3
PART
Basic Electronics

This page intentionally left blank

19
CHAPTER

Introduction to Semiconductors

SINCE THE 1960S, WHEN THE TRANSISTOR BECAME COMMON IN CONSUMER DEVICES, *SEMICONDUCTORS* have acquired a dominating role in electronics. The term *semiconductor* arises from the ability of these materials to conduct some of the time, but not all the time. The conductivity can be controlled to produce effects such as amplification, rectification, oscillation, signal mixing, and switching.

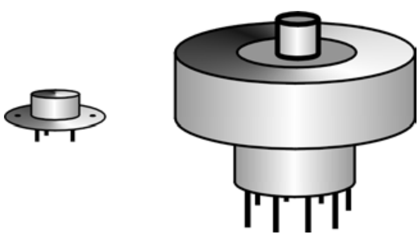
The Semiconductor Revolution

Decades ago, *vacuum tubes*, also known as *electron tubes*, were the only devices available for use as amplifiers, oscillators, detectors, and other electronic circuits and systems. A typical tube (called a *valve* in England) ranged from the size of your thumb to the size of your fist. They are still used in some power amplifiers, microwave oscillators, and video display units.

Tubes generally require high voltage. Even in modest radio receivers, 100 V to 200 V dc was required when tubes were employed. This mandated bulky power supplies, and created an electrical shock hazard. Nowadays, a transistor of microscopic dimensions can perform the functions of a tube in most situations. The power supply can be a couple of AA cells or a 9-V transistor battery.

Even in high-power applications, transistors are smaller and lighter than tubes. Figure 19-1 is a size comparison drawing between a transistor and a vacuum tube for use in an AF or RF power amplifier.

Integrated circuits (ICs), hardly larger than individual transistors, can do the work of hundreds or even thousands of vacuum tubes. An excellent example of this technology is found in personal computers and the peripheral devices used with them.



19-1 A power-amplifier transistor (at left) is much smaller than a vacuum tube of comparable power-handling capacity (right).

Semiconductor Materials

Various elements, compounds, and mixtures can function as semiconductors. The two most common materials are *silicon* and a compound of gallium and arsenic known as *gallium arsenide* (often abbreviated GaAs). In the early years of semiconductor technology, *germanium* formed the basis for many semiconductors; today it is seen occasionally, but not often. Other substances that work as semiconductors are *selenium*, *cadmium* compounds, *indium* compounds, and the oxides of certain metals.

Silicon

Silicon (chemical symbol Si) is widely used in diodes, transistors, and integrated circuits. Generally, other substances, or *impurities*, must be added to silicon to give it the desired properties. The best quality silicon is obtained by growing crystals in a laboratory. The silicon is then fabricated into *wafers* or *chips.*

Gallium Arsenide

Another common semiconductor is the compound gallium arsenide. Engineers and technicians call this material by its acronym-like chemical symbol, GaAs, pronounced "gas." If you hear about "gasfets" and "gas ICs," you're hearing about gallium-arsenide technology.

GaAs devices require little voltage, and will function at higher frequencies than silicon devices because the charge carriers move faster through the semiconductor material. GaAs devices are relatively immune to the effects of ionizing radiation such as X rays and gamma rays. GaAs is used in light-emitting diodes (LEDs), infrared-emitting diodes (IREDs), laser diodes, visible-light and infrared (IR) detectors, ultra-high-frequency (UHF) amplifying devices, and a variety of integrated circuits.

Selenium

Selenium exhibits conductivity that varies depending on the intensity of visible light or IR radiation that strikes it. All semiconductor materials exhibit this property, known as *photoconductivity*, to some degree; but in selenium the effect is especially pronounced. For this reason, selenium is useful for making *photocells.* Selenium is also used in certain types of *rectifiers.* A rectifier is a component or circuit that converts ac to pulsating dc.

A significant advantage of selenium is the fact that it is electrically rugged. Selenium-based components can withstand brief *transients*, or spikes, of abnormally high voltage, better than components made with most other semiconductor materials.

Germanium

Pure elemental germanium is a poor electrical conductor. It becomes a semiconductor only when *impurities* are added. Germanium was used extensively in the early years of semiconductor technology. Some diodes and transistors still use it.

A germanium diode has a low voltage drop (0.3 V, compared with 0.6 V for silicon and 1 V for selenium) when it conducts, and this makes it useful in some situations. But germanium is easily destroyed by heat. Extreme care must be used when soldering the leads of a germanium component.

Metal Oxides

Certain metal oxides have properties that make them useful in the manufacture of semiconductor devices. When you hear about MOS (pronounced "moss") or CMOS (pronounced "sea moss") technology, you are hearing about *metal-oxide semiconductor* and *complementary metal-oxide semiconductor* devices, respectively.

An advantage of MOS and CMOS devices is the fact that they need almost no power to function. They draw so little current that a battery in a MOS or CMOS device lasts just about as long as it would on the shelf. Another advantage is high speed. This allows operation at high frequencies in RF equipment, and makes it possible to perform many switching operations per second for use in computers.

Certain types of transistors, and many kinds of ICs, make use of this technology. In integrated circuits, MOS and CMOS allow for a large number of discrete diodes and transistors on a single chip. Engineers would say that MOS/CMOS has *high component density.*

The biggest problem with MOS and CMOS technology is the fact that the devices are easily damaged by static electricity. Care must be used when handling components of this type. Technicians working with MOS and CMOS components must literally ground themselves by wearing a metal wrist strap connected to a good earth ground. Otherwise, the electrostatic charges that normally build up on their bodies can destroy MOS and CMOS components when equipment is constructed or serviced.

