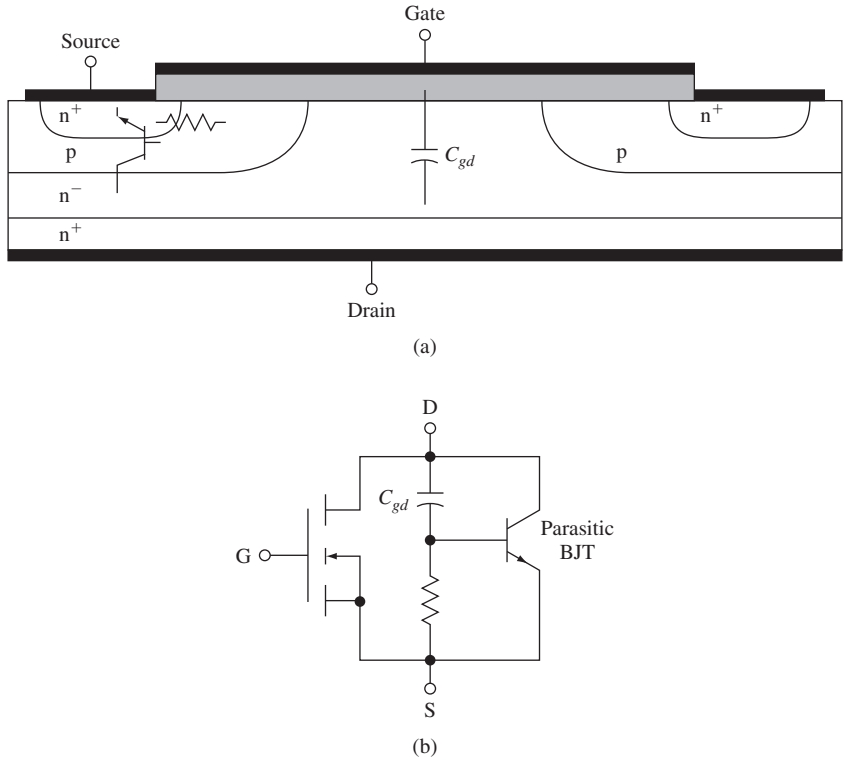


Figure 15.28 | (a) Cross section of vertical MOSFET showing parasitic BJT and distributed resistance; (b) equivalent circuit of MOSFET and parasitic BJT with distributed parameters.

The BJT should be cutoff at all times, which means the source-to-body voltage (emitter-to-base voltage) should be as close to zero as possible. We see from the geometries of Figures 15.20 and 15.21 that the source ohmic contact also goes across the p-type body region so that this junction voltage is zero during steady-state operation of the transistor. However, the BJT may be turned on during high-speed switching of the MOSFET.

Figure 15.28b shows that the base and the collector of the parasitic BJT are connected by the gate-to-drain capacitance. A parasitic or distributed resistance also connects the base to the emitter of the BJT. When the MOSFET is being turned off, the drain-to-source voltage increases and induces a current in the gate-to-drain capacitance in the direction from the parasitic collector terminal to the parasitic base terminal. This induced current may be large enough to induce a voltage in the parasitic resistance that is sufficient to forward bias the base–emitter junction and therefore turn the BJT on. The turned on BJT may then induce a large drain current that can cause burnout of the MOSFET. This breakdown mechanism is known as *snapback breakdown* and has been discussed briefly in Section 11.4.1.

The current–voltage characteristics are shown in Figure 11.22. Devices can be designed to minimize the parasitic or distributed base–emitter resistance to minimize this problem.

15.6 | THE THYRISTOR

One of the important applications of electronic devices is in switching between an off or blocking state to an on or low-impedance state. Thyristor is the name given to a general class of semiconductor pnpn switching devices that exhibit bistable regenerative switching characteristics. We have considered the transistor, which may be switched on with the application of a base drive or a gate voltage. The base drive or gate voltage must be applied as long as the transistor is to remain on. There are a number of applications in which it is useful to have a device remain in a blocking state until switched to the low-impedance state by a control signal, which then does not necessarily have to remain on. These devices are efficient in switching large currents at low frequencies, such as industrial control circuits operating at 60 Hz.

A **S**emiconductor **C**ontrolled **R**ectifier (SCR) is the common name given to a three-terminal thyristor. The SCR (sometimes referred to as a silicon controlled rectifier) is a four-layer pnpn structure with a gate control terminal. As with most semiconductor devices, there are several variations of the device structure. We consider the basic SCR operation and limitations, and then discuss some variations of the basic four-layer device.

15.6.1 The Basic Characteristics

The four-layer pnpn structure is shown in Figure 15.29a. The upper p region is called the anode and the lower n region is called the cathode. If a positive voltage is applied to the anode, the device is said to be forward biased. However, the junction J_2 is reverse biased so that only a very small current exists. If a negative voltage is applied to the anode, then junctions J_1 and J_3 are reverse biased—again only a very small current will exist. Figure 15.29b shows the I–V characteristics for these conditions. The voltage V_p is the breakdown voltage of the J_2 junction. For properly designed devices, the blocking voltage can be several thousand volts.

