

Second Edition

John G. Kassakian
David J. Perreault
George C. Verghese
Martin F. Schlecht



Principles of Power Electronics

Principles of Power Electronics

Second Edition

Substantially expanded and updated, the new edition of this classic textbook provides unrivaled coverage of the fundamentals of power electronics.

It includes:

- Comprehensive and up-to-date coverage of foundational concepts in circuits, magnetics, devices, dynamic models, and control, establishing a strong conceptual framework for further study.
- Extensive discussion of contemporary practical considerations, enhanced by real-world examples, preparing readers for any design scenario from low-power dc/dc converters to multi-megawatt ac machine drives.
- New topics including SiC and GaN wide-bandgap materials, superjunction MOSFET and IGBT devices, advanced magnetics design, multi-level and switched-capacitor converters, RF converter circuits, and EMI.
- Over 300 new and revised end-of-chapter problems, designed to enhance and expand understanding of the material, with solutions for instructors.

Unique in its breadth and depth, and providing a range of flexible teaching pathways for instructors at multiple levels, this is the definitive guide to power electronics for graduate and senior undergraduate students in electrical engineering, and practicing electrical engineers.

John G. Kassakian is Professor of Electrical Engineering Emeritus at the Massachusetts Institute of Technology. He is the founding President of the IEEE Power Electronics Society, a Fellow of the IEEE, a member of the US National Academy of Engineering, and has taught, conducted research, and consulted in power electronics for over 45 years.

David J. Perreault is Ford Professor of Engineering at the Massachusetts Institute of Technology, with over 25 years of experience in power electronics research and teaching. He is a Fellow of the IEEE, and a member of the US National Academy of Engineering.

George C. Verghese is Henry Ellis Warren Professor of Electrical and Biomedical Engineering at the Massachusetts Institute of Technology, with over 40 years of research and teaching experience. He is an MIT MacVicar Faculty Fellow for outstanding contributions to undergraduate education, and a Fellow of the IEEE.

Martin F. Schlecht is the founder of SynQor, a supplier of high-performance power conversion solutions, and prior to that was Professor of Electrical Engineering at the Massachusetts Institute of Technology for 15 years. He has over 40 years of research, teaching, and industrial practice in power electronics.

"*Principles of Power Electronics* was a landmark in power electronics pedagogy when it was first published more than three decades ago. It is thrilling to see how these distinguished authors have not only thoroughly brought the book's core contents up to date, but also expanded its coverage to include several new topics that are increasingly important for students to learn as they prepare to enter this exciting field. Bravo!"

Thomas Jahns, *University Wisconsin – Madison*

"The second edition of the *Principles of Power Electronics* makes this classic book even more valuable. The book teaches power electronics from the ground up, providing the formal framework to learn its fundamentals and many advanced topics. This highly accessible book is an excellent text for a foundational course in power electronics. A must-have for both beginners and experienced practitioners."

Khurram Afridi, *Cornell University*

"The new edition of *Principles of Power Electronics* is a must for anyone in the field of power electronics, from the student learner to a working professional. The coverage is comprehensive, with detailed explanations backed up by well-chosen worked examples and illustrative problems. The treatment of magnetics, analysis, and design is particularly strong and a welcome addition."

Gerard Hurley, *University of Galway*

Principles of Power Electronics

SECOND EDITION

John G. Kassakian

Massachusetts Institute of Technology

David J. Perreault

Massachusetts Institute of Technology

George C. Verghese

Massachusetts Institute of Technology

Martin F. Schlecht

SynQor Inc.

 **CAMBRIDGE**
UNIVERSITY PRESS



CAMBRIDGE
UNIVERSITY PRESS

Shaftesbury Road, Cambridge CB2 8EA, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314-321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi – 110025, India

103 Penang Road, #05-06/07, Visioncrest Commercial, Singapore 238467

Cambridge University Press is part of Cambridge University Press & Assessment,
a department of the University of Cambridge.

We share the University's mission to contribute to society through the pursuit of
education, learning and research at the highest international levels of excellence.

www.cambridge.org

Information on this title: www.cambridge.org/highereducation/isbn/9781316519516

DOI: 10.1017/9781009023894

Second edition © John G. Kassakian, David J. Perreault, George C. Verghese, and
Martin F. Schlecht 2024

This publication is in copyright. Subject to statutory exception
and to the provisions of relevant collective licensing agreements,
no reproduction of any part may take place without the written
permission of Cambridge University Press & Assessment.

First published by Pearson College Div, 1991

Second edition published by Cambridge University Press & Assessment 2024 (version 2, February 2024)

Printed in Great Britain by CPI Group (UK) Ltd, Croydon CR0 4YY, February 2024

A catalogue record for this publication is available from the British Library.

Library of Congress Cataloging-in-Publication Data

Names: Kassakian, John G., author. | Perreault, David J., author. |

Verghese, George C., author. | Schlecht, Martin F., author.

Title: Principles of power electronics / John G. Kassakian,

David J. Perreault, George C. Verghese, Martin F. Schlecht.

Description: Second edition. | New York : Cambridge University Press, 2023.

| Includes bibliographical references and index.

Identifiers: LCCN 2023003028 (print) | LCCN 2023003029 (ebook) |

ISBN 9781316519516 (hardback) | ISBN 9781009023894 (epub)

Subjects: LCSH: Power electronics.

Classification: LCC TK7881.15 .K37 2023 (print) | LCC TK7881.15 (ebook) |

DDC 621.31/7–dc23/eng/20230126

LC record available at <https://lcn.loc.gov/2023003028>

LC ebook record available at <https://lcn.loc.gov/2023003029>

ISBN 978-1-316-51951-6 Hardback

Additional resources for this publication at www.cambridge.org/Kassakian_et_al

Cambridge University Press & Assessment has no responsibility for the persistence or accuracy of
URLs for external or third-party internet websites referred to in this publication
and does not guarantee that any content on such websites is, or will remain,
accurate or appropriate.

To our students,
who have been our best teachers,

and to Daniel Perreault Nakajima,
in memoriam.

Contents

Preface	page xv
Acknowledgments	xviii
1 Introduction	1
1.1 Power Electronic Circuits	1
1.2 Power Semiconductor Switches	2
1.3 Transformers	5
1.4 Nomenclature	7
1.5 Bibliographies	8
1.6 Problems	8
Part I Form and Function	
2 Form and Function: An Overview	11
2.1 Functions of a Power Circuit	11
2.2 AC/DC Converters	13
2.3 DC/DC Converters	18
2.4 AC/AC Converters	20
2.5 Influence of Switch Implementation	22
Problems	24
3 Introduction to Rectifiers	27
3.1 Power in Electrical Networks	27
3.2 Single-Phase Half-Wave Rectifier	29
3.3 AC-Side Reactance and Current Commutation	35
3.4 Measures and Effects of Distortion	38
3.5 Bridge Rectifiers	44
Notes and Bibliography	49
Problems	50
4 Phase-Controlled Converters	57
4.1 Single-Phase Configurations	58
4.2 Phase Control with AC-Side Reactance	64
4.3 Inversion Limits	67

Notes and Bibliography	71
Problems	72
5 Pulse-Width-Modulated DC/DC Converters	76
5.1 The DC/DC Converter Topology	77
5.2 The Canonical Switching Cell	82
5.3 Direct Converter	82
5.4 Indirect Converter	89
5.5 Other PWM DC/DC Converters	92
5.6 Choice of Capacitor and Inductor Values	95
5.7 Semiconductor Device Stresses	103
5.8 Three-Level Flying-Capacitor Converter	104
5.9 Converter Operation with Discontinuous Conduction	106
Notes and Bibliography	108
Problems	109
6 Switched-Capacitor and Related Converters	116
6.1 Switched-Capacitor DC/DC Converter	116
6.2 Two-State Switched-Capacitor Converter	118
6.3 Switch Implementation	121
6.4 Other Switched-Capacitor DC/DC Converters	123
6.5 Other Kinds of Switched-Capacitor Converters	131
Notes and Bibliography	132
Problems	133
7 Isolated Pulse-Width-Modulated DC/DC Converters	136
7.1 Single-Ended Isolated Forward Converter	136
7.2 Double-Ended Isolated Forward Converter	145
7.3 Dual Active-Bridge Converter	150
7.4 Flyback Converter	153
7.5 Other PWM Isolated Converters	154
7.6 Effects of Transformer Leakage Inductance	155
7.7 Converters with Multiple Outputs	158
Notes and Bibliography	158
Problems	159
8 Single-Phase Switched-Mode DC/AC Converters	167
8.1 Basic Variable-Frequency Bridge Converter	167
8.2 Harmonic Reduction	172
8.3 Pulse-Width-Modulated DC/AC Converters	179
8.4 Current Control of Inverters	187

8.5	Multi-level Converters	190
8.6	Transformer-Isolated DC/AC Converters	192
8.7	Other DC/AC Converter Topologies	194
8.8	Power Balance in Single-Phase DC/AC Converters	196
8.9	Switched-Mode Rectifiers and Power Factor Correction	197
	Notes and Bibliography	202
	Problems	203
9	Polyphase Sources and Converters	207
9.1	Polyphase Sources	207
9.2	Three-Phase Sources	209
9.3	Introduction to Polyphase Rectifier Circuits	213
9.4	Phase-Controlled Three-Phase Converters	215
9.5	Commutation in Polyphase Rectifiers	216
9.6	Three-Phase Inverters	221
9.7	Space-Vector Representation and Modulation for Three-Phase Systems	229
9.8	Multi-Level Inverters	237
	Notes and Bibliography	242
	Problems	243
10	Resonant Converters	249
10.1	Review of Second-Order System Behavior	250
10.2	Quality Factor	254
10.3	Resonant Converter Analysis	256
10.4	Soft Switching of Resonant Converters	260
10.5	Resonant DC/DC Converters	264
10.6	Radio-Frequency Converters	271
	Notes and Bibliography	284
	Problems	285
11	AC/AC Converters	289
11.1	Introduction to AC/AC Converters	289
11.2	Energy Storage Requirements in a DC-Link Converter	290
11.3	The Naturally Commutated Cycloconverter	294
11.4	An Isolated High-Frequency-Link Cycloconverter	299
11.5	Solid-State Transformer	302
11.6	Matrix Converter	305
	Notes and Bibliography	306
	Problems	306

Part II Dynamic Models and Control

12 Dynamic Models and Control: An Overview	313
12.1 Dynamic Behavior of Power Converters	314
12.2 Dynamic Models	322
12.3 Averaged-Circuit Models	323
12.4 Linearized Models	329
12.5 Feedback Control	332
Notes and Bibliography	342
Problems	343
13 Averaged-Circuit and State-Space Models	346
13.1 Averaged-Circuit Models	346
13.2 Generalizing Circuit Averaging to the Fundamental Component	350
13.3 Continuous-Time State-Space Models	354
13.4 Discrete-Time or Sampled-Data Models	365
13.5 Generalized State-Space Models	368
13.6 Models for Controllers and Interconnected Systems	372
Notes and Bibliography	373
Problems	374
14 Linear Models and Feedback Control	378
14.1 Linearization	378
14.2 Linearizing an Averaged-Circuit Model	379
14.3 Linearizing Continuous-Time State-Space Models	383
14.4 Analysis of Continuous-Time LTI Models	385
14.5 Piecewise LTI Models	394
14.6 Linearizing Discrete-Time Generalized State-Space Models	395
14.7 Analysis of Discrete-Time LTI Models	399
14.8 Feedback Control Design	405
14.9 Multi-Loop Control	414
14.10 State Feedback	422
14.11 Digital Control	425
Notes and Bibliography	428
Problems	429

Part III Components and Devices

15 Components and Devices: An Overview	437
15.1 Practical Semiconductor Switches	438
15.2 Practical Energy Storage Elements	438
15.3 Semiconductor Devices	439

15.4	Capacitors	452
15.5	Inductors and Transformers	455
	Problems	457
16	Review of Semiconductor Devices	461
16.1	Elementary Physics of Semiconductors	461
16.2	Simple Analysis of a Diode	472
16.3	Bipolar Junction Transistor	478
16.4	MOSFET	483
16.5	The Safe Operating Area	486
	Notes and Bibliography	487
	Problems	488
17	Power Semiconductor Devices	494
17.1	Bipolar Diode	494
17.2	Switch Transitions in a pin Diode	504
17.3	The Schottky Barrier Diode	507
17.4	Power BJT	511
17.5	Power MOSFET	517
17.6	Superjunction MOSFET	523
17.7	IGBT	524
17.8	Silicon Carbide and Gallium Nitride Devices	528
17.9	Thyristor	530
17.10	Datasheet	538
	Notes and Bibliography	541
	Problems	542
18	Introduction to Magnetics	545
18.1	Inductor	546
18.2	Saturation, Hysteresis, and Residual Flux	555
18.3	Losses In Magnetic Components	557
18.4	Transformers	566
18.5	Magnetic Material Properties	575
	Notes and Bibliography	576
	Problems	577
19	Magnetic Component Modeling	582
19.1	Mathematical Representation: Inductance Matrix	582
19.2	Circuit Representations for Two-Port Transformers	584
19.3	Determining Transformer Parameters	587
19.4	Multi-Winding Model	588

19.5	Extending the Magnetic Circuit Model Concept	597
19.6	Physical Electric Circuit Models for Magnetic Systems	603
	Notes and Bibliography	607
	Problems	608
20	Introduction to Magnetics Design	612
20.1	Filter Inductor Design and Core Factor	613
20.2	Thermal Constraints in Magnetics Design	618
20.3	Inductor Energy Storage Limits	621
20.4	AC Magnetics Sizing and Core Area Product	624
20.5	Performance Factor for Magnetic Materials	634
20.6	An Iterative Inductor Design Algorithm	636
	Notes and Bibliography	638
	Problems	640
21	Magnetics Loss Analysis and Design	643
21.1	Magnetic Diffusion	643
21.2	Winding Loss Calculation	655
21.3	Core Loss	675
	Notes and Bibliography	684
	Problems	686
Part IV	Practical Considerations	
22	Practical Considerations: An Overview	693
22.1	Gate and Base Drives	693
22.2	Thyristor Commutation Circuits	695
22.3	Snubbers, Clamps, and Soft Switching	697
22.4	Thermal Considerations	700
	Problems	701
23	Gate and Base Drives	704
23.1	MOSFET and IGBT Gate Drives	705
23.2	Bipolar Transistor Base Drives	713
23.3	Thyristor Gate Drives	716
	Notes and Bibliography	719
	Problems	719
24	Snubber Circuits, Clamps, and Soft Switching	724
24.1	Turn-Off Snubber	725
24.2	Turn-On Snubber	730

24.3	Combined Turn-On/Turn-Off Snubber	733
24.4	Alternative Placements of the Snubber Circuit	735
24.5	Dissipation in Snubber Circuits	736
24.6	Soft Switching	739
	Notes and Bibliography	748
	Problems	750
25	Thermal Modeling and Heat Sinking	753
25.1	Static Thermal Models	754
25.2	Thermal Interfaces	758
25.3	Transient Thermal Models	762
	Notes and Bibliography	767
	Problems	768
26	Electromagnetic Interference and Filtering	771
26.1	EMI Specifications and Measurement	771
26.2	Filter Design	773
26.3	Common-Mode and Differential-Mode Variables	781
26.4	Parasitics and Circuit Layout	787
	Notes and Bibliography	789
	Problems	790
	Index	793

In this chapter we describe power electronics and present a brief introduction to semiconductor switching devices and magnetic components. An introduction to these circuit elements is necessary because we use them in Part I, although we do not discuss them in detail until Part III. We also introduce in this chapter nomenclature that we use throughout the book.

1.1 Power Electronic Circuits

The dominant application of electronics today is to process information. The computer industry is the biggest user of semiconductor devices, and consumer electronics, including cameras and cell phones, is second. While all these applications require power (from a wall plug or a battery), their primary function is to process information; for instance, to take the digital optical signal produced by a cell phone camera and transform it into a photographic image. Power electronic circuits, on the other hand, are principally concerned with processing energy. They convert electrical energy from the form supplied by a source to the form required by a load. For example, the part of a computer that takes the ac mains voltage and changes it to the low-voltage dc required by the logic chips is a power electronic circuit (often abbreviated as *power circuit*). In many applications, the conversion process concludes with mechanical motion. In these cases the power circuit converts electric energy to the form required by the electromechanical transducer, such as a dc or ac motor.

Efficiency is an important concern in any energy processing system, for the difference between the energy into the system and the energy out is usually converted to heat. Although the cost of energy is sometimes a consideration, the most unpleasant consequence of generating heat is that it must be removed from the system. This consideration alone largely dictates the size of power electronic apparatus. A power circuit must, therefore, be designed to operate as efficiently as possible. The efficiency of very large systems exceeds 99%. High efficiency is achieved by using power semiconductor devices (where their voltage is nearly zero when they are on, and their current is nearly zero when they are off) to minimize their dissipation.[†] The only other components in the basic power circuit are inductors, capacitors, and transformers, so the ideal power circuit is lossless.

[†] Exceptions, such as linear voltage regulators, are so few that we do not consider them explicitly in this book.

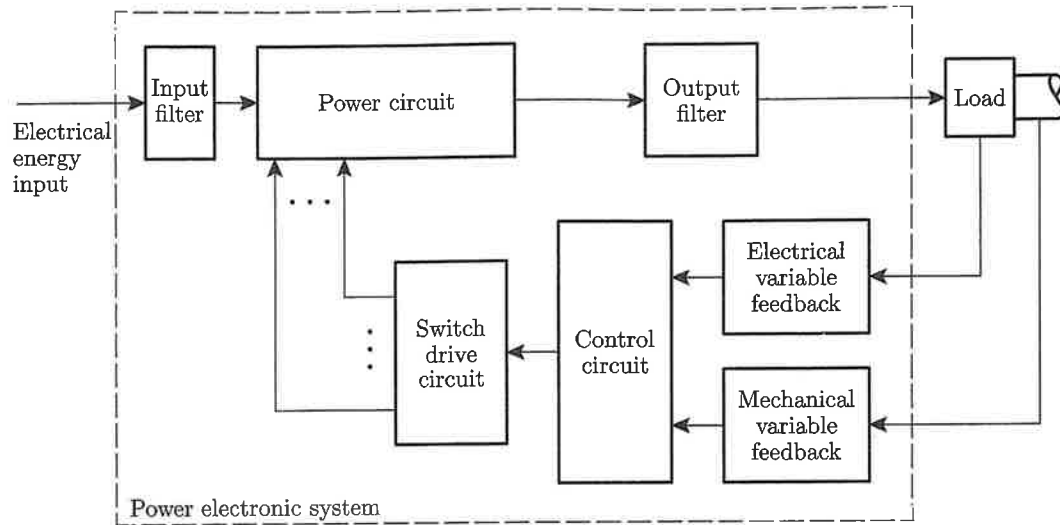


Figure 1.1 A block diagram of a typical power electronic system.

A power electronic system typically consists of much more than a power circuit, as shown in Fig. 1.1. Switching creates waveforms with harmonics that may be undesirable because they interfere with proper operation of the load or other equipment, so filters are often employed at the inputs and outputs of the power circuit. The system load, which may be electrical or electromechanical, is controlled via the feedback of electrical and/or electromechanical variables to a control circuit. This control circuit processes the feedback signals and drives the switches in the power circuit according to the demands of these signals. The system also includes mechanical elements, such as heatsinks and structures to support the physically large components of the power circuit.

In our circuit drawings, we use a capacitor symbol with a curved plate to differentiate it from the symbol for a contactor. We also find it a more elegant symbol than a pair of parallel plates. For electrolytic capacitors, which are polarized, the curved plate represents the negative terminal. We do not worry about this in our diagrams as the type of capacitor is irrelevant.

1.2 Power Semiconductor Switches

The basic semiconductor devices used as switches in power electronic circuits are the bipolar and Schottky diodes, the bipolar junction transistor (BJT), the metal-oxide-semiconductor field-effect transistor (MOSFET), the insulated-gate bipolar transistor (IGBT), and a class of latching bipolar devices known as thyristors, the most common of which is the silicon controlled rectifier (SCR). Their circuit symbols and operating regions in the v - i plane are shown in Fig. 1.2. We discuss these and other hybrid devices in detail in Part III.

What follows is a brief description of the salient operating characteristics of each device shown in Fig. 1.2. This information allows us to present the basic operation of power electronic circuits without first mastering Part III.

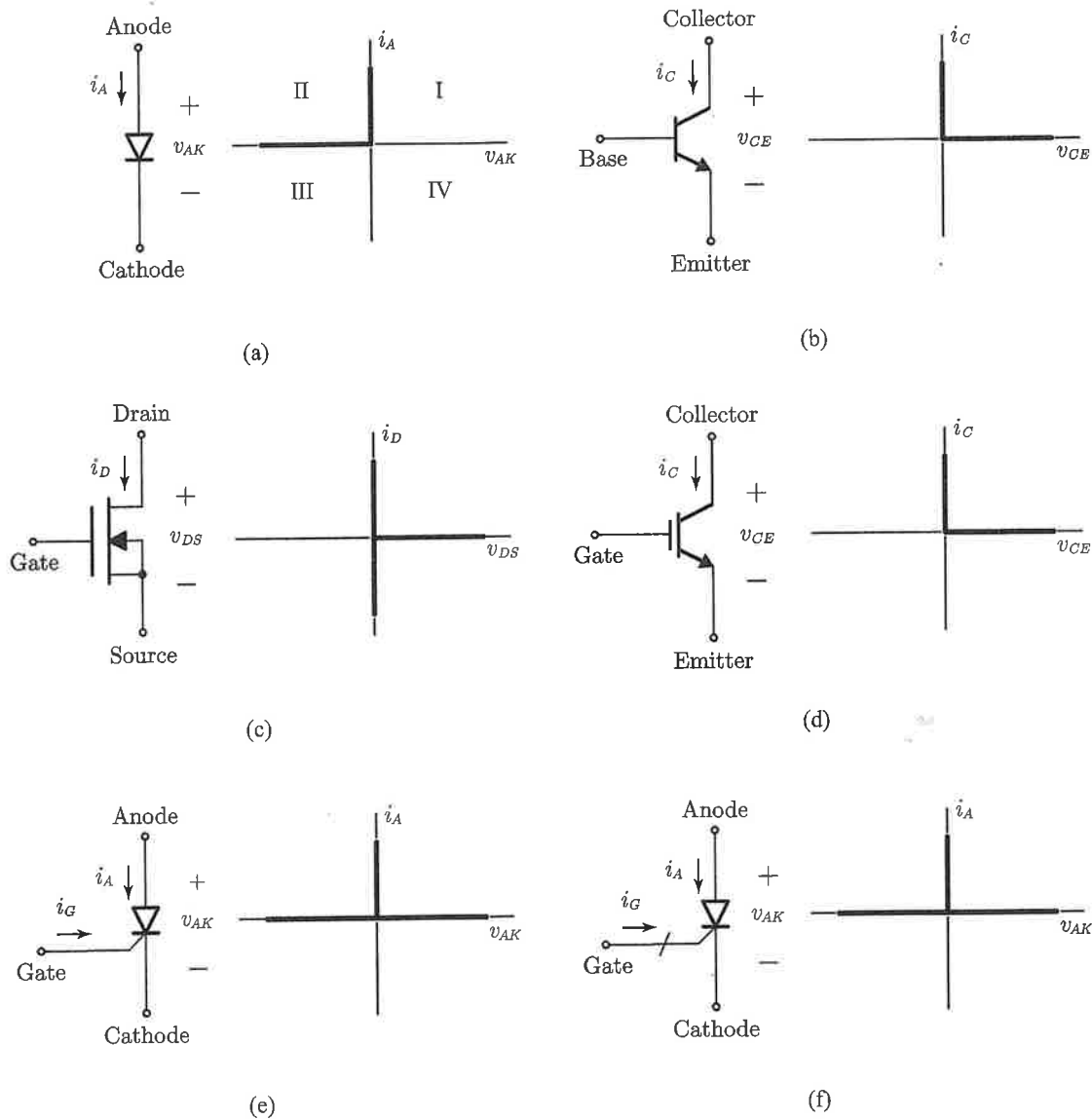


Figure 1.2 Circuit symbols and operating regions for power semiconductor devices: (a) diode; (b) npn bipolar junction transistor (BJT); (c) (n-channel) power metal-oxide-semiconductor field-effect transistor (MOSFET); (d) insulated gate bipolar transistor (IGBT); (e) silicon controlled rectifier (SCR); (f) gate turn-off thyristor (GTO).

1.2.1 Diode

The diode, whose symbol and variable definitions are shown in Fig. 1.2(a), is an uncontrollable semiconductor switch. It is uncontrollable because whether it is on or off is determined by the voltages and currents in the network, not by any action we can take. When on, its anode current, i_A , is positive. When off, its anode-cathode voltage, v_{AK} , is negative.[†] The diode switches in response to the behavior of its terminal variables. If it is off and the circuit causes v_{AK} to try

[†] The use of “K” instead of “C” reflects the Greek origin of the word cathode, or *kathodos*, meaning “way down,” that is, the negative terminal.

to go positive, the diode will turn on. If on, the diode will turn off if the circuit tries to force i_A to go negative.

1.2.2 Transistor

Transistors, whether of the bipolar or MOS type, are fully controllable switches. They possess a third terminal (the *base* terminal for the BJT and the *gate* terminal for the MOSFET and IGBT) from which we can turn the device on and off. Control of the BJT requires a base current, while the MOSFET and IGBT require a gate voltage. The symbols and terminal variables for the npn BJT, n-channel power MOSFET, and IGBT are shown in Fig. 1.2(b), (c), and (d) respectively.