Doping and Charge Carriers

For a semiconductor material to have the properties necessary in order to function as electronic components, impurities are usually added. The impurities cause the material to conduct currents in certain ways. The addition of an impurity to a semiconductor is called *doping.* Sometimes the impurity is called a *dopant.*

Donor Impurities

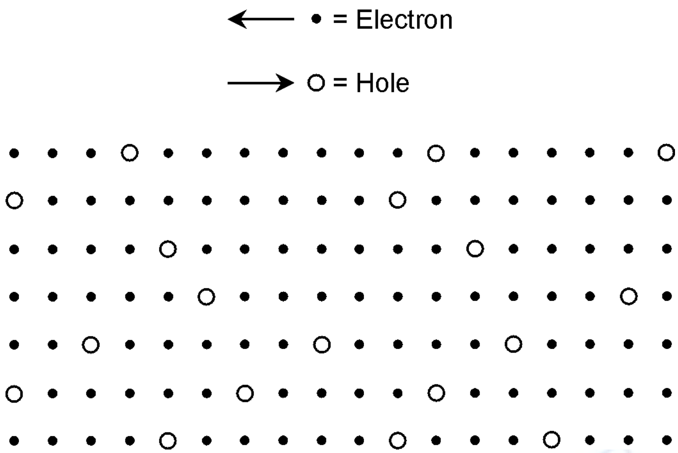
When an impurity contains an excess of electrons, the dopant is called a *donor impurity.* Adding such a substance causes conduction mainly by means of electron flow, as in an ordinary metal such as copper or aluminum. The excess electrons are passed from atom to atom when a voltage exists across the material. Elements that serve as donor impurities include antimony, arsenic, bismuth, and phosphorus. A material with a donor impurity is called an *N-type semiconductor*, because electrons have negative (N) charge.

Acceptor Impurities

If an impurity has a deficiency of electrons, the dopant is called an *acceptor impurity.* When a substance such as aluminum, boron, gallium, or indium is added to a semiconductor, the material conducts by means of *hole flow.* A *hole* is a missing electron—or more precisely, a place in an atom where an electron should be, but isn't. A semiconductor with an acceptor impurity is called a *P-type semiconductor*, because holes have, in effect, a positive (P) charge.

Majority and Minority Carriers

Charge carriers in semiconductor materials are either electrons, each of which has a unit negative charge, or holes, each of which has a unit positive charge. In any semiconductor substance, some



19-2 Pictorial representation of hole flow. Solid black dots represent electrons, moving in one direction. Open circles represent holes, moving in the opposite direction.

of the current takes the form of electrons passed from atom to atom in a negative-to-positive direction, and some of the current occurs as holes that move from atom to atom in a positive-to-negative direction.

Sometimes electrons account for most of the current in a semiconductor. This is the case if the material has donor impurities, that is, if it is of the N type. In other cases, holes account for most of the current. This happens when the material has acceptor impurities, and is thus of the P type. The dominating charge carriers (either electrons or holes) are called the *majority carriers.* The less abundant ones are called the *minority carriers.* The ratio of majority to minority carriers can vary, depending on the way in which the semiconductor material has been manufactured.

Figure 19-2 is a simplified illustration of electron flow versus hole flow in a sample of N-type semiconductor material, where the majority carriers are electrons and the minority carriers are holes. The solid black dots represent electrons. Imagine them moving from right to left in this illustration as they are passed from atom to atom. Small open circles represent holes. Imagine them moving from left to right in the illustration. In this particular example, the positive battery or power-supply terminal (or "source of holes") would be out of the picture toward the left, and the negative battery or power-supply terminal (or "source of electrons") would be out of the picture toward the right.

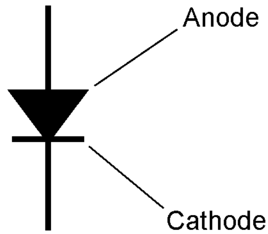
The P-N Junction

Merely connecting up a piece of semiconducting material, either P or N type, to a source of current can be interesting, and a good subject for science experiments. But when the two types of material are brought together, the boundary between them, called the *P-N junction*, behaves in ways that make semiconductor materials truly useful in electronic components.

The Semiconductor Diode

Figure 19-3 shows the schematic symbol for a *semiconductor diode*, formed by joining a piece of P-type material to a piece of N-type material. The N-type semiconductor is represented by the short, straight line in the symbol, and is called the *cathode.* The P-type semiconductor is represented by the arrow, and is called the *anode.*

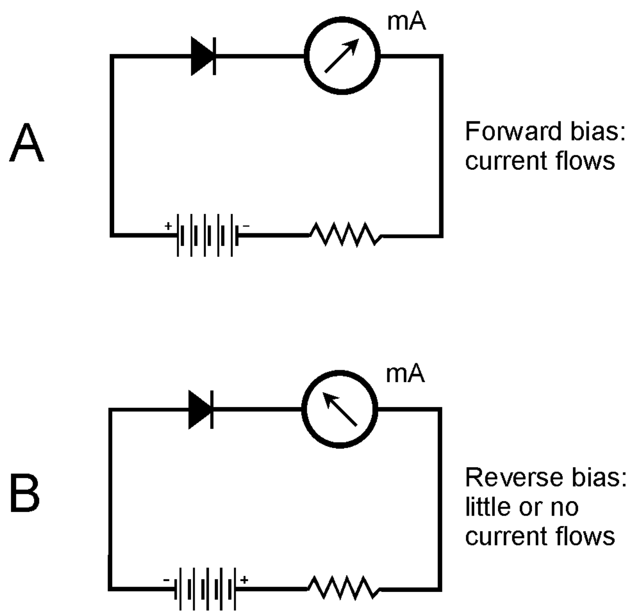
19-3 Schematic symbol for a semiconductor diode.