To consider the characteristics of the device as it goes into its conducting state, we can model the structure as coupled npn and pnp bipolar transistors. Figure 15.30a shows how we can split the four-layer structure and Figure 15.30b shows the two-transistor equivalent circuit with the associated currents. Since the base of the pnp device is the same as the collector of the npn transistor, the base current I_{B1} must in fact be the same as the collector current I_{C2}. Similarly, since the collector of the pnp transistor is the same as the base of the npn device, the collector current I_{C1} must be the same as the base current I_{B2}. In this bias configuration, the B–C of the pnp and the B–C of the npn devices are reverse biased, while the B–E junctions are both forward biased. The parameters α_1 and α_2 are the common base current gains of the pnp and npn transistors, respectively.

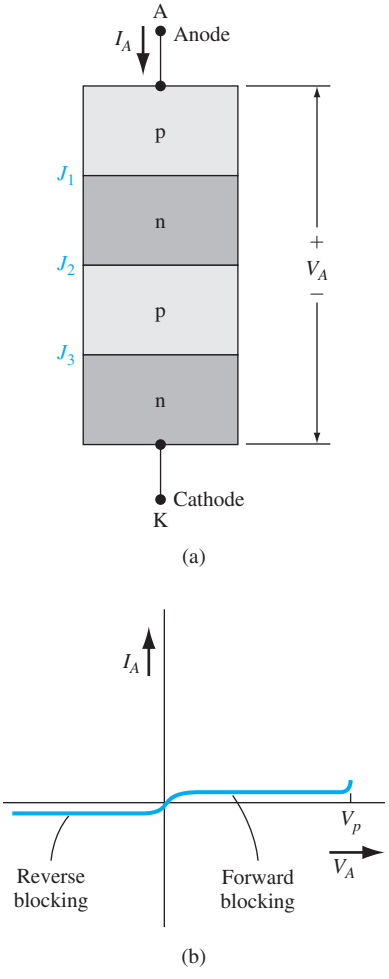


Figure 15.29 | (a) The basic four-layer pnpn structure. (b) The initial current–voltage characteristic of the pnpn device.

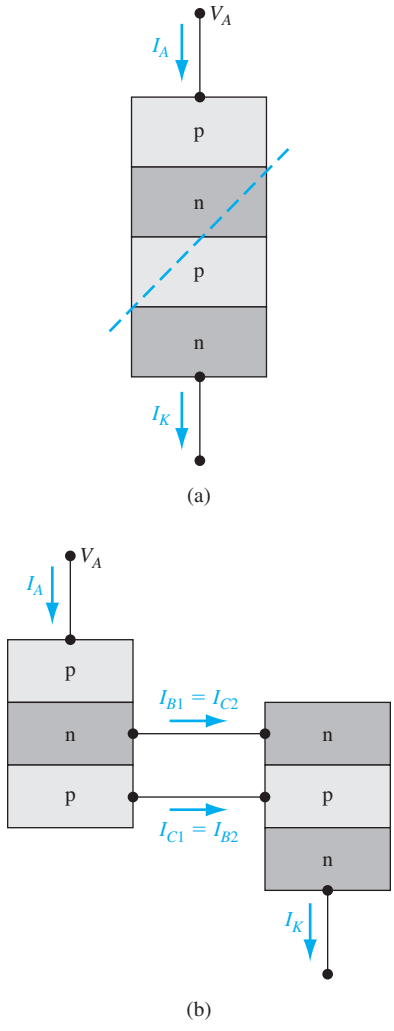


Figure 15.30 | (a) The splitting of the basic pnpn structure. (b) Two two-transistor equivalent circuit of the four-layer pnpn device.

We can write

$$I_{C1} = \alpha_1 I_A + I_{C01} = I_{B2} \tag{15.11a}$$

and

$$I_{C2} = \alpha_2 I_K + I_{C02} = I_{B1} \tag{15.11b}$$

where I_{C01} and I_{C02} are the reverse B–C junction saturation currents in the two devices. In this particular configuration, $I_A = I_K$ and $I_{C1} + I_{C2} = I_A$. If we add Equations (15.11a) and (15.11b), we obtain

$$I_{C1} + I_{C2} = I_A = (\alpha_1 + \alpha_2)\, I_A + I_{C01} + I_{C02} \tag{15.12}$$

The anode current I_A, from Equation (15.12), can be found as

$$I_A = \frac{I_{C01} + I_{C02}}{1 - (\alpha_1 + \alpha_2)} \tag{15.13}$$

As long as $(\alpha_1 + \alpha_2)$ is much smaller than unity, the anode current is small, as we have indicated in Figure 15.29b.

