POWER DEVICES

RECTIFIERS

The semiconductor rectifier (1) was the first semiconductor device developed for power circuit applications. It is a semiconductor device specifically designed to rectify alternating current; that is, to exhibit a very low resistance to current flow in one direction and a very high resistance in the other direction.

P–i–N Rectifier

In the past, only *P–i–N* rectifiers were available for use in power circuits. The first devices were made by using germanium. However, the high leakage current in germanium devices associated with its small energy bandgap (0.66 eV) led to their replacement by silicon *P–i–N* rectifiers. Since the 1950s, the performance of silicon *P–i–N* rectifiers has been continually improving due to the optimization of the device structure and lifetime control (2) that is used to adjust the switching speed. The basic *P–i–N* rectifier structure is shown in Fig. 1. The doping concentration and the thickness of the *i*-region (N^- region) are designed to support the required reverse blocking voltage. In order to support large reverse blocking voltage, it is necessary to use a low doping concentration as well as a large thickness for the *i*-region. In the on-state, minority carrier holes are injected from the P^+ anode and electrons injected from the N^+ cathode; they are equal in number to maintain charge neutrality. This phenomenon of injecting a high concentration of holes and electrons into the *i*-region, called conductivity modulation, is an extremely important effect that allows the transporting of high currents through the *P–i–N* rectifiers with breakdown voltages of up to 5000 V. In the steady-state reverse blocking state, the reverse current is only the leakage current due to minority carrier generation in the depletion region, and the *P–i–N* rectifier exhibits a very high resistance to current flow.

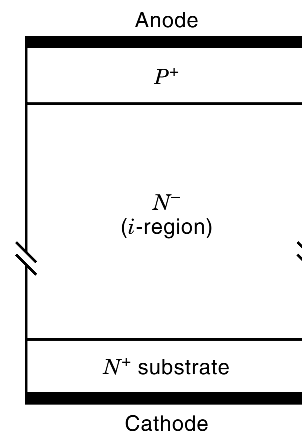

Figure 1. Basic structure of the *P–i–N* rectifier. The lightly doped thick *i*-region is designed to support the required reverse voltage. The heavily doped p^+ anode and n^+ cathode inject carriers into the *i*-region to modulate its conductivity during forward conduction.

As described, the injection of a high concentration of minority carriers into the *i*-region can increase the conductivity of the *i*-region. However, the injected carriers also create problems during switching of the *P–i–N* rectifier. When the voltage across the device reverses polarity, the injected electron-hole population (called stored charge) must be removed before the formation of a depletion region can occur to support the reverse blocking voltage. This leads to a reverse recovery current. The presence of such a reverse recovery transient leads to power dissipation that limits the maximum switching frequency of *P–i–N* rectifiers and degrades the reliability of the applied circuits, an additional concern is the large voltage overshoot caused by the di/dt of the reverse recovery current flowing through the stray inductance in the circuit. When the switching frequency of a power circuit increases, the turn-off di/dt must be increased. It has been found that this causes an increase in both the peak reverse recovery current and the ensuing reverse recovery di/dt. If this reverse recovery di/dt is large, an increase in the breakdown voltage of all the circuit components becomes essential. Raising the breakdown voltage capability causes an increase in the forward voltage drop of power devices, which in turn degrades the system efficiency because of a higher conduction loss. Many methods of lifetime control have been developed to reduce the minority carrier lifetime in the *i*-region to decrease the switching loss

J. Webster (ed.), *Wiley Encyclopedia of Electrical and Electronics Engineering*. Copyright © 1999 John Wiley & Sons, Inc.

of the P–i–N rectifier, but they also lead to an increase in the on-state voltage drop. It is therefore customary to perform a trade-off between on-state and turn-off losses when designing P–i–N rectifiers.

Another drawback of the P–i–N rectifier is the forward voltage overshoot during its turn-on. The forward voltage overshoot in a P–i–N rectifier arises from the existence of the high resistance i-region. Under steady-state current conduction, the i-region resistance is drastically reduced by hole–electron population injected by the N^+ and the P^+ regions. However, during turn-on under high di/dt condition, the current rises at a faster rate than the diffusion of the minority carriers injected from the junction. A high voltage drop develops across the i-region for a short period of time until the minority carriers can swamp out the i-region resistance.

Schottky Barrier Diode

In order to eliminate the reverse recovery problem associated with P–i–N rectifiers, the Schottky barrier diode (SBD, Schottky rectifier) was developed in the 1970s. The basic structure of the SBD is shown in Fig. 2. It consists of a metal-semiconductor rectifying contact with an N-region designed to support the required reverse blocking voltage. When a positive bias is applied to the metal with respect to the N-type semiconductor, forward conduction in the SBD occurs by thermal emission of majority carrier electrons across a lowered metal-semiconductor barrier. The on-state voltage drop of the SBD therefore consists of the sum of the voltage drop across the barrier and the ohmic voltage drop across the N-region. There is no conductivity modulation of the N-region in the SBD structure, because the minority carrier injection is negligible. The N-region resistance depends on the reverse blocking voltage. For low breakdown voltages (<100 V), the doping concentration of the N-region lies between 5×10^{15} and 1×10^{16} cm^{-3} and its thickness can be made to less than 10 μm. This leads to a relatively small voltage drop in the N-region. For the case of a typical SBD with a reverse breakdown voltage of 50 V and a barrier height of 0.8 V, the forward voltage drop is about 0.554 V at a forward conduction current density of 100 A/cm^2. This lower forward voltage drop compared to a P–i–N rectifier (whose forward voltage drop at 100 A/cm^2 is about 0.9 V) and its faster switching speed (because of the absence of large reverse recovery current observed in a P–i–N rectifier) make the SBD attractive in low-voltage switching applications.

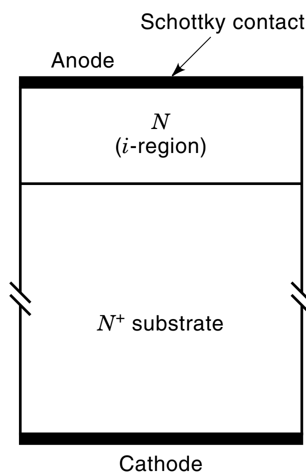

Figure 2. Basic structure of the Schottky barrier diode. A specially selected anode metalization (such as tungsten, aluminum) forms a Schottky barrier with the lightly doped N-region.

With the increase of the reverse blocking capability, the forward voltage drop of the SBD will increase rapidly and will approach that of a P–i–N rectifier when the reverse blocking capability is increased to 200 V. In addition, the Schottky rectifier has a larger reverse leakage current, which also increases with the increase of the temperature, and a soft blocking characteristic. These make the silicon Schottky rectifier generally unacceptable for use in high-voltage applications.

As discussed previously, the silicon Schottky rectifier eliminates the reverse recovery problem that limits its high-frequency application, but the forward voltage drop of the silicon Schottky rectifier increases rapidly with the increase of the reverse blocking capability. A much superior power rectifier can be created by using SBD contact if the resistance of the N-region can be reduced while achieving the same blocking capability. An approach for achieving this is to replace silicon with a wide-band gap semiconductor. Based on this fundamental analysis, it was demonstrated that a gallium arsenide (GaAs) Schottky rectifier has a better forward voltage drop than a silicon P–i–N diode for blocking voltages of up to 500 V.

The mobility of electrons in gallium arsenide at low field is larger than in silicon by a factor of 5.6. In addition, because of the larger energy band gap, the critical electric field for breakdown in gallium arsenide is higher than in silicon. These two facts result in a reduction in the specific on-resistance (the on-resistance per unit area) of the N-region by a factor of 13. For a typical Schottky barrier height of 0.8 V, the gallium arsenide Schottky rectifiers are expected to have a lower forward drop than silicon P–i–N rectifiers for breakdown voltages of up to about 500 V at a typical operating current density of 100 to 200 A/cm^2. In this voltage range, the gallium arsenide Schottky rectifiers offer a clear advantage over P–i–N rectifiers due to faster switching speed because of the absence of the reverse recovery current. The fabrication of gallium arsenide Schottky barrier power rectifiers can be accomplished by using aluminum or titanium Schottky barrier contacts. These devices are now commercially available.

An even more promising material for power rectifiers is silicon carbide (SiC) because of its much higher critical electric field at breakdown. The on-state voltage drop of the silicon carbide Schottky barrier power rectifier is superior to that of the silicon P–i–N power rectifier for a blocking voltage up to 3000 V. The silicon carbide Schottky barrier power rectifiers have also been found to have excellent reverse recovery and reverse bias leakage characteristics even at high operating temperatures. They are likely to replace silicon P–i–N rectifiers in high-voltage power electronic circuits in the next decade.

JBS Rectifiers

With the trend toward lower operating voltages for very large scale integration (VLSI) chips, there is an increasing demand to reduce the forward voltage drop in rectifiers. The forward voltage drop of a Schottky rectifier can be reduced by decreas-

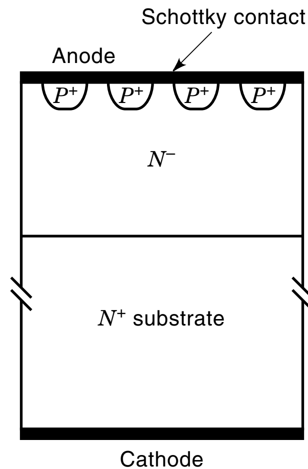

Figure 3. Cross-sectional view of the junction barrier controlled Schottky (JBS) rectifier structure. The p^+ junction grid is designed so that its depletion layers do not pinch-off under zero and forward bias conditions of the rectifier, but intersect with each other underneath the Schottky contact when the reverse bias exceeds a few volts.

ing the Schottky barrier height. Unfortunately, a low barrier height results in a severe increase in leakage current and a reduction in maximum operating temperature. Further, Schottky power rectifiers fabricated with barrier heights of less than 0.7 eV have been found to exhibit an extremely soft breakdown characteristic, which makes them prone to failure.

The junction barrier controlled Schottky (JBS) (3) rectifier is a Schottky rectifier structure with a *P-N* junction grid integrated into its *N*-region. A cross-section of the JBS structure is provided in Fig. 3. The junction grid is designed so that its depletion layers do not pinch-off under zero and forward bias conditions of the rectifier, but intersect with each other under the Schottky contact when the reverse bias exceeds a few volts.

Under reverse blocking states, after the depletion layer pinches off, a potential barrier is formed in the channel between the two grids, and further increase of the applied reverse voltage is supported by it with the depletion layer extending toward the N^+ substrate. Therefore, the potential barrier shields the Schottky contact from the applied voltage. This shielding prevents the Schottky barrier lowering phenomenon and eliminates the large increase in leakage current observed in conventional Schottky rectifiers. During on-state operation, there are multiple conductive channels under the Schottky contact through which current can flow. Because of the suppressed leakage current, the Schottky barrier height used in JBS rectifiers can be decreased compared to that of conventional Schottky rectifiers. This has allowed a reduction in the forward voltage drop while maintaining acceptable reverse blocking characteristics. For the same leakage current, the JBS rectifier has been found to provide a forward voltage of 0.25 V, compared with about 0.5 V for the Schottky rectifier.

In the design of the JBS rectifier, the lowest on-state voltage drop can be obtained by making the width of the junction diffusion window as small as possible. This minimizes the dead space below the junction where the current does not flow. The best JBS rectifier characteristics can therefore be expected when submicron lithography is used to pattern the diffusion windows for the P^+ regions.

MPS Rectifiers

The merged *P–i–N*/Schottky (MPS) (4) rectifier has a structure similar to that of the JBS rectifier, as shown in Fig. 4. However, the operating physics and applications of the two rectifiers are quite different. In the JBS rectifier, there is no injection of minority carrier holes from the *P-N* junction, and the on-state voltage drop is less than 0.5 V. In the MPS rectifier, the *N*-region is designed for supporting a high reverse blocking voltage, and the forward biasing *P-N* junction becomes necessary in the on-state. The forward bias of the *P-N* junction produces the injection of holes into the *N*-region and results in conductivity modulation in the *N*-region in a manner similar to the *P–i–N* rectifier, which drastically reduces the resistance of the *N*-region to current flow. This also allows larger current flow via the Schottky region due to a lower series resistance. Because of the existence of the Schottky region, the injection level required to reduce the resistance in the MPS is not as large as that observed in the *P–i–N* rectifier. As a consequence, the stored charge in the MPS rectifier is much smaller than that in the *P–i–N* rectifier when they have the same on-state voltage drop. Consequently, MPS rectifiers have a better reverse recovery characteristic and a superior trade-off curve between the on-state and the turn-off loss than those for *P–i–N* rectifiers.

Emitter Short Diode

The emitter short diode (ESD) (5,6) is an improved *P–i–N* diode structure proposed for high-voltage and high-speed applications. Figure 5 shows the basic structure of the ESD.

In the ESD structure, additional N^+ short regions are formed on a lightly doped *p*-emitter layer to reduce the hole injection efficiency and attain asymmetric carrier profiles for the N^--region in the on-state by control the junction depth of the N^+ regions. To block a high reverse voltage, a high impurity concentration for the *p*-emitter is needed. However, to

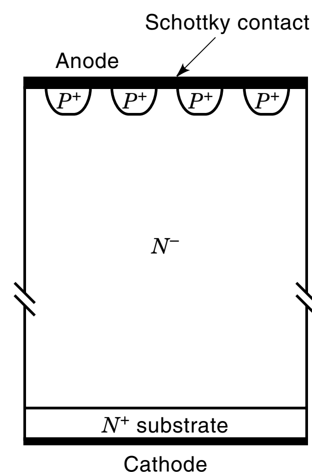

Figure 4. Basic structure of the merged *P–i–N*/Schottky (MPS) rectifier. The N^- region is designed to support a high reverse blocking voltage. The purpose of the p^+ region is to enhance the minority carrier injection when the forward voltage is higher than 0.7 V. This p^+ region also reduces the leakage current.

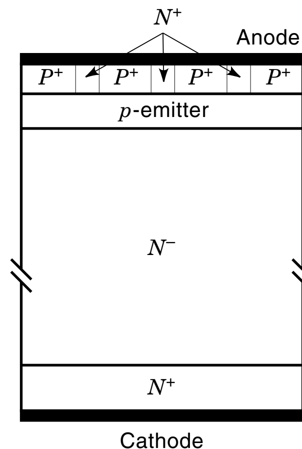

Figure 5. Basic structure of the emitter short diode (ESD). Its N^+ short regions reduce the emitter injection efficiency of holes by clamping the peak values of the electron concentration underneath these N^+ regions.

reduce emitter injection efficiency, a shallow junction and low impurity concentration for the p-emitter are needed. Consequently, the minimum impurity dose of the shallow p-emitter should be chosen carefully to block the desired high reverse voltage and avoid the depletion region reaching the N^+ short layers. To reduce the lateral resistance in the p-emitter layer under N^+ short regions and avoid the parasitic transistor effect during reverse recovery, the N^+ areas should be formed in a fine pattern.

The ESD reduces the emitter injection efficiency and controls the reverse recovery current behavior by using a shallow p-emitter combined with the emitter short structure. It attains half of the reverse recovery current, half of the reverse recovery time, and one-fourth of the reverse charge, when compared to a conventional P–i–N diode. Its leakage current is also as low as a conventional P–i–N diode, even at 125°C.

Synchronous Rectifier

With the increasing requirement in applications such as computers for power supplies for even lower supply voltage, the conduction loss in the output rectifier becomes the biggest source of power loss in switching power supplies. Even the commonly used Schottky diodes have a relatively large voltage drop and, hence, a large power loss in such low-output-voltage applications. Consequently, low-voltage metal-oxide-semiconductor field-effect transistors (MOSFETs), which operates in the third quadrant, with a very low on-state resistance and fast switching speed can be used to replace the diodes in the output stage. Because the gate signal to the low-voltage power MOSFET is provided in synchronism with the drain-source voltage to maintain low on-resistance in one direction and blocking state in another direction, the low-voltage power MOSFET in these applications is called a synchronous rectifier (SR) (1,7,8). Low-voltage power MOSFETs are successfully used as SR because of their linear V–I characteristic. The conduction loss can therefore be reduced to a very low value by paralleling more MOSFETs. The SR is also fast because it is a majority carrier device.

SWITCHES

A switch, in semiconductor terms, is a device that has two states for current flow in the same direction—a low-impedance state (ON state) and a high-impedance state (OFF state). Switching between these two states can be controlled by voltage, current, temperature, or light.

Bipolar Power Transistor

Bipolar power transistors have been commercially available for more than 30 years. They were favored for low and medium power applications because of their faster switching capability. The rating for bipolar power transistors grew steadily until the end of the 1970s, when power MOSFET started to appear in the market.