In the diode as shown in Figure 19-3, electrons can move easily in the direction opposite the arrow, and holes can move easily in the direction in which the arrow points. But current cannot, under most conditions, flow the other way. Electrons normally do not move with the arrow, and holes normally do not move against the arrow.

If you connect a battery and a resistor in series with the diode, you'll get a current to flow if the negative terminal of the battery is connected to the cathode and the positive terminal is connected to the anode, as shown in Fig. 19-4A. No current will flow if the battery is reversed, as shown in Fig. 19-4B. (The resistor is included in the circuit to prevent destruction of the diode by excessive current.)

It takes a specific, well-defined minimum applied voltage for conduction to occur through a semiconductor diode. This is called the *forward breakover voltage.* Depending on the type of material, the forward breakover voltage varies from about 0.3 V to 1 V. If the voltage across the junction is not at least as great as the forward breakover voltage, the diode will not conduct, even when it is connected as shown in Fig. 19-4A. This effect, known as the *forward breakover effect* or the *P-N junction threshold effect*, can be of use in circuits designed to limit the positive and/or negative peak voltages that signals can attain. The effect can also be used in a device called a *threshold detector*, in which a signal must be stronger than a certain amplitude in order to pass through.



19-4 Series connection of a battery, a resistor, a current meter, and a diode. At A, forward bias results in a flow of current. At B, reverse bias results in no current.

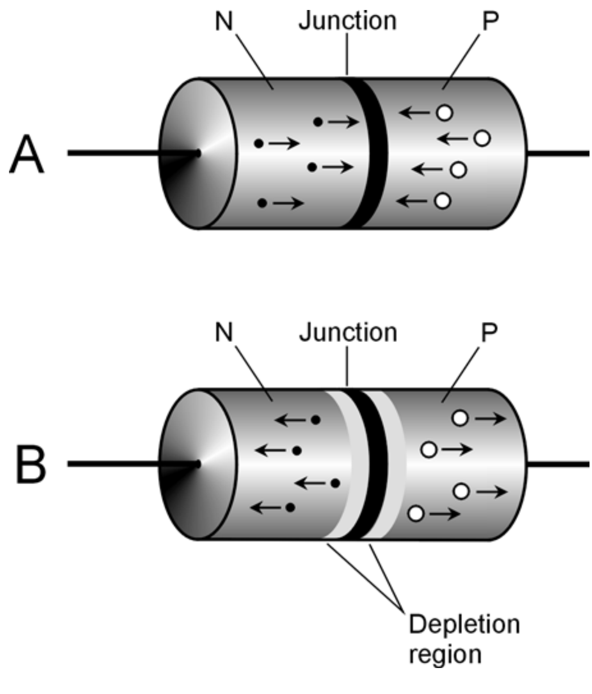
How the Junction Works

When the N-type material is negative with respect to the P type, as in Fig. 19-4A, electrons flow easily from N to P. The N-type semiconductor, which already has an excess of electrons, receives more; the P-type semiconductor, with a shortage of electrons, has some more taken away. The N-type material constantly feeds electrons to the P type in an attempt to create an electron balance, and the battery or power supply keeps robbing electrons from the P-type material. This condition is illustrated in Fig. 19-5A, and is known as *forward bias.* Current can flow through the diode easily under these circumstances.

When the battery or dc power-supply polarity is switched so the N-type material is positive with respect to the P type, the situation is called *reverse bias.* Electrons in the N-type material are pulled toward the positive charge pole, away from the P-N junction. In the P-type material, holes are pulled toward the negative charge pole, also away from the P-N junction. The electrons are the majority carriers in the N-type material, and the holes are the majority carriers in the P-type material. The charge therefore becomes depleted in the vicinity of the P-N junction, and on both sides of it, as shown in Fig. 19-5B. This zone, where majority carriers are deficient, is called the *depletion region.* A shortage of majority carriers in any semiconductor substance means that the substance cannot conduct well. Thus, the depletion region acts like an electrical insulator. This is why a semiconductor diode will not normally conduct when it is reverse-biased. A diode is, in effect, a one-way current gate—usually!

Junction Capacitance

Some P-N junctions can alternate between conduction (in forward bias) and nonconduction (in reverse bias) millions or billions of times per second. Other junctions are slower. The main limiting



19-5 At A, forward bias of a P-N junction. At B, reverse bias of the same junction. Solid black dots represent electrons. White dots represent holes. Arrows indicate direction of charge-carrier movement.

factor is the capacitance at the P-N junction during conditions of reverse bias. As the *junction capacitance* of a diode increases, maximum frequency at which it can alternate between the conducting state and the nonconducting state decreases.

The junction capacitance of a diode depends on several factors, including the operating voltage, the type of semiconductor material, and the cross-sectional area of the P-N junction. If you examine Fig. 19-5B, you might get the idea that the depletion region, sandwiched between two semiconducting sections, can play a role similar to that of the dielectric in a capacitor. This is true! In fact, a reverse-biased P-N junction actually is a capacitor. Some semiconductor components, called *varactor diodes*, are manufactured with this property specifically in mind.

