A brief survey of power semiconductor devices

The most fundamental challenge in power semiconductor design is obtaining a high breakdown voltage, while maintaining low forward voltage drop and on-resistance. A closely related issue is the longer switching times of high-voltage low-on-resistance devices. The trade-off between breakdown voltage, on-resistance, and switching times, is a key distinguishing feature of the various power devices.

The breakdown voltage of a reverse-biased *p-n* junction and its associated depletion region is a function of doping level: obtaining a high breakdown voltage requires low doping concentration, and hence high resistivity, in the material on at least one side of the junction. This high-resistivity region is usually the dominant contributor to the on-resistance of the device, and hence high-voltage devices must have higher on-resistance than low-voltage devices. In *majority carrier* devices, including the MOSFET and Schottky diode, this accounts for the first-order dependence of on-resistance on rated voltage. However, *minority carrier* devices, including the diffused-junction *p-n* diode, the bipolar junction transistor (BJT), the insulated-gate bipolar transistor (IGBT), and the thyristor family (SCR, GTO, MCT), exhibit another phenomenon known as *conductivity modulation*. When a minority-carrier device operates in the on-state, minority carriers are injected into the lightly-doped high-resistivity region by the forward-biased *p-n* junction. The resulting high concentration of minority carriers effectively reduces the apparent resistivity of the region, reducing the on-resistance of the device. Hence, minority-carrier devices exhibit lower on-resistances than comparable majority-carrier devices.

However, the advantage of decreased on-resistance in minority-carrier devices comes with the disadvantage of decreased switching speed. The conducting state of any semiconductor device is controlled by the presence or absence of key charge quantities within the device, and the turn-on and turn-off switching times are equal to the times required to insert or remove this controlling charge. Devices operating with conductivity

modulation are controlled by their injected minority carriers. The total amount of controlling minority charge in minority-carrier devices is much greater than the charge required to control an equivalent majority-carrier device. Although the mechanisms for inserting and removing the controlling charge of the various devices can differ, it is nonetheless true that, because of their large amounts of minority charge, minority-carrier devices exhibit switching times which are significantly longer than those of majority-carrier devices. In consequence, majority-carrier devices find application at lower voltage levels and higher switching frequencies, while the reverse is true of minority-carrier devices.

Modern power devices are fabricated using up-to-date processing techniques. The resulting small feature size allows construction of highly interdigitated devices, whose unwanted parasitic elements are less significant. The resulting devices are more rugged and well-behaved than their predecessors.

A detailed description of power semiconductor device physics and switching mechanisms is beyond the scope of this book. Selected references on power semiconductor devices are listed in the bibliography [8-18].

Power diodes

As discussed above, the diffused-junction *p-n* diode contains a lightly-doped or intrinsic high-resistivity region, which allows a high breakdown voltage to be obtained. As illustrated in Fig. 4.24(a), this region comprises one side of the p-n^- junction (denoted n^-); under reverse-biased conditions, essentially all of the applied voltage appears across the depletion region inside the n^- region. On-state conditions are illustrated in Fig. 4.24(b). Holes are injected across the forward-biased junction, and become minority carriers in the n^- region. These minority carriers effectively reduce the apparent resistivity of the n^- region via conductivity modulation. Essentially all of the forward current $i(t)$ is comprised of holes which diffuse across the *p-n* region, and then recombine with

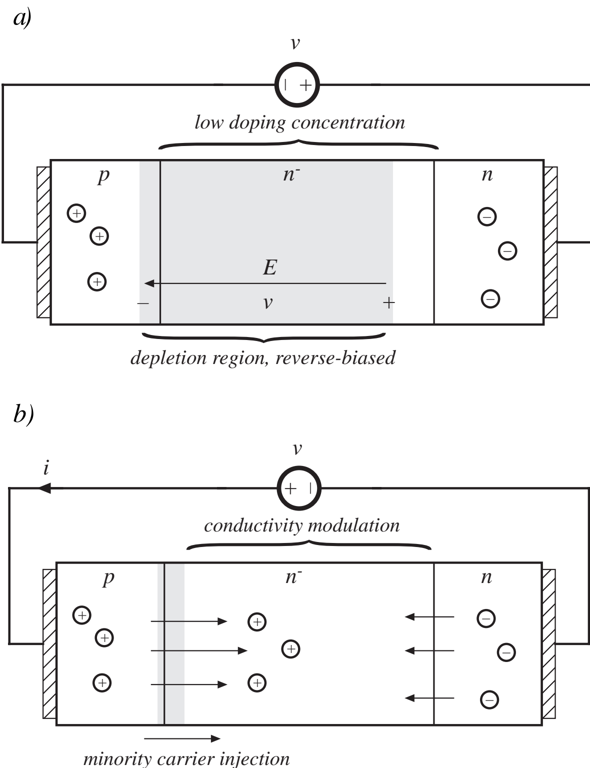

Fig. 4.24. Power diode: (a) under reverse-bias conditions, (b) under forward-bias conditions.

electrons from the n region.

Typical switching waveforms are illustrated in Fig. 4.25. The familiar exponential *i-v* characteristic of the *p-n* diode is an equilibrium relation. During transients, significant deviations from the exponential characteristic are observed; these deviations are associated with changes in the stored minority charge. As illustrated in Fig. 4.25, the diode operates in the off-state during interval *(1)*, with zero current and negative voltage. At the beginning of interval *(2)*, the current increases to some positive value. This current charges the effective capacitance of the reverse-biased diode, supplying charge to the depletion region and increasing the voltage $v(t)$. Eventually, the voltage becomes positive, and the diode junction becomes forward-biased. The voltage may rise to a peak value of several volts, or even several tens of volts, reflecting the somewhat large resistance of the lightly-doped n^- region. The forward-biased p-n^- junction continues to inject minority charge into the n^- region. As the total minority charge in the n^- region increases, conductivity modulation of the n^- region causes its effective resistance to decrease, and hence the forward voltage drop $v(t)$ also decreases. Eventually, the diode reaches equilibrium, in which the minority carrier injection rate and recombination rate are equal. During interval *(3)*, the diode operates in the on-state, with forward voltage drop given by the diode static *i-v* characteristic.

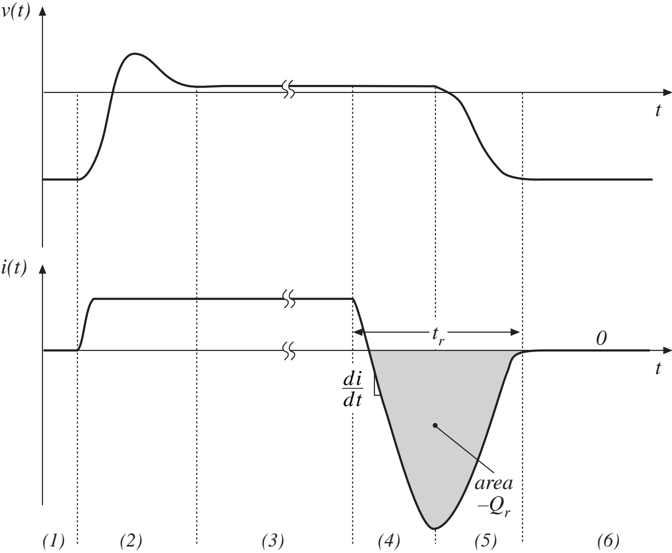

Fig. 4.25. Diode voltage and current waveforms. Interval (1): off state. Interval (2): turn-on transition. Interval (3): on state. Intervals (4) and (5): turn-off transition. Interval (6): off state.

