current (I_p) rise with a small applied forward bias, the decreasing current with *increasing* forward bias to a minimum (valley) current (I_v), and finally a normal increasing current with further increasing voltage. That portion of the characteristic between I_p and I_v represents an *ac negative resistance* which permits the use of this device in circuits requiring ac amplification (par. 3–6 through 3–34). The internal conduction mechanism that causes this phenomenon is discussed in paragraphs 2–24 through 2–28.

b. The most important effect of heavy doping of semiconductor materials is the effect of the quantity of doping on the width of the depletion region. The heavier the doping, the narrower the depletion region. Figure 2–20A shows a *p-n* junction formed by lightly doped semiconductor materials. For clarity, the holes, excess electrons, and electron-pair bonds are not shown. If heavier doping is used (Fig. 2–20B), the width of the depletion region is reduced. If the two types of semiconductor materials contain different amounts of impurities, the portion of the depletion region in the more heavily doped material is narrower than that in the lightly doped material (Fig. 2–20C). The reason for these conditions is that the barrier (electric field of the depletion region) is formed by a given number of donor ions on one side of the barrier and an *equal* number of acceptor ions on the other side of the barrier. The heavier the impurity concentration in the semiconductor material, the smaller the depth of penetration on that side of the junction necessary to establish a given number of uncompensated ions.

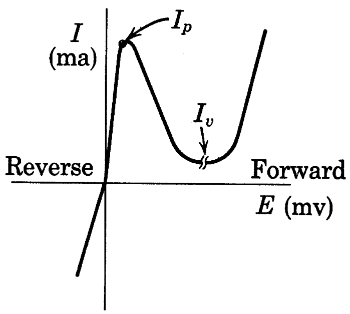


FIG. 2–19. Current-voltage chart of heavily doped (tunnel) diode

c. Because of the heavy doping used in tunnel diodes, the depletion region is only *one millionth of an inch* wide. This fact is important in explaining the current-voltage characteristic (Fig. 2–19) of the tunnel diode.

2–24. p_T-n Junction, Zero Bias

a. Figure 2–21 shows a junction formed by two heavily doped semiconductor materials. As in the case of the normal *p-n* junction (par. 2–20), a depletion region is formed even with zero bias. Note, however, that the depletion region is very narrow; compare Fig. 2–21 with Fig. 2–16. This ultra-thin depletion region is referred to throughout this text as a p_T-n junction.

b. Note that a *large* number of majority holes and excess electrons are at the same electron-energy levels (par. 2–20). The same condition is true of the minority carriers. If there is any movement of carriers across the

