# 12 / NEGATIVE-RESISTANCE DEVICES

A number of devices which find extensive application in pulse and switching circuitry are most conveniently characterized in terms of a volt-ampere curve which displays, over a limited range, a negative incremental resistance. In this chapter we describe the physical principles which account for this characteristic in the tunnel diode, the unijunction transistor, the *p-n-p-n* diode, the silicon controlled switch, and the thyristor. In the following chapter, circuits are constructed with these negative-resistance devices and it is demonstrated that bistable, monostable, and astable operation are possible.

## 12-1 THE TUNNEL DIODE

A *p-n* junction diode of the type discussed in Sec. 6-1 has an impurity concentration of about 1 part in $10^8$. With this amount of doping the width of the depletion layer, which constitutes a potential barrier at the junction, is of the order of 5 microns ($5 \times 10^{-4}$ cm). This potential barrier restrains the flow of carriers from the side of the junction where they constitute majority carriers to the side where they constitute minority carriers. If the concentration of impurity atoms is greatly increased, say to 1 part in $10^3$, then the device characteristics are completely changed. This new diode was announced in 1958 by Esaki,[1] who also gave the correct theoretical explanation for its volt-ampere characteristic, which is depicted in Fig. 12-1. The width of the junction barrier varies inversely as the square root of impurity concentration and therefore is reduced from 5 microns to about 100 Å ($10^{-6}$ cm). This thickness is only about one-fiftieth the wavelength of visible light. Classically, a particle must have an energy





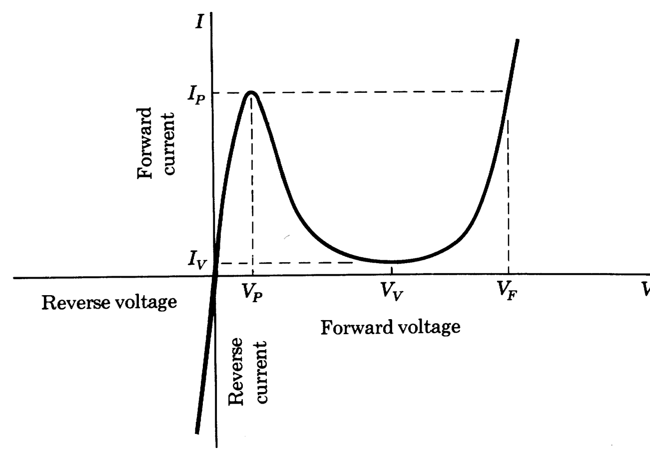


Fig. 12-1 Volt-ampere characteristic of a tunnel diode.

at least equal to the height of a potential barrier if it is to move from one side of the barrier to the other. However, for barriers as thin as those estimated above in the Esaki diode, quantum mechanics dictates that there is a large probability that an electron will penetrate *through* the barrier. The quantum-mechanical behavior is referred to as "tunneling," and hence these high-impurity-density *p-n* junction devices are called "tunnel diodes." This same tunneling effect is responsible for high-field emission of electrons from a cold metal and for radioactive emissions.

As a consequence of the tunneling effect and the band structure of heavily doped semiconductors the volt-ampere characteristic of Fig. 12-1 is obtained.[2,3] The device is an excellent conductor in the reverse direction (*p* side of junction negative with respect to the *n* side). Also, for small forward voltages (up to 50 mV for Ge) the resistance remains small (of the order of 5 Ω). At the *peak current* $I_P$ corresponding to the voltage $V_P$ the slope $dI/dV$ of the characteristic is zero. If $V$ is increased beyond $V_P$, then the current decreases. As a consequence the dynamic conductance $g = dI/dV$ is negative. The tunnel diode exhibits a *negative-resistance characteristic* between the peak current $I_P$ and the minimum value $I_V$, called the *valley current*. At the *valley voltage* $V_V$ at which $I = I_V$ the conductance is again zero, and beyond this point the resistance becomes and remains positive. At the so-called *peak forward voltage* $V_F$ the current again reaches the value $I_P$. For larger voltages the current increases beyond this value. The portion of the characteristic beyond $V_V$ is caused by the injection current in an ordinary *p-n* junction diode. The remainder of the characteristic is a result of the tunneling phenomenon in the highly doped diode.

For currents whose values are between $I_V$ and $I_P$ the curve is triple-valued, because each current can be obtained at three different applied voltages. It is this multivalued feature which makes the tunnel diode useful in pulse and

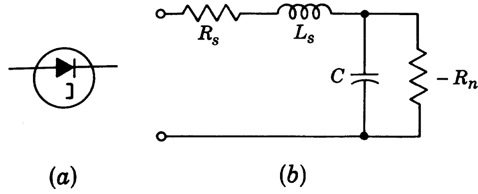


Fig. 12-2 (a) Symbol for a tunnel diode (Ref. 4); (b) small-signal model in the negative-resistance region.

digital circuitry (Chap. 13). Note that whereas the characteristic in Fig. 12-1 is a multivalued function of current, it is a single-valued function of voltage. Each value of $V$ corresponds to one and only one current. Hence, the tunnel diode is said to be *voltage-controllable*. The vacuum-tube tetrode is another negative-resistance device belonging to the voltage-controllable class. On the other hand, there also exist negative-resistance devices whose characteristics are multivalued functions of voltage but are single-valued with respect to current. These *current-controllable* devices (for example, the unijunction transistor, the *p-n-p-n* diode, etc.) are discussed later in this chapter.

The standard circuit symbol[4] for a tunnel diode is given in Fig. 12-2*a*. The small-signal model for operation in the negative-resistance region is indicated in Fig. 12-2*b*. The negative resistance $-R_n$ has a minimum at the point of inflection between $I_P$ and $I_V$. The series resistance $R_s$ is ohmic resistance. The series inductance $L_s$ depends upon the lead length and the geometry of the diode package. The junction capacitance $C$ depends upon the bias and is usually measured at the valley point. Typical values for these parameters for a tunnel diode of peak current value $I_P = 10$ mA are $-R_n = -30\ \Omega$, $R_s = 1\ \Omega$, $L_s = 5$ nH, and $C = 20$ pF.

Our principal interest in the tunnel diode is its application as a very high speed switch. Since tunneling takes place at the speed of light, the transient response is limited only by total shunt capacitance (junction plus stray wiring capacitance) and peak driving current. Switching times of the order of a nanosecond are reasonable, and times as low as 50 psec have been obtained.

The most common commercially available tunnel diodes are made from germanium or gallium arsenide. It is difficult to manufacture a silicon tunnel diode with a high ratio of peak-to-valley current $I_P/I_V$. Table 12-1 summarizes the important static characteristics of these devices. The voltage values in this table are determined principally by the particular semiconductor used

**TABLE 12-1 Typical tunnel-diode parameters**

| | Ge | GaAs | Si |
|---|---|---|---|
| $I_P/I_V$ | 8 | 15 | 3.5 |
| $V_P$, V | 0.055 | 0.15 | 0.065 |
| $V_V$, V | 0.35 | 0.50 | 0.42 |
| $V_F$, V | 0.50 | 1.10 | 0.70 |



and are almost independent of the current rating. Note that gallium arsenide has the highest ratio $I_P/I_V$ and the largest voltage swing $V_F - V_P \approx 1.0$ V as against 0.45 V for germanium.

The peak current $I_P$ is determined by the impurity concentration (the resistivity) and the junction area. A spread of 20 percent in the value of $I_P$ for a given tunnel-diode type is normal, but tighter-tolerance diodes are also available. For computer applications, devices with $I_P$ in the range of 1 to 100 mA are most common. However, it is possible to obtain diodes whose $I_P$ is as small as 100 $\mu$A or as large as 100 A.