The BJT and IGBT can carry current in only one direction. For the npn BJT and IGBT shown in the figure, this direction is $i_C > 0$. The power MOSFET, because of its physical structure, can carry current in either direction. The unique structure of the power MOSFET results in a diode, known as the *body diode*, between its source and drain, represented by the arrowhead in the symbol. While this diode allows negative drain current conduction, this negative current can be supported with a lower forward drop by applying a gate signal to the device to turn on the channel and use it instead of the body diode to conduct the current.

When off, all three transistors can support only one polarity of voltage, which, for the transistors shown, are $v_{CE} > 0$ and $v_{DS} > 0$. These voltage and current polarities are reversed for the pnp BJT and the p-channel MOSFET. For reasons discussed in Part III, npn and n-channel devices are the most commonly used types of power transistors. Because of their ease and efficiency of control, and their low on-state voltage, MOSFETs and IGBTs are the dominant form of switch used in power electronic circuits.

1.2.3 Thyristor

The only members of the thyristor family that we describe in this introduction are the SCR and the gate turn-off thyristor (GTO), whose circuit symbols are shown in Fig. 1.2(e) and (f).

The SCR is a switch that in some ways can be thought of as a “semi-controllable” diode. If no signal is applied to the gate, the device will remain off, independent of the polarity of v_{AK} . To turn the SCR on, a brief pulse of current, i_G , is applied to the gate terminal during a time when $v_{AK} > 0$. This initiates a regenerative turn-on process that quickly latches the SCR in the on state, in which $v_{AK} \approx 0$ and the gate no longer has any control over the device. When in this on state, the SCR can conduct only positive i_A . It turns off when i_A tries to go negative. So once on, the SCR behaves as a diode. In summary, the SCR is a diode whose turn-on can be inhibited by not applying a gate pulse.

The GTO is an SCR which has been constructed to enable a negative gate current to turn it off in the presence of $i_{AK} > 0$. The turn-off gain, defined as the ratio of I_{AK} to the negative gate current required to turn the device off, is typically in the low single digits. Two types of GTOs are available: asymmetrical GTOs, which cannot block a high reverse voltage, and symmetrical GTOs, which can.

1.3 Transformers

Transformers are a prominent feature of power electronic circuits. We treat them extensively in Part III, but the following introduction to their behavior permits us to use them as circuit elements in Parts I and II.

Transformers are employed to provide electric isolation and the step-up or step-down of ac voltages and currents. The *ideal transformer* shown in Fig. 1.3(a) has two windings of N_1 and N_2 turns. Dots indicate the orientation of the windings. If a voltage is applied to one winding so that the dot is positive, the dotted ends of all the other windings (only one in this case) are also positive. If its terminal variables are defined relative to the dots, as shown in Fig. 1.3(a), the ideal transformer has the following terminal relationships:

$$\frac{v_1}{v_2} = \frac{N_1}{N_2}, \quad (1.1)$$

$$\frac{i_1}{i_2} = -\frac{N_2}{N_1}. \quad (1.2)$$

From (1.1) and (1.2), we deduce that $v_1 i_1 = -v_2 i_2$; that is, the instantaneous power into one port is equal to the instantaneous power out of the other. The ideal transformer neither dissipates nor stores energy. A straightforward application of the above relations also shows that if an impedance of value Z_1 is connected to terminals 1–1', an impedance of value $Z_2 = (N_2/N_1)^2 Z_1$ is measured at terminals 2–2'.

A transformer is ideal if it obeys (1.1) and (1.2), but no practical transformer is ideal. In most transformers, the principal departures from ideal result in some voltage and current being “lost” in the transformation, so terminal variables are not precisely related by (1.1) and (1.2). A model that represents these effects is shown in Fig. 1.3(b). Some of the terminal current i'_1 is shunted through the *magnetizing inductance* L_μ and is called the *magnetizing current*. This is the current required to establish the necessary flux in the transformer core. So whereas i_1 and i_2 are still related by (1.1) and (1.2), the real terminal currents i'_1 and i'_2 are not. Similarly, the real terminal voltages v'_1 and v'_2 differ from v_1 and v_2 by the drops across $L_{\ell 1}$ and $L_{\ell 2}$, which are called *leakage*

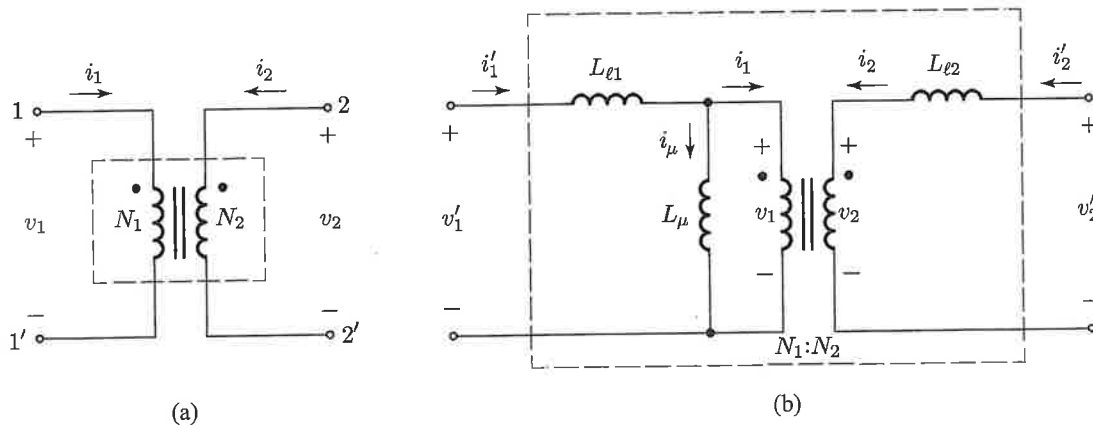


Figure 1.3 (a) The ideal transformer model. (b) A more practical model, with the effects of magnetizing inductance (L_μ) and leakage ($L_{\ell 1}$ and $L_{\ell 2}$) included.

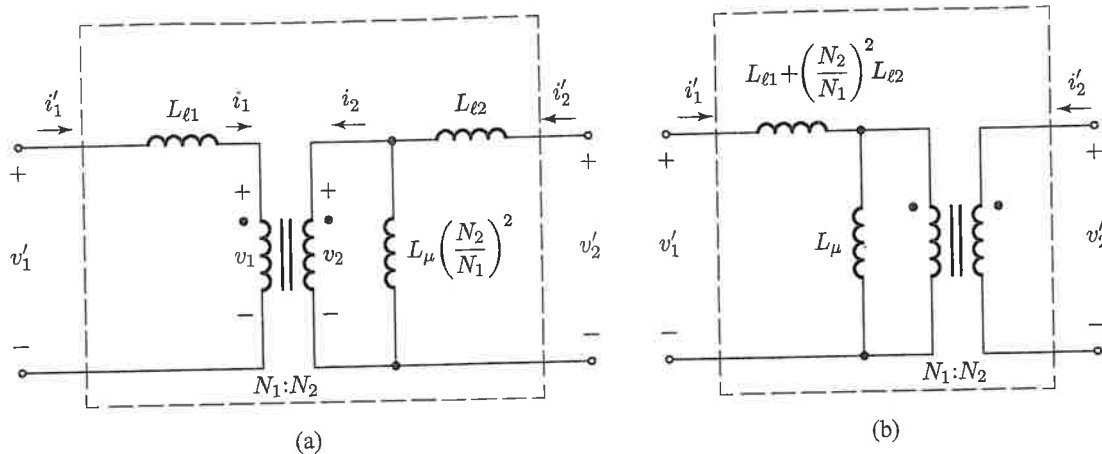


Figure 1.4 (a) The model of Fig. 1.3(b), with the magnetizing inductance placed on the N_2 side of the ideal transformer. (b) The model of Fig. 1.3(b), simplified by reflecting $L_{\ell 2}$ through the ideal transformer and combining it with $L_{\ell 1}$.

inductances. These represent flux linking one winding of the transformer but not the other. In Chapter 18 we describe the physical origins of these effects in much greater detail.

Figure 1.3(b) shows L_μ across the winding N_1 . We can, however, *reflect* it through the ideal transformer so that it appears across the N_2 winding, as shown in Fig. 1.4(a). Sometimes we do this because the result is analytically more convenient to use. Although two leakage inductances, one for each winding, are shown in Fig. 1.3(b), they are often combined by reflecting one through the ideal transformer. If the voltage drop across this inductor is small relative to the voltage across L_μ , then L_μ can be moved inside this reflected inductance without introducing much of an error, and the two leakage inductances can be combined. The resulting approximate model is shown in Fig. 1.4(b).

Another useful model transformation is to reflect the entire circuit on one side of the ideal transformer to the other side. A transformation of this kind is shown in Fig. 1.5. There, the N_2 side circuit, C_o and R_o , has been “brought through” the ideal transformer. Of course, the isolation function is lost in the transformation, which makes the technique inappropriate for the analysis of some circuits.

We can calculate or measure the leakage and magnetizing inductances of transformers, and we sometimes construct transformers to have specific values for these parameters. And, even though we have been discussing only two-winding transformers, similar but somewhat more complicated considerations apply to the modeling of transformers with more than two windings. Other practical considerations, such as the resistance of the windings or losses in the core, are represented by the addition of appropriate elements to the model of Fig. 1.3(b).

Figures 1.3(b) and 1.4 show the schematic transformer representation that we use throughout this book: the circuit model being used to describe a transformer is enclosed in a dashed box. The model frequently has an ideal transformer as one of its elements, represented by windings with adjacent double bars. Some schematic conventions utilize the double bars to represent an iron core, but we use the bars to indicate the coupled windings of an ideal transformer when

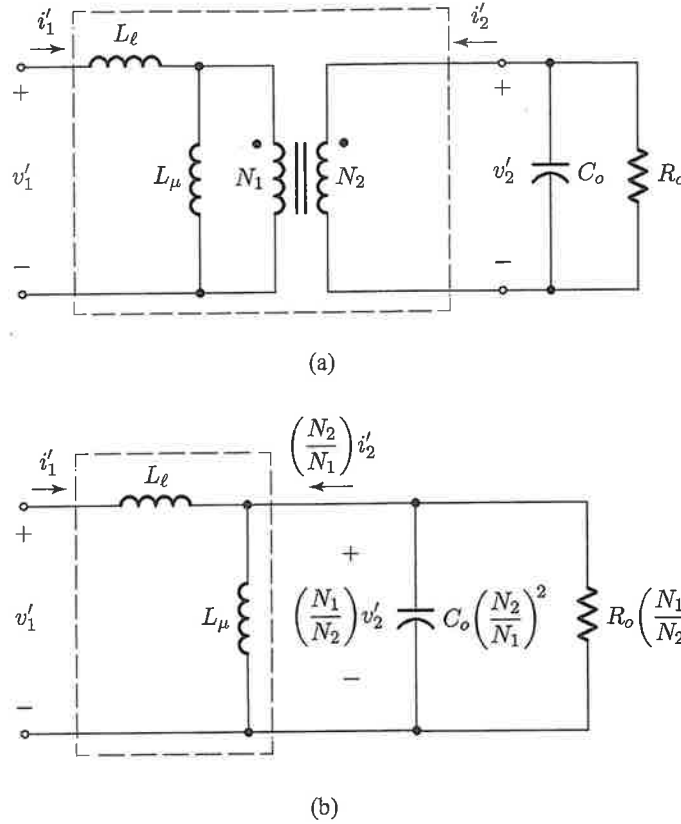


Figure 1.5 (a) A transformer with an RC load on the N_2 side. (b) The circuit of (a) with all the N_2 side components reflected to the N_1 side so that the ideal transformer can be eliminated from the transformer model.

it appears inside a dashed box. This convention avoids ambiguity and schematic clutter when more than two windings are involved.

1.4 Nomenclature

Because we discuss several different kinds of variables, we establish some notation and definitions now to permit a quick entry into the subject, and to avoid confusion later.

- Variables that may be time dependent are represented by lowercase names, such as v_1 . When necessary for clarification, the time dependence is explicitly indicated, as in $v_1(t)$.
- Variables that are held constant are represented by uppercase names, such as V_1 or V_{dc} .
- The *average value* or *dc component* of a variable that varies periodically in time is denoted by angle brackets around the variable, for example $\langle v_o(t) \rangle = V_o$. Since this average value is a constant, it is represented by an uppercase name.
- The *local average* is defined for a possibly non-periodic waveform, say $x(t)$. It is the average over a time window of fixed length that moves with t ; the window is usually chosen to be short

relative to the timescales of interest. The particular version defined in Chapter 12 is denoted by an overbar, so $\bar{x}(t)$.

- Perturbations around a constant value are indicated by a tilde, for instance $\tilde{v}_C = v_C - V_C$.
- The *root mean square* (rms) value of a periodic variable $v(t)$ with period T is denoted by V_{rms} and defined as $V_{\text{rms}} = \sqrt{(1/T) \int_0^T v^2(\tau) d\tau}$. The integral may be taken over any contiguous interval of length T .
- Harmonic components of a (non-sinusoidal) periodic waveform are indicated by an additional subscript representing the harmonic number, for example $v_a = v_{a_1} + v_{a_3} + v_{a_5} + \dots$.
- A sinusoidal function $v(t)$ can be written in the form

$$v(t) = V \cos(\omega t + \phi) = \text{Re} (V e^{j\phi} e^{j\omega t}) = \text{Re}(\hat{V} e^{j\omega t}),$$

where “Re” denotes the real part of the complex number that follows it. The *complex amplitude* of $v(t)$ is the complex number $\hat{V} = V e^{j\phi}$, and (particularly when plotted in the complex plane) is also termed the *phasor* associated with the sinusoid.

1.5 Bibliographies

We include an annotated bibliography at the end of most chapters. It provides sources of additional information on topics that you might want to pursue further.

1.6 Problems

Each chapter includes end-of-chapter problems that are designed to introduce variations or extensions of the chapter’s material, or to provide students the opportunity to examine more closely concepts or circuits introduced in the chapter. They are not simply exercises in applying formulas. Further practice and confidence in applying and extending the material in a chapter can be obtained by filling in details of the derivations and examples, or by developing parallel results for circuits or converters in the same family or category as those explicitly treated in the chapter.

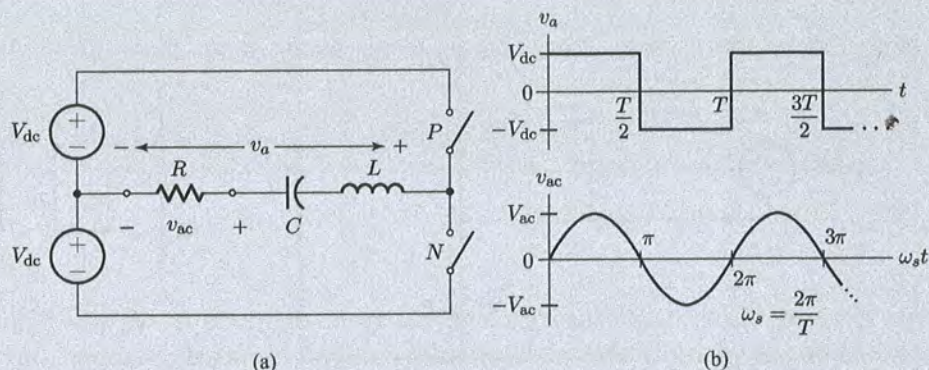


Figure 2.5 (a) A half-bridge dc/ac converter topology employing a series-tuned filter on the ac side. (b) The relationship between the unfiltered voltage, v_a , and the output voltage, v_{ac} , when the filter is tuned to the switching frequency ω_s .

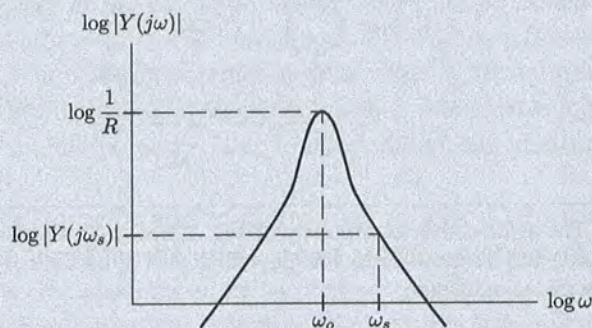


Figure 2.6 Admittance as a function of frequency for the RLC circuit on the ac side of Fig. 2.5(a).

The topology of Fig. 2.5(a) is known as a *series-resonant converter*. It is one member of a family of dc/ac converters called *resonant converters*, treated in more detail in Chapter 10. Designers use them principally to obtain clean, high-frequency ac waveforms, often for induction heating applications. Note that the technique of power control that we just discussed results in a varying output frequency, so any application of this type of control must not require a precise output frequency.

2.3 DC/DC Converters

Used extensively in power supplies for electronic equipment, dc/dc converters control the flow of energy between two dc systems. The dc/dc converter takes the dc output of the ac/dc converter and transforms it to the different dc voltages required by the electronics, 5 V and ± 15 V, for example. These converters are also used in battery-powered equipment and to control the speed of dc motors in many traction applications, such as battery-powered forklifts or trains operating from a dc third rail or catenary. In relatively high-power applications, such as traction, the dc/dc converter is known as a *chopper*.

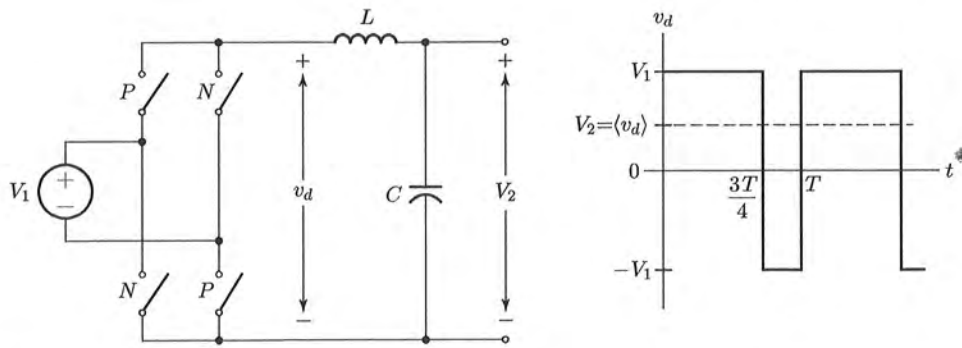


Figure 2.7 Basic dc/dc converter topology. The ratio V_2/V_1 can vary between ± 1 .

2.3.1 Basic Topology

The basic dc/dc converter topology is shown in Fig. 2.7. The striking feature of this circuit is its similarity to the ac/dc converter topology shown in Fig. 2.1(b). In fact, the only difference is the control of switches. In the dc/dc converter, the switches are controlled to produce a voltage v_d that contains a nonzero dc component. This component is then extracted by the low-pass LC filter to produce the output dc voltage V_2 .

As before, we alternately operate the P and N switches. Instead of leaving each set closed for exactly half the cycle, as we did in the ac/dc converter, here we control them to have asymmetric on-times, so that v_d contains a dc component. For example, if we leave the P switches on for three quarters of the switching period T , the voltage v_d shown in Fig. 2.7 results. It has an average value of $V_1/2$. In this circuit we can obtain an average voltage between V_1 and $-V_1$ by adjusting the relative conduction periods of the two sets of switches. The fraction of its switching period during which a switch is on is known as the *duty ratio* D of the switch. Here, the duty ratio of the P switches is 0.75, and that of the N switches is 0.25.

Example 2.3 A Simplified DC/DC Converter Topology

We can simplify the topology of Fig. 2.7 if we require an output voltage of only one polarity, for example, $0 < V_2 < V_1$. The circuit of Fig. 2.8, which contains only two switches, is sufficient to do the job. As we turn these switches on and off sequentially, the voltage waveform they create steps between V_1 and 0.

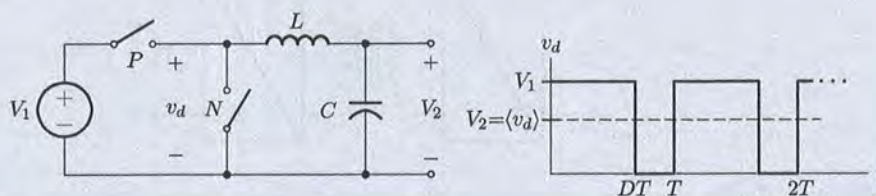


Figure 2.8 A simplified dc/dc converter topology limited to an output voltage of $0 < V_2 < V_1$.

Assume that we leave switch P on for time DT , where T is the switching period and D is the duty ratio of the P switch. We then turn switch P off and switch N on for the rest of the period, that is, for a time $(1 - D)T$. The output voltage V_2 , the average value of v_d , is then DV_1 . We can alter this average value by changing the duty ratio D , but the lower limit of the output voltage is zero.

What did we gain from this simplification? One benefit is the reduction in the number of switches required. Another, which can be shown through a harmonic analysis, is that the ac components of v_d in this circuit are smaller than those in v_d for the bridge circuit of Fig. 2.7. We can therefore use smaller filter elements to achieve the same level of ripple in the waveforms presented to the external system.

Because the frequency of the desired output (dc) of the converter of Fig. 2.7 is much less than the switching frequency ($1/T$), this is a member of a class of converters known as *high-frequency switching converters*. We discuss high-frequency dc/dc converters in Chapter 5. As you will see in Chapter 8, another member of this class is the high-frequency dc/ac converter.

We can also achieve the transformation from dc to dc in quite a different way, using a dc/ac converter to create an ac waveform from the dc at one port, and then an ac/dc converter to transform this ac waveform back to dc at the other port. At first glance this approach appears to be wasteful because it requires so many switches and filter elements. It does, however, give us a point in the circuit where the waveforms are ac. As shown in Fig. 2.9, we can install a transformer at this point to provide electrical isolation and to make use of its turns ratio when the difference between the two external voltages is large. We discuss this type of converter, called an *isolated high-frequency converter*, in Chapter 7.

2.4 AC/AC Converters

An ac/ac converter converts an ac waveform of one amplitude, phase, and/or frequency to another ac waveform with different parameters. One of three distinct topological approaches can be used, depending on system requirements. The first and simplest can be used to change the amplitude parameters of an ac waveform (the rms value, for instance). It is known as an *ac controller*, and functions by simply taking symmetric “bites” out of the input waveform. The second can be used if the output frequency is much lower than the input source frequency, allowing the use of a *cycloconverter*, which approximates the desired output waveform by

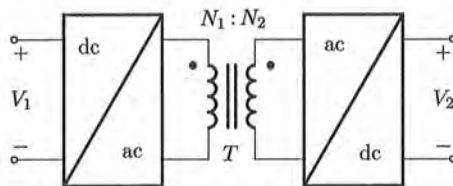


Figure 2.9 A dc/dc converter consisting of two cascaded ac/dc converters. This configuration permits the use of an isolating transformer T , as shown.

5

Pulse-Width-Modulated DC/DC Converters

Pulse-width-modulated switching converters are power circuits in which the semiconductor devices switch between on and off states at a rate that is fast compared to the frequencies of the input and output waveforms. Control of the converter is exercised by varying the ratio of on-time to total switching period of the controlling switches, thereby controlling the width of the pulses applied to the output. This is called *pulse-width modulation* or *PWM* control.

Certain characteristics of these converters differentiate them from the converters described in other chapters of Part I. First, unlike the rectifiers and inverters of Chapters 3 and 4, the semiconductor devices in pulse-width-modulated high-frequency converters do not usually use the reversal of external waveforms to turn themselves off. Second, unlike the resonant converters of Chapter 10, the switching frequency is high enough that the inductors and capacitors in the circuit act as low-pass filters to the switching-frequency components; that is, these components are greatly attenuated in the output.

High-frequency switching converters are used most often as interfaces between dc systems of different voltage levels. Our discussion in this chapter and in Chapters 6 and 7 is therefore in the context of this application. The converters here and in Chapter 7 are known as PWM dc/dc converters, and examples of their use are the power supplies in computers and other electronic equipment, where typically a rectifier is followed by a dc/dc converter.

High-frequency switching converters can also be used to interface between dc and ac systems. Although we look at the unique aspects of this application in Chapter 8, much of what we say about dc/dc converters in this chapter applies to dc/ac converters as well. The reason is that in the dc/ac application, the switching frequency is usually so much higher than the ac-output (or -input) frequency that we can consider the variables at the ac port to be constant over many switching periods. In other words, for times on the order of the switching period, we can view the dc/ac converter as a dc/dc converter.

In this chapter we present the structure and operation of high-frequency switching converters that use both inductors and capacitors to store energy. Chapter 6 discusses another class of dc/dc converters that use only capacitors. Rather than just categorizing a large number of circuits, we develop the fundamental concepts and topological relationships on which all such circuits are based. Again, you will see that the specification of a topology does not simultaneously specify which port is the input and which is the output. Only when we specify the implementation and control of the switches do the ports obtain identities as input or output.

5.1 The DC/DC Converter Topology

The simplest form of a switching dc/dc converter is that of Fig. 5.1. The switch opens and closes at a frequency $1/T$, with the ratio of the on-time to the period defined as D , the *duty ratio*. The resulting load voltage v_2 is a *chopped* version of the input – a series of pulses having an amplitude of V_1 and an average, or dc, value of DV_1 . But this dc value comes with a substantial amount of ripple, which is present not only in the load voltage v_2 but also in the source current i_1 . Few applications can make use of this dc component in the presence of so much distortion. The high frequencies contained in the ripple can cause both conducted and radiated interference with other apparatus, such as computers or communications equipment. Moreover, some loads, such as many electronic circuits, function properly only if operated from a dc power supply with little ripple. Therefore, a dc/dc converter generally has terminal voltages and currents that deviate only minutely and instantaneously from their dc values. This restriction means that we have to modify the elementary topology of Fig. 5.1. We develop these modifications by first exploring the topological constraint imposed by the requirement of nearly constant terminal variables.