The turn-off transient is initiated at the beginning of interval *(4)*. The diode remains forward-biased while minority charge is present in the vicinity of the diode p-n^- junction. Reduction of the stored minority charge can be accomplished either by active means, via negative terminal current, or by passive means, via recombination. Normally, both mechanisms occur simultaneously. The charge Q_r contained in the negative portion of the diode turn-off current waveform is called the *recovered charge*. The portion of Q_r occurring during interval *(4)* is actively-removed minority charge. At the end of interval *(4)*, the stored minority charge in the vicinity of the p-n^- junction has been removed, such that the diode junction becomes reverse-biased and is able to block negative voltage. The depletion region effective capacitance is then charged during interval *(5)* to the negative off-state

voltage. The portion of Q_r occurring during interval *(5)* is charge supplied to the depletion region, as well as minority charge that is actively removed from remote areas of the diode. At the end of interval *(5)*, the diode is able to block the entire applied reverse voltage. The length of intervals *(4)* and *(5)* is called the *reverse recovery time* t_r. During interval *(6)*, the diode operates in the off-state. The diode turn-off transition, and its influence on switching loss in a PWM converter, is discussed further in section 4.3.2.

Diodes are rated according to the length of their reverse recovery time t_r. *Standard recovery* rectifiers are intended for 50Hz or 60Hz operation; reverse recovery times of these devices are usually not specified. *Fast recovery* rectifiers and *ultra-fast recovery* rectifiers are intended for use in converter applications. The reverse recovery time t_r, and sometimes also the recovered charge Q_r, are specified by manufacturers of these devices. Ratings of several commercial devices are listed in Table 4.1.

Schottky diodes are essentially majority-carrier devices whose operation is based on the rectifying characteristic of a metal-semiconductor junction. These devices exhibit negligible minority stored charge, and their switching behavior can be adequately modeled simply by their depletion-region capacitance and equilibrium exponential *i-v* characteristic. Hence, an advantage of the Schottky diode is its fast switching speed. An even more important advantage of Schottky diodes is their low forward voltage drops, especially in devices rated 45V or less. Schottky diodes are restricted to low breakdown voltages; very few commercial devices are rated to block 100V or more. Their off-state reverse currents are considerably higher than those of *p-n* junction diodes. Characteristics of several commercial Schottky rectifiers are also listed in Table 4.1.

Table 4.1. Characteristics of several commercial power rectifier diodes

Part number	*Rated max voltage*	*Rated avg current*	V_F *(typical)*	t_r *(max)*
Fast recovery rectifiers				
1N3913	400V	30A	1.1V	400ns
SD453N25S20PC	2500V	400A	2.2V	2µs
Ultra-fast recovery rectifiers				
MUR815	150V	8A	0.975V	35ns
MUR1560	600V	15A	1.2V	60ns
RHRU100120	1200V	100A	2.6V	60ns
Schottky rectifiers				
MBR6030L	30V	60A	0.48V	
444CNQ045	45V	440A	0.69V	
30CPQ150	150V	30A	1.19V	

Another important characteristic of a power semiconductor device is whether its on-resistance and forward voltage drop exhibits a positive temperature coefficient. Such devices, including the MOSFET and IGBT, are advantageous because multiple chips can be easily paralleled, to obtain high-current modules. These devices also tend to be more rugged and less susceptible to hot-spot formation and second-breakdown problems. Diodes cannot be easily connected in parallel, because of their negative temperature coefficients: an imbalance in device characteristics may cause one diode to conduct more current than the others. This diode becomes hotter, which causes it to conduct even more of the total current. In consequence, the current does not divide evenly between the paralleled devices, and the current rating of one of the devices may be exceeded. Since BJT's and thyristors are controlled by a diode junction, these devices also exhibit negative temperature coefficients and have similar problems when operated in parallel. Of course, it is possible to parallel any type of semiconductor device; however, use of matched dice, a common thermal substrate, and/or external circuitry may be required to cause the on-state currents of the devices to be equal.

Metal-Oxide-Semiconductor Field-Effect Transistor (MOSFET)

The power MOSFET is a modern power semiconductor device having gate lengths approaching one micron. The power device is comprised of many small parallel-connected enhancement-mode MOSFET cells, which cover the surface of the silicon die. A cross-section of one cell is illustrated in Fig. 4.26. Current flows vertically through the silicon wafer: the metallized drain connection is made on the bottom of the chip, while the metallized source connection and polysilicon gate are on the top surface. Under normal operating conditions, in which $v_{ds} \geq 0$, both the *p-n* and *p-n*$^-$ junctions are reverse-biased. Fig. 4.27(a) illustrates operation of the device in the off-state. The applied drain-to-source voltage then appears across the depletion region of the *p-n*$^-$ junction. The n^- region is lightly doped, such that the desired breakdown voltage rating is attained. Fig. 4.27(b) illustrates operation in the on-state, with a sufficiently large positive gate-to-source voltage. A channel then forms at the surface of the *p* region, underneath the gate. The drain current flows through the n^- region, channel, *n* region, and

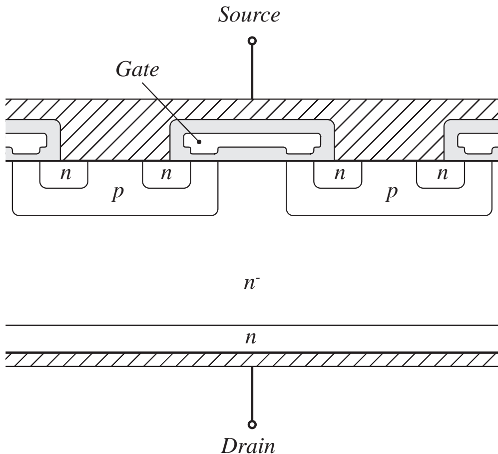

Fig. 4.26. DMOS n-channel power MOSFET structure. Crosshatched regions are metallized contacts. Shaded regions are insulating silicon dioxide layers.

out through the source contact. The on-resistance of the device is the sum of the resistances of the n^- region, the channel, the source and drain contacts, etc. As the breakdown voltage is increased, the on-resistance becomes dominated by the resistance of the n^- region. Since there are no minority carriers to cause conductivity modulation, the on-resistance increases rapidly as the breakdown voltage is increased to several hundred volts and beyond.