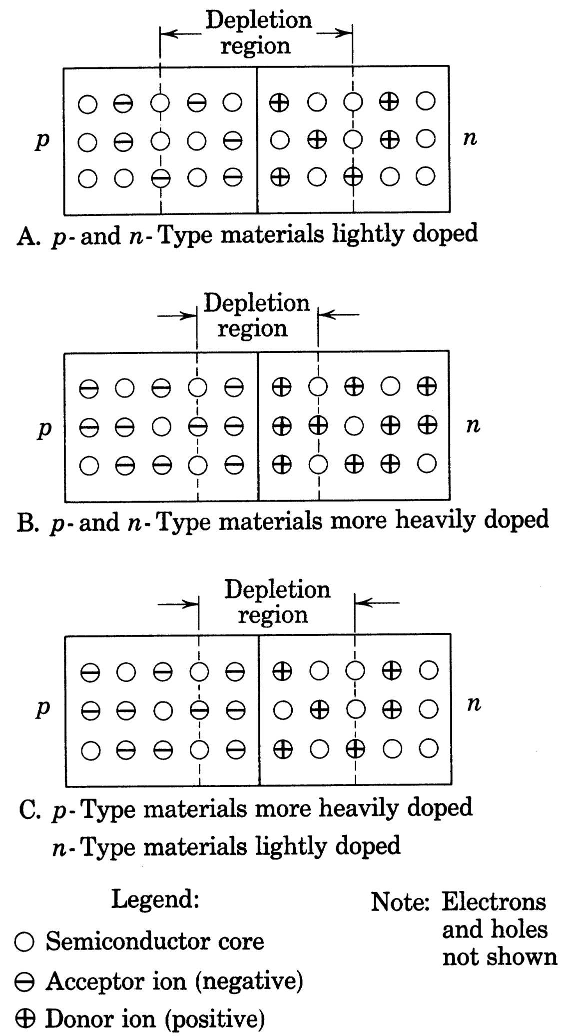
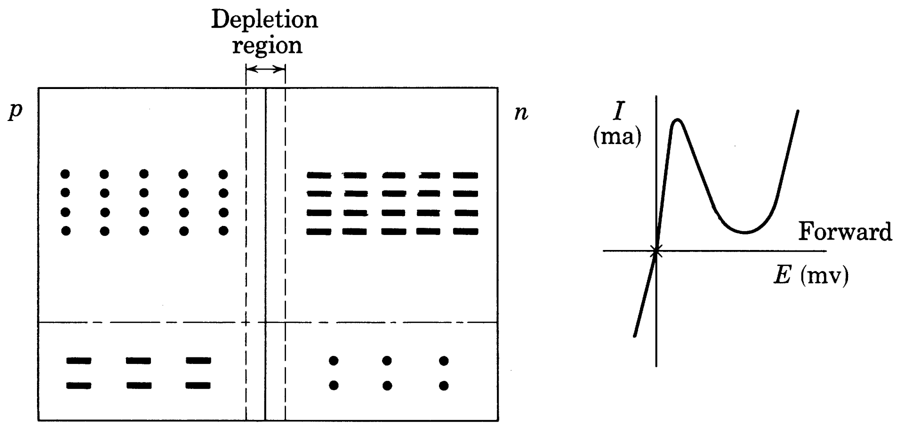


FIG. 2–20. *p-n* junctions, showing variation of depletion-region width with amount of doping of semiconductor materials

depletion region due to thermal energy, the net current flow will be zero because an equal number of like charges will flow in opposite directions. This zero net current flow is marked by an × on the current-voltage chart.

2–25. Junction with Forward Bias and Peak Current

a. Figure 2–22 shows a p_T-n junction with a small forward bias (about 50 mv) applied. The negative potential of the battery connected to the *n*-type material slightly raises the electron-energy levels of the majority electrons. The positive potential of the battery connected to the p-type material slightly lowers the electron-energy level of the majority holes. The depletion region is also slightly narrowed. For normal p-n junctions no net current flow would occur since approximately 0.5 to 1 volt forward bias is required. However, a substantial current flow is measured externally in the direction of the arrow. The magnitude of current is marked by an × on the current-voltage chart.



Legend:

- • Hole (positive)
- ▬ Excess electron (negative)

Majority carriers above long-short dash line
Minority carriers below long-short dash line

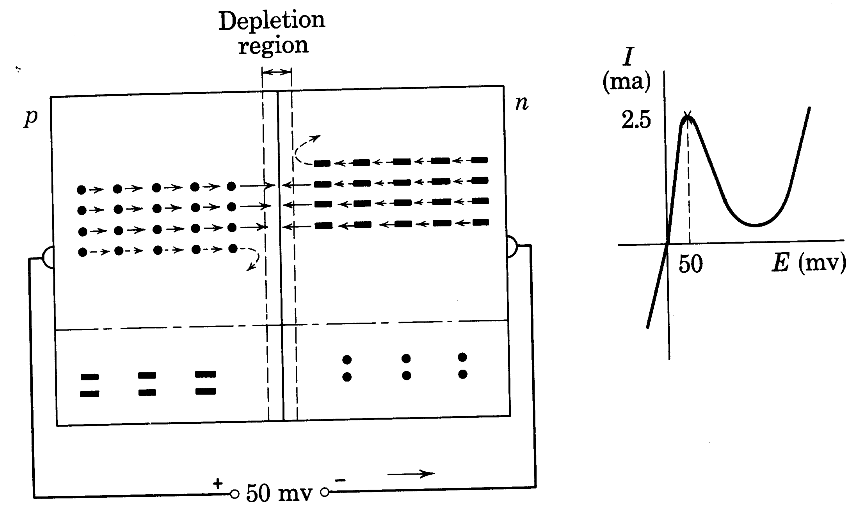
FIG. 2–21. p_T-n junction, showing narrow depletion region, energy level of carriers, and current-voltage chart