The junction capacitance of a diode can be varied by changing the reverse-bias voltage, because this voltage affects the width of the depletion region. The greater the reverse voltage, the wider the depletion region gets, and the smaller the capacitance becomes.

Avalanche Effect

Sometimes, a diode conducts when it is reverse-biased. The greater the reverse-bias voltage, the more like an electrical insulator a P-N junction gets—up to a point. But if the reverse bias rises past a specific critical value, the voltage overcomes the ability of the junction to prevent the flow of current, and the junction conducts as if it were forward-biased. This phenomenon is called the *avalanche effect* because conduction occurs in a sudden and massive way, something like a snow avalanche on a mountainside.

The avalanche effect does not damage a P-N junction (unless the voltage is extreme). It's a temporary thing. When the voltage drops back below the critical value, the junction behaves normally again.

Some components are designed to take advantage of the avalanche effect. In other cases, the avalanche effect limits the performance of a circuit. In a device designed for voltage regulation, called a *Zener diode*, you'll hear about the *avalanche voltage* or *Zener voltage* specification. This can range from a couple of volts to well over 100 V. Zener diodes are often used in voltage-regulating circuits.

For *rectifier diodes* in power supplies, you'll hear or read about the *peak inverse voltage* (PIV) or *peak reverse voltage* (PRV) specification. It's important that rectifier diodes have PIV ratings great enough so that the avalanche effect will not occur (or even come close to happening) during any part of the ac cycle.

Quiz

Refer to the text in this chapter if necessary. A good score is at least 18 correct. Answers are in the back of the book.

1. The term *semiconductor* arises from
   (a) resistor-like properties of metal oxides.
   (b) variable conductive properties of some materials.
   (c) the fact that electrons conduct better than holes.
   (d) insulating properties of silicon and GaAs.

2. Which of the following is *not* an advantage of semiconductor devices over vacuum tubes?
   (a) Smaller size
   (b) Lower working voltage

   (c) Lighter weight
   (d) Ability to withstand high voltage spikes

3. Of the following substances, which is the most commonly used semiconductor?
   (a) Germanium
   (b) Galena
   (c) Silicon
   (d) Copper

4. GaAs is
   (a) a compound.
   (b) an element.
   (c) a mixture.
   (d) a gas.

5. A disadvantage of MOS devices is the fact that
   (a) the charge carriers move fast.
   (b) the material does not react to ionizing radiation.
   (c) they can be damaged by electrostatic discharges.
   (d) they must always be used at high frequencies.

6. Selenium works especially well in
   (a) photocells.
   (b) high-frequency detectors.
   (c) RF power amplifiers.
   (d) voltage regulators.

7. Of the following, which material allows the lowest forward voltage drop in a diode?
   (a) Selenium
   (b) Silicon
   (c) Copper
   (d) Germanium

8. A CMOS integrated circuit
   (a) can only work at low frequencies.
   (b) requires very little power to function.
   (c) requires considerable power to function.
   (d) can only work at high frequencies.

9. The purpose of doping is to
   (a) make the charge carriers move faster.
   (b) cause holes to flow.

(c) give a semiconductor material specific properties.
(d) protect devices from damage in case of transients.

10. A semiconductor material is made into N type by
    (a) adding an acceptor impurity.
    (b) adding a donor impurity.
    (c) injecting protons.
    (d) taking neutrons away.
11. Which of the following does not result from adding an acceptor impurity?
    (a) The material becomes P type.
    (b) Current flows mainly in the form of holes.
    (c) Most of the carriers have positive electric charge.
    (d) The substance acquires an electron surplus.
12. In a P-type material, electrons are
    (a) the majority carriers.
    (b) the minority carriers.
    (c) positively charged.
    (d) entirely absent.
13. Holes move from
    (a) minus to plus.
    (b) plus to minus.
    (c) P-type to N-type material.
    (d) N-type to P-type material.
14. When a P-N junction does not conduct even though a voltage is applied, the junction is
    (a) reverse-biased at a voltage less than the avalanche voltage.
    (b) overdriven.
    (c) biased past the breaker voltage.
    (d) in a state of avalanche effect.
15. Holes flow the opposite way from electrons because
    (a) charge carriers flow continuously.
    (b) they have opposite electric charge.
    (c) they have the same electric charge.
    (d) Forget it! Holes flow in the same direction as electrons.
16. If an electron is considered to have a charge of −1 unit, then a hole can be considered to have
    (a) a charge of −1 unit.
    (b) no charge.

(c) a charge of +1 unit.
(d) a charge that depends on the semiconductor type.

17. When a P-N junction is forward-biased, conduction will not occur unless
    (a) the applied voltage exceeds the forward breakover voltage.
    (b) the applied voltage is less than the forward breakover voltage.
    (c) the junction capacitance is high enough.
    (d) the depletion region is wide enough.
18. If the reverse bias exceeds the avalanche voltage in a P-N junction,
    (a) the junction will be destroyed.
    (b) the junction will insulate; no current will flow.
    (c) the junction will conduct current.
    (d) the capacitance will become extremely low.
19. Avalanche voltage is routinely exceeded when a P-N junction acts as a
    (a) current rectifier.
    (b) variable resistor.
    (c) variable capacitor.
    (d) voltage regulator.
20. Which of the following does not affect the junction capacitance of a diode?
    (a) the cross-sectional area of the P-N junction
    (b) the width of the depletion region
    (c) the phase of an applied ac signal
    (d) the reverse-bias voltage