The p-n^- junction is called the *body diode*; as illustrated in Fig. 4.27(c), this junction forms an effective diode in parallel with the MOSFET channel. The body diode can become forward-biased when the drain-to-source voltage $v_{ds}(t)$ is negative. This diode is capable of conducting the full rated current of the MOSFET. However, most MOSFETs are not optimized with respect to the speed of their body diodes, and the large peak currents which flow during the reverse recovery transition of the body diode can cause device failure. Several manufacturers produce MOSFETs which contain fast recovery body diodes; these devices are rated to withstand the peak currents during the body diode reverse recovery transition.

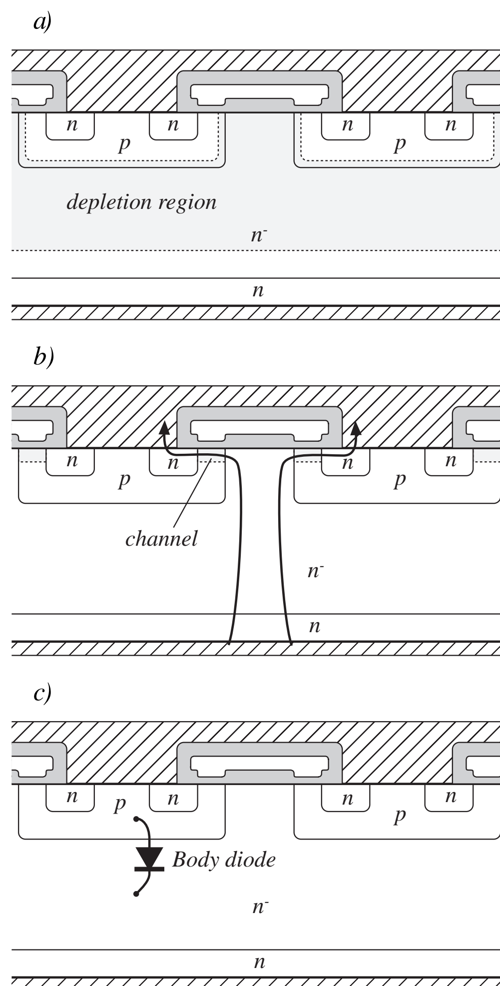

Fig. 4.27. Operation of the power MOSFET: (a) in the off-state, v_{ds} appears across the depletion region in the n^- region; (b) current flow through the conducting channel in the on-state; (c) body diode due to the p-n^- junction.

Typical n-channel MOSFET static switch characteristics are illustrated in Fig. 4.28. The drain current is plotted as a function of the gate-to-source voltage, for various values of drain-to-source voltage. When the gate-to-source voltage is less than the threshold voltage V_{th}, the device operates in the off-state. A typical value of V_{th} is 3V. When the gate-to-source voltage is greater than 6 or 7V, the device operates in the on-state; typically, the gate is driven to 12 or 15V to ensure minimization of the forward voltage drop. In the on-state, the drain-to-source voltage V_{DS} is roughly proportional to the drain current I_D. The MOSFET is able to conduct peak currents well in excess of its average current rating, and the nature of the static characteristics is unchanged at high current levels. Logic-level power MOSFETs are also available, which operate in the on-state with a gate-to-source voltage of

5V. A few p-channel devices can be obtained, but their properties are inferior to those of equivalent n-channel devices.

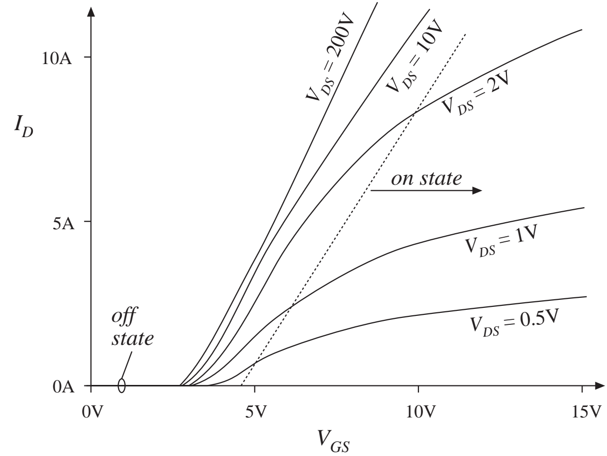

Fig. 4.28. Typical static characteristics of a power MOSFET. Drain current I_D is plotted vs. gate-to-source voltage V_{GS}, for various values of drain-to-source voltage V_{DS}.

The on-resistance and forward voltage drop of the MOSFET have a positive temperature coefficient. This property makes it relatively easy to parallel devices. High current MOSFET modules are available, containing several parallel-connect chips.

The major capacitances of the MOSFET are illustrated in Fig. 4.29. This model is sufficient for qualitative understanding of the MOSFET switching behavior; more accurate models account for the parasitic junction field-effect transistor inherent in the DMOS geometry. Switching times of the MOSFET are determined essentially by the times required for the gate driver to charge these capacitances. Since the drain current is a function of the gate-to-source voltage, the rate at which the drain current changes is dependent on the rate at which the gate-to-source capacitance is charged by the gate drive circuit. Likewise, the rate at which the drain voltage changes is a function of the rate at which the gate-to-drain capacitance is charged. The drain-to-source capacitance leads directly to switching loss in PWM converters, since the energy stored in this capacitance is lost during the transistor turn-on transition. Switching loss is discussed in section 4.3.

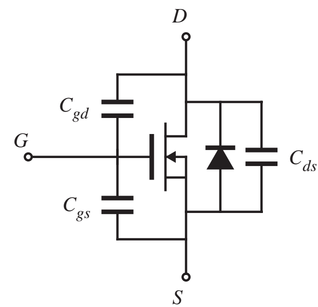

Fig. 4.29. MOSFET equivalent circuit which accounts for the body diode and effective terminal capacitances.

The gate-to-source capacitance is essentially linear. However, the drain-to-source and gate-to-drain capacitances are strongly nonlinear: these incremental capacitances vary as the inverse square root of the applied capacitor voltage. For example, the dependence of the incremental drain-to-source capacitance can be written in the form

$$C_{ds}(v_{ds}) = \frac{C_0}{\sqrt{1 + \frac{v_{ds}}{V_0}}} \tag{4-4}$$

where C_0 and V_0 are constants that depend on the construction of the device. These capacitances can easily vary by several orders of magnitude as v_{ds} varies over its normal operating range. For $v_{ds} >> V_0$, Eq. (4-4) can be approximated as

$$C_{ds}(v_{ds}) \approx C_0 \sqrt{\frac{V_0}{v_{ds}}} = \frac{C_0'}{\sqrt{v_{ds}}} \qquad (4\text{-}5)$$

These expressions are used in section 4.3.3 to determine the switching loss due to energy stored in C_{ds}.

