IGBTs (Insulated Gate Bipolar Transistor)

Description

This document describes the basic structures, ratings, and electrical characteristics of IGBTs. It also provides usage considerations for IGBTs.

Table of Contents

List of Figures

List of Tables

1. Device structure and characteristics of IGBTs

An Insulated Gate Bipolar Transistor (IGBT) is a device that combines the MOSFET's advantages of high input impedance and high switching speed[*1] with the bipolar transistor's advantage of high conductivity characteristics (i.e., low saturation voltage).

Like MOSFETs and bipolar transistors, the IGBT is also used as an electronic switch.

*1 The IGBT provides a relatively high switching speed although it is slower than the power MOSFET.

1.1. Basic structure of the IGBT

Figure 1.1 shows the basic structure and an equivalent circuit of an IGBT. The IGBT has a structure similar to that of the MOSFET. Basically, a MOSFET has an n^+–n^- substrate whereas an IGBT has a p^+–n^+–n^- substrate. Therefore, IGBTs and MOSFETs are fabricated using similar processes.

The equivalent circuit of an IGBT indicates that a thyristor[*2] structure is formed by the coupling of PNP and NPN transistors. However, as shown by its structure, the thyristor is designed not to function since the base and the emitter of the NPN transistor is short-circuited by an aluminum wire (via a resistor in the P-base layer). Therefore, an IGBT and its principle of operation can be viewed as equivalent to an inverted Darlington configuration realized by an enhancement N-channel MOSFET as input stage and a PNP transistor as output stage.

Since an IGBT has a monolithic structure consisting of a MOSFET and a PNP transistor, its operation is characterized by the conductivity modulation of the n^- region in addition to the operation represented by the equivalent circuit. Conductivity modulation occurs in the n^- region because of holes (i.e., minority carriers) injected into the n^- region from the p^+–n^+ region. Conductivity modulation causes a decrease in the drain-to-source resistance of the MOSFET. Owing to conductivity modulation, the IGBT has a very low on-state voltage drop (low saturation voltage) that is difficult to achieve with a high-voltage MOSFET.
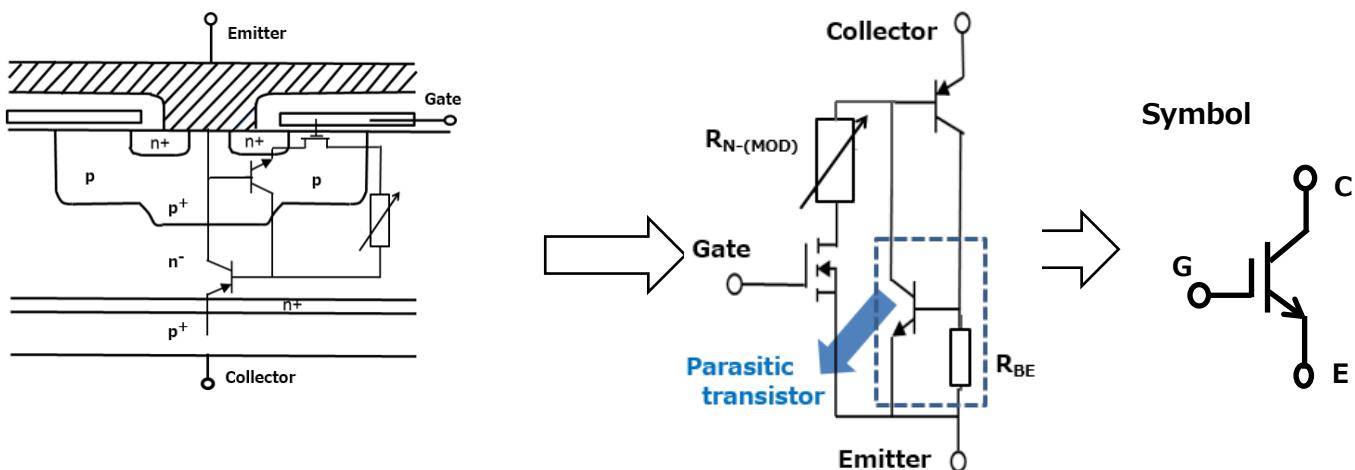
Figure 1.1 Basic structure and equivalent circuit of an IGBT

*2 Thyristor

Typically, a thyristor is a bistable switch (with on and off states) having three PN junctions.

The current that flows between the anode and the cathode (#2 and #3) can be controlled via the current applied to the gate electrode (#1), as shown in Figure 1.2.

The gate current switches on the lower transistor. Then, the upper transistor switches on, drawing a base current from the collector of the lower transistor. This produces a positive feedback in which each transistor causes the other to remain on.

Once the thyristor is on, it cannot be turned off simply by removing the current from the gate. Instead, it must be switched off from its load side (for example, by inverting the voltage polarity).
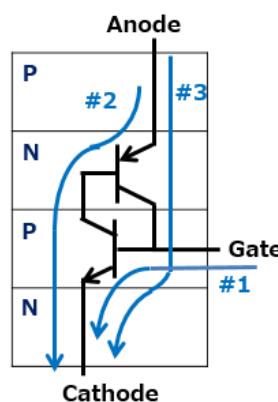
Figure 1.2 Basic structure of a thyristor

From the equivalent circuit, the saturation voltage ($V_{CE(sat)}$) of the IGBT can be expressed as:

$V_{CE(sat)} = V_{BE} + I_{MOS}(R_{N-(MOD)} + R_{ch})$ ----- Equation 1

V_{BE}:   Base-emitter voltage of the PNP transistor

I_{MOS}:   Drain current of the MOSFET

$R_{N-(MOD)}$: Resistance of the n⁻ region after conductivity modulation

R_{ch}:   Channel resistance of the MOSFET

Let the collector current and the DC current gain of the PNP transistor be $I_{C(PNP)}$ and $h_{FE(PNP)}$ respectively. Then, I_{MOS} is calculated as:

$I_{MOS} = I_{C(PNP)} / h_{FE(PNP)}$ ----- Equation 2

The total current of the IGBT (I_{IGBT}) is:

$I_{IGBT} = I_{MOS} + I_{C(PNP)}$

Equation 1 indicates that the saturation voltage ($V_{CE(sat)}$) of an IGBT depends greatly on I_{MOS}, which is a direct function of the h_{FE} of the PNP transistor as shown by Equation 2. Because there is a trade-off between $h_{FE(PNP)}$ and switching characteristics, the h_{FE} of the PNP transistor greatly affects the trade-off between the saturation voltage and the switching characteristics of an IGBT.

1.2. Comparison of different types of transistors (bipolar transistors, MOSFETs, and IGBTs)

Table 1.1 compares the structures and characteristics of IGBTs, bipolar power transistors, and power MOSFETs.

Table 1.1 Comparison of different types of transistors

Good>Fair>Poor

	Bipolar Transistor	IGBT	Power MOSFET
Structure			
Carrier	Electron and Hole	Electron and Hole	Electron
Drive	Base current	Gate voltage	Gate voltage
Current ability	Fair	Good	Poor
On-state voltage	Fair	Good	Poor
Operation frequency	Poor (Low)	Fair (~20kHz)	Good (~300kHz)
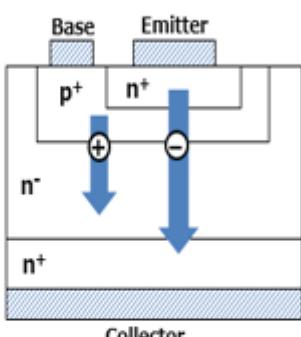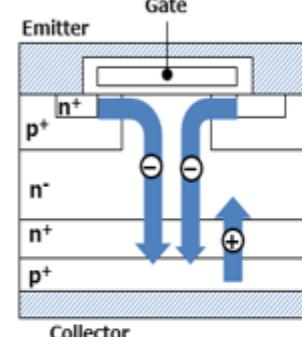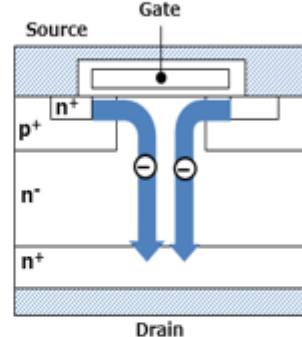
· Bipolar transistors (NPN transistors)

See the device structure of a bipolar transistor shown in Table 1.1. In a typical operation, the base-emitter junction is positively biased while voltage is applied between the collector and the emitter. This causes a current (collector current) to flow from the collector to the emitter.

In an NPN transistor, carriers (electrons) are injected from the emitter (n^+ region) into the base region (p^+ region). A majority of these electrons cross the PN junction into the collector to form the collector current. The remainder of the electrons recombine with holes to form the base current. (The direction of current is opposite to the flow of electrons.) The base region is so thin that a majority of electrons injected from the emitter diffuse across the base to the collector without recombining with holes in the base. This causes the transistor to turn on.

When the transistor is in the on state, carrier density of the n^- region remarkably increases and resistance remarkably decreases in the case that the collector's electrical potential is lower than the base's electrical potential. The dopant concentration in the n^- region changes and resistance varies. This phenomenon is called conductivity modulation.

When a zero or negative voltage is applied across the base-emitter junction, the base current becomes zero, decreasing carrier concentration in the device. As a result, both the base-emitter and collector-emitter junctions are reverse-biased, and the transistor is turned off.

· N-channel MOSFETs

See the device structure of an N-channel MOSFET shown in Table 1.1. In a typical operation, the gate-source junction is positively biased while voltage is applied between the drain and the source. Then, the gate attracts electrons, inducing a conductive channel in the substrate below the oxide film, which electrically connects the source region to the drift region, allowing a current to flow. When a zero or negative voltage is applied to the gate, the channel disappears, turning off the MOSFET.

The drain-source current can be controlled via the voltage applied to the gate electrode. The current drive capability of the gate control circuit can be as low as necessary to charge and discharge the MOSFET gate capacitance. This makes it possible to considerably reduce the size and power loss of a gate control circuit of the MOSFET, compared to that of a bipolar transistor.

Since carrier is not accumulated in the drift region, MOSFETs can switch faster than bipolar transistors. The maximum allowable drain-source voltage (V_{DSS}) of a MOSFET can be increased by increasing the dopant concentration and thickness of the drift region. A downside of this is an increased on-resistance, which cause a substantial increase in conduction loss.

· IGBTs

As described above, an IGBT is constructed similarly to a MOSFET, except that the IGBT has an additional p^+ region on the drain side. Consequently, an IGBT has four alternating layers (p-n-p-n).

IGBT current flows via the internal PNP transistor by turning on the internal MOSFET while voltage is applied between the collector and the emitter.

2018-09-01

This causes conductivity modulation in which holes are injected from the p^+ collector region to the n^- region, reducing the resistance across the n^- region. When a zero or negative voltage is applied to the gate, the internal MOSFET channel disappears, turning off the IGBT. At this time, the PN junction between the p-base and n^- drift regions is reverse-biased, interrupting the injection of holes from the p^+ collector region to the n^- region. At turn-off, current continues flowing until the carriers accumulated by the conductivity modulation exit the drift region or disappear as a result of recombination. (This current is called tail current.)

Like the MOSFET, the collector-emitter current of an IGBT can be controlled via the voltage applied to the gate. Therefore, the current drive capability of the gate control circuit can be as low as necessary to charge and discharge the IGBT gate capacitance. As is the case with the MOSFET, this makes it possible to considerably reduce the size and power loss of the gate control circuit of an IGBT, compared to that of a bipolar transistor.

