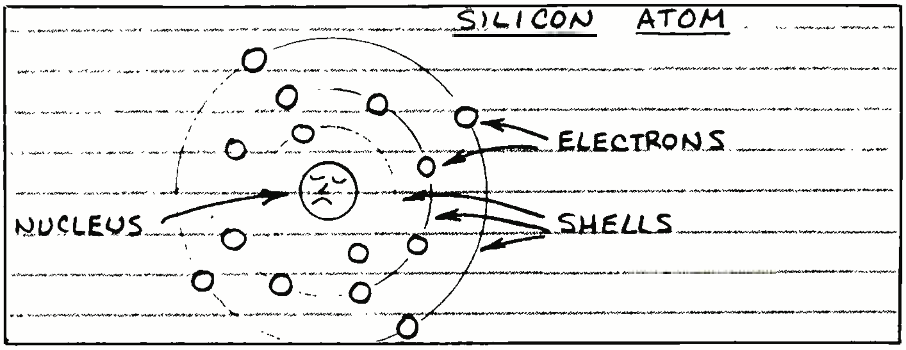
How To Get Started In Electronics

Part I: Inside Semiconductors

By Forrest M. Mims III



Electronics experimenting is rewarding, and challenging. There are few other pursuits that are more satisfying, especially when you finish building a project, turn it on, and find you've created from a meaningless pile of components something really useful. Thus, you've become a creator with an end-product to show for your efforts.

Electronics experimenters often have different objectives. You might wish to simply convert an idea for a circuit into a finished project just for the fun of seeing it take shape and operate as you designed it to do. Or you might have a need for a device that isn't available commercially, or you might want to make something that is available commercially but you feel you can build at lower cost.

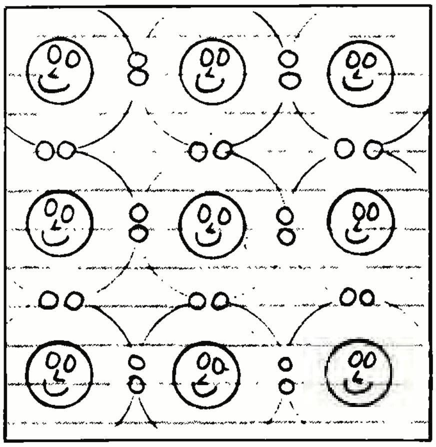
Whatever your reasons for wanting to build projects or experiment with your own circuit ideas, you'll need some idea of how to use the various elements available to create practical circuits. To be able to do this, you must understand how these devices work and how they interact.

As the title of this series implies, the objective of this series is to introduce you to these elements starting with semiconductors.

Silicon Recipes

There are many different semiconducting materials, but silicon, the main ingredient of sand, is the most popular ingredient.

A silicon atom has only four electrons in its outermost shell, but it would like to have eight. Therefore, a silicon atom will link up with four of its neighbors to *share* electrons as follows:



A cluster of silicon atoms sharing outer electrons forms a regular arrangement called a crystal. To keep things simple, in the above drawing of a silicon crystal, only the outer electrons are shown.

Pure silicon isn't very useful. That's why silicon makers spice up their silicon recipes with a dash of phosphorous, boron, or other goodies. This is called "doping" the silicon. Then silicon has very useful electrical properties indeed!

P and N Spiced Loaf. Boron, phosphorous, and certain other atoms can join with silicon atoms to form crystals. Here's the catch: A boron atom has only *three* electrons in its outer shell, while a phosphorus atoms has *five* electrons. Silicon with extra phosphorus electrons is called *n*-type silicon (n = negative). Silicon with electron-deficient boron atoms is called *p*-type (p = positive) silicon.

P-type Silicon. A boron atom in a cluster of silicon atoms leaves a vacant electron opening called a "hole." It's possible for an electron from a nearby atom to "fall" into the hole. Therefore, the hole has *moved* to a new location. Remember, holes can move through silicon (just as bubbles move through water).

N-Type Silicon. A phosphorus atom in a cluster of silicon atoms donates an extra electron that can move through the crystal with comparative ease. In other words, *n*-type silicon can carry an electrical current. But so can *p*-type silicon! Holes are what "carry" the current.