20
CHAPTER

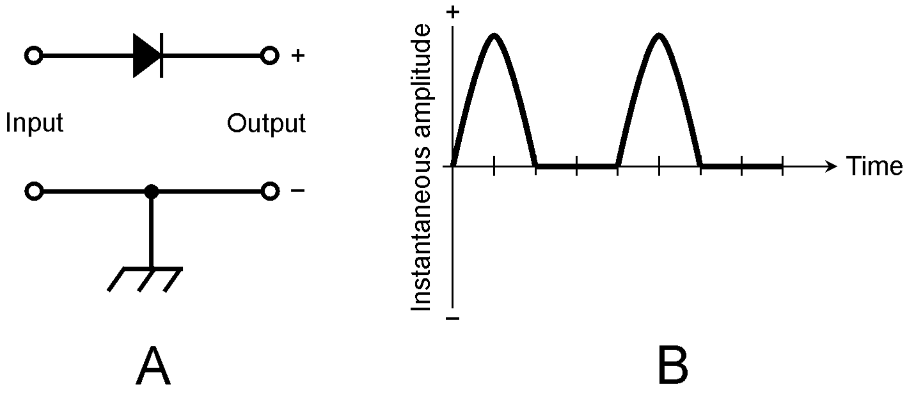
How Diodes Are Used

IN THE EARLY YEARS OF ELECTRONICS, NEARLY ALL DIODES WERE VACUUM TUBES. TODAY, MOST ARE made from semiconductors. Contemporary diodes can do almost everything that the old ones could, and also some things that people in the tube era could only dream about.

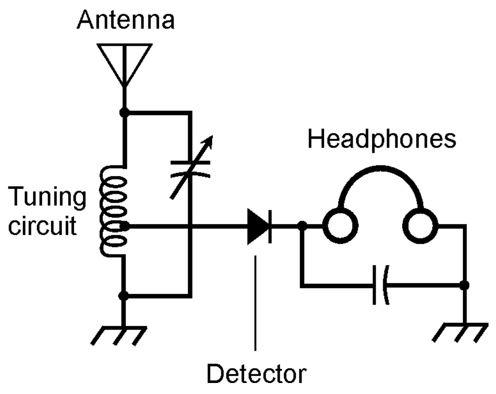
Rectification

The hallmark of a *rectifier diode* is that it passes current in only one direction. This makes it useful for changing ac to dc. Generally speaking, when the cathode is negative with respect to the anode, current flows; when the cathode is positive relative to the anode, there is no current. The constraints on this behavior are the forward breakover and avalanche voltages, as you learned about in Chap. 19.

Examine the circuit shown at A in Fig. 20-1. Suppose a 60-Hz ac sine wave is applied to the input. During half the cycle, the diode conducts, and during the other half, it doesn't. This cuts off half of every cycle. Depending on which way the diode is hooked up, either the positive half or the negative half of the ac cycle will be removed. Drawing B in Fig. 20-1 shows a graph of the output



20-1 At A, a half-wave rectifier circuit. At B, the output of the circuit shown at A when an ac sine wave is applied to the input.



20-2 Schematic diagram of a crystal-set radio receiver.

of the circuit at A. Remember that electrons flow from negative to positive, against the arrow in the diode symbol.

The circuit and wave diagram of Fig. 20-1 show a *half-wave rectifier* circuit. This is the simplest possible rectifier. That's its chief advantage over other, more complicated rectifier circuits. You'll learn about the various types of rectifier diodes and circuits in Chap. 21.

Detection

One of the earliest diodes, existing even before vacuum tubes, was actually a primitive semiconductor device. Known as a *cat whisker*, it consisted of a fine piece of wire in contact with a small piece of the mineral *galena.* This strange-looking contraption had the ability to act as a rectifier for extremely weak RF currents. When the cat whisker was connected in a circuit such as the one shown in Fig. 20-2, the result was a device capable of picking up amplitude-modulated (AM) radio signals and producing audio output that could be heard in the headset.

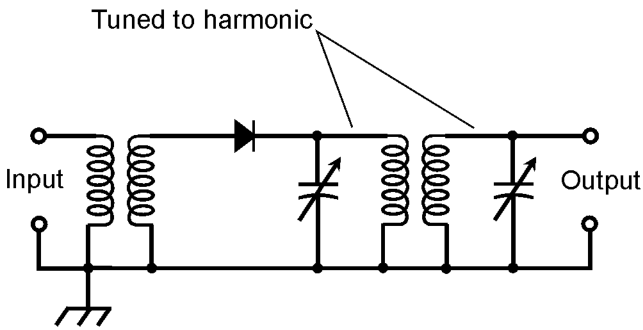
The galena, sometimes called a "crystal," gave rise to the nickname *crystal set* for this primitive radio receiver. You can still build a crystal set today, using a simple RF diode, a coil, a tuning capacitor, a headset, and a long-wire antenna. Notice that there's no battery! The audio is provided by the received signal alone.

The diode in Fig. 20-2 acts to recover the audio from the radio signal. This process is called *detection*; the circuit is called a *detector* or *demodulator.* If the detector is to be effective, the diode must be of the proper type. It must have low junction capacitance, so that it can work as a rectifier (and not as a capacitor) at radio frequencies. Some modern RF diodes are microscopic versions of the old cat whisker, enclosed in a glass case with axial leads.

Frequency Multiplication

When current passes through a diode, half of the cycle is cut off, as shown in Fig. 20-1B. This occurs no matter what the frequency, from 60-Hz utility current through RF, as long as the diode capacitance is not too great.