The peak point ($V_P$, $I_P$), which is in the tunneling region, is not a very sensitive function of temperature. Commercial diodes are available[3] for which $I_P$ and $V_P$ vary by only about 10 percent over the range $-50$ to $+150$°C. The temperature coefficient of $I_P$ may be positive or negative, depending upon the impurity concentration and the operating temperature, but the temperature coefficient of $V_P$ is always negative. The valley point $V_V$, which is affected by injection current, is quite temperature-sensitive. The value of $I_V$ increases rapidly with temperature and at 150°C may be two or three times its value at $-50$°C. The voltages $V_V$ and $V_F$ have negative temperature coefficients of about 1.0 mV/°C, a value only about half that found for the shift in voltage with temperature of a *p-n* junction diode or transistor. These values apply equally well to Ge or GaAs diodes. Gallium arsenide devices show a marked reduction of the peak current if operated at high current levels in the forward injection region. However, it is found empirically[3] that negligible degradation results if, at room temperature, the average operating current $I$ is kept small enough to satisfy the condition $I/C \leq 0.5$ mA/pF, where $C$ is the junction capacitance. Tunnel diodes are found to be several orders of magnitude less sensitive to nuclear radiation than are transistors.

The advantages of the tunnel diode are low cost, low noise, simplicity, high speed, environmental immunity, and low power. The disadvantages of the diode are its low output-voltage swing and the fact that it is a two-terminal device. Because of the latter feature, there is no isolation between input and output, and this leads to serious circuit-design difficulties. Hence, a transistor (an essentially unilateral device) is usually preferred for frequencies below about 1 GHz (a kilomegacycle per second) or for switching times longer than several nanoseconds. The tunnel diode and transistor may be combined advantageously, as indicated in Sec. 13-11.

## 12-2 THE BACKWARD DIODE

A tunnel diode designed to have a small peak current ($I_P$ of the order of $I_V$) may be used to advantage, in the reverse direction, for purposes for which the conventional diode is employed in the forward direction. The volt-ampere characteristic of such a "tunnel rectifier" is shown in Fig. 12-3. Because

this device is a better conductor in the reverse than in the forward direction it is also called a "backward diode" or simply a "back diode." In the neighborhood of zero voltage, in response to either a forward-biasing or reverse-biasing voltage, the tunnel diode responds with a current which is large in comparison to the corresponding current in a conventional diode. These large currents are a result of the tunneling effect. In the back diode, the current due to tunneling is large only in the reverse direction. For this reason the back diode is also called the "unitunnel diode."

The high-conduction portion of the volt-ampere characteristic of Fig. 12-3 is in the third quadrant. Since this portion of the characteristic corresponds to the region of forward conduction in a conventional diode, it is customary to plot the back diode with the voltage and current scales both reversed. In the back diode, the "forward direction" of applied voltage is actually the direction where the $p$ side of the diode is negative with respect to the $n$ side. The appearance of the characteristics as normally supplied by manufacturers may be seen by turning the page upside down.

The merits of the back diode are made clear in Fig. 12-4, where the "forward" characteristics at various temperatures of a typical silicon back diode are compared with the forward characteristics of a conventional diode. We note that the temperature sensitivity of the back diode is appreciably less than the sensitivity of the conventional diode. The back diode has a sensitivity of about $-0.1$ mV/°C (both silicon and germanium), compared, as we have seen, with about $-2$ mV/°C for the conventional diode. We observe further that while the conventional silicon diode has a break point, at room temperature, between 0.6 and 0.7 V, the back diode has a break point at 0 V. The back diode is therefore very useful when the rectifying action of a diode is required in connection with small-amplitude waveforms. Suppose, by way of example, that a sinusoidal signal is applied to a rectifying circuit which consists of a diode and resistor in series. If the signal has an amplitude of, say, 200 mV, and the diode is a conventional device (silicon or germanium), the diode will hardly conduct at any point in the cycle and the efficiency of rectification will be very poor. With a back diode the efficiency will be greatly improved.

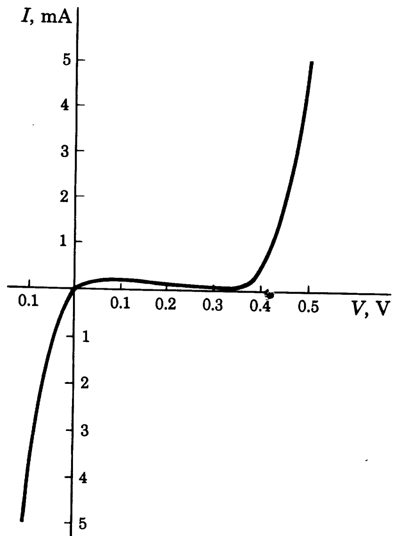


Fig. 12-3 A typical germanium backward-diode characteristic.



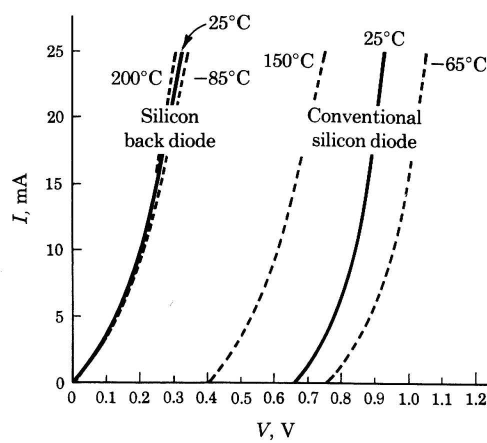


Fig. 12-4 The characteristics of a typical silicon back diode at various temperatures compared with the corresponding characteristics of a conventional silicon diode. (Courtesy of Hoffman Semiconductor.)

## 12-3 THE UNIJUNCTION TRANSISTOR[5]

The construction of this device is indicated in Fig. 12-5*a*. A bar of high-resistivity $n$-type silicon of typical dimensions $8 \times 10 \times 35$ mils, called the *base B*, has attached to it at opposite ends two ohmic contacts $B1$ and $B2$. A 3-mil aluminum wire, called the *emitter E*, is alloyed to the base to form a $p$-$n$ rectifying junction. This device was originally described in the literature as the *double-base diode* but is now commercially available under the designation

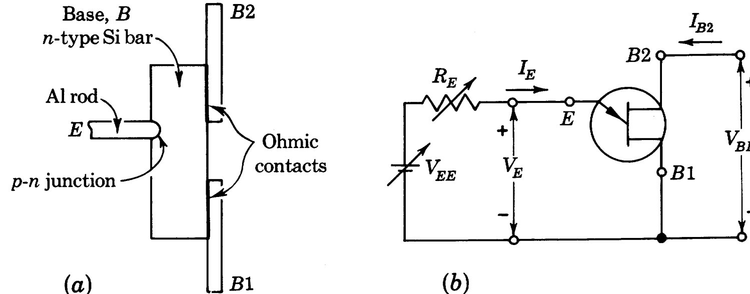


Fig. 12-5 Unijunction transistor. (*a*) Constructional details; (*b*) circuit symbol (Ref. 4).

*unijunction transistor* (UJT). The standard symbol for this device is shown in Fig. 12-5*b*. Note that the emitter arrow is inclined and points toward $B1$, whereas the ohmic contacts $B1$ and $B2$ are brought out at right angles to the line which represents the base.

As usually employed, a fixed interbase potential $V_{BB}$ is applied between $B1$ and $B2$. The most important characteristic of the UJT is that of the input diode between $E$ and $B1$. If $B2$ is open-circuited so that $I_{B2} = 0$, then the input volt-ampere relationship is that of the usual *p-n* junction diode as given in Eq. (6-1). In Fig. 12-6 the input current-voltage characteristics are plotted for $I_{B2} = 0$ and also for a fixed value of interbase voltage $V_{BB}$. The latter curve is seen to have the current-controlled negative-resistance characteristic which is single-valued in current but may be multivalued in voltage. A qualitative explanation of the physical origin of the negative resistance will now be given.