The common base current gains, α_1 and α_2, are very strong functions of collector current as we discussed in Chapter 12. For small values of V_A, the collector current in each device is just the reverse saturation current, which is very small. The small collector current implies that both α_1 and α_2 are much smaller than unity. The four-layer structure maintains this blocking condition until the junction J_2 starts into breakdown or until a current is induced in the J_2 junction by some external means.

Consider, initially, the condition when the applied anode voltage is sufficiently large to cause the J_2 junction to start into avalanche breakdown. This effect is shown in Figure 15.31a. The electrons generated by impact ionization are swept into the n_1 region, making the n_1 region more negative, and the holes generated by impact ionization are swept into the p_2 region, making the p_2 region more positive. The more negative voltage of the n_1 region and the more positive voltage of the p_2 region means that the forward-bias junction voltages V_1 and V_3 both increase. The increase in the respective B–E junction voltages causes an increase in current, which results in an increase in the common-base current gains α_1 and α_2, causing a further increase in

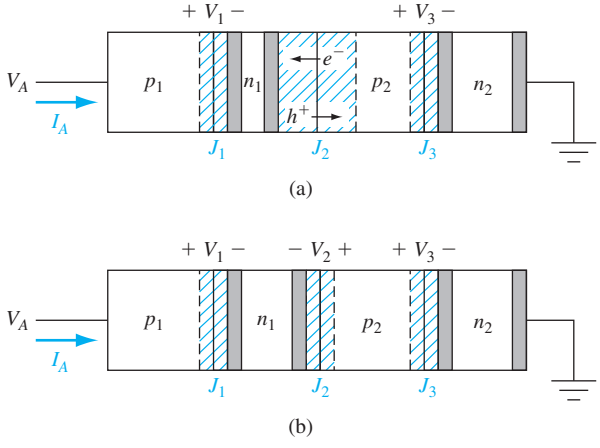


Figure 15.31 | (a) The pnpn device when the J_2 junction starts into avalanche breakdown. (b) The junction voltages in the pnpn structure when the device is in the high-current, low-impedance state.

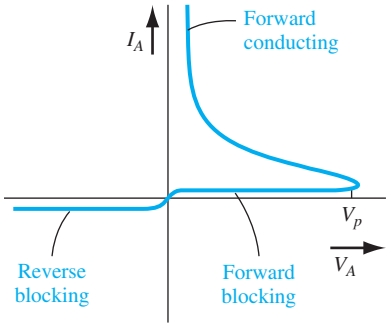


Figure 15.32 | The current–voltage characteristics of the pnpn device.

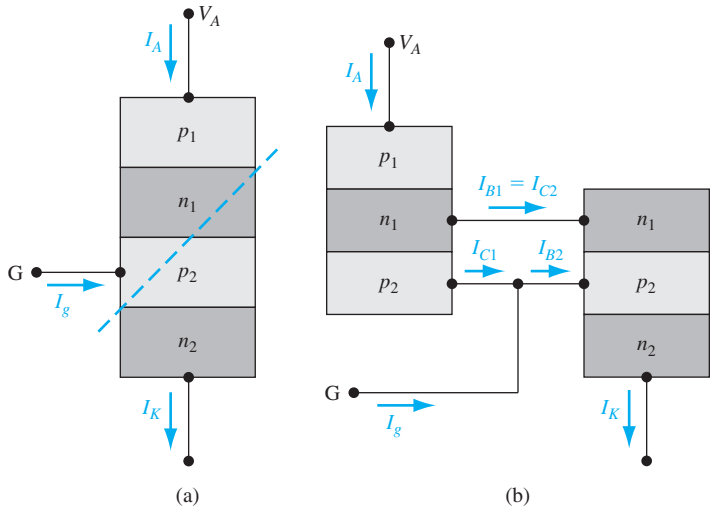


Figure 15.33 | (a) The three-terminal SCR. (b) The two-transistor equivalent circuit of the three-terminal SCR.

I_A as seen in Equation (15.13). We now have a regenerative positive feedback situation, so the current I_A will increase very rapidly.

As the anode current I_A increases and $\alpha_1 + \alpha_2$ increases, the two equivalent bipolar transistors are driven into saturation and the junction J_2 becomes forward biased. The total voltage across the device decreases and is approximately equal to one diode drop as shown in Figure 15.31b. The current in the device is limited by the external circuit. If the current is allowed to increase, ohmic losses may become important so that the voltage drop across the device may increase slightly with current. The I_A versus V_A characteristic is shown in Figure 15.32.

15.6.2 Triggering the SCR

In the last section, we considered the case when the four-layer pnpn device is turned on by the avalanche breakdown process in the center junction. The turn-on condition can also be initiated by other means. Figure 15.33a shows three-terminal SCR in which the third terminal is the gate control. We can determine the effect of the gate current by reconsidering Equations (15.11a) and (15.11b).