IGBTs are ideal for high-voltage applications since their on-state voltage can be reduced due to the conductivity modulation effect. However, it takes time for the accumulated carriers to exit the drift region at turn-off. Therefore, MOSFETs, which do not have carrier accumulation, can switch faster than IGBTs.

2. Different types of IGBTs and their structures

IGBTs can be divided into planar-gate and trench-gate IGBTs according to their gate structures, and into punch-through, non-punch-through, and thin-wafer punch-through IGBTs according to their vertical structures.

2.1. Gating structures

Figure 2.1 compares planar-gate and trench-gate IGBTs.

In a planar-gate IGBT, a gate is formed on the chip surface. In contrast, in a trench-gate IGBT, the gate electrode is buried in a trench that runs through the n^+-emitter and p-base regions. The trench-gate structure can greatly increase the cell density and therefore reduce the channel voltage drop, compared to the planar-gate structure. A JFET[*3] is formed between channels in a planar-gate IGBT whereas no JFET exists in a trench-gate IGBT. Consequently, the trench-gate IGBT has no voltage drop due to a JFET, making it possible to reduce on-state voltage considerably, compared to the planar-gate IGBT.

Planar-gate structure   **Trench-gate structure**

R_{ch}: Channel resistance
R_{JFET}: JFET resistance
R_{drift}: Resistance across the drift region
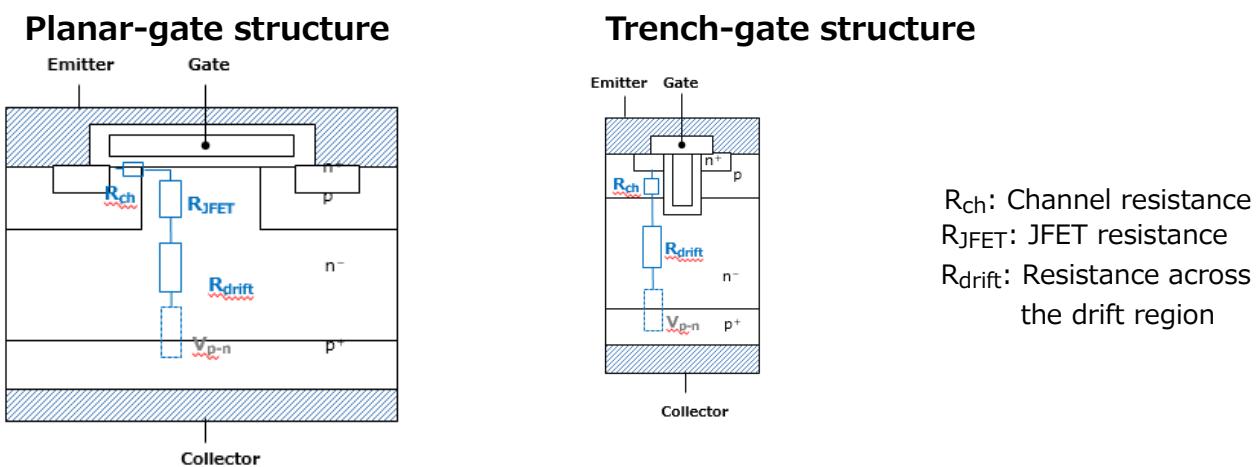
Figure 2.1 Planar-gate structure vs. trench-gate structure

*3 JFET (junction field-effect transistor)

When a reverse bias voltage is applied across the gate and source (drain) terminals, the width of the depletion layer increases, impeding the gate-source current and switching off a JFET.
The JFET-like structure in an IGBT does not act as a JFET.
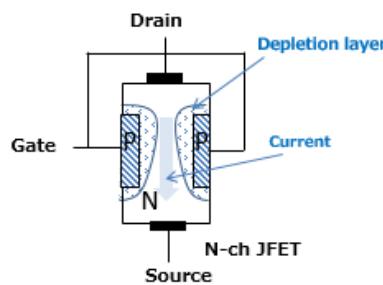
Figure 2.2 N-channel junction field-effect transistor

2.2. IGBT classification based on the vertical structure

IGBTs can be divided into punch-through (PT) and non-punch-through (NPT) IGBTs based on the depletion of the drift region that occurs as a result of voltage application across the collector and the emitter. A PT IGBT has a depletion layer that extends across the drift region into an n^+ low-resistance layer whereas the depletion layer in an NPT IGBT does not punch through the drift region and thus eliminates the need for an n^+ low-resistance layer.

The punch-through structure fabricated using a thin-wafer process is called a thin-wafer punch-through structure or a field-stop punch-through structure.

Figure 2.3 compares the PT, NPT, and thin-wafer PT structures.

Reverse-conducting (RC) IGBTs are fabricated using the thin-wafer PT technology to form an n^+ region through the p^+ region.

2.2.1. Punch-through (PT) IGBTs

Generally, the punch-through (PT) structure is comprised of n^- (drift), n^+ (buffer), and p^+ (anode) regions. The n^- region is optimized to meet the withstand voltage requirement.

PT IGBTs, which are fabricated using epitaxial wafers, have a thick p^+ collector region with high dopant concentration. In the conducting state, large amounts of carriers are injected from the collector to achieve conductivity modulation and thereby reduce the on-state voltage. A downside of the PT IGBT is that, at turn-off, current (called tail current) continues flowing until carriers exit the n^- drift region or recombine, increasing switching loss. To reduce switching loss, crystal defects are formed in the n^- drift region in order to make the accumulated carriers exit the n^- drift region faster. This technique is called lifetime control.

Generally, PT IGBTs have a high collector-emitter threshold voltage in the conducting state. Further reduction in the $V_{CE(sat)}$ of PT IGBTs has been a challenge (or obstacle). (See Figure 2.4)

2.2.2. Non-punch-through (NPT) IGBTs

The non-punch-through (NPT) structure is designed with a thick n^- drift region so that the depletion layer remains within the n^- drift region even in the maximum electric field. Therefore, NPT IGBTs do not have an n^+ buffer region that is required in PT IGBTs to reduce the expansion of the depletion layer. (The n^+ buffer region in the PT IGBT is also used to control the amount of holes injected from the p^+ anode.)

In addition, the NPT IGBT has a thin p^+ region. And by varying dopant concentration of this p^+ region, the amount of carrier injection is controlled. This eliminates the need for lifetime control, which is required for PT IGBTs to reduce the turn-off time. Although the NPT IGBT generally has lower switching loss than the PT IGBT, the thicker n^- drift region of the NPT IGBT causes an increase in on-state voltage, particularly in the high collector current region.

On the other hand, the NPT IGBT has much lower $V_{CE(sat)}$ threshold than the PT IGBT. Therefore, the NPT IGBT generally has lower on-state voltage in the low to rated current region. (See Figure 2.4)

2.2.3. Thin-wafer PT (thin-PT) IGBTs

Thin-PT IGBTs combine the advantages of both PT and NPT processes. Specifically, thin-PT IGBTs provide the low operating resistance characteristic of the PT process because of the optimized drift region as well as the high-speed switching and forward threshold characteristics of the NPT process.

The PT IGBT is fabricated using an epitaxial wafer to use a high-concentration p region as a collector region whereas the NPT process forms a collector region at the backside of the device via ion implantation. As a result, NPT IGBTs can accurately control the amount of holes injected from the collector without the need for lifetime control. However, since NPT IGBTs do not have an n^+ buffer region with high dopant concentration that alters the electric field distribution, they must be designed in such a manner as to prevent the depletion layer from punching through the collector region. Consequently, NPT IGBTs need a thicker drift region than PT IGBTs. To avoid this problem, by adopting the thin wafer technology and optimum design of the n^+ buffer region, the thin-PT IGBT was realized low $V_{CE(sat)}$ and high switching performance.
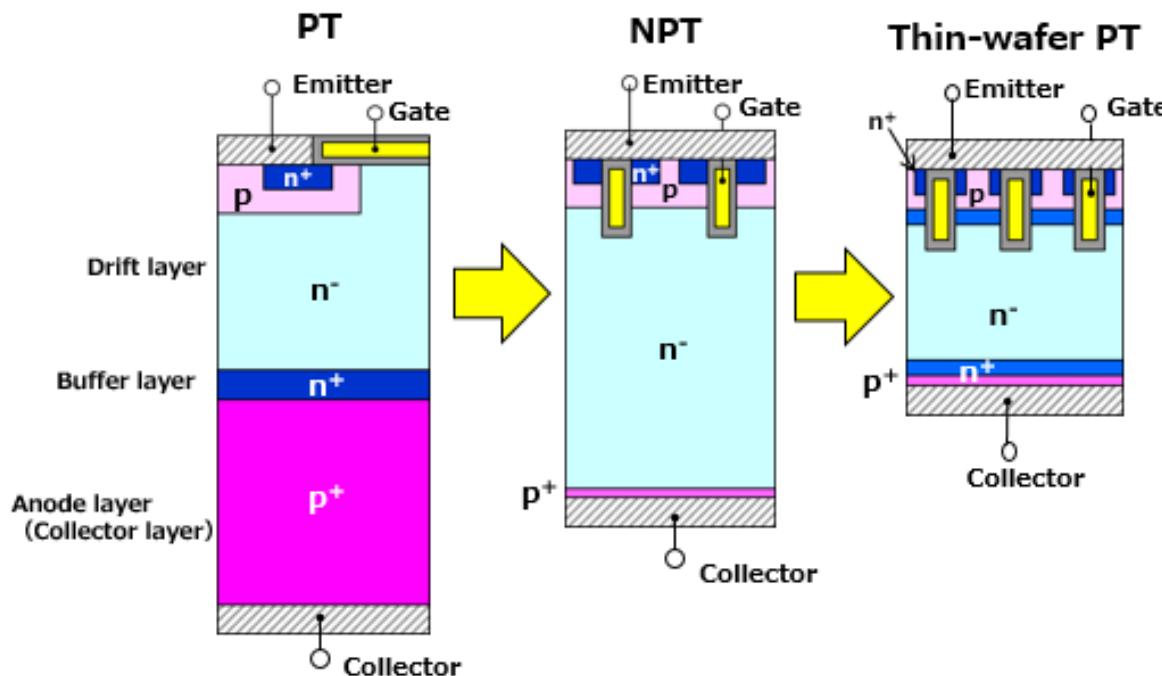
Figure 2.3 IGBTs with different vertical structures
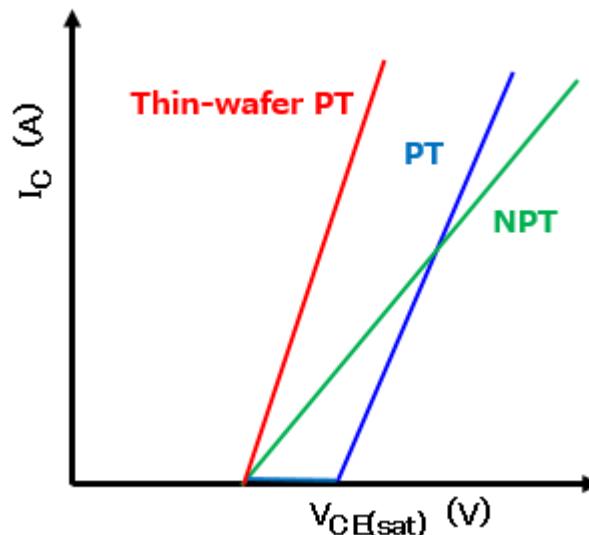
**Figure 2.4 I_C-$V_{CE(sat)}$ Characteristic Image**