If $I_E = 0$ then the silicon bar may be considered as an ohmic resistance $R_{BB}$ between base leads. Usually $R_{BB}$ lies in the range between 5 and 10 K. Between $B1$ (or $B2$) and the $n$ side of the emitter junction the resistance is $R_{B1}$ (or $R_{B2}$, respectively), so that $R_{BB} = R_{B1} + R_{B2}$. Under this condition of zero or very small emitter current the voltage on the $n$ side of the emitter junction is $\eta V_{BB}$, where $\eta \equiv R_{B1}/R_{BB}$ is called the *intrinsic stand-off ratio*. This parameter is specified by the manufacturer and usually lies between 0.5 and 0.75. If $V_E$ is less than $\eta V_{BB}$, then the *p-n* junction is reverse-biased and the input current $I_E$ is negative. As indicated in Fig. 12-6, the maximum value of this negative current is the reverse saturation current $I_{EO}$, which is of the order of only 10 $\mu$A even at 150°C. If $V_{EE}$ is increased beyond $\eta V_{BB}$ the input diode becomes forward-biased and $I_E$ goes positive. However, as already noted in connection with Fig. 12-6, the current remains quite small until the forward bias equals the cutin voltage $V_\gamma$ ($\approx$ 0.6 V), and then increases very rapidly with small increases in voltage. We must now take account of the conductivity modulation of the base region due to $I_E$.

The emitter current increases the charge concentration between $E$ and $B1$ because holes are injected into the *n*-type silicon. Since conductivity

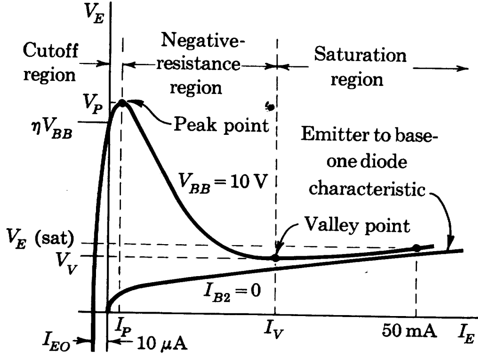


**Fig. 12-6 Idealized input characteristic of a unijunction transistor. (Courtesy of General Electric Company.)**



is proportional to charge density, the resistance $R_{B1}$ decreases with increasing emitter current. Hence, for voltages above the threshold value $V_\gamma$, as $I_E$ is increased (by either increasing $V_{EE}$ or decreasing $R_E$ in Fig. 12-5*b*) the voltage $V_E$ between $E$ and $B1$ decreases because of the decrease in the value of the resistance $R_{B1}$. Since the current is increasing while the voltage is decreasing, then this device possesses a negative resistance.

After the emitter current has become very large compared with $I_{B2}$, then we may neglect $I_{B2}$. Hence, for very large $I_E$ the input characteristic asymptotically approaches the curve for $I_{B2} = 0$. As indicated in Fig. 12-6, this behavior results in a minimum or valley point where the resistance changes from negative to positive. For currents in excess of the valley current $I_V$ the resistance remains positive. This portion of the curve is called the *saturation region*. The voltage at $I_E = 50$ mA is arbitrarily called the *saturation voltage* $V_E$(sat) and is of the order of 3 V.

At the maximum voltage or peak point $V_P$ the current is very small ($I_P \approx 25$ $\mu$A), and hence the region to the left of the peak point is called the *cutoff region*. For many applications the most important parameter is the peak voltage $V_P$, which, as explained above, is given by

$$V_P = \eta V_{BB} + V_\gamma \qquad (12\text{-}1)$$

In Sec. 6-1 we noted that $V_\gamma \approx 0.6$ and decreases about 2 mV/°C. (Both of these facts are approximated by replacing $V_\gamma$ by $200/T$, where $T$ is the junction temperature in degrees Kelvin.) Since the temperature coefficient of $R_{B1}$ is the same as that of $R_{B2}$, then $\eta = R_{B1}/(R_{B1} + R_{B2})$ should be independent of temperature. Experimentally it is found that the temperature coefficient of $\eta$ is less than 0.01 percent/°C and may be either positive or negative. To illustrate that $V_P$ is quite insensitive to temperature, assume $\eta = 0.5$, $V_{BB} = 20$ V, and a temperature change from 25 to 125°C. At 25°C, $V_P = 10.6$ V. At 125°C the change due to $V_\gamma$ is 0.2 V and due to $\eta$ is a maximum of 0.2 V. Hence $V_P$ will decrease no more than 4 percent over the 100°C increase in temperature. The peak voltage can be made even much less sensitive to temperature by adding a small resistance $R_2$ in series with the $B2$ lead (Prob. 12-12).

The peak current $I_P$ varies inversely with the interbase voltage $V_{BB}$ and decreases with increasing temperature. Typically, a peak current of 10 $\mu$A at 25°C will reduce to about 6 $\mu$A at 125°C and increase to 12 $\mu$A at $-55$°C.

A family of input characteristics for commercially available UJTs is indicated in Fig. 12-7*a*. Note that the peak voltage increases linearly with $V_{BB}$, and observe also that the valley is very broad. Hence it is difficult to give the exact value of the valley current $I_V$, but the valley voltage $V_V$ can be determined fairly accurately. The valley current has about the same temperature coefficient as the peak current.

A family of output characteristics are given in Fig. 12-7*b*. The straight line for $I_E = 0$ indicates that with the emitter open-circuited the *n*-type silicon bar is essentially ohmic. The reciprocal slope of this line gives $R_{BB}$.

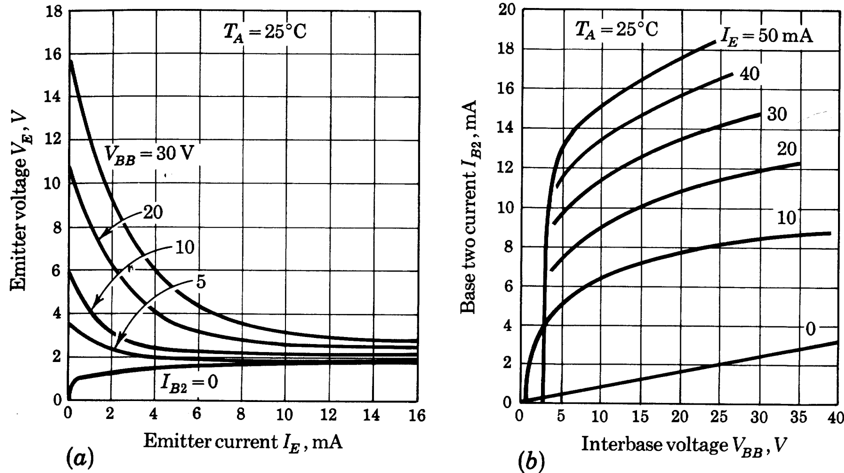


**Fig. 12-7** Unijunction characteristics for types 2N489 through 2N494. (*a*) Input; (*b*) output. (Courtesy of General Electric Company.)

For $I_E = 50$ mA the drop across $R_{B1}$ is 3 V even if $I_{B2} = 0$. As $I_{B2}$ is increased, $R_{B1}$ decreases, and the decreased drop across $R_{B1}$ offsets the increase in voltage across $R_{B2}$. Hence the voltage $V_{BB}$ remains almost constant for small values of $I_{B2}$ (up to about 10 mA) and then rises with $B2$ current.

We shall find when we discuss applications of the unijunction transistor that the input characteristics in Fig. 12-7*a* are much more important than the output curves of Fig. 12-7*b*. The most useful features of the UJT are its stable firing voltage $V_P$ which depends linearly on $V_{BB}$, the low (microampere) firing current $I_P$, the stable negative-resistance characteristic, and the high pulse-current capability.