The circuit in Fig. 5.2(a) shows a converter connecting two systems whose terminal voltages and currents are dc with the values shown. Whatever is in the box has to produce terminal variables that are free from pulsations or ripple. The difference between input and output voltages, $V_i - V_o = 50$ V, must drop across an element connected between the ports, and the box must contain a shunt element to provide a path for the difference between input and output currents, $I_o - I_i = 5$ A. This minimal connection of elements is shown in Fig. 5.2(b). Note that the power absorbed by the series element is equal to the power supplied by the shunt element, so the box is lossless – but the energy must be transferred from the shunt to the series element. It is difficult to imagine a way to implement this simple two-element topology.

A circuit element that can support both a nonzero average voltage and a nonzero average current without dissipating energy is a switch. Furthermore, we can control the average value of its variables by varying the ratio of its on to off times. The problem with using switches as the shunt and series elements, however, is that instantaneous values of the input current and output voltage will differ from their average values. We can, however, extract their average values by using low-pass filters. By defining the converter to include these filters, we create an interface that produces the desired terminal variables.

Figure 5.3 shows switches as the shunt and series elements in our example converter, connecting the dc input voltage, $V_i = 200$ V, to the dc output current, $I_o = 20$ A. The switches

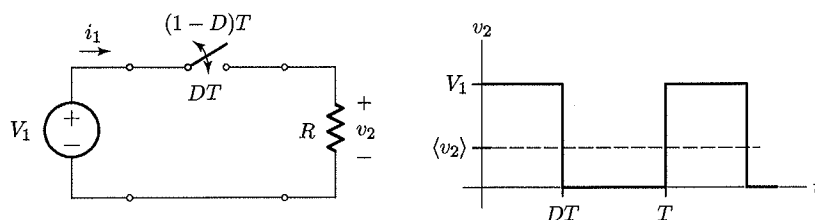


Figure 5.1 The simplest form of a switching dc/dc converter.

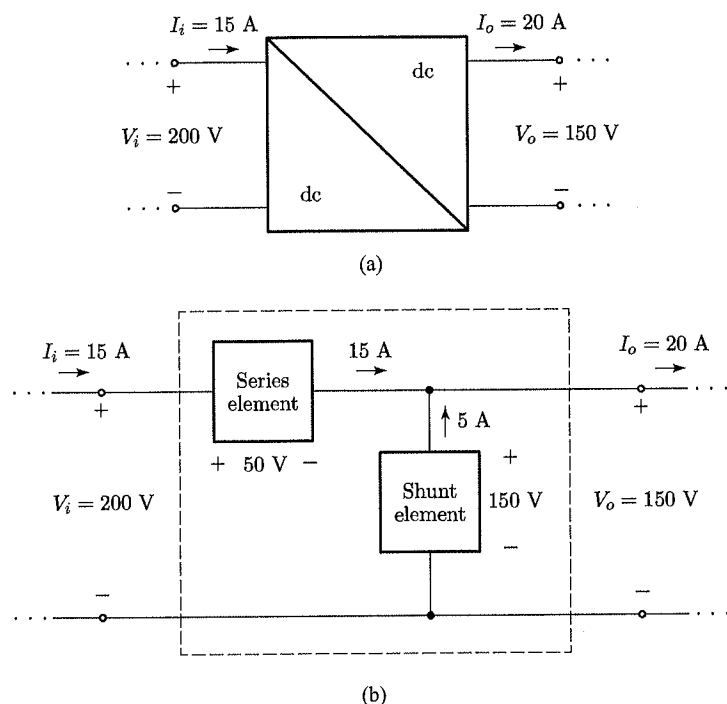


Figure 5.2 (a) A converter connecting two dc systems. (b) The minimal necessary topology for performing the conversion function of (a).

operate at a constant switching frequency $1/T$, and are controlled to be complementary. That is, when one is closed, the other is open (so that a path is always provided for the load current and the input voltage is never shorted). The waveforms of v_{S_1} and v_{S_2} show that switches S_1 and S_2 are on for times DT and $(1 - D)T$, respectively, giving average values of

$$\langle v_{S_1} \rangle = (1 - D)V_i \quad \text{and} \quad \langle v_{S_2} \rangle = DV_i.$$

Because $\langle v_o \rangle = \langle v_{S_2} \rangle = 150$ V, we must control the switches so that $D = 0.75$. The average voltage across the series switch is 50 V, or the difference between the average input and output voltages that we said was necessary when discussing Fig. 5.2. Similarly, we can show that the average current in the shunt switch is the difference between the average input and output currents, as required, or

$$\langle i_{S_2} \rangle = \langle i_i \rangle - I_o = -(1 - D)I_o = -5 \text{ A}. \quad (5.1)$$

We assume that the input voltage contains no ripple. However, the input current contains substantial ripple caused by switching, as shown in Fig. 5.3. Although the output current is ripple-free, the output voltage is not. To obtain the desired ripple-free input current and output voltage, we must insert low-pass filters at the input and output. In their simplest form, these filters consist of a shunt capacitor at the input and a series inductor at the output. By making these elements very large, small values of external network input impedance and output conductance will result in acceptably small ripples in all the terminal variables. The resulting high-frequency

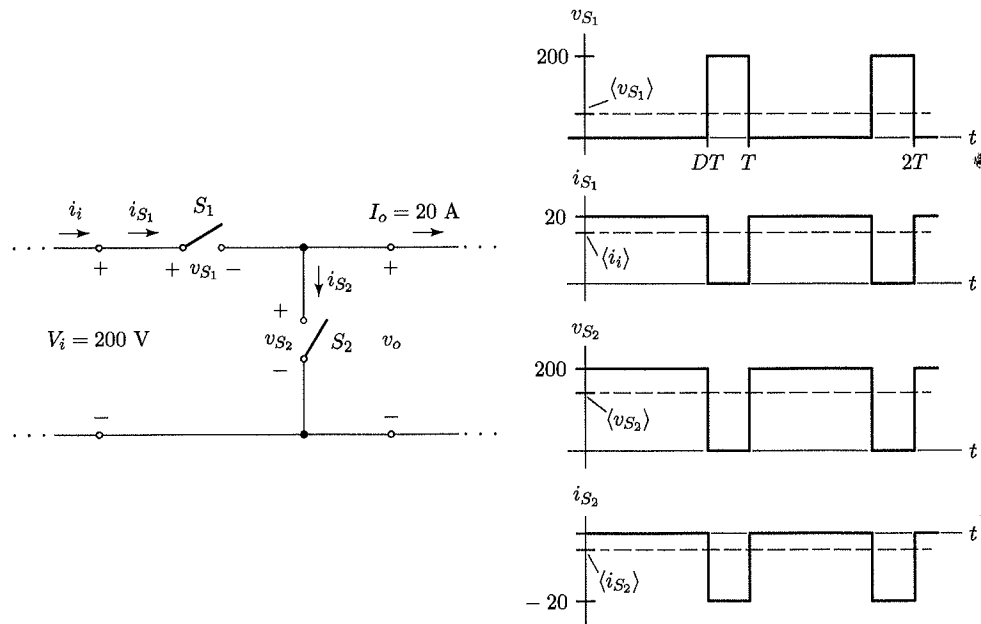


Figure 5.3 Two dc systems connected by a lossless interface, using switches as the shunt and series elements.

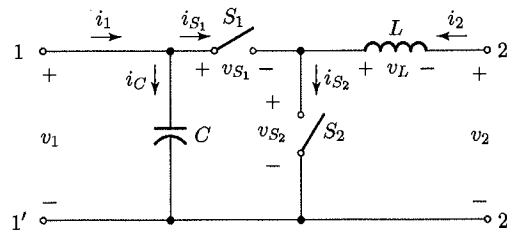


Figure 5.4 Simplest topology for a high-frequency dc/dc converter with low-pass filtering at its ports.

dc/dc converter topology is shown in Fig. 5.4. We have not identified input and output ports, because energy can flow in either direction, depending on how we control the switches.

Example 5.1 Input Filter Effectiveness

Figure 5.5 shows a converter whose external input circuit is modeled by a voltage source V_i behind a resistor R , which might represent wiring resistance or the internal resistance of a battery. We assume that the converter inductor L is large enough to make the output current constant at I_o . What is the peak-peak amplitude of the ripple on the input current i_i ?

If the switches are operating at a fixed duty ratio, D , we can replace the circuit to the right of the capacitor by an equivalent source i_e , as shown in Fig. 5.5(b). Furthermore, being concerned only with the ripple component of i_i , we can replace i_e with its ac component i'_e . The circuit that we analyze is shown in Fig. 5.6(a), where i'_i is the ac component of the input current, and the voltage source has been modeled as an incremental short.

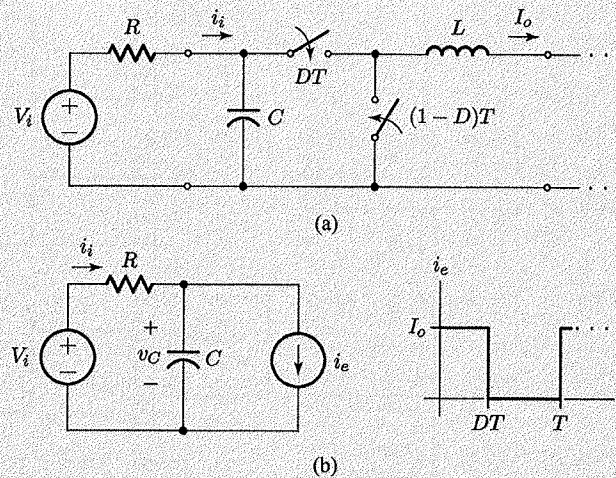


Figure 5.5 (a) Switching converter with a low-pass filter to eliminate ripple from the current flowing through the input circuit. (b) The switches, inductor, and external output network replaced by an equivalent source, i_e .

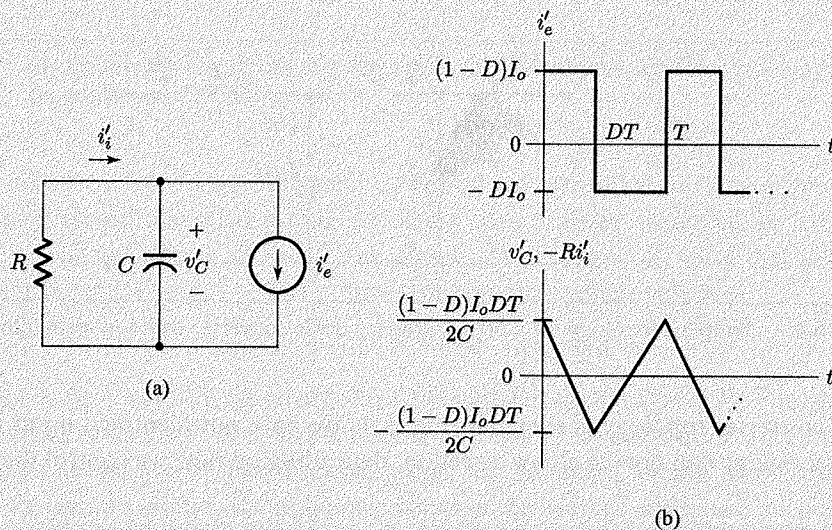


Figure 5.6 (a) Equivalent circuit for the calculation of the ripple component of the input current in the circuit of Fig. 5.4. (b) Branch variables for the circuit of (a), assuming that $RC \gg T$.

If the low-pass filter is effective, that is, $\tau = RC \gg T$, or $R \gg 1/\omega C$, then almost all the ripple current will pass through C . Therefore the capacitor voltage ripple is triangular with a peak-peak amplitude of

$$\Delta v_C = \frac{(1-D)I_o}{C} DT, \quad (5.2)$$

as shown in Fig. 5.6(b). The peak-peak amplitude of the input ripple current will thus be

$$\Delta i_l = \frac{\Delta v_C}{R} = \frac{(1-D)I_oDT}{RC}. \quad (5.3)$$

From (5.3) we see that if $C \rightarrow \infty$ and R is nonzero, then $\Delta i_l \rightarrow 0$.

In many cases, we add additional energy storage elements to the external systems to enhance filtering. For instance, if the external network connected to port 1–1' in Fig. 5.4 has a very low ac impedance, C will have to be very large. However, if we increase the impedance of the external network by placing an inductor in series with it, C can be much smaller. Similarly, if the external impedance at port 2–2' is very high, L will have to be large unless we place a capacitor in parallel with the port to reduce its impedance. These additional elements on each side of the converter make the filters second-order. Third-, fourth-, and even-higher-order filters are used if the need to attenuate the ac components is great enough. These elements do not change the basic function of the original capacitor and inductor, so we do not include them in the topologies of this chapter.

5.1.1 Average Value Analysis

The average values of currents and voltages in a circuit obey Kirchhoff's current and voltage laws (KVL and KCL), providing a straightforward analysis technique for determining the relationship among average variables in converters. To illustrate, consider Fig. 5.4. Writing KVL for the average voltages around the loop formed by v_1 , S_1 , v_L , and v_2 gives

$$-\langle v_1 \rangle + \langle v_{S_1} \rangle + \langle v_L \rangle + \langle v_2 \rangle = 0.$$

Recognizing that in a periodic steady state the average value of the voltage across an inductor is zero,

$$-\langle v_1 \rangle + \langle v_{S_1} \rangle + \langle v_2 \rangle = 0.$$

Referring to Fig. 5.3, with $\langle v_1 \rangle = V_1$ and $\langle v_2 \rangle = V_2$,

$$\langle v_{S_1} \rangle = (1-D)V_1$$

so

$$\begin{aligned} -V_1 + (1-D)V_1 + V_2 &= 0 \\ V_2 &= DV_1. \end{aligned}$$

A similar analysis can be done for the currents, recognizing that the average value of the current through a capacitor in a periodic steady state is zero. Since the basic circuit has three nodes, two KCL equations are required. We choose the nodes between S_1 and C , and S_1 and S_2 :

$$\begin{aligned} \langle i_l \rangle - \langle i_{S_1} \rangle - \langle i_C \rangle &= 0, \\ \langle i_{S_1} \rangle - \langle i_{S_2} \rangle + \langle i_2 \rangle &= 0, \\ \langle i_l \rangle &= I_1 = \langle i_{S_1} \rangle, \\ \langle i_2 \rangle &= I_2 = \langle i_{S_2} \rangle - I_1, \\ \langle i_{S_2} \rangle &= I_2(1-D), \\ I_2 &= -\frac{1}{D}I_1. \end{aligned}$$

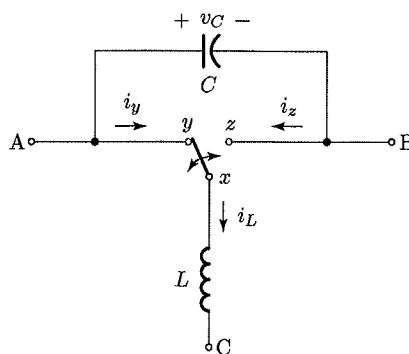


Figure 5.7 The canonical switching cell, the basic building block of many high-frequency dc/dc converters.

5.2 The Canonical Switching Cell

We refer to the minimal practical dc/dc converter topology developed in Section 5.1 and shown in Fig. 5.4 as the *canonical switching cell*.[†] We redraw it symmetrically in Fig. 5.7, using a single-pole double-throw switch, which satisfies the condition that the switches be neither on nor off simultaneously. The canonical cell is the basic building block for many high-frequency switching converters. The distinctions between these converters arise mainly from the way in which the external systems are connected to the cell. These connections determine both the input/output conversion ratios and the levels of current and voltage stress imposed on the cell's components.

We may connect the three-terminal canonical cell of Fig. 5.7 between two dc systems in one of three ways. The connections differ with respect to the cell node that we make common to both ports. Because of the symmetry of the cell, however, the use of node A or B as a common node results in indistinguishable topologies. Thus there are only two unique connections to consider. If node A or B is common, the resulting topology is that of Fig. 5.8(a), which is identical to Fig. 5.4. Figure 5.8(b) shows the topology that results from the use of node C as the common terminal.

In the connection of Fig. 5.8(a), there is a direct dc path between the input and output ports when switch S_{xy} is conducting. For this reason we call this a *direct converter* connection. In the connection of Fig. 5.8(b), there is no dc path between the ports in any switch state. Therefore we call this an *indirect converter* connection. The use of the names *direct* and *indirect* highlights the essential difference among various specific high-frequency dc/dc converter circuits. The presence of transformers or additional filter elements does not negate this distinction.

5.3 Direct Converter

We now consider the direct converter of Fig. 5.8(a) in detail. We are particularly interested in the dc conversion ratio and the implementation of the switches. Note that we cannot define the

[†] E. Landsman, "A Unifying Derivation of Switching DC-DC Converter Topologies," *Power Electronics Specialists Conference*, IEEE, 1979, pp. 239-243.

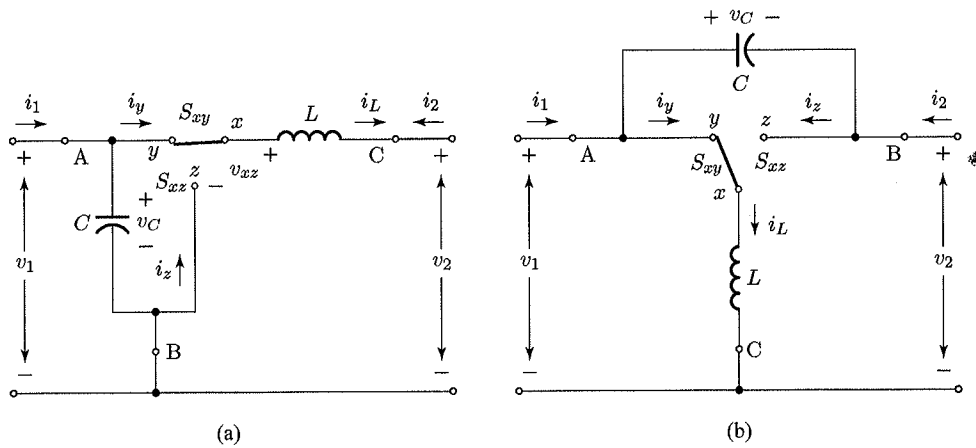


Figure 5.8 The two unique connections of the canonical cell between two dc systems: (a) a direct converter; (b) an indirect converter.

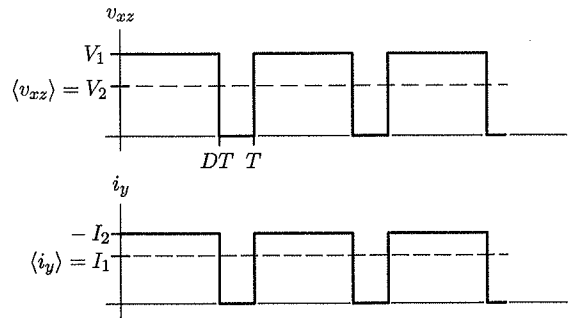


Figure 5.9 Waveforms of the switch variables v_{xz} and i_y for the direct converter of Fig. 5.8(a).

ports as input or output until we specify the semiconductor devices to be used as the switches, and possibly the external networks and controls as well.

5.3.1 DC Conversion Ratio of the Direct Converter

In the discussion that follows we assume that L and C are large enough to eliminate switching-frequency components from the terminal variables v_1 , i_1 and v_2 , i_2 . Therefore we specify the terminal variables using uppercase letters to indicate dc quantities. For example, $i_1(t) = \langle i_1(t) \rangle = I_1$.

The values of the conversion ratios, V_2/V_1 and $-I_1/I_2$, depend on the duty ratio of the switches, and we can determine them as above using average KVL and KCL relations, or from the waveforms of v_{xz} and i_y shown in Fig. 5.9. Here we express the conversion ratios in terms of the duty ratio D of the series switch. As the inductor cannot support an average voltage, $V_2 = \langle v_{xz} \rangle = DV_1$. And as the capacitor cannot carry an average current, $I_1 = \langle i_y \rangle = D(-I_2)$. Solving these constraints for the conversion ratios gives

$$\frac{V_2}{V_1} = D \quad \text{and} \quad \frac{I_2}{I_1} = -\frac{1}{D}. \quad (5.4)$$

From (5.4) we can show that the average terminal powers are equal and opposite, that is, $V_1 I_1 = -V_2 I_2$. We expected this result because the ideal canonical cell is lossless. Note that we still have not specified the direction of power flow; the product $V_1 I_1$ is not necessarily positive, and $V_2 I_2$ is not necessarily negative.

So far, we have assumed that the converter is switching at a constant frequency, $1/T$, as we vary D to control the conversion ratio. Known as *constant frequency control*, it is only one of the ways by which we can operate the high-frequency converter. We can also control the conversion ratio by turning the series switch on for a fixed time and varying its off time, called *constant on-time control*. Alternatively, we could hold the off-time of the series switch constant and vary its on time, called *constant off-time control*. Constant on-time and off-time control both result in changes in the switching frequency as we vary the conversion ratio. In all three cases, the conversion ratio depends only on the percentage of the period the switches are in one state or the other, which represents a form of pulse-width modulation.

Because of the relationship between the two external voltages given by (5.4), the direct converter has come to be known as either a *buck* (or *down*) *converter* or a *boost* (or *up*) *converter*, depending on the direction of power; a buck converter has power sourced from the high-voltage side (V_1), a boost from the low-voltage side (V_2). These are the common names when we discuss these circuits, but they do not describe two separate converters. As we show in Section 5.3.2, the desired directions of power flow only affect how we implement the switches.

Note that these names are based on how the *voltages* of the two external systems are related. Although in many applications the voltage variable is of prime concern, in some situations the current variable is more important. In these cases, the circuit we called a voltage buck converter is actually a current boost converter, and the voltage boost converter is a current buck converter. Thus the popular names of these circuits can be misleading unless you understand the commonality of topology and function implied by the single (direct converter) connection of the canonical cell.

5.3.2 Implementation of Switches and Power Flow

To determine the direction of power in a dc/dc converter, we generally need to specify switch implementations, external networks, and the controls.

When we replace the generalized switches we have used so far with semiconductor devices, we also constrain the directions of their current and voltage, depending on the device. The networks external to the converter terminals may or may not be able to source or sink energy. And in a situation where both the switches and networks allow bidirectional power flow, how the switches are controlled determines in which direction it flows. Here we treat the simplest case in which the switch implementation is sufficient to determine the direction of power flow.

The buck converter To illustrate how switch type can determine the direction of power flow, let's consider the direct converter of Fig. 5.8(a) operating in the buck mode. That is, power is flowing from the higher-voltage side (V_1) to the lower-voltage side (V_2). In this case, $V_1 > 0$,

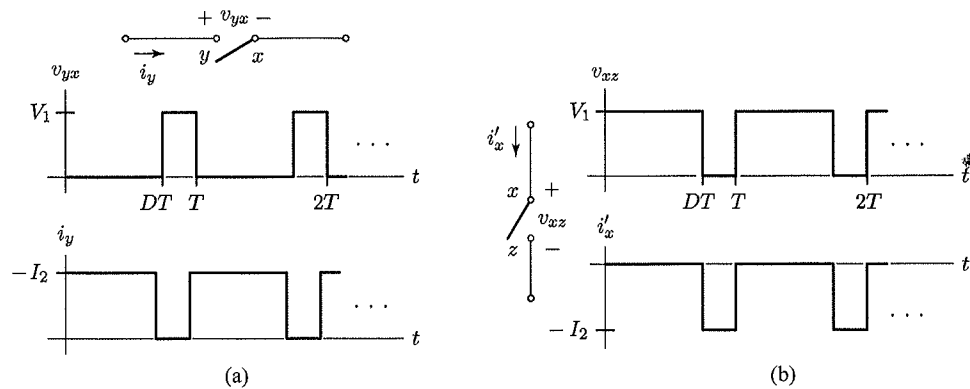


Figure 5.10 Voltage and current waveforms for the series and shunt switches in the direct converter of Fig. 5.8(a) when it is operating in the buck mode: (a) series switch; (b) shunt switch.

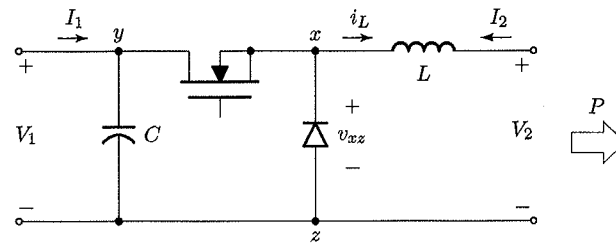


Figure 5.11 Switch implementation for the direct buck converter.

$I_1 > 0$, $V_2 > 0$, $I_2 < 0$; the switch voltages and currents are shown in Fig. 5.10, where the shunt switch component of i_x is designated as i'_x ($i_x = -i_z$ in Fig. 5.8(a)).

The current and voltage of the series switch, i_y and v_{yx} , are both positive, conditions satisfied by a transistor. The current and voltage of the shunt switch, i_x and v_{xz} , are of opposite polarities, as are those of a diode. It may not be immediately apparent that an uncontrollable device is suitable for the shunt switch. However, the diode is forced on and off by the series switch. When the switch closes, V_1 reverse biases the diode, and when it opens, the continuity of current in L forces the diode to conduct. The function and operation of the shunt switch is identical to that of the freewheeling diode in the half-wave rectifier of Fig. 3.5.

The buck converter – with the series and shunt switches replaced by a transistor and diode, respectively – is shown in Fig. 5.11. Although the integral diode in the MOSFET (known as the *body diode*) would permit reverse current in the device, the shunt diode, being uncontrollable, provides no means for reversing this current.

Example 5.2 A Buck Converter with Common Positives

At times we need to connect two dc systems so that their positive terminals are common. In this example, we design a buck converter for such an application.