Figure 6 shows the cross-sectional view of a bipolar power transistor cell and its circuit symbol. Although the operating physics for bipolar power transistors is essentially the same as that for signal transistors, their characteristics differ because they need to support a high collector voltage in the forward blocking mode. The high voltage capability of the bipolar power transistor is obtained by incorporating a high resistive, thick N^- region into the collector structure. In addition, the base region must be carefully designed to prevent punch-through breakdown. These differences strongly influence the current gain of the device.

Another distinguishing feature of bipolar power transistors is that they operate at relatively high current densities in the saturation region when both the emitter and collector junction are forward biased. This produces high-level injection not only in the base but also in the collector region. The high-level injection results in severe degradation of the current gain. In general, bipolar power transistors have a low current gain at typical operating current levels. Since the bipolar power transistor is fundamentally a current controlled device, with the magnitude of the collector current determined by the base drive current, a low current gain means a bulky and expensive control circuit requiring many discrete components in applications. Consequently, one of the most critical design goals

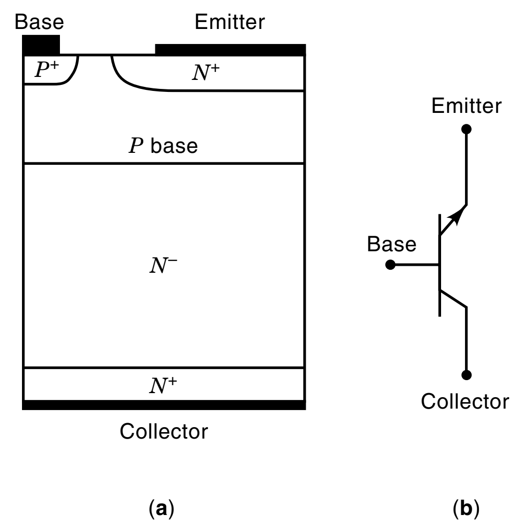

Figure 6. (a) Cross-sectional view of a bipolar power transistor cell and (b) its circuit symbol.

has been to improve the current gain in order to reduce the complexity, size, and weight of the base control circuit. Unfortunately, achieving a high current gain conflicts with the achievement of high breakdown voltage. Further, the fall-off in current gain at typical operating current densities due to high injection level leads to a gain of less than ten. The current gain can be improved by using the Darlington power transistor but this has the disadvantage of considerably increasing the on-state voltage drop.

In addition, the devices are prone to failure due to the second breakdown phenomenon. This occurs because of the affinity for the formation of a local region in the emitter through which the current tends to constrict itself. It appears on the output characteristic of the power bipolar transistor as a precipitous drop in the collector-emitter voltage at large collector currents. As the collector voltage drops, there is a significant increase in the collector current and a substantial increase in the power dissipation. What makes this situation particularly dangerous for the power bipolar transistor is that the dissipation is concentrated in highly localized regions where the local temperature may grow very quickly to unacceptably high values because of the positive feedback relationship between the current and the temperature within the power bipolar transistor. If this situation is not terminated quickly, device destruction results. This positive feedback relationship between the current and the temperature also means that power bipolar transistors are difficult to parallel. For these reasons, the bipolar power transistor has been displaced by the power MOSFET for high-speed, low-power applications in the 1980s, and for medium-power applications by the IGBT in the 1990s.

Darlington Power Transistor

Figure 7 shows the cross-sectional view of a monolithic Darlington power transistor (8) and its equivalent circuit. In this structure, two transistors have a common collector connection, and the emitter of the drive transistor (T_1) is connected to the base terminal of the output transistor (T_2). The base drive current (I_B) is supplied to the drive transistor. This current turns “on” the transistor T_1, which then provides the base drive current for transistor T_2. Consequently, the current gain of the Darlington power transistor, β, is approximately equal to the product of the current gain of the drive transistor (β_1) and the current gain of the output transistor (β_2). That is

$$\beta = \beta_1 \beta_2$$

However, in order to turn on the output transistor T_2, it is necessary to raise the potential of the collector. Thus, the on-state voltage drop of the Darlington power transistor is higher than that of a single bipolar power transistor.

The Darlington power transistor was the only switching power device with a fast switching speed to deliver several hundred amperes of current and up to 1000 V before the appearance of IGBTs. These power Darlington transistors were also called giant transistors (GTRs). Darlington power transistor can be in the form of a monolithic device, such as that shown in Fig. 7; it can also be formed by multichip packaging technique, commonly known as power module.

Thyristor

A thyristor is loosely defined as a device having a four-layer *P-N-P-N* structure, leading to bistable behavior, that can be switched between a high-impedance, low-current OFF state and a low-impedance, high-current ON state. In the past, the thyristor was also commonly called the semiconductor controlled rectifier (SCR) or the silicon controlled rectifier (SCR). The basic structure of the thyristor and its equivalent circuit are illustrated in Fig. 8.

As shown in Fig. 8(a), a thyristor consists of four semiconductor layers (*P-N-P-N*). The N^- layer is lightly doped and supports a high voltage when the device is in its blocking state, with junction J_1 or J_2 reverse biased. Thyristors hence offer both forward and reverse blocking capability of comparable magnitude. This makes them well suited for ac circuit applications. Thyristors are now available with ratings of 10 kV and 6 kA. These devices are manufactured from single 10 to 12.5-cm diameter wafers by using a matured deep diffusion process with gallium and aluminum as dopants. High breakdown voltages are realized by using positive and negative bevel etching techniques at the edge of the wafer. An SCR typically only has a few gate fingers, so device turn-off relies on forced commutation in ac circuits.

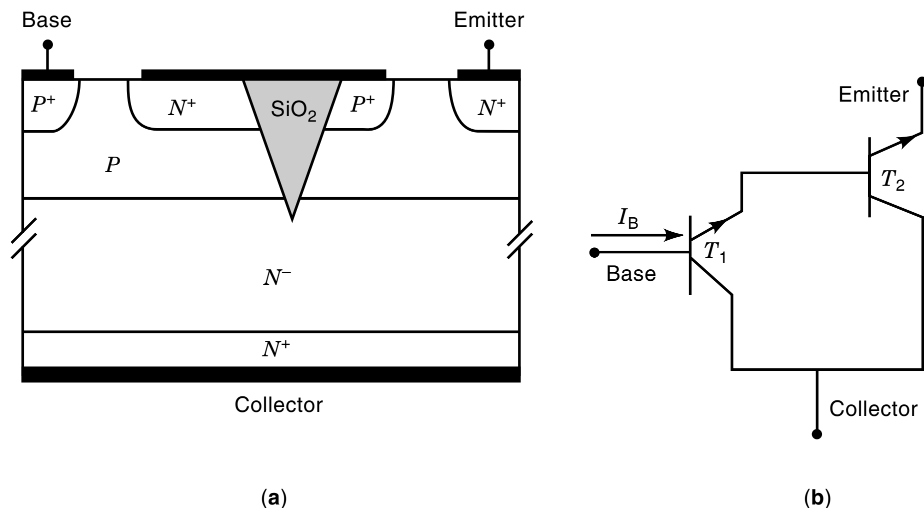

Figure 7. (a) Cross-sectional view of the monolithic Darlington power transistor and (b) its equivalent circuit.

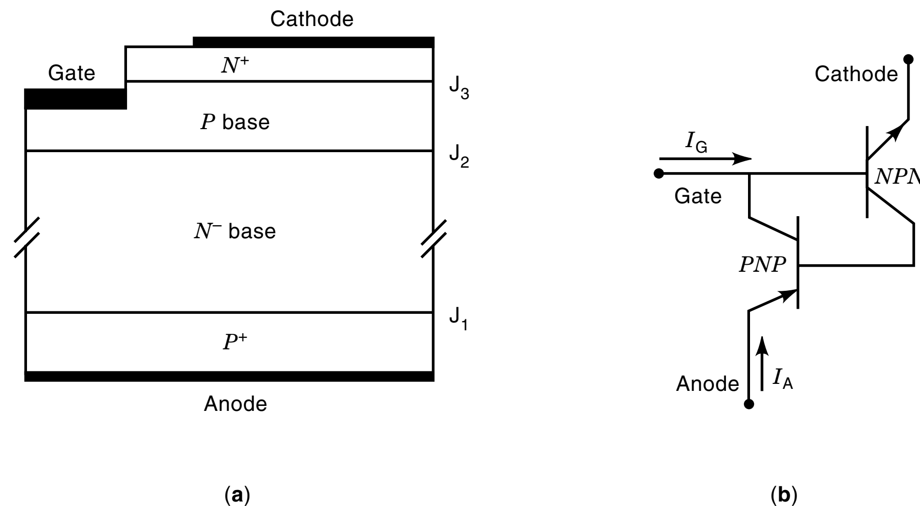

Figure 8. (a) Basic structure of the thyristor and (b) its equivalent circuit.

From Fig. 8(b) it is clear that a thyristor can be bisected into a *P-N-P* and an *N-P-N* bipolar transistor, with each base connected to the other's collector. Consequently, the anode current I_A can be described by

$$I_A = \frac{\alpha_{NPN} I_G + (I_{co1} + I_{co2})}{1 - (\alpha_{NPN} + \alpha_{PNP})}$$

where α_{NPN} and α_{PNP} are current gains of the *N-P-N* and the *P-N-P* bipolar transistors, respectively. I_{co1} and I_{co2} are leakage currents of the *N-P-N* and the *P-N-P* bipolar transistors, respectively. In the blocking state, the sum of $\alpha_{NPN} + \alpha_{PNP}$ is much small than unity and I_G is zero so that the anode current can be kept quite small. If the sum of $\alpha_{NPN} + \alpha_{PNP}$ approaches unity, the anode current will be arbitrarily large.

For turning on the SCR, a small triggering current is required at the gate. This gate drive current serves to turn on the *N-P-N* bipolar transistor and increases the current gain. Once the current gains of the two transistors within the thyristor structure become sufficiently large, the two transistors can provide the base drive currents for each other and the thyristor enters self-sustaining mode. This mode is referred to as regenerative mode, or as latch-up state.

Because thyristors operate in the latch-up mode, they have very low conduction loss. However, it is difficult to turn off a thyristor. In the case of a conventional thyristor structure, the device is primarily used in ac circuits where the anode voltage periodically reverses to force the current to decrease to zero.

A light-triggered thyristor is a thyristor that can be directly triggered via an optical signal. It is useful for high-power systems, such as the high-voltage direct-current (HVDC) transmission system, because the control system is isolated from the power stage.

TRIAC

A TRIAC (triode ac switch) (1) is a semiconductor device with bi-directional voltage blocking capability and bi-directional current conduction capability. Fig. 9 shows a basic TRIAC structure with a single control gate electrode and its output characteristic. It is clear that this structure has two back-to-back thyristors integrated monolithically in an antiparallel configuration.

GTO Thyristor

The gate turn-off thyristor (GTO) is similar to the SCR in device structure but has gate current turn-off capability. Figure 10 shows the vertical cross-sectional view of the GTO. It has the basic *P-N-P-N* four-layer structure of the SCR. In order to have more efficient gate-controlled turn-off, during the on-state the base current of the *N-P-N* bipolar transistor and the collector current of the *P-N-P* bipolar transistor have to be minimized. This demands that $\alpha_{NPN} \gg \alpha_{PNP}$. Consequently, in the GTO structure, the thickness of the *P* base layer is generally somewhat smaller than that in an SCR structure. Another significant difference between a GTO and an SCR is that a lot of gate fingers are placed next to the cathode emitter in an interdigital fashion. The basic goal is to maximize the periphery of the cathode and minimize the distance from the gate to the center of the cathode region.

In the GTO, a negative gate current, I_G, alone is able to turn off the thyristor without forced commutation. A large reverse gate drive current extracts charges from the base region of the upper transistor and disrupts the self-sustaining current conduction mechanism in the GTO. I_A/I_G is defined as the turn-off gain, β_{off}, where I_A is the anode current. For successful turn-off, β_{off} has to satisfy

$$\beta_{off} \leq \frac{\alpha_{NPN}}{\alpha_{PNP} + \alpha_{NPN} - 1}$$

Typical values for turn-off gain range between 5 and 10.

The main advantage of the GTO thyristor is the elimination of an external circuit for forced commutation, which provides increasing flexibility in circuit applications. Another advantage is a smaller turn-off time and the capability for the higher speed operation compared to that of the SCR. The disadvantage is the large gate currents required for turn-on and turn-off of the GTO. The GTO is the only commercially avail-

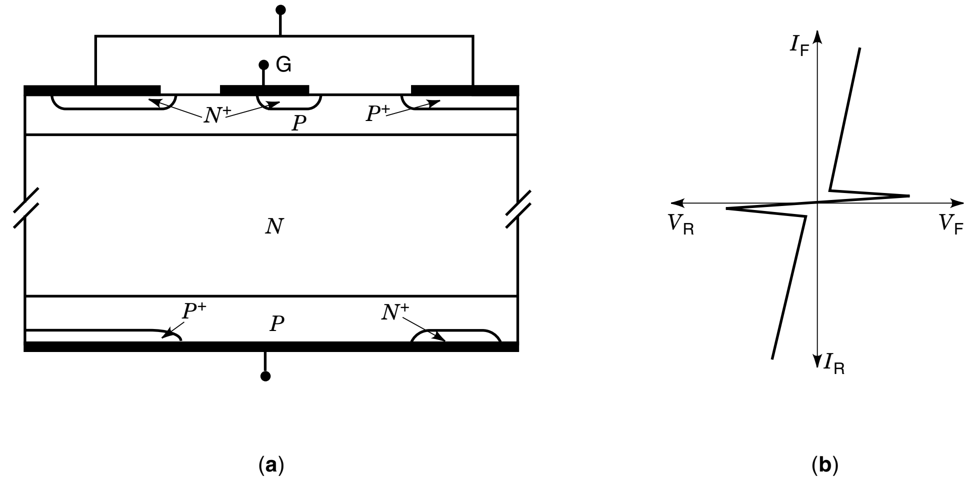

Figure 9. (a) Basic structure of a TRIAC and (b) its output characteristic.

able device today with the ability to block 8 kV and control over 6 kA for applications such as traction control. The fabrication process of the GTO is similar to that of the SCR.

The main problem that limits the application of GTOs is that they require a complicated and expensive gate drive to turn off GTOs. Not only a large negative gate current is required, but also a high di_G/dt for that negative gate current. The latter is crucial to ensure that each segment in the GTO turns off uniformly. Otherwise, the GTO could be destroyed permanently due to a process called current filamentation, in which a slow gate turn-off current only turns off some of the GTO cells, forcing all of the anode current to crowd around a few cells or even a single cell, hence destroying the device through a very high localized power dissipation. A large snubber is therefore routinely used in GTO applications to solve this problem.

SIT

The static-induction transistor (SIT) (9) was introduced in 1972 and began to be produced in the market in the mid-1980s as a switching power device. Several structures of the SIT are shown in Fig. 11. The buried gate structure is the original proposed scheme, while the planar gate structure and the recessed gate structure are more popular. In an SIT structure, the gate and source regions are highly interdigital. Thousands of these basic gate-source cells are connected in parallel to make up a single SIT. The most critical parameters in an SIT are the spacing between gates and the channel doping level. Since most SITs are designed as "normally on" devices, the doping is chosen such that the depletion regions from the gates do not merge and there exists a narrow neutral channel opening at zero gate bias. The gates in SITs are formed normally by *P-N* junctions, but the SIT operations can also be generalized to include metal (Sckottky) gates, or even metal-insulator-semiconductor (MIS) gates. Normally off SITs are also being fabricated by using very narrow channel design.

The SIT is basically a junction field-effect-transistor (JFET) or metal-semiconductor field-effect-transistor (MESFET) with super-short channel length, and with multiple channels connected in parallel. As a result of short channel length, punch-through occurs with high drain bias even if the transistor is originally turned off (static induction is equivalent to punch-through). The output characteristics of a normally on SIT are shown in Fig. 12. These characteristics are quite different from those of a bipolar transistor and are often referred to as triode-like characteristics because of their resemblance to the *I–V* characteristics of a vacuum triode.

As shown in Fig. 12, when a positive bias is applied to the drain, the normally on SIT is in the on-state when the gate-source voltage is zero. The current conduction is drift in nature and is similar to a JFET. When a reverse bias is applied to the gate, the depletion layers widen, and pinch off the channel. The depletion layers set up a potential barrier to the flow of the drain electron current. As a consequence, there will be no flow of current between drain and source as long as the drain-source voltage is kept small. This is the off-state of the SIT. As the drain to source bias voltage, V_{DS}, increases, the potential barrier to the drain current flow gets smaller and smaller. When V_{DS} is large enough to suppress the potential barrier set up by the gate-source bias voltage, current begins to flow again and increases with the increase of the V_{DS}.