## 12-4 THE FOUR–LAYER DIODE[6]

Another device which exhibits a negative resistance and which finds extensive applications in switching circuits is represented in Fig. 12-8, together with its circuit symbol. The device consists of four layers of silicon doped alternately

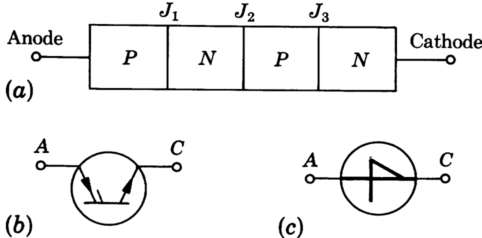


**Fig. 12-8** (*a*) A four-layer *p-n-p-n* diode; (*b*) standard circuit symbol (Ref. 4); (*c*) alternative symbol (not recommended).



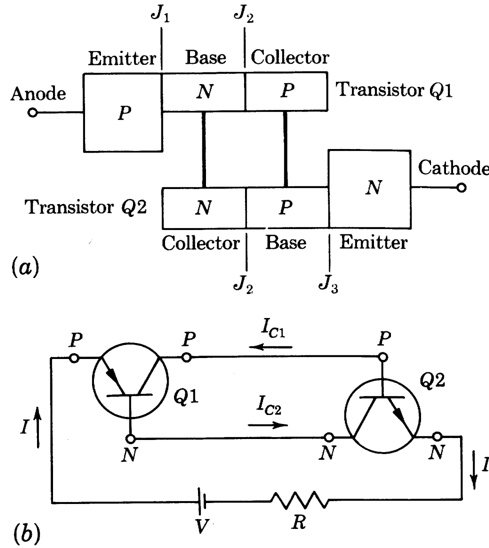


**Fig. 12-9** (*a*) The *p-n-p-n* diode is redrawn to make it appear as two interconnected "transistors." (*b*) The two interconnected transistors are supplied current from a source through a resistor.

with *p*- and *n*-type impurities. Because of this structure it is called a *p-n-p-n* (often pronounced "pinpin") diode or switch. The terminal $P$ region is the anode, or $p$ emitter, and the terminal $N$ region is the cathode, or $n$ emitter. When an external voltage is applied to make the anode positive with respect to the cathode, junctions $J_1$ and $J_3$ are forward-biased and the center junction $J_2$ is reverse-biased. The externally impressed voltage appears principally across the reverse-biased junction, and the current which flows through the device is small. As the impressed voltage is increased, the current increases slowly until a voltage called the *firing* or *breakover* voltage $V_{BO}$ is reached where the current increases abruptly and the voltage across the device decreases sharply. At this breakover point the *p-n-p-n* diode switches from its OFF (also called *blocking*) state to its ON state.

In Fig. 12-9*a*, the *p-n-p-n* switch has been split into two parts which have been displaced mechanically from one another but left electrically connected. This splitting is intended to illustrate that the device may be viewed as two transistors back to back. One transistor is a *p-n-p* type, whereas the second is an *n-p-n* type. The $N$ region that is the base of one transistor is the collector of the other, and similarly for the adjoining $P$ region. The junction $J_2$ is a common collector junction for both transistors. In Fig. 12-9*b* the arrangement in Fig. 12-9*a* has been redrawn using transistor-circuit symbols, and a voltage source has been impressed through a resistor across the switch, giving rise to a current $I$. Collector currents $I_{C1}$ and $I_{C2}$ for transistors $Q1$ and $Q2$ are indicated. In the active region the collector current is given by Eq. (6-13),

$$I_C = -\alpha I_E + I_{CO} \tag{12-2}$$

with $I_E$ the emitter current, $I_{CO}$ the reverse saturation current, and $\alpha$ the short-circuit common-base forward current gain. We may apply Eq. (12-2), in turn, to $Q1$ and $Q2$. Since $I_{E1} = +I$ and $I_{E2} = -I$, we obtain

$$I_{C1} = -\alpha_1 I + I_{CO1} \tag{12-3}$$

$$I_{C2} = \alpha_2 I + I_{CO2} \tag{12-4}$$

For the *p-n-p* transistor $I_{CO1}$ is negative, whereas for the *n-p-n* device $I_{CO2}$ is positive. Hence, we write $I_{CO2} = -I_{CO1} \equiv I_{CO}/2$. Setting equal to zero the sum of the currents into transistor $Q1$ we have

$$I + I_{C1} - I_{C2} = 0 \tag{12-5}$$

Combining Eqs. (12-3) through (12-5) we find

$$I = \frac{I_{CO2} - I_{CO1}}{1 - \alpha_1 - \alpha_2} = \frac{I_{CO}}{1 - \alpha_1 - \alpha_2} \tag{12-6}$$

We observe that as the sum $\alpha_1 + \alpha_2$ approaches unity Eq. (12-6) indicates that the current $I$ increases without limit; that is, the device breaks over. Such a development is not unexpected in view of the regenerative manner in which the two transistors are interconnected. The collector current of $Q1$ is furnished as the base current of $Q2$, and vice versa. When the *p-n-p-n* switch is operating in such a manner that the sum $\alpha_1 + \alpha_2$ is less than unity, the switch is in its OFF state and the current $I$ is small. When the condition $\alpha_1 + \alpha_2 = 1$ is attained, the switch transfers to its ON state. The voltage across the switch drops to a low value and the current becomes large, being limited by the external resistance in series with the switch.

The reason why the device can exist in either of two states is that at very low currents $\alpha_1$ and $\alpha_2$ may be small enough so that $\alpha_1 + \alpha_2 < 1$, whereas at larger currents the $\alpha$'s increase, thereby making it possible to attain the condition $\alpha_1 + \alpha_2 = 1$. Thus, as the voltage across the switch is increased from zero, the current starts at a very small value and then increases because of avalanche multiplication (not avalanche breakdown) at the reverse-biased junction. This increase in current increases $\alpha_1$ and $\alpha_2$. *When the sum of the small-signal avalanche-enhanced alphas equals unity, $\alpha_1 + \alpha_2 = 1$, breakover occurs.* At this point, the current is large, and $\alpha_1$ and $\alpha_2$ might be expected individually to attain values in the neighborhood of unity. If such were the case, then Eq. (12-6) indicates that the current might be expected to reverse. What provides stability to the ON state of the switch is that in the ON state the center junction becomes forward-biased. Now all the transistors are in saturation and the current gain $\alpha$ is again small. Thus, stability is attained by virtue of the fact that the transistors enter saturation to the extent necessary to maintain the condition $\alpha_1 + \alpha_2 = 1$.

In the ON state all junctions are forward-biased, and so the total voltage across the device is equal very nearly to the algebraic sum of these three saturation junction voltages. The voltage drop across the center junction $J_2$ is in a direction opposite to the voltages across the junctions $J_1$ and $J_3$. This feature serves additionally to keep quite low (of the order of 0.7 V) the total voltage drop across the switch in the ON state.



The operation of the *p-n-p-n* switch depends, as we have seen, on the fact that at low currents, the current gain $\alpha$ may be less than one-half, a condition which is necessary if the sum of two $\alpha$'s is to be less than unity. This characteristic of $\alpha$ is not encountered in germanium but is distinctive of silicon, where it results from the fact that at low currents an appreciable part of the current which crosses the emitter junction is caused by recombination of holes and electrons in the transition region rather than the injection of minority carriers across the junction from emitter to base.[7] In germanium it is not practicable to establish $\alpha_1 + \alpha_2 < 1$. Therefore germanium structures incline to settle immediately in the ON state and have no stable OFF state. Accordingly, germanium *p-n-p-n* switches are not available. We shall see in the discussion below that the *p-n-p-n* structure and mechanism are basic to a large number of other switching devices and in all but one (Sec. 12-9) the material employed is silicon. Thus we shall encounter silicon controlled rectifiers (SCR) but no germanium controlled rectifiers, etc.

## 12-5 *p-n-p-n* CHARACTERISTICS

The volt-ampere characteristic of a *p-n-p-n* diode, not drawn to scale, is shown in Fig. 12-10. When the voltage is applied in the reverse direction, the two outer junctions of the switch are reverse-biased. At an adequately large voltage, breakdown will occur at these junctions, as indicated, at the "reverse avalanche" voltage $V_{RA}$. However, no special interest is attached to operation in this reverse direction.