Characteristics of several commercially-available power MOSFETs are listed in Table 4.2. The gate charge Q_g is the charge which the gate drive circuit must supply to the MOSFET to raise the gate voltage from zero to some specified value (typically 10V), with a specified value of off-state drain-to-source voltage (typically 80% of the rated V_{DS}). The total gate charge is the sum of the charges on the gate-to-drain and the gate-to-source capacitance. The total gate charge is to some extent a measure of the size and switching speed of the MOSFET.

Unlike other power devices, MOSFETs are usually not selected on the basis of their rated average current. Rather, on-resistance and its influence on conduction loss are the limiting factors, and MOSFETs may be operated at average currents somewhat less than the rated value.

MOSFETs are usually the device of choice at voltages less than or equal to approximately 400-500V. At these voltages, the forward voltage drop is competitive or superior to the forward voltage drops of minority-carrier devices, and the switching speed is approximately one order of magnitude faster. Typical switching times are in the range 50ns – 200ns. At voltages greater than 400-500V, minority-carrier devices having lower forward voltage drops, such as the IGBT, are usually preferred. The only exception is in applications where the high switching speed overrides the increased cost of silicon required to obtain acceptably low conduction loss.

Table 4.2. Characteristics of several commercial n-channel power MOSFETs

Part number	*Rated max voltage*	*Rated avg current*	R_{on}	Q_g *(typical)*
IRFZ48	60V	50A	0.018Ω	110nC
IRF510	100V	5.6A	0.54Ω	8.3nC
IRF540	100V	28A	0.077Ω	72nC
APT10M25BNR	100V	75A	0.025Ω	171nC
IRF740	400V	10A	0.55Ω	63nC
MTM15N40E	400V	15A	0.3Ω	110nC
APT5025BN	500V	23A	0.25Ω	83nC
APT1001RBNR	1000V	11A	1.0Ω	150nC

Bipolar Junction Transistor (BJT)

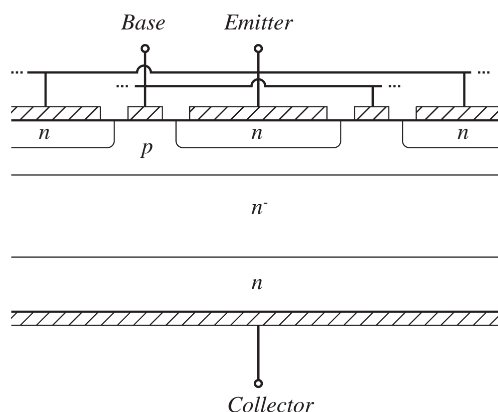

Fig. 4.30 Power BJT structure. Crosshatched regions are metallized contacts.

A cross-section of an NPN power BJT is illustrated in Fig. 4.30. As with other power devices, current flows vertically through the silicon wafer. A lightly-doped n^- region is inserted in the collector, to obtain the desired voltage breakdown rating. The transistor operates in the off-state (cutoff) when the *p-n* base-emitter junction and the p-n^- base-collector junction are reverse-biased; the applied collector-to-emitter voltage then appears essentially across the depletion region of the p-n^- junction. The transistor operates in the on-state (saturation) when both junctions are forward-biased; substantial minority charge is then present in the *p* and n^- regions. This minority charge causes the n^- region to exhibit a low on-resistance via the conductivity modulation effect. Between the off-state and the on-state is the familiar active region, in which the *p-n* base-emitter junction is forward-biased and the p-n^- base-collector junction is reverse-biased. When the BJT operates in the active region, the collector current is proportional to the base region minority charge, which in turn is proportional (in equilibrium) to the base current. There is additionally a fourth region of operation known as *quasi-saturation*, occurring between the active and saturation regions. Quasi-saturation occurs when the base current is insufficient to fully saturate the device; hence, the minority charge present in the n^- region is insufficient to fully reduce the n^- region resistance, and high transistor on-resistance is observed.

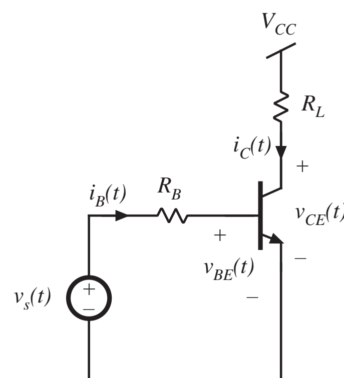

Fig. 4.31. Circuit for BJT switching time example.

Consider the simple switching circuit of Fig. 4.31. Figure 4.32 contains waveforms illustrating the BJT turn-on and turn-off transitions. The transistor operates in the off-state during interval *(1)*, with the base-emitter junction reverse-biased by the source voltage $v_s(t) = -V_{s1}$. The turn-on transition is initiated at the beginning of interval *(2)*, when the source voltage changes to $v_s(t) = +V_{s2}$. Positive current is then supplied by source v_s to the base of the BJT. This current first charges the capacitances of the depletion regions of the reverse-biased base-emitter and base-collector junctions. At the end of interval *(2)*, the base-emitter voltage exceeds zero sufficiently for the base-emitter junction to become forward-biased. The length of interval *(2)* is called the

turn-on delay time. During interval *(3)*, minority charge is injected across the base-emitter junction from the emitter into the base region; the collector current is proportional to this minority base charge. Hence during interval *(3)*, the collector current increases. Since the transistor drives a resistive load R_L, the collector voltage also decreases during interval *(3)*. This causes the voltage to reduce across the reverse-biased base-collector depletion region (Miller) capacitance. Increasing the base current I_{B1} (by reducing R_B or increasing V_{s2}) increases the rate of change of both the base region minority charge and the charge in the Miller capacitance. Hence, increased I_{B1} leads to a decreased turn-on switching time.

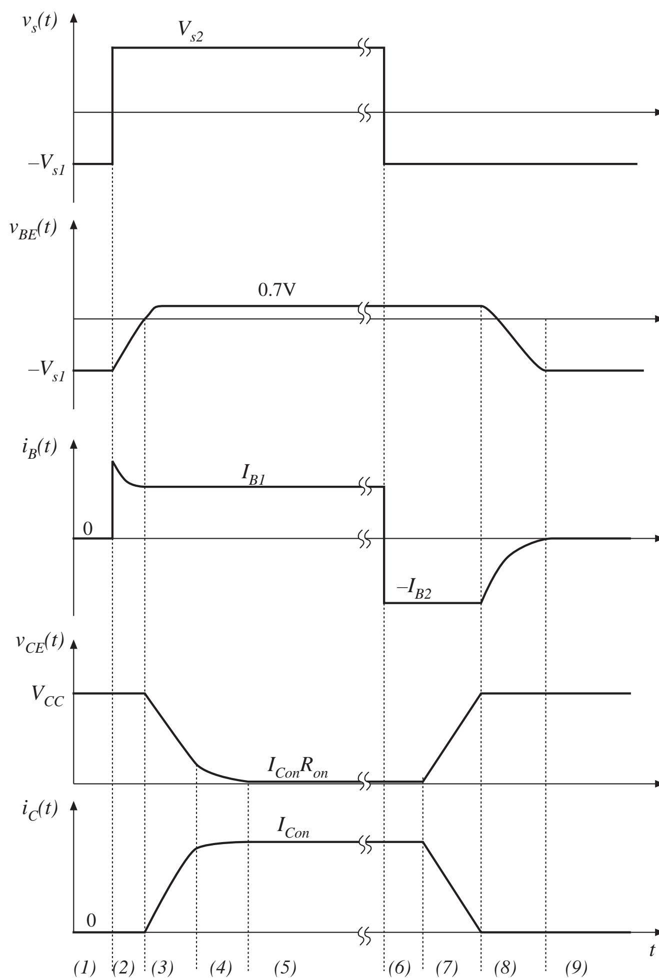

Fig. 4.32. BJT turn-on and turn-off transition waveforms.