b. The phenomenon of high current with such small bias is referred to as *electron tunneling. Electrons will penetrate an ultra-thin barrier that normally they could not penetrate provided that there are vacancies (holes) on the opposite side of the barrier at the same electron-energy level as the penetrating electrons.* Note that the top row of majority electrons that move toward the barrier are deflected *back* into the n-type material because no vacancies (holes) of equal electron-energy level are present in the p-type material. The bottom three rows of majority electrons and the top three rows of majority holes do combine and cause current flow externally in the usual manner.

c. The majority electrons that penetrate the barrier appear in the p-type

material at the same energy level they had in the n-type material. This action is comparable to a man walking through a *tunnel* in a mountain that he could not possibly climb over. Thus the term *electron tunneling;* the electrons penetrate a barrier which normally they could not penetrate and appear at the same energy level on the other side of the barrier.

d. No current flow is caused by the minority carriers.



Legend:

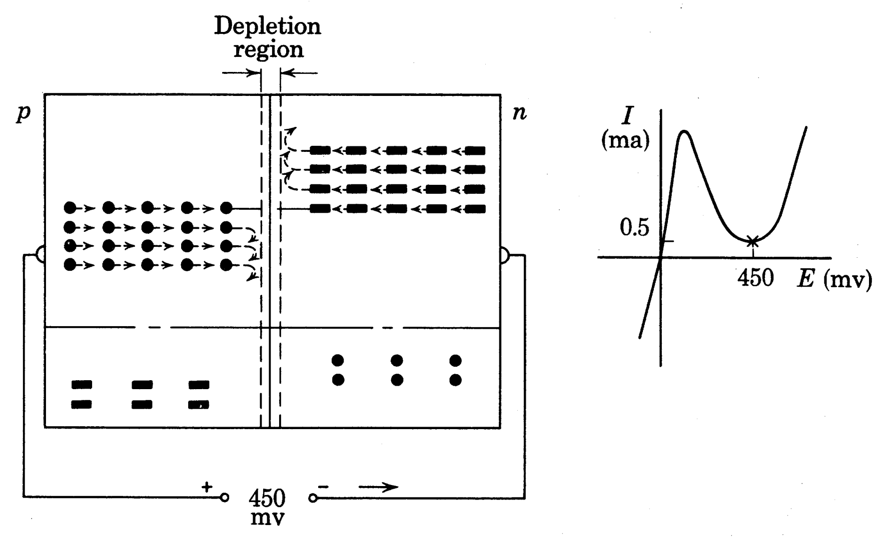
- • Hole (positive)
- ▬ Excess electron (negative)

Majority carriers above long-short dash line
Minority carriers below long-short dash line

FIG. 2–22. p_T-n junction, slight forward bias, showing electron tunneling and peak current on current-voltage chart

2–26. p_T-n Junction with Forward Bias and Valley Current

a. Figure 2–23 shows a p_T-n junction with forward bias increased [from 50 mv (par. 2–25)] to 450 mv. The negative potential of the battery further increases the electron-energy level of the majority electrons; the positive potential further decreases the electron-energy level of the majority holes. Note that only one row of majority electrons and majority holes are at the same energy level. Very little electron tunneling (par. 2–25) occurs and only a small current (marked by × on the current-voltage chart) can be measured in the external circuit. This current is referred to as the



Legend:

- • Hole (positive)
- ▬ Excess electron (negative)

Majority carriers above long-short dash line

Minority carriers below long-short dash line

FIG. 2–23. p_T-n junction, increased forward bias, showing little electron tunneling and valley current on current-voltage chart

valley current. Most of the majority electrons and holes are deflected back from the barrier.

b. No current flow is caused by the minority carriers.