When a forward voltage is applied, only a small forward current will flow until the voltage attains the breakover voltage $V_{BO}$. The corresponding current is $I_{BO}$. If the voltage $V$ which is applied through a resistor as in Fig. 12-9*b* is increased beyond $V_{BO}$, the diode will switch from its OFF (blocked) state to its ON (saturation) state and will operate in the saturation region. The device is then said to *latch*. If the voltage is now reduced, the switch will remain ON until the current has decreased to $I_H$. This current and the corresponding voltage $V_H$ are called the *holding* or *latching* current and voltage, respectively. The current $I_H$ is the minimum current required to hold the switch in its ON state.

There are available *p-n-p-n* switches with voltages $V_{BO}$ in the range from tens of volts to some hundreds of volts. The current $I_{BO}$ is of the order, at most, of some hundreds of microamperes. In this OFF range up to breakover, the resistance of the switch is the range from some megohms to several hundred megohms.

The holding current varies, depending on the type, in the range from several milliamperes to several hundred milliamperes. The holding voltage is

found to range from about 0.5 V to about 20 V. The incremental resistance $R_H$ in the saturation state is rarely in excess of 10 Ω and decreases with increasing current. At currents of the order of amperes (which can be sustained briefly under pulsed operation), the incremental resistance may drop to as low as some tenths of an ohm.

The switching parameters of the four-layer diode are somewhat temperature-dependent. A decrease in temperature from room temperature to −60°C has negligible effect on $V_{BO}$, but a temperature increase to +100°C will decrease $V_{BO}$ by about 10 percent. $I_H$ decreases substantially with increase in temperature and increases to a lesser extent with decrease in temperature.

**Rate Effect** We can see that the breakover voltage of a *p-n-p-n* switch depends on the rate[8] at which the applied voltage rises. In Fig. 12-11 we have represented the switch in the OFF state as a series combination of three diodes, two forward-biased and the center one reverse-biased. Across this latter diode we have placed a capacitance which represents the transition capacitance across this reverse-biased junction. When the applied voltage $v$ increases slowly enough so that the current through $C$ may be neglected, we must wait until the avalanche-increased current through $D2$ (which is also the current through $D1$ and $D3$) increases to the point where the current gains satisfy the condition $\alpha_1 + \alpha_2 = 1$. When, however, $v$ changes rapidly, so that the capacitor voltage changes at the rate $dv_C/dt$, a current $C\,dv_C/dt$ passes

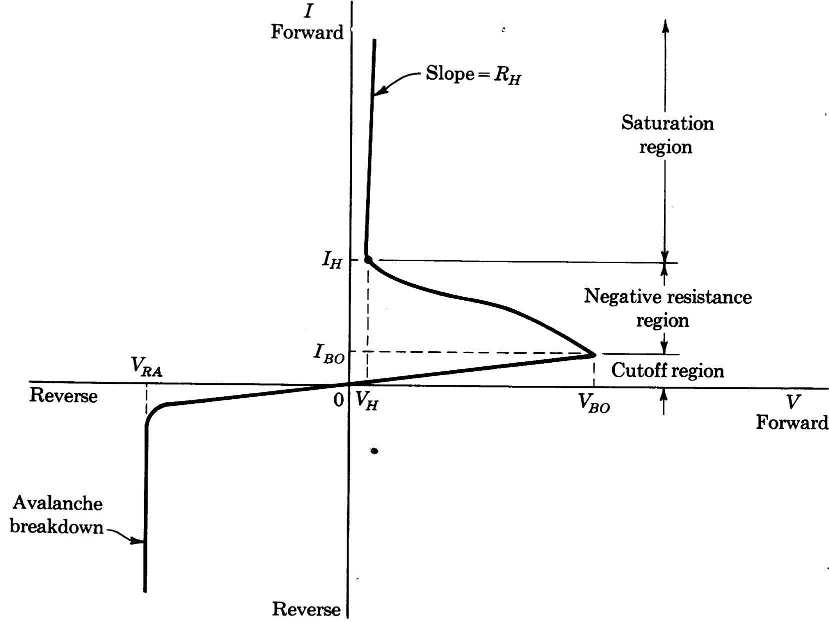


Fig. 12-10 Volt-ampere characteristic of the *p-n-p-n* diode.



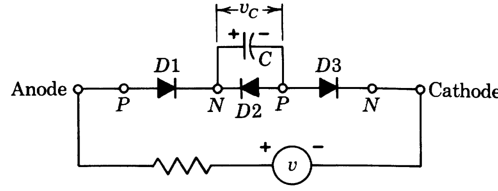


Fig. 12-11 *p-n-p-n* diode in OFF state to illustrate the origin of the rate effect.

through $C$ and adds to the current in $D1$ and $D3$. The current through $D2$ need not be as large as before to attain breakover, and switching takes place at a lower voltage. The capacitance at the reverse-biased junction may lie in the range of some tens of picofarads to over 100 pF, and the reduction in switching voltage may well make itself felt for voltage rates of change $dv_C/dt$ of the order of tens of volts per microsecond.

The discussion above suggests that $V_{BO}$ would continue to become smaller as $dv_C/dt$ increases. Actually, as the rate $dv_C/dt$ increases and becomes very rapid the reduction of switching voltage may become much less pronounced. The reason for this apparently anomalous situation is that, before switching takes place, not only must the switch current reach a certain level but also a time interval is required for the redistribution of base charge in the two "transistors" to allow the end junctions to function as emitting junctions of a transistor. If the applied voltage rises rapidly enough it may well have reached the d-c breakover voltage before this redistribution of base charges has been completed. This matter of stored base charge is considered in detail in Chap. 20, where there is also a discussion of matters relating to the speed with which a transistor can be turned ON and OFF. For the present we simply note that for typical *p-n-p-n* switches the time required to complete the transition from OFF to ON is about 0.1 μsec, and the time to complete the reverse transition is about 0.2 μsec.

## 12-6 THE SILICON CONTROLLED SWITCH[9,10]

The structure of the silicon controlled switch (SCS) consists of four alternate *p*- and *n*-type layers, as in the four-layer diode. In the SCS (also called a *p-n-p-n transistor* or *n-p-n-p switch*) connections are made available to the inner layers, which are not accessible in the diode. The circuit symbols for the SCS are shown in Fig. 12-12. The terminal connected to the $P$ region nearest the cathode is called the *cathode gate*, or *p base*, and the terminal connected to the $N$ region nearest the anode is called the *anode gate*, or *n base*. In very many switch types both gates are not brought out. Where only one gate terminal is available it is ordinarily the cathode gate. These three-terminal devices are available under a variety of commercial names (Sec. 12-8).

The usefulness of the gate terminals rests on the fact that currents introduced into one or both gate terminals may be used to control the anode-to-cathode breakover voltage. Such behavior is to be expected on the basis

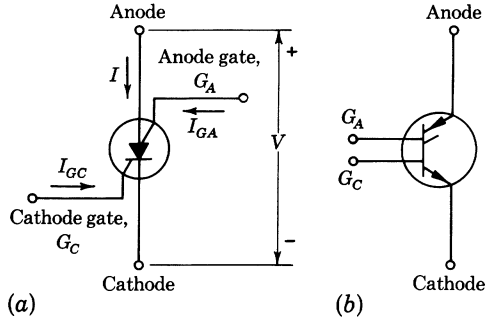


**Fig. 12-12** (*a*) Circuit symbol used by most manufacturers for the SCS; (*b*) alternative symbol (Ref. 4).

of the earlier discussion of the condition $\alpha_1 + \alpha_2 = 1$ which establishes the firing point. If the current through one or both outer junctions is increased as a result of currents introduced at the gate terminals, then $\alpha$ increases and the breakover voltage will be decreased. In Fig. 12-13 the volt-ampere characteristic of an SCS is shown for various cathode-gate currents. We observe that the firing voltage is a function of the gate current, decreasing with increasing gate current and increasing when the gate current is negative and consequently in a direction to reverse-bias the cathode junction. The current after breakdown may well be larger by a factor of a thousand than the current before breakdown. When the gate current is very large, breakover may occur at so low a voltage that the characteristic has the appearance of a simple *p-n* diode.