Again, we consider the general direct converter of Fig. 5.8(a). The buck constraint requires that $v_1 i_1 > 0$ and $v_2 i_2 < 0$. The common positive requirement implies $v_1 < 0$ and $v_2 < 0$. Therefore $i_1 < 0$ and $i_2 > 0$. Referring to Fig. 5.10 and noting the polarity for the terminal variables in this application, you can see that the series switch is still a transistor and that the shunt switch is still a diode. However, since their directions of conduction are the reverse of those for the devices in the converter of Fig. 5.11, they are connected in the opposite sense. The resulting converter with common positives is shown in Fig. 5.12.

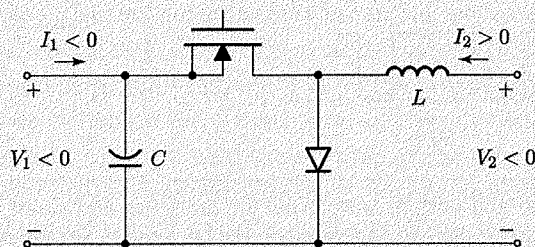


Figure 5.12 A buck converter configured so that the positive terminals of the input and output ports are common.

We can also derive the converter of Fig. 5.12 by simply transforming the circuit of Fig. 5.11. We do so by noting that we can change the position of the inductor and the series switch to the lower branches in their respective loops without changing the operation of the circuit. The resulting circuit is shown in Fig. 5.13. Now, by simply flipping the circuit over, we have the result shown in Fig. 5.12. An advantage of the circuit of Fig. 5.13 is that the gate of the MOSFET is referenced to the negative terminal of the source V_1 , eliminating the need for level shifting of the gate drive.

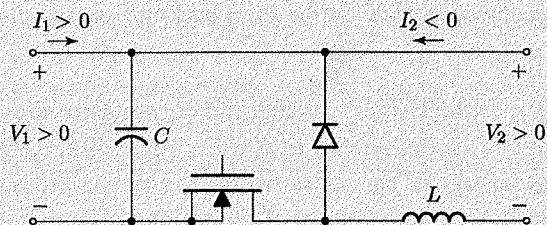


Figure 5.13 A buck converter with common positives derived by changing the positions of the inductor and series switch in the circuit of Fig. 5.12.

The boost converter We now consider the switch implementation necessary to create a boost converter. We choose the switches so that the power in the direct converter of Fig. 5.8(a) flows from the lower-voltage side (V_2) to the higher-voltage side (V_1). We do so by simply reversing the polarities of either the terminal currents or the terminal voltages, relative to the buck converter we just discussed.

If we choose to reverse the polarity of the terminal currents, leaving positive voltages V_1 and V_2 , the switch current polarities of Fig. 5.10 are reversed. In this case, the shunt switch becomes a transistor and the series switch can be a diode. The result is the direct boost converter shown in Fig. 5.14. Note that the controllable switch is now the shunt switch instead of the series switch.

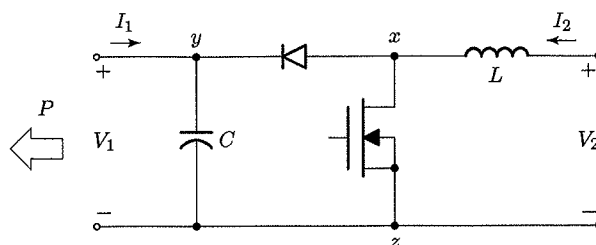


Figure 5.14 Switch implementation for the direct boost converter.

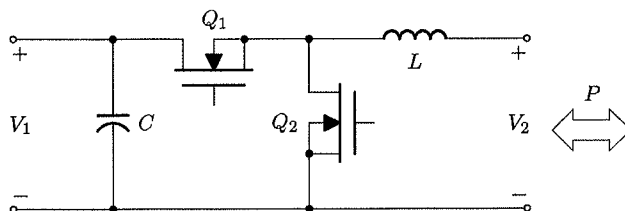


Figure 5.15 A direct converter configured with MOSFETs to provide bilateral terminal currents.

Direct converter with bilateral energy flow Bilateral power can be implemented in the direct converter by replacing the diode with another MOSFET, as shown in Fig. 5.15. For operation in the buck mode, that is, power flowing from V_1 to V_2 , Q_1 controls D , and the body diode of Q_2 can function as the shunt element. However, when on, a power MOSFET can conduct negative current, which makes it possible to synchronously control Q_2 so its current is conducted through the channel, bypassing the body diode. The benefit of this is that the source-drain drop is lower than it would be if the body diode were conducting, providing improved efficiency. For power flowing from V_2 to V_1 , that is, operation in the boost mode, Q_2 controls D and Q_1 is controlled synchronously. Using power MOSFETs in this manner is the basis of *synchronous rectifiers*.

Example 5.3 Switch Implementation for Bilateral Terminal Voltages

Some applications of dc/dc converters require power to flow in either direction. Two examples are a dc motor that is actively braked, and a magnetic resonance imaging (MRI) machine magnet, which must create a precisely controlled time-varying magnetic field. In both cases the terminal voltages change sign but the current polarities remain unchanged (due to the inability to create a step change in a magnetic field absent a voltage impulse) in order to remove or introduce energy into the magnetic field system. That is, with reference to Fig. 4.5(b), the converters operate in quadrants I and IV as seen from their source terminals (V_1 , I_1). How do we implement the switches in a direct converter designed for such applications?

We again consider the direct converter switch variables of Fig. 5.10, noting that if the terminal currents do not change sign, the switches will still only be required to conduct in one direction. However, because the direction of power is reversed by inverting the terminal voltages, the switches must now be capable of blocking both voltage polarities. A switch made up of a diode in series with a transistor has the required characteristic, and the resulting bilateral direct converter is shown in Fig. 5.16. Note that we must now control both switches explicitly and make sure that they are not off simultaneously. However, when energy flows from left to right, that is, when we are supplying energy to the magnetic system and operating in

quadrant I ($V_1 > 0$, $I_1 > 0$, $V_2 > 0$, $I_2 < 0$), we can hold Q_2 on continuously, because D_2 will do the switching. Similarly, when discharging the magnet or motor, power flows from right to left ($V_1 < 0$, $I_1 > 0$, $V_2 < 0$, $I_2 < 0$) we can hold Q_1 on continuously.

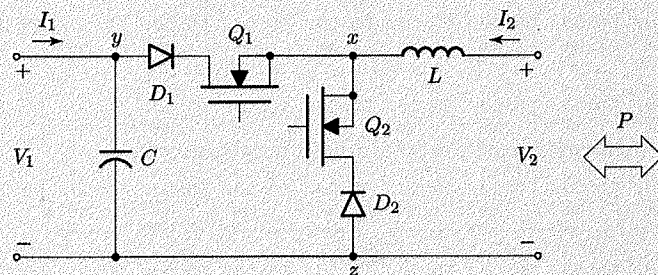


Figure 5.16 Switch implementation in the direct converter to provide bidirectional power flow with voltage reversal (operation in quadrants I and IV).

Note that the diodes in the circuit of Fig. 5.16 are on the drain sides of Q_1 and Q_2 . Therefore we can connect the sources of Q_1 and Q_2 . This approach simplifies the level-shifting and gate drive circuit by referencing the drives for both transistors to the same point.

5.3.3 About the Duty Ratio

In Section 5.1 we defined the duty ratio for the direct converter as the fraction of the switching period during which the series switch is on. Actually, common practice is to define the duty ratio as the fraction of the period during which the *controllable* switch (the transistor) conducts, regardless of whether it is in the series or shunt position. When we use this definition for D , (5.4), written for the boost converter of Fig. 5.14, becomes

$$\frac{V_2}{V_1} = 1 - D \quad \text{and} \quad \frac{I_2}{I_1} = -\frac{1}{1 - D}. \quad (5.5)$$

We sometimes refer to the quantity $1 - D$ as D' . We also commonly express the conversion ratio for a dc/dc converter in the form of the output over the input. Therefore we can write the voltage conversion ratio for the boost converter as

$$\frac{V_{\text{out}}}{V_{\text{in}}} = \frac{1}{1 - D} = \frac{1}{D'}. \quad (5.6)$$

Finally, the conventional conversion ratio expressions are ambiguous when power can flow in either direction, because we cannot label either of the external systems as the input or the output, and both switches are controllable. In this case we normally use the terms and expressions that are appropriate for the direction power is flowing at the time.

5.4 Indirect Converter

The indirect converter connection of the canonical cell is shown in Fig. 5.8(b). As you will see in this section, the indirect converter exhibits operating characteristics that are very different from those of the direct converter just discussed.

5.4.1 The DC Conversion Ratio of the Indirect Converter

An important distinction between direct and indirect connections of the canonical cell is the relative polarity of their actual (measured) terminal voltages and currents, as defined in Fig. 5.8. The voltages are of the same polarity in the direct connection, but they are of opposite polarities in the indirect connection. In contrast, the terminal currents are of opposite signs in the direct converter, but are of the same sign in the indirect converter.

We can derive the polarity relationships between the terminal variables of the indirect converter of Fig. 5.8(b) by considering the inductor voltage and the capacitor current. The voltage across the inductor is V_1 when switch S_{xy} is closed and V_2 when S_{xz} is closed, so one of these two voltages must be positive and the other negative to satisfy the condition that the average inductor voltage be zero. Similarly, the capacitor current is $-I_2$ when S_{xy} is closed and I_1 when S_{xz} is closed, so both I_1 and I_2 must be of the same polarity to satisfy the condition that the average capacitor current be zero.

If we leave switch S_{xz} on for a long time, both V_2 and I_1 become zero, regardless of the values of V_1 and I_2 . Therefore both the ratios V_2/V_1 and I_1/I_2 become zero. However, if we leave switch S_{xy} on for a long time, these ratios become minus and plus infinity, respectively. As we inferred for the direct converter, when the two switches are alternating, these ratios will be somewhere between their two extremes. For the indirect converter, the magnitude of the conversion ratio ranges from zero to infinity, compared with zero to one or one to infinity for the direct converter.

The simplest way to determine the exact dependence of the conversion ratios on the duty ratio is to set the average inductor voltage to zero by balancing its positive and negative volt-time integrals. If we define D to be the fraction of the switching period T during which S_{xy} is closed, the average inductor voltage is zero when

$$V_1 D T = -V_2 (1 - D) T,$$

which gives a voltage conversion ratio of

$$\frac{V_2}{V_1} = -\frac{D}{1 - D}. \quad (5.7)$$

Similarly, we can determine the current conversion ratio by setting the average capacitor current to zero, that is,

$$I_2 D T = I_1 (1 - D) T,$$

which gives a current conversion ratio of

$$\frac{I_2}{I_1} = \frac{1 - D}{D}. \quad (5.8)$$

We can also obtain (5.8) from (5.7) by setting the net power into the converter to zero.

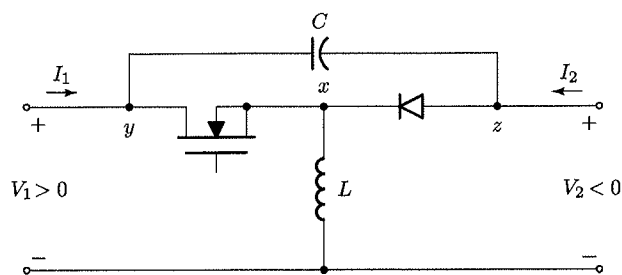


Figure 5.17 Switch implementation for the indirect converter with power flowing from left to right and $V_1 > 0$ (implying I_1 and I_2 are also > 0).

We commonly refer to the indirect converter as a *buck/boost* converter, because the output voltage (or respectively current) can be higher or lower than the input, for either direction of power.

5.4.2 Implementation of Switches

If we consider the switch voltages and currents as we did for the direct converter, we can determine appropriate semiconductor devices to use for the indirect converter under different conditions. Again, we must specify the direction(s) of power flow and the polarity of one of the terminal voltages or currents. For power flowing left to right in the circuit of Fig. 5.8(b) (there is no higher or lower voltage side in the indirect converter) and $V_1 > 0$, switch S_{yx} conducts positive current and blocks positive voltage (note that $V_2 < 0$ for the assumed conditions on power and V_1). Therefore this switch can be a transistor. Switch S_{zx} , on the other hand, conducts positive current but blocks negative voltage, which are conditions met by a diode. The resulting indirect converter circuit is shown in Fig. 5.17.

Example 5.4 Indirect Converter with Power Flowing Right to Left

One way to make an indirect converter with power flowing from right to left in the circuit of Fig. 5.8(b) is to take the converter of Fig. 5.17 and flip it about the inductor. In this case, the voltage at the left terminal pair becomes negative. If the requirement were that this voltage be positive, we would invert the voltage and current for both switches and turn the diode and transistor around, as shown in Fig. 5.18.

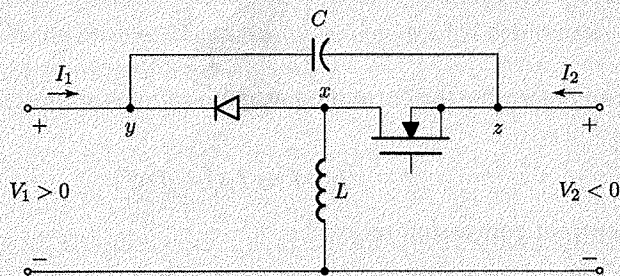


Figure 5.18 Indirect converter switch implementation for power flowing from right to left and $V_1 > 0$.

5.4.3 A Variation of the Basic Indirect Converter Topology

There is an important variation of the basic indirect converter topology. This variation leaves the conversion ratios given by (5.7) and (5.8) unchanged, and the voltage and current waveforms at the semiconductor devices are unaffected. The only thing that we change is deployment of the capacitor to filter the high-frequency switch currents.

The capacitor and inductor in the indirect converter of Figs. 5.17 and 5.18 perform their high-frequency filtering functions in the following ways. The ac components of the switch currents must circulate through C to prevent their appearing at the terminals. But note that because we can model L as an open circuit at the switching frequency, the series connection of the external networks is in parallel with C . Therefore the high-frequency impedance of C must be much smaller than the sum of the high-frequency impedances of the external networks. Similarly, the ac components of the switch voltages must appear across L if they are not to appear at the terminals. In this case, we model the capacitor as a short circuit at the switching frequency, which results in the parallel connection of the two external networks appearing in series with L . Thus the high-frequency impedance of L must be much larger than the parallel combination of the high-frequency impedances of the external networks. (We derive these conditions in more detail in Section 5.6.)

The implications of these conditions are: if the capacitor is to be of reasonable size, at least one of the external networks must have a high ac impedance; if the inductor is to be small, at least one of the external networks needs to have a low ac impedance. In the ideal case, one of the external systems has a high ac impedance and the other a low one.

Unfortunately, both external systems often have a large impedance, particularly when the switching frequency is high, and long (inductive) connection wires are used. Under these conditions the inductor may be unreasonably large. However, if we modify the topology by placing an additional capacitor across one of the two ports, as shown in Fig. 5.19(a), we make one of the external systems appear to have a low ac impedance. Note that the purpose of this capacitor C_1 is to allow most of the ac switch voltage to drop across a reasonably sized inductor. This capacitor does not help the original capacitor C_{12} filter the ac switch current. This switch current continues to flow through C_{12} , because one of the two external systems still has a large ac impedance.

At the switching frequency, the three nodes of the canonical cell are shorted together by the two capacitors of Fig. 5.19(a). We can achieve the same result with the arrangement of capacitors shown in Fig. 5.19(b). If we treat the inductor as an open circuit at the switching frequency, we can see that both capacitors now help filter the ac switch currents. Each carries the full ac switch current if its ac impedance is small compared to the impedance of its corresponding external system. However, across each appears a dc voltage equal only to that at its port, rather than the sum of the terminal voltages, which is what the capacitor in the indirect topology of Fig. 5.16 must withstand. Each of the two capacitors is therefore smaller than the original. In fact, under certain assumptions, we can show the sum of these two capacitors' peak energies to be equal to that of the original capacitor (see Section 5.6.5). Using the names *up/down* or *buck/boost* to identify a converter circuit usually implies the circuit variation of Fig. 5.19(b). In practice the configuration of Fig. 5.19(b) is preferred to that of Fig. 5.18(a) because it eliminates capacitive coupling between the two converter ports.

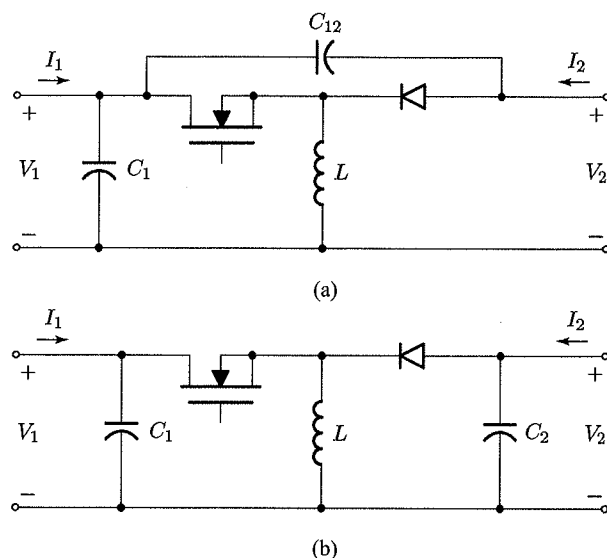


Figure 5.19 Variations of the basic indirect converter topology to provide ac terminal voltage filtering for two external networks having high ac impedances. (a) Addition of capacitor C_1 to the basic indirect topology. (b) Alternative placement for the capacitors in (a). This circuit is commonly referred to as the *buck/boost* converter.

5.5 Other PWM DC/DC Converters

There are numerous other PWM dc/dc converter topologies and variants. These can be arrived at through circuit transformations on canonical cell converters (e.g., through topological duality, addition and rearrangement of components, and addition of circuit blocks); as realizations of the polarity of other, more complex, canonical cells; through intuition; and even by exhaustive search. In this section we introduce two of the more useful and widely applied such topologies. We illustrate both techniques that can be used to synthesize a topology suitable for a particular goal, and methods that can be used to analyze any such topology.

5.5.1 Ćuk Converter

Consider further the canonical cell indirect converter of Fig. 5.17. When both external systems have a low ac impedance, an inconveniently large capacitor must be used to remove ac ripple from the terminal currents. We can, however, increase the impedance of one of the external networks by placing an inductance in series with it, such as L_1 in Fig. 5.20(a). This approach results in a return to the ideal situation in which an indirect converter interfaces one high- and one low-impedance dc network. Although we can now reduce the size of the cell capacitor, this additional inductor has no influence on the size of the original cell inductor L_{12} . The parallel impedance of the two external systems is still low because at least one of them has a low ac impedance, so most of the ac switch voltage will still appear across L_{12} . So far we have simply traded off capacitor size for an additional inductor.

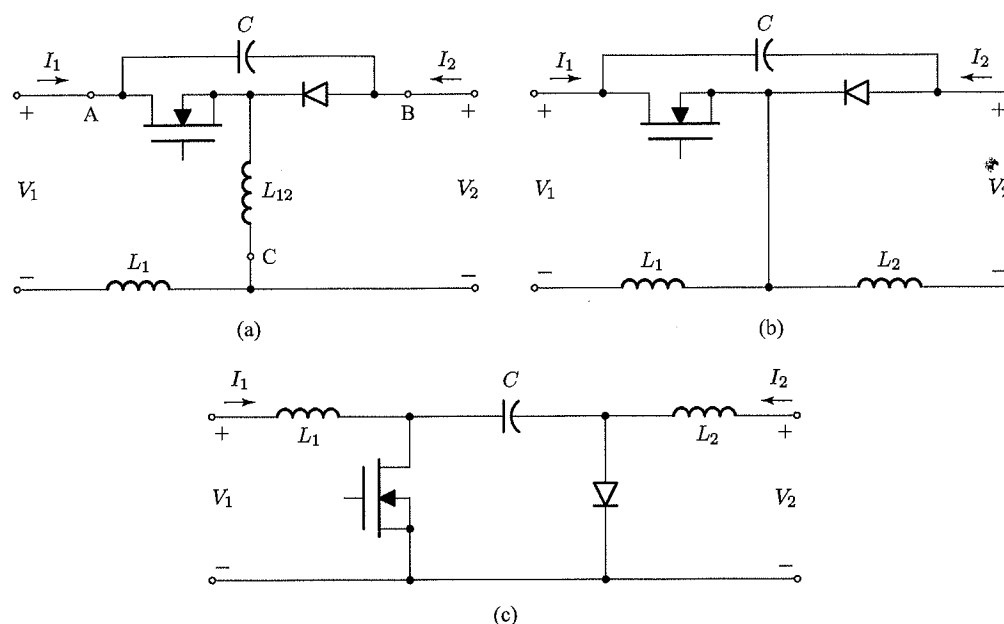


Figure 5.20 Indirect converter topology variations designed to provide ac terminal current filtering for two external networks having low ac impedances. (a) Addition of inductor L_1 to the basic indirect converter topology. (b) Alternative placement for the inductor in (a). (c) The circuit of (b) redrawn in the form commonly known as the Ćuk converter.

Of the three branches stemming from node C in Fig. 5.20(a), two have inductors that, if they have been properly sized to eliminate ac voltage ripple from the converter terminals, we can model as open circuits at the switching frequency. Therefore the ac current flowing through the third branch is also zero. We can achieve the same result if we arrange the two inductors as L_1 and L_2 , shown in Fig. 5.20(b). Now both inductors contribute to filtering the ac switch voltage. Each must have an ac impedance that is large compared to the impedance of its corresponding external system. But each carries only the terminal current rather than the sum of the terminal currents that L_{12} must carry. Therefore each inductor is smaller than L_{12} . For the same ripple currents, the sum of the peak stored energies in these two inductors is equal to the peak energy in L_{12} .

Figure 5.20(b) shows the inductors connected in series with the negative terminals of the converter; however, positioning them in series with the positive terminals, as shown in Fig. 5.20(c), is more practical. This arrangement changes nothing functionally, but it results in a circuit node common to the input, output, and source of the transistor, which is both convenient and often required. This variant of the indirect converter is commonly known as the Ćuk converter. It is worth noting that one can also arrive at the Ćuk converter by applying topological duality techniques to the indirect converter of Fig. 5.19(b).

The Ćuk converter has some attractive features, starting with the desirable filtering properties described above. In addition, it employs a switch referenced to circuit common, which can greatly simplify level shifting, gate drive, and sensing. Moreover, the two inductors of the Ćuk

converter can sometimes be wound on the same core and/or coupled magnetically to improve filtering of input or output waveforms (at the expense of design complexity and manufacturing requirements). In part because of these features, the Ćuk converter has found use despite having a higher magnetic component count as compared to the conventional buck/boost converter.

5.5.2 SEPIC Converter

Here we introduce another widely used indirect converter topology, the “Single-Ended Primary Inductor Converter” known as the SEPIC. It is essentially a cascade of a boost and buck/boost converter using common switching elements. If we take the direct boost converter of Fig. 5.14 and insert a voltage source equal to the input source value V_i between the switch and the diode, as shown in Fig. 5.21, the conversion ratio can be calculated by setting the average value of the inductor voltage v_{L_1} to zero. The result is

$$V_o = V_i \left(\frac{D}{1-D} \right). \quad (5.9)$$

That is, we now have the conversion ratio characteristics of a buck/boost converter but without voltage inversion. We would like to replace the inserted source with a capacitor, but as the circuit is presently configured, the diode current, which has a dc component, passes through the source. Simply replacing the source with a capacitor is prevented by the presence of this dc current. However, if we can provide an appropriate shunt component between the capacitor and the diode to carry the diode's dc current component, a capacitor would function as did the source. We note that the average voltage between the source-diode node and ground, $\langle v_x \rangle$, is zero, so this shunt element must carry a dc current without producing a dc voltage. These two requirements are satisfied by an inductor. The resulting circuit is shown in Fig. 5.22.

We now calculate the conversion ratio of our new circuit to see if our replacements for the source do what we expected. Recognizing that the average inductor voltages in periodic steady state must be zero, we get $v_C = V_i$. Since the average voltage across L_2 over the cycle must be zero in periodic steady state,

$$DV_i + (1-D)(-V_o) = 0, \quad (5.10)$$

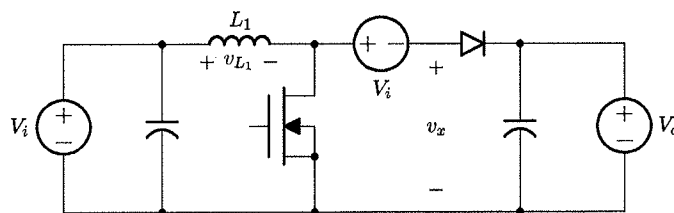


Figure 5.21 The boost converter of Fig. 5.14 with a voltage source of value V_i inserted between the diode and the switch.

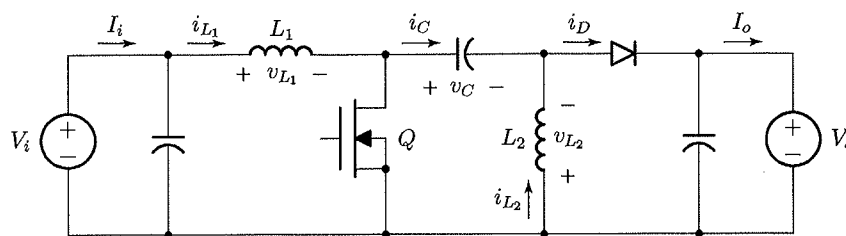


Figure 5.22 The SEPIC converter.

which yields the result sought above:

$$V_o = \frac{D}{(1-D)} V_i. \quad (5.11)$$

The SEPIC converter, which we developed from the direct connection of the canonical cell, is an indirect converter. Its primary advantages over the buck/boost converter of Fig. 5.19(b) are that the input and output voltages are not inverted, and that the MOSFET is referenced to ground. Its disadvantages are that since all the energy is transferred through the coupling capacitor, this capacitor is quite large and must be rated to carry high currents. Also, the circuit's four energy storage elements make control of transient deviations from periodic operation more of a challenge. (As we'll see in Part II, the transfer functions needed for control design will have four poles.)