Near or at the end of interval *(3)*, the base-collector p-n^- junction becomes forward-biased. Minority carriers are then injected into the n^- region, reducing its effective resistivity. Depending on the device geometry and the magnitude of the base current, a *voltage tail* (interval *(4)*) may be observed as the apparent resistance of the n^- region decreases via conductivity modulation. The BJT reaches on-state equilibrium at the beginning of interval *(5)*, with low on-resistance and with substantial minority charge present in both the n^- and p regions. This minority charge significantly exceeds the amount necessary to support the active region conduction of the collector current I_{Con}; its magnitude is a function of $I_{B1} - I_{Con}/\beta$, where β is the active-region current gain.

The turn-off process is initiated at the beginning of interval *(6)*, when the source voltage changes to $v_s(t) = -V_{s1}$. The base-emitter junction remains forward-biased as long as minority carriers are present in its vicinity. Also, the collector current continues to be $i_C(t) = I_{Con}$ as long as the minority charge exceeds the amount necessary to support the active region conduction of I_{Con}, i.e., as long as *excess charge* is present. So during

interval *(6)*, a negative base current flows equal to $-I_{B2} = (-V_{s1} - v_{BE}(t))/R_B$. This negative base current actively removes the total stored minority charge. Recombination further reduces the stored minority charge. Interval *(6)* ends when all of the excess minority charge has been removed. The length of interval *(6)* is called the *storage time*. During interval *(7)*, the transistor operates in the active region. The collector current $i_C(t)$ is now proportional to the stored minority charge. Recombination and the negative base current continue to reduce the minority base charge, and hence the collector decreases. In addition, the collector voltage increases, and hence the base current must charge the Miller capacitance. At the end of interval *(7)*, the minority stored charge is equal to zero, the base-emitter junction can become reverse-biased. The length of interval *(7)* is called the turn-off time or fall time. During interval *(8)*, the reverse-biased base-emitter junction capacitance is discharged to voltage $-V_{s1}$. During interval *(9)*, the transistor operates in equilibrium, in the off-state.

It is possible to turn the transistor off using $I_{B2} = 0$; for example, we could let V_{s1} be approximately zero. However, this leads to very long storage and turn-off switching times. If $I_{B2} = 0$, then all of the stored minority charge must be removed passively, via recombination. From the standpoint of minimizing switching times, the base current waveform of Fig. 4.33 is ideal. The initial base current I_{B1} is large in magnitude, such that charge is inserted quickly into the base, and the turn-on switching times are short. A compromise value of equilibrium on-state current I_{Bon} is chosen, to yield a reasonably low collector-to-emitter forward voltage drop, while maintaining moderate amounts of excess stored minority charge and hence keeping the storage time reasonably short. The current $-I_{B2}$ is large in magnitude, such that charge is removed quickly from the base and hence the storage and turn-off switching times are minimized.

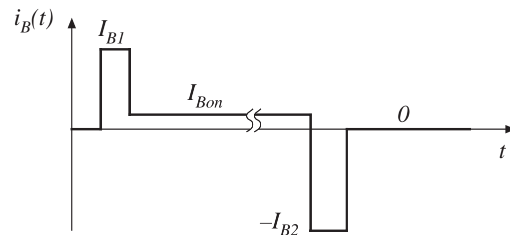

Fig. 4.33. Ideal base current waveform for minimization of switching times.

Unfortunately, in most BJT's, the magnitudes of I_{B1} and I_{B2} must be limited because excessive values lead to device failure. As illustrated in Fig. 4.34, the base current flows laterally through the *p*-region. This current leads to a voltage drop in the resistance of the *p* material, which influences the voltage across the base-emitter junction. During the turn-off transition, the base current $-I_{B2}$ causes the base-emitter junction voltage to be greater in the center of base region, and smaller at the edges near the base contacts. This causes the collector current to focus near the center of the base region. In a similar fashion, a large I_{B1} causes the collector current to crowd near the edges of the base region during the turn-on transition. Since the collector-to-emitter voltage and collector current are

simultaneously large during the switching transitions, substantial power loss can be associated with current focusing. Hence hot spots are induced at the center or edge of the base region. The positive temperature coefficient of the base-emitter junction current (corresponding to a negative temperature coefficient of the junction voltage) can then lead to thermal runaway and device failure. Thus, to obtain reliable operation, it may be necessary to limit the magnitudes of I_{B1} and I_{B2}. It may also be necessary to add external *snubber networks* which the reduce the instantaneous transistor power dissipation during the switching transitions.

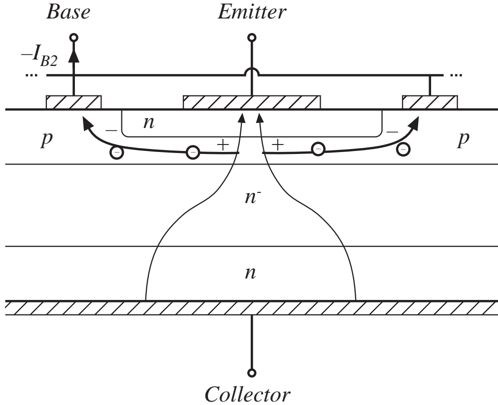

Fig. 4.34. A large I_{B2} leads to focusing of the emitter current away from the base contacts, due to the voltage induced by the lateral base region current.