Figure 15.33b again shows the two-transistor equivalent circuit including the gate current. We can write

$$I_{C1} = \alpha_1 I_A + I_{C01} \tag{15.14a}$$

and

$$I_{C2} = \alpha_2 I_K + I_{C02} \tag{15.14b}$$

We now have $I_K = I_A + I_g$ and we can still write $I_{C1} + I_{C2} = I_A$. Adding Equations (15.14a) and (15.14b), we find that

$$I_{C1} + I_{C2} = I_A = (\alpha_1 + \alpha_2)I_A + \alpha_2 I_g + I_{C01} + I_{C02} \tag{15.15}$$

Solving for I_A, we find

$$I_A = \frac{\alpha_2 I_g + (I_{C01} + I_{C02})}{1 - (\alpha_1 + \alpha_2)} \tag{15.16}$$

We can think of the gate current as the flow of holes into the p_2 region. The additional holes increase the potential of this region, which increases the forward-biased B–E voltage of the npn bipolar transistor, and the transistor action. The transistor action of the npn increases the collector current I_{C2}, which starts the transistor action of the pnp bipolar transistor, and the entire pnpn device can be turned on into its low-impedance state. The gate current required to switch the SCR into its on condition is typically in the milliamp range. SCR can be turned on with a small gate current, which can control hundreds of amperes of anode current. The gate current can be turned off and the SCR will remain in its conducting state. The gate loses control of the device once the SCR is triggered into its conducting state. The current–voltage characteristics of the SCR as a function of gate current is shown in Figure 15.34.

A simple application of an SCR in a half-wave control circuit is shown in Figure 15.35a. The input signal is an ac voltage and a trigger pulse will control the turn-on of the SCR. We assume that the trigger pulse occurs at time t_1 during the ac voltage cycle. Prior to t_1, the SCR is off so that the current in the load is zero; thus, there is a zero output voltage. At $t = t_1$, the SCR is triggered on and the input voltage appears across the load (neglecting the voltage drop across the SCR). The SCR turns off when the anode-to-cathode voltage becomes zero even though the trigger pulse has been turned off prior to this time. The time at which the SCR is triggered during the voltage cycle can be varied, changing the amount of power delivered to the load. Full-wave control circuits can be designed to increase efficiency and degree of control.

The gate allows control of the turn-on of the SCR. However, the four-layer pnpn structure can also be triggered on by other means. In many integrated circuits, parasitic pnpn structures exist. One such example is the CMOS structure that we considered in Chapter 10. A transient ionizing radiation pulse can trigger the parasitic

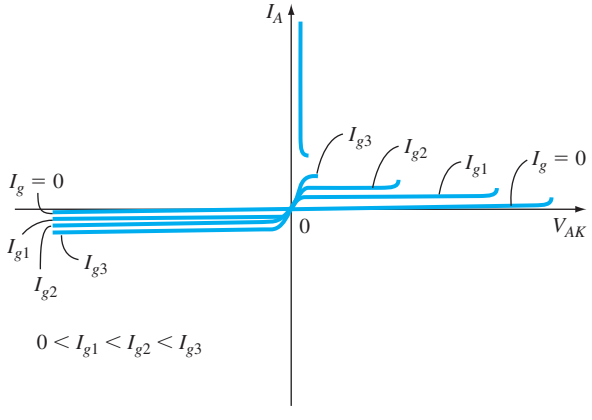


Figure 15.34 | Current–voltage characteristics of an SCR.

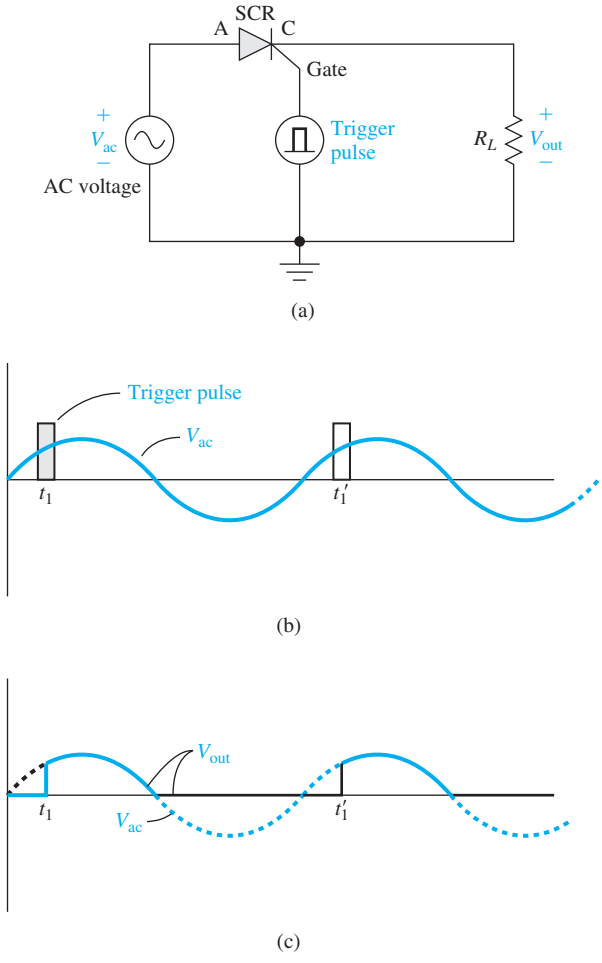


Figure 15.35 | (a) Simple SCR circuit. (b) Input ac voltage signal and trigger pulse. (c) Output voltage versus time.