There are, of course, other means to achieve a non-inverting buck/boost function, such as by cascading a buck and a boost stage to form a *four-switch buck/boost* converter, as explored in Problem 5.14. Such an approach requires more switches (not all ground referenced) than a SEPIC converter, but requires only a single magnetic component. Part of the art of power electronics design is selecting a topology (from among myriad possibilities) that best meets the design requirements in a particular application. It is thus important to be able to both analyze and synthesize designs to meet a particular goal.

5.6 Choice of Capacitor and Inductor Values

So far in our discussion of dc/dc converters, we have assumed that the high-frequency filter elements (e.g., L and C of the canonical cell) are "large enough" to reduce the switching-frequency ripple in the terminal variables to an acceptable level. We now consider the specification of these components in more detail, assuming a second-order filter, that is, a single capacitor and inductor. This subject is important because the physical size of a converter is strongly influenced by the size of its energy storage elements. Furthermore, comparisons among different converter circuits are often based on ripple amplitude for a given amount of energy storage. In this section we discuss the source of ripple in terminal currents and voltages, develop models for analyzing ripple for various converter topologies, and specify the relative sizes of L and C for direct and indirect converters.

In a typical power circuit the filter elements are made large enough to limit ac ripple in the terminal variables to small values. As a result, the capacitor is nearly an ac short circuit and the inductor nearly an ac open circuit. In what follows, therefore, we make the simplifying assumptions that the inductor is an ac open circuit when discussing ripple *current*, and the capacitor is an ac short circuit when discussing ripple *voltage*. These assumptions permit us to make straightforward first-order estimates of the ripple in the terminal variables.

5.6.1 Ripple-Frequency Model for the Direct Converter

We first consider the source and distribution of ripple current in the direct converter of Fig. 5.8(a), repeated here as Fig. 5.23(a) for convenience. With L assumed to be an open circuit at the ripple frequencies, there is no ripple in i_2 , and the ac components of i_y and i_z must be equal and opposite. We designate the ac component of i_y as i'_y . For the purpose of calculating the ripple component of i_1 , denoted by i'_1 , we now replace the circuit to the right of C in Fig. 5.23(a) with an equivalent current source of value i'_y . The resulting ripple-current model is shown in Fig. 5.23(b). Note that the ripple current i'_1 depends on how i'_y divides between C and the impedance of the external network, Z_1 .

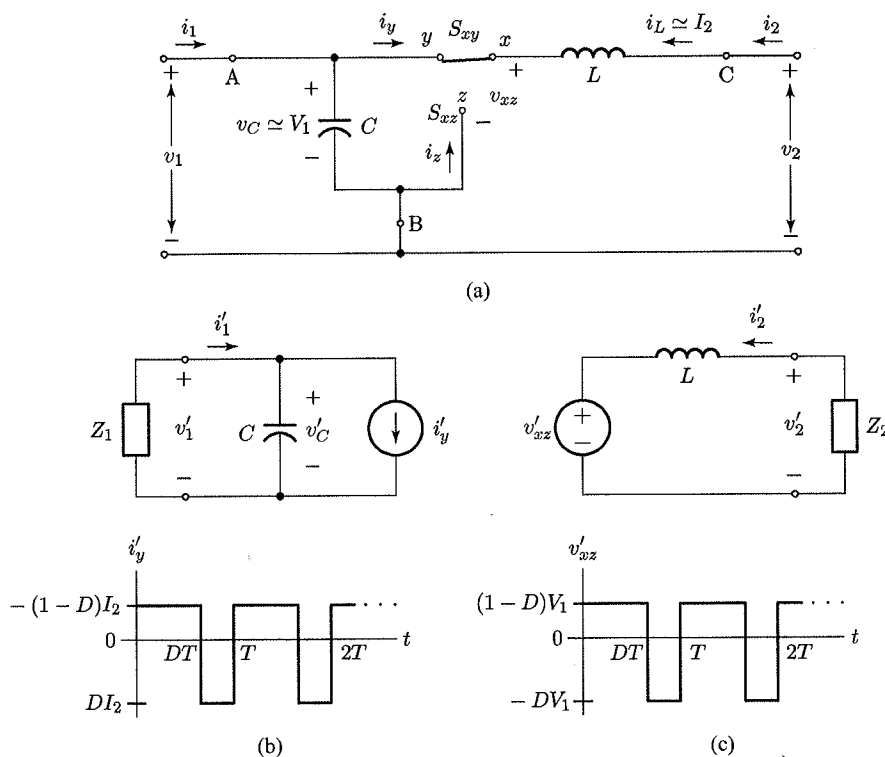


Figure 5.23 (a) The direct converter connection of the canonical cell. (b) First-order ripple-current model. (c) First-order ripple-voltage model for this converter.

We now develop a model that we can use to calculate the ripple voltage at port 2 of the direct converter. Based on the assumption that the capacitor is an ac short circuit at the switching frequency, the ac component of the voltage v_{xz} , which we denote by v'_{xz} , is as shown in Fig. 5.23(c). We now replace the circuit to the left of L in Fig. 5.23(a) by an equivalent voltage source of value v'_{xz} . The resulting ripple-voltage model is shown in Fig. 5.23(c). We calculate the first-order voltage ripple at port 2 by dividing v'_{xz} between the impedance of L and that of the external network Z_2 .

5.6.2 Ripple-Frequency Model for the Indirect Converter

For the indirect converter of Fig. 5.8(b), repeated here as Fig. 5.24(a), the interactions of the capacitor and the inductor with the impedances of the external systems are more complicated than they are for the direct converter. In the direct converter, the high-frequency impedance

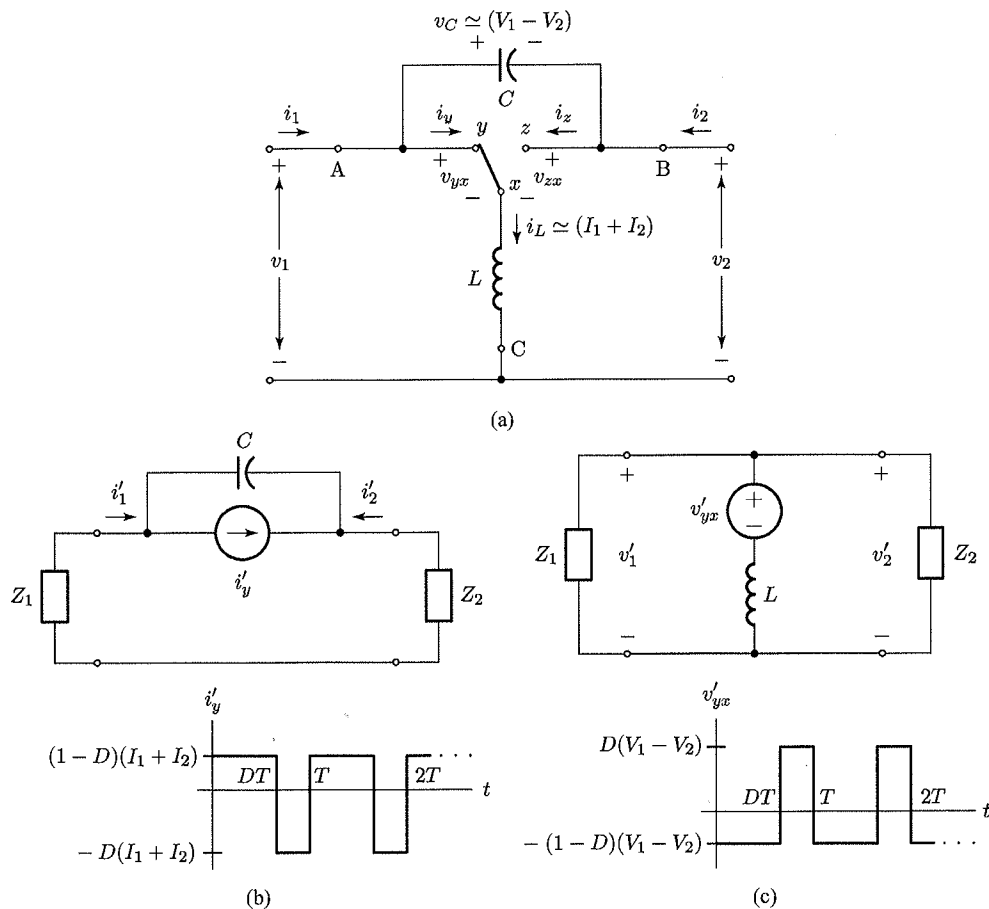


Figure 5.24 (a) The indirect converter connection of the canonical cell. (b) First-order ripple-current model. (c) First-order ripple-voltage model.

of only one external system influences the filtering effectiveness of the L or C in the canonical cell. For the indirect converter, we must simultaneously consider the impedance of both external systems to determine the ripple in the terminal variables.

We first consider the effect of capacitor value on the terminal ripple currents. We assume that the inductor is large enough for its ripple current to be negligible, permitting us to treat it as an infinite impedance (open circuit) at the switching frequency. Again, $i'_y = -i'_z$. The result is the switching-frequency ripple-current model shown in Fig. 5.24(b). We derive the waveform of i'_y by recognizing that the inductor current is $I_1 + I_2$ and is switched back and forth between nodes y and z . The ac current i'_y has two paths through which it can flow. One is through the capacitor, and the other is through the *series* connection of the two external system impedances. Therefore the high-frequency impedance of the capacitor must be small compared to the sum of the external impedances, $Z_1 + Z_2$, if we are to keep the switching-frequency current from flowing through the external network. The ripple current that does show up at the converter terminals has the same amplitude at both ports.

Now we consider the requirements on the inductor, assuming that the capacitor is large enough that its voltage ripple is negligible and the capacitor can be treated as a short circuit in the switching frequency model. This assumption forces $v'_{yx} = v'_{zx}$, putting nodes y and z at the same potential in the ac model. We can therefore connect these nodes, and in turn connect them to the inductor through a source of value v'_{yx} , creating the voltage-ripple model of Fig. 5.24(c). (In deriving the waveform of v'_{yx} , you should remember that if $V_1 > 0$, then $V_2 < 0$ for the indirect converter.) The source v'_{yx} is in series with the inductor and the *parallel* combination of the two external network impedances. Therefore the high-frequency impedance of the inductor must be large compared to that of Z_1 and Z_2 in parallel if the switching-frequency voltage is not to appear at the converter terminals. To the extent that it does appear, the ripple voltage is the same at both ports.

5.6.3 Minimum L and C for the Direct Converter

Up to this point we have assumed that L was large enough that its current could be considered constant, and that C was large enough that its voltage could be considered constant. We now address the problem of determining how large L and C have to be in order to justify these approximations. Any ripple in i_L or v_C shows up directly as first-order ripple in i_2 or v_1 , respectively, in Fig. 5.23(a). It would have a second-order effect on i'_1 or v'_2 by adding ripple to the flat portions of the waveforms of the ripple sources i'_y and v'_{xz} in Fig. 5.23(b) and (c). Another reason for calculating the first-order capacitor voltage and inductor current ripples is that they are independent of the external system impedances. The values for L and C that result from specifying a maximum ripple are thus parameters that we can compare for different converters without regard to the external systems.

In the calculations that follow we again use the first-order approximations that $i_1 \approx I_1$ when calculating the ripple in v_1 , and $v_2 \approx V_2$ for calculating the ripple in i_2 . When the series switch is off in the direct converter of Fig. 5.23(a), the current flowing into the capacitor is I_1 , and during this interval (of length $(1 - D)T$) the capacitor voltage v_1 will increase, giving a peak-peak ripple amplitude of

$$\Delta v_1 = \frac{1}{C} \int_0^{(1-D)T} I_1 dt = \frac{I_1(1-D)T}{C}. \quad (5.12)$$

During this same interval, the voltage across the inductor equals V_2 , and the inductor current $i_L = -i_2$ decreases by an amount equal to its peak-peak ripple amplitude:

$$-\Delta i_L = \Delta i_2 = \frac{1}{L} \int_0^{(1-D)T} V_2 dt = \frac{V_2(1-D)T}{L}. \quad (5.13)$$

With specified limits on the terminal current and voltage ripple amplitudes Δi_2 and Δv_1 , we can use (5.12) and (5.13) to determine minimum values of L and C :

$$C \geq \frac{I_1(1-D)T}{\Delta v_1}, \quad (5.14)$$

$$L \geq \frac{V_2(1-D)T}{\Delta i_2}. \quad (5.15)$$

After we have determined values for L and C from these expressions, we can determine the maximum energy stored in each element, which is useful because the stored energy is an indication of the element's size and cost. To express the stored energy in a useful way, we first define the *ripple ratios* $\mathcal{R}_C$ and $\mathcal{R}_L$ for the capacitor voltage and the inductor current, respectively. These are the ratios of the peak ripple amplitude to the dc (average) value of the capacitor voltage and inductor current, or

$$\mathcal{R}_C = \frac{\Delta v_1/2}{V_1}, \quad (5.16)$$

$$\mathcal{R}_L = \frac{\Delta i_2/2}{I_2}. \quad (5.17)$$

By combining (5.14) with (5.16), and (5.15) with (5.17), we can find the minimum peak stored energy in the capacitor and inductor:

$$E_C = \frac{1}{2} C V_{1p}^2 = \left(\frac{1-D}{4} \right) \left(\frac{P_o}{f_s} \right) \frac{(1 + \mathcal{R}_C)^2}{\mathcal{R}_C}, \quad (5.18)$$

$$E_L = \frac{1}{2} L I_{2p}^2 = \left(\frac{1-D}{4} \right) \left(\frac{P_o}{f_s} \right) \frac{(1 + \mathcal{R}_L)^2}{\mathcal{R}_L}. \quad (5.19)$$

In these expressions, f_s is the switching frequency $1/T$; V_{1p} and I_{2p} are the peak values of the capacitor voltage and inductor current, respectively; and P_o is the average power of the converter, that is, $P_o = I_1 V_1 = -I_2 V_2$. The value P_o/f_s is the amount of energy transferred from input to output during one cycle of the switching frequency. As we decrease the power or increase the switching frequency, the energy storage requirement goes down. Equations (5.18) and (5.19) show that, for a fixed ripple ratio, as the value of $1-D$ gets larger, the peak energy storage gets proportionally larger. Therefore, as the difference between the input and output voltage or current becomes greater, the energy storage requirement increases proportionally.

Example 5.5 Specifying L and C Values for a Direct Converter

The circuit of Fig. 5.25(a) shows a direct (buck) converter between a source of 101 V having an internal resistance of 0.5Ω , a resistive load of 2Ω , and a load voltage of 20 V. What values of L and C will ensure that the input and output currents have ripples with peak-to-peak amplitudes of no more than 5% of their

dc values if $f_s = 500$ kHz? If these minimum values of L and C are used in the circuit, what are the peak-to-peak terminal ripple-voltage amplitudes?

We can determine the average input-port variables $\langle v_1 \rangle$ and $\langle i_1 \rangle$ by equating the powers at the input and output ports of the converter:

$$\begin{aligned}\langle v_1 \rangle \langle i_1 \rangle &= \frac{\langle v_2 \rangle^2}{2} = 200 \text{ W}, \\ \langle v_1 \rangle &= 101 - (0.5) \langle i_1 \rangle, \\ \langle i_1 \rangle &= 2 \text{ A}, \\ \langle v_1 \rangle &= 100 \text{ V}.\end{aligned}$$

The duty ratio $D = 20/100 = 0.2$ and $T = 2 \mu\text{s}$.

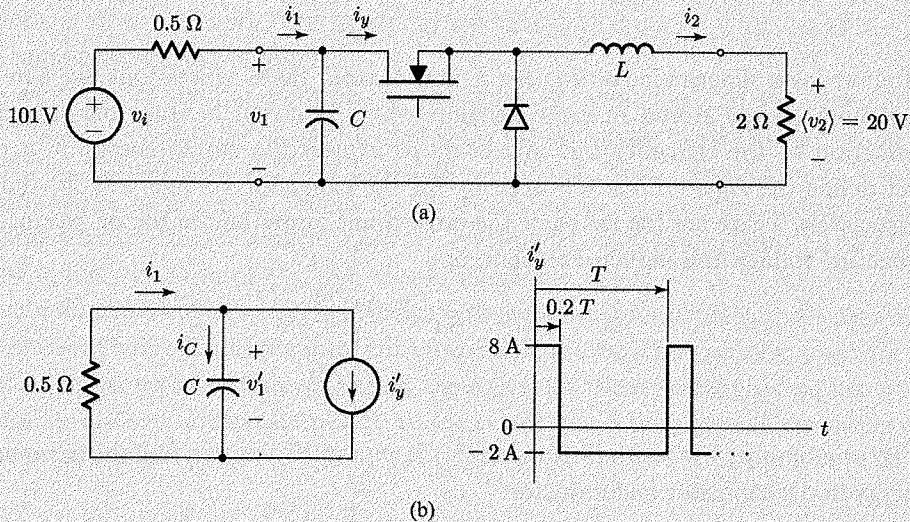


Figure 5.25 (a) Buck converter circuit. (b) Input ripple-current model for the circuit in (a).

We can easily calculate the value of L from (5.15):

$$L \geq \frac{20(0.8) (2 \times 10^{-6})}{0.05(10)} = 64 \mu\text{H}.$$

To determine the value of C , we must first use the input ripple-current model of Fig. 5.25(b) to calculate i'_1 . Because $i'_1 \ll i'_y$ with the proper values of L and C , we can make the first-order approximation that $i_C \approx -i'_y$. Therefore the peak-to-peak ripple voltage on C , $\Delta v'_1$, is

$$\begin{aligned}|\Delta v'_1| &\approx \frac{1}{C} \int_0^{DT} i'_y dt = \frac{(0.2)(8) (2 \times 10^{-6})}{C} = \frac{6.4 \times 10^{-7}}{C}, \\ \Delta i'_1 &= \frac{\Delta v'_1}{0.5} \leq .05 \langle i_1 \rangle = 0.1 \text{ A}, \\ C &= \frac{6.4 \times 10^{-7}}{0.1} = 64 \mu\text{F}.\end{aligned}\tag{5.20}$$

We can now calculate the first-order ripple in the terminal voltages. Because the load is resistive, the output voltage ripple as a percentage of V_2 is the same as the ripple-current specification. That is,

$$\frac{\Delta v'_2}{V_2} = \frac{\Delta i'_2}{I_2} = 5\%.$$

The input voltage ripple given by (5.20) is

$$\frac{\Delta v'_1}{V_1} = \frac{0.05}{100} = 0.05\%.$$

In practice one would include a capacitor at the output to reduce the size of the inductor and to provide the converter with better dynamic performance.

5.6.4 Minimum L and C for the Indirect Converter

We use the same approach to determine the capacitor and inductor values for the indirect converter of Fig. 5.8(b). The resulting values of L and C are

$$L \geq \frac{V_1 D T}{\Delta i_L}, \quad (5.21)$$

$$C \geq \frac{I_1(1-D)T}{\Delta v_C}. \quad (5.22)$$

In the indirect converter the average capacitor voltage is the sum of $|V_1|$ and $|V_2|$, and the average inductor current is the sum of $|I_1|$ and $|I_2|$. Expressed in terms of these terminal variables, the ripple ratios for the capacitor and inductor are:

$$\mathcal{R}_C = \frac{\Delta v_C/2}{|V_1| + |V_2|}, \quad (5.23)$$

$$\mathcal{R}_L = \frac{\Delta i_L/2}{|I_1| + |I_2|}. \quad (5.24)$$

We can again express the peak stored energy in the capacitor and inductor in terms of the ripple ratios:

$$E_C = \left(\frac{1}{4}\right) \left(\frac{P_o}{f_s}\right) \frac{(1 + \mathcal{R}_C)^2}{\mathcal{R}_C}, \quad (5.25)$$

$$E_L = \left(\frac{1}{4}\right) \left(\frac{P_o}{f_s}\right) \frac{(1 + \mathcal{R}_L)^2}{\mathcal{R}_L}. \quad (5.26)$$

Note that, unlike for the direct converter, these storage requirements do not depend on the duty ratio. We can again minimize them by choosing a ripple ratio equal to unity.

The peak energy storage requirements of the indirect converter are typically larger than they are for the direct converter. Only when the direct converter operates at a duty ratio near zero are they equal. For example, at an operating point of $D = 0.5$ for the direct converter, the indirect converter requires twice as much storage. Indirect conversion also imposes increased semiconductor stresses in a converter, as detailed below.

5.6.5 Calculations for the Buck/Boost and Ćuk Converters

The buck/boost and Ćuk variations of the indirect converter use two capacitors and two inductors, respectively. For these converters we compare the total capacitive or inductive peak energy storage requirements to the corresponding values calculated for the indirect converter. The requirement placed on the inductor in the buck/boost variant is the same as that for the inductor in the basic indirect converter. The requirements on the capacitors in the Ćuk variant and the basic circuit are identical. Let's consider the buck/boost circuit of Fig. 5.19(b).

The terminal voltage ripples in the indirect converter are related by the relative sizes of the external impedances connected to the two ports of the converter, as discussed in Section 5.6.2. In the analysis that follows, we assume that the ripple amplitudes at the two ports are proportional to their dc voltages. That is, the ripple ratio at each port is the same as the ripple ratio of the capacitor in the basic indirect converter against which we are comparing the stored energy requirements of the buck/boost variant. As a result, the two capacitors in the buck/boost converter each have ripple ratios equal to that of the single capacitor in the basic indirect converter, or $\mathcal{R}_C$.

In the buck/boost circuit of Fig. 5.19(b), the capacitor at the left-hand port, C_1 , is charged in the same way as the capacitor in the direct converter. It therefore stores the same peak energy, which is given by (5.18):

$$E_{C_1} = \left(\frac{1-D}{4} \right) \left(\frac{P_o}{f_s} \right) \frac{(1+\mathcal{R}_C)^2}{\mathcal{R}_C}. \quad (5.27)$$

The ripple voltage on C_2 is

$$\Delta v_{C_2} = \frac{DTI_2}{C_2}, \quad (5.28)$$

giving a minimum capacitance of

$$C_2 \geq \frac{DTI_2}{2V_2\mathcal{R}_C}. \quad (5.29)$$

The resulting peak energy stored in C_2 is

$$E_{C_2} = \left(\frac{D}{4} \right) \left(\frac{P_o}{f_s} \right) \frac{(1+\mathcal{R}_C)^2}{\mathcal{R}_C}. \quad (5.30)$$

The sum of E_{C_1} and E_{C_2} is exactly equal to the peak energy stored in the single capacitor of the basic indirect converter and given by (5.25).

A similar analysis done on the Ćuk variant of the indirect converter shown in Fig. 5.19(c) gives the same results. That is, the total energy storage requirement is identical to that for the basic indirect converter if we postulate the same ripple requirements at the ports – current ripple in this case.

5.7 Semiconductor Device Stresses

Semiconductor switches must be rated to carry the peak current I_p and withstand the peak voltage V_p presented to them by the power circuit. These ratings not only affect the cost of a switch, but they also affect various device performance parameters, such as storage time, current gain, and switching speed. Therefore they are meaningful parameters for comparing the attributes of different circuits. The product of the two peak stresses $V_p I_p$ is frequently useful for making comparisons among the various topologies. We call this product the *switch-stress parameter*.

For both the direct and indirect converters of Fig. 5.8, the peak voltage at the switches is equal to the maximum capacitor voltage. Similarly, the peak current carried by either switch is equal to the maximum inductor current. In the case of the buck/boost variant of the indirect converter of Fig. 5.19(b), the peak switch voltage is the sum of the two peak capacitor voltages. For the Čuk variant of Fig. 5.20(c), the peak current is the sum of the two peak inductor currents.

For the direct converter circuit of Fig. 5.8(a), the stress parameter for both switches is

$$V_p I_p = V_1(1 + \mathcal{R}_C)I_2(1 + \mathcal{R}_L) = P_o \frac{(1 + \mathcal{R}_C)(1 + \mathcal{R}_L)}{D}. \quad (5.31)$$

For the indirect converter of Fig. 5.8(b), the stress parameter is

$$\begin{aligned} V_p I_p &= (|V_1| + |V_2|)(1 + \mathcal{R}_C)(|I_1| + |I_2|)(1 + \mathcal{R}_L) \\ &= P_o \frac{(1 + \mathcal{R}_C)(1 + \mathcal{R}_L)}{D(1 - D)}. \end{aligned} \quad (5.32)$$

Note that for both converters, the stress parameter is smallest when the conversion ratio is unity, that is, when $D=1$ in the direct converter and when $D=0.5$ in the indirect converter.

We call the switch-stress parameters normalized to the power P_o the *switch-stress factors*. Figure 5.26 shows them for both direct (buck or boost) and indirect converters under the condition of zero ripple ratios. We need to plot only the stresses for output-to-input voltage ratios between zero and one because of circuit symmetry. We obtain the switch-stress factors for ratios between one and ∞ by exchanging the input and output roles of the ports. For example, the factors for a conversion ratio of 2.5 are the same as those for a conversion ratio of 0.4 ($= 1/2.5$). When the conversion ratio is unity, the stresses on the switches in the indirect converter are four times greater than they are for the direct converter. At a 50% conversion ratio, the difference is more than 2. Even at a 25% ratio the difference is greater than 1.5. Clearly, we would not choose to use an indirect converter unless we needed either a negative output voltage or *both* buck and boost conversion ratios from the same converter. The relative energy storage requirements for the two converters, as derived in Sections 5.6.3 and 5.6.4, also support this decision.