The breakover voltage may be increased by applying reverse voltage to the cathode junction or equivalently by injecting a reverse current into the gate terminal. So long as firing is determined by the condition $\alpha_1 + \alpha_2 = 1$

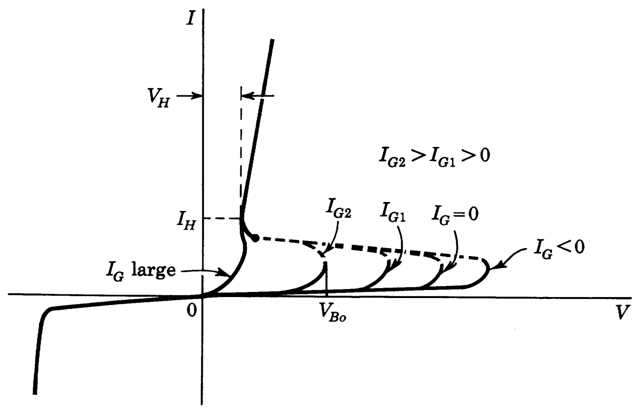


**Fig. 12-13** Volt-ampere characteristics of a three-terminal SCS illustrating that the forward breakover voltage is a function of the cathode-gate current. (Not drawn to scale.)



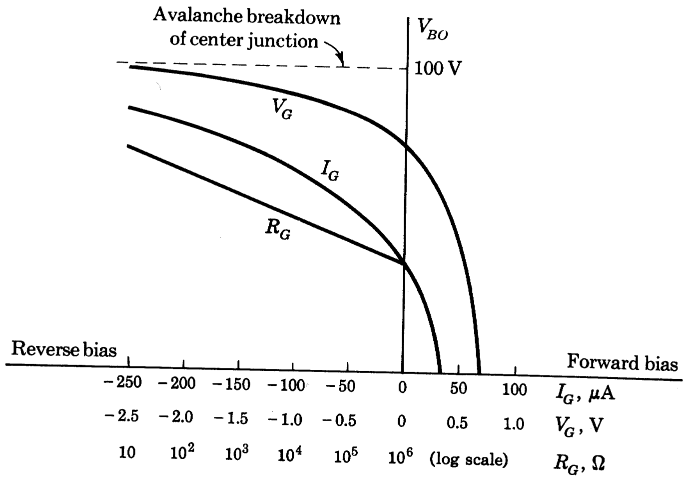


**Fig. 12-14** Typical variation of breakover voltage $V_{BO}$ in an SCS as a function of cathode-gate bias $V_G$, gate current $I_G$, and gate resistance $R_G$.

the breakover voltage $V_{BO}$ will be increased. Eventually, however, a point is reached where firing is the result of an avalanche breakdown of the center junction. In this case there is no additional increase in $V_{BO}$ even if the cathode junction is further reverse-biased. The form of the variation of breakover voltage for a typical SCS is sketched in Fig. 12-14 for three separate circumstances of bias provided by a voltage source $V_G$, a current source $I_G$, or a resistor $R_G$ connected from the cathode gate to the cathode. Note that very large values of $R_G$ are equivalent to $I_G = 0$, and very small values of $R_G$ to $V_G = 0$. These curves are temperature-dependent, the breakover voltage at fixed $I_G$ decreasing with an increase in temperature.

From Fig. 12-14 we see that in the forward-bias region $V_{BO}$ changes very rapidly with $V_G$ or $I_G$. Hence we may expect great variability in the firing voltage from unit to unit. For this reason the manufacturer does not supply curves of the sort indicated in Fig. 12-14. Instead he provides information concerning the forward voltage or current (as a function of temperature), which, even at an anode-to-cathode voltage of only a few volts, is "guaranteed to cause triggering in all units." For the SCS depicted in Fig. 12-14 this maximum gate firing signal is 30 μA or 0.6 V d-c or a pulse of this magnitude (with a minimum pulse width, as discussed in the following section).

Three commonly employed methods of providing bias to a silicon controlled switch are shown in Fig. 12-15. In Fig. 12-15*a* a resistor $R_G$ is connected from gate to cathode. In Fig. 12-15*b* the resistor is returned to a negative supply voltage to raise the firing voltage. The diode is employed to limit the possible back-biasing voltage across the cathode junction. This precaution is necessary since the maximum allowable reverse voltage at the

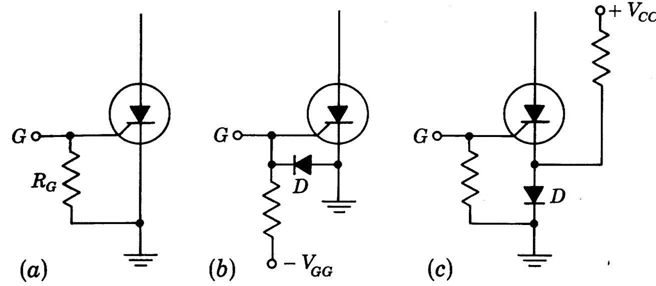


**Fig. 12-15 Biasing methods for the SCS. The switch is fired by applying a signal (either d-c or pulse) to gate input terminal $G$.**

$n$-emitter junction is often not more than a few volts. In Fig. 12-15$c$ the voltage drop across the forward-biased diode provides negative bias for the switch. The positive supply serves to provide current to keep the diode forward-biased.

Suppose that a supply voltage is applied through a load resistor between anode and cathode of a silicon controlled switch. Consider that the bias is such that the applied voltage is less than breakover voltage. Then the switch will remain OFF and may be turned ON by the application to the gate of a triggering current or voltage adequate to lower the breakover voltage to less than the applied voltage. The switch having been turned on, it latches, and it is found to be impractical to stop the conduction by reverse-biasing the gate. For example, it may well be that the reverse gate current for turn-off is nearly equal to the anode current. Ordinarily the most effective and commonly employed method for turn-off is, temporarily at least, to reduce the anode voltage below the holding voltage $V_H$ or equivalently to reduce the anode current below the holding current $I_H$. The gate will then again assume control of the breakover voltage of the switch.

## 12-7 SCS CHARACTERISTICS

Silicon controlled switches are ordinarily available in types that allow continuous anode currents up to about 1.0 A and maximum breakover voltages in the range from about 30 to 200 V.

The anode current which flows when reverse voltage is applied between anode and cathode is small, being of the order of several thousandths of a microampere in a unit with low leakage to about 1 $\mu$A in other units. The magnitude of this current is quite comparable to the current which flows when the anode-cathode voltage is in the forward direction but the switch is in an OFF state. Both these currents, forward and reverse, increase with temperature in a manner similar to the reverse saturation current in a transistor.



The ratio of the continuous allowable anode current to the forward gate current required to turn the switch ON even at low anode-to-cathode voltage is rarely less than several hundred, and in specially sensitive switches may attain values approaching 50,000. Thus, triggering currents of tens of microamperes may be enough to turn ON a switch which will then carry continuously hundreds of milliamperes. The impedance seen looking into the cathode-gate terminals is that of a forward-biased silicon diode. As we observe from Fig. 6-2, no appreciable diode current flows in a silicon diode except for voltages in excess of about 0.6 V. Accordingly, we find that required triggering voltages are usually of the order of 0.6 V. The firing current and voltage required decrease with increasing temperature.

The holding voltage at room temperature is approximately 1.0 V and the holding current lies in the range from less than a milliampere to several tens of milliamperes depending on the size of the unit. Both the holding voltage and holding current decrease with temperature. The incremental resistance between anode and cathode on the ON state is usually less than 1 Ω and may be as low as several tenths of an ohm. The holding current is affected by the gate bias. Increasing this bias increases the holding current because the more negative bias diverts out of the gate terminal some of the internal feedback current that the switch requires to maintain itself in the ON state.