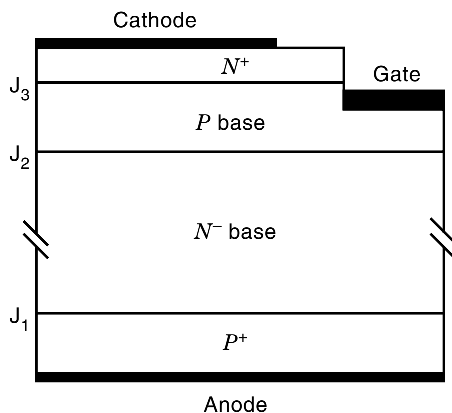

Figure 10. Vertical cross-sectional view of the GTO structure. The structure of the GTO is similar to that of the thyristor, but it has a much narrower emitter cell width than that of the thyristor so as to improve its gate-controlled turn-off capability.

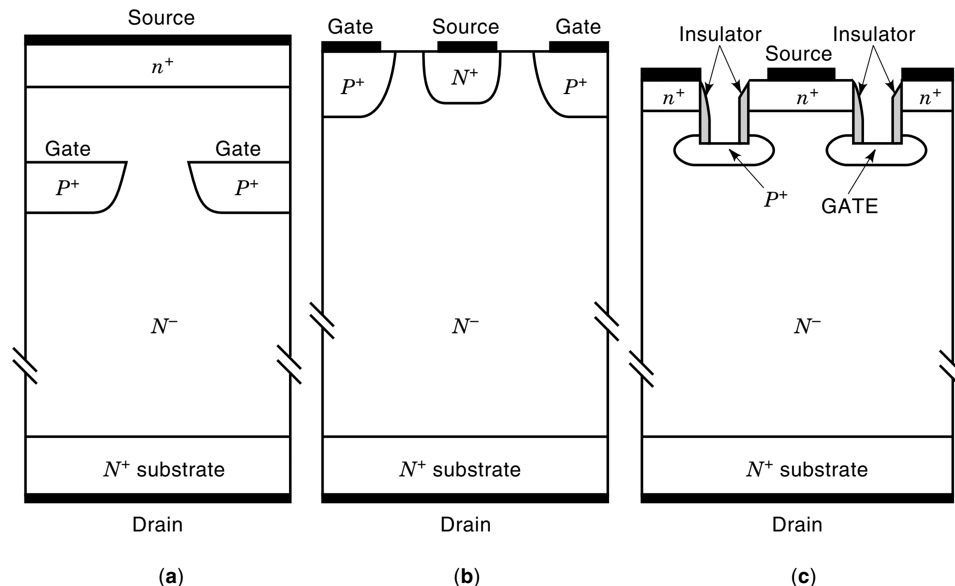

Figure 11. Structures of static-induction transistor (SIT) with (a) buried gate, (b) planar gate, and (c) recessed gate.

The main attractiveness of an SIT is the combination of high-voltage and high-speed capability. Its cut-off frequency can be up to 2 GHz (9). As an audio power amplifier, the SIT has low noise, low distortion, and low output impedance. It can be used in high-power oscillators of microwave equipment, such as broadcasting transmitters and microwave ovens. As switching power devices, SITs are limited by their normally on characteristic and because power MOSFETs, developing concurrently with SITS, are superior to SITs in switching power applications due to their fast speed, high input impedance, and normally off characteristic. SITs are also difficult in scaling up to high voltages, not only because the conduction loss will increase, but also because a large negative gate bias is needed to block higher voltage. A factor commonly defined as forward-blocking voltage gain is defined as the change of the blocking voltage V_{DS} induced by the change of V_G for the same drain current. High-voltage SITs typically have a blocking gain of less than 20.

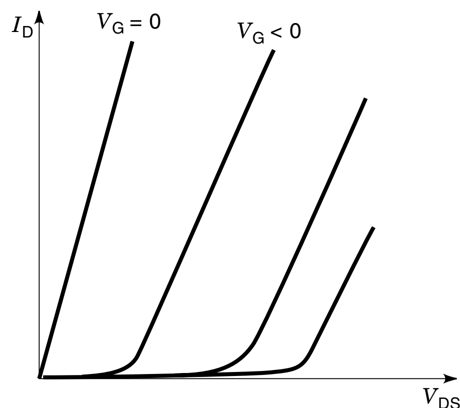

Figure 12. Output characteristics of the normally-on SIT. These I–V curves are very different from those of a bipolar transistor and similar to a triode. The SIT typically conducts current when the gate voltage is zero. This type of device is called a "normally on" device.

SITH

The static induction thyristor (SITH) is also called the field controlled thyristor (FCT). The SITH was introduced in the mid-1970s with the aim to reduce the conduction loss of high-voltage SITs. Although the SITH is commercially available as a power device, its performance is superseded by the IGBTs.

The structure of the SITH is similar to that of the SIT with a P^+ anode replacing the N^+ drain. The basic structures of the SITH with planar gates, buried gates, and double gates are shown in Fig. 13 (a), (b), and (c), respectively. It is clear that the SITH consists of a P–i–N diode with part of the channel surrounded by closely spaced junction grids or gates. There are two types of SITH—normally on SITH and normally off SITH. In the normally on SITH, pinch-off of the channel does not occur with zero gate voltage, and a high current can flow. In the normally off SITH, the depletion regions of the nearby gate merge, and the pinch-off occurs at zero gate voltage. The output characteristics for a normally on SITH are shown in Fig. 14.

In the normally on SITH, at zero gate bias or small positive gate bias, the depletion regions around the gates do not pinch off the gap completely. The current conduction from anode to cathode is similar to that of a P–i–N diode. At a forward biased voltage V_{AK}, electrons are injected from the cathode and holes are injected from the anode, and they are equal in number to maintain charge neutrality. These excess electrons and holes increase the conductivity of the N^- layer. Note that although the output characteristics are similar in shape to those of the SIT, the P^+ anode can inject holes and enable conductivity modulation, resulting in a lower forward-voltage drop or lower on-resistance.

With a larger reverse gate bias, the depletion layers extend, pinch-off of the channel is introduced, and a barrier for electrons is formed. This barrier limits the electron supply and becomes the controlling factor for the overall current. Without an ample electron supply, the hole current reduces to leakage generation current and becomes insignificant. The SITH enters the forward blocking state.

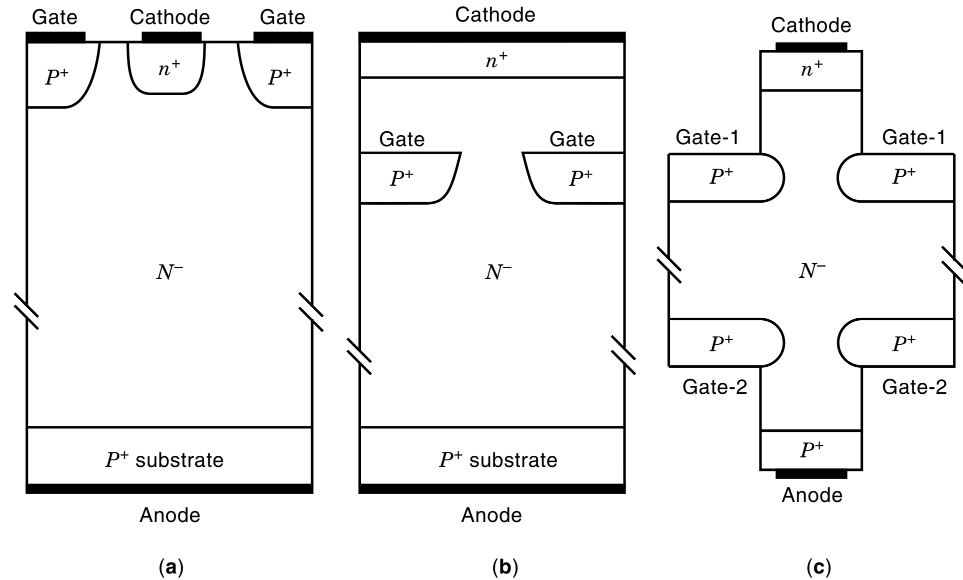

Figure 13. Structure of the static induction thyristor (SITH) with (a) planar gates, (b) buried gates, and (c) double gates.

In the SITH structure, the channel barrier height can be influenced by the gate voltage as well as by the anode voltage. A large forward anode bias V_{AK} can lower this barrier height. This dependence of the barrier height on the forward anode bias is called static induction. Static induction current is basically a punch-through current due to the thin and small barrier in the direction of current flow.

One useful parameter for the SITH is the forward-blocking voltage gain, which is defined as the change of the blocking voltage V_{AK} induced by the change of V_G for the same anode current. The forward-blocking voltage gain depends on the structure of the gates and the channel doping.

One of the advantages of the SITH when compared with GTOs is its higher speed of operation due to a faster turn-off process. During turn-off, the reverse gate bias can extract the excess minority carriers (holes) quickly. The excess electrons, being majority carriers in the N^- region, can be swept away quickly by the drift process. The hole current contributes to an instantaneously large gate current, and a small gate resistance is critical to avoid gate debiasing. An alternate technique to reduce the turn-off time is to reduce the minority carrier lifetime by lifetime control technique. The penalty for using this technique is a larger forward voltage drop.

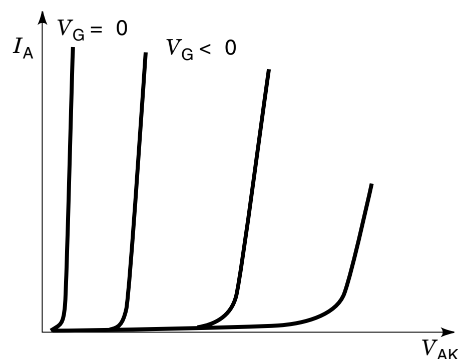

Figure 14. Output characteristics of the normally-on SITH.

In the SITH structure, the planar gates structure has a lower gate resistance since a metal contact can be deposited directly over it. This results in a smaller debiasing effect during the turn-off process when there is a substantial current through the gate. The advantage of the buried gates structure is a higher forward blocking voltage gain resulting from a more efficient use of the cathode area and a more effective gate control of the current. The double-gates SITH is capable of higher speed than the single-gate structure, but it has a more complicated fabrication process.

Because of the fast turn-off capability, SITHs with an operating frequency up to 100 kHz are possible. With a high forward-blocking voltage gain of up to 700, stable operation at high temperature, and large dI/dt and dV/dt capabilities, the SITHs have been applied mainly in power source conversion such as ac to dc converters, dc to ac converters, and chopper circuits. Other applications of the SITH include pulse generation, induction heating, lighting of fluorescent lamps, and driving pulsed lasers.

The main problem for the SITHs is similar to that of the GTOs, in that a large and expensive gate drive circuit has to be provided. Because of the nature of their planar and shallow junction process, the SITHs have not reached the power ratings of GTOs. They are, therefore, seriously challenged by the IGBTs in the medium power range, because IGBTs have many of the SITHs advantages plus a simple control interface owing to their high input impedance.

Power MOSFET

Prior to the development of the power metal-oxide-semiconductor field-effect-transistors (power MOSFETs), the most favorable device available for high-speed, medium-power switching applications was the power bipolar transistor. But the power bipolar transistors exhibit several fundamental drawbacks in their operating characteristics, such as that they are current controlled devices and are difficult to parallel. In order to suppress these performance limitations, the power MOSFET was developed in the 1970s due to the advancement of VLSI technology.

In the power MOSFET, the control signal is applied to a metal (or polysilicon) gate electrode that is separated from the semiconductor surface by an intervening insulator (typically silicon dioxide). Thus, the power MOSFET has a very high input impedance in steady-state, and it is classified as a voltage controlled device that can be controlled using integrated circuits because of the small gate currents that are required to charge and discharge the input gate capacitance. Even during the switching of the devices between the ON and OFF states, the gate current is small at typical operating frequencies of less than 100 kHz because it serves only to charge and discharge the input gate capacitance. When the operating frequency becomes high (>100 kHz), this capacitance current can become significant, but it is still possible to integrate the control circuit due to the low gate bias voltages (typically 5 to 15 V) required to drive the device into its on-state with a low forward voltage drop.

In comparison with the bipolar transistor, the power MOSFET is a unipolar device, it therefore has a very fast switching speed due to the absence of minority carrier injection. The switching time for the MOSFET is dictated by the ability to charge and discharge the input capacitance rapidly. This feature is particularly attractive in circuits operating at high frequency, where switching power loss is dominant. Further, the power MOSFET has superior ruggedness and has been found to display an excellent safe-operating area (i.e., they can withstand the simultaneous application of high current and voltage without undergoing destructive failure).

These characteristics of power MOSFETs make them important candidates for many applications such as high-frequency power conversion and lamp ballasts. Three discrete vertical channel power MOSFET structures are described.

VVMOS (VMOS). The vertical V-shaped groove MOSFET (VVMOS or VMOS) was the first commercial structure of power MOSFET developed in the 1970s. The VVMOS structure is shown in Fig. 15. This structure is based on the V-shaped groove by anisotropically etching in a ⟨100⟩ silicon substrate within which the gate is located along the ⟨111⟩ planes extending through the *P* layer. It can be fabricated by first performing an unpatterned *P*-region diffusion followed by the N^+ source region diffusion. A V-shaped groove extending through these diffusions is then formed by using preferential etching with potassium hydroxide based solutions. For an aluminum gate VMOS, the gate oxidation layer is grown and the gate electrode is then deposited and patterned. The channel region for this structure is formed along the walls of the V-groove. This structure can therefore provide short channel length without fine lithography capability, and the vertical current flow also maximizes the total current for a given surface area.

For the *N*-channel structure shown, when a positive bias larger than the threshold voltage of the MOSFET is applied to the gate electrode, an inversion layer forms along the V-groove and the channel is turned on. When a positive bias voltage is applied to the drain, electrons flow from the N^+ source via the MOSFET channel into the drift region and are then collected by the N^+ drain. When the gate voltage is lower than the threshold voltage of the MOSFET, the device is in the forward blocking state and can support a high drain voltage across the *P*-body/*N*-drift junction.

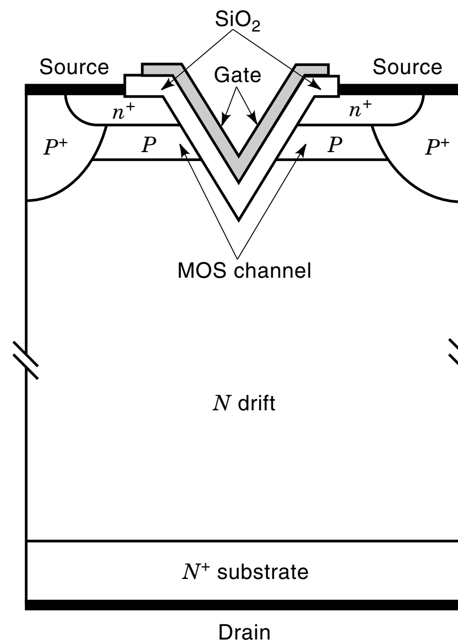

Figure 15. Cross-section of the VVMOS structure. The V-shaped groove is formed by anisotropically etching an ⟨100⟩ silicon substrate. The MOSFET channel region is formed along the walls of the V-groove and can therefore provide short channel length without fine lithography capability. The vertical current flow also maximizes the total current for a given surface area. The drawback of this structure is that a strong electric field peak exists at the V-groove corner when a high drain-voltage is applied.

Because the gate of the VVMOS is located in the V-groove formed by the preferential etching, there are instabilities in the threshold voltage during manufacturing. In addition, the sharp tip at the bottom of the V-groove creates a high electric field during the forward blocking state, which degrades its breakdown voltage. Furthermore, the mobility of the carrier on a ⟨111⟩ etched slope is somewhat lower. For these reasons, the VVMOS has been displaced by the VDMOS structure based on the double-diffusion MOS (DMOS) process.

VDMOS. The vertical double-diffusion MOSFET (VDMOS) is the most popular power MOSFET structure. Figure 16 shows the cross-section of one-half of a VDMOS cell structure. This VDMOS is fabricated by using planar diffusion technology with a refractory gate, such as polysilicon, as a mask. The *P*-region and the N^+ source region are diffused through a common window defined by the edge of the polysilicon gate. The name for this device is derived from this *double-diffusion process.* The *P*-region is driven deeper than the N^+ source region, hence defining the surface channel region. This process enables fabrication of a submicrometer channel length without resorting to high-resolution lithography. A commercial VDMOS chip contains millions of such VDMOS cells in parallel, with a common gate control.