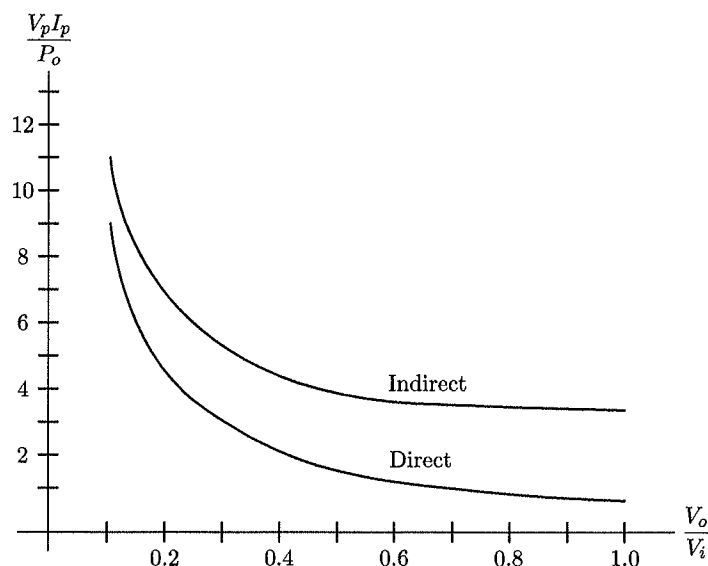


Figure 5.26 Switch-stress factors for direct and indirect converters.

5.8 Three-Level Flying-Capacitor Converter

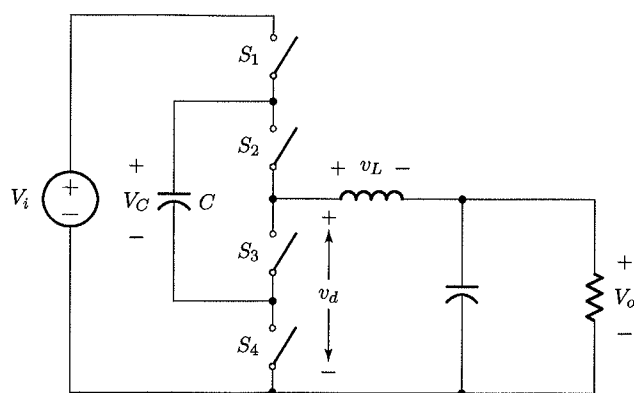
If large variation of the conversion ratio is necessary, both the energy storage requirements and switch-stress factors can be reduced by varying the converter input voltage to avoid extreme duty ratios. One circuit technique for accomplishing this is through the use of *multi-level* power conversion. An example of this is the *flying-capacitor* multi-level (FCML) converter, a three-level version of which is shown in Fig. 5.27(a). The FCML itself is just one of a whole variety of multi-level converters that attain improved performance and reduced individual device stresses at the cost of increased complexity. Multi-level conversion will be further treated in Chapters 8 and 9, as it is often employed in the context of dc/ac inverters.

The circuit of Fig. 5.27(a) can have four possible switch states $S_{k,\ell}$, where the subscripts indicate which switches are closed. These states yield four different voltages v_d :

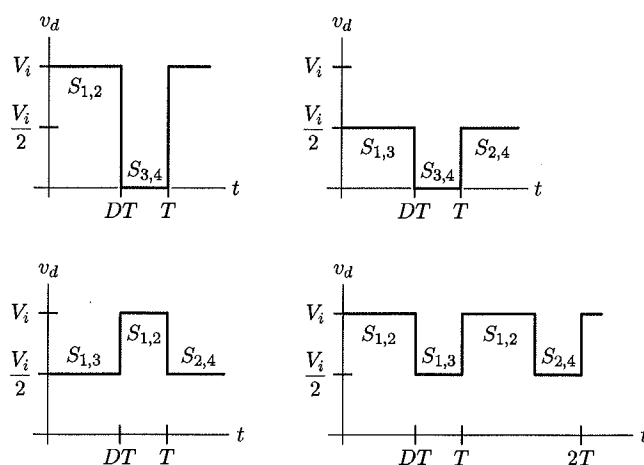
- state 1 $S_{1,2}: v_d = V_i$,
- state 2 $S_{3,4}: v_d = 0$,
- state 3 $S_{1,3}: v_d = V_i - V_C$,
- state 4 $S_{2,4}: v_d = V_C$.

If V_C is controlled to be $V_i/2$, states 3 and 4 produce $V_i/2$ and the filter input voltage can be V_{dc} , $V_i/2$, or 0. The converter can thus be operated to produce four different waveforms for v_d , as shown in Fig. 5.27(b).

Switching between states 1 and 2 produces a conventional buck converter. However, for $V_o \leq V_i/2$, considerable performance improvement can be obtained by employing the flying capacitor to reduce the voltage at the filter input to $V_i/2$ and switching among states 1, 2, 3, and 4 as



(a)



(b)

Figure 5.27 (a) A flying-capacitor buck converter producing an output voltage V_o . (b) The three possible voltages v_d if $V_C = V_i/2$.

shown in the middle waveform in Fig. 5.27(b). Alternating between states 1,3 and 2,4 assures C is charged and discharged equally to maintain $V_C = V_i/2$. In this operating mode the filter sees twice the switching frequency of the individual switches and $\Delta v_L = V_i/2$. Compared to a conventional buck converter with an output voltage $\leq V_i/2$, the three-level flying-capacitor converter produces twice the ripple frequency and half the inductor ripple voltage. The result is a smaller inductor for the same output ripple current. And although there are four switches, each changes state only twice per period T , as do the switches in a conventional buck. However, each switch sees $V_i/2$, not V_i , so switching losses are reduced by half per switch, giving total switching losses comparable to those of the conventional converter. If V_C can be precharged to $V_i/2$ without subjecting the devices to V_i at start-up, the switches can be rated for $V_i/2$ instead of V_i , further reducing losses, as lower-voltage devices with lower $R_{ds(on)}$ can be used. Because of these advantages and despite their complexity, flying-capacitor multi-level converters have

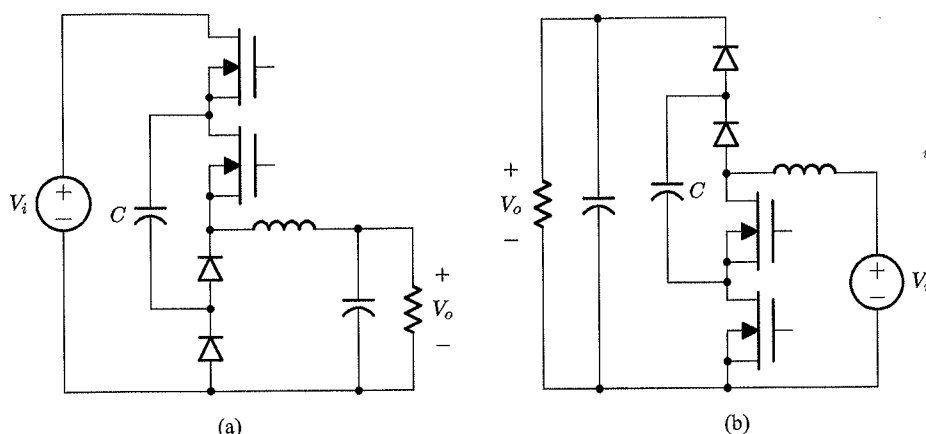


Figure 5.28 The circuit of Fig. 5.27 with switch implementations for (a) a buck converter and (b) a boost converter.

found uses in applications from low-voltage integrated CMOS power converters to high-voltage industrial machine drives.

Theoretically, V_C is maintained at its preset voltage by the charging symmetry of states 3 and 4, but any variations in switch timings or device rise and fall times will cause V_C to drift from its setpoint value. Therefore, some means to maintain $V_C = V_i/2$ must be provided. Feeding back V_C to control the relative on-times of states 3 and 4 is one technique frequently employed.

Although the circuit shown in Fig. 5.27 is a buck converter, depending on the switch implementations it can also be configured as a boost converter, as shown in Fig. 5.28.

Flying capacitors can also be used to produce multiple levels in dc/ac inverters, as discussed in Chapter 8. In addition they are the foundation of *switched-capacitor converters* presented in Chapter 6.

5.9 Converter Operation with Discontinuous Conduction

So far we have implicitly assumed that the peak inductor current ripple was smaller than the dc component of the inductor current. The total current therefore was always positive, and the diode was forced to be on when the transistor was off. However, if the dc component of the current is smaller than the peak ripple, the total current will fall to zero during the time when the diode is on, that is, during the period $1 - D$. The diode will then turn off, and the inductor current will remain zero until the transistor is turned on again. When this sequence occurs, we say that the converter is operating in the *discontinuous conduction mode* (DCM).

Figure 5.29 shows the current and voltage waveforms of the direct converter of Fig. 5.11 operating in DCM. In drawing these waveforms, we assumed an additional capacitor across the output terminal pair, so that v_2 is constant at V_2 . Note that when the inductor current is zero, the voltage across the inductor is also zero, and the diode is reverse biased by the voltage $v_{xz} = V_2$. The average output voltage (which in this case is V_2) is still equal to $\langle v_{xz} \rangle$,

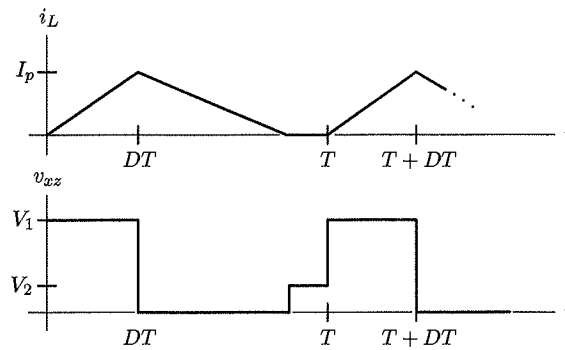


Figure 5.29 Waveforms of i_L and v_{xz} for the buck converter of Fig. 5.11 operating in the discontinuous conduction mode.

but its value in this mode of operation is higher than (5.4) predicts for the continuous conduction mode.

In practice, owing to switch and other capacitances, the voltage at the switch node x in Fig. 5.11 can ring with the inductor during the portion of the cycle when both switches are off. This can be a practical concern as regards electromagnetic interference (EMI), and is considered a disadvantage of DCM operation.

For the direct converter operating in continuous conduction, the voltage across the inductor is set at all times by the state of the switches. The average of this voltage must be zero, so the duty ratio specifies, by (5.4) and (5.5), the input/output conversion ratios. In discontinuous conduction, however, the inductor voltage averaged over a switching cycle is always zero, independent of D . The reason is that discontinuous conduction, by definition, ensures this condition. Therefore the duty ratio does not directly control the voltage conversion ratio. Instead, given V_1 and V_2 , the duty ratio controls the average current in the inductor, and hence the dc components of i_1 and i_2 , or I_1 and I_2 . That is, in discontinuous conduction the current and voltage conversion ratios are functions of the duty ratio, but the functional dependencies are *not* those of (5.4). We must specify the external systems in order to determine these conversion ratios.

Example 5.6 Direct Converter in Discontinuous Conduction

The converter of Fig. 5.30 is operating in DCM and supplying a constant voltage V_2 to a resistive load in parallel with a large capacitor. What is the value of V_2 ?

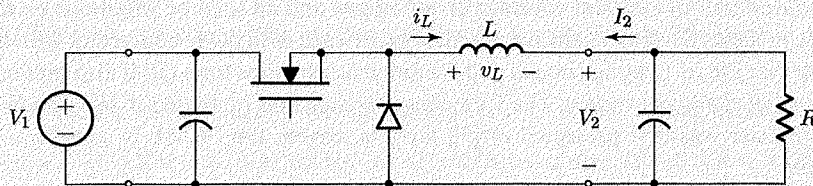


Figure 5.30 A buck converter analyzed under the condition of discontinuous conduction.

When we turn the transistor on for a time DT , the inductor current increases linearly from zero to a peak value of

$$i_{Lp} = \frac{(V_1 - V_2)DT}{L}. \quad (5.33)$$

When we then turn the transistor off, the voltage across the inductor is $-V_2$, and the time required to reduce its current to zero is

$$\Delta t = \frac{Li_{Lp}}{V_2} = \left(\frac{V_1}{V_2} - 1 \right) DT. \quad (5.34)$$

The average value of i_L is therefore

$$\langle i_L \rangle = \frac{i_{Lp}(DT + \Delta t)}{2T} = -\langle i_2 \rangle = \frac{V_2}{R}. \quad (5.35)$$

When we combine (5.33), (5.34), and (5.35), we obtain

$$V_2^2 + V_2 \left(\frac{V_1 RD^2 T}{2L} \right) - \left(\frac{V_1^2 RD^2 T}{2L} \right) = 0, \quad (5.36)$$

which we can solve for V_2 .

Designers often avoid operation in DCM, at least under nominal conditions. One reason is that the peak rms current stresses on the semiconductor devices are high compared to their values in a converter operating at the same power level in the continuous conduction mode. If the load varies over a wide range, however, the converter may enter the discontinuous mode at the low-power end. When this happens, the response of the converter to changes in the duty ratio is altered. In this case, we must make sure that the control circuit continues to work properly.

Positive aspects of DCM operation include the fast transient response resulting from the small inductor value and high ripple, simplified converter dynamics due to the inductor current returning to zero each cycle, and the very wide power range that can be spanned with a small inductor.

Notes and Bibliography

The evolution of the canonical cell is described in [1]. Landsman's approach is insightful and clear, and the paper represents the first time the fundamental similarity among various dc/dc converter topologies is identified; subsequent papers have identified and applied other, more complex, canonical cells to realize dc/dc converters. An excellent coverage of topologies and all sorts of variations on them can be found in [2], which reflects Severns' extensive practical experience with dc/dc converters. Mitchell's book [3] does a more concise job of treating circuits and is more mathematical than [2]. It also presents a good discussion of control of dc/dc converters. The Ćuk converter was first presented in [4]. The flying-capacitor multi-level converter was first presented in [5]; for this reason, the FCML is sometimes called a "Meynard converter."

1. E. Landsman, "A Unifying Derivation of Switching dc-dc Converter Topologies," in *IEEE Power Electronics Specialists Conference (PESC)*, pp. 239-243, 1979.

2. R. P. Severns and G. Bloom, *Modern DC-to-DC Switchmode Power Converter Circuits*, New York: Van Nostrand, 1985.
3. D. M. Mitchell, *DC-DC Switching Regulator Analysis*, New York: McGraw-Hill, 1988.
4. R. D. Middlebrook and S. Čuk, "A New Optimum Topology Switching dc-to-dc Converter," in *IEEE Power Electronics Specialists Conference (PESC)*, pp. 160-179, 1977.
5. T. A. Meynard and H. Foch, "Multi-Level Conversion: High-Voltage Choppers and Voltage-Source Inverters," *IEEE Power Electronics Specialists Conference (PESC)*, pp. 397-403, 1992.

PROBLEMS

- 5.1** Proceeding along the lines of Example 5.1, determine the amplitude of the ripple on the voltage v_o in Fig. 5.31. Assume that the input capacitor C is large enough that the input voltage is constant at value V_i and that the Norton equivalent resistance R_N of the external load network is small.

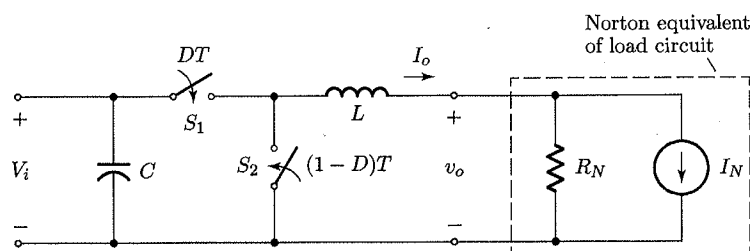


Figure 5.31 The dc/dc converter whose output voltage ripple is calculated in Problem 5.1.

- 5.2** Figure 5.32 shows a buck converter supplying 5 V to a load of 0.1Ω from a 25 V source having an internal resistance of 0.2Ω . Determine the duty ratio D at which the converter is operating.

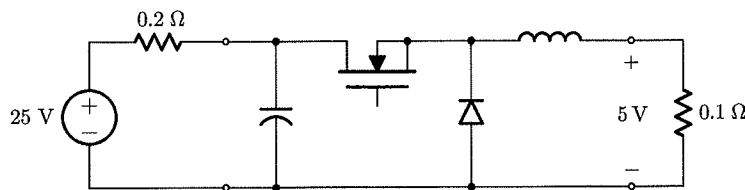


Figure 5.32 The buck converter whose operating duty ratio is calculated in Problem 5.2.

- 5.3** With reference to Fig. 5.21 it was stated that $\langle v_x \rangle = 0$. Show that this is true.
- 5.4** Derive a switch implementation for the direct boost converter so that the input and output circuits share a common positive terminal.
- 5.5** Reconsider Example 5.3 and develop a bilateral direct converter in which the terminal currents reverse direction but the terminal voltage polarities remain unchanged.

Part III

Components and Devices

A power circuit is typically composed of only a few kinds of components (other than its source and load): switches, and energy storage elements such as capacitors and inductors (or transformers). In its ideal form, each of these components is lossless and capable of operating at any frequency. In the ideal switch, for instance, the voltage across it is zero when it is on, the current flowing through it is zero when it is off, and the transition between these two states occurs infinitely fast. Although we did not explicitly discuss the means by which the actual switches are turned on and off (a topic covered in Chapter 23), an ideal switch responds infinitely fast to its drive signal and requires zero drive power. And, of course, there is no limit to the off-state voltage or the on-state current of the ideal switch. Similarly, the ideal capacitor and the ideal inductor are purely energy storage elements, and they each store energy in only one field type. The ideal capacitor contains no magnetic field energy, and the ideal inductor contains no electric field energy. Furthermore, neither ideal element dissipates any energy.

In reality, the components we use fall short of this ideal description, and so our power circuits are neither 100% efficient nor capable of operating at an infinitely high switching frequency. Improvements in power circuit behavior depend on the development of components with characteristics that are closer to ideal, and design techniques that compensate for the non-idealities. You thus need to understand the fabrication and material technologies on which power electronic components depend. Only through such an understanding can you properly apply power circuit components, as well as track and predict developments in the field.

As an example of the impact of advances in components, consider how the power supply for low-dc-voltage electronic loads has evolved. In the early 1950s, semiconductor power diodes became a practical replacement for vacuum tube (or selenium) rectifiers.[†] As the on-state voltage of a semiconductor diode (about 1 V) was much lower than that of a vacuum rectifier (about 30 V), the power supply efficiency increased significantly. By the mid 1960s, the switching speeds of bipolar power transistors had improved enough to make high-frequency dc/dc converters operating in the 10–20 kHz range possible. This improvement allowed replacement of linear-regulator technology with the lighter, smaller, and more efficient switching regulator. By the late 1970s, the power MOSFET became economically practical and permitted efficient operation in the 200 kHz to 1 MHz range, which further reduced the size of the required switching power supply. The more recent availability of semiconductor devices using wide-bandgap materials, such as SiC and GaN, and improved magnetic materials, have enabled switching frequencies

[†] Not many readers of this book will have experienced the joy of replacing the cat whisker and galena crystal with a germanium diode, and marveling at the increased volume of the local radio station.

in the 10s and even 100s of MHz. In each case, a component improvement permitted the “breakthrough” in power electronic technology.

15.1 Practical Semiconductor Switches

Typical current and voltage waveforms for a semiconductor switching device are shown in Fig. 15.1. When the switch is on, it exhibits a forward voltage drop V_{on} , which gives a conduction loss $p = V_{on}I$ as current I flows through it. Similarly, when the switch is off, it carries a small *leakage current* that produces an off-state loss. The leakage current is usually small enough that we can neglect this off-state loss. But when the switch is a Schottky diode, or when the power circuit tries to impose a voltage across a switch that is greater than the switch’s *breakdown voltage*, the leakage current and consequent off-state loss can be significant.

A semiconductor switch also needs time to make the transition from one state to the other. During this *switching time*, $v_{sw}i_{sw} \neq 0$ and the device dissipates energy. The total energy lost in the turn-on and turn-off transitions, multiplied by the switching frequency f_{sw} , gives the net *switching loss* for the device.

15.2 Practical Energy Storage Elements

We can explain some of the dissipation that occurs in real energy storage elements by the resistances of the capacitor’s leads and plates, and of the inductor’s winding. However, calculating the loss is not simple because we must take into account the *skin effect*, which dictates where ac current flows in a conductor, and the *proximity effect*, which dictates how much circulating, or *eddy*, current is induced in nearby wires or other conductive material. In general, to reduce

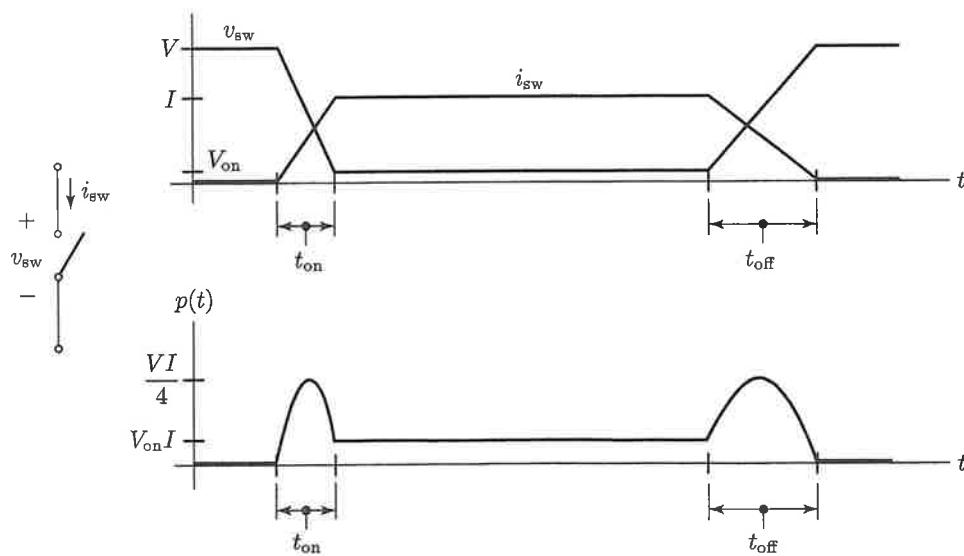


Figure 15.1 Voltage, current, and dissipation waveforms for a semiconductor switch.

this I^2R dissipation, we must make the energy storage element physically large to incorporate more conductor area. However, when the conductors are large enough for the skin and proximity effects to dominate the total loss, conductor fabrication becomes more complicated.

In addition to I^2R dissipation, the *change* of stored energy in the magnetic or dielectric materials also causes losses in inductors and capacitors. Generally, these losses are a nonlinear function of ac field strength, frequency, dc bias, and temperature. Both inductors and capacitors can be made from a variety of materials, which implies trade-offs. One material might have less loss at a given frequency than another, but have a lower permeability (or permittivity), a lower maximum field strength, or a higher cost.

Dissipation is not the only troublesome imperfection of inductors and capacitors. Because these elements can be used as filters, we also care about whether their impedances deviate from the anticipated $j\omega L$ or $1/j\omega C$. Capacitors have a parasitic series inductance resulting from the loop formed by their plates, their leads, and their connection to other components. Inductors have a parasitic parallel capacitance caused by the proximity of two windings at different potentials. We can neglect these parasitic components for both elements if the frequency is low enough. However, at very high frequencies the “capacitor” acts like an inductor and the “inductor” acts like a capacitor. As a result, a filter having these components does not behave as expected above a certain frequency.

15.3 Semiconductor Devices

In this section our purpose is to outline the main power semiconductor devices available today – diodes, transistors, and thyristors – and to discuss their terminal characteristics. In Chapter 16 we explain in more detail the physics underlying the behavior of these devices, and in Chapter 17 we discuss the unique characteristics of their design for power applications.

At this point, however, you should know that current in a semiconductor is carried by both positive charges (holes) and negative charges (electrons). When one of these charge carriers dominates the total concentration of carriers, we call it the *majority carrier*. We call the other type of carrier the *minority carrier*. Devices whose behavior is dominated by the distribution of majority carriers (*majority-carrier devices*) – the field-effect transistor, for example – can be switched much more quickly than can those devices that rely on the change of minority-carrier distributions (*minority-carrier devices*), such as the bipolar junction transistor. The carrier concentrations also determine the resistance of the different regions of a device.

15.3.1 Diodes

A diode is a switch whose state cannot be explicitly controlled, but is instead determined by the behavior of the circuit in which it is used. If the circuit tries to impose a positive voltage across the diode, it will turn on, and if the circuit tries to impose a negative current through the diode, it will turn off. Types of semiconductor diodes include the bipolar diode (based on a p–n semiconductor junction) and the Schottky diode (based on a metal–semiconductor junction). The schematic symbols and the i – v characteristics for these devices are shown in Fig. 15.2.