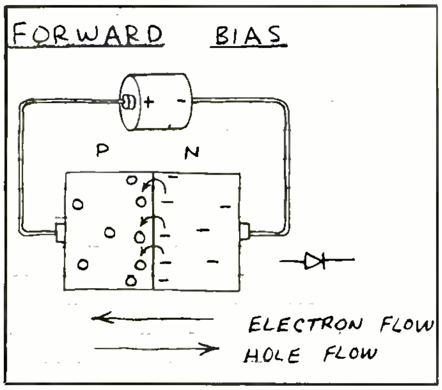
The Diode

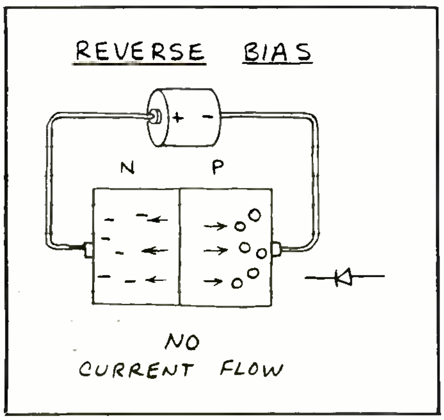
Both *n*- and *p*-type silicon conduct electricity. The resistance of both types is determined by the proportion of holes or surplus electrons. Therefore, both types can function as resistors. And they will conduct electricity in any direction.

By forming some *p*-type silicon in a chip of *n*-type silicon, electrons will flow through the silicon in only one direction. This is the principle of the *diode*. The *p-n* interface is called the *pn junction.*

How the Diode Works

Here's a simplified explanation of how a diode conducts electricity in one direction (forward) while blocking the flow of current in the opposite direction (reverse).

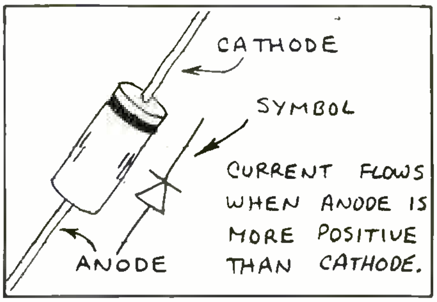




In A, the charge from the battery *repels* holes and electrons toward the junction. If the voltage exceeds 0.6 volt (silicon), electrons will cross the junction and combine with holes, allowing a current to flow.

In B, the charge from the battery *attracts* holes and electrons away from the junction. Under this condition, no current can flow.

Diodes are commonly enclosed in small glass cylinders. A dark band around one end of the cylinder identi-



fies the *cathode* terminal. The opposite terminal is the anode. The above drawing illustrates both the physical packaging of the diode and its schematic symbol.

You already know from the foregoing discussion that a diode is like an electronic one-way valve. It's important to understand some additional aspects of diode operation. Here are some key ones:

1. A diode will *not* conduct until the forward voltage reaches a certain threshold point. For silicon diodes, this is about 0.6 volt.

2. If the forward current becomes excessive, the semiconductor chip may crack or melt! And the contacts may separate. If the chip melts, the diode may suddenly conduct in both directions. The resulting heat may vaporize the chip!

3. Too much reverse voltage will cause a diode to conduct in the wrong direction. Since this voltage is fairly high, the sudden current surge may zap the diode.

Diode Types & Uses

Many different kinds of diodes are available. Here are some of the major types in common usage:

Small Signal. These diodes transform low-current ac to dc, detect (demodulate) radio signals, multiply voltage, perform logic operation, absorb voltage spikes, etc.

Power Rectifier. Functionally identical to small-signal diodes, power rectifiers can handle much more current. They are installed in large metal packages that soak up excess heat and transfer it to a metal heat sink. They are used mainly in power supplies.

Zener Diode. This specialty rectifier is designed to have a specific *reverse breakdown* (conduction) *voltage.* This means zener diodes can function like a voltage-sensitive switch. Zener diodes having breakdown voltages (V_Z of from 2 to 200 volts are available.

Light-Emitting Diode. All diodes emit some electromagnetic radiation when forward baised. Diodes made from certain semiconductors (like gallium arsenide phosphide) emit *considerably* more radiation than silicon diodes. They're called light-emitting diodes (LEDs).

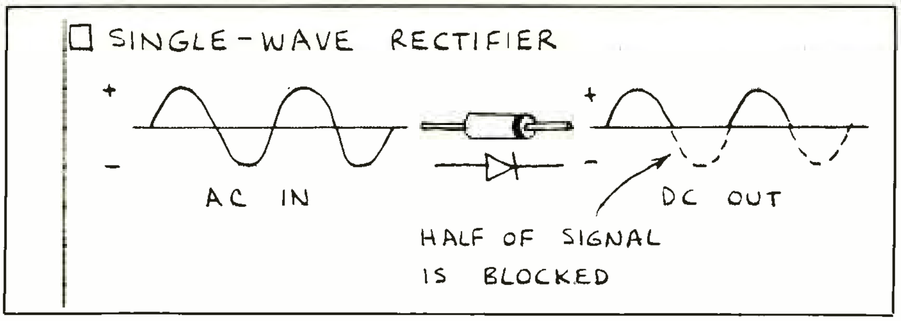
Photodiode. All diodes also *respond* to some degree when illuminated by light. Diodes designed specifically to detect light are called photodiodes. They include a glass or plastic window through which the light enters. Often they have a large, exposed junction region. Silicon makes good photodiodes.