Steady-state characteristics of the BJT are illustrated in Fig. 4.35. In Fig. 4.35(a), the collector current I_C is plotted as a function of the base current I_B, for various values of collector-to-emitter voltage V_{CE}. The cutoff, active, quasi-saturation, and saturation regions are identified. At a given collector current I_C, to operate in the saturation region with minimum forward voltage drop, the base current I_B must be sufficiently large. The slope dI_C/dI_B in the active region is the current gain β. It can be seen that β decreases at high current —near the rated current of the BJT, the current gain decreases rapidly and hence it is difficult to fully saturate the device. Collector current I_C is plotted as a function of collector-to-emitter voltage V_{CE} in Fig. 4.35(b), for various values of I_B. The breakdown voltages BV_{sus}, BV_{CEO} and BV_{CBO} are illustrated. BV_{CBO} is the avalanche breakdown voltage of the base-collector junction, with the emitter open-circuited or

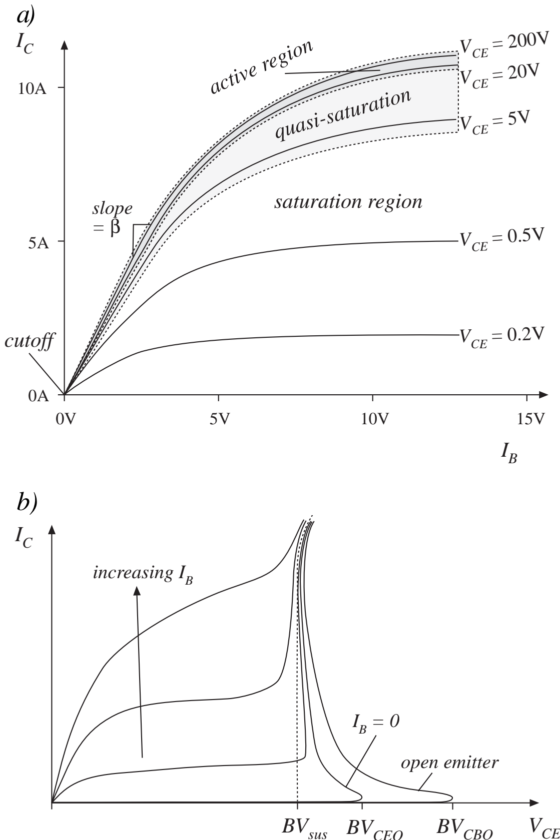

Fig. 4.35. BJT static characteristics: (a) I_C vs. I_B, illustrating the regions of operation; (b) I_C vs. V_{CE}, illustrating voltage breakdown characteristics.

with sufficiently negative base current. BV_{CEO} is the somewhat smaller collector-emitter breakdown voltage observed when the base current is zero; as avalanche breakdown is approached, free carriers are created which have the same effect as a positive base current and which cause the breakdown voltage to be reduced. BV_{sus} is the breakdown voltage observed with positive base current. Because of the high instantaneous power dissipation, breakdown usually results in destruction of the BJT. In most applications, the off-state transistor voltage must not exceed BV_{CEO}.

High-voltage BJT's typically have low current gain, and hence Darlington-connected devices, Fig. 3.36, are common. If transistors Q_1 and Q_2 have current gains β_1 and β_2, respectively, then the Darlington-connected device has the substantially increased current gain $\beta_1 + \beta_2 + \beta_1\beta_2$. In a monolithic Darlington device, transistors Q_1 and Q_2 are integrated on the same silicon wafer. Diode D_1 speeds up the turn-off process, by allowing the base driver to actively remove the stored charge of both Q_1 and Q_2 during the turn-off transition.

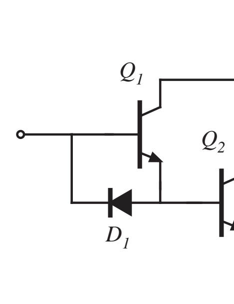

Fig. 4.36. Darlington-connected BJTs, including diode for improvement of turn-off times.

At voltage levels below 500V, the BJT has been almost entirely replaced by the MOSFET in power applications. It has also recently begun to be displaced in higher voltage applications, where new designs utilize faster IGBT's or other devices.

Insulated Gate Bipolar Transistor (IGBT)

A cross-section of the IGBT is illustrated in Fig. 4.37. Comparison with Fig. 4.26 reveals that the IGBT and power MOSFET are very similar in construction. The key difference is the *p* region connected to the collector of the IGBT. So the IGBT is a modern four-layer power semiconductor device having a MOS gate.

The function of the added *p* region is to inject minority charges into the n^- region while the device operates in the on-state, as illustrated in Fig. 4.37. When the IGBT conducts, the p-n^- junction is forward-biased, and the minority charges injected into the n^- region cause conductivity modulation. This reduces the on-resistance of the n^- region, and allows high-

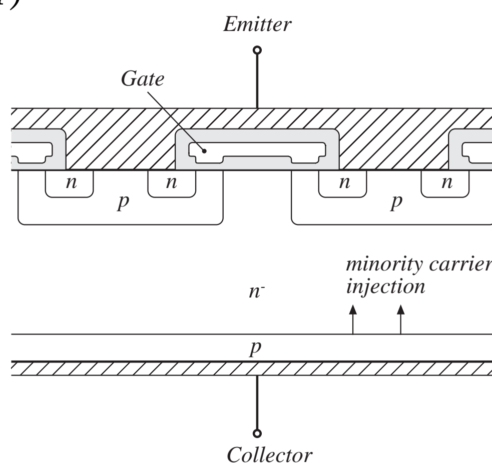

Fig. 4.37. IGBT structure. Crosshatched regions are metallized contacts. Shaded regions are insulating silicon dioxide layers.

voltage IGBTs to be constructed which have low forward voltage drops. IGBTs rated as low as 600V and as high as 1700V are available (in 1995), and even higher voltage ratings are anticipated. The forward voltage drops of these devices are typically 2-4V, much lower than would be obtained in equivalent MOSFETs of the same silicon area.

Several schematic symbols for the IGBT are in current use; the symbol illustrated in Fig. 4.38(a) is the most popular. A two-transistor equivalent circuit for the IGBT is illustrated in Fig. 4.38(b). The IGBT functions effectively as an n-channel power MOSFET, cascaded by a PNP emitter-follower BJT. The physical locations of the two effective devices are illustrated in Fig. 4.39. It can be seen that there are two effective currents: the effective MOSFET channel current i_1, and the effective PNP collector current I_2.

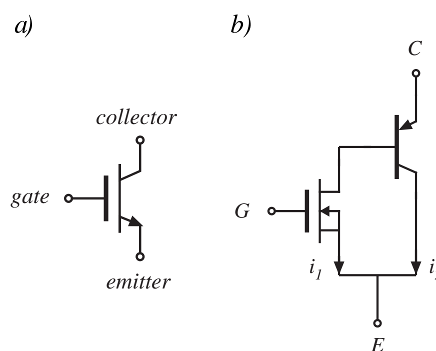

Fig. 4.38. The IGBT: (a) schematic symbol, (b) equivalent circuit.

The price paid for the reduced voltage drop of the IGBT is its increased switching times, especially during the turn-off transition. In particular, the IGBT turn-off transition exhibits a phenomenon known as *current tailing*. The effective MOSFET can be turned off quickly, by removing the gate charge such that the gate-to-emitter voltage is negative. This causes the channel current i_1 to quickly become zero. However, the PNP collector current i_2 continues to flow as long as minority charge is present in the n^- region. Since there is no way to actively remove the stored minority charge, it slowly decays via recombination. So i_2 slowly decays in proportion to the minority charge, and a current tail is observed. The length of the current tail can be reduced by introduction of recombination centers in the n^- region, at the expense of a somewhat increased on-resistance. The current gain of the effective PNP transistor can also be minimized, causing i_1 to be greater than i_2. Nonetheless, the turn-off switching time of the IGBT is significantly longer than that of the MOSFET, with typical turn-off times in the range 0.5µs – 5µs. Switching loss induced by IGBT current tailing is discussed in section 4.3.1. The switching frequencies of PWM converters containing IGBTs are typically in the range 1-30kHz.