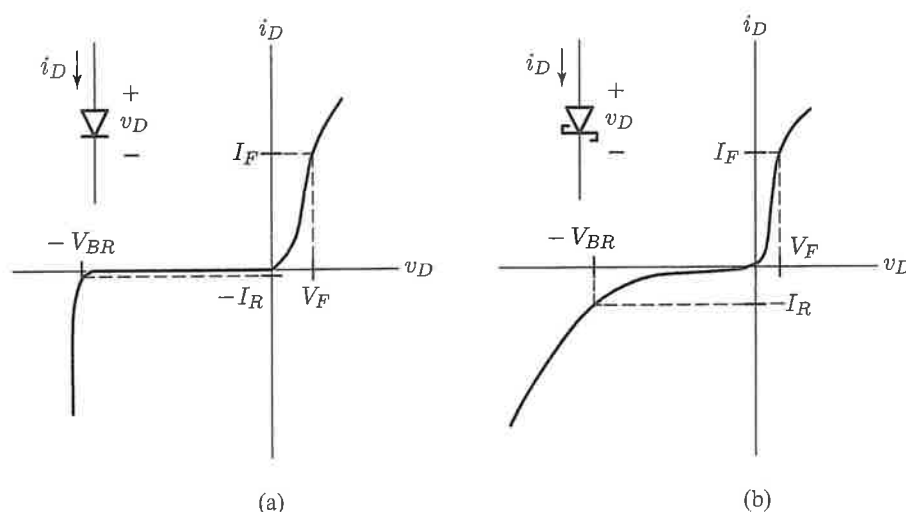


Figure 15.2 Symbols and i - v characteristics of power semiconductor diodes. (a) Bipolar junction diode. (b) Schottky diode.

At the rated current I_F , the forward voltage V_F is typically 1 V for a silicon bipolar power diode, and 0.6 V or less for a silicon Schottky diode. Diodes fabricated on wide-bandgap materials such as SiC and GaN have higher forward voltages. Note that above I_F the forward voltage increases at a faster rate with current than it does below I_F . The reason is that the resistive drop in the diode, which increases linearly with current, becomes comparable to, and then exceeds, the junction voltage, which increases logarithmically with current. The reverse voltage rating of a power diode, V_{BR} , sometimes called the *peak inverse voltage* (PIV), is determined by reverse leakage current. For a bipolar diode, V_{BR} can be in excess of 10 kV, and is usually specified at a leakage current I_R between 1 μ A and 10s of milliamps, depending on the I_F rating of the diode. For reverse voltages greater than V_{BR} , the leakage current increases rapidly. For the same die size and voltage, the leakage current of a Si Schottky diode is much higher than that of a Si bipolar diode. For both diodes, leakage current increases exponentially with temperature.

Bipolar Diode In addition to the presence of a nonzero forward drop and leakage current, the bipolar diode departs from ideal in its switching behavior. Figure 15.3 shows typical voltage and current waveforms during turn-on and turn-off of a bipolar diode in a buck converter. Note that when the device first begins to carry current, its forward voltage rises to V_{fp} , known as the *forward recovery voltage*, a value higher than the static curve of Fig. 15.2(a) predicts. Typical values for V_{fp} range between 5 V and 20 V. This phenomenon, which is a function of the on-state current, I_o in this case, occurs because the excess charge requires time to build up to its final value in the device, and before it does, the resistance of the diode is very high. During the *forward recovery time* t_{fr} , the power dissipation in the diode is much higher than predicted from the diode's static characteristics. Also, this transient voltage is impressed across other devices in the power circuit (the transistor in this converter, for example), and this additional voltage stress must be considered when rating these other devices.

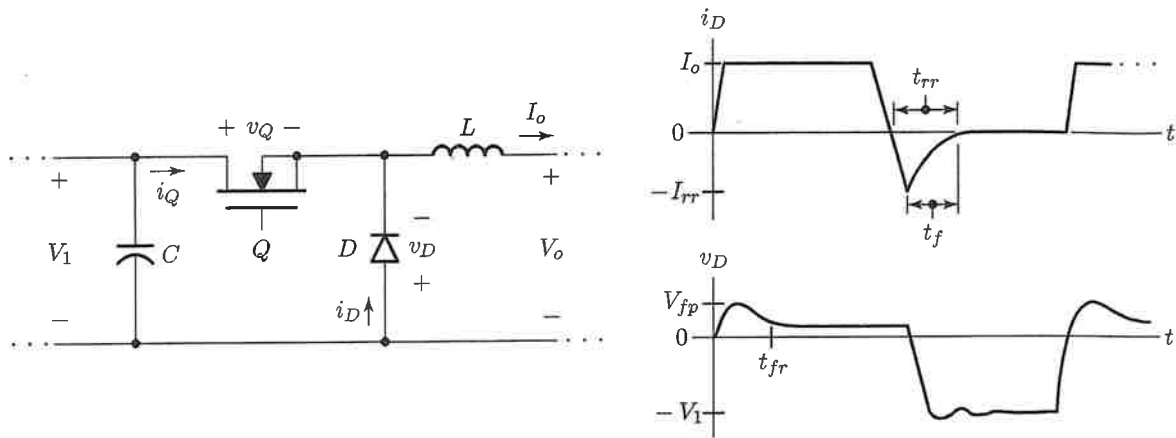


Figure 15.3 Switching waveforms for a bipolar diode used in a buck converter.

As the waveforms in Fig. 15.3 also show, when the bipolar diode turns off, its current first goes negative; then its reverse voltage begins to fall and the current returns to zero. This phenomenon is called *reverse recovery*, and it lasts for a time called the *reverse recovery time* t_{rr} . It occurs because the excess carriers in the device must be removed before the junction can withstand a reverse voltage. Another practical consequence of reverse recovery is that until the reverse current peaks at $-I_{rr}$ and starts to decrease to zero, the diode is still on. Therefore reverse recovery limits the maximum frequency at which we can use a diode.

Large values of $-I_{rr}$ increase the current stress on other semiconductor devices in the power circuit. During time t_f , which can be very short compared to t_{rr} , di/dt is high, and the parasitic inductance through which the reverse current flows imposes a high $L di/dt$ voltage across these other devices. For this reason, we sometimes refer to diodes in terms of their “snappiness.” A snappy diode is one whose current quickly drops to zero after I_{rr} is reached; in other words, the di/dt of these diodes is very high. In a non-snappy, or *soft-recovery*, diode, the recovery current takes a relatively long time to return to zero from $-I_{rr}$. That is, it exhibits a relatively small recovery di/dt .

In general, the higher the reverse voltage rating of a bipolar diode, the longer its forward and reverse recovery times. Reducing these recovery times (by reducing minority-carrier lifetimes) increases the forward voltage of the device. For this reason, diodes designed for low-frequency applications, such as 50 or 60 Hz rectifiers, have lower forward drops than diodes designed for high-frequency dc/dc or PWM converters. For instance, a typical 200 V *rectifier-grade* diode has a reverse recovery time of 50 μ s and a forward drop of 1.2 V at its rated current. A typical 200 V, 30 A, *fast-recovery* diode has a reverse recovery time as low as 10s of nanoseconds, but a forward drop of 1.6 V.

Schottky Barrier Diode A Schottky diode does not generate significant excess carrier concentrations when conducting. Instead it carries current primarily by the drift of majority carriers. For this reason in the ideal case it does not have the forward or reverse recovery behavior exhibited by a bipolar diode, although its junction space-charge layer capacitance must still be charged and discharged. As the reverse voltage rating of a Si Schottky is raised above 100 V,

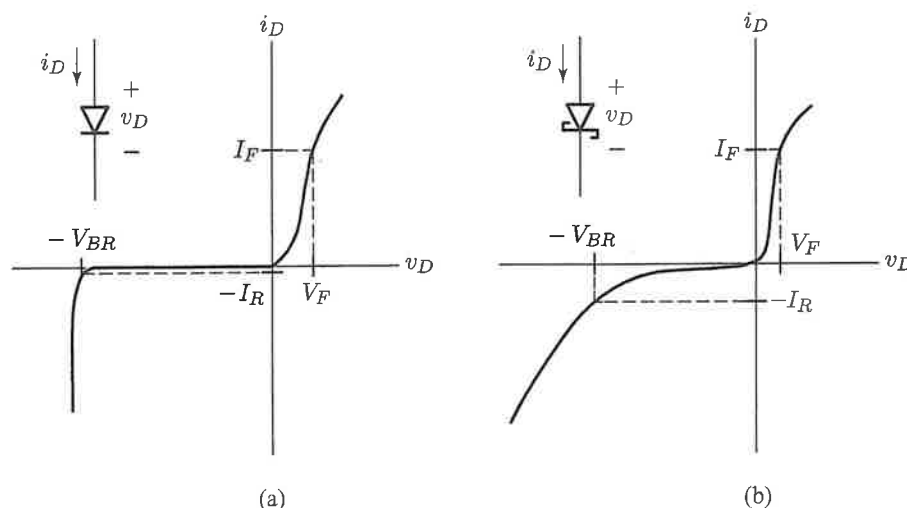
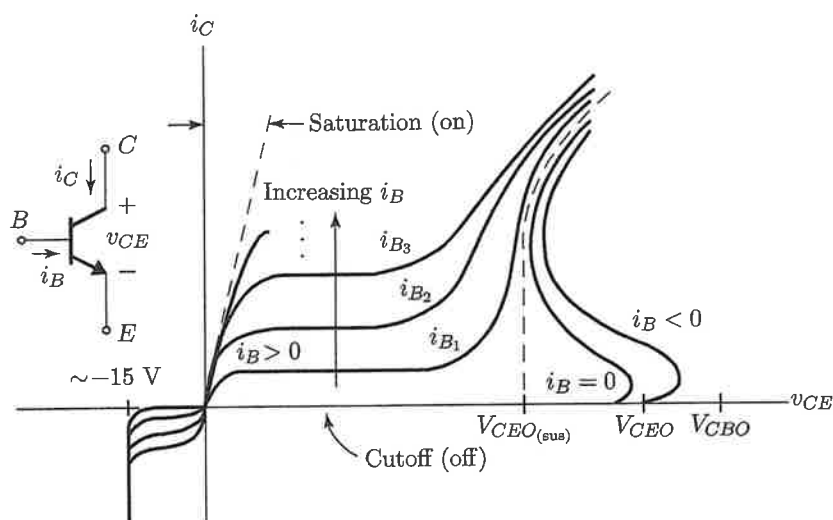


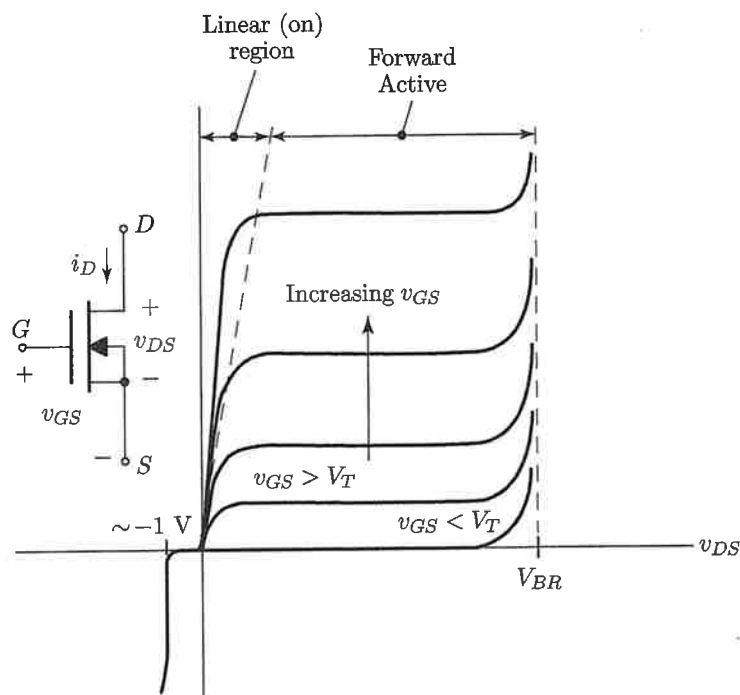
Figure 15.2 Symbols and i - v characteristics of power semiconductor diodes. (a) Bipolar junction diode. (b) Schottky diode.

At the rated current I_F , the forward voltage V_F is typically 1 V for a silicon bipolar power diode, and 0.6 V or less for a silicon Schottky diode. Diodes fabricated on wide-bandgap materials such as SiC and GaN have higher forward voltages. Note that above I_F the forward voltage increases at a faster rate with current than it does below I_F . The reason is that the resistive drop in the diode, which increases linearly with current, becomes comparable to, and then exceeds, the junction voltage, which increases logarithmically with current. The reverse voltage rating of a power diode, V_{BR} , sometimes called the *peak inverse voltage* (PIV), is determined by reverse leakage current. For a bipolar diode, V_{BR} can be in excess of 10 kV, and is usually specified at a leakage current I_R between 1 μ A and 10s of milliamps, depending on the I_F rating of the diode. For reverse voltages greater than V_{BR} , the leakage current increases rapidly. For the same die size and voltage, the leakage current of a Si Schottky diode is much higher than that of a Si bipolar diode. For both diodes, leakage current increases exponentially with temperature.

Bipolar Diode In addition to the presence of a nonzero forward drop and leakage current, the bipolar diode departs from ideal in its switching behavior. Figure 15.3 shows typical voltage and current waveforms during turn-on and turn-off of a bipolar diode in a buck converter. Note that when the device first begins to carry current, its forward voltage rises to V_{fp} , known as the *forward recovery voltage*, a value higher than the static curve of Fig. 15.2(a) predicts. Typical values for V_{fp} range between 5 V and 20 V. This phenomenon, which is a function of the on-state current, I_o in this case, occurs because the excess charge requires time to build up to its final value in the device, and before it does, the resistance of the diode is very high. During the *forward recovery time* t_{fr} , the power dissipation in the diode is much higher than predicted from the diode's static characteristics. Also, this transient voltage is impressed across other devices in the power circuit (the transistor in this converter, for example), and this additional voltage stress must be considered when rating these other devices.



(a)



(b)

Figure 15.4 Schematic symbols and static i - v characteristics for: (a) the npn BJT; and (b) the n-channel MOSFET.

which is lower than V_{CBO} because transistor action is amplifying the collector leakage current. In either case, as the collector leakage current grows, the device voltage folds back to the *sustaining* voltage $V_{CEO(sus)}$ – the voltage at which the leakage current amplified by the transistor action is in balance with, or sustains, the collector current.

Darlington Connection Because the current gain of a power BJT is so low, the connection of two BJTs as shown in Fig. 15.5(a) is often used to reduce the base drive requirement of the main

however, an important fraction of its total current is carried by minority carriers. As this happens, the device begins to exhibit recovery characteristics similar to those of a bipolar diode. When applied conservatively, Schottky diodes exhibit a forward drop of about 0.4 V less than that of comparably rated silicon bipolar junction diodes.

15.3.2 Transistors

Two fundamental types of transistors are the bipolar junction transistor (BJT) and the field-effect transistor (FET), among which is the metal-oxide-semiconductor field-effect transistor (MOSFET). A third device, the insulated-gate bipolar transistor (IGBT), is an integrated combination of these two. The discussion that follows describes transistors fabricated in Si, but all three of these transistor types are also available in the wide-bandgap (WBG) materials SiC and GaN. The structure and characteristics of these transistors, particularly those made from GaN, differ in some respects from those for Si. These WBG devices are discussed in Chapter 17.

The schematic symbols and the i - v characteristics for the BJT and MOSFET are shown in Fig. 15.4.[‡] A significant distinction between the two devices is that the BJT is controlled by its base current, while the MOSFET by its gate voltage. The BJT is a minority-carrier device, so like the bipolar diode, excess charge must be supplied or removed to turn it on or off. As a majority-carrier device, the MOSFET does not have stored minority carriers, but it does have junction and oxide capacitances that must be charged or discharged to turn the device on or off. Therefore, neither the BJT nor the MOSFET can be switched instantaneously.

Note that both the BJT and MOSFET have a maximum permissible off-state voltage. They both also have a nonzero on-state voltage that depends on how much current is flowing and how hard the transistor is driven. Between the on and off states is a region of operation referred to as the *forward active region*. For a fixed base current (respectively, gate voltage), the collector (respectively, drain) current of both transistors remains roughly constant over the full range of the device voltage in this region. Because power transistors are used as switches that are either on or off, the forward active region of operation is of interest only during switching, when it must be traversed.

Bipolar Junction Transistor To be in the on state, the BJT requires the presence of base current i_B . The ratio of on-state collector current i_C to the base current i_B necessary to support this value of i_C is an important parameter called the *dc current gain*, h_{FE} (or β_F). Unlike signal-level transistors, for which $h_{FE} > 200$ is common, values for h_{FE} of 5 to 20 at their rated current are typical for power transistors.

Base current also affects the off-state characteristics of a BJT, as shown in Fig. 15.4(a). If the base current is sufficiently negative (resulting from reverse biasing the base-emitter junction during turn-off), the collector voltage can go as high as the collector-base junction breakdown voltage V_{CBO} .[§] If $i_B = 0$ in the off state, the collector voltage can go only as high as V_{CEO} ,

[‡] The linear and forward active regions of MOSFET operation are also referred to as the triode and saturation regions, respectively.

[§] This subscript notation defines the terminal conditions for the parameter. For example, V_{CBO} is the collector-base voltage with the emitter open.

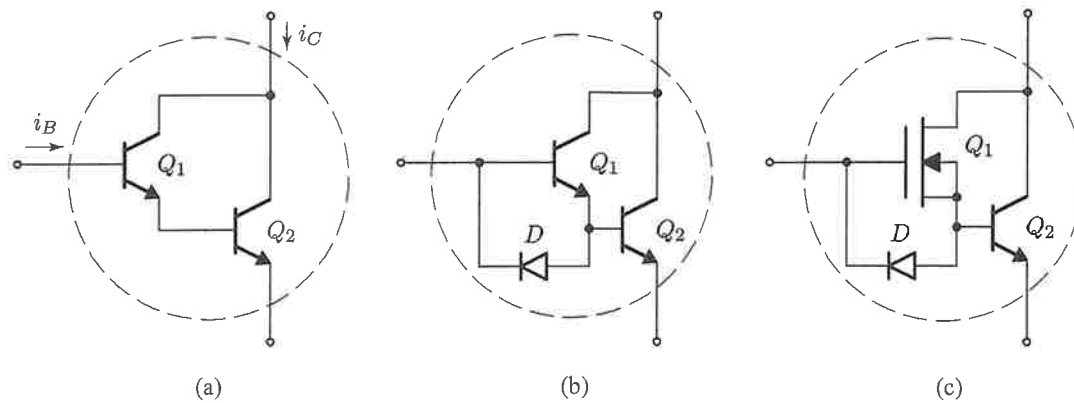


Figure 15.5 The Darlington connection. (a) The basic connection of two bipolar transistors. (b) The connection of (a) with the addition of a diode to improve its turn-off speed. (c) A MOSFET connected as the drive transistor Q_1 .

device. This is known as a Darlington connection and results in a composite device having an effective current gain that is approximately the product of the individual gains, that is, $h_{FE} = I_C/I_B \approx h_{FE1}h_{FE2}$. It is a particularly useful connection for power transistors because the base current necessary to drive the pair is considerably less than that needed to drive Q_2 alone. The smaller base current allows a reduction in the current rating of the base drive circuit, simplifying its design.

The Darlington connection has two disadvantages. One is that the on-state drop of the pair is higher than the on-state voltage of Q_2 alone. The reason is that the lowest on-state voltage of the pair is the hard-saturation voltage of Q_1 plus the base-emitter voltage of Q_2 , whereas Q_2 , if used alone, has a hard-saturation voltage equal to its base-emitter voltage *minus* its base-collector voltage. This condition is easy to see for signal-processing transistors, in which the saturation voltage of Q_2 alone could be on the order of 0.2 V, whereas the Darlington pair can have an on-state voltage of no less than 0.8 V – the saturated collector-emitter voltage of Q_1 plus the base-emitter voltage of Q_2 . However, the saturation voltage of a power device usually has an appreciable contribution from the resistance of the collector region, making the difference between a single device and a Darlington pair less dramatic. A Darlington pair exhibits an on-state voltage of between 2 and 5 V, depending on its current and voltage ratings.

The second disadvantage of the Darlington connection is that when the drive circuit tries to turn the pair off, it can pull current from the base of Q_1 only. The base current of Q_2 goes to zero when Q_1 turns off, and therefore the charge stored in the base of Q_2 is not forcibly removed. Instead, it decays to zero by the process of recombination, resulting in a long turn-off time. An effective solution to this second problem is to add a diode between the bases of Q_1 and Q_2 , as shown in Fig. 15.5(b), so that the drive circuit has access to the base of Q_2 during turn-off, when $I_B < 0$ and the diode is on. In a monolithic Darlington connection, the diode is usually a separate device mounted in the same package and wire-bonded to the die containing the Darlington pair.

We can also use a MOSFET for Q_1 , as shown in Fig. 15.5(c), resulting in a composite device with the MOSFET's ease of drive and the BJT's on-state drop. The larger the die area of the MOSFET, the smaller its on-state resistance, and the lower the on-state voltage of the pair.

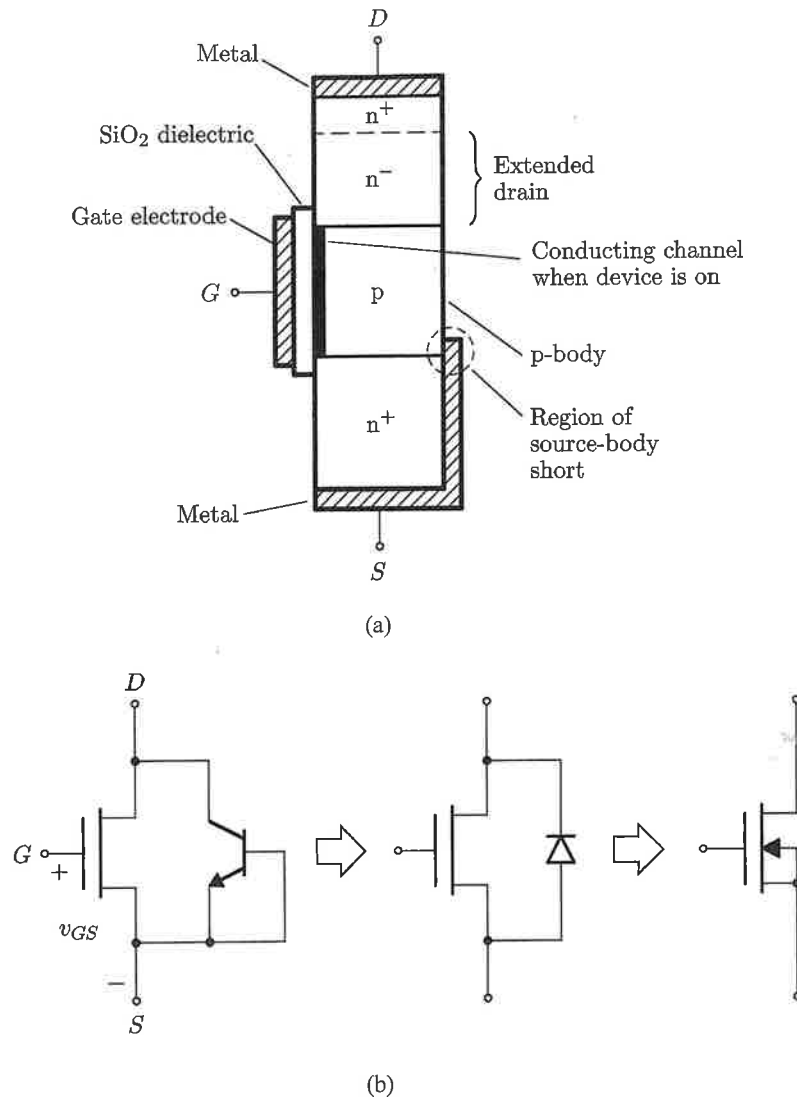


Figure 15.6 The n-channel power MOSFET. (a) Simplified structure. (b) The parasitic npn transistor with the source contact creating a base–emitter short, the parasitic body diode, and the MOSFET symbol showing the presence of the source–body short.

A monolithic derivative of this structure is the IGBT, which has largely displaced the Darlington connection.

MOSFET Figure 15.6(a) shows a simple representation of the *n-channel* vertical MOSFET structure typically used in discrete devices. “Vertical” refers to the direction of the drain current, which is orthogonal to the device substrate. When a positive voltage v_{GS} is applied from gate to source, a conducting n-type channel is formed beneath the gate electrode in the p-region, which is known as the *body* of the device. This channel connects the drain to the source, turning the device on.

The MOSFET turns on when v_{GS} exceeds the *threshold voltage* V_T , which is typically 2–5 V for Si devices. The gate acts like a capacitor, so when the MOSFET is either on or off, the gate power

is zero, making possible the use of gate drive circuits that are simple and efficient compared to those necessary to drive a BJT. Energy is required to switch the MOSFET, however, as the process is equivalent to charging or discharging a capacitor. For this reason, the gate drive circuit is required to supply energy at an appreciable average power if the MOSFET is operated at high frequencies.

When on, the MOSFET acts like a resistor of value $R_{DS(on)}$, which is composed of two parts. The first is the *channel resistance* arising from the conducting channel created by the electric field beneath the gate electrode. The second is the resistance of the relatively long region comprising the drain.

Unlike the channel resistance, whose value varies with v_{GS} , the drain resistance is of constant value. Therefore $R_{DS(on)}$ for most MOSFETs is independent of v_{GS} for v_{GS} higher than a few volts above the threshold voltage V_T . The most significant part of $R_{DS(on)}$ for devices rated above about 100 V results from the lightly doped drain region, called the *extended drain*, which determines their maximum V_{DS} rating, so its length for these devices is strongly related to the transistor's breakdown voltage.

The three silicon layers comprising the MOSFET of Fig. 15.6(a) create a parasitic npn transistor structure. To prevent this parasitic device from affecting the behavior of the MOSFET, the p-body and n^+ source regions are shorted by the source contact, as shown. The result is the parasitic bipolar diode across the MOSFET shown in Fig. 15.6(b). This *body diode* is often useful in bridge circuits. The MOSFET symbol indicates the presence of the source-body short and the resulting body diode. The direction of the arrow indicates that the body is p-type and the channel is n-type.