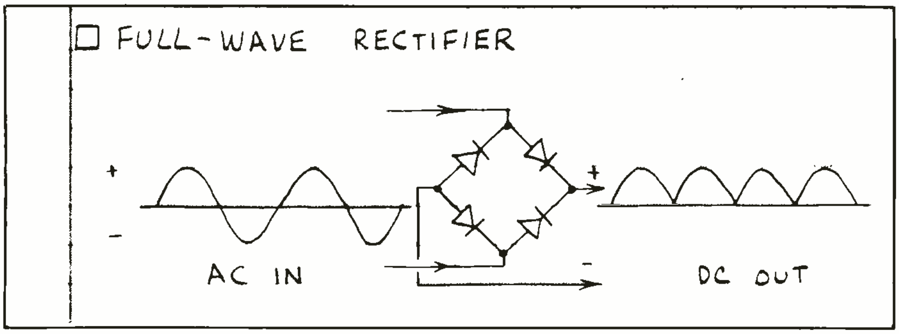
Various types of diodes are used in many applications. For now, here are two of the most important roles for small-signal diodes and rectifiers:

The first is a single-wave rectifier circuit in which an undulating (ac) signal or voltage is rectified into a single-polarity (dc) signal or voltage.

The second is the full-wave rectifier circuit in which a four-diode "network" (or bridge rectifier) rectifies *both* halves of an ac signal.

Note the difference in output from the two circuits. In the half-wave circuit only one alternation of the input generates an output, while *both* alter-





ing through the other two leads. This means transistors can be used as amplifiers and switches. There are two main families of transistors: bipolar and field-effect. We will concentrate here only on bipolar transistors.

Add a second junction to a *pn*-junction diode and you get a three-layer silicon "sandwich." The sandwich can be either *npn* or *pnp*. Either way, the middle layer acts like a faucet or gate that controls the amount of current moving through the three layers.

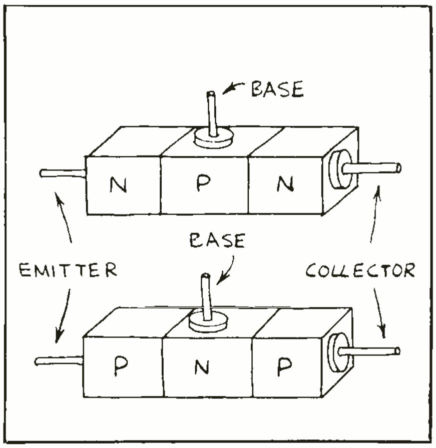
The three layers of a bipolar transistor are the emitter, base, and collector. The base is very thin and has fewer doping atoms than the emitter and collector. Therefore, a very small emitter-base current will cause a much larger emitter-collector current to flow.

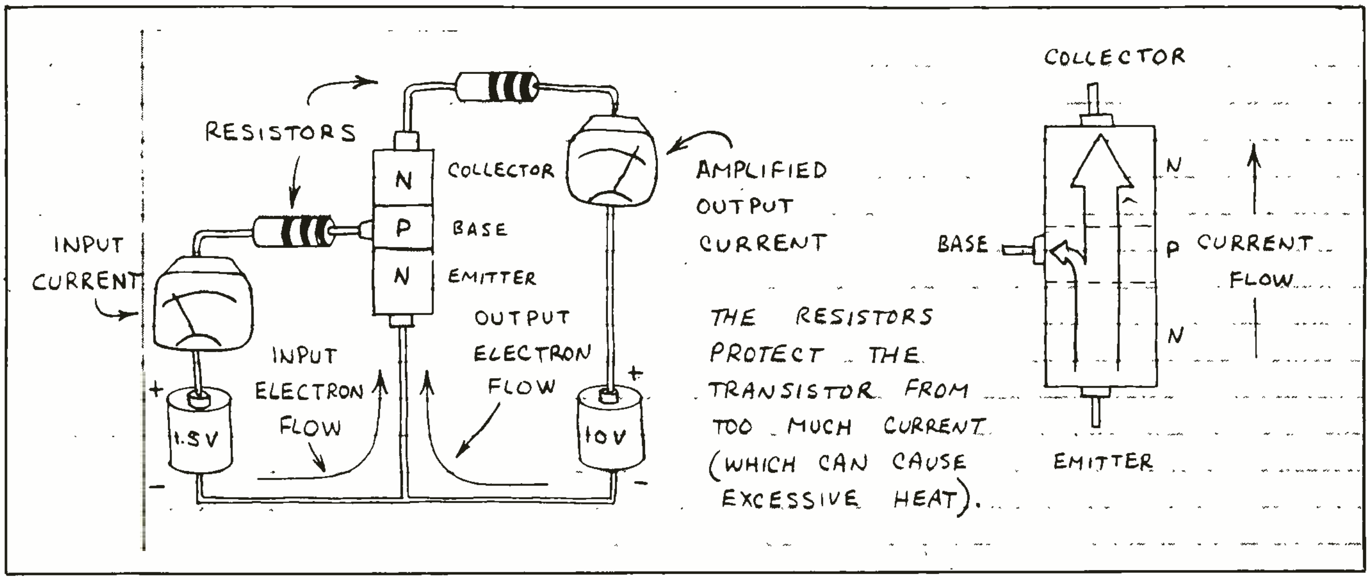
Diodes and transistors share several key features: The base-emitter junction (or diode) will not conduct until the forward voltage exceeds 0.6 volt. Too much current will cause a transistor to become hot and operate improperly. If a transistor is hot (not just warm) when touched, disconnect the power to it! Finally, too much current or voltage may damage or permanently destroy the semiconductor chip that forms the transistor. Even if the chip itself isn't harmed, its tiny connection wires may melt or