2.2.4. Reverse-conducting (RC) IGBTs

RC IGBTs have an n^+ region that runs through the p^+ collector region to form a diode. Although other types of IGBTs are often connected in anti-parallel with a freewheeling diode (FWD), the RC IGBT integrates both an IGBT and an FWD in a single chip.
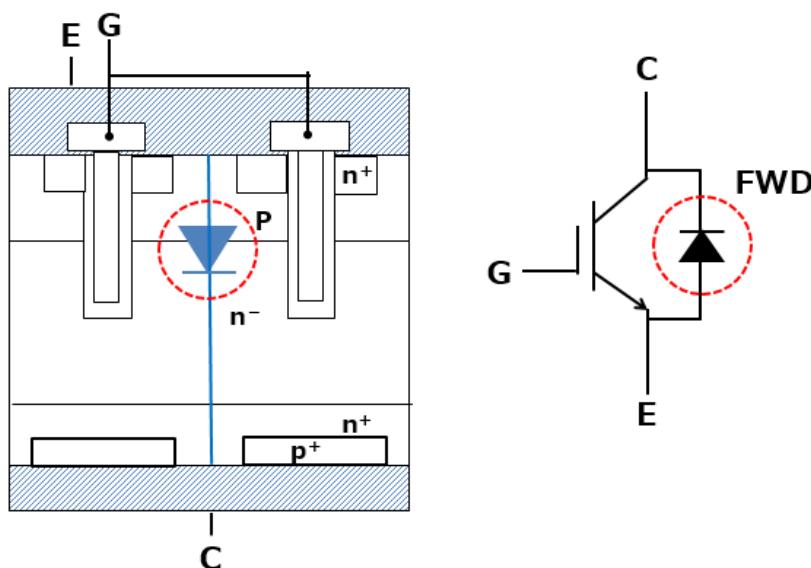

Figure 2.5 Structure of the RC IGBT

3. Ratings of discrete IGBTs

3.1. Maximum ratings of IGBTs

For IGBTs, the maximum allowable current, voltage, power dissipation, and other parameters are specified as maximum ratings.

In creating a semiconductor circuit design, understanding maximum ratings is crucial to ensure that semiconductor devices operate within the target period of time and with sufficient reliability.

One of the characteristics of semiconductor devices is that their electrical characteristics are very sensitive to temperature. Therefore, the maximum ratings are determined, considering the temperature rise of a device. When a voltage applied to an IGBT is constant, its electrical conductivity and thereby leakage current increase as the ambient temperature increases. The increased leakage current causes a further temperature rise. This cycle enters into a positive feedback loop, eventually destroying the device in the worst case.

The maximum ratings of IGBTs must not be exceeded to ensure the expected useful life and reliability. The maximum ratings are limited by the materials, circuit designs, and manufacturing conditions used and therefore differ from device to device.

The absolute maximum ratings are the highest values that must not be exceeded during operation even instantaneously.

In addition, it is difficult to apply two or more absolute maximum ratings simultaneously. In other words, the rated value of each item is specified under conditions without the influence of other items. For example, when voltage and current within the maximum ratings are applied simultaneously, a problem may arise.

Exposure to a condition exceeding a maximum rating sometimes causes permanent degradation of the electrical characteristics. Care should be exercised as to supply voltage bounces, variations in the characteristics of circuit components, possible exposure to stress higher than the maximum ratings during circuit adjustment, changes in ambient temperature, and input signal fluctuations.

The maximum ratings of IGBTs are specified with respect to base, collector, and emitter currents, terminal-to-terminal voltages, collector power dissipation, junction temperature, storage temperature and so on. These parameters are interrelated and cannot be considered separately. In some cases, they also depend on external circuit conditions.

3.2. Absolute maximum ratings

All parameters are specified at 25°C ambient unless otherwise noted.

Different IGBTs have slightly different types of ratings. For details, see the technical datasheets for individual IGBTs.

Table 3.1 Absolute maximum ratings

Characteristic	Symbol	Definition
		(For ratings, see the relevant technical datasheets.)
Collector-emitter voltage	V_{CES}	The maximum allowable voltage between the collector and the emitter when the gate and the emitter are zero-biased (i.e., shorted)
Gate-emitter voltage	V_{GES}	The maximum allowable voltage between the gate and the emitter when the collector and the emitter are zero-biased (i.e., shorted)
Collector current (DC)	I_C	Maximum allowable DC collector current
Collector current (pulsed)	I_{CP}	Maximum allowable pulsed collector current
Diode forward current[*4]	I_F	The maximum DC current that can flow through the freewheeling diode in the forward direction
Diode forward current (pulse/100 μs)[*4]	I_{FP}	The maximum pulsed current that can flow through the freewheeling diode in the forward direction
Short-circuit withstand time[*5]	t_{sc}	The maximum allowable period of time during which the device can be short-circuited under the specified conditions and after which it normally returns to the off state
Collector power dissipation (T_C =25°C)	P_C	The maximum allowable power dissipation
Junction temperature	T_j	The maximum junction temperature at which an IGBT can operate
Storage temperature	T_{stg}	The maximum temperature at which an IGBT may be stored without current or voltage application
Tightening torque	TOR	The maximum torque that can be applied to screws when mounting an IGBT on a cooling surface with the specified screws

*4: These ratings apply to IGBTs with a freewheeling diode (FWD).

*5: This rating applies to the IGBTs that are commonly used for applications in which the IGBTs may be used in short-circuit mode.

3.3. Thermal resistance characteristics

Table 3.2 Thermal resistance characteristics

Characteristic	Symbol	Definition
		(For ratings, see the relevant technical datasheets.)
Junction-to-case thermal resistance[*6]	$R_{th(j-c)}$	Thermal resistance from the device junction to the package case

*6:   This parameter is provided for IGBTs with a built-in diode.

  Thermal resistance is a value that indicates the resistance of an object or material to a heat flow. It is expressed as a temperature rise caused by power dissipation per unit of time. The symbols commonly used for thermal resistance are R_{th} and θ, and its unit of measure is °C/W.

  Figure 3.1 shows a junction-to-case transient thermal impedance $(r_{th(j-c)})$[*7] curve. The value, which is fixed regardless of time, is thermal resistance R_{th}. (The sign in parentheses expresses the place that defines the value. "(j-c)" means "between a junction and a case.")

A rise in junction temperature caused by a single pulse is calculated as:

$\Delta T_j = r_{th(j-c)(t)} \times P_C$

where, ΔT_j is a rise in junction temperature, P_C is the power applied, and $r_{th(j-c)(t)}$ is the thermal resistance at time t (period during which power is applied).
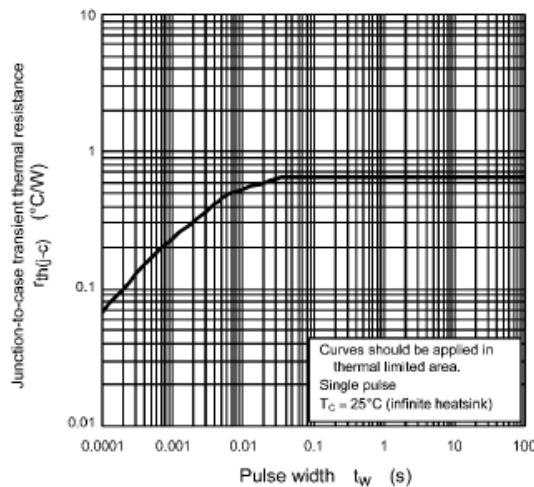
Figure 3.1 Transient thermal impedance curve

*7: transient thermal impedance $(r_{th(j-c)})$

 $r_{th(j-c)}$ means a temperature rise by power dissipation per (in the time of a rectangular wave pulse) unit of time. The symbols commonly used for thermal resistance are R_{th} and θ, and its unit of measure is °C/W.

  It is also called the transient thermal impedance Z_{th}.

3.4. Electrical characteristics

3.4.1. Static characteristics

All parameters are specified at 25°C ambient unless otherwise noted.

Test conditions differ for different IGBTs. For details, see the technical datasheets for individual IGBTs.

Table 3.3 Static characteristics

Characteristic	Symbol	Definition
		(For ratings, see the relevant technical datasheets.)
Gate leakage current	I_{GES}	The gate-emitter current that flows when the rated voltage is applied across the gate and the emitter with the collector and emitter electrodes shorted
Collector cut-off current	I_{CES}	The collector-emitter current that flows when the rated voltage is applied across the collector and the emitter with the gate and emitter electrodes shorted
Gate-emitter cut-off voltage	$V_{GE(OFF)}$	The gate-emitter voltage at which the rated collector current flows when the rated collector-emitter voltage is applied
Collector-emitter saturation voltage	$V_{CE(sat)}$	The collector-emitter voltage when the rated gate-emitter voltage is applied and the collector current is at the rated value
Diode forward voltage[8]	V_F	The emitter-collector voltage when the rated forward current is applied to the freewheeling diode of an IGBT

*8: This parameter applies to IGBTs with a built-in freewheeling diode (FWD).

3.4.2. Dynamic characteristics

All parameters are specified at 25°C ambient unless otherwise noted.

Test conditions differ for different IGBTs. For details, see the technical datasheets for individual IGBTs.

Table 3.4 Dynamic characteristics

Characteristic		Symbol	Definition (For ratings, see the relevant technical datasheets.)
Input capacitance		C_{ies}	The capacitance between the gate and the emitter at the rated gate-emitter voltage, collector-emitter voltage, and frequency
Switching time[*9]	Turn-on delay time	$t_{d(on)}$	The time it takes for the collector current to reach 10% of the rated value at turn-on from the time at which the gate voltage is 10% of the rated value (inductive load)
	Rise time	t_r	The time it takes for the collector current to rise from 10% to 90% of its rated value at turn-on
	Turn-on time	t_{on}	The time it takes for the collector current to reach either 90% of its rated value (in the case of a resistive load) or 10% of the drain-source voltage (in the case of an inductive load) at turn-on from the time at which the gate voltage is 10% of the rated value
	Turn-off delay time	$t_{d(off)}$	The time it takes for the drain-source voltage to reach 10% at turn-off from the time at which the gate voltage is 90% of the maximum rated value
	Fall time	t_f	The time it takes for the collector current to fall from 90% to 10% of its rated value at turn-off
	Turn-off time	t_{off}	The time it takes for the collector current to reach 10% of its rated value at turn-off from the time at which the gate voltage is 90% of the maximum rated value
Switching loss (Turn-on loss)		E_{on}	The amount of energy lost during turn-on until the collector-emitter voltage reaches the rated value
Switching loss (Turn-off loss)		E_{off}	The amount of energy lost during turn-off until the collector-emitter voltage reaches the rated value
Reverse recovery time[*10]		t_{rr}	The period of time during which reverse recovery current flows through the freewheeling diode under the rated conditions that is connected in anti-parallel between the collector and the emitter

*9: This parameter is specified under either resistive- or inductive-load conditions, depending on the intended use of each IGBT. See the datasheets for individual IGBTs as the specified parameters and their conditions differ between resistive and inductive loads.

*10: This parameter applies to IGBTs with a built-in freewheeling diode (FWD).

(1) Capacitance characteristics

In an IGBT, the gate is insulated by a thin silicon oxide. Therefore, an IGBT has capacitances between the collector, gate, and emitter terminals as shown in Figure 3.2.