2–27. p_T-n Junction with Forward Bias and Injection Current

a. Figure 2–24 shows the p_T-n junction with forward bias increased [from 450 mv (par. 2–26)] to 500 mv. The negative potential of the battery further raises the electron-energy level of the majority electrons; the positive potential further decreases the electron-energy level of the holes. All of the majority electrons have gained sufficient energy to penetrate the barrier and combine with the holes of low electron-energy level. The resultant current flow is the same as that which occurs in a normal p-n junction (par. 2–22) and is referred to as *injection* current to distinguish it from *tunnel* current.

b. The forward-bias voltage at which the injection current occurs in a p_T-n junction is approximately the same voltage at which injection current

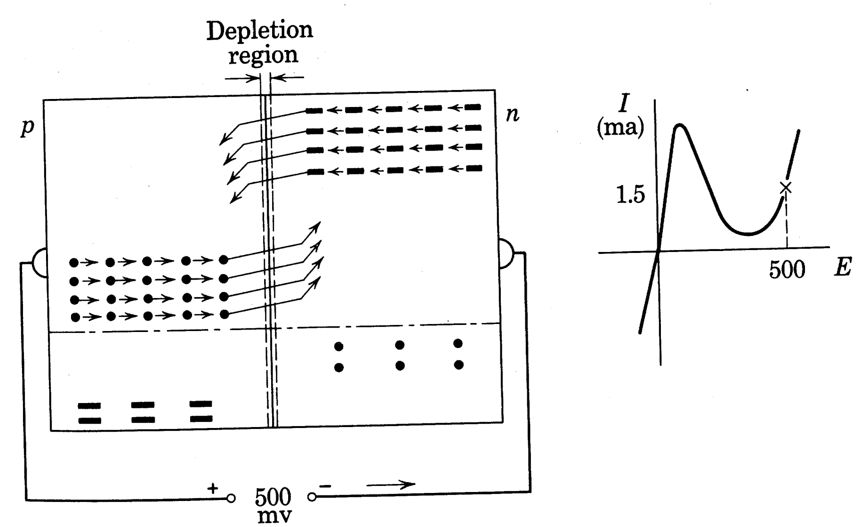
would occur in a p-n junction made of the same semiconductor material.

c. The current flow from zero bias through the peak current down to the valley current is called tunnel current. Beyond the valley current, it is called injection current.

2–28. p_T-n Junction, Reverse Biased

a. Figure 2–25 shows a p_T-n junction reverse biased. The majority electrons are greatly reduced in electron-energy level and the majority holes are greatly increased in electron-energy level. No current flow due to majority carriers occurs.

b. The application of negative potential to the p-type material raises the electron-energy level of the minority electrons originally present. It also increases the number of minority electrons by causing electrons from electron-pair bonds to break their bonds and become excess electrons. An equal number of holes are also formed, and these join the original number of majority holes. Compare the numbers of majority holes and minority electrons in Figs. 2–24 and 2–25. This generation of holes and excess elec-



Legend:

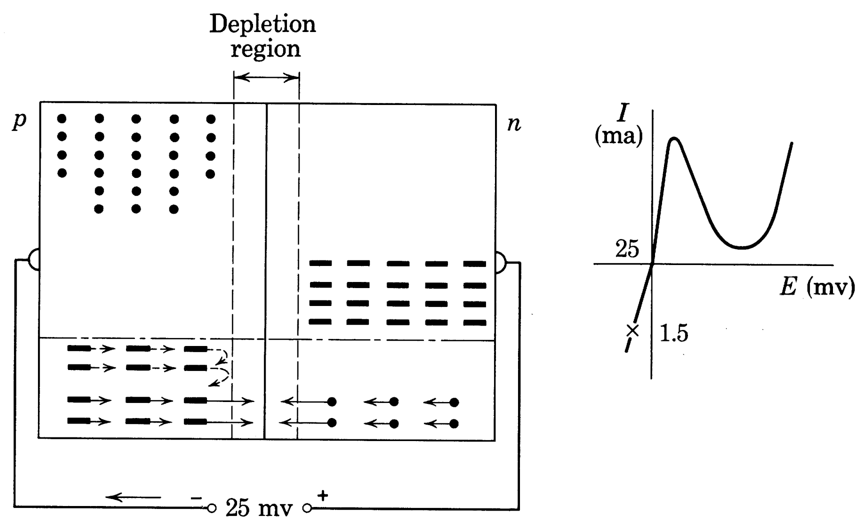
- • Hole (positive)
- ▬ Excess electron (negative)

Majority carriers above long-short dash line

Minority carriers below long-short dash line

FIG. 2–24. p_T-n junction, forward biased to cause injection current as marked on current-voltage chart trons occurs only because the semiconductor is heavily doped and acts very much like an ordinary conductor.

c. Although the width of the depletion region has been slightly increased by the reverse bias, the increase is not sufficient to prevent electron tunneling between the minority electrons and the minority holes that are at



Legend:
- • Hole (positive)
- ▬ Excess electron (negative)

Majority carriers above long-short dash line
Minority carriers below long-short dash line

Fig. 2–25. p_T-n junction, reversed biased, showing resultant current marked by × on current-voltage chart

the same electron-energy level. As a result, a relatively heavy reverse-bias current (marked by an × on the current-voltage chart) flows in the direction of the arrow with only 25 mv applied.

2–29. Peak and Valley Currents

a. The peak and valley currents of p_T-n junctions made from a particular semiconductor material always occur at the same peak voltage and valley voltage, respectively. Figure 2–26A shows the forward current-voltage chart of two tunnel diodes, each made of germanium. The differences in magnitude in peak currents shown for curve 1 and curve 2 can be achieved by

increasing or decreasing the amount of doping, respectively, or by varying the size of the junction area. Units varying in peak currents from 10 μa to 10 amp have been constructed. However, most present applications use units having peak currents of from 1 to 50 ma. The ratio of peak current to valley current ordinarily remains the same. The valley current is usually 10% of the peak current. Compare the magnitudes of peak to valley currents for each curve shown. The voltage at which the injection current equals the peak current is referred to as the injection-current (or forward) voltage. For germanium units this voltage occurs approximately between 500 mv and 600 mv.

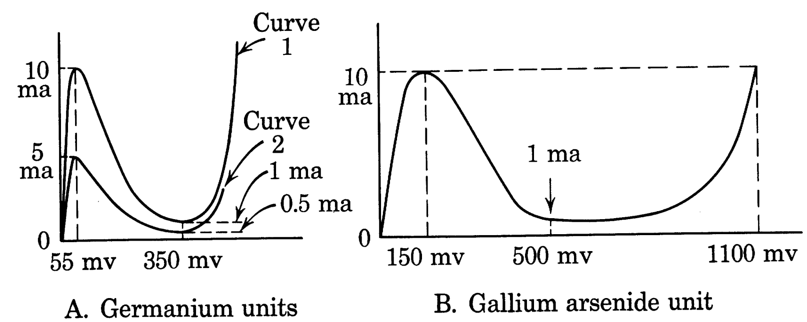


Fig. 2–26. Voltage-current charts for germanium and gallium arsenide tunnel diodes

b. For gallium arsenide unit (Fig. 2–26B), the peak current occurs at 150 mv, the valley current at 500 mv, and the injection current voltage at 1100 mv. This spread of approximately one volt between peak-tunnel voltage and injection-current voltage promises a wider use for gallium arsenide units, particularly in computer switching applications.

c. The fact that the peak-tunnel current and the valley current occur at the same respective voltages for a tunnel diode made of a particular material might have been expected from the theory of electron tunneling and injection (pars. 2–24 through 2–28). These conduction mechanisms depend upon specific energy levels of current carriers, which, in turn, are affected by the applied voltage and the forces existing between the orbiting electron and its nucleus.