nations produce outputs in the full-wave circuit. This difference offers significant advantages, in the case of the full-wave circuit, when the output must be filtered to obtain a smooth (minimal-variation) dc voltage with minimum ripple.

The Transistor

Transistors are semiconductor devices with three leads. A very small current or voltage at one lead can control a much larger current flow-





separate from the chip. *Never* connect a transistor backwards!

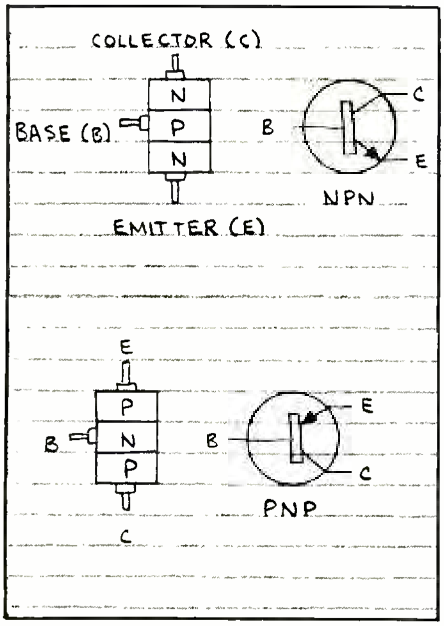
Many different types of transistors are available. Typical examples are as follows:

Small-Signal and Switching. Small-signal transistors are used to amplify low-level signals. Switching transistors are designed to be operated fully on or off. Some transistors can both amplify *and* switch equally well.

Power. These transistors are used in high-power amplifiers and power supplies. Large size and exposed metal surfaces keep them cool.

High-Frequency. These transistors operate at radio, television, and microwave frequencies. The base region is *very* thin, and the actual chip is very small.

The schematic symbols for the *npn* and *pnp* transistors is as follows:



Note that the arrow in the emitter leads always points in the direction of hole flow. If the arrow points away from the base, the emitter is made of *n*-type material, which identifies it as an *npn* transistor. Conversely, if the arrow points toward the base, the base is *n*-type material, which means the emitter and collector are both *p*-type material and, thus, the transistor is a *pnp* device.

Using Bipolar Transistors

When the base of an *npn* transistor is grounded (0 volt), no current flows from the emitter to the collector and the transistor is "off." If the base is forward-biased by at least 0.6 volt, a current will flow from the emitter to the collector and the transistor is "on." When operated in only these two modes, the transistor functions as a switch. If the base is forward-biased, the emitter-collector current will follow variations in a much smaller base current. The transistor then functions as an amplifier.

This discussion applies to a transistor in which the emitter is the ground connection for both the input and the output, and is called the *common-emitter* circuit. Some simlified common-emitter circuits are show below.

•**Transistor Switch** (A)

Only two inputs are possible with this circuit: ground (0 volt) and positive battery voltage (+V). Therefore, the transistor is off or on. A typical base resistance is 5000 to 10,000 ohms. If the resistor is replaced by a wire, the lamp can be switched on and off from a considerable distance.

•**Transistor dc Amplifier** (B)

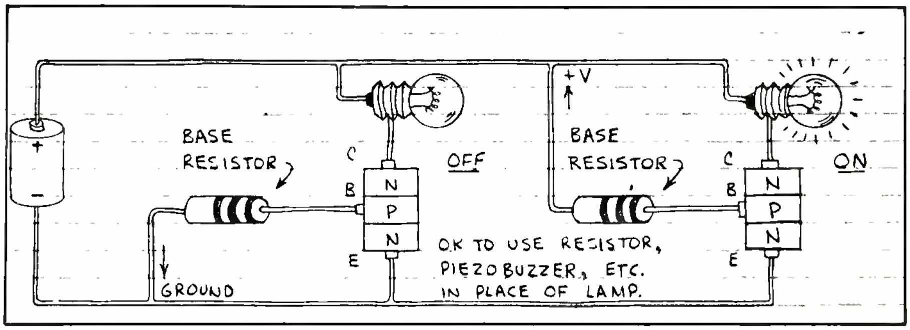
The variable resistor forward-biases the transistor *and* controls the input (base-emitter) current. The meter indicates the output (collector-emitter) current. The series resistor protects the meter from excessive current that can damage it.