The gate-collector capacitance (C_{gc}) and the gate-emitter capacitance (C_{ge}) are mainly determined by the structure of the gate electrode whereas the collector-emitter capacitance (C_{ce}) is determined by the capacitance of the vertical PN junction.

The input capacitance ($C_{ies}=C_{gc}+C_{ge}$) is an important parameter for switching characteristics and for the design of a gate drive circuit. The gate drive circuit must be capable of charging and discharging the input capacitance of an IGBT, and the charge and discharge times greatly affect its switching characteristics.

Input capacitance, $C_{ies} = C_{gc} + C_{ge}$
Output capacitance, $C_{oes} = C_{ce} + C_{gd}$
Reverse transfer capacitance, $C_{rss} = C_{gc}$
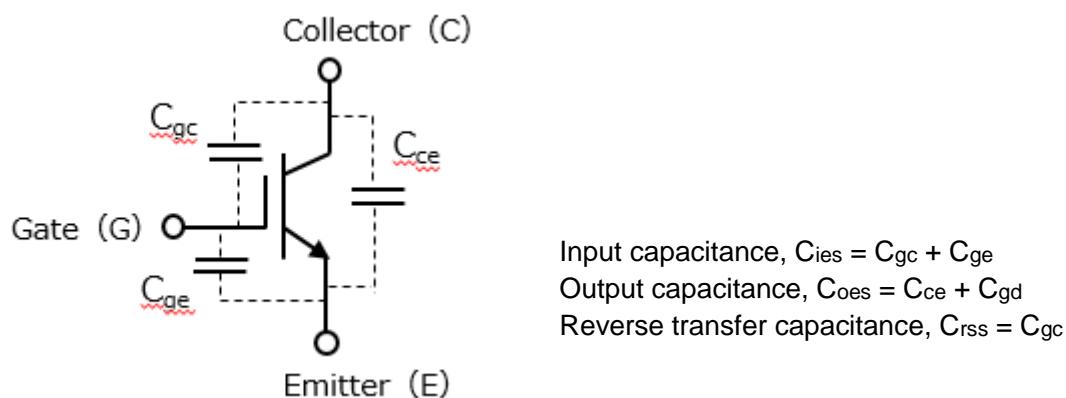
Figure 3.2 IGBT capacitances

(2) Switching time Resistive loads

Figure 3.3 shows the switching time test circuit and switching waveforms of an IGBT under resistive-load conditions.

Switching waveforms under resistive-load conditions

Switching time test circuit for resistive loads

Figure 3.3 Switching time test circuit and switching waveforms under resistive-load conditions

2018-09-01

Inductive loads

Figure 3.4 shows the switching time test circuit and switching waveforms of an IGBT under inductive-load conditions. Since the freewheeling diode in the test circuit is connected in parallel with an inductive load, the I_C waveform during turn-on includes the diode's reverse recovery current.

Switching time test circuit for inductive loads   Switching waveforms under inductive-load conditions
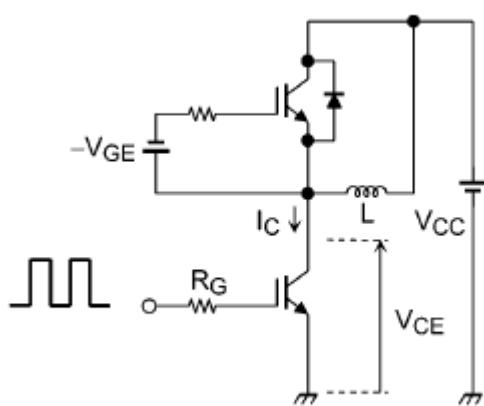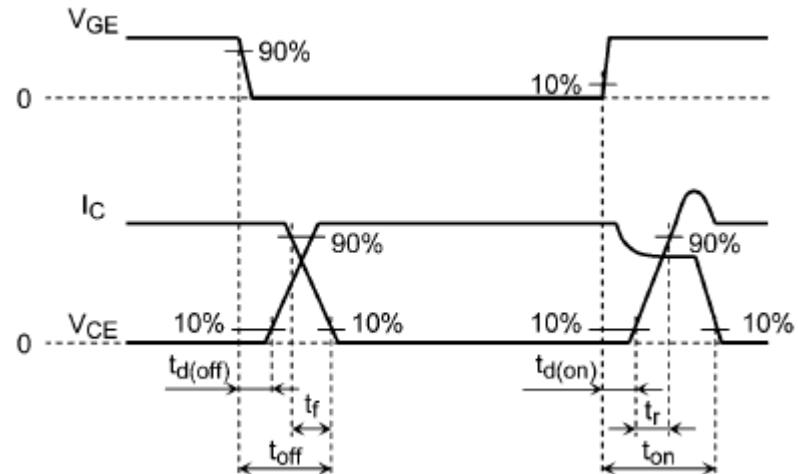

Figure 3.4 Switching time test circuit and switching waveforms under inductive-load conditions

(3) Switching loss

Switching loss is specified under inductive-load conditions for the IGBTs that are suitable for applications in which switching loss is important such as motor drivers. Figure 3.5 shows the periods of time during which switching loss is measured.
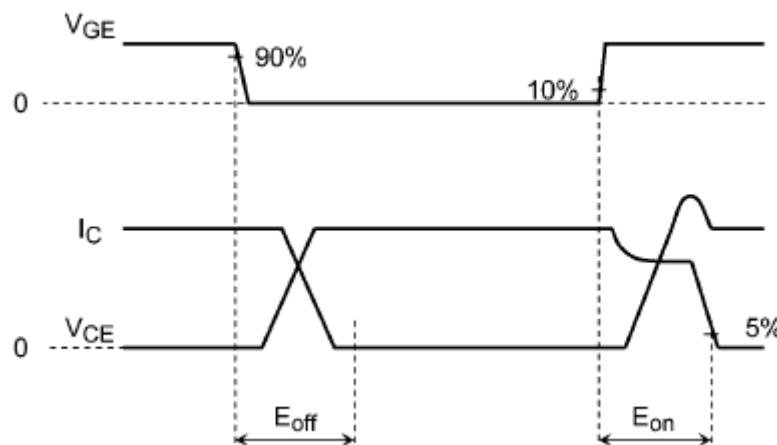
Figure 3.5 Measurement of switching loss under inductive-load conditions

4. Performance characteristics curves

4.1. Static characteristics

Figure 4.1 shows the relationships between the collector current (I_C) and the collector-emitter voltage (V_{CE}) at different gate-emitter voltages (V_{GE}). In other words, these curves indicate the collector current capability at a given gate-emitter voltage. While the collector current is flowing, the power loss is determined by the collector current and the collector-emitter voltage. The I_C-V_{CE} curves are therefore very important in creating a circuit design. It should be noted that the I_C-V_{CE} curves vary with temperature.

Increasing the gate-emitter (V_{GE}) voltage reduces $V_{CE(sat)}$. When IGBTs are used for switching applications, it is desirable to set V_{GE} at around 15 V in order to minimize $V_{CE(sat)}$.

As shown in Figure 4.1, the collector current increases sharply at a V_{CE} of 0.7 V or so. This occurs because the base-emitter voltage (V_{BE}) of the PNP transistor between the collector and the emitter of an IGBT's equivalent circuit (i.e., forward voltage across the PN junction) is added to V_{CE}.
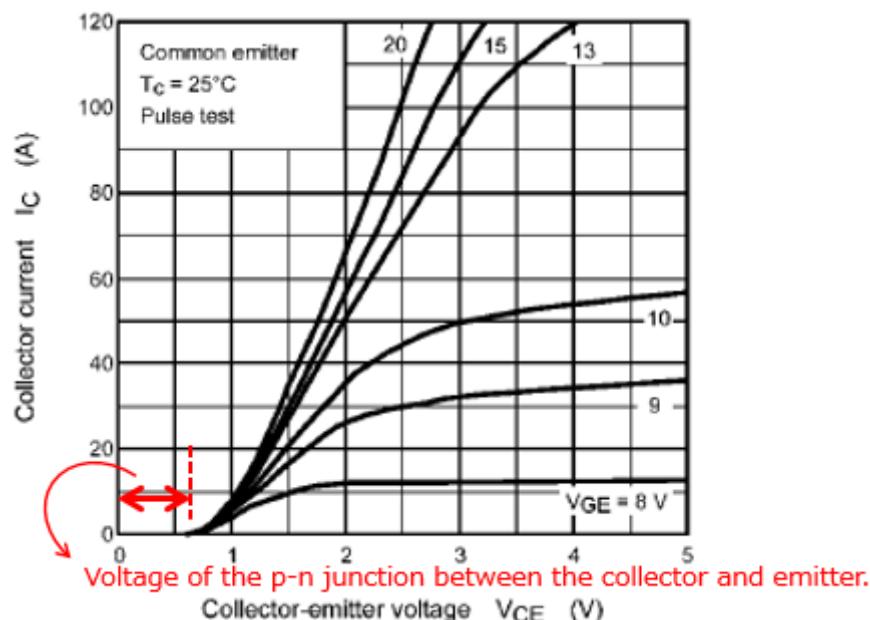
Figure 4.1 I_C-V_{CE} curves of an IGBT

4.2. Safe operating area (SOA)

4.2.1. Concept

The safe operating area (SOA) is defined as the voltage and current conditions over which an IGBT can be expected to operate without self-damage or degradation and often shows multiple curves for different pulse widths. Separate SOA graphs may be provided for different temperature conditions.

The SOA of an IGBT is limited by the absolute maximum ratings such as the maximum voltage, maximum current, and maximum collector power dissipation. However, IGBTs for switching applications could be degraded or damaged even when they are used within individual maximum ratings. Therefore, IGBT applications must be designed in such a manner that the operating points of all IGBTs fall within their SOA curves.

There are two safe operating areas: forward-bias safe operating area (FBSOA) and reverse-bias safe operating area (RBSOA). Since different regions of SOAs are limited by different designs for IGBTs, it

is necessary to refer to the datasheets for the IGBTs used.

4.2.2. Forward-bias safe operating area (FBSOA)

The forward-bias safe operating area of an IGBT is limited by its absolute maximum ratings such as the maximum voltage, maximum current, and maximum collector power dissipation as well as the secondary breakdown.

IGBT applications should be designed, taking the FBSOA into account.

The FBSOA of an IGBT can be divided into the following four regions:

1. Region limited by the maximum collector current rating
2. Region limited by collector power dissipation (thermal breakdown)
3. Region limited by secondary breakdown
4. Region limited by the maximum collector-emitter voltage rating
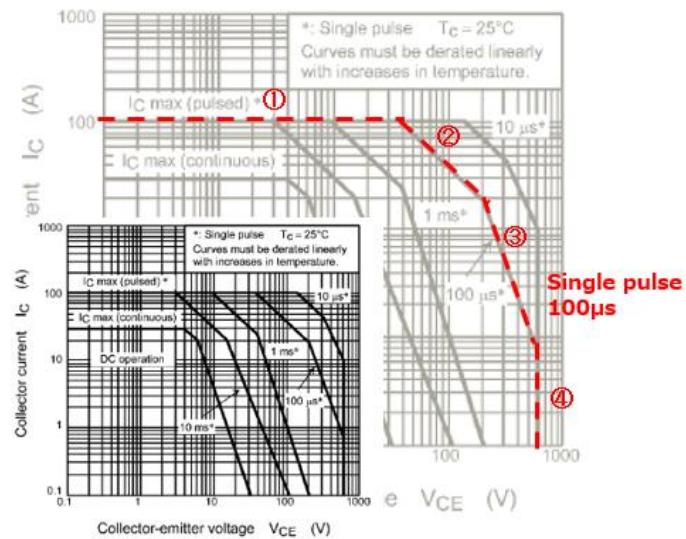
Figure 4.2 Safe operating area

4.2.3. Secondary breakdown (S/B)

Secondary breakdown (S/B) of an IGBT occurs because $V_{GE(OFF)}$ has negative thermal dependence. As patterning becomes finer with the improvement of IGBT characteristics, the S/B phenomenon becomes more pronounced.