For an *N*-channel VDMOS, the device operates with a positive voltage applied to the drain. When the gate voltage is lower than the threshold voltage of the DMOS, the device can support a high drain voltage across the *P*-region/*N*-drift region junction. The forward blocking capability is determined by the breakdown characteristic of the *P*-region/*N*-drift junc-

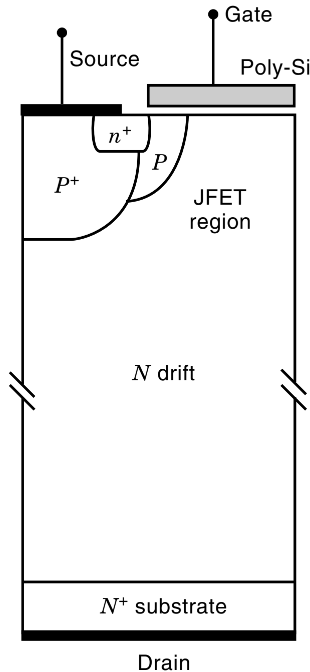

Figure 16. Cross-section of the VDMOS structure. This VDMOS is fabricated by using planar diffusion technology with a refractory gate as a mask. The double-diffusion process forms the channel laterally and enables fabrication of a submicrometer channel length without resorting to high-resolution lithography.

tion. The voltage blocking capability of the VDMOS is better than that of the VVMOS, because the field distribution in the cell region is essentially one-dimensional. When the gate voltage is higher than the threshold voltage of the DMOS, the surface channel is turned on, and electrons will flow from the source via the channel into the drift region and then bend 90° before being collected by the drain.

The conduction loss of the VDMOS is specified by the on-resistance. The on-resistance is an important device parameter because it determines the maximum current rating. The specific on-resistance, defined as the on-resistance per unit area, is a preferable parameter in the design of the VDMOS. The on-resistance of the VDMOS consists of the N^+ source resistance, the channel resistance, the accumulation layer resistance under the polysilicon gate, the JFET region resistance, the drift region resistance, the substrate resistance, and the contact resistance.

A trade-off between the forward blocking voltage and the specific on-resistance exists for the VDMOS and other unipolar devices such as the SIT. A high forward blocking voltage needs a lightly doped and thick drift region, hence creating a large specific on-resistance. This trade-off relationship can be described by:

$$R_{\text{drift, specific}} = \frac{4BV^2}{\epsilon_s \mu_n E_C^3}$$

where ϵ_s is the dielectric constant of the silicon, μ_n is the electron mobility, BV is the breakdown voltage, and E_c is the critical electric field at avalanche breakdown. Because of this trade-off, the VDMOS usually are designed to operate at high-voltage, low-current levels or low-voltage, large-current levels.

The forward voltage drop or the on-resistance in the VDMOS increases with the increase of temperature; therefore, VDMOSs can be easily paralleled. This characteristic of VDMOSs makes them important candidates for many applications. The VDMOS is a unipolar device; current conduction occurs via transport of majority carriers in the drift region without the minority carrier injection required in bipolar transistor operation. Thus, the VDMOS has a faster switching speed than bipolar transistors. This feature is particularly attractive in circuits operating at high frequencies, where switching power losses are dominant.

UMOS (Trench-Gate MOSFET, UMOSFET). The UMOS structure is shown in Fig. 17. The name for this structure is derived from the U-shaped groove formed in the gate region by using reactive ion etching (RIE). The fabrication of this structure can be performed by following the same sequence as the VVMOS structure with the V-groove replaced by the U-groove. The U-groove structure has a higher channel density than either the VMOS or DMOS structures, which allows significant reduction in the on-resistance of the device. The technology for the fabrication of this structure was derived from the trench etching technique developed for memory cells in DRAMs.

In the UMOS structure, the UMOS cell size can be made relatively small (6 μm) when compared with the DMOS cell (20 μm) for the same design rules. This results in an increase in the channel density (channel width per square centimeter of device area). In the UMOS structure, no JFET region exists. The on-resistance of the UMOS consists of the N^+ source resistance, the channel resistance, the drift region resistance, the N^+ substrate resistance, and the contact resistance. Un-

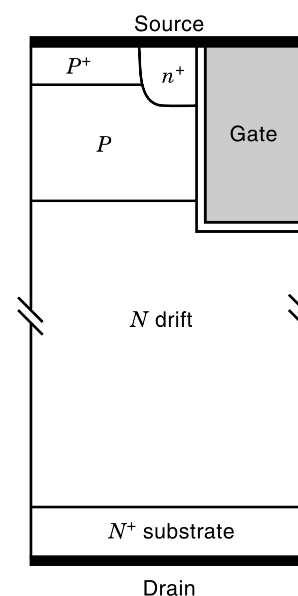

Figure 17. Cross-section of the UMOS structure. A U-shaped gate region is formed by using reactive ion etching (RIE) followed by gate oxidation and polysilicon refill. THe UMOSFET has a higher channel density than either the VMOS or DMOS structure, resulting in the lowest on-resistance per unit silicon area.

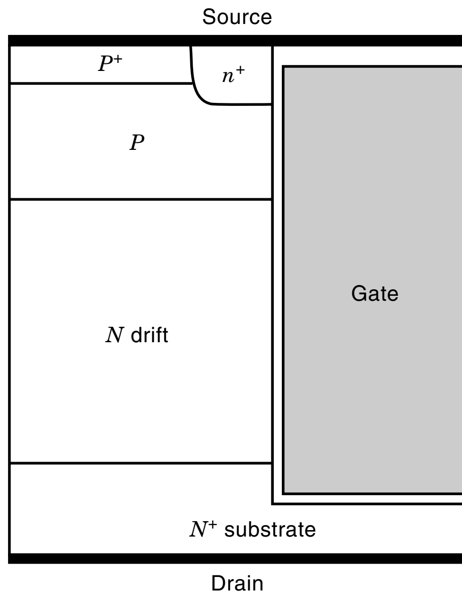

Figure 18. Cross-section of the deep-trench UMOS. This structure extends the trench down to the N^+ substrate, and the drift region resistance component of the on-resistance is reduced by the parallel current flow path of an accumulation layer on the sidewall of the trench.

like the DMOS structure, there is no optimum design for the UMOS cell. In this case, it is beneficial to reduce the mesa and trench width as much as possible. As these dimensions becomes smaller, the channel resistance contribution becomes smaller because the channel density increases. Therefore, the UMOS has the lowest specific on-resistance in power MOSFETs. Using UMOS structure, however, will not reduce the on-resistance of the drift region, hence, UMOS structure is only beneficial for low voltage power MOSFETs.

Figure 18 shows a modified UMOS structure with a deep trench, which has a very low specific on-resistance approaching the limits for silicon FET performance (10). In this structure, the trench extends down to the N^+ substrate, and the drift region resistance component of the on-resistance is reduced by the parallel current flow path created by the formation of an accumulation layer on the sidewall of the trench. But it must be noted that the blocking voltage of this structure is limited to less than 30 V by the high electric field created in the gate oxide by the extension of the trench into the N^+ substrate.

Low breakdown voltage power MOSFETs (<30 V) are also successfully used as synchronous rectifiers because their conduction losses can be reduced to even lower than those of the SBDs because of their linear V–I relationships (5).

IGBT

The name insulated-gate bipolar transistor (IGBT) comes from its operation based on an internal interaction between an insulated-gate FET (IGFET) and a bipolar transistor. It has also been called previously an IGT (insulated-gate transistor), an IGR (insulated-gate rectifier), a COMFET (conductivity-modulated field-effect transistor), a GEMFET (gain-enhanced MOSFET), a BiFET (bipolar FET), and an injector FET. IGBTs have been successfully used since they were first demonstrated in 1982.

The IGBT is an important power switch used in converters with ratings up to several hundred kilowatts. A cross-section of the planar DMOS-technology based IGBT structure is shown in Fig. 19(a). It is clear from Fig. 19(a) that the IGBT structure is similar to that of the VDMOS from the fabrication point of view. This has made its manufacturing relatively easy immediately after conception and its power rating has grown at a rapid pace due to the ability to scale up both the current and the blocking voltage.

The equivalent circuit for the IGBT, shown in Fig. 19(b) consists of a wide-base *P-N-P* bipolar transistor driven by a short-channel MOSFET. This *P-N-P* transistor has a long base region, and therefore a very low current gain. A parasitic *N-P-N* transistor also exists in the IGBT, which forms a parasitic thyristor with the *P-N-P* transistor. In the IGBT structure, when a positive bias voltage larger than the threshold voltage of the DMOS is applied to the gate electrode, an inversion layer is formed along the *P*-base surface of the DMOS, and the DMOS channel is turned on. When a positive bias is applied to the collector, electrons flow from the N^+ emitter via the DMOS channel into the N^- region. This provides the base drive current for the wide base vertical *P-N-P* transistor in the IGBT structure. Since the emitter junction (J_1) for this bipolar transistor is forward biased, the P^+ substrate injects holes into the N^- base region. When the positive bias on the collector terminal of the IGBT is increased, the injected hole concentration increases and reduces the resistance of the N^- region. Consequently, the IGBT can operate at much higher current densities than the VDMOS even when it is designed to support high blocking voltages.

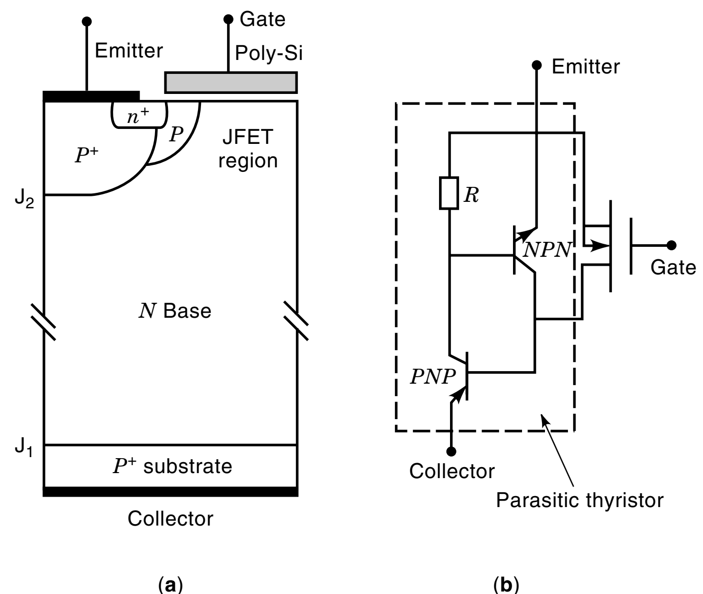

Figure 19. (a) Cross-section of the IGBT structure and (b) its equivalent circuit. The IGBT structure is similar to that of the VDMOS except that a p^+ collector is used to replace the n^+ drain in the VDMOS. This p^+ collector can inject holes into the N base to modulate the conductivity and improve the forward current capability during the forward conduction. The introduction of such p^+ layer results in a vertical *P-N-P* transistor, which in turn forms a parasitic thyristor with the parasitic *N-P-N* transistor.

As long as the gate bias is sufficiently large to produce a strong inversion layer charge of electrons at the N^- base region surface, the IGBT's forward conduction characteristic looks like that of a *P–i–N* diode. Therefore, the IGBT can also be considered as a *P–i–N* diode in series with a MOSFET. However, if the DMOS channel or JFET channel becomes pinched-off and the electron current saturates, the hole current also saturates due to the saturation of the base drive current for the *P-N-P* transistor. Consequently, the device operates with current saturation in its active region with a gate-controlled output current. This current saturation characteristic is useful for applications in which the device is required to sustain a short-circuit condition.

When the gate voltage is lower than the threshold voltage of the DMOS, the inversion layer cannot sustain and the electron current via the DMOS channel is terminated. The IGBT then operates in the forward blocking mode. A large voltage can then be supported by the reverse biased *P*-base/*N*-base region junction (J_2). Figure 20 shows the typical output characteristics of the IGBT. Because of the existence of the P^+ collector junction, IGBT also has a reverse blocking capability. Commercial IGBTs, however, are mostly using asymmetric structures, and their reverse blocking voltages are typical very low.

The IGBT was the first commercially successful device based on combining the physics of MOS-gate control with bipolar current conduction. Because of the injection of a high concentration of holes from the P^+ substrate into the N^- drift, the conductivity of the long N^- region is modulated and the IGBT exhibits *P–i–N* diode-like on-state characteristic with a low forward voltage drop. Thus, the IGBT exhibits excellent current-carrying capability with forward conduction current densities 20 times higher than that of a power MOSFET and 5 times greater than that of a bipolar transistor operating at a current gain of 10. Since the input signal for the IGBT is a voltage applied to the MOS-gate, the IGBT has the high input impedance of the power MOSFET and can be classified as a voltage-controlled device. However, unlike the power MOSFET, the switching speed of the IGBT is limited by the time taken to remove the stored charges in the *N*-region due to the injection of holes during on-state current conduction. The turn-off time for the IGBT is dictated by the conduction modulation of the *N*-region and the minority carrier lifetime. The latter can be controlled by a lifetime control process, such as electron irradiation. Although the lifetime control process can be successful in reducing the turn-off time, it was found that there is a trade-off between the on-state voltage drop (conduction loss) and the turn-off time (switching loss). A shorter minority carrier lifetime makes the switching loss of the IGBT lower, but the shorter minority carrier lifetime also results in a higher conduction loss.

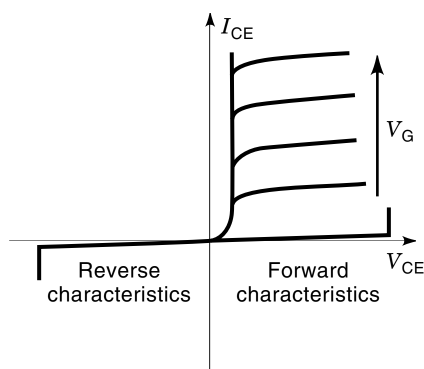

Figure 20. Output characteristics of the IGBT.

One of the problems encountered when operating the IGBT at high current levels has been the latch-up of the parasitic *P-N-P-N* thyristor structure inherent in the device structure. Latch-up of this thyristor can occur, causing losses of gate-controlled current conduction. Since the current gains of the *N-P-N* and *P-N-P* transistors increase with increasing temperature, the latching current decreases with increasing temperature. This effect is also aggravated by an increase in the resistance of the *P*-base with temperature due to a decrease in the mobility of holes. Many methods have been explored to suppress the latch-up of the parasitic thyristor, such as the use of a deep P^+ diffusion (see p. 454 of Reference 1), a shallow P^+ diffusion (see p. 456 of Reference 1), or a self-aligned sidewall diffusion of N^+ emitter (11, 12). The objective of these methods is to reduce the gain of the parasitic *N-P-N* transistor to minimal.

Traditionally, IGBTs are fabricated on a lightly doped epitaxial substrate, such as the one shown in Fig. 19(a). Because of the difficulty of growing the lightly doped epitaxial layer, the breakdown voltage of this type of IGBT is limited to below 1200 V. To benefit from such a design, an *N* buffer layer is normally introduced between the P^+ substrate and the N^- epitaxial layer, so that the whole N^--region can be depleted when the device is blocking the off-state voltage, and the electric field shape inside the N^- region is close to rectangular. This type of design is referred to as punch-through IGBT (PT IGBT), as shown in Fig. 21(a). The PT structure allows it to support the same forward blocking voltage with about half the thickness of the N^- base region of the *P-N-P* transistor, resulting in a greatly improved trade-off relationship between the forward voltage drop and the turn-off time. Thus, the PT structure together with lifetime control is preferred for IGBTs with forward blocking capabilities of up to 1200 V.

For higher blocking voltages, the thickness of the *N*-base region becomes too large for cost-effective epitaxial growth. Another type of design, the non-punch-through IGBT (NPT IGBT, as shown in Fig. 21(b)), is gaining popularity (13). In the NPT IGBTs, devices are built on an N^- wafer substrate that serves as the N^- base region. The collector is implanted from the backside of the wafer and no field stopping *N* buffer layer is applied to the NPT IGBT. In this concept, the shape of the electric field is triangular in the forward blocking state, which makes a longer N^- base region necessary to achieve the same breakdown voltage as compared with the PT IGBT. However, the NPT IGBT offers some advantages over the PT IGBT. For instance, the injection efficiency from the collector side can be more easily controlled and devices with voltage ratings as high as 4 kV can be realized. Further, by optimizing the injection efficiency of carriers from the P^+ collector layer and the transport factor of carriers in the N^- base, the trade-off between the forward voltage drop and the turn-off time for the NPT IGBT can be improved to become similar to that of the PT type IGBT (14). NPT IGBT is now widely used in high-voltage IGBT design where no lifetime control is used.