four-layer device by generating electron–hole pairs, particularly in the J_2 junction, producing a photocurrent. The photocurrent is equivalent to a gate current in an SCR so the parasitic device can be switched into its conducting state. Again, once the device is switched on, it will remain in its conducting state even when the radiation ceases. An optical signal can also trigger the device in the same manner by generating electron–hole pairs.

Another triggering mechanism in the pnpn device is dV/dt triggering. If the forward-bias anode voltage is applied rapidly, the voltage across the J_2 junction will also change quickly. This changing reverse-biased J_2 junction voltage means that the space charge region width is increasing; thus, electrons are being removed from the n_1 side of the junction and holes are being removed from the p_2 side of the junction.

If dV/dt is large, the rate of removal of these carriers is rapid, which leads to a large transient current that is equivalent to a gate current and can trigger the device into a low-impedance conducting state. In SCR devices, a dV/dt rating is usually specified. However, in parasitic pnpn structures, the dV/dt triggering mechanism is a potential problem.

15.6.3 SCR Turn-Off

Switching the four-layer pnpn structure from its conducting state to its blocking state can be accomplished if the current I_A is reduced below the value creating the $\alpha_1 + \alpha_2 = 1$ condition. This critical I_A current is called the holding current. If a parasitic four-layer structure is triggered into the conducting state, the effective anode current in the device must be reduced below the corresponding holding current in order to turn off the device. This requirement essentially implies that all power supplies must be turned off in order to bring the parasitic device back into its blocking state.

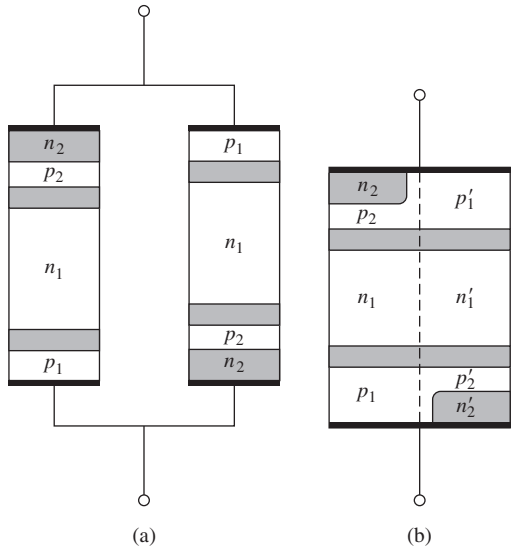
The SCR can be triggered on by supplying holes to the p_2 region of the device. The SCR can perhaps be turned off by removing holes from this same region. If the reverse gate current is large enough to bring the npn bipolar transistor out of saturation, then the SCR can be switched from the conducting state into the blocking state. However, the lateral dimensions of the device may be large enough so that nonuniform biasing in the J_2 and J_3 junctions occurs during a negative gate current and the device will remain in the low-impedance conducting state. The four-layer pnpn device must be specifically designed for a turn-off capability.

15.6.4 Device Structures

Many thyristor structures have been fabricated with specific characteristics for specific applications. We consider a few of these types of device to gain an appreciation for the variety of structures.

Basic SCR There are many variations of diffusion, implantation, and epitaxial growth that can be used in the fabrication of the SCR device. The basic structure is shown in Figure 15.36. The p_1 and p_2 regions are diffused into a fairly high resistivity n_1 material. The n^+ cathode is formed and the p^+ gate contact is made. High thermal conductivity materials can be used for the anode and cathode ohmic contacts to aid in heat dissipation for high-power devices. The n_1 region width may be on the order of 250 μm in order to support very large reverse-biased voltages across the J_2 junction. The p_1 and p_2 regions may be on the order of 75 μm wide, while the n^+ and p^+ regions are normally quite thin.

Bilateral Thyristor Since thyristors are often used in ac power applications, it may be useful to have a device that switches symmetrically in the positive and negative cycles of the ac voltage. There are a number of such devices, but the basic concept is to connect two conventional thyristors in antiparallel as shown in Figure 15.37a. The integration of this concept into a single device is shown in Figure 15.37b. Symmetrical n regions can be diffused into a pnp structure. Figure 15.37c shows the current–voltage characteristics in which the triggering into the conduction mode



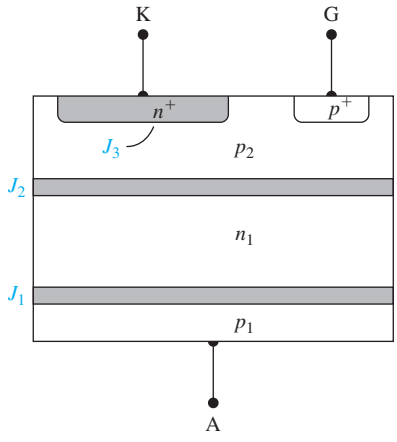


Figure 15.36 | The basic SCR device structure.