S/B originally referred to a phenomenon in bipolar transistors, namely, the occurrence of negative resistance (current concentration) during operation at high voltage or large current. Current concentration causes local heating, resulting in a small hot spot. Furthermore, current increases and concentrates at the hotspots.

This cycle is called thermal runaway. It degrades or causes breakdown of the device. In this respect, S/B of an IGBT in an SOA is the same as that of a bipolar transistor. But S/B of the IGBT is not caused by internal parasitic bipolar transistor operation. Although the term "S/B" may be inappropriate for the IGBT, it is applied to the IGBT, having been inherited from the bipolar transistor.

4.2.4. The mechanism of secondary breakdown (S/B)

The S/B mechanism of an IGBT is considered in the following:

Electric charges are induced between the insulated gate oxide and the interface when the gate is biased positively. It makes an inverted channel.

Once the inverted channel has been made, a current starts flowing through the collector to the emitter of the IGBT. The gate voltage at this point is $V_{GE(OFF)}$. The IGBT controls the collector current by the change of the electric charge induced between the insulated gate oxide and the interface.

The higher the temperature rises, the lower the $V_{GE(OFF)}$ becomes. This is because electric charge is generated easily at high temperature.

Channel temperature is inversely proportional to the difference between gate voltage V_{GE} and $V_{GE(OFF)}$ (i.e., V_{GE}-$V_{GE(OFF)}$). The larger this difference, the more electric charge is generated, and the higher the density becomes. Hence, channel resistance decreases.

The higher temperature leads to the generation of more electric charge, causing $V_{GE(OFF)}$ to become lower and V_{GE}-$V_{GE(OFF)}$ to become higher. Consequently, channel resistance decreases.

Based on the above explanation, the S/B mechanism is described in the following:

(1) Increase of temperature of the IGBT decreases gate-emitter cut-off voltage $V_{GE(OFF)}$, and channel resistance decreases.

(2) Current concentrates in the channel with reduced resistance, causing a further temperature rise, which results in a further decrease in the gate threshold voltage $V_{GE(OFF)}$.

(3) The current becomes increasingly concentrated through positive feedback. This leads to destruction of the device.

Based on the above explanation, the S/B line on an SOA is determined by changing the channel resistance as the temperature increases. Slopes of lines are different between the thermal restricted region (region #2 in Figure 4.2) and the S/B region (region #3 in Figure 4.2).

4.2.5. Reverse-bias safe operating area (RBSOA)

The reverse-bias safe operating area (RBSOA) shows the current and voltage conditions over which an IGBT can be turned off without self-damage. A transition from the on state to the off state gives rise to a voltage surge due to a circuit's stray inductance. IGBT applications must be designed in such a manner that the operating loci of the IGBT's turn-off cut-off current and the resulting surge voltage must fall within the RBSOA (for example, by reducing circuit stray inductance, adding a surge suppression circuit, and reducing the turn-off speed).

Figure 4.3 shows the reverse-bias safe operating area (RBSOA) of an IGBT, which is bound by the following limits:

1. Region limited by the maximum collector current rating, I_{CP} (1 ms)

2. Region limited by the characteristics intrinsic to a device, which are determined by the device design and other factors

3. Region limited by the maximum collector-emitter voltage rating (V_{CES})
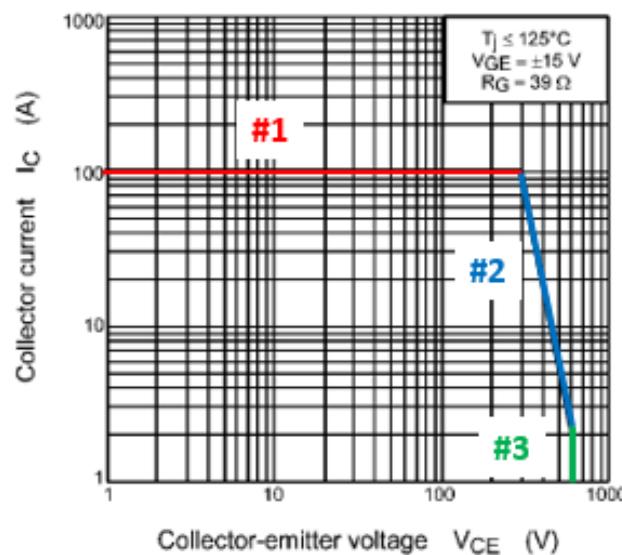
Figure 4.3 Reverse-bias safe operating area
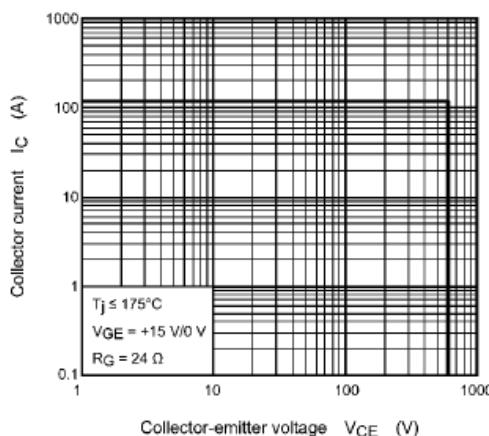
Figure 4.4 Reverse-bias safe operating area

For products for motor driving applications, wider RBSOA is applied. (Figure 4.4)

5. Applications
5.1. Thermal design
5.1.1. Heat dissipation in a steady state

The permissible power dissipation of an IGBT is determined by the ambient temperature (T_a), the IGBT's maximum junction temperature (T_{jmax}), and heat dissipation conditions.

A thermal energy flow can be modeled by analogy to an electrical circuit. Using this model, the thermal energy flow from the junction of an IGBT to the ambient air is derived from thermal resistances (see Appendix A) and thermal capacitances. Figure 5.1 shows an equivalent thermal circuit in a thermally steady state.

Figure 5.1 Equivalent thermal circuit
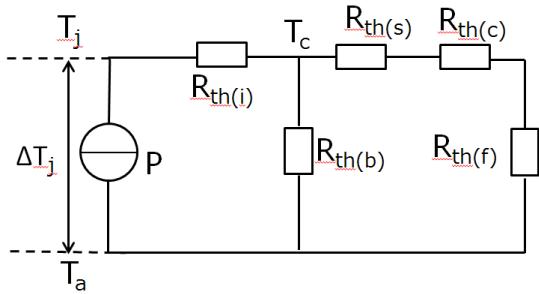
$R_{th(i)}$: Internal thermal resistance ($R_{th(j-c)}$: Junction-to-case thermal resistance)

$R_{th(b)}$: External (case-to-ambient) thermal resistance

$R_{th(s)}$: Thermal resistance of an insulation plate

$R_{th(c)}$: Contact thermal resistance (between case and thermal fin)

$R_{th(f)}$: Thermal resistance from a thermal fin to ambient air

From the equivalent circuit of Figure 5.1, the junction-to-ambient thermal resistance $R_{th(j-a)}$ can be calculated as follows:

$$R_{th(j-a)} = R_{th(i)} + \frac{R_{th(b)}\left(R_{th(s)} + R_{th(c)} + R_{th(f)}\right)}{R_{th(b)} + R_{th(s)} + R_{th(c)} + R_{th(f)}} \qquad \text{------------} \qquad \text{Equation 1}$$

If no heat sink is used, $R_{th(j-a)}$ is simplified to:

$$R_{th(j-a)} = R_{th(i)} + R_{th(b)} \qquad \text{--} \qquad \text{Equation 2}$$

In Figure 5.1, the $R_{th(b)}$ of the IGBT is considerably larger than $R_{th(i)}$, $R_{th(c)}$, $R_{th(s)}$, and $R_{th(f)}$. Therefore, Equation 1 can be simplified to:

$$R_{th(j-a)} = R_{th(i)} + R_{th(S)} + R_{th(C)} + R_{th(f)} \qquad \text{-----------------------} \qquad \text{Equation 3}$$

Hence, when thermal energy is produced at the junction due to a power dissipation of P_C watts, the resulting junction temperature (T_j) is expressed as follows:

$$T_j = P_c\left(R_{th(i)} + R_{th(S)} + R_{th(C)} + R_{th(f)}\right) + T_a \qquad \text{------------------} \qquad \text{Equation 4}$$

When $T_a = 25°C$, the collector power dissipation (P_c) of the IGBT's equivalent circuit shown in Figure 5.1 is:

$$P_{Cmax(Ta=25°C)} = \frac{T_{jmax} - T_{a(Ta=25°C)}}{R_{th(j-a)}} \qquad \text{----------------------------} \qquad \text{Equation 5}$$

The maximum collector power dissipation rating of an IGBT is specified at $T_c = 25°C$ ($T_c = T_a$, infinite heat sink), and its thermal resistance is specified as junction-to-case thermal resistance, $R_{th(j-c)}$[13].

$$P_{Cmax(Tc=25°C)} = \frac{T_{jmax} - T_{c(Tc=25°C)}}{R_{th(j-c)}} \quad \text{-----------------------------} \quad \text{Equation 6}$$

[13] $R_{th(j-c)}$ is shown as $R_{th(i)}$ in Figure 5.1.

5.1.2. Heat dissipation in a transient state

Generally, the thermal impedance of a transistor is modeled as a distributed constant circuit as shown in Figure 5.2.
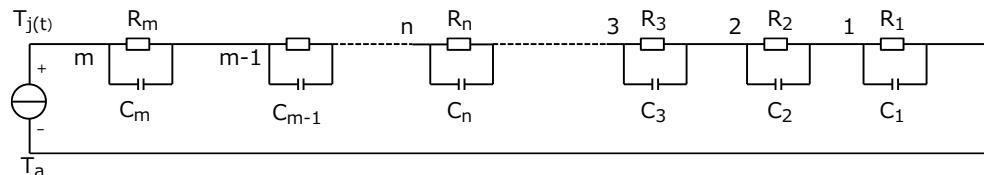
Figure 5.2 Transient thermal impedance model

When the pulse dissipation $P_{j(t)}$ shown in Figure 5.3 is applied to the circuit of Figure 5.2, a change in junction temperature $T_{j(t)}$ that appears at the mth parallel RC circuit under stable thermal conditions can be calculated as follows:

1) In the region where $P_{j(t)} = P_0$:

$$T_{j(t)} = \sum_{n-1}^{m}\{(P_0 R_n) - T_{n(min)}\}\left\{1 - e^{\left(\frac{-t}{C_n R_n}\right)}\right\} + T_{n(min)}$$
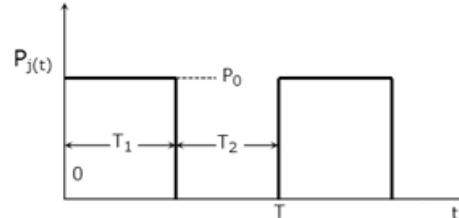
2) In the region where $P_{j(t)} = 0$:

$$T_{j(t)} = \sum_{n-1}^{m}\left\{T_{n(max)} e^{\left(\frac{-t}{C_n R_n}\right)}\right\}$$
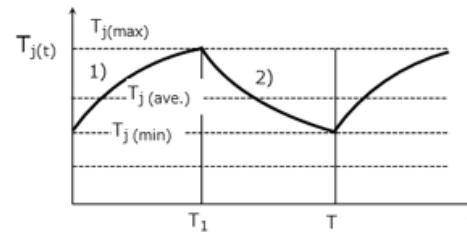
Figure 5.3 Temperature change due to a pulse train