Generally speaking, the current tail in the NPT IGBT is longer than the PT IGBT, but the NPT IGBT is more robust

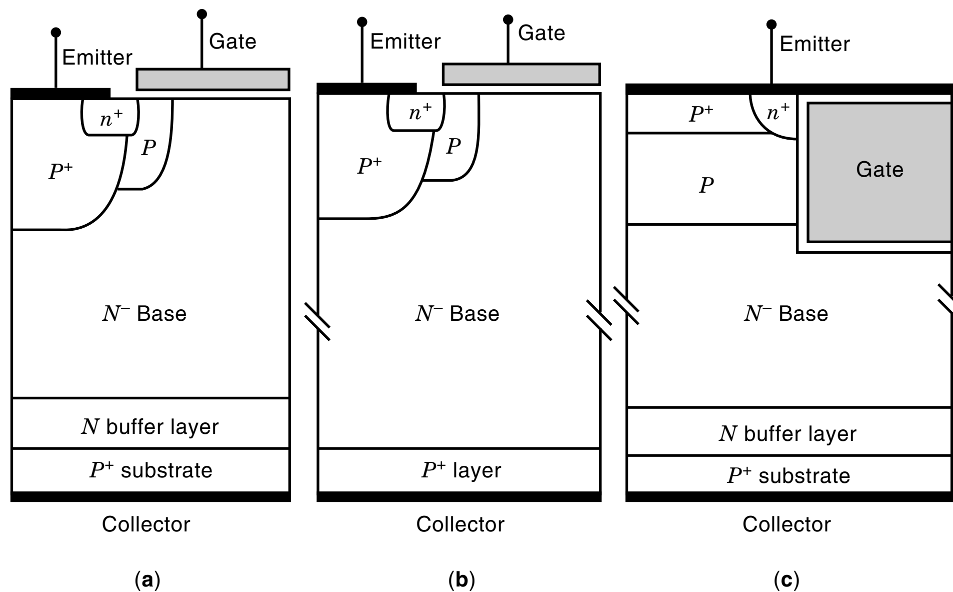

Figure 21. (a) PT IGBT structure, (b) NPT IGBT structure, and (c) UMOS gate PT IGBT structure.

than the PT IGBT, particularly under a short-circuit condition.

The trench gate IGBT (Trench IGBT, TIGBT, and UMOS-gate IGBT) structure (15) is shown in Fig. 21(c). With the UMOS structure in place of the DMOS gate structure in the IGBT, the channel density is greatly increased and the JFET region between the two adjacent *P*-base region is eliminated. In addition, the electron-hole concentration is enhanced at the bottom of the trench because an *N*-type accumulation layer forms. This creates a catenary-type carrier distribution profile in the IGBT which resembles that obtained in a thyristor or *P*–*i*–*N* diode (16). These improvements lead to a large reduction in the on-state voltage drop until it approaches that of a *P*–*i*–*N* diode, hence approaching the theoretical limitation of a silicon device. The latching current density of the UMOS IGBT structure is superior to that of the DMOS structure. This is attributed to the improved hole current flow path in the UMOS structure. As shown in Fig. 21(c), the hole current flow can take place along a vertical trajectory in the UMOS structure, while in the DMOS structure hole current flow occurs below the N^+ emitter in the lateral direction. The resistance for the hole current which causes the latch-up is determined only by the depth of the N^+ emitter region. A shallow P^+ layer can be used, as shown in the figure, to reduce this resistance. As a consequence, the safe operating area (SOA) of the UMOS IGBT structure is superior to that of the DMOS IGBT structure. Further, because of a very strong percentage of electron current flow in the trench gate IGBT, the turn-off speed of the trench-based IGBT is generally faster than the DMOS-based IGBT. To obtain better trade-off between conduction and switching losses, several improved structures and technologies, such as the carrier injection enhancement (17), local lifetime control by proton irradiation (18), and p^+/p^- collector region (19), have been proposed for the high-voltage trench IGBT. It can be anticipated that trench gate IGBTs will replace the DMOS IGBT structures in the future.

MCT

The MOS controlled thyristor (MCT) (20) is a newer commercially available semiconductor power switch that also combines the physics of MOSFET and bipolar conduction. It is basically a thyristor with two MOSFETs built into the gate structure. One of the two MOSFETs, the ON-FET, is responsible for turning the MCT on, and the other MOSFET, the OFF-FET, is responsible for turning the MCT off. There are two types of MCTs, the *N*-MCT and the *P*-MCT, and both combine the low on-state losses and large current handling capability of a thyristor structure with the advantages of MOSFET-controlled turn-on and turn-off and relatively fast switching speed. MCTs provide both an easy gate drive, due to the high input impedance of the MOS gate, and a low forward voltage drop, due to the strong conductivity modulation effect of the thyristor structure. They are expected to compete with IGBTs and GTOs in high power applications.

A cross-sectional view of a single cell of an *N*-MCT is shown in Fig. 22(a). A complete *N*-MCT is composed of thousands of these cells fabricated integrally on the same silicon wafer connected in parallel to achieve the desired current rating. The ON-FET density in an MCT can also be adjusted to a suitable percentage of the overall cell density. From Fig. 22, it is clear that the MCT has a four layer *P-N-P-N* thyristor structure and a MOS gate controlling both the turn-on and turn-off FETs. When the gate bias is zero or negative, the OFF-FET is turned on to short the upper transistors in the emitter junction. The MCT exhibits a high forward blocking voltage by supporting the voltage across the reverse biased junction J_2. When a positive bias is applied to the gate electrode, the ON-FET channel turns on and electrons are supplied to the N^- base of the *P-N-P* transistor. This results in the injection of holes from the P^+ anode into the N^- base region, and they are collected at the reverse biased junction J_2. The current in the *P* upper base created by the collection of holes across junction J_2 acts as the base drive current of the *N-P-N* transistor, which turns on the *N-P-N* transistor then triggers the regenerative feedback mechanism between the two coupled transistors within the thyristor structure. The thyristor can therefore be turned on by the application of a positive gate voltage to the MOS electrode.

The MCT can be truned off when a negative bias is applied to the gate electrode because a *P*-channel is formed by the

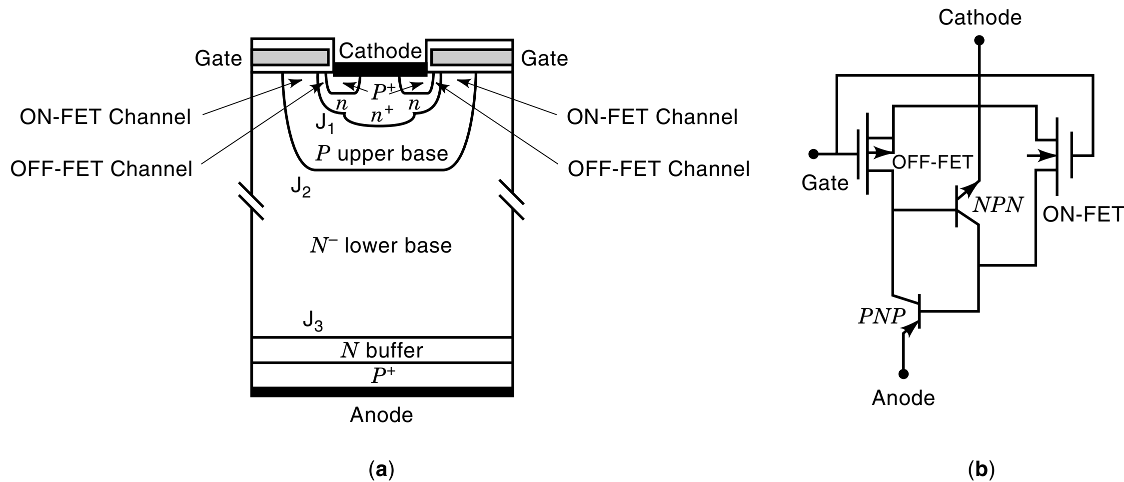

Figure 22. (a) Cross-sectional view of the MCT cell and (b) its equivalent circuit. The MCT is a five-layer semiconductor device. It is basically a thyristor with two MOSFET channels underneath the gate. One of the two MOSFETs, the ON-FET, is responsible for turning the MCT on, when the gate voltage is positive (such as +15 V), and the other MOSFET, the OFF-FET, is responsible for turning the MCT off, when the gate voltage is negative (such as −15 V).

inversion of the *N*-emitter surface. This provides a path for the flow of holes from the *P*-base region into the cathode contact that bypasses the N^+-emitter/*P*-base junction. The holes that are flowing into the *P*-base region when the thyristor was operating in its on-state can then be diverted via the *P*-channel MOSFET into the cathode electrode. This will reduce the current gain of the *N*-*P*-*N* transistor. If the resistance of the *P*-channel MOSFET is significantly low, a sufficient number of holes are diverted to the cathode by the *P*-channel MOSFET. The latch-up condition would then be broken and the thyristor would be turned off successfully.

As mentioned previously, there are two types of MCTs, the *N*-MCT and the *P*-MCT. The MCT with an *N*-channel ON-FET is called an *N*-MCT and the MCT with a *P*-channel ON-FET is called a *P*-MCT. A *P*-MCT can turn off higher currents because of a higher electron mobility in the turn-off *N*-channel MOSFET compared to that of the *N*-MCT. Thus, the first two generations of MCTs developed were *P*-MCTs. On the other hand, applications normally require an *N*-MCT, because its SOA is larger than the *P*-MCT's and its bias configuration is compatible with an *N*-channel IGBT.

An alternative *N*-MCT structure that uses an *N*-channel MOSFET for turn-off is shown in Fig. 23 (21). It has an *N*-channel OFF-FET and its bias configuration is compatible with an *N*-channel IGBT. In this *N*-MCT, a floating ohmic contact (FOC) is used to form the bridge of transferring hole current to electron channel current in the event of device turn-off.

Figure 24 shows the basic structure of a trench MCT (TMCT) cell based on trench technology (22). The TMCT can have a much smaller cell pitch than its planar counterpart, so it has a much more uniform turn-off across the chip. The TMCT has a turn-on cell to turn-off cell ratio of 1 without losing silicon area. Thus, the turn-on capability of the TMCT is also expected to be much better than the planar MCT.

Although MCTs combine the advantages of the high input impedance of the MOS gate and the low forward voltage drop of the thyristor, there are some drawbacks limiting the development of the MCT. One is that the MCT has the current filamentation problem during device turn-off that can destroy the device because of internal regenerative action and the negative temperature coefficient for the on-state voltage within the thyristor. Even if a uniform turn-off is possible, the OFF-FET channel resistance will limit the maximum turn-off current and hence the SOA of the MCT. Further, the MCT lacks the ability to saturate the anode current level, making it a fundamentally different device than the IGBT from an application viewpoint. Consequently, a number of new MOS-gated thyristors are currently being developed and studied,

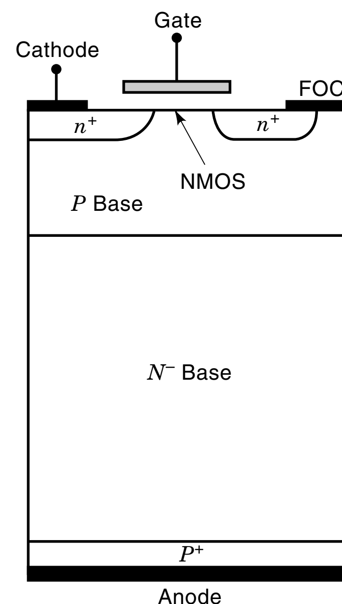

Figure 23. Cross-sectional view of the *N*-MCT cell structure. The *N*-channel MOSFET shorts the upper base via the floating ohmic contact (FOC) to the cathode, resulting in MOS-controlled turn-off.

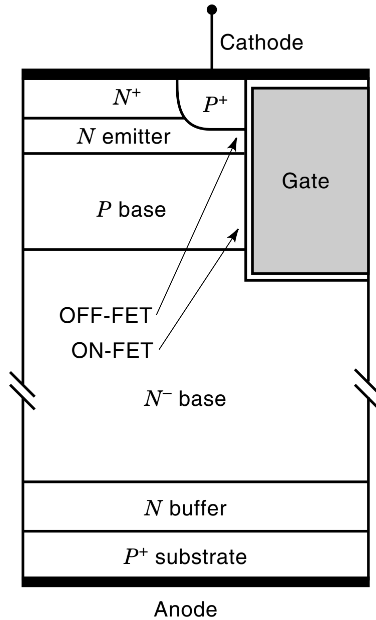

Figure 24. The basic structure of the trench MCT cell.

and commercialization of some of these newer devices is expected in the future (23,24).

Table 1 (25) lists the major electric characteristics of the switches. These switches are all based on silicon material.

INTEGRABLE LATERAL POWER DEVICES

LDMOS

The lateral double-diffusion MOSFET (LDMOS) was one of the first integrable lateral power devices developed and finds wide use in power ICs (PICs) (26,27). A schematic diagram of the cross section of the LDMOS transistor is shown in Fig. 25. The use of the self-aligned double-diffusion process results in a relatively short channel. A lightly doped, thin drift region is used to support a high drain-to-source breakdown voltage. Although the operation mechanism of the LDMOS is the same as that of the VDMOS, the LDMOS has all three electrodes on the substrate surface and there is no need of excessive epitaxial growth. In addition, it is also easily integrated with CMOS circuitry. These reasons make the LDMOS attractive in monolithic PICs, where interconnection between the LDMOS and low voltage analog and digital ICs can be easily achieved.

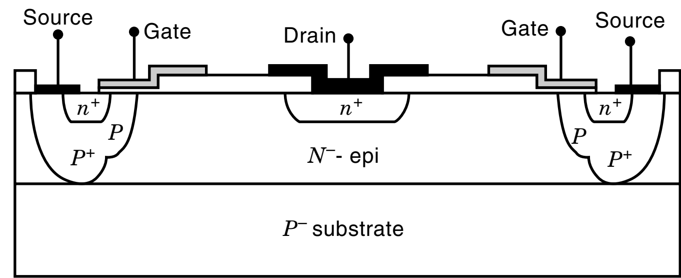

Figure 25. Cross-sectional view of the RESURF LDMOS. The LDMOS is a lateral version of the VDMOS, with all three electrodes on the substrate surface. There is no need of excessive epitaxial growth and the high voltage is supported laterally. The LDMOS is used mostly in power integrated circuits (PICs).

Significant efforts have been directed toward increasing the breakdown voltage and reducing the on-resistance of the LDMOS. The important parameters related to a required breakdown voltage are the charge in the drift layer per unit area, the length of the drift region, and the substrate doping

Table 1. Comparison among Commercially Available Power Semiconductor Switches

Switch	Number of Junctions	Control Type	Carrier Conduction Type	Highest Voltage Rating (V)	Largest Current Rating (A)	Typical Switching Frequency (Hz)	Major Limitation Factor
BJT	2	Current	Bipolar	1.5 k	800	15 k	• Low current gain • Second breakdown • Difficult to parallel
Darlington	3	Current	Bipolar	1.5 k	150	10 k	• High forward voltage drop • Second breakdown
SCR	3	Current	Bipolar	10 k	6 k	1 k	• No turn off capability • Low operating frequency
TRIAC	3	Current	Bipolar	1 k	300	400	• No turn off capability • Low operating frequency
GTO	3	Current	Bipolar	8 k	6 k	5 k	• Complex gate drive • Current filamentation
SIT	1	Current	Unipolar	1.2 k	300	200 k	• High on-resistance • Normal-on device • Low blocking gain
SITH	2	Current	Bipolar	1.5 k	500	20 k	• Normal-on device • Complex gate drive
Power MOS	2	Voltage	Unipolar	1.6 k	400	200 k	• High on-reistance
IGBT	3	Voltage	Bipolar	3.5 k	1.2	30 k	• Latch-up
MCT	4	Voltage	Bipolar	1.2 k	k 100	20 k	• Maximum turn-off current • Current filamentation • No current saturation

density. The use of charge control technology, namely, RESURF (reduced surface field) (28) enabled an increase in breakdown voltage or a corresponding reduction in its on-resistance per unit area (specific on-resistance) by almost a factor of 2.

Figure 25 also illustrates the RESURF principle. As shown in Fig. 25, a thin, lightly doped N^- type epitaxial layer is located on a P^- substrate. When the total charge in the N^- region between the P^+ source region and the N^+ drain region is large, the surface electric field near the channel reaches the critical electric field before the N^- epitaxial layer is fully depleted. Therefore, the surface limited breakdown occurs. In the opposite case, when the charge of the N^- epitaxial layer is carefully controlled so that the surface electric field is always lower than that in the bulk, and the superficial N^- layer is fully depleted prior to reaching the critical electric field, the full bulk breakdown value is achieved. The RESURF technology permits realization of lateral power transistors with breakdown voltages of up to 1200 V (29).

In high-frequency applications, the LDMOS offers the desired high switching speed, in the order of a few tens of nanoseconds, with no significant storage time. This is because current transport occurs solely by majority carriers. As a radio-frequency (RF) power device (for power amplification, not for power conversion), sub-micrometer channel LDMOS with very small gate-to-source is being used due to its low cost and high efficiency (30).

LIGBT

The lateral insulated gate bipolar transistor (LIGBT) (26,27), the lateral version of the IGBT, is another promising integrable power device. Its cross-sectional view is shown in Fig. 26(a).

In the LIGBT structure, a lightly doped drift region is also needed to support a high forward blocking voltage. The use of the RESURF technology enabled an increase in breakdown voltage similar to that of an LDMOS. An *N*-buffer layer is needed in the LIGBT to prevent vertical punch-through breakdown of the vertical *P*-*N*-*P* transistor formed by the P^+ collector and the P^- substrate. The operating mechanism of the LIGBT is similar to that of the IGBT. Because a high concentration of holes is injected from the P^+ collector into the N^- region, the conductivity of the long N^- region is modulated and the LIGBT has a specific on-resistance that is lower by about a factor of 5 to 10 than that of the LDMOS transistor. As a consequence, the LIGBT results in substantial reduction of die size for the same power handling capability in comparison with the LDMOS, which is an important factor in power ICs.