The output wave from the diode looks much different than the input wave. This condition is known as *nonlinearity.* Whenever there is nonlinearity of any kind in a circuit—that is, whenever the output waveform is shaped differently from the input waveform—there are harmonics in the output. These are waves at integer multiples of the input frequency. (If you've forgotten what harmonics are, refer to Chap. 9.)



20-3 A frequency-multiplier circuit using a semiconductor diode.

Often, nonlinearity is undesirable. Then engineers strive to make the circuit *linear*, so the output waveform has exactly the same shape as the input waveform. But sometimes harmonics are desired. Then nonlinearity is introduced deliberately to produce *frequency multiplication.* Diodes are ideal for this purpose. A simple frequency-multiplier circuit is shown in Fig. 20-3. The output *LC* circuit is tuned to the desired *n*th harmonic frequency, nf_o, rather than to the input or fundamental frequency, f_o.

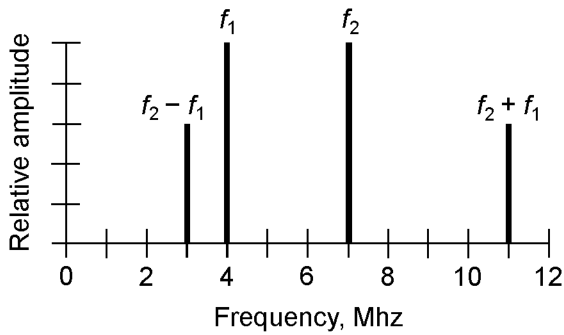
For a diode to work as a frequency multiplier, it must be of a type that would also work well as a detector at the same frequencies. This means that the component should act like a rectifier, but not like a capacitor.

Signal Mixing

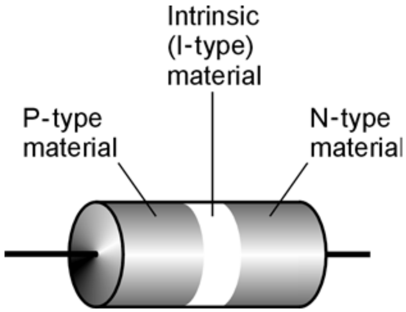
When two waves having different frequencies are combined in a nonlinear circuit, new waves are produced at frequencies equal to the sum and difference of the frequencies of the input waves. Diodes can provide this nonlinearity.

Suppose there are two signals with frequencies f_1 and f_2. For mathematical convenience, let's assign f_2 to the wave with the higher frequency, and f_1 to the wave with the lower frequency. If these signals are combined in a nonlinear circuit, new waves result. One of them has a frequency of $f_2 + f_1$, and the other has a frequency of $f_2 - f_1$. These sum and difference frequencies are known as *beat frequencies.* The signals themselves are called *mixing products* or *heterodynes* (Fig. 20-4).

Figure 20-4, incidentally, is an illustration of a *frequency domain* display. The amplitude (on the vertical scale or axis) is shown as a function of the frequency (on the horizontal scale or axis). This sort of display is what engineers see when they look at the screen of a lab instrument known as a *spectrum analyzer.* In contrast, an ordinary oscilloscope displays amplitude (on the vertical scale or axis) as a function of time (on the horizontal scale or axis). The oscilloscope provides a *time domain* display.



20-4 Spectral (frequency-domain) illustration of signal mixing.



20-5 The PIN diode has a layer of intrinsic (I type) semiconductor material at the P-N junction.

Switching

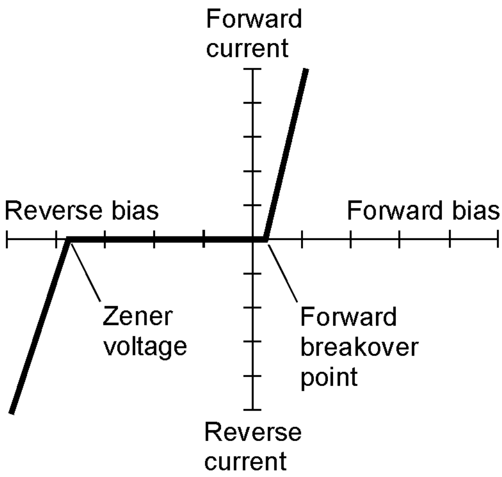
The ability of diodes to conduct with forward bias, and to insulate with reverse bias, makes them useful for switching in some electronic applications. Diodes can perform switching operations much faster than any mechanical device.

One type of diode, made for use as an RF switch, has a special semiconductor layer sandwiched in between the P-type and N-type material. The material in this layer is called an *intrinsic* (or *I-type*) *semiconductor.* The *intrinsic layer* (or *I layer*) reduces the capacitance of the diode, so that it can work at higher frequencies than an ordinary diode. A diode with an I-type semiconductor layer sandwiched in between the P- and N-type layers is called a *PIN diode* (Fig. 20-5).

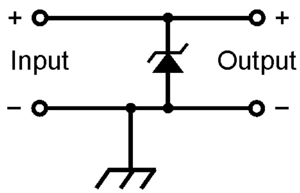
Direct-current bias, applied to one or more PIN diodes, allows RF currents to be effectively channeled without using relays and cables. A PIN diode also makes a good RF detector, especially at very high frequencies.

Voltage Regulation

Most diodes have an avalanche breakdown voltage that is much higher than the reverse bias ever gets. The value of the avalanche voltage depends on how a diode is manufactured. *Zener diodes* are specially made so they exhibit well-defined, constant avalanche voltages.