   For typical transistors, the actual $P_{j(t)}$ value can be approximated by substituting 4 for n. However, if the C and R values are indefinite, it is difficult to calculate T_j. Therefore, T_{jpeak} is generally calculated using transient thermal impedance as follows:

Figure 5.4 shows an example of typical transient thermal impedance characteristics. Suppose that a single rectangular pulse (with a pulse width of t and a peak value of P_0) is applied. From the figure, we read the transient thermal impedance at a pulse width of t, and then use Equation 7 to calculate T_{jpeak}.

$$T_{jpeak} = r_{th(t)} P_O + T_a \quad \text{---} \quad \text{Equation 7}$$

When a repetitive pulse train with a cyclic period of T is applied as shown in Figure 5.3, T_{jpeak} is given by Equation 8 using the principle of superposition.
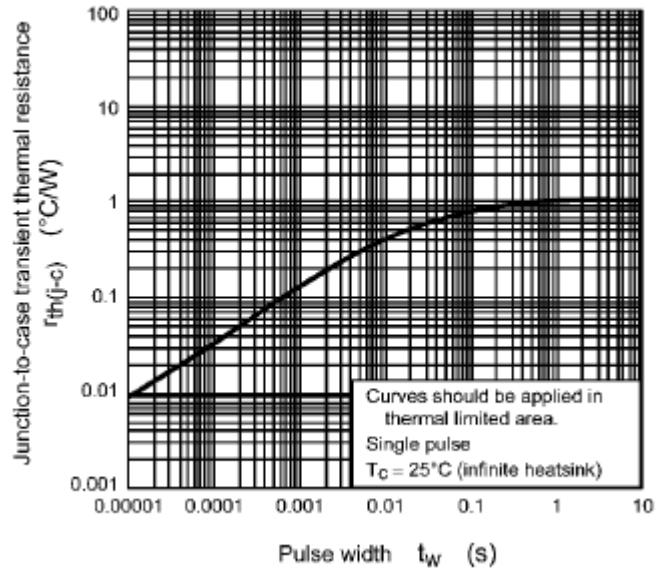
Figure 5.4 Transient thermal impedance

(a) Repetitive application of dissipation

$$P_{ave.} = \frac{1}{T} \int_0^T P_C\, dt = \frac{t}{T} P_O$$

(b) Approximated losses

$$P_O - P_{ave.} = P_O - \frac{t}{T} P_O = P_O \left(1 - \frac{t}{T} \right)$$

$$P_{ave.} = \frac{t}{T} P_O$$

(c) Using the principle of superposition

(d) Junction temperature
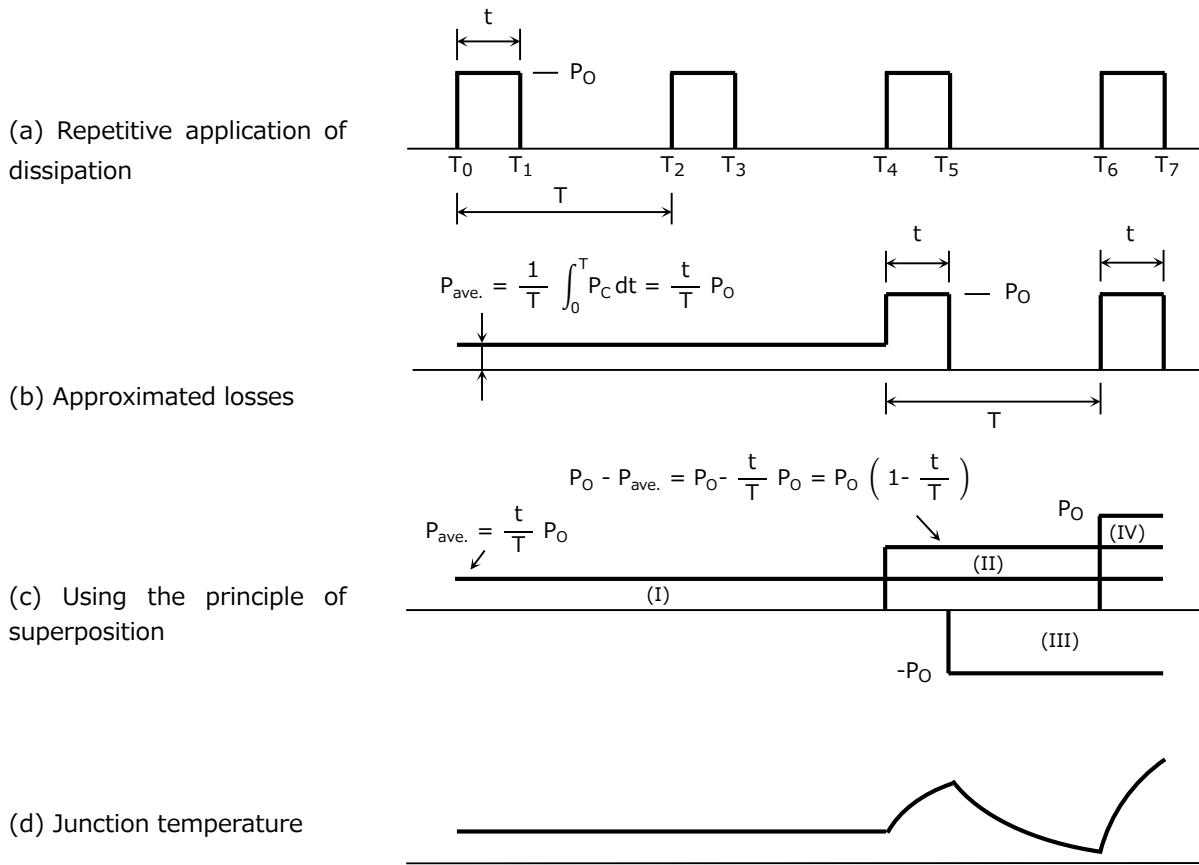
Figure 5.5 Calculating junction temperature using the principle of superposition

$$\Delta T_{j(I)} = P_O \frac{t}{T} R_{th(j-a)}$$

$$\Delta T_{j(II)} = P_O \left(1 - \frac{t}{T} \right) r_{th(T+t)}$$

$$\Delta T_{j(III)} = -P_O\, r_{th(T)}$$

$$\Delta T_{j(IV)} = P_O\, r_{th(t)}$$

$$T_{j\,peak} = P_O \left[\frac{t}{T} R_{th(j-a)} + \left(1 - \frac{t}{T} \right) r_{th(T+t)} - r_{th(T)} + r_{th(t)} \right] + T_a$$

-------------------- Equation 8

Equation 8 is applicable only to the thermally limited region of the SOA where no current concentration occurs due to secondary breakdown. Great care should be exercised in the thermal design for a pulsed power application to ensure that T_{jpeak} given by Equation 8 does not exceed the maximum rated junction temperature of the transistor.

The above description assumes that a rectangular waveform is applied to a transistor. However, for actual transistor applications, a rectangular $P_{j(t)}$ waveform is seldom used. In such cases, approximate the power loss waveform to a rectangular wave as shown in Figure 5.5 and use Equation 8 to estimate T_{jpeak}.

Sine and triangular waves can be approximated to rectangular waves as shown in Figure 5.6.

To obtain a rectangle with an area equal to a half-sine or triangular area, multiply the peak value of P_P by 0.7 in the case of (a) and (b), and multiple the pulse width by 0.91 for (a) and by 0.71 for (b).

In the case of (c) and (d), use the same peak value of P_P, and multiply the pulse width by 0.63 for (c) and by 0.5 for (d).
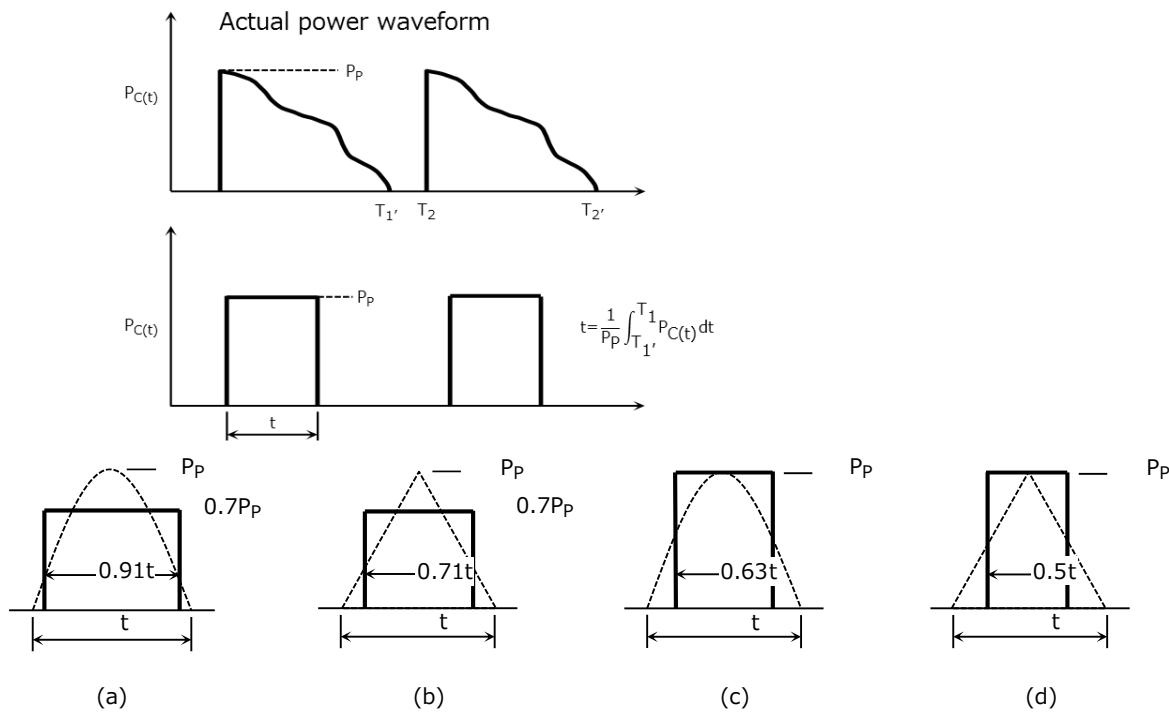
Figure 5.6 Approximating sine and triangular waves to rectangular waves

5.2. IGBT losses

Power losses of an IGBT mainly consist of conduction and switching losses. Conduction loss occurs while an IGBT is in the on state whereas switching loss occurs during turn-on and turn-off of transitions. In IGBTs with a diode, the diode also causes a loss, depending on the operation of the circuit in which they are used.

Figure 5.7 shows the switching waveform of an IGBT and the periods of time during which conduction and switching losses occur.

Special care should be taken as to power losses since they cause the heating of an IGBT.

5.2.1. Conduction loss

Conduction loss (P_{cond}) is calculated as collector current (I_C) multiplied by collector-emitter saturation voltage ($V_{CE(sat)}$):

$$P_{cond} = V_{CE(sat)} \times I_C$$

Conduction loss can be reduced by selecting an IGBT with lower $V_{CE(sat)}$ at the required collector current (I_C).