However, during turn-off, the minority carriers injected by the P^+ collector at the on-state, called the storage charge, reduce the switching speed compared to the LDMOS. Storage charge effect improves the dc performance of the LIGBT, but degrades its switching performance and limits its usefulness at the high operating frequencies. Another disadvantage of the LIGBT is the existence of substrate current due to its bipolar current conduction mechanism which may cause interference with the neighboring analog and digital ICs.

Similar to the IGBT, the LIGBT can improve its switching speed by lifetime control technique (2) at the expense of higher on-resistance. But in the PIC, reducing the switching loss of the LIGBT by lowering the minority carrier lifetime is difficult because it can also degrade the characteristics of other devices on the chip. An alternative is to use the shorted collector structure, as shown in Fig. 26(b), which provides an efficient way to remove excess carriers at turn-off, hence reducing the switching loss (31).

In a shorted collector LIGBT, the device operates like a conventional LDMOS at low drain currents. As the current increases, the voltage drop across the *N*-buffer layer resistance underneath the P^+ collector increases, and eventually the P^+ collector becomes forward biased. Holes will then be injected into the N^--base region from the P^+ collector. As soon as the injected hole density becomes comparable to the N^--base region doping, conductivity modulation takes place and the on-resistance is reduced. Because the P^+ region is shorted to the N^+ region, the emitter efficiency of the P^+ collector is lower than that of the conventional LIGBT, and the minority carrier concentration in the N^--base region is also lower. Further, because the N^- base is connected to the collector contact by the N^+ region, excess electrons can be removed rapidly by the N^+ drain, resulting in a faster turn-off process.

The LIGBT is susceptible to latch-up in the same way as the discrete IGBT because of an inherent parasitic *P*-*N*-*P*-*N* thyristor in the device. At high current levels, sufficient voltage drop occurs across the emitter/body junction and the parasitic *N*-*P*-*N* transistor turns on. The collector current of the *N*-*P*-*N* constitutes the base current of the lateral *P*-*N*-*P* tran-

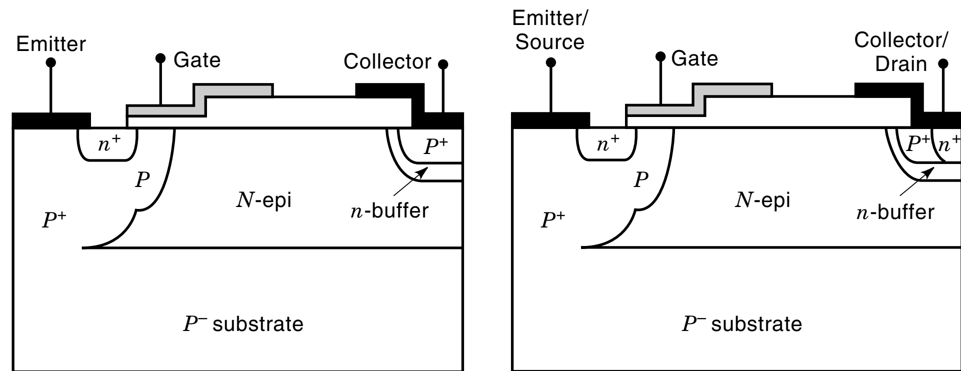

Figure 26. Cross-sectional view of (a) an LIGBT and (b) a shorted collector LIGBT.

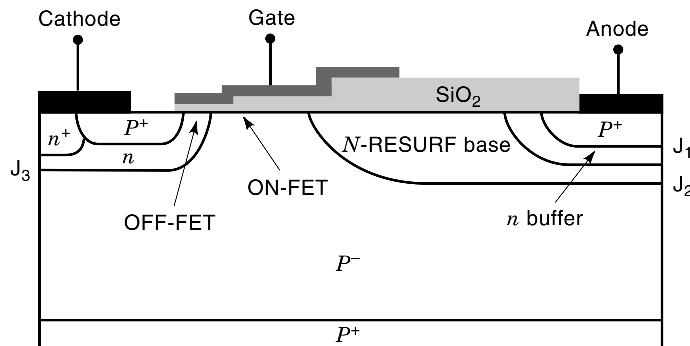

Figure 27. Cross-sectional view of an LMCT structure.

sistor. When the sum of the current gain of the two transistors reaches unity, latch-up occurs and gate control is lost. In the RESURF LIGBT, a parasitic vertical *P-N-P* transistor also exists that diverts some of the hole current into the substrate, and the latch-up threshold is increased. Another method to suppress latch-up is to lower the gain of the *P-N-P* transistor by using an *N*-buffer layer. The buffer layer can also be used to limit carrier injection by controlling the collector emitter efficiency. This increases both latch-up current and switching speed and provides a convenient way to trade off speed and forward voltage drop.

Latch-up can be either static, as discussed previously, or dynamic, that is, during switching. Dynamic latch-up occurs because of a rapid increase in the displacement current as well as the change in electron and hole current components near the emitter side. Consequently, the rate at which the device is turned on is critical to dynamic latch-up.

LMCT

The lateral MOS-controlled thyristor (LMCT) is another integrable lateral power device, the lateral type of the MCT. The cross-sectional view of an LMCT is shown in Fig. 27 (32).

The LMCT shown in Fig. 27 is built on a P^-/P^+ substrate and has a *P*-channel DMOS transistor to turn it off and an *N*-channel lateral MOS transistor to turn it on. An efficient utilization of area is obtained by using the RESURF technique to achieve high breakdown voltage. By controlling the charge in the ion-implanted *N*-RESURF base layer, it becomes fully depleted at maximum blocking voltage and the breakdown occurs in the bulk at the *N*-RESURF/P^- substrate junction (J_2). The *N*-buffer layer has a higher doping concentration than the *N*-RESURF layer and is used to prevent punch-through breakdown between the anode and the P^- substrate in the off-state.

The LMCT structure turns on by biasing the gate positively with respect to the cathode and turning on the lateral NMOS. Electrons flow into the *N*-RESURF base layer via the NMOS channel. These electrons form the base current of the *P-N-P* transistor and turn the anode junction (J_1) on. Because of the *P-N-P-N* thyristor in the LMCT, at a certain current level, regenerative action takes place and the LMCT latches up. Turn-off is achieved by biasing the gate negatively with respect to the cathode and turning the *P*-channel DMOS transistor on, which effectively shorts the base-emitter junction (J_3) of the *N-P-N* transistor and diverts its base current to the cathode contact.

Because the LMCT operates at the latch-up state in its on-state, it has much better current-carrying capability than the LDMOS and the LIGBT. However, the LMCT has a poorer current turn-off capability. Its maximum controllable turn-off current is limited by the *P*-channel DMOS transistor's channel resistance. Other variations of the LMCT also exist, which are being studied to improve performance and process compatibility (33).

ISOLATION TECHNOLOGIES FOR POWER ICS

In order to obtain high performance, low cost, small size, and high reliability of electronic equipment, the development of monolithic power ICs (PICs) has been promoted for several years. For integrating both power devices and low-voltage control ICs on the same silicon substrate, isolating the high and low voltage components is essential.

Junction Isolation

The junction isolation (JI, or P-N isolation) is the most commonly used isolation technology for PICs. With the junction isolation technique, the silicon islands where the various components are integrated are separated through reverse-biased junctions realized by the diffusion of *P* regions through the entire depth of the *N*-type epitaxial layer grown on a *P*-type substrate [Fig. 28(a)]. This technique is the most widely used because it offers the best compromise between cost and versatility. For thick epitaxial layers, up and down diffusion is used to form the isolation islands [Fig. 28(b)]. Because the isolation diffusion must extend through the entire epitaxial layer in the junction isolation, the thickness of the epitaxial layer is limited. However, blocking voltages of up to 1200 V are possible by applying RESURF technique.

The major drawback of junction isolation is that it uses a significant amount of silicon space. Moreover, it introduces an extra *P-N* junction, which has capacitance and causes leakage current to the substrate. Under unfavorable conditions, JI can also introduce parasitic components such as a *P-N-P* transistor.

Dielectric Isolation

Dielectric isolation (DI) allows the realization of silicon islands completely surrounded by oxide. DI has advantages such as low parasitic capacitance to the substrate, the absence of leakage current, and a reduction in the size of high-voltage components. However, this method requires a more complex and costly manufacturing process. Moreover, since

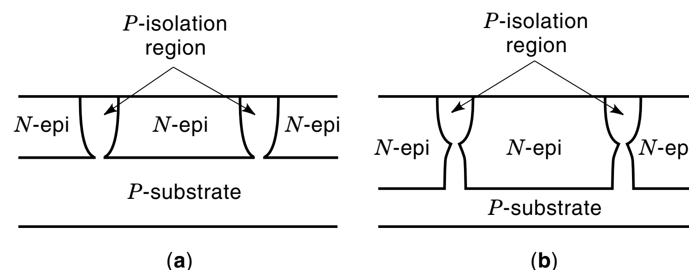

Figure 28. Junction isolation techniques: (a) down diffusion in thin epitaxial layer and (b) up and down diffusion in thick epitaxial layer.

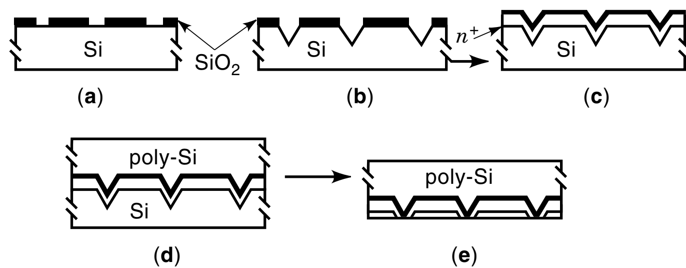

Figure 29. Fabrication process flow of an EPIC type wafer: (a) oxidation and photoetching, (b) V-groove etching, (c) N^+ deposition and isolation oxidation, (d) polysilicon epitaxial growth, and (e) grinding and polishing.

oxide is a poor conductor of heat, it limits the integration of high-power devices.

The conventional dielectric isolation process (epitaxial passivated integrated circuit, EPIC) is shown in Fig. 29. Although EPIC is the dielectric isolation technology for power ICs in mass production, it has a problem with wafer warpage caused by its thick polysilicon layer, which must be solved before the wafer size can be increased and the minimum device patterning size can be lowered. Therefore, many advanced DI technologies are also being developed.

SOI Isolation Technology. The silicon on insulator (SOI) isolation technology is one of the dielectric isolation technologies developed for ICs aimed at high-speed, high-level integration, and low power consumption. Although the origin of the SOI can be traced back to the 1934 patent of Oscar Heil on MIS structure, it is only in the 1980s that the SOI material became an evolution of the silicon and not an exotic revolutionary material. SOI material has been successfully introduced in production at the begin of the 1990s in some applications where limited volumes of wafer are required. This first IC market helped the SOI development. Among all the techniques proposed in the 1980s to perform the SOI structure, only three are still competing and are serious challengers of standard silicon: SDB (silicon direct bonding) (27), SIMOX (synthesis by implanted oxygen) (34), and smart cut (35).

Power devices fabricated on an SOI substrate have attracted a lot of attentions in the area of smart power integrated circuits. The reason for this is the many advantages offered by SOI over a conventional bulk substrate. V-groove etching, trench, or LOCOS (local oxidation of the silicon) isolation processes between adjacent devices on SOI offer true dielectric isolation and allow simple integration of power and logic devices on the same substrate. It is also possible to achieve significant improvements in breakdown voltage and switching speed with an SOI substrate. These advantages can be attributed to the excellent insulating properties of silicon dioxide in these devices. The buried oxide helps sustain a high electric field, which results in high breakdown voltage, and confines the carriers, which reduces the switching time of minority carrier devices.

On the other hand, the buried oxide underneath the device is also a good thermal insulator. The thermal conductivity of silicon dioxide is only 1.4 W/K-m compared to 140 W/K-m for bulk silicon. This significant difference impedes the dissipation of the heat generated inside the device. Therefore, the temperature rise inside an SOI device can be much higher than that in a bulk device.

In the SOI technologies, the SDB dielectric isolation technology and the SIMOX are promising for high-voltage power IC applications.

SDB Dielectric Isolation Technique. The SDB (silicon direct bonding) dielectric isolation technique, which is also called BESOI (bond and etch back SOI) or DISDB (dielectric isolation by silicon wafer direct bonding) is widely used to prepare starting SOI material. The use of SDB technology has made SOI technology more viable and cost-effective. The key process steps are shown in Fig. 30. Starting from two silicon wafers, at least one with an oxide layer on top, these two wafers are bonded together using van der Walls forces. Subsequent annealing increases the mechanical strength of the bonded interface by the chemical reaction which can occur at this interface. One of the substrates is then thinned down to the required thickness from several hundred microns by mechanical grinding and polishing.

The SDB dielectric isolation technique is a promising candidate for power ICs, because thick silicon islands, which are required for handling high current and high voltage, can be easily fabricated. The adjacent devices isolation on the SDB wafer can be provided by etching V-grooves, by RIE (reactive ion etching) trench isolation, or even by LOCOS if the SOI layer is thin enough.

The SDB dielectric isolation technique is used not only in the isolation between lateral power devices such as LDMOSs and low-voltage integrated circuits, it is also used in the isolation between vertical power devices such as IGBTs and low-voltage integrated circuits. There are two methods for this application, as shown in Fig. 31 and Fig. 32.

Figure 31 shows the process flow of the first method for the SDB wafer to be used in isolating PICs. First, in the SDB wafer, the silicon and oxide films where the power device is to be formed are removed by wet etching. Then, the etched place is buried with Si epitaxial growth and the lapping and polishing are performed. Next, lateral isolation regions are formed by the conventional steps of groove etching (trench etching or V-groove etching), thermal oxidation for isolation film formation, and refilling the groove with polysilicon.

Another method is shown in Fig. 32. First, in the wafer A, the region where the low-voltage controlling circuit is formed is masked and etched. After thermal oxidation is carried out to form an isolation film, the wafer is polished until the opti-

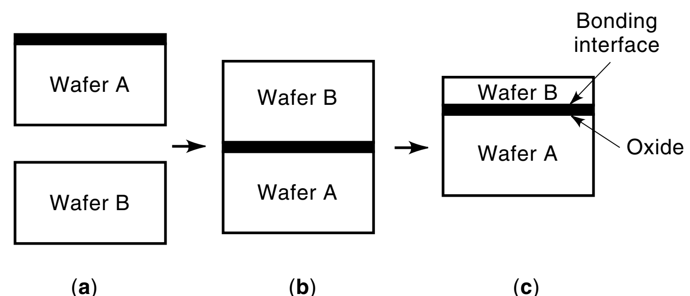

Figure 30. Fabrication process flow of a SDB wafer: (a) oxidation of slice A, (b) cleaning and bonding, and (c) annealing, then grinding and polishing of wafer B.

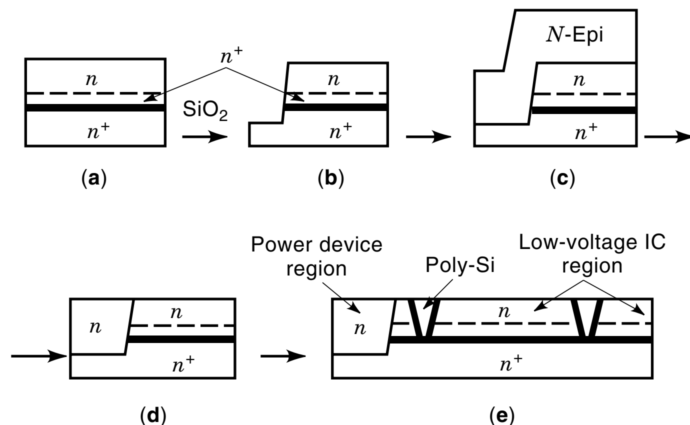

Figure 31. Method one for using SDB in power ICs: (a) SDB wafer, (b) silicon and oxide film etching, (c) epitaxial growth, (d) lapping and surface polishing, and (e) lateral isolation.

cally flat Si–SiO_2 coexistent surface is exposed. Then, wafer bonding is performed. After that, the side of the bonded wafer on which devices are fabricated is ground and polished to a thickness of several microns. Finally, lateral isolation regions are formed by conventional steps of groove etching, thermal oxidation for isolation film formation, and refilling of the groove with polysilicon.