SECTION VI. TUNNELTRONS

2–30. Tunneling Through Ultra-Thin Insulators

a. In experiments conducted at the General Electric Company, a scientist, Ivar Giaever, discovered that the phenomenon of tunneling occurs in devices other than semiconductors. In these experiments films of such conductors as aluminum, lead, and tin were separated by ultra-thin insulators—aluminum oxide, tantalum oxide, and nickel oxide (Fig. 2–27A). Two conductors separated by an insulator normally form a capacitor even if the insulator is only one ten-thousandth of an inch thick. In these experiments the insulating film was only 10 to 100 atoms thick. Furthermore, the tunnel effect (using the materials cited above) occurred only if the

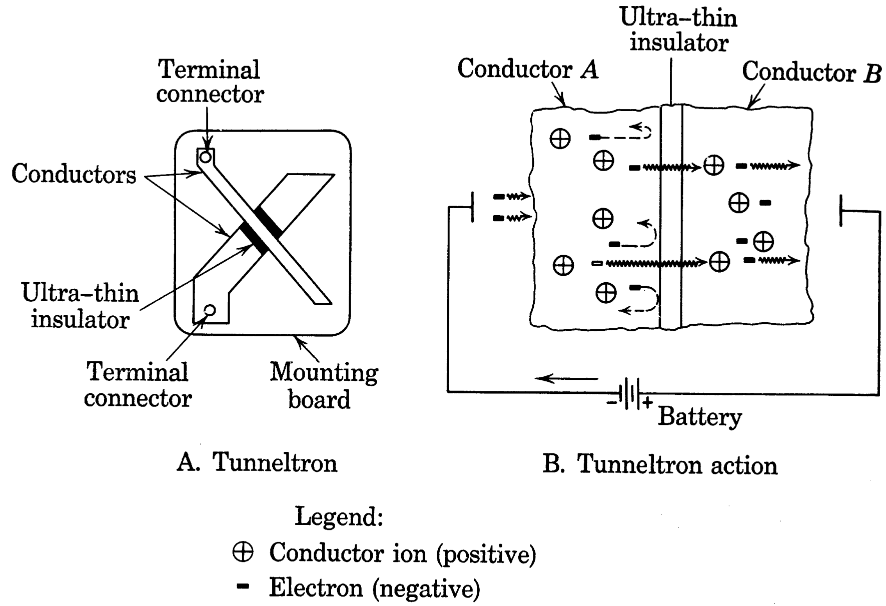


FIG. 2–27. Physical construction of tunneltron and tunneltron action

conductors were made superconducting (negligible in resistance) by cooling to liquid helium temperatures. The resultant device has been referred to as a *tunneltron*. (*Note:* Subsequent announcements by scientists at the Republic Aviation Corporation indicate the development of tunneltrons that operate and display the tunneling phenomenon at room temperatures. These devices use films of titanium as the conductors and a film of glass as the insulator.)

b. A diagrammatic representation of the theory of operation of the tunneltron is shown in Fig. 2–27B. All tunneling action takes place at the point where the three films of material overlap. A battery is connected across the tunneltron at the terminal connectors. For discussion purposes the conductor connected to the negative terminal of the battery is arbitrarily designated conductor *A*; the other conductor is designated conductor *B*. Assuming in this case that the conducting material used is aluminum, then the positive conductor ion represents the nucleus of the aluminum atom and twelve of its thirteen orbiting electrons. When one of the three (valence) outer orbit electrons (Fig. 2–2B), moves away from the atom, the remainder of the atom (now a positive ion) exerts a strong holding influence on the other two valence electrons.