**Rate Effect** Silicon controlled switches suffer from the same *rate effect* as do four-layer diodes (page 464). The inclination of a switch to fire prematurely because of the rate effect may be suppressed by operating the switch with a larger reverse bias on the gate and by reducing the resistance between gate and cathode. Both these measures, of course, reduce the sensitivity of the switch to an externally impressed triggering signal. An additional effective method is to bypass the gate to the cathode through a small capacitance. This component will shunt current past the cathode-gate junction in the presence of a rapidly varying applied voltage, but will have no effect on the d-c operation of the switch.

When both gate terminals $G_A$ and $G_C$ are externally available, the circuit[11] of Fig. 12-16 may be employed to suppress the rate effect entirely. In this circuit, the anode gate has been returned to the supply voltage through a resistor $R_{GA}$. The switch $S$ is not essential to the circuit and has been included only to facilitate the discussion to follow.

Assume initially that switch $S$ is open and that the SCS is OFF. The resistance $R_{GA}$ is large enough not to affect materially the voltages on the various layers of the silicon switch. The capacitor $C$ represents the capacitance across the center junction of the switch when this junction is reverse-biased. When a triggering signal is applied at the cathode gate $G_C$, the voltage at the anode $A$ drops, as does also the voltage at the anode gate $G_A$. The SCS may now be reset to the OFF state by closing switch $S$, since, with $S$ closed, the anode voltage and current are reduced below the holding values. As long as switch $S$ is closed the anode must remain at ground voltage even though the

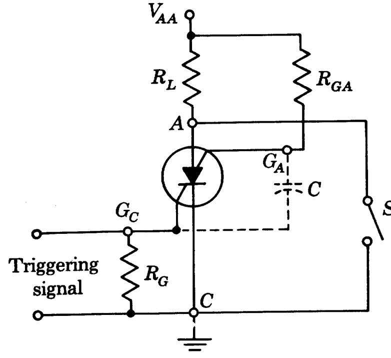


**Fig. 12-16 A silicon-controlled-switch circuit which eliminates the rate effect.**

SCS is OFF. The anode gate, however, is not so restrained and will begin to rise toward $V_{AA}$ with a time constant nominally equal to $R_{GA}C$ (for $R_G \ll R_{GA}$). The rising anode-gate voltage will charge the capacitor $C$. It will also reverse-bias the anode-to-anode-gate junction because $A$ is at ground potential and $G_A$ is rising toward $V_{AA}$.

If the switch $S$ is now opened, the anode voltage will rise abruptly, being limited in its speed $dv/dt$ only by whatever incidental capacitance may be present at the anode. If it were not for the fact that the anode gate $G_A$ is already at the supply voltage, this rapid rise in anode voltage might cause a premature firing. As noted in Sec. 12-5, premature firing results from the current which passes through the two forward-biased junctions and the capacitance $C$ of the reverse-biased junction. In the present case, however, this rapid rise in voltage is transmitted through a reverse-biased rather than a forward-biased anode junction. Consequently a smaller cathode junction current flows. Furthermore, so long as the anode junction is reverse-biased the negative-resistance *p-n-p-n* characteristic which is responsible for the switching cannot develop since forward-biased anode and cathode junctions are required. The net result is that, provided enough time is allowed after the closing of switch $S$ for the anode-gate voltage to rise, the switch is entirely free of rate effect at the anode. In a typical case $R_{GA} \approx 100$ K and $C \approx 5$ pF, so that the time constant is $R_{GA}C = 0.5$ μsec.

The connection of $R_{GA}$ to the supply $V_{AA}$, as in Fig. 12-16, will suppress the rate effect against a rapid rise in anode voltage such as would result from the opening of switch $S$. However, with switch $S$ open the rate effect might again make itself felt if a positive transient appears at the anode supply voltage. In this case the capacitor $C$ would restrain the anode gate from rising as rapidly as the anode, and the anode junction would not be reverse-biased. This situation may be corrected by simply returning the resistor $R_{GA}$ to a supply voltage higher than the anode supply.

**Gate ON and OFF Times** The process by which the SCS changes state occupies a finite time interval. When a triggering signal is applied to a gate to turn a switch ON, a time interval, the *turn-on* time, elapses before the transition is completed. This turn-on time decreases with increasing amplitude of trigger signal, increases with temperature, and increases also with increasing anode current. If the triggering signal is a pulse, then, to be effective, not only must the pulse amplitude be adequate but the pulse duration must be at least as long as a critical value called *gate time to hold*. Otherwise, at the termination of the gating pulse the SCS will fall back to its original state. A similar situation applies in driving the switch OFF by dropping the anode voltage. At a minimum, the anode voltage must drop below the holding voltage. If, however, the anode voltage is driven in the reverse direction the *turn-off* time may thereby be reduced. The turn-off time increases with temperature and with increasing magnitude of anode current. Further, the anode voltage must be kept below the maintaining voltage for an interval at least as long as a critical value called the *gate recovery time* if the transition is to persist after the anode voltage rises.



In fast units, all of the time intervals are in the range of tenths of microseconds, whereas in slower units these times may be as long as several microseconds. In general, the time required to turn a switch OFF is longer than the time required to turn it ON. The time intervals involved are required to allow for the establishment and dissipation of stored charge in the base regions, very much as in the case of the transistor. The matter of charge storage in transistors is discussed in detail in Chap. 20.

## 12-8 ADDITIONAL FOUR–LAYER DEVICES[12]

The *Binistor*, *Trigistor*, and *Transwitch* (commercial names) are *p-n-p-n* silicon devices which are in almost every respect identical to the SCS described above. They differ only in that, by design, they are more readily turned off by a negative triggering signal at the gate terminal. Their sensitivity to a turn-off trigger is an order of magnitude or so smaller than to a turn-on signal. Voltage ratings, temperature dependence, and switching-speed characteristics are not unlike those described for the SCS. Bistable operation of these switches results when the gate is driven by a train of alternate positive and negative pulses. When the speed attainable with these devices is adequate, they offer the great convenience of providing a bistable circuit which involves only a single active device, a minimum of components, and the feasibility of triggering to either state at one triggering input terminal.

The silicon controlled rectifier (SCR)[10] is a three-terminal silicon controlled switch with the principal difference that the rectifier is mechanically larger and designed to operate at higher currents and voltages. Currents in excess of 100 A and operating voltages approaching 1,000 V are possible. The rectifiers are used to control large amounts of power, whereas the switches are intended for low-level logic and switching applications. The switches are characterized not only by smaller voltage and current ratings but also by comparatively low leakage and holding currents. In addition, the switches are designed to require smaller triggering signals and have triggering characteristics which are more uniform from sample to sample of a given type.

## 12-9 THE THYRISTOR[13]

This germanium three-layer device is similar in construction to a mesa diffused-junction *p-n-p* transistor. The essential difference between a transistor and a thyristor is in the nature of the collector contact. In the thyristor the contact consists of a nickel tab soldered at high temperature to the germanium of the collector with an alloy of lead, tin, and indium. At low currents this combination behaves as an ohmic contact which serves simply to collect the holes that have diffused across the base and entered the collector region. The special nature of this contact is that at high currents it acts to inject electrons into the collector. Thus, at high currents the collector tab behaves like an *n*-type layer, making the transistor, in effect, a four-layer device. Typical collector characteristics for a thyristor are shown in Fig. 12-17. The dashed portions represent regions of negative resistance. Observe that the curves for low currents are those of a conventional *p-n-p* transistor, whereas for higher currents the characteristics are generally similar to the silicon controlled switch.