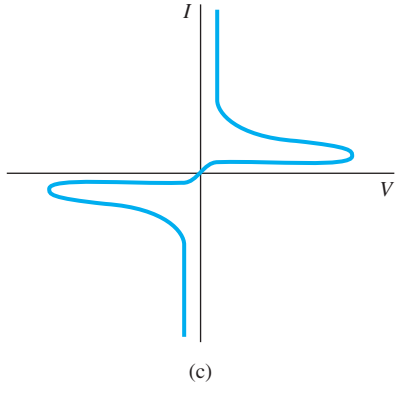


Figure 15.37 | (a) The antiparallel connection of two thyristors to form a bilateral device. (b) The bilateral thyristor as an integrated device. (c) The current–voltage characteristics of the bilateral thyristor. *(From Ghandhi [7].)*

would be due to breakdown triggering. The two terminals alternately share the role of anode and cathode during successive half cycles of the ac voltage.

Triggering by a gate control is more complex for this device since a single gate region must serve for both of the antiparallel thyristors. One such device is known as a *triac*. Figure 15.38a shows the cross section of such a device. This device can be triggered into conduction by gate signals of either polarity and with anode-to-cathode voltages of either polarity.

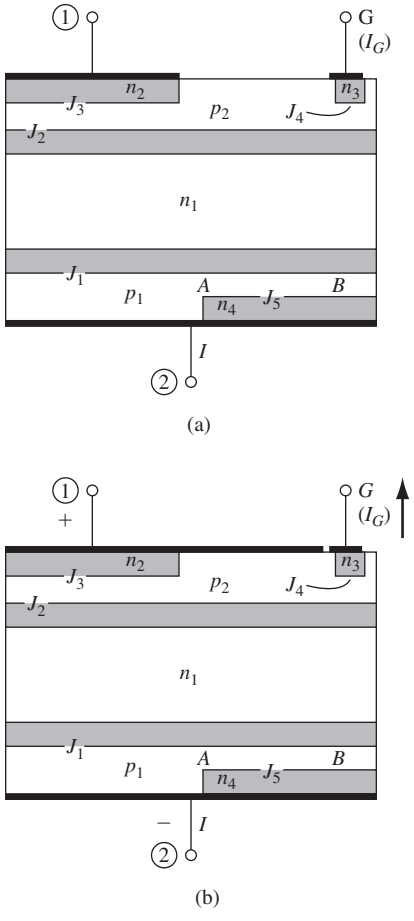


Figure 15.38 | (a) The triac device. (b) The triac with a specific bias configuration. *(From Ghandhi [7].)*

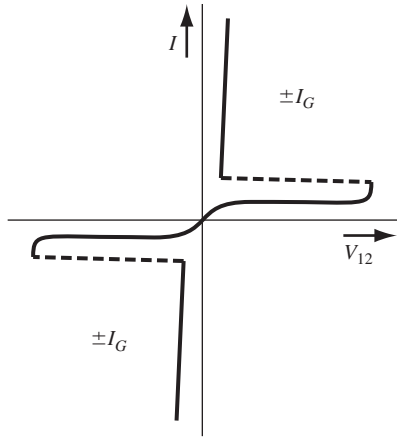


Figure 15.39 | The current–voltage characteristics of the triac.

One particular gate control situation is shown in Figure 15.38b. Terminal 1 is positive with respect to terminal 2, and a negative gate voltage is applied with respect to terminal 1, so the gate current is negative. This polarity arrangement induces the current I_1 and the junction J_4 becomes forward biased. Electrons are injected from n_3, diffuse across p_2, and are collected in the n_1 region. In this case $n_3p_2n_1$ behaves like a saturated transistor. The collected electrons in n_1 lower the potential of n_1 with respect to p_2. The current across the p_2n_1 junction increases, which can trigger the $p_2n_1p_1n_4$ thyristor into its conducting mode.

We can show that the other combinations of gate, anode, and cathode voltages will also trigger the triac into conduction. Figure 15.39 shows the terminal characteristics.