c. With battery voltage applied as shown (Fig. 2–27B) one electron from each aluminum atom in conductor *A* moves away from the negative terminal of the battery and toward the ultra-thin insulating barrier. The electrons in conductor *B* have a tendency to move toward the positive terminal of the battery. The majority of the electrons in conductor *A* that reach the insulating barrier are *reflected* back into conductor *A*. Some of the electrons penetrate the barrier and enter conductor *B*. Although the majority of the electrons are reflected back, the electrons are so numerous that an appreciable current flows. For each electron that penetrates the barrier, one electron enters conductor *A* from the negative terminal of the battery; one electron leaves conductor *B* and enters the positive terminal of the battery.

d. At very low voltages, current increases with applied voltage until a current peak is reached. Further increase in voltage results in lower current values until a current minimum occurs, after which current increases with applied voltage. The net result of this action gives rise to an ac negative resistance identical with that obtained with the tunnel diode.

2–31. Advantages and Disadvantages of Tunneltron

a. The major disadvantage of the tunneltron developed by Mr. Giaever is that one or both of the conductors must be in a state of superconductivity. Superconductivity is achieved by reducing the temperature of the conductor to near liquid helium temperatures. This requirement can be met by operating the tunneltron in a liquid helium refrigerator which in many electronic applications would not be practical economically. However, this major disadvantage is overcome if the tunneltron is made of titanium and glass.

b. One important property of superconducting metals is the ability to make them nonsuperconducting by the application of strong external magnetic fields. Because the ac negative resistance displayed by the tunneltron depends upon superconductivity of metals, it is obvious that variation of the ac negative resistance can be achieved by using magnetic fields. Such variation of ac negative resistance is not possible with the tunnel diode. However, specific circuits employing this important characteristic of the tunneltron have not yet been devised.

c. Manufacturing processes for depositing films of metals in sandwiches of conductor-insulator-conductor as required by tunneltrons are standard and extremely inexpensive. It is expected therefore that costwise, tunneltrons can compete well with tunnel diodes.

d. At the present time the tunneltron is undergoing extensive study in an effort to determine specific and practical applications. At this writing no practical and commercial circuit applications of the tunneltron have been published. It is expected, however, that its properties will permit its application to low-noise amplifiers, high-frequency oscillators, and switching and memory circuits. Even if the tunneltron never emerges from the laboratory, its discovery has added greatly to the knowledge of tunneling phenomenon. If it does take its place among other practical active circuit devices, the discussions of ac negative resistance as presented in subsequent chapters of this text will be equally applicable to tunneltrons as well as tunnel diodes. Furthermore, it is expected that the tunneltron can, with minor circuit changes, substitute for the tunnel diode in many applications.

2–32. Comparison of Tunneltron and Tunnel Diode Actions

a. The basic tunneling concepts in the tunneltron and in the tunnel diode are identical; i.e., that electrons will move through an ultra-thin barrier provided there are vacancies (positive areas) on the other side of the barrier at the same energy level as the electrons. In the tunnel diode, the vacancy is a hole, a positive area in an electron-pair bond that appears to move freely. In the tunneltron the vacancy is a positive ion, an atom that has lost an electron and is not considered to be mobile. The hole is the result of the loss of an electron by a semiconductor (germanium, silicon, etc.) atom to an acceptor atom. The hole, therefore, although mobile, actually represents at any given instant of time a positive semiconductor ion. The positive area vacancies in the tunneltron and the tunnel diode then represent positive ions. The term *hole,* however, is reserved to designate a vacancy in the electron-pair bond structure of a crystal lattice network and is not applied to all positive ions.

b. The tunnel diode depends for its ultra-thin barrier on the depletion region between a heavily doped *p*-type semiconductor and a heavily doped *n*-type semiconductor; the barrier is created by an array of uncompensated negative acceptor ions on one side of a junction and uncompensated positive donor ions on the other side of the junction. The tunneltron depends for its ultra-thin barrier on a film of insulating material only several atoms thick.