20-6 Current through a Zener diode as a function of the bias voltage.



20-7 Connection of a Zener diode for voltage regulation.

Suppose a certain Zener diode has an avalanche voltage, also called the *Zener voltage*, of 50 V. If reverse bias is applied to the P-N junction, the diode acts as an open circuit as long as the bias is less than 50 V. But if the reverse-bias voltage reaches 50 V—even for a brief instant of time—the diode conducts. This effectively prevents the reverse-bias voltage from exceeding 50 V.

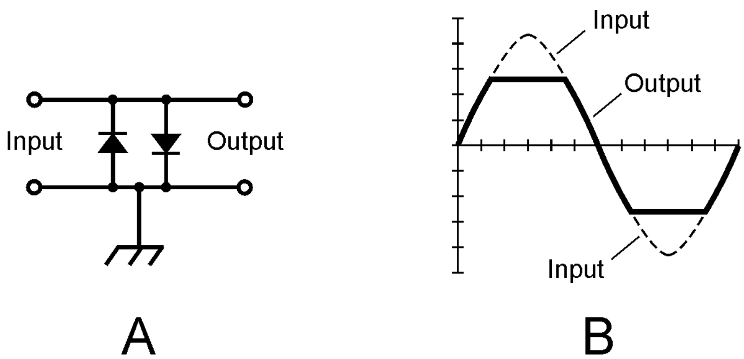
The current through a Zener diode, as a function of the voltage, is shown in Fig. 20-6. The Zener voltage is indicated by the abrupt rise in reverse current as the reverse-bias voltage increases. A simple Zener-diode voltage-limiting circuit is shown in Fig. 20-7. Note the polarity of the diode: the cathode is connected to the positive pole, and the anode is connected to the negative pole.

Amplitude Limiting

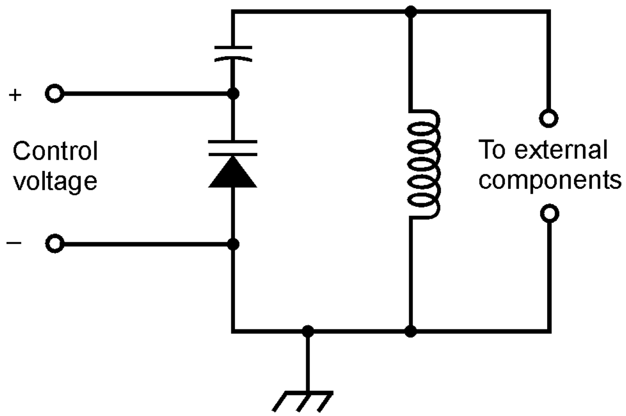
In Chap. 19, you learned that a diode will not conduct until the forward-bias voltage is at least as great as the forward breakover voltage. There's a corollary to this: a diode will always conduct when the forward-bias voltage reaches or exceeds the forward breakover voltage, when the device is conducting current in the forward direction. In the case of silicon diodes this is approximately 0.6 V. For germanium diodes it is about 0.3 V, and for selenium diodes it is about 1 V.

This phenomenon can be used to advantage when it is necessary to limit the amplitude of a signal, as shown in Fig. 20-8. By connecting two identical diodes back-to-back in parallel with the signal path (A), the maximum peak amplitude is limited, or *clipped*, to the forward breakover voltage of the diodes. The input and output waveforms of a clipped signal are illustrated at B. This scheme is sometimes used in radio receivers to prevent "blasting" when a strong signal comes in.

The downside of the *diode limiter* circuit, such as the one shown in Fig. 20-8, is the fact that it introduces distortion when clipping occurs. This might not be a problem for reception of digital signals, for frequency-modulated signals, or for analog signals that rarely reach the limiting voltage. But for amplitude-modulated signals with peaks that rise past the limiting voltage, it can cause trouble.



20-8 At A, connection of two diodes to act as an ac limiter. At B, illustration of sine-wave peaks cut off by the action of the diodes in an ac limiter.



20-9 Connection of a varactor diode in a tuned circuit.

Frequency Control

When a diode is reverse-biased, there is a region at the P-N junction with dielectric (insulating) properties. As you know from Chap. 19, this is called the depletion region, because it has a shortage of majority charge carriers. The width of this zone depends on several things, including the reverse-bias voltage.

As long as the reverse bias is less than the avalanche voltage, varying the bias affects the width of the depletion region. This in turn varies the junction capacitance. This capacitance, which is always small (on the order of picofarads), varies inversely with the square root of the reverse-bias voltage, as long as the reverse bias remains less than the avalanche voltage. Thus, for example, if the reverse-bias voltage is quadrupled, the junction capacitance drops to one-half; if the reverse-bias voltage is decreased by a factor of 9, then the junction capacitance increases by a factor of 3.

Some diodes are manufactured especially for use as variable capacitors. Such a device is known as varactor diode, as you learned in Chap. 19. Varactors are used in a special type of circuit called a *voltage-controlled oscillator* (VCO). Figure 20-9 is a simple example of the *LC* circuit in a VCO, using a coil, a fixed capacitor, and a varactor. This is a parallel-tuned circuit. The fixed capacitor, whose value is large compared with that of the varactor, serves to keep the coil from short-circuiting the control voltage across the varactor. Notice that the symbol for the varactor has two lines on the cathode side.