It turns out to be feasible to turn the thyristor OFF as well as ON by the application of a triggering signal at the base. The thyristor is like the SCS in that a larger triggering signal (by a factor of about 10) is required to turn it OFF than to turn it ON. The thyristor has the advantage of being faster. Turn-on times of 25 nsec and turn-off times of 100 nsec are not uncommon. Since the device material is germanium instead of silicon, the holding collector-to-emitter voltage ($\approx 0.3$ V) is smaller than in an SCS.

The name *thyristor* is intended to indicate a transistor replacement for the thyratron (the gas-filled hot-cathode triode). All the switches discussed above, however, are analogous to the thyratron and are replacing it in most applications. These semiconductor devices have the advantages of smaller triggering requirements, microsecond switching, no heater power, no vacuum seal, and smaller size. In addition, unlike the thyratron, some of the switches may be turned off with a gate signal. In terms of the concepts introduced in this chapter a thyratron may be described as a device possessing a current-controlled negative-resistance characteristic which latches upon the application of a positive pulse of voltage on its grid, but which can be turned off only by reducing its anode current below the value required to maintain the ionization.

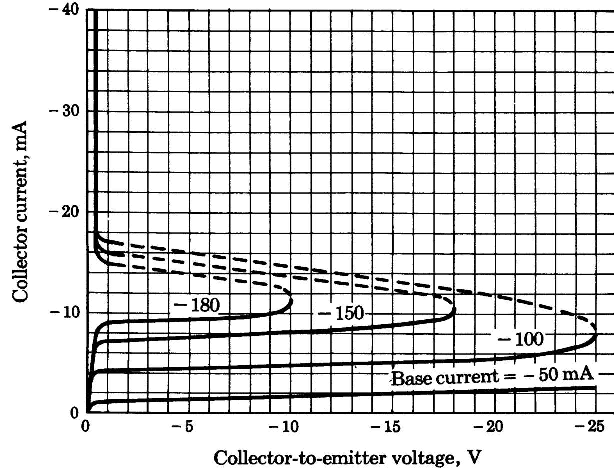


Fig. 12-17 Typical volt-ampere characteristics of a 2N1214 thyristor for various values of base current. (Courtesy of Radio Corporation of America.)



## 12-10 AVALANCHE–MODE OPERATION OF TRANSISTORS[14]

We observe in Figs. 6-18 and 6-20 that when the collector-to-emitter voltage of a transistor is raised to the level where avalanche breakdown occurs, the volt-ampere characteristic of the transistor exhibits a current-controlled negative-resistance characteristic. The similarity between Fig. 6-18 and Fig. 12-13 for the silicon controlled switch is readily apparent. Thus a transistor may be induced to make an abrupt switching transition from an OFF state to an ON state much in the manner of any of the other current-controlled devices we have considered in this chapter. There are two important differences between an avalanche-mode-operated transistor and the other switches. The first difference concerns the switching time. The OFF to ON switching time of the other switches is in the range of about 0.1 $\mu$sec (100 nsec), whereas the switching time for an avalanche transistor may be of the order of a few nanoseconds. The only competitors for speed with an avalanche transistor are the tunnel diode and snap-off diode (Sec. 20-7). The ON to OFF switching time of an avalanche transistor is quite long in comparison with the OFF to ON time, and such turn-off switching is not used when speed is at a premium. The normal procedure for turning OFF an avalanche transistor is to drop the collector voltage and, correspondingly, the collector current to a point where the current is no longer able to sustain the avalanche discharge, that is, to a point below the holding current.

The second important difference concerns the holding voltage of the avalanche transistor. In the four-layer switch the maintaining voltage is of the order of 1 V down to several tenths of a volt. Thus when the switch is conducting, the power dissipation may be quite small. In the avalanche transistor the holding voltage may be comparable to the breakdown voltage. Thus in a transistor biased to avalanche at, say, 100 V, the latching voltage may be 50 V. As a consequence, when the transistor is ON the dissipation may be comparatively large and thus limit the ON time of the transistor. For example, if an avalanche transistor is used to generate a train of pulses, and even if the pulse duration is short, being of the order of tens of nanoseconds, the allowable dissipation may severely limit the pulse repetition frequency.

Many transistors originally intended for other purposes will "avalanche," that is, respond with speed in avalanche-mode operation. Generally, those transistors do best which were designed for fast switching applications. In

addition, there are available a small number of transistor types such as the Texas Instruments 2N3033 which have been specifically designed for high-current, high-speed, avalanche-mode operation.

## REFERENCES


1. Esaki, L.: New Phenomenon in Narrow Ge *p-n* Junctions, *Phys. Rev.*, vol. 109, p. 603, 1958.
2. Millman, J.: "Electronic Devices and Models," McGraw-Hill Book Company, New York, in preparation.
3. "Tunnel Diode Manual, TD-30," Radio Corporation of America, Semiconductor and Materials Division, Somerville, N.J., 1963.
   "Tunnel Diode Manual," 1st ed., General Electric Company, Semiconductor Products Dept., Liverpool, N.Y., 1961.
4. MIL-STD-15-1A, May, 1963.
5. Sylvan, T. P.: *Appl. Note* 90.10, General Electric Company, Syracuse, N.Y., May, 1961.
6. Harding, M., and R. Windecker: Small Signal Planar PNPN Switch, Texas Instruments Company, Inc., August, 1963.
   Shockley, W., and J. F. Gibbons: Introduction to the 4-layer Diode, *Semicond. Prod.*, 1958.
   Shockley, W.: The Four Layer Diode, *Electron. Ind.*, vol. 16, no. 8, pp. 58–60, August, 1957.
   Gibbons, J. F.: A Critique of the Theory of *p-n-p-n* Devices, *IEEE Trans. Electron Devices*, vol. ED-11, no. 9, pp. 406–413, September, 1964.
7. Moll, J. L., M. Tannenbaum, J. M. Goldey, and N. Holonyak: *p-n-p-n* Switches, *Proc. IRE*, vol. 44, pp. 1174–1182, 1956.
   Sah, C. T., R. N. Noyce, and W. Shockley: Carrier Generation and Recombination in *p-n* Junctions and *p-n* Junction Characteristics, *Proc. IRE*, vol. 45, pp. 1228–1243, September, 1957.
8. Rate Effect, the Voltage-Current Characteristics of Four-layer Diodes at High Frequencies, *Appl. Data Note*, Shockley Transistor Corporation, May, 1959.
9. "Transistor Manual," 7th ed., chap. 16, General Electric Company, Syracuse, N.Y., 1964.
   Stasior, R. A.: Silicon Controlled Switches, *Appl. Note* 90.16, General Electric Company, Syracuse, N.Y., September, 1964.
10. "Silicon Controlled Rectifier Manual," 3d ed., General Electric Company, Auburn, N.Y., 1964.
11. Stasior, R. A.: How to Suppress Rate Effect in PNPN Devices, *Electronics*, vol. 37, no. 2, pp. 30–33, Jan. 10, 1964.





12. A Survey of Some Basic Trigistor Circuits, *Bull.* D410-02, Solid State Products, Inc., Salem, Mass., March, 1960.
    The Transwitch, Circuit Design Information and Application Notes, *Bull.* AN-1357A, March, 1960, and The Binistor, Circuit Design Information and Application Notes, *Bull.* AN-1360A, August, 1960, Transitron Electronic Corp., Wakefield, Mass.
13. RCA Thyristors, *Bull.* 1CE-208, Radio Corporation of America, Somerville, N.J., March, 1960.
    Mueller, C. W., and J. Hilibrand: The "Thyristor"—A New High-speed Switching Transistor, *IRE Trans. PGED*, vol. ED-5, pp. 2–5, January, 1958.
14. Henebry, W. M.: Avalanche Transistor Circuits, *Rev. Sci. Instr.*, vol. 32, no. 11, pp. 1198–1203, November, 1961.
    Miller, S. L., and J. J. Ebers: Alloyed Junction Avalanche Transistors, *Bell System Tech. J.*, vol. 34, pp. 883–902, September, 1955.