For applications in which IGBTs are turned on and off repeatedly, conduction loss can be calculated by multiplying P_{cond} by the duty cycle.

5.2.2. Switching loss

The switching loss of an IGBT can be approximated from its switching times. However, to ensure accuracy, it is necessary to measure switching loss under the actual operating conditions.

Since the switching times of an IGBT are affected by the collector current, gate resistance, and temperature, switching loss is also affected by these parameters. Turn-on loss energy (E_{on}) and turn-off loss energy (E_{off}) can be calculated in mJ as an integral of the product of the collector current (I_C) and the collector-emitter voltage (V_{CE}) values over a period of switching transitions. For IGBTs used in applications sensitive to switching losses, E_{on} and E_{off} are shown in their datasheets (Figure 5.8).

For applications in which IGBTs are turned on and off repeatedly, P_{on} and P_{off} can be calculated in watts by multiplying E_{on} (mJ) and E_{off} (mJ) by the switching frequency.
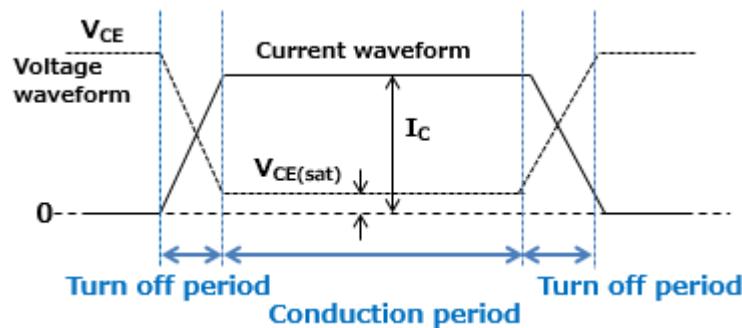
Figure 5.7 Switching waveform
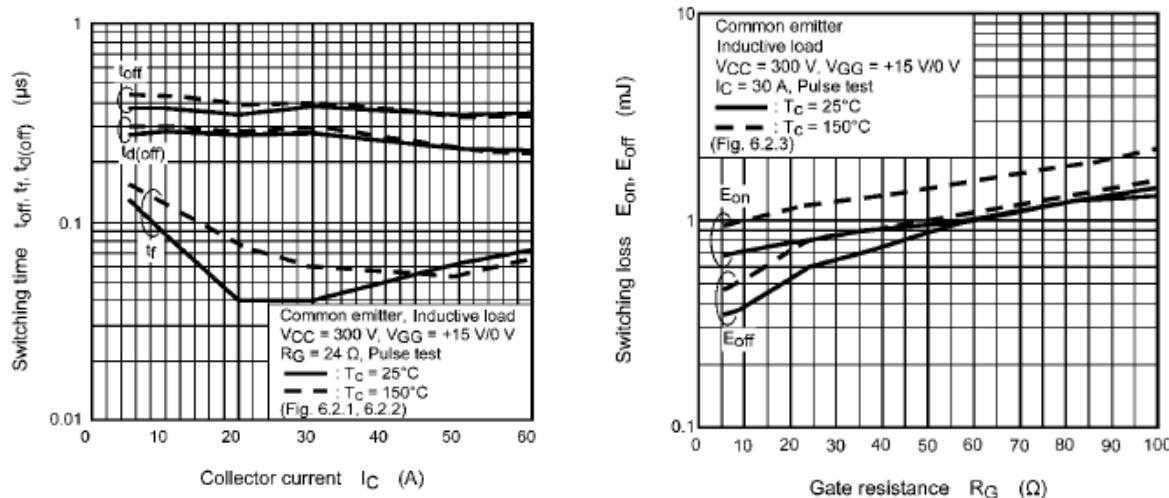
Figure 5.8 Switching losses

5.3. Gate drive

5.3.1. Gate voltage

The recommended forward gate bias of the IGBT is +15 V. In some cases, reverse gate bias is applied to IGBTs to increase their operating noise margin.

Too low forward gate bias leads to some problems such as the inability to obtain sufficient collector current and an increase in power losses due to an increase in the collector-emitter saturation voltage ($V_{CE(sat)}$).

Figure 5.9 shows the I_C-V_{CE} curves of an IGBT. At the recommended forward gate bias of +15 V, the IGBT provides sufficient collector current while maintaining the collector-emitter saturation voltage ($V_{CE(sat)}$) at a low level.
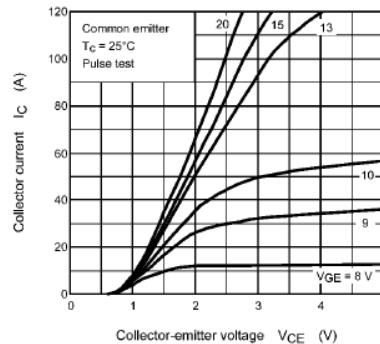
Figure 5.9 I_C-V_{CE} curves

5.3.2. Gate resistance

The gate charge current of an IGBT is constrained by gate resistance. Reducing gate resistance increases gate charge current and thus reduces the time required to charge the gate. This increases the switching speed and reduces switching loss. However, if the gate wire has inductance, gate oscillation or noise might occur. For motor applications, the turn-on di/dt and the turn-off surge voltage may cause problems. To avoid problems, it is necessary to select appropriate gate resistors. Figure 5.10 shows examples of switching time vs. gate resistance (R_G) curves.
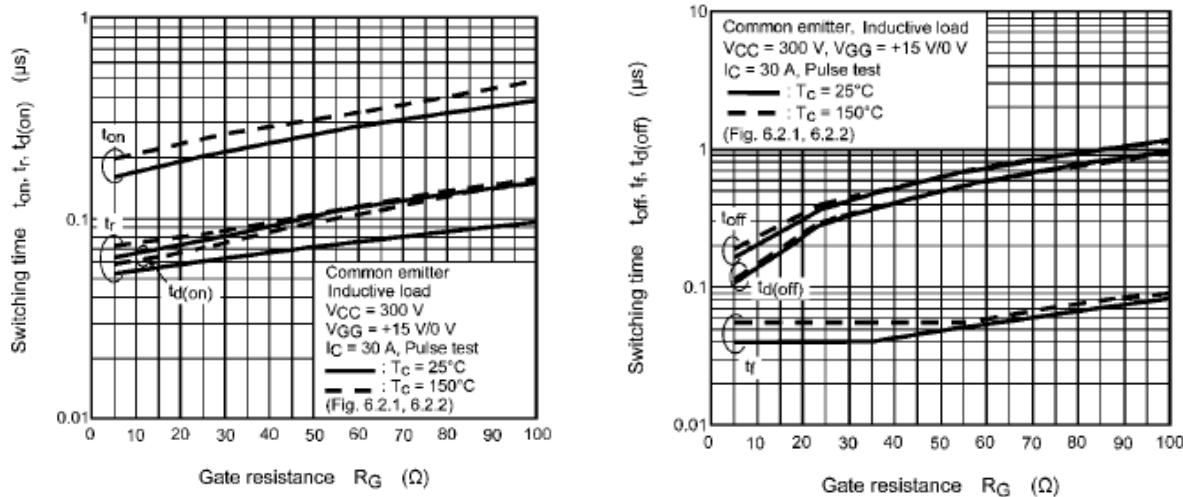
Figure 5.10 Switching time vs. gate resistance (R_G) curves

5.3.3. Dead time

Inverters for motor drive applications need dead time to prevent cross conduction in which the upper-arm and lower-arm IGBTs provide a direct short-circuit across power supply and GND lines. Dead time is a period of time during which both the upper-arm and lower-arm IGBTs are off. Dead time is inserted to force off both upper-arm and lower-arm IGBTs to ensure that they are never on at the same time while they are switching.

5.3.4. Gate drive current

The output stage of a gate drive circuit must have a current drive capability greater than the peak gate current ($I_{G(peak)}$) of the driven IGBTs.

As shown in Figure 5.11, $I_{G(peak)}$ is determined by the following equation from the values of a resistor at the signal source (R_S), an external gate resistor (R_G), an IGBT's internal resistance (r_g), and gate drive voltage (V_{GG}).

$$I_{G(peak)} = V_{GG}/(R_S+R_G+r_g)$$

In reality, the actual peak gate current ($I_{G(peak)}$) tends to be smaller than the result of the above equation due to a rise delay of the gate current, which is affected by several factors, including the rise time of a gate drive circuit and the inductance of the gate wire between the gate drive circuit and the IGBT.

In general applications, r_g is smaller than the optimally selected R_G, and R_S is small. Therefore, the above equation can be approximated as $I_{G(peak)}=V_{GE}/R_G$. V_{GE} is the gate-emitter voltage. If the gate drive circuit does not have a current capability greater than $I_{G(peak)}$, the switching speed of the IGBT might be adversely affected.
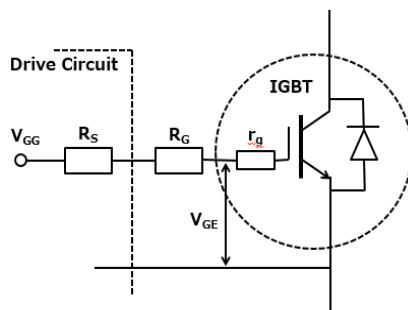
Figure 5.11 Resistance components of the gate drive circuit

5.3.5. Miscellaneous information

Gate drivers are selected from driver ICs, photocouplers, hybrid ICs, and other types of devices according to the application of IGBTs. Since the upper-arm and lower-arm IGBTs comprising a motor drive inverter have different GND levels, the gate drivers for the upper and lower arms need separate power supplies. Usually, a circuit that generates an upper-arm gate drive power supply from the lower-arm gate drive power supply or a high-voltage driver IC is used to meet the above requirement. (See Appendix B)

5.4. Parallel connection

Direct paralleling of IGBTs could cause the gate-emitter voltage and the collector current to oscillate.

The oscillation of the collector current might make IGBTs uncontrollable or cause an increase in power losses. To prevent oscillation, it is effective to attach separate gate resistors to the gate electrode of each IGBT as shown in Figure 5.12.
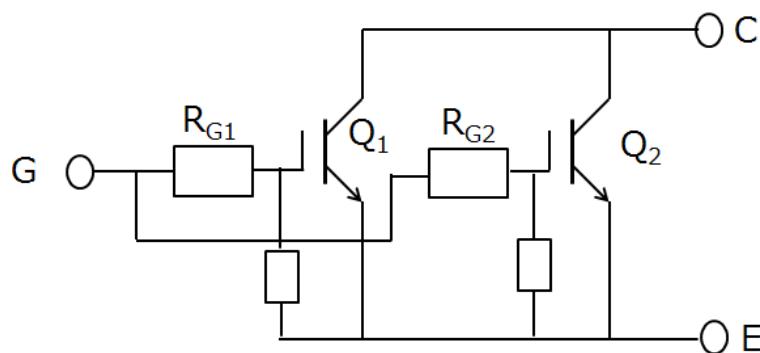
Figure 5.12 Gate resistors for parallel IGBTs

5.5. Surge voltage protection

In switching applications, a surge voltage (V_{surge}) could be applied to IGBTs at turn-off due to wire stray inductance. The surge voltage (V_{surge}) resulting from stray inductance is calculated as follows:

$$V_{surge} = -L_s \times di/dt + V_{CC}$$

(V_{CC}: Supply voltage)

IGBTs might be destroyed if V_{surge} exceeds their maximum rated voltage.