2–33. Summary

a. Atoms are composed of positively charged particles called protons, negatively charged particles called electrons, and uncharged particles called neutrons.

b. A conductor is a material that has many loosely held electrons. Examples are silver, copper, and aluminum.

c. An insulator is a material that has few loosely held electrons. Examples are rubber, glass, and porcelain.

d. A semiconductor is a material, the resistivity of which is between those of conductors and insulators. Examples are germanium, silicon, and gallium arsenide.

e. A crystal is a material with atoms arranged in a specific pattern.

f. The properties of *polycrystalline* materials, e.g., copper and silver, are quite different from those of *single* crystal materials; single crystal materials are prepared for use in electron semiconductor devices. Germanium and silicon may be processed as single crystal materials.

g. Electrons shared by adjacent atoms in a crystal form electron-pair bonds.

h. *n*-Type germanium contains donor impurities. Donor impurities are materials that have *five* valence electrons, one of which cannot form an electron-pair bond. This electron is called an excess electron.

i. Arsenic, antimony, and phosphorus are examples of donor materials.

j. *p*-Type germanium contains acceptor impurities. Acceptor impurities are materials that have three valence electrons. Because *four* valence electrons are required to form and complete all adjacent electron-pair bonds, a hole is created.

k. Aluminum, gallium, boron, and indium are examples of acceptor impurities.

l. A hole can be considered a positive charge which diffuses or drifts through a crystal. The drift of holes constitutes a current.

m. A depletion (space charge) region occurs at a *p-n* junction. The potential difference across the depletion region is called a barrier. The width of the barrier is the width of the depletion region. The potential difference is called the height of the barrier.

n. Forward bias of a *p-n* junction causes heavy current (flow of majority carriers). Reverse bias causes very low current (flow of minority carriers).

o. p_T-n Junctions are formed by heavily doped *p*- and *n*-type semiconductor materials which result in ultra-thin (one millionth of an inch) depletion regions.

p. Electron tunneling will occur only through ultra-thin barriers provided that excess electrons on one side of the barrier and holes (vacancies) on the other side are at the same electron-energy level.

q. The most important property of the tunnel diode current-voltage chart is the region between the peak current (I_p) and the valley current (I_v), wherein current *decreases* with *increasing* voltage.

r. Decreasing current with increasing voltage represents an *ac negative resistance*, usable in electronic circuits to achieve gain.

s. The tunneltron, a *sandwich* of two conductors and an ultra-thin insulator, also displays the phenomenon of electron tunneling.

Chapter 3

AMPLIFICATION AND OSCILLATION USING AC NEGATIVE RESISTANCE

SECTION I. HISTORY AND DEVICES

3–1. General

Devices and circuit arrangements (other than tunnel diodes) displaying ac negative resistance and employed for amplification and oscillation were studied and applied as early as 1918. (See References, Appendix A.) One of the most popularly known devices is the thyratron, a gas-filled triode, which displays an *N*-shaped voltage-current characteristic. A circuit arrangement using an ordinary pentode electron tube (par. 3–2) and used as an oscillator also displays negative resistance; this circuit is called a *transitron*. A number of solid-state devices also display negative-resistance characteristics; these devices are discussed briefly in paragraph 3–3. None of these devices have found extensive use in electronics or offer promise of extensive use because not one has a combination of all the advantages of tunnel diodes (par. 1–4); namely, high-frequency oscillation, high-speed switching, simplicity of construction, and ease of manufacturing.

3–2. Transitron Oscillator

a. Figure 3–1A is a schematic diagram of a transitron oscillator. In this circuit, resistor *R*1 develops the ac suppressor voltage coupled from the screen grid through capacitor *C*2. Capacitor *C*2 is a blocking capacitor having negligible reactance at the oscillating frequency. Capacitor *C*1 and resistor *R*2 develop cathode bias. Coil *L*1 and capacitor *C*3 form a tank circuit and primarily determine the frequency of oscillation. Capacitor *C*4 couples the signal to the following stage.

b. The most important aspects of this circuit are:

1. The control grid is connected directly to the cathode. The cathode current, therefore, is a fixed quantity.