In a working circuit, the variable resistor may be in series with a second component whose resistance varies with temperature, light, moistur etc. When the input signal chang rapidly, an ac amplifier such as t! one shown below is used.

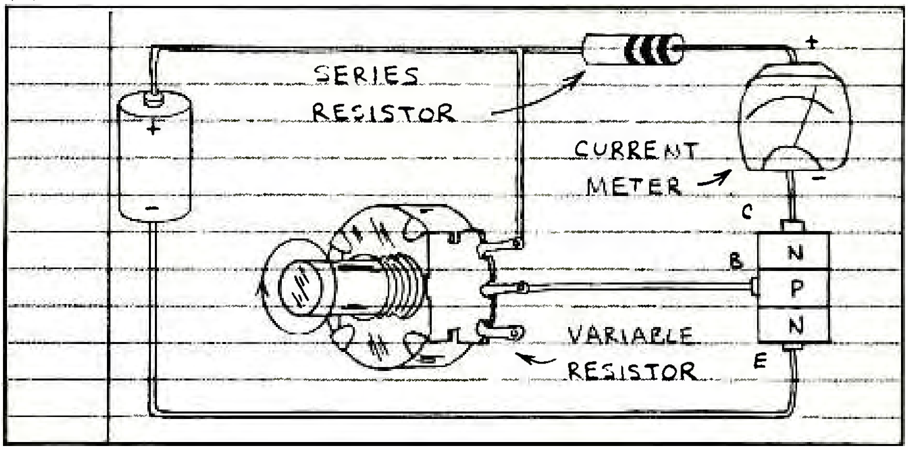
Transistor AC Amplifier (C)

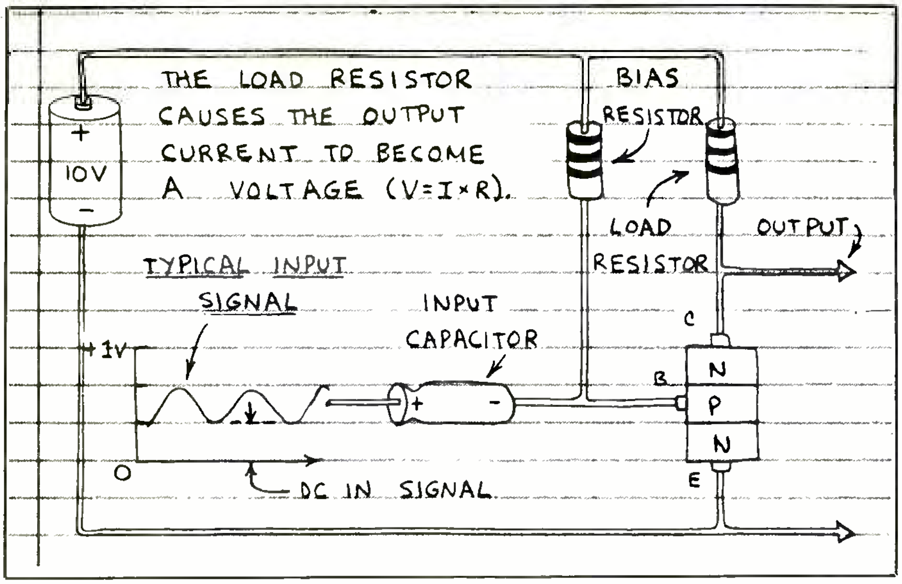
This is the simplest of several basic ac amplifiers. The input capacitor blocks any dc in the input signal. The bias resistor is selected to give an output voltage of about half the battery voltage. The amplified signal "rides" on this steady output voltage and varies above and below it. Without

(A)



(B)





(C)

the bias resistor, only the positive half of the input signal *above* 0.6 volt will be amplified. This will cause severe distortion.

About the Direction of Current Flow

An electrical current is the movement of electrons through a conductor or semiconductor. Since electrons move from a negatively charged to a positively charged region, why does the arrowhead in a diode symbol point in the opposite direction? There are two reasons:

1. Beginning with Benjamin Franklin, it was traditionally assumed electricity flows from a positively charged to a negatively charged region. The discovery of the electron changed that. But most electrical circuit diagrams today still follow the old tradition in which the positive power supply connection is placed *above* the negative connection, as if gravity somehow influences the flow of current in an electrical circuit.

2. In a semiconductor, holes flow in the direction opposite that of electron flow. It's therefore common to refer to *positive* current flow in semiconductors.