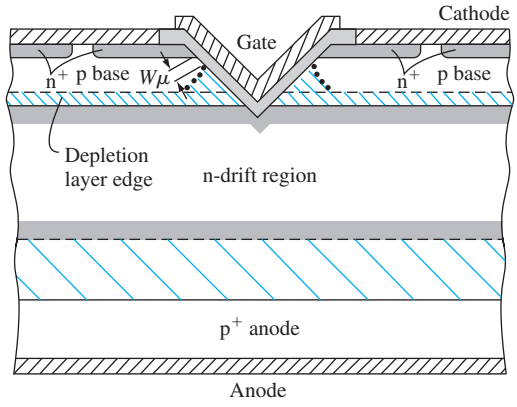


Figure 15.40 | The V groove MOS gated thyristor.
(From Baliga [1].)

MOS Gated Thyristor The operation of a MOS gated thyristor is based upon controlling the gain of the npn bipolar transistor. Figure 15.40 shows a V-groove MOS gated thyristor. The MOS gate structure must extend into the n-drift region. If the gate voltage is zero, the depletion edge in the p-base remains essentially flat and parallel to the junction J_2; the gain of the npn transistor is low. This effect is shown in the figure by the dashed line. When a positive gate voltage is applied, the surface of the p base becomes depleted—the depletion region in the p base adjacent to the gate is shown by the dotted line. The undepleted base width W_μ of the npn bipolar device narrows and the gain of the device increases.

At a gate voltage approximately equal to the threshold voltage, electrons from the n^+ emitter are injected through the depletion region into the n-drift region. The potential of the n-drift region is lowered, which further forward biases the p^+ anode to n-drift junction voltage, and the regenerative process is initiated. The gate voltage required to initiate turn-on is approximately the threshold voltage of the MOS device. One advantage of this device is that the input impedance to the control terminal is very high; relatively large currents can be switched with very small capacity coupled gate currents.

MOS Turn-Off Thyristor The MOS turn-off thyristor can both turn on and turn off the anode current by applying a signal to a MOS gate terminal. The basic device structure is shown in Figure 15.41. By applying a positive gate voltage, the n^+pn bipolar transistor can be turned on as just discussed. Once the thyristor is turned on, the device can be turned off by applying a negative gate voltage: the negative gate voltage turns on the p-channel MOS transistor that effectively short circuits the B–E junction of the n^+pn bipolar transistor. Holes that now enter the p-base have an alternative path to the cathode. If the resistance of the p-channel MOSFET becomes low enough, all current will be diverted away from the n^+p emitter and the n^+pn device will effectively be turned off.

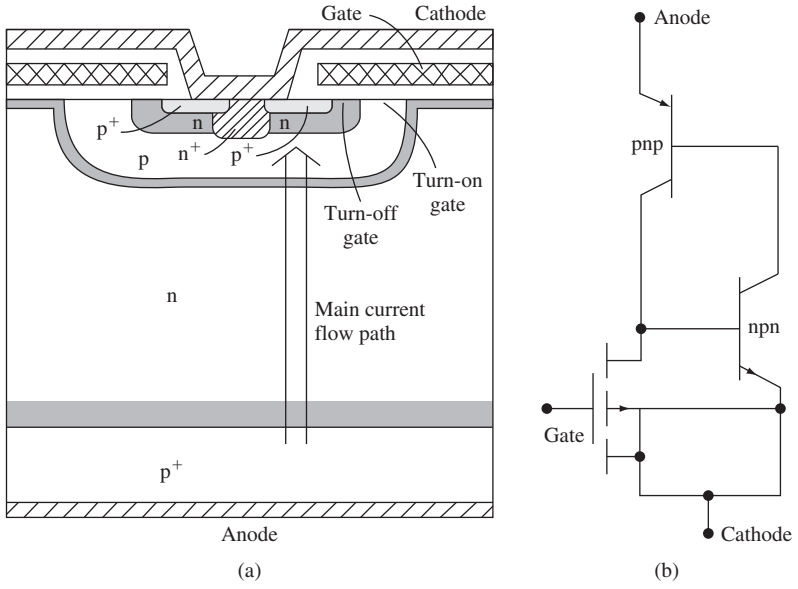


Figure 15.41 | (a) The MOS turn-off thyristor. (b) Equivalent circuit for the MOS turn-off thyristor.
(From Baliga [1].)

15.7 | SUMMARY

- The concept of a negative differential resistance in the *I–V* characteristic of the tunnel diode is used in the design of a microwave tunnel diode oscillator. The expression for the maximum resistance cutoff frequency is derived.
- The operation of a microwave GUNN diode oscillator is based on the concept of negative differential mobility.
- The IMPATT diode oscillator uses injection and drift time delays to create a region of differential negative resistance.
- The power BJT has a vertical configuration and an interdigitated base–emitter surface structure. The collector drift region (doping and width) determines the rated blocking voltage of the BJT, while the base width must be sufficiently wide to avoid punch-through breakdown at the rated blocking voltage.
- A power BJT is characterized by the maximum rated collector current, maximum rated voltage, and maximum rated power dissipation. These three parameters define the SOA of the transistor.
- A power MOSFET has a vertical configuration and an interdigitated gate–source surface structure. Two specific devices considered are the DMOS and VMOS structures. The drain-drift region (doping and width) determines the rated blocking voltage of the MOSFET, while the channel length of the base (body) must be sufficiently wide to avoid punch-through breakdown at the rated blocking voltage.
- A power MOSFET is characterized by the maximum rated drain current, maximum rated voltage, and maximum rated power dissipation. These three parameters define the SOA of the transistor.