V_{surge} can be decreased by reducing di/dt and stray inductance. Since reducing di/dt compromises the switching performance of IGBTs, it is necessary to minimize stray inductance.

Snubber circuits can suppress a surge voltage. However, since snubber circuits adversely affect the switching performance of IGBTs, you need to understand the characteristics of different types of snubber circuits and select an appropriate one.

Examples of snubber circuits

Figure 5.13 shows three examples of snubber circuits.

(a) Snubber   (b) RDC snubber   (c) Separate RDC
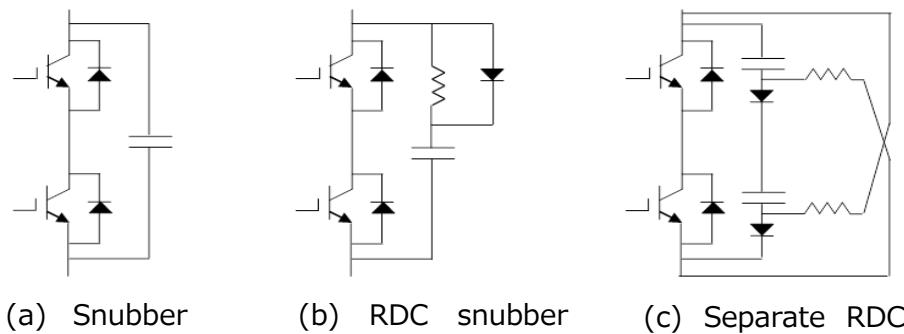
Figure 5.13 Examples of snubber circuits

The following briefly describes the circuit shown in Figure 5.13(c).

Figure 5.14 shows separate RDC snubbers for the upper and lower arms and the turn-off waveform of the IGBTs. ΔV_1 represents the surge voltage that cannot be suppressed owing to the inductance (L_2) of the snubber circuit. In other words, ΔV_1 is equal to the turn-off voltage generated by di/dt×L_2 at turn-off. C_s is calculated from the following equation:

$$1/2 \times L_1 \times (I_c)^2 = 1/2 \times C_s \times (\Delta V_2)^2$$
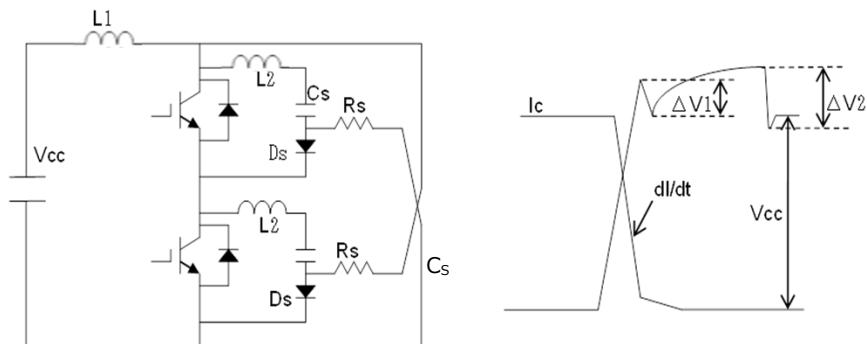
L_1: Inductance of the main wiring L_2: Inductance of the snubber circuit

R_s: Snubber resistor D_s: Snubber diode C_s: Snubber capacitor

Figure 5.14 Separate RDC snubbers for the upper and lower arms and turn-off waveform

In some cases, the snubber circuits shown in (a) and (c) are used in combination.

Appendixes

Appendix A: Thermal resistance

The thermal resistance values shown in the equivalent thermal circuit of Figure 5.1 can be explained as follows:

(1) Junction-to-case thermal resistance (internal thermal resistance): $R_{th(i)}$

The internal thermal resistance $R_{th(i)}$ from the junction of a transistor to the case depends on the structure and material of the transistor and differs from transistor to transistor.

To measure internal thermal resistance, the case of the transistor must be cooled to maintain a constant temperature. When the case temperature T_c is held at 25°C, the maximum allowable power dissipation, P_{Cmax}, of a transistor can be calculated as follows:

$$P_{C\,max} = \frac{T_{j\,max} - T_C}{R_{th(i)}} = \frac{T_{j\,max} - 25}{R_{th(i)}} \quad (W)$$

In the datasheets for high-power transistors, the maximum allowable collector power dissipation P_{Cmax} is specified either at $T_C = 25°C$ or assuming the use of an infinite heat sink. P_{Cmax} is determined by the internal thermal resistance of the transistor.

(2) Contact thermal resistance: $R_{th(c)}$

Contact thermal resistance $R_{th(c)}$ varies according to the condition of the contact surface between the case of a transistor and a heat sink. This condition is greatly affected by factors such as the evenness, coarseness, and area of contact, as well as the tightening of the transistor onto the heat sink. The influence of the coarseness and unevenness of the contact surface can be reduced by applying silicone grease or attaching silicone rubber.

(3) Insulation plate's thermal resistance: $R_{th(s)}$

If it is necessary to provide electrical insulation between a transistor and a heat sink, an insulation plate must be inserted between them. The thermal resistance of this insulation plate $R_{th(s)}$ varies with the materials, thickness, and area of the plate and is not negligible.

For packages insulated by mold resin, the thermal resistance specified for a transistor includes the insulator's thermal resistance $R_{th(s)}$.

(4) Heat sink's thermal resistance: $R_{th(f)}$

The thermal resistance of a heat sink can be considered as the distributed thermal resistance of a heat path from the surface of a heat sink to the ambient air. The thermal resistance of a heat sink depends on the condition of the ambient, a difference in temperature between the heat sink and the ambient air, and the effective area of the heat sink. It is difficult to mathematically express $R_{th(f)}$. Actually, $R_{th(f)}$ is obtained by measurement.

Figure 5.15 shows an example of thermal resistance data measured for a transistor standing

vertically at the center of a heat sink.

Various heat sinks are available from many vendors. The optimal heat sinks should be selected, referring to their technical datasheets.
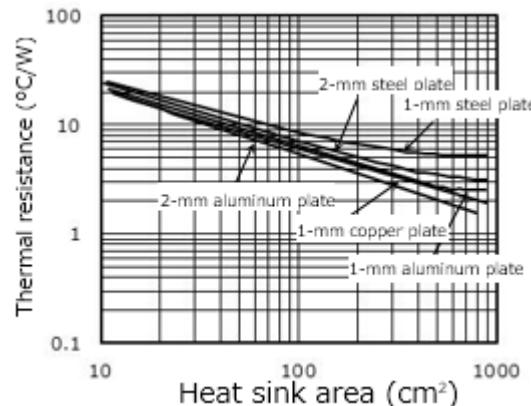
Figure 5.15 Heat sink surface area vs. thermal resistance

Appendix B: Power supply circuit

Charge pump

In Figure 5.16, the NMOS and PMOS transistors operate with mutually exclusive logic signals.

#1) While the NMOS transistor is on, C_1 is charged through the V_{CCL}-D_1-C_1-NMOS loop.

#2) This turns on the PMOS transistor, causing V_{CCH} to be formed through the C_1-D_2-V_{CCH}-PMOS-C_1 loop.

V_{CCH} is placed in such a manner as to be superimposed on the main power supply (V_{DD}). Gate driver circuits are configured via a freewheeling diode (FWD) of the main switching transistors. Since the drain of the high-side MOSFET acts as a reference (on the lower-voltage side) for V_{CCH}, the high side of a charge pump can also operate from a single power supply. In practice, however, charge pumps are not used for high-voltage applications since many of the devices including the PMOS and NMOS transistors need a breakdown voltage higher than $V_{DD}+V_{CCH}$.
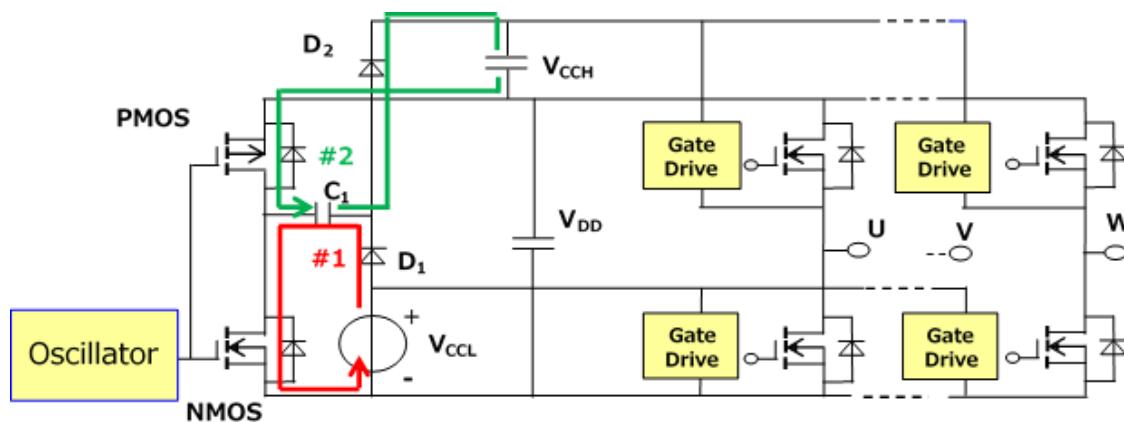
Figure 5.16 Charge pump

Bootstrap circuit

In Figure 5.17, the power supplies for the high-side gate drivers are primarily formed by the loops $V_{CC} \rightarrow$ (resistor) $\rightarrow$ D $\rightarrow C_1$ (or C_2 or C_3) $\rightarrow Q_2$ (or Q_4 or Q_6) $\rightarrow V_{CC}$. The midpoint voltage of each phase needs to be equal to the lower-side voltage of V_{CC} in order for C_1 to C_3 to be charged. This means C_1, C_2, and C_3 are charged while Q_2, Q_4, and Q_6 are on or while currents are flowing back through the FWDs. The charge voltages might decrease significantly, depending on the high-side and low-side on-off timings and the ratios of C_1 to C_3 to the input capacitances of the corresponding MOSFETs.
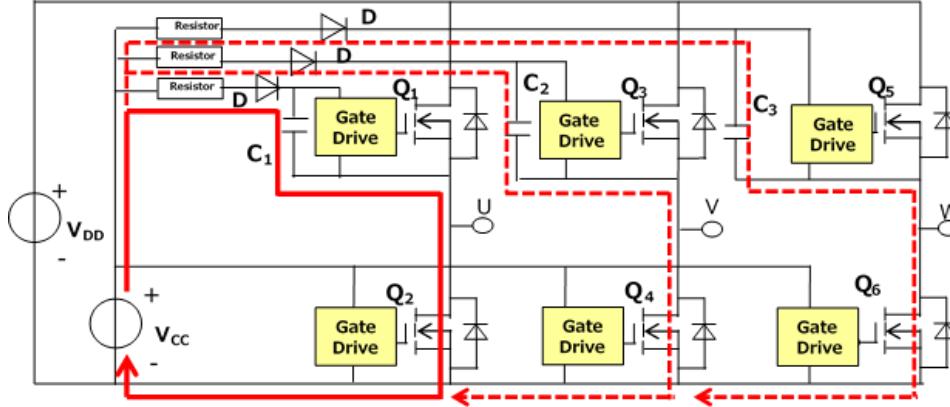
Figure 5.17 Bootstrap circuit

RESTRICTIONS ON PRODUCT USE