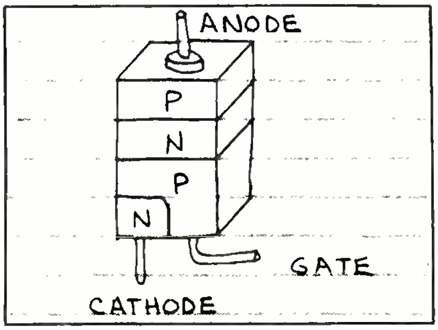
For accuracy here, "current flow" refers to *electron* flow. But we're stuck with symbols that indicate the direction of *hole* flow.

The Thyristor

Like transistors, thyristors are semiconductor devices with three leads. Also similarly, a small current at one lead will allow a much larger current to flow through the other two leads. Unlike transistors, however, thyristors do not amplify fluctuating signals. The controlled current is either fully on or fully off. Hence, thyristors are often referred to as "solid-state switches."

Silicon-Controlled Rectifiers

Commonly called SCRs, these devices make up one of two basic categories of thyristors. They are similar to bipolar transistors with a fourth layer and, therefore, three junctions, as in the following drawing.



The physical construction and schematic symbol for the SCR are shown in (D). Frequently, single-letter designations are used to identify the leads: A for anode, K (or C) for cathode, and G for gate.

If the anode (A) of an SCR is made more positive than the cathode (K), the two outermost junctions are forward biased. The middle *pn* junction, however, is reverse biased, and current cannot flow. A small gate current forward biases the middle *pn* junction and allows a much larger current to flow through the device. The SCR stays on even if the gate current is removed! (until power is disconnected.)

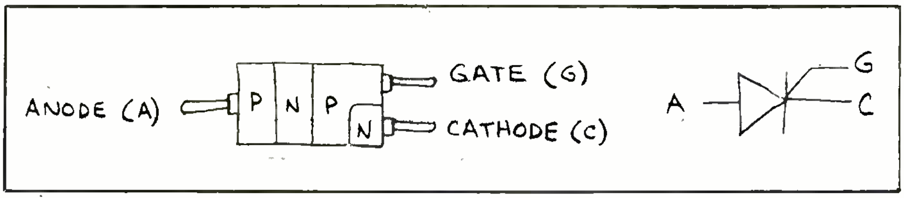
Circuit E shows how an SCR is used to switch on an incandescent lamp. (Other devices can also be controlled.) A single touch of the indicated switch applies current to the G terminal and causes the SCR to "fire." This, in turn, turns on the lamp. Since the SCR is a *dc* device, the current will continue to flow until the switch to the left of the lamp is momentarily pressed.

SCRs are available in a wide range of current and voltage ratings. Low-current SCRs switch up to 1 ampere at up to 100 volts; medium-current SCRs switch up to 10 amperes at up to several hundred volts; and high-current SCRs have the ability to switch up to 2500 amperes at up to several thousand volts!

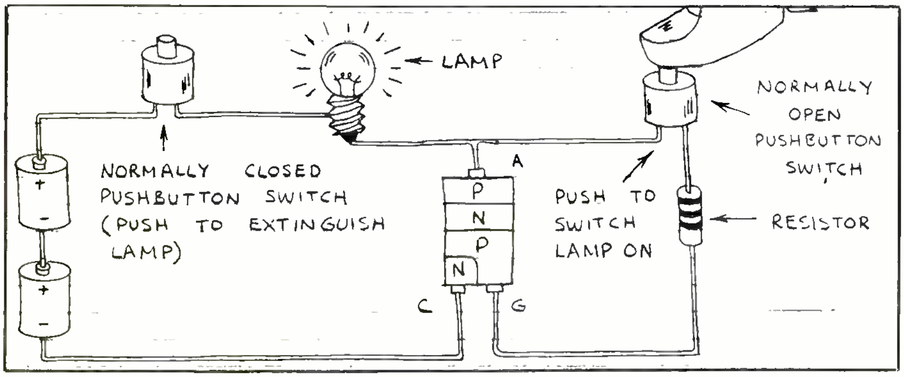
Triacs

The triac is equivalent to two SCRs connected in reverse parallel. This means it can switch both direct *and* alternating current (*dc* and *ac*). Notice that the triac has *five* layers and an extra *n*-type region and that all three leads contact two layers.

Though the triac has a gate terminal similar to that in an SCR, it does not have separate anode and cathode terminals for the simple reason that each of the other terminals connects to both an anode *and* a cathode. Therefore, these terminals are identi-



(D)



(E)

fied as MT1 and MT2, where MT stands for "main terminal." (F)

When used to switch alternating current, the traic stays on only when the gate receives current. Remove the gate current, and the triac switches off when the ac passes through 0 volt.

The arrangement shown in (G) demonstrates how a triac can switch on a lamp powered by household line current. (Motors and other devices can also be controlled).