- The "on resistance" of a MOSFET has a positive temperature coefficient so that the power MOSFET is more stable versus temperature than a power BJT. This characteristic allows MOSFETs to be fabricated in parallel to increase the current capability of the device.
- The thyristor refers to a general class of pnpn switching devices that can be switched between a high-impedance, low-current state and a low-impedance, high-current state. These devices exhibit a bistable regenerative positive feedback switching characteristic.
- The basic pnpn device can be modeled as coupled npn and pnp bipolar transistors. In the "on" state, both bipolar transistors are driven into saturation, creating the high-current, low-voltage condition. In the "off" or blocking state, large voltages can be applied to the device and the current is essentially zero.
- The turn-on characteristics of the thyristor can be controlled through a gate control terminal. The three-terminal thyristors are referred to as semiconductor controlled rectifiers (SCRs).

GLOSSARY OF IMPORTANT TERMS

double-diffused MOSFET (DMOS) A power MOSFET in which the source and channel regions are formed using a double diffusion process.

HEXFET The structure of a power MOSFET in which many individual MOSFETs are placed in parallel in a hexagonal configuration.

maximum rated current The maximum allowed current in a power transistor such that proper operation is maintained.

maximum rated power The maximum allowed power dissipation in a power transistor such that no permanent damage is done to the transistor.

maximum rated voltage The maximum allowed applied voltage to a power transistor such that breakdown is not initiated.

negative differential mobility A region in the drift velocity versus electric field characteristic of a semiconductor material in which the drift velocity decreases with an increase in the electric field.

negative differential resistance A region in the *I*–*V* characteristic of a device in which the current decreases while the voltage increases.

on resistance The effective resistance between source and drain of a power MOSFET.

safe operating area The allowed current–voltage regions of operation for a power transistor bounded by the maximum rated current, maximum rated voltage, and maximum power.

second breakdown A breakdown effect in a power BJT in which high temperature causes a thermal runaway process.

SCR (semiconductor controlled rectifier) The common name given to a three-terminal thyristor.

thyristor The name given to a general class of semiconductor pnpn switching devices exhibiting bistable regenerative switching characteristics.

transferred-electron effect The phenomenon in which conduction electrons are scattered from a lower energy, high-mobility band into a higher energy, low-mobility band.

triac The name of a bilateral three-terminal thyristor.

V-groove MOSFET (VMOS) A power MOSFET in which the channel region is formed along a V-shaped groove formed in the surface of the semiconductor.

CHECKPOINT

After studying this chapter, the reader should have the ability to:

- Explain how a region of negative differential resistance is developed in the *I–V* characteristic of the tunnel diode.
- Discuss the concept of negative differential mobility in GaAs and discuss how this phenomenon leads to the generation of domains in a GUNN diode.
- Discuss the operation of an IMPATT diode oscillator.
- Sketch the cross section of a power BJT and discuss the voltage and current limitations of the device.
- Discuss the reason the current gain of a power BJT is generally smaller than that of a small switching BJT.
- Sketch the safe operating area of a power BJT.
- Describe the reason for and the operation of a Darlington configuration.
- Sketch the cross section of the DMOS and VMOS power MOSFET structures.
- Sketch the safe operating area of a power MOSFET.
- Describe why the "on resistance" of a power MOSFET has a positive temperature coefficient.
- Describe the switching characteristics of a pnpn device.
- Describe the switching characteristics of a semiconductor controlled rectifier.

REVIEW QUESTIONS

1. Describe how a negative differential resistance region in the *I–V* characteristic of the tunnel diode is generated.
2. Describe how a negative differential mobility region in the drift velocity versus electric field characteristic in GaAs is developed.
3. Describe how a negative differential resistance characteristic is produced in the IMPATT diode.
4. Why is the doping concentration in the collector drift region low and why is the drift region width large in a power BJT?
5. Why does a power BJT have an interdigitated base–emitter structure?
6. Sketch the safe operating area of a power BJT.
7. Discuss how a DMOS structure of a power MOSFET is formed.
8. Discuss the voltage limitation of a power MOSFET.
9. Define the "on resistance" of a power MOSFET and show that the on resistance has a positive temperature coefficient.
10. Discuss how the gate terminal of a semiconductor controlled rectifier can control the switching characteristics.

PROBLEMS

Section 15.1 Tunnel Diode

15.1 Sketch the energy band diagrams of a tunnel diode in which both the n and p regions are degenerately doped for the case of (*a*) zero bias, (*b*) $0 < V < V_p$, (*c*) $V_p < V < V_v$, and (*d*) $V > V_v$.