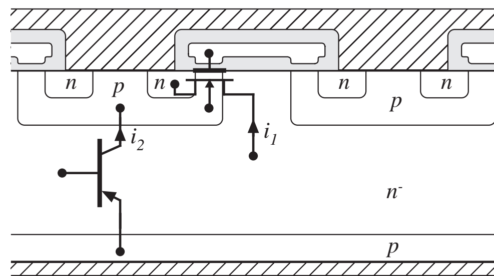

Fig. 4.39. Physical locations of the effective MOSFET and PNP components of the IGBT.

The added p-n^- diode junction of the IGBT is not normally designed to block significant voltage. Hence, the IGBT has negligible reverse voltage-blocking capability.

Since the IGBT is a four-layer device, there is the possibility of SCR-type latchup, in which the IGBT cannot be turned off by gate voltage control. Recent devices are not susceptible to this problem. These devices are quite robust, hot spot and current crowding problems are nonexistent, and the need for external snubber circuits is minimal.

The on-state forward voltage drop of the IGBT can be modeled by a forward-biased diode junction, in series with an effective on-resistance. The temperature coefficient of the IGBT forward voltage drop is complicated by the fact that the diode junction voltage has a negative temperature coefficient, while the on-resistance has a positive temperature coefficient. Fortunately, near rated current the on-resistance dominates, leading to an overall positive temperature coefficient. In consequence, IGBTs can be easily connected in parallel, with a modest current derating. Large modules are commercially available, containing multiple parallel-connected chips.

Characteristics of several commercially-available single-chip IGBTs and multiple-chip IGBT modules are listed in Table 4.3.

Table 4.3. Characteristics of several commercial IGBTs

Part number	*Rated max voltage*	*Rated avg current*	V_F *(typical)*	t_f *(typical)*
Single-chip devices				
HGTG32N60E2	600V	32A	2.4V	0.62µs
HGTG30N120D2	1200V	30A	3.2A	0.58µs
Multiple-chip power modules				
CM400HA-12E	600V	400A	2.7V	0.3µs
CM300HA-24E	1200V	300A	2.7V	0.3µs

Thyristors (SCR, GTO, MCT)

Of all conventional semiconductor power devices, the silicon controlled rectifier (SCR) is the oldest, has the lowest cost per rated kVA, and is capable of controlling the greatest amount of power. Devices having voltage ratings of 5000-7000V and current ratings of several thousand amperes are available. In utility dc transmission line applications, series-connected light-triggered SCRs are employed in inverters and rectifiers which interface the ac utility system to dc transmission lines which carry roughly 1kA and 1MV. A single large SCR fills a silicon wafer that is several inches in diameter, and is mounted in a hockey-puck-style case.

The schematic symbol of the SCR is illustrated in Fig. 4.40(a), and an equivalent circuit containing NPN and PNP BJT devices is illustrated in Fig. 4.40(b). A cross section of the silicon chip is illustrated in Fig. 4.41. Effective transistor Q_1 is composed of the n, p, and n^- regions, while effective transistor Q_2 is composed of the p, n^-, and p regions as illustrated.

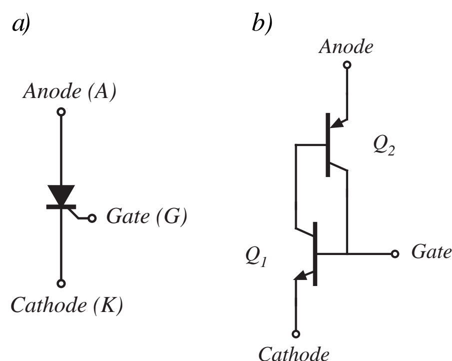

Fig. 4.40. The SCR: (a) schematic symbol, (b) equivalent circuit.

The device is capable of blocking both positive and negative anode-to-cathode voltages. Depending on the polarity of the applied voltage, one of the p-n^- junctions is reverse-biased. In either case, the depletion region extends into the lightly-doped n^- region. As with other devices, the desired voltage breakdown rating is obtained by proper design of the n^- region thickness and doping concentration.

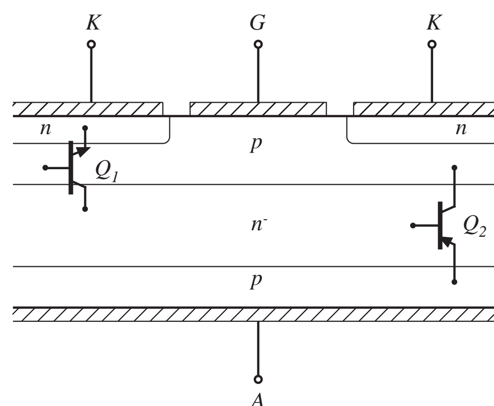

Fig. 4.41. Physical locations of the effective NPN and PNP components of the SCR.

The SCR can enter the on-state when the applied anode-to-cathode voltage v_{AK} is positive. Positive gate current i_G then causes effective transistor Q_1 to turn on; this in turn supplies base current to effective transistor Q_2, and causes it to turn on as well. The effective connections of the base and collector regions of transistors Q_1 and Q_2 constitutes a positive feedback loop. Provided that the product of the current gains of the two transistors is greater than one, then the currents of the transistors will increase regeneratively. In the on-state, the anode current is limited by the external circuit, and both effective transistors operate fully saturated. Minority carriers are injected into all four regions, and the resulting conductivity modulation leads to very low forward voltage drop. In the on-state, the SCR can be modeled as a forward-biased diode junction in series with a low-value on-resistance. Regardless of the gate current, the SCR is latched in the on-state: it cannot be turned off except by application of negative anode current or negative anode-to-cathode voltage. In phase controlled converters, the SCR turns off at the zero crossing of the converter ac input or output waveform. In forced commutation converters, external commutation circuits force the controlled turn-off of the SCR, by reversing either the anode current or the anode-to-cathode voltage.

Static i_A-v_{AK} characteristics of the conventional SCR are illustrated in Fig. 4.42. It can be seen that the SCR is a voltage-bidirectional two-quadrant switch. The turn-on transition is controlled actively via the gate current. The turn-off transition is passive.

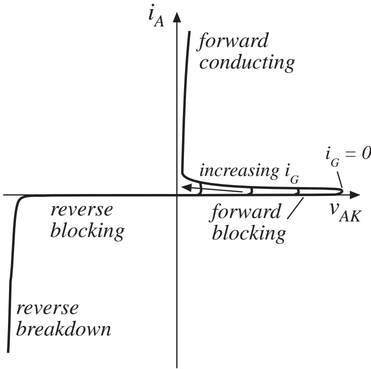

Fig. 4.42. Static i_A-v_{AK} characteristics of the SCR.