Like SCRs, triacs are categorized according to the current they can switch. But triacs don't have the very high power capability of high-current SCRs.

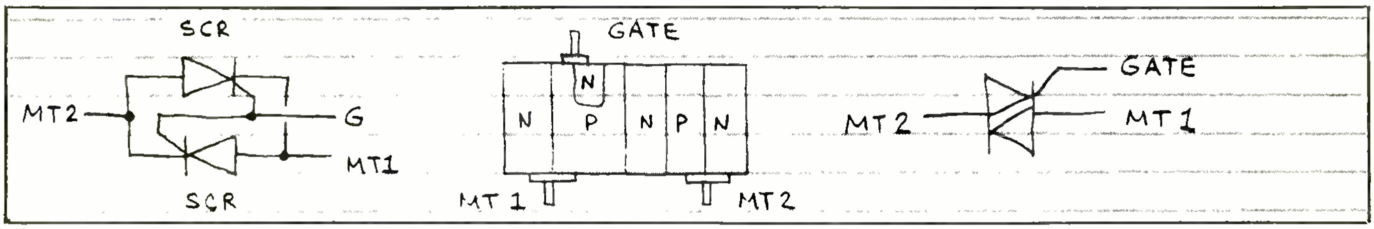
Semiconductor Light Sources

When bombarded by light, heat, electrons, and other forms of energy, most semiconductors will emit visible or infrared light. The best semiconductor light sources are the family of *pn* junction diodes.

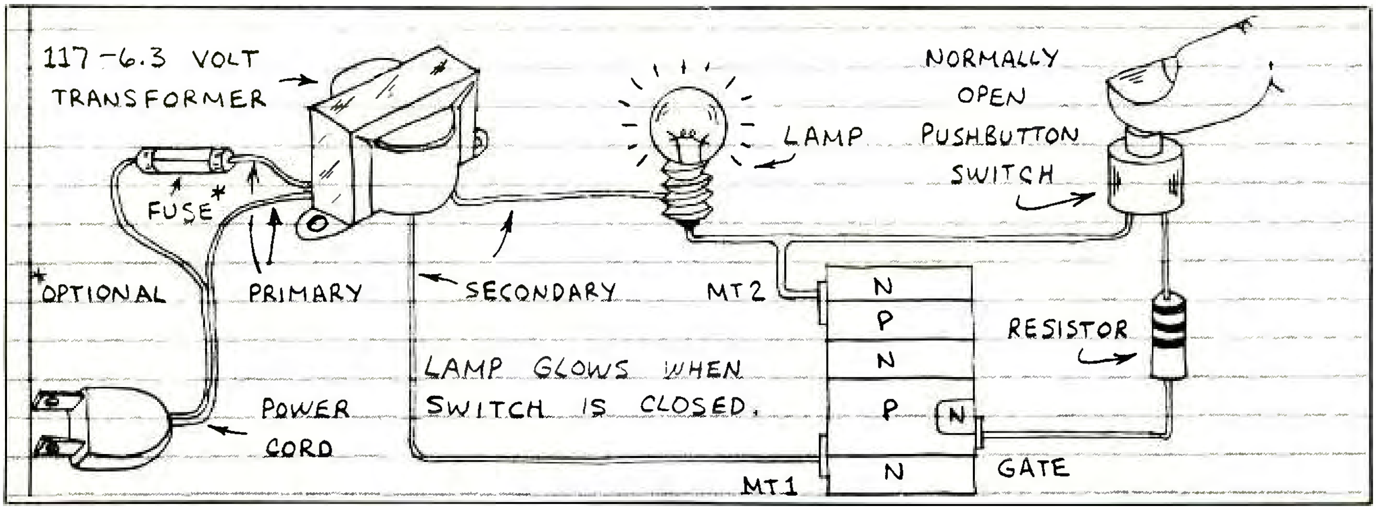
The light-emitting diode (LED) converts an electrical current directly into light. Therefore, the LED is more efficient than many other light sources. Inexpensive visible-light LEDs are used as panel indicators. Certain red LEDs are used to transmit information. Many kinds of LED readouts are capable of displaying digits and characters. They are more rugged than liquid-crystal displays (LCDs), but they use more current during operation.