SIMOX. The SIMOX (synthesis by implanted oxygen, separation by implanted oxygen, or selective implantation with oxygen) technique is considered to be one of the most advanced and promising SOI technologies for high-density CMOS circuits. The key processes of the SIMOX technology are shown in Fig. 33. First, oxygen ions are implanted into the silicon underneath the initial silicon surface. Then, a postimplantation annealing regenerates the crystalline quality of the silicon layer remaining over the oxide. This annealing also drives the chemical reaction that forms the stoichiometric oxide buried in the silicon wafer. Although the ideal annealing conditions are not fully identified yet, it is known that good SIMOX must be annealed at about 1320°C, for 6 h, in argon ambient containing 1% of oxygen.

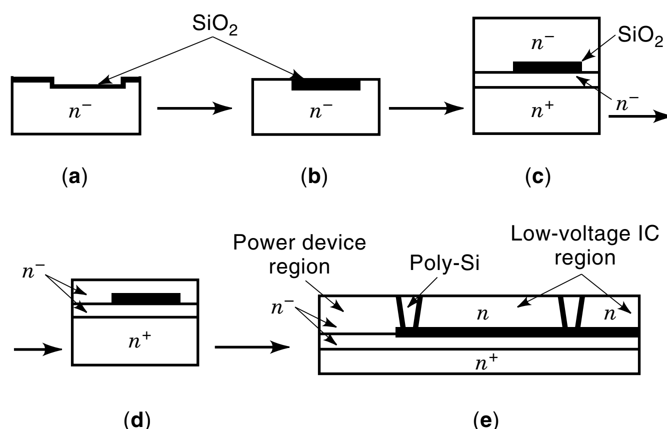

Figure 32. Method two for using SDB in power ICs: (a) etching and oxidation, (b) grinding and polishing, (c) bonding, (d) grinding and polishing, and (e) lateral isolation.

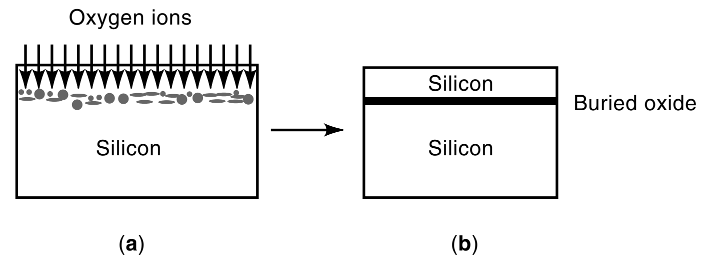

Figure 33. Fabrication process of a SIMOX wafer: (a) oxygen ion implantation and (b) high-temperature annealing.

A number of SIMOX variants have been explored, as shown in Fig. 34. The thin or thick buried oxide (BOX) can be fabricated by using a lower implanted dose or a higher implanted dose. The thin and thick silicon film over the BOX can be obtained by using lower implanted energy or higher implanted energy. The double SIMOX structure is fabricated by two sequential oxygen implants. The thicker silicon layer can be achieved by using epitaxy technology. The interrupted BOX and the totally isolated island SIMOX can be processed by masked implantation.

The main disadvantage of the SIMOX technology is the need of >1300°C annealing, which could be a limitation for the 300-mm wafers.

The SIMOX technology is also considered to be one of the most promising dielectric isolation technologies for PICs. Figure 35 shows a PIC using SIMOX technology. As shown in Fig. 35, the wafer can be selectively implanted with oxygen ions and forms local dielectric isolation. This local SIMOX technology with the epitaxy and the trench technology offers the unique opportunity to integrate monolithic devices with vertical current path-like VDMOS or IGBT with lateral low-voltage control circuits by means of rather standard VLSI process steps. In contrast to other dielectric isolation technologies, no “exotic” process steps like selective epitaxy or mechanical back-lapping and surface polishing are necessary.

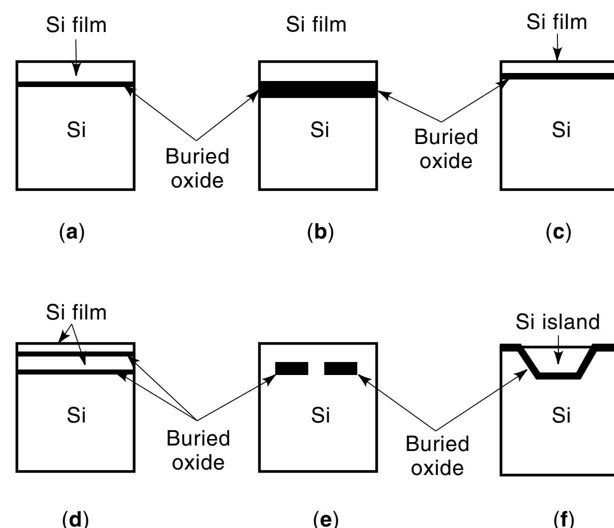

Figure 34. Various types of SIMOX structure: (a) thin BOX, (b) thick BOX, (c) thin Si film, (d) double SIMOX, (e) interrupted BOX, and (f) totally isolated island SIMOX.

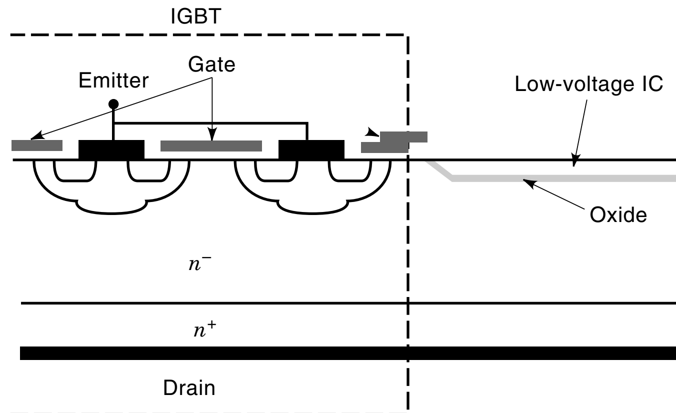

Figure 35. One possible way of using SIMOX technology in power IC.

Further, since the SIMOX technology involves forming a local buried oxide layer, improved smart power discrete devices are possible using SIMOX technology and they are being studied and developed.

SPSDB Technique. The single-silicon polysilicon direct bonding (SPSDB) technique (36–38), which is also called laminated dielectric isolation (LDI or laminated DI) is a new SDB isolation method used for power ICs. Figure 36 shows the cross-sectional view of the SPSDB wafer. The SPSDB wafer has inverse V-groove isolation regions with a narrow isolation width of about 5 μm, which is independent of breakdown voltage. Consequently, the SPSDB technique has a very high packing density. Further, the SPSDB wafer has a simple fabrication process and is suited to mass production. Furthermore, the same design rules as those for the EPIC DI wafer can be utilized.

The process of the SPSDB is based on those of the EPIC and the SDB. Similar to the process of the conventional EPIC, first the ⟨100⟩ silicon wafer is oxide-masked and the V-grooves are etched with a preferential etching solution like KOH. When the grooves are completed, the etching stops automatically. The wafer is then oxidized and subsequently covered with a polysilicon layer about 80-μm thick. Next, the polysilicon layer is lapped and polished to about 10 μm thickness. The polished wafer is treated with NH_4OH–H_2O_2 solution at 70°C, rinsed with deionized water, and dried using a spin-dryer. Then, a single crystal silicon wafer is placed onto the polished polysilicon surface. These wafers are bonded at 1100°C for 2 h in an oxidizing atmosphere. Finally, the slice is inverted and the original single crystal substrate is lapped and polished until the silicon islands are isolated from each other.

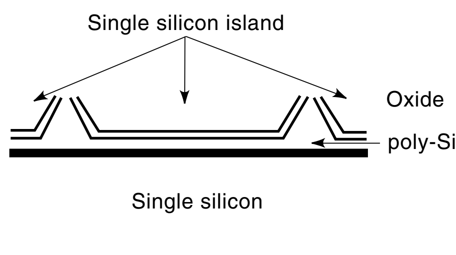

Figure 36. Cross-sectional view of the SPSDB wafer.

The SPSDB wafer has an unchanging warpage height during high-temperature heat treatments and has a high bonding strength comparable to that of the thermal oxidizing layer interface.

SiC-BASED POWER DEVICES

In recent years, silicon carbide (SiC) has received increased attention as a potential material for power devices operating at high temperatures, high power levels, and high frequencies due to its unique material properties (39,40). Silicon carbide has a bandgap about three times wider than that of silicon (3.0 eV for 6H–SiC and 3.25 eV for 4H–SiC), high avalanche breakdown electric field of 2~4 × 10^6 V/cm, high saturated electron drift velocity of 2 × 107 cm/s, and high thermal conductivity of 4.9 W/cm-K. The high breakdown electric field allows the use of much higher doping and thinner layers for a given blocking voltage than silicon devices, resulting in much lower specific on-resistance for unipolar devices. Further, high thermal conductivity and high saturated electron drift velocity also make SiC especially attractive in the power device arena.

With the arrival of commercial single crystal substrates of 6H– and 4H–SiC and the ability to grow high-quality SiC epitaxial layers, the silicon carbide process has developed rapidly and the fabrication of power devices has become viable.

Among the SiC substrates, the two SiC polytypes 6H and 4H are having the biggest impact on power devices. Although the 6H–SiC has the best single crystal quality of the established polytypes, the 4H–SiC is more attractive for power devices than the 6H–SiC. The reason for this is that the electron mobility in the 4H–SiC is two times that of the 6H–SiC in the direction perpendicular to the *c*-axis and almost 10 times that of the 6H–SiC in the direction parallel to the *c*-axis.

Compared with silicon and gallium arsenide materials, SiC has a lower mobility in the inversion layer and very small diffusion rates for dopants. These factors are limiting the pace of SiC power device development.

Schottky and *P–i–N* Junction Diodes

High voltage SiC Schottky rectifiers are already commercially available. Figure 37 shows the cross-sectional view of a SiC Schottky barrier diode (SBD, or Schottky diode) with high-resistance edge termination. This structure consists of an N^+-doped substrate with backside ohmic contact, a lightly doped epitaxial layer, and a topside Schottky contact with a high-resistance termination. The Schottky diode is fabricated by evaporating a high work function metal, such as titanium, nickel, or gold, onto the lightly doped epitaxial layer to form the Schottky contact and by depositing a metal onto the back of the N^+ substrate to form the back ohmic contact. The high resistivity edge termination is achieved by implanting argon, which damages the exposed semiconductor to create a high resistance region. This process is self-aligned to the Schottky contact because the Schottky metal acts as a mask preventing damage under the contact (41).

Because of the higher breakdown electric field, the epitaxial layer of the SiC SBD can have a higher doping and thinner

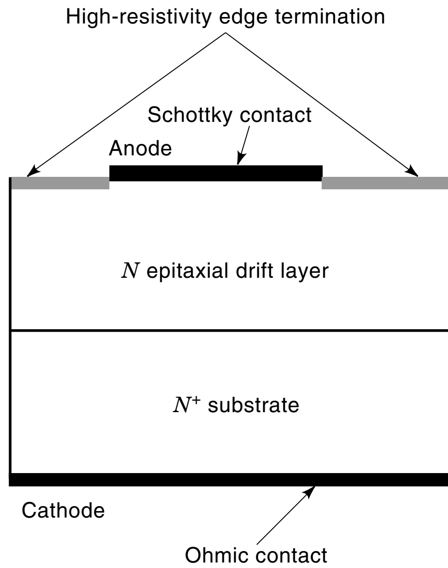

Figure 37. Cross-sectional view of a SiC Schottky diode with high-resistivity edge termination. The SiC Schottky diode consists of an N^+ doped substrate with backside ohmic contact, a lightly doped N epitaxial layer, and a topside Schottky contact surrounded by a high-resistance termination.

drift layer at the same blocking capability when compared with that of a gallium arsenide (GaAs) SBD and Si SBD. The specific on-resistance of a 1000-V 4H–SiC SBD is 15 times lower than that of a 1000 V GaAs SBD and over 200 times lower than that of a 1000 V silicon SBD. However, due to the high electron mobility of the GaAs material, a GaAs SBD has a lower specific on-resistance than an SiC SBD at block voltages lower than 200 V.

SBDs have also been found to have excellent reverse recovery and reverse bias leakage characteristics even at high operating temperatures. They are likely to replace silicon P–i–N rectifiers in high voltage power electronic circuits.

A somewhat more complex device is a SiC P–i–N diode, as shown in Fig. 38 (42). A high concentration N^+ SiC epitaxial

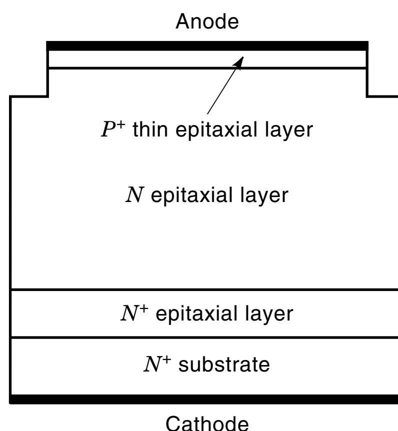

Figure 38. Cross-sectional view of the SiC P–i–N diode. In the SiC P–i–N diode structure, a high concentration N^+ SiC epitaxial layer is grown on the N^+ SiC wafer, and then an N-type epitaxial drift region and a high concentration P^+ thin epitaxial layer are grown. A mesa edge termination is formed by using RIE (reactive ion etching) technology to block reverse voltage.

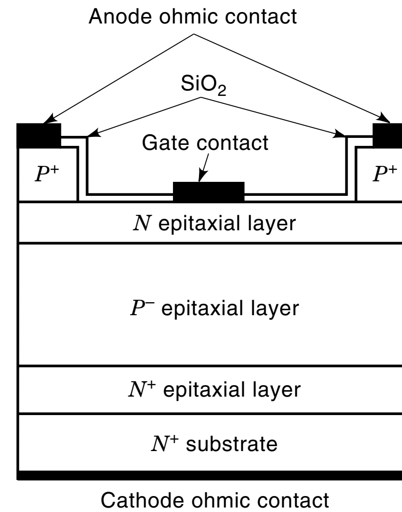

Figure 39. Cross-sectional view of the SiC N-P-N-P thyristor. In the SiC thyristor structure, epitaxy is used to grow all semiconductor layers. The RIE is used to define the gate contact. Because an N^+ substrate is used, the resulting thyristor is a p-type thyristor.

layer is grown on the N^+ SiC wafer, and then an N-type epitaxial drift region and a high concentration P^+ thin epitaxial layer are grown. A mesa edge termination is formed by using RIE (reactive ion etching) technology to block reverse voltage. This etching process was self-aligned in that the aluminum etch mask also acts as the top contact to the P^+ layer in the device.

Thyristors

For very high voltage (5 to 10 kV) applications, such as traction control and high-voltage dc transmission, silicon bipolar devices have much lower on-state losses than silicon unipolar devices. The same is expected to be true for SiC bipolar devices. At these very high voltages, a single SiC thyristor could replace a stack of silicon thyristors and thereby achieve a lower forward voltage drop (43). In addition it is expected that properly designed SiC bipolar devices, which take advantage of the high breakdown field of SiC, will have lower voltage drops than silicon bipolar devices. The most promising SiC thyristor structure to date has been an N-P-N-P device in 4H–SiC, as shown in Fig. 39. This structure utilized a mesa structure, with all of the doping being done in situ during epitaxy. The device periphery was terminated using an RIE mesa.

MOSFET

SiC vertical power MOSFETs have a strong advantage over those made in silicon because the drift layer may use a 10 times higher doping level and one-tenth the thickness for a given breakdown voltage because of the much higher breakdown electric field of the SiC material. Ultimately, this could translate into specific on-resistances as low as 1/300th that of an equivalent Si device. Because dopant diffusion rates in SiC material are very small, the UMOS process was considered to be the most suitable for making SiC power MOSFETs because the UMOS process can rely on epitaxy to form the channel region (44).

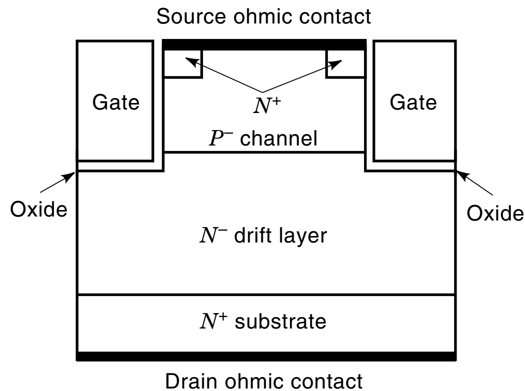

Figure 40. Cross-sectional view of the SiC UMOSFET. An N^- epitaxial drift layer is grown on the N^+ substrate, and then a P-type channel layer is epitaxially grown. N^+ source regions are formed by using implantation into the p-type channel layer. RIE is used to form the trenches on either side of the p-type layer.