During the turn-off transition, the rate at which forward anode-to-cathode voltage is reapplied must be limited, to avoid retriggering the SCR. The turn-off time t_q is the time required for minority stored charge to be actively removed via negative anode current, and for recombination of any remaining minority charge. During the turn-off transition, negative anode current actively removes stored minority charge, with waveforms similar to diode turn-off transition waveforms of Fig. 4.25. Thus, after the first zero crossing of the anode current, it is necessary to wait for time t_q before reapplying positive anode-to-cathode voltage. It is then necessary to limit the rate at which the anode-to-cathode voltage increases, to avoid retriggering the device. *Inverter-grade* SCRs are optimized for faster switching times, and exhibit smaller values of t_q.

Conventional SCR wafers have large feature size, with coarse or nonexistent interdigitation of the gate and cathode contacts. The parasitic elements arising from this large feature size lead to several limitations. During the turn-on transition, the rate of increase of the anode current must be limited to a safe value. Otherwise, cathode current focusing can occur, which leads to formation of hot spots and device failure.

The coarse feature size of the gate and cathode structure is also what prevents the conventional SCR from being turned off by active gate control. One might apply a negative gate current, in an attempt to actively remove all of the minority stored charge and to reverse-bias the *p*-*n* gate-cathode junction. The reason that this attempt fails is illustrated in Fig. 4.43. The large negative gate current flows laterally through the adjoining the *p* region, inducing a voltage drop as shown. This causes the gate-cathode junction voltage to be smaller near the gate contact, and relatively larger away from the gate contact. The negative gate current is able to reverse-bias only the portion of the gate-cathode junction in the vicinity of the gate contact; the

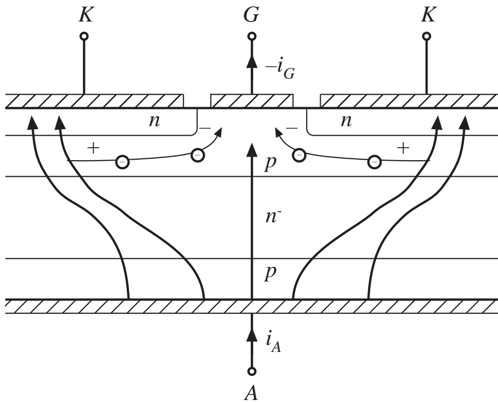

Fig. 4.43. Negative gate current is unable to completely reverse-bias the gate-cathode junction. The anode current focuses away from the gate contact.

remainder of the gate-cathode junction continues to be forward-biased, and cathode current continues to flow. In effect, the gate contact is able to influence only the nearby portions of the cathode.

The *Gate Turn Off Thyristor*, or GTO, is a modern power device having small feature size. The gate and cathode contacts highly interdigitated, such that the entire gate-cathode *p-n* junction can be reverse-biased via negative gate current during the turn-off transition. Like the SCR, a single large GTO can fill an entire silicon wafer. Maximum voltage and current ratings of commercial GTOs are lower than those of SCRs.

The turn-off gain of a GTO is the ratio of on-state current to the negative gate current magnitude required to switch the device off. Typical values of this gain are 2-5, meaning that several hundred amperes of negative gate current may be required to turn off a GTO conducting 1000A. Also of interest is the maximum controllable on-state current. The GTO is able to conduct peak currents significantly greater than the rated average current; however, it may not be possible to switch the device off under gate control while these high peak currents are present.

The *MOS-Controlled Thyristor*, or MCT, is a recent power device in which MOSFETs are integrated onto a highly-interdigitated SCR, to control the turn-on and turn-off processes. Like the MOSFET and IGBT, the MCT is a single-quadrant device whose turn-on and turn-off transitions are controlled by a MOS gate terminal. Commercial MCTs are p-type devices. Voltage-bidirectional two-quadrant MCTs, and n-type MCTs, are also possible.

A cross-section of an MCT containing MOSFETs for control of the turn-on and turn-off transitions is illustrated in Fig. 4.44. An equivalent circuit which explains the operation of this structure is given in Fig. 4.45. To turn the device on, the gate-to-anode voltage is driven negative. This forward-biases p-channel MOSFET Q_3, forward-biasing the base-emitter junction of BJT Q_1. Transistors Q_1 and Q_2 then latch in the on-state. To turn the device off, the gate-to-anode voltage is driven positive. This forward-biases n-channel MOSFET Q_4, which in turn reverse-biases the base-emitter junction of BJT Q_2. The BJTs then turn off. It is important that the on-resistance of the n-channel MOSFET be small enough that sufficient influence on the

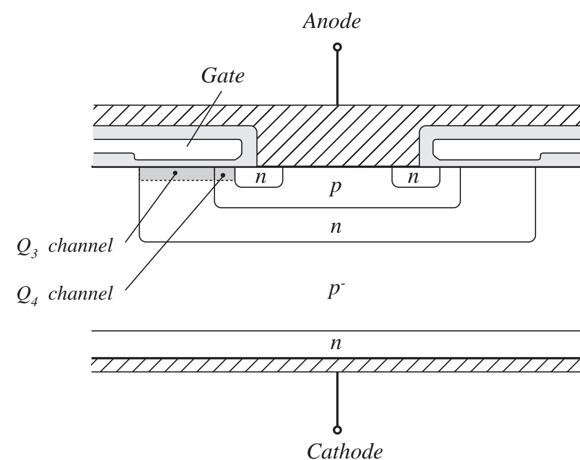

Fig. 4.44. MCT structure. Crosshatched regions are metallized contacts. Lightly shaded regions are insulating silicon dioxide layers.

cathode current is exerted —this limits the maximum controllable on-state current (i.e., the maximum current that can be interrupted via gate control).

High-voltage MCTs exhibit lower forward voltage drops and higher current densities than IGBTs of similar voltage ratings and silicon area. However, the switching times are longer. Like the GTO, the MCT can conduct considerable surge currents; but again, the maximum current which can be interrupted via gate control is limited. To obtain a reliable turn-off transition, external snubbers are required to limit the peak anode-to-cathode voltage. A sufficiently fast gate-voltage rise time is also required. To some extent, the MCT is still an emerging device —future generations of MCTs may exhibit considerable improvements in performance and ratings.

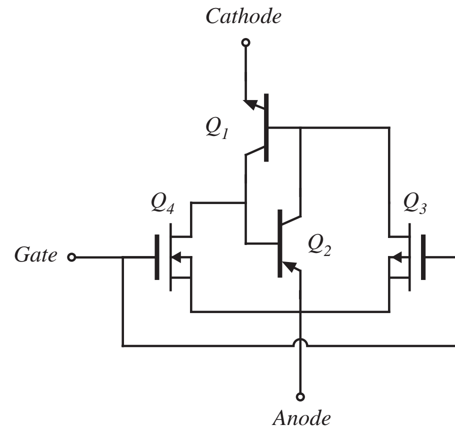

Fig. 4.45. Equivalent circuit for the MCT.