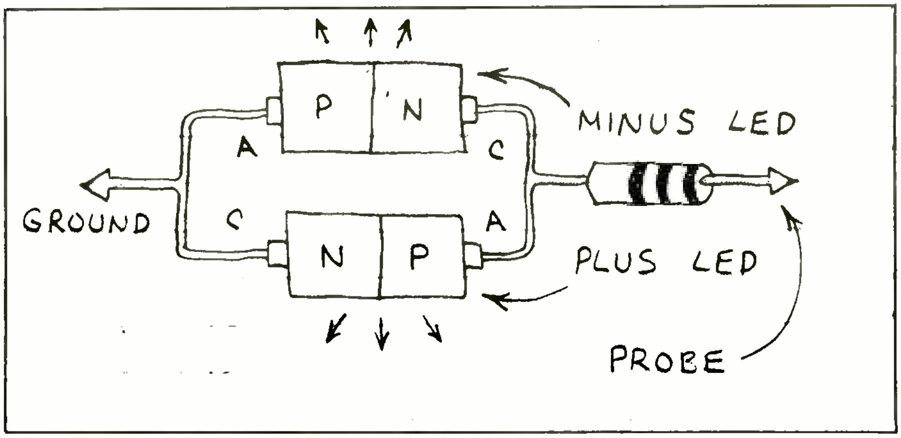
Infrared LEDs should be called "infrared-emitting diodes." They are used to transmit information. A special kind of infrared LED is the diode laser.

(F)



(G)





LEDs can be powered by continuous current or by brief pulses of current. When operated continuously, the current can be varied to change the light output.

Two different-colored LEDs can be connected in reverse parallel to serve as a polarity indicator.

The schematic symbol for the LED is the same as for the ordinary diode, except that it is accompanied by a pair or arrows pointing away from the symbol.

Semiconductor Light Detectors

Energy entering a semiconductor crystal can generate a current in the crystal. This is the basis of semiconductor light detectors. There are two major classes of semiconductor light detectors—those with and those without *pn* junctions. The latter are mentioned here only to make you aware of their existence. The remainder of this discussion will deal with the *pn*-junction detector.

Pn-junction light detectors form the largest family of photonic semiconductors. Most are made from silicon and can detect both visible light and near-infrared.

Photodiodes are specifically designed for light detection. They are used in cameras, intrusion alarms, lightwave communications, etc.

A photodiode will create a hole/electron pair at a *pn* junction. A current will flow if the two sides of the junction are connected. When operated in the photovoltaic mode, the photodiode becomes a current source when it is illuminated. In photoconductive operation, the photodiode is reverse-biased. A current flows when the *pn* junction is illuminated. When dark, a tiny current, called the "dark current," will flow.

There are many case styles used for photodiodes. Most have plastic housings, built-in lenses and filters, etc. The most important distinction is the size of the semiconductor chip.

Small-area photodiodes have very fast response times when used in the reverse-biased photoconductive mode. Large-area photodiodes, though slower responding, provide high sensitivity to photo energy.

The schematic symbol for the photodiode is the same as for the LED, except that the arrows point *toward* the symbol.

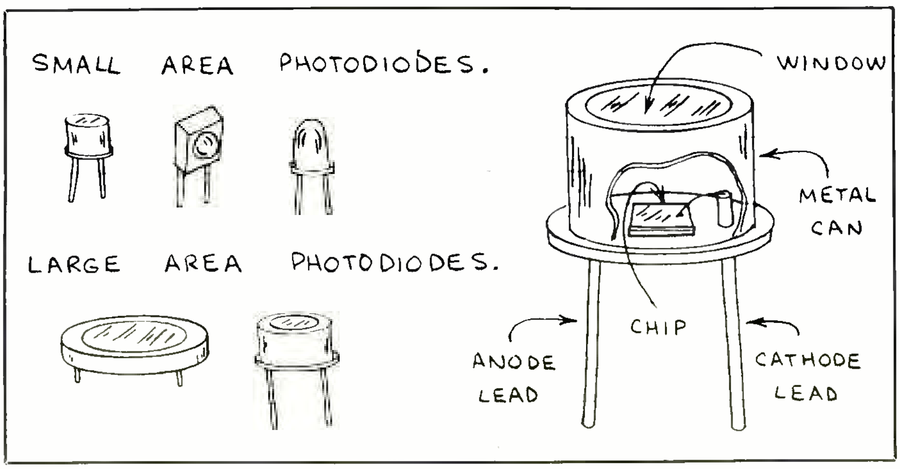
Photodiodes are commonly used to detect fast pulses of near-infrared, as in lightwave communications. The following circuit demonstrates how the photodiode might be used in a light-meter arrangement. Response of this circuit is very linear.

Phototransistors are specifically designed to take advantage of the light-sensitive property in all transistors. The most common phototransistor is an *npn* device with a large, exposed base region. Light entering the base replaces the base-emitter current of ordinary *npn* transistors. Therefore, a phototransistor directly amplifies variations in the light.

Two types of *npn* phototransistors are available. One is an *npn* device, while the other includes a second *npn* transistor to provide more amplification. The latter is called a "photodarlington" transistor and is very sensitive, though it's slower than the ordinary *npn* phototransistor.

The symbols commonly used for phototransistors are as follows:

Phototransistors are often used to detect fluctuating (ac) light signals. They are also used to detect steady (dc) light, as in this circuit in which a photodarlington is used to energize a relay. **ME**



Continued next month, where our focus will be on integrated circuits.