The cross-sectional view of an SiC UMOSFET structure is shown in Fig. 40. An N^- epitaxial drift layer is grown on the N^+ substrate, and then a P-type channel layer is grown. N^+ source regions are formed by using implantation into the P-type channel layer. RIE is used to form the trenches on either side of the P-type layer. After the gate oxide (SiO_2) is grown and annealed, ohmic contacts are formed on the source and drain areas. Finally, the gate metal and interconnect metal are defined. During operation, current flows from the N^+ source contacts through an inversion channel layer to the N^- drift layer and they are collected by the N^+ drain. The current flow from source to drain is controlled by the voltage on the gate electrode.

It was found, however, that there is a high electric field at the corners of the trenches, which restricts the breakdown voltage of the UMOSFET far lower than its theoretical breakdown voltage. Further, the side-wall inversion channel mobilities in the SiC UMOSFET are lower than those in planar SiC MOSFETs, which leads to a severe increase in the on-resistance of the device. A planar high-voltage SiC MOSFET using double-implants has also been reported (45), which avoids both of these problems by forming the inversion channel on the silicon. Shown in Fig. 41, the DMOS structure is formed by using multiple energy boron and nitrogen implants. Both implants are activated simultaneously at 1600°C for 30 min in an argon ambient. Then, the wafer is thermally oxidized to obtain the gate oxide, and the polysilicon is deposited or aluminum is thermally evaporated to form the gate electrode.

6H–SiC UMOS IGBT has also been experimentally demonstrated (44). Its structure is similar to that of the SiC UMOS MOSFET, shown in Fig. 40, except that a p^+ collector substrate is used to replace the n^+ substrate of the MOSFET. The SiC IGBT has better forward conduction capability than that of the SiC MOSFET at high blocking voltage (>1000 V) and high operating temperature (>200°C) due to the conductivity modulation.

SIT

Although the SiC SIT (static induction transistor) has a structure resembling that of the UMOSFET, as shown in Fig. 42 (46), the operation mechanism is significantly different. The SiC SIT is a vertical device with an ohmic source contact on the top and an ohmic drain contact on the back of the wafer. Between the N^+ source and N^+ drain regions is an N^- epitaxial drift layer whose doping is one of the factors that determines the device breakdown voltage and pinch-off voltage. Trenches are etched to define the channel region, and Schottky gate contacts are formed in the bottom and along the sidewalls of the trench. Majority carriers flow from the source contact to the drain contact through the N-type channel region. By applying a negative voltage to the gate contact, the current flow can be modulated and even decreased to zero when depletion regions under each gate contact meet in the middle of the channel.

The SiC SIT is ideally siuited to high-power microwave devices owing to the remarkable transport properties, very high breakdown field strength, and thermal conductivity of SiC.

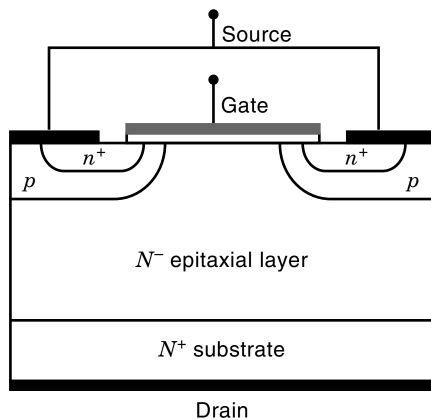

Figure 41. Structure of the planar SiC MOSFET. An N^- epitaxial drift layer is first grown on the N^+ substrate. The DMOS structure is formed by using multiple high energy boron (p region) and nitrogen implants (n region).

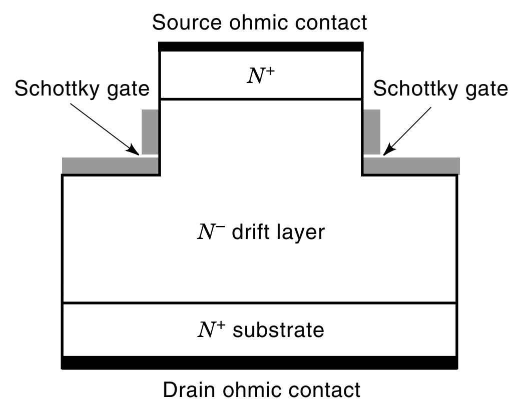

Figure 42. Cross-sectional view of the SiC SIT. An N^- epitaxial drift layer is grown on the N^+ substrate, and then an N^+ layer is grown. Trenches are etched to define the channel region, and Schottky gate contacts are formed in the bottom and along the sidewalls of the trench.

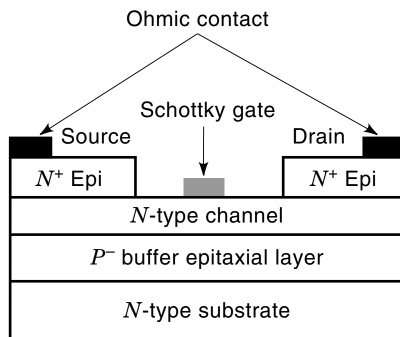

Figure 43. Cross-sectional view of the RF SiC MESFET. In the SiC RF MESFET structure, all semiconductor layers are epitaxially grown. The RIE is used to define the Schottky gate.

The SiC SIT is being developed as a discrete power microwave transistor for operation at frequencies up to S-band.

RF MESFET

The cross-sectional view of an RF SiC MESFET is shown in Fig. 43 (47). This device is a lateral device with both source and drain contacts on the top surface of the wafer. The MESFET epitaxial structure consists of an undoped P-buffer layer, N-type channel layer, and N^+ contact layer. The majority of carriers flow laterally from source to drain, confined to the N-type channel by the P^- buffer layer and controlled by the Schottky gate electrode.

For RF Si LDMOS, GaAs MESFET, and SiC MESFET, the device parameters that are important in different power densities are low field electron mobility, breakdown electric field, and electron saturation velocity. At a doping density of 1×10^{17} cm^{-3} the electron mobility of 4H–SiC is 560 cm^2/V-s, which is slightly lower than that of Si (800 cm^2/V-s) and significantly lower than that of GaAs (4900 cm^2/V-s). On the other hand, the breakdown electric field of 4H–SiC is about 10 times that of Si and GaAs, and the saturated drift velocity is 2 times that of Si and GaAs. Consequently, at low voltages, GaAs MESFETs, which have the highest electron mobility, have the highest power density. The higher power density of SiC MESFETs is only achieved at drain voltages higher than those normally used with either Si or GaAs devices.

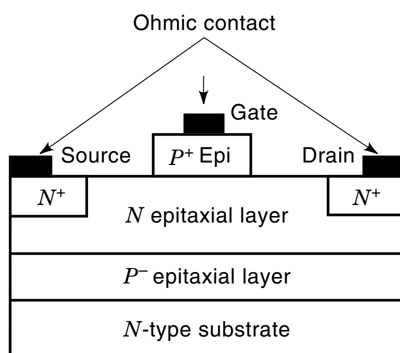

Figure 44. Cross-sectional view of the SiC JFET. In the SiC RF JFET structure, a p^- epitaxial layer is grown on the N-type substrate, and then an N-type epitaxial layer and high concentration P^+ epitaxial layer are grown. The P^+ mesa is formed by using RIE technology and N^+ source and drain regions are formed by using ion-implantation.

RF JFET

High-frequency SiC JFETs are of interest for high-temperature RF applications because a much lower gate leakage current can be obtained with a P-N junction at high temperature than with a Schottky gate (48). The cross-section of a SiC RF JFET (shown in Fig. 44) is similar to that of the RF MESFET, except a P^+ SiC epitaxial region with an ohmic contact on top is used in place of a Schottky contact, and ion-implanted N^+ source and drain contact regions are used in place of the N^+ epitaxial region.

BIBLIOGRAPHY

1. B. J. Baliga, *Power Semiconductor Devices,* Boston: PWS Publishing, 1996.
2. B. J. Baliga and E. Sun, Comparison of gold, platinum, and electron irradiation for controlling lifetime in power rectifier, *IEEE Trans. Electron Devices,* **ED-24**: 685–688, 1977.
3. M. Mehrotra and B. J. Baliga, Very low forward drop JBS rectifiers fabricated using submicron technology, *IEEE Trans. Electron Devices,* **ED-40**: 2131–2132, 1993.
4. L. Tu and B. J. Baliga, Controlling the characteristics of the MPS rectifier by variation of area of Schottky region, *IEEE Trans. Electron Devices,* **ED-40**: 1307–1315, 1993.
5. M. Kitagawa, K. Matsushita, and A. Nakagawa, High-voltage emitter short diode (ESD), *Japan. J. Appl. Phys.,* **35**: 5998–6002, 1997.
6. H. Schlangenotto et al., Improved recovery of fast power diodes with self-adjusting, *p* emitter efficiency, *IEEE Electron Device Lett.,* **10**: 322–324, 1989.
7. J. Blanc, Practical application of MOSFET synchronous rectifiers, *13th Int. Telecommun. Energy Conf.,* **INTELEC-91**: 495–501, 1991.
8. N. Mohan, T. M. Undeland, and W. P. Robbins, *Power Electronics,* 2nd ed., New York: Wiley, 1995.
9. K. K. Ng, *Complete Guide to Semiconductor Devices,* New York: McGraw-Hill, 1995.
10. T. Syau, P. Venkatraman, and B. J. Baliga, Comparison of ultralow specific on-resistance UMOSFET structure: The ACCUFET, EXTFET, INVFET, and conventional UMOSFET's, *IEEE Trans. Electron Devices,* **ED-41**: 800–808, 1994.
11. M. Mori, Y. Nakano, and T. Tanaka, An insulated gate bipolar transistor with a self-aligned DMOS structure, *IEEE Int. Electron Devices Meet. Dig.,* **IEDM-88**: 1988, pp. 813–816.
12. T. P. Chow et al., A self-aligned short process for insulated-gate bipolar transistor, *IEEE Trans. Electron Devices,* **ED-39**: 1317–1321, 1992.
13. G. Miller and J. Sack, A new concept for non-punch through IGBT with MOSFET like switching characteristics, *Conf. Rec. IEEE Power Electron. Specialists Conf.,* 1989, pp. 21–25.
14. T. Laska, G. Miller, and J. Niedermeyer, 2000-V non-punch through IGBT with high ruggedness, *Solid-State Electron.,* **35**: 681–685, 1992.
15. H. R. Chang et al., Insulated gate bipolar transistor (IGBT) with trench gate structure, *IEEE Int. Electron Devices Meet. Dig.,* **IEDM-87**: 1987, pp. 674–677.

16. M. Harada et al., 600-V trench IGBT in comparison with planar IGBT, *Int. Symp. 1994 IEEE Int. Symp. Power Semicond. Devices and IC's,* **ISPSD-94**: 1994, pp. 411–416.
17. I. Omura et al., Carrier injection enhancement effect of high voltage MOS devices, *1997 IEEE Int. Symp. Power Semicond. Devices and ICs,* **ISPSD-97**: 217–220, 1997.
18. S. Eicher et al., Advanced lifetime control for reducing turn-off switching loss of 4.5 kV IEGT devices, *1998 IEEE Int. Symp. Power Semicond. Devices and ICs,* **ISPSD-98**: 39–42, 1998.
19. E. Suekawa et al., High voltage IGBT (HV-IGBT) having p^+/p^- collection region, *1998 IEEE Int. Symp. Power Semicond. Devices and ICs,* **ISPSD-98**: 249–252, 1998.
20. V. A. K. Temple, MOS controlled thyristors (MCT's), *IEEE Int. Electron Devices Meet. Dig.,* **IEDM-84**: 1984, pp. 282–285.
21. Q. Huang et al., Analysis of *n*-channel MOS controlled thyristors, *IEEE Trans. Electron Devices,* **ED-38**: 1612–1618, 1991.
22. A. Q. Huang, Analysis of the inductive turn-off of double gate MOS controlled thyristor, *IEEE Trans. Electron Devices,* **ED-43**: 1029–1032, 1996.
23. B. J. Baliga, Trends in power semiconductor devices, *IEEE Trans. Electron Devices,* **ED-43**: 1727–1731, 1996.
24. A. Q. Huang, A unified view of the MOS gated thyristors, *Solid-State Electronics,* submitted.
25. *Power Semiconductors,* ed. 36, Englewood: DATA Digest, An IHS group company, 1996.
26. P. A. H. Hart (ed.), *Bipolar and Bipolar-MOS Integration,* Amsterdam, The Netherlands: Elsevier Science B. V., 1994.
27. B. Murari, F. Bertotti, and G. A. Vignola (eds.), *Smart Power ICs,* New York: Springer, 1995.
28. J. A. Appels and H. M. J. Vaes, High voltage thin layer devices (RESURF devices), *IEEE Int. Electron Devices Meet. Dig.,* **IEDM-79**: 1979, pp. 238–241.
29. A. W. Ludikuize, A versatile 700-1200 V IC process for analog and switching applications, *IEEE Trans. Electron Devices,* **ED-38**: 1582–1589, 1991.
30. A. Wood, C. Dragon, and W. Burger, High performance silicon LDMOS technology for 2-GHz RF power amplifier applications, *IEEE Int. Electron Devices Meet. Dig.,* **IEDM-96**: 1996, pp. 87–90.
31. M. R. Simpson et al., Analysis of the lateral insulated gate transistor, *IEEE Int. Electron Devices Meet. Dig.,* **IEDM-85**: 1985, pp. 740–743.
32. M. N. Darwish, A new lateral MOS controlled thyristor, *IEEE Electron Device Lett.,* **11**: 256–257, 1990.
33. A. Q. Huang, Lateral insulated gate *P–i–N* transistor (LIGPT)—A new MOS gate lateral power device, *IEEE Electron Device Lett.,* **17**: 297–299, 1996.
34. H. Haddara (ed.), *Characterization Methods for Submicron MOSFETs,* Boston: Kluwer, 1995.
35. M. Bruel, B. Aspar, and A. J. Auberton-Herve, Smart cut: A new silicon on insulator material technology based on hydrogen implantation and wafer bonding, *Jpn. J. Appl. Phys., part 1,* **36**: 1636–1641, 1997.
36. Y. Inoue, Y. Sugawara, and S. Kurita, Characteristics of new dielectric isolation wafers for high voltage powe ICs by single-Si poly-Si direct bonding (SPSDB) technique, *IEEE Trans. Electron Devices,* **ED-42**: 356–358, 1995.
37. W. G. Easier et al., Polysilicon to silicon bonding in laminated dielectrically isolated (LDI) wafers, *Proc. 1st Int. Symp. Semicond., Wafer Bonding,* 1991, pp. 223–229.
38. Y. Sugawara, Y. Inoue, and S. Kurita, New dielectric isolation for high voltage power ICs by single silicon poly silicon direct bonding (SPSDB) technique, *1992 IEEE Int. Symp. Power Semicond. Devices and ICs,* **ISPSD-92**: 1992, pp. 316–319.
39. C. E. Weitzel et al., Silicon carbide high-power devices, *IEEE Trans. Electron Devices,* **ED-43**: 1732–1739, 1996.
40. J. W. Palmour et al., Silicon carbide for power devices, *1997 IEEE Int. Symp. Power Semicond. Devices and ICs,* **ISPSD-97**: 1997, pp. 25–32.
41. M. Bhatnagar, P. Mclarty, and B. J. Baliga, Silicon-carbide high-voltage (400 V) Schottky barrier diodes, *IEEE Electron Device Lett.,* **13**: 501–503, 1992.
42. P. G. Neudeck and C. Fazi, Positive temperature coefficient of breakdown voltage in 4H–SiC *P-N* junction rectifiers, *IEEE Electron Device Lett.,* **18**: 96–98, 1997.
43. J. W. Palmour et al., Silicon carbide substrates and power devices, in H. Goronkin and U. Mishra (eds.), *Compound Semiconductors 1994,* Bristol, UK: *Inst. Phys. Pub.,* **141**: 377–382, 1994.
44. N. Ramungul et al., A fully planarized 6H-SiC UMOS insulated-gate bipolar-transistor, *54th Annu. Device Res. Conf.,* 1996, pp. 24–26.
45. J. N. Shenoy, J. A. Cooper, and M. R. Melloch, High-voltage double-implanted power MOSFET's in 6H-SiC, *IEEE Electron Device Lett.,* **18**: 93–95, 1997.
46. R. R. Siergiej et al., High power 4H–SiC static induction transistors, *IEEE Int. Electron Devices Meet. Dig.,* **IEDM-95**: 1995, pp. 353–356.
47. C. E. Weitzel, Comparison of Si, GaAs, and SiC RF MESFET power densities, *IEEE Electron Device Lett.,* **16**: 451–453, 1995.
48. C. E. Weitzel et al., SiC microwave power MESFET's and JFET's, in H. Goronkin and U. Mishra (eds.), *Compound Semiconductors 1994,* Bristol, UK: IOP Publishing, **141**: 389–394, 1994.

ALEX Q. HUANG
BO ZHANG
Virginia Polytechnic Institute and State University

POWER DIODES. See DIODES FOR POWER ELECTRONICS.

POWER DISTRIBUTION SYSTEMS. See UNDERGROUND DISTRIBUTION SYSTEMS.