Oscillation and Amplification

Under certain conditions, diodes can be made to produce microwave RF signals. Three types of diodes that can do this are *Gunn diodes*, *IMPATT diodes*, and *tunnel diodes.*

Gunn Diodes

A Gunn diode can produce up to 1 W of RF power output, but more commonly it works at levels of about 0.1 W. Gunn diodes are usually made from gallium arsenide. A Gunn diode oscillates because of the *Gunn effect*, named after J. Gunn of International Business Machines (IBM), who first observed it in the 1960s. A Gunn diode doesn't work like a rectifier, detector, or mixer. Instead, the oscillation takes place as a result of a quirk called *negative resistance.*

Gunn-diode oscillators are often tuned using varactor diodes. A Gunn-diode oscillator, connected directly to a microwave horn antenna, is known as a *Gunnplexer.* These devices are popular with amateur-radio experimenters at frequencies of 10 GHz and above.

IMPATT Diodes

The acronym *IMPATT* comes from the words *imp*act *a*valanche *t*ransit *t*ime. This, like negative resistance, is a rather esoteric phenomenon. An *IMPATT diode* is a microwave oscillating device like a Gunn diode, except that it uses silicon rather than gallium arsenide.

An IMPATT diode can be used as an amplifier for a microwave transmitter that employs a Gunn-diode oscillator. As an oscillator, an IMPATT diode produces about the same amount of output power, at comparable frequencies, as a Gunn diode.

Tunnel Diodes

Another type of diode that will oscillate at microwave frequencies is the *tunnel diode*, also known as the *Esaki diode.* It produces enough power so it can be used as a local oscillator in a microwave radio receiver, but not much more.

Tunnel diodes work well as amplifiers in microwave receivers, because they generate very little unwanted noise. This is especially true of gallium arsenide devices.

Energy Emission

Some semiconductor diodes emit radiant energy when a current passes through the P-N junction in a forward direction. This phenomenon occurs as electrons fall from higher to lower energy states within atoms.

LEDs and IREDs

Depending on the exact mixture of semiconductors used in manufacture, visible light of almost any color can be produced by diodes when bias is applied to them in the forward direction. Infrared-emitting devices also exist. The most common color for a *light-emitting diode* (LED) is bright red. An *infrared-emitting diode* (IRED) produces energy at wavelengths slightly longer than those of visible red light.

The intensity of the radiant energy from an LED or IRED depends to some extent on the forward current. As the current rises, the brightness increases, but only up to a certain point. If the current continues to rise, no further increase in brilliance takes place. The LED or IRED is then said to be in a state of *saturation.*

Digital Displays

Because LEDs can be made in various different shapes and sizes, they are ideal for use in digital displays. You've seen digital clock radios that use them. They are common in car radios. They make good indicators for "on/off," "a.m./p.m.," "battery low," and other conditions.

In recent years, LED displays have been largely replaced by *liquid crystal displays* (LCDs). The LCD technology has advantages over LED technology, including lower power consumption and better visibility in direct sunlight. However, LCDs require backlighting when the ambient illumination is low.

Communications

Both LEDs and IREDs are useful in communications because their intensity can be modulated to carry information. When the current through the device is sufficient to produce output, but not enough to cause saturation, the LED or IRED output follows along with rapid current changes. Analog and digital signals can be conveyed over light beams in this way. Some modern telephone systems make use of modulated light, transmitted through clear fibers. This is known as *fiber-optic* technology.

Special LEDs and IREDs produce *coherent radiation.* These are called *laser diodes.* The rays from these diodes aren't the intense, parallel beams that most people imagine when they think about lasers. A laser LED or IRED generates a cone-shaped beam of low intensity. But it can be focused into a parallel beam, and the resulting rays have some of the same advantages found in larger lasers, including the ability to travel long distances with little decrease in their intensity.

Photosensitive Diodes

Virtually all P-N junctions exhibit conductivity that varies with exposure to radiant electromagnetic energy such as IR, visible light, and UV. The reason that conventional diodes are not affected by these rays is that they are enclosed in opaque packages. Some *photosensitive diodes* have variable dc resistance that depends on the intensity of the electromagnetic rays. Other types of diodes produce their own dc in the presence of radiant energy.

Silicon Photodiodes

A silicon diode, housed in a transparent case and constructed in such a way that visible light can strike the barrier between the P-type and N-type materials, forms a *silicon photodiode.* A reverse-bias voltage is applied to the device. When radiant energy strikes the junction, current flows. The current is proportional to the intensity of the radiant energy, within certain limits.

Silicon photodiodes are more sensitive at some wavelengths than at others. The greatest sensitivity is in the *near infrared* part of the spectrum, at wavelengths just a little bit longer than the wavelength of visible red light.

When radiant energy of variable intensity strikes the P-N junction of a reverse-biased silicon photodiode, the output current follows the light-intensity variations. This makes silicon photodiodes useful for receiving modulated-light signals of the kind used in fiber-optic communications systems.

The Optoisolator

An LED or IRED and a photodiode can be combined in a single package to get a component called an *optoisolator.* This device, the schematic symbol for which is shown in Fig. 20-10, creates a modulated-light signal and sends it over a small, clear gap to a receptor. An LED or IRED converts an electrical signal to visible light or IR; a photodiode changes the visible light or infrared back into an electrical signal.

When a signal is electrically coupled from one circuit to another, the two stages interact. The input impedance of a given stage, such as an amplifier, can affect the behavior of the circuits that feed power to it. This can lead to various sorts of trouble. Optoisolators overcome this effect, because the coupling is not done electrically. If the input impedance of the second circuit changes, the impedance that the first circuit sees is not affected, because it is simply the impedance of the LED