Figure 15.7 shows typical voltage and current waveforms for the MOSFET as it turns on and off with a clamped inductive load. Such a load creates the most severe switching behavior, because the switched voltage cannot begin to fall (at turn-on) or rise (at turn-off) until the

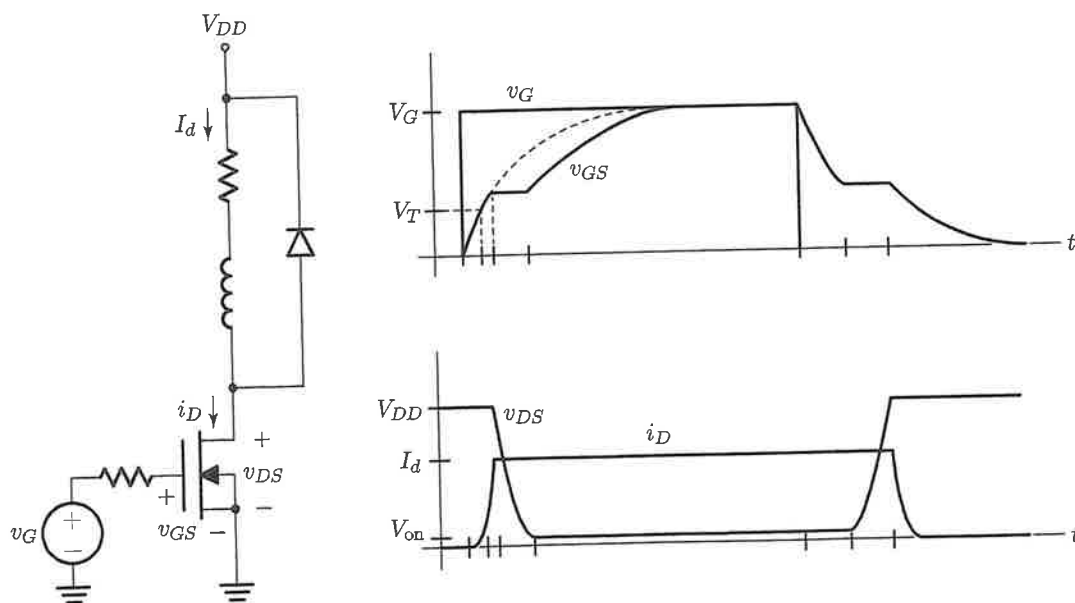


Figure 15.7 Typical waveforms for a power MOSFET with a clamped inductive load.

current has commuted from the diode to the transistor, or from the transistor to the diode. This is the condition used in device datasheets to specify turn-on and turn-off times. Note that the device does not begin to change state until the gate voltage equals V_T (typically 3–5 V). During each transition, as the drain voltage either falls or rises to its final value, the gate voltage stays constant at approximately the threshold voltage V_T . This plateau in the gate voltage waveform is called the *Miller effect*, which we explain in more detail in Chapter 23. The speed at which a MOSFET turns on or off is determined by the rate at which its parasitic capacitances can be charged or discharged, which is limited by the gate driver and the internal gate resistance.

Insulated Gate Bipolar Transistor The IGBT is a monolithically integrated Darlington-like connection of a MOSFET driving a BJT, as shown in Fig. 15.8(a). It has the benefits of the drive simplicity of a MOSFET and the high voltage capability of a BJT. Because of the way the two transistors are integrated, we can protect the MOSFET from the full off-state voltage of the IGBT by the space-charge layer at the collector junction of the BJT. The result is that a much smaller area on the device die is dedicated to the MOSFET. Note that because the two

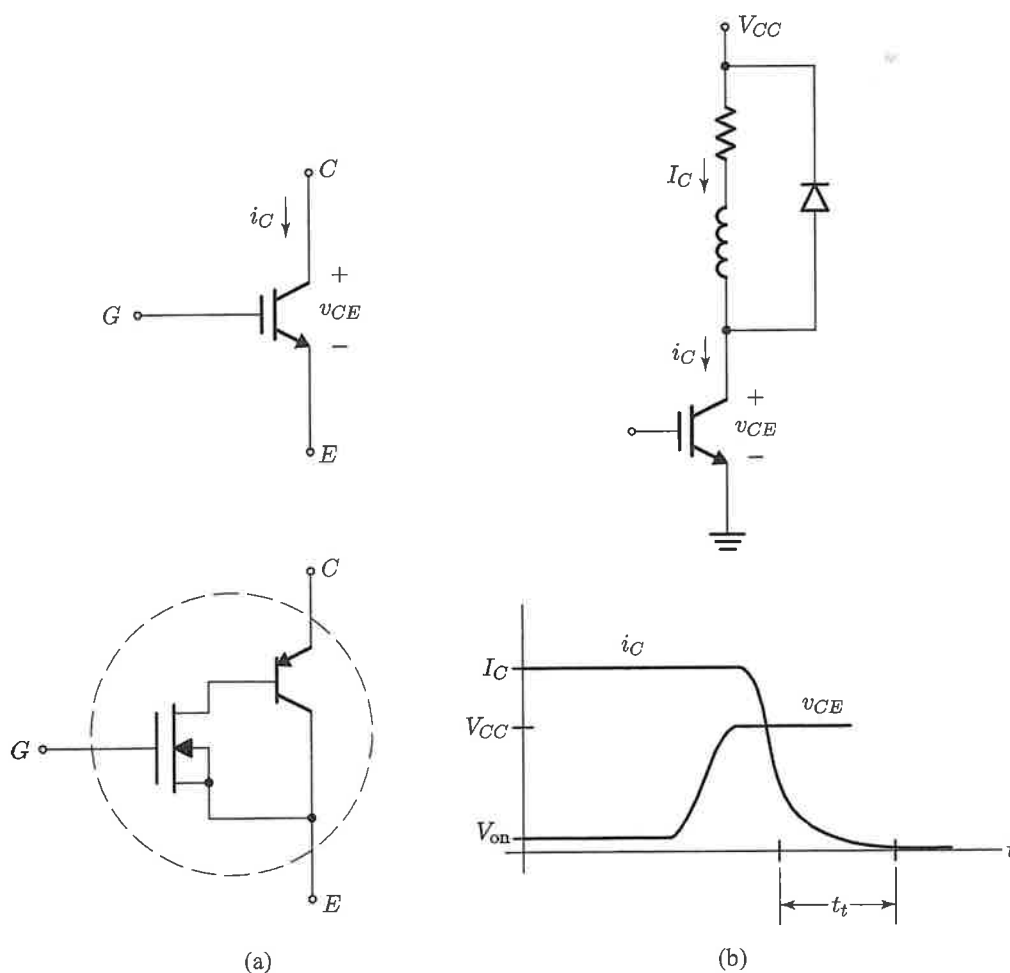


Figure 15.8 (a) The IGBT and its equivalent connection of MOSFET and BJT. (b) Waveforms at turn-off showing the slow “tailing” effect during t_t .

transistors are of opposite polarity (an n-channel MOSFET and a pnp BJT), the gate is driven with respect to the collector of the BJT.

The turn-on transition time of the IGBT is that of its BJT. The turn-off is much slower than would be possible with the BJT alone, however, because there is no way to aid the turn-off process by pulling negative current from the base of the BJT. Figure 15.8(b) shows how the collector current makes a quick change when the MOSFET first turns off, but then slowly tails to zero. The initial step corresponds to that fraction of the device current that had been flowing through the MOSFET. The tail during time t_t is the characteristic of a BJT turning off with an open-circuit base.

15.3.3 Thyristor

The name *thyristor* is generic for devices composed of four silicon layers (p-n-p-n) and exhibiting a regenerative internal mechanism that latches the device in the on state. As a family, thyristors are capable of switching voltages over 9 kV and currents over 8 kA (rms). (Such a device weighs over 5 kg.)

Members of the thyristor family used as switches in power circuits include the silicon-controlled rectifier (SCR), the triode ac switch (TRIAC), and the gate turn-off thyristor (GTO). As you have already seen in Chapter 4, the SCR is only semi-controllable; it can be turned on at a specified time, but it must be turned off by the action of the circuit in which it is connected. The GTO, on the other hand, can be turned both on and off from its control (gate) terminal, but it cannot block reverse voltage. The TRIAC is similar to the SCR, but can conduct and withstand voltage in either direction. In many respects, it can be thought of as a pair of SCRs connected in anti-parallel.

Silicon-Controlled Rectifier The most common form of the thyristor is the SCR. Figure 15.9(a) shows its schematic symbol and i - v characteristic. If $v_{AK} > 0$, a current applied to the gate turns the SCR on. This action initiates an internal regenerative mechanism that latches the device in the on state, even if the gate current is then removed. The gate power required to switch the SCR is therefore very small.

The voltage V_{BO} in the i - v characteristic is the *forward breakover voltage*. When subjected to a forward voltage above this value, the leakage current at junction J_1 acts similarly to gate current and is sufficient to initiate the internal regenerative mechanism in the absence of gate current, and the SCR turns on. Other mechanisms to trigger on the SCR are: a high dv_{AK}/dt , which produces a current across J_1 as its junction capacitance charges; a high temperature, which increases the J_1 leakage current; and optical injection of carriers into the p-gate region. Devices triggered optically, known as opto-SCRs, are commercially available, but their dv_{AK}/dt rating is quite low. The on-state voltage of an SCR is typically between 1 and 3 V.

When the SCR turns off, it displays the same reverse recovery phenomenon as a bipolar diode, and there is a period in the turn-off process during which $i_A < 0$, as illustrated in Fig. 15.9(b). Even after $i_A = 0$, however, the SCR is incapable of blocking a forward voltage until the charge remaining in the device has decayed below a critical value. The time required for this process is called the *circuit commutated turn-off time*, t_q , and ranges from about 5 to 100 μ s. Reapplying a

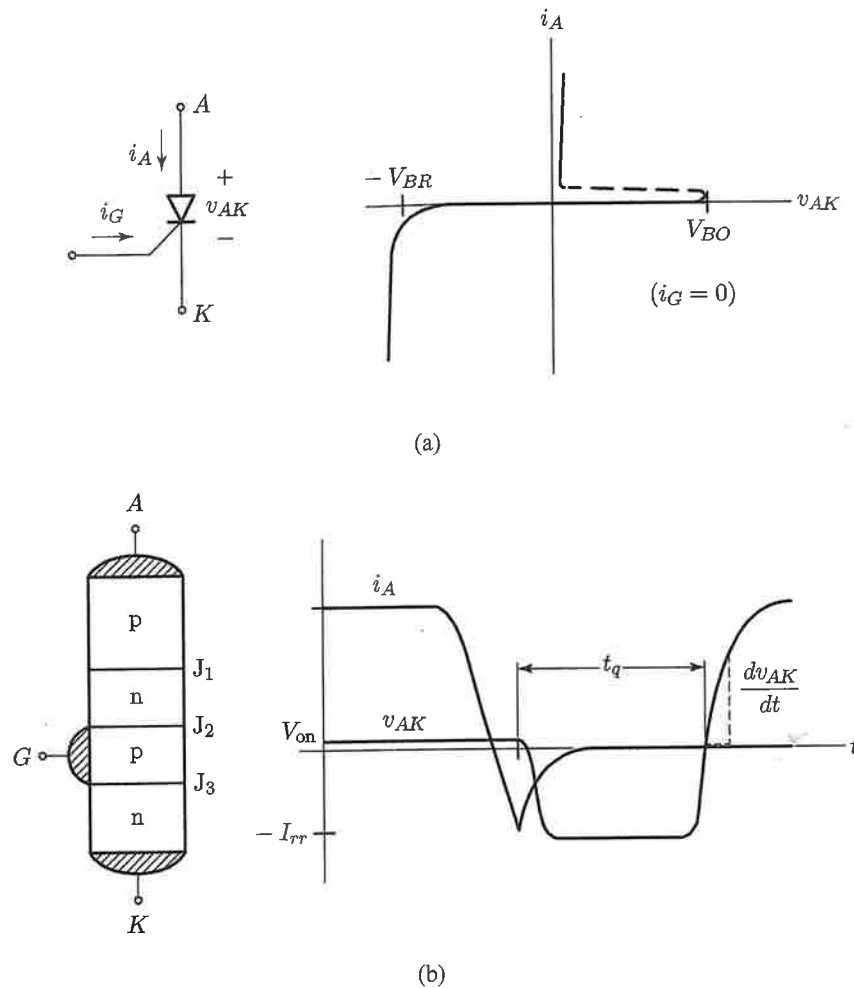


Figure 15.9 The SCR. (a) Schematic symbol and its static $i-v$ characteristic. (b) Reverse recovery, showing the minimum off-state time t_q .

forward voltage to the SCR before t_q risks retriggering the regenerative process and turning the device on again.

There are ways to reduce t_q by adjusting the SCR's design, although an additional consequence is a higher on-state voltage. For this reason there are two standard grades of SCRs: *rectifier grade* and *inverter grade*. Rectifier-grade SCRs have a t_q in the 100 μs range and an on-state voltage between 1 and 2 V. Inverter-grade SCRs have a t_q of between 5 and 30 μs , and an on-state voltage between 2 and 3 V. While inverter-grade SCRs have historically been applied to motor drives or other PWM applications, and to high-frequency resonant converters, their use has been supplanted in most of these applications by the MOSFET or IGBT. Today the primary use of the SCR is in line-frequency rectifiers, as discussed in Chapters 4 and 9.

There is another special version of the SCR, called the *asymmetrical SCR* (ASCR), in which the reverse blocking capability of the device has been sacrificed for a much shorter t_q and a higher dv_{AK}/dt rating. The ASCR is useful in power circuits that have a diode in anti-parallel with each controllable switch, and therefore require no reverse blocking capability.

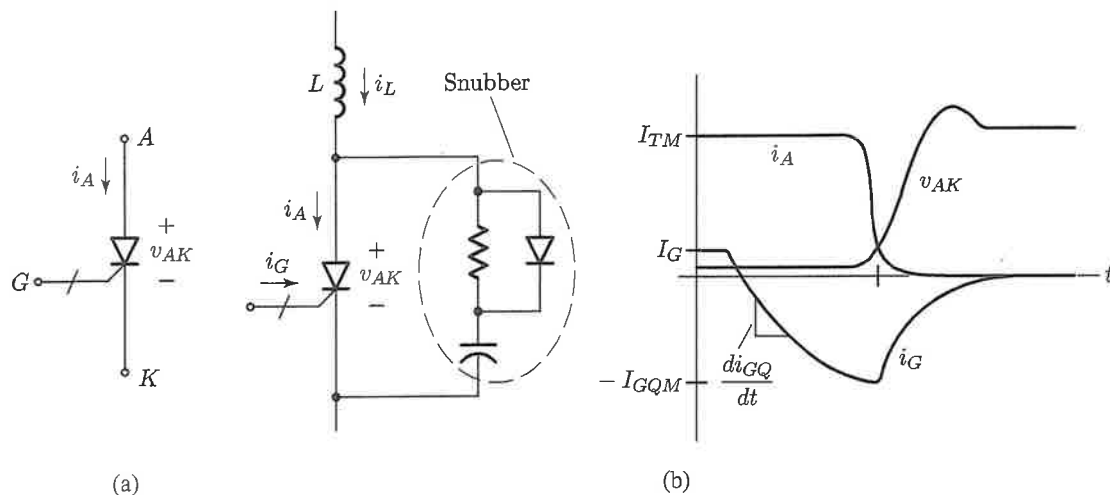


Figure 15.10 (a) The GTO symbol. (b) GTO switching waveforms when a snubber circuit is used to provide an alternate path for i_L .

Gate Turn-Off Thyristor The GTO is an SCR in which special provisions made during fabrication enable a large negative gate current to stop the internal regenerative process, turning the device off. Two parameters are compromised in the GTO: forward drop, which is typically 1 V more than a comparably rated SCR, and reverse blocking capability, which is typically less than 50 V.

Figure 15.10(a) shows the schematic symbol for the GTO and terminal variables during the turn-off process. Like the SCR, it has maximum dv_{AK}/dt and di_A/dt ratings. It also has a maximum anode current, I_{TM} , that can be turned off from the gate terminal. The *turn-off gain* of the GTO is defined as the ratio of I_{TM} to the peak negative gate current required to turn it off. Gains on the order of 3–4 are typical, so the gate drive must be capable of providing a large negative current I_{GQM} to turn off an anode current of I_{TM} . This negative current pulse has a minimum specified negative rate of rise, $-di_{GQ}/dt$. During its on state the GTO requires a minimum gate current I_G to prevent the device from turning off at low anode currents.

The mechanism employed to turn off a GTO is to force the anode current into regions of the device area having low regenerative gain. Therefore, as this current-focusing process is taking place, the anode current must decrease in order to avoid localized regions of high current density and device failure. Any parasitic or load inductance in series with the GTO will prevent i_A from decreasing at the necessary rate. For this reason, GTOs are always employed with a snubber circuit, as shown in Fig. 15.10(b), to provide an alternative path for the current in the circuit inductance L . This snubber also determines the rate of rise of v_{AK} and its overshoot.

Triode AC Switch The TRIAC, whose symbol and $i-v$ characteristic are shown in Fig. 15.11, is a thyristor capable of bilateral conduction. Because it is a slow device, having a dv/dt limit on the order of 100 V/ μ s, it is used almost exclusively for electric utility (50/60 Hz) applications such as light dimming. The TRIAC is bilateral, so we do not use the terms “anode” and “cathode” to identify its terminals. Instead, we identify the switched terminals as main terminals 1 and 2 (MT₁ and MT₂). We use MT₁ as the reference terminal for the gate and MT₂ voltages.

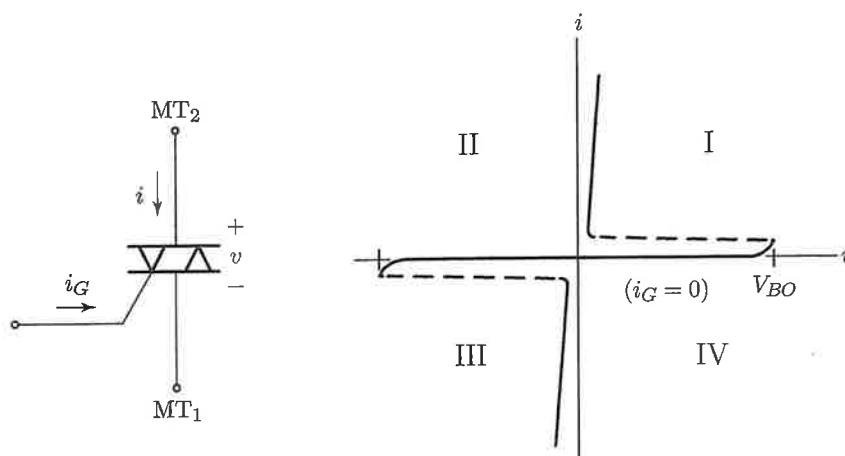


Figure 15.11 Symbol and i - v characteristic for the TRIAC.

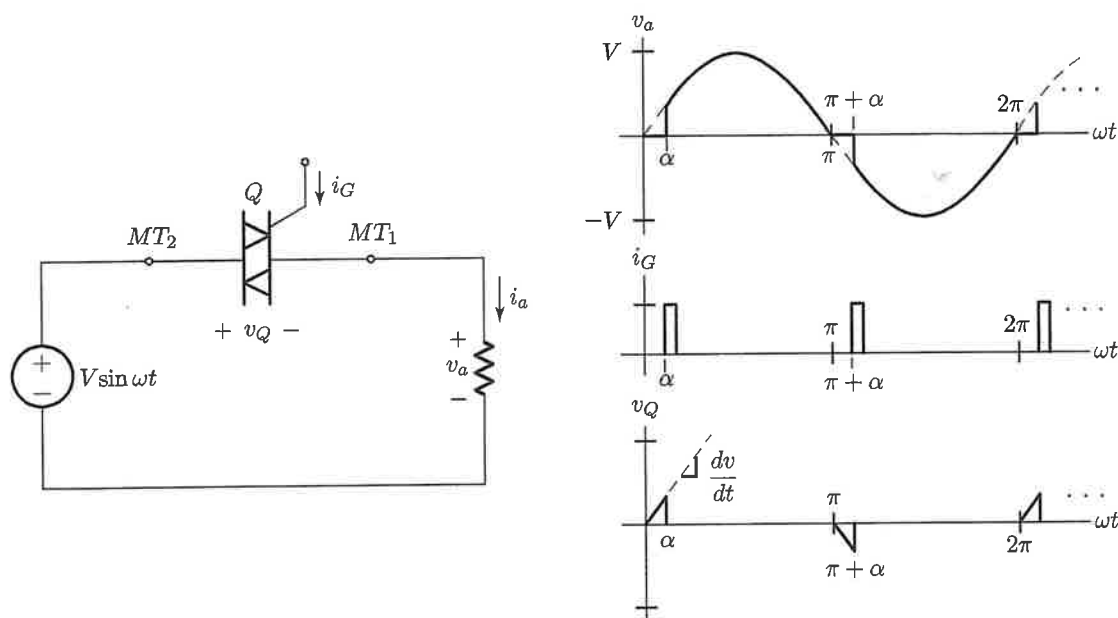


Figure 15.12 The TRIAC used as an ac controller to illustrate the commutating dv/dt to which it is subjected.

We can turn the TRIAC on with positive or negative gate current – and with MT_2 either positive or negative. The TRIAC conducts in quadrant I of its i - v plane if MT_2 is positive when gate current is applied, and in quadrant III if MT_2 is negative. The four possible triggering modes are therefore I+, I–, III+, and III–, with the “+” and “–” referring to the direction of the gate current. The triggering sensitivity of the TRIAC is different in each of the triggering modes, with I+ and III– being the most sensitive (requiring the least gate current) and therefore the most frequently used. In any of the modes, the gate characteristic, measured between G and MT_1 , is similar to a forward-biased diode.

In practically all applications the TRIAC must block voltage immediately after turning off, so we replace the SCR parameter t_q with a parameter called the *commutating dv/dt* for the TRIAC. The meaning of this parameter is illustrated in Fig. 15.12, which shows TRIAC Q being used

as an ac controller. When Q turns off at the zero-crossing of i_a , its voltage v_Q rises at an initial rate ωV . This rate of rise is the commutating dv/dt to which the TRIAC is subjected. It must be less than the maximum value specified by the manufacturer, otherwise the TRIAC will turn on again before the next gate current pulse. This limitation makes TRIACs unsuitable for use at frequencies above approximately 400 Hz.

15.4 Capacitors

Many types of capacitors are used in power circuits. They differ in the dielectric employed and in their physical form. The choice of capacitor type depends on the particular application. In addition to its capacitance value, a capacitor designed for power circuit applications usually has a maximum rms current specification because of internal dielectric and I^2R heating.

The dielectric material supporting the electric field in a capacitor is characterized by two parameters: its *dielectric strength*, which is the value of the electric field at which the dielectric breaks down; and its *dielectric constant*, K . The capacitance value achievable with a material is proportional to K , which is the permittivity of the material normalized to ϵ_o , the permittivity of free space: $K = \epsilon/\epsilon_o$.

An important parameter characterizing capacitors is the *dissipation factor* D , which is the ratio of the real part of the capacitor's admittance to its imaginary part. The admittance is frequency dependent, as is D , and therefore is usually presented in graphic form. A nonzero dissipation factor means that the capacitor exhibits losses. These losses are a function of construction method and material properties of the dielectric.

15.4.1 Film Capacitors

For applications that require large currents but relatively little capacitance, such as in snubbers or resonant tanks, a *film capacitor* is a common choice. It is simply two layers of metal foil separated by insulating film. This long dielectric-conductor sandwich is usually rolled into a cylinder. The two foil electrodes are offset so that only one sticks out beyond the dielectric at either end of the cylinder. Terminals are then connected to the edges of these electrodes.

The dielectric film is a thin, insulating sheet of plastic. Although many plastics can be used, the ones most commonly used in power electronic capacitors are polypropylene and polyester (Mylar). Polycarbonate is a very thermally stable dielectric with a very low dissipation factor, but is no longer generally available, as the major manufacturer of the film has stopped producing it. Polypropylene has more than twice the dielectric strength of polycarbonate (about 16 kV/mm for polypropylene versus 7.5 kV/mm for polycarbonate); polyester sits between (at about 11 kV/mm). The electrode can also be vacuum-deposited on the dielectric film to produce *metalized film*, yielding an electrode thickness of 0.01–0.05 μm ; alternatively, a separate sheet of metal foil, 5–15 μm thick, can be used. Capacitors using the *metal foil* construction are capable of carrying considerably higher currents than those made with metalized film.

The dielectric properties of polypropylene are very stable when subjected to changes in temperature. The material's dielectric constant ($K \approx 2.3$) changes only a few percent when

the temperature changes from 25 °C to 105 °C. Polycarbonate has a slightly higher dielectric constant ($K \approx 3.0$) and can be used at temperatures as high as 125 °C. The dielectric loss in both materials is very low, and is stable under temperature variations (a ΔD of approximately 0.02% for polypropylene and 0.1% for polycarbonate over a frequency range of 1–100 kHz is typical). The film capacitor's lead and electrode resistance are the dominant sources of loss.

The parasitic inductance of a film capacitor depends on the capacitor package's geometry and size. Typical resonant frequencies (where $\omega L = 1/\omega C$) range from 100 kHz for a 50 μF capacitor to 10 MHz for a 0.01 μF capacitor.

15.4.2 Electrolytic Capacitors

Electrolytic capacitors offer far greater capacitance per unit volume than do film capacitors. The trade-off is a much higher series resistance and an inability to withstand a reverse voltage of significant value. Because of their unipolar voltage limitation, they are most often used in dc filters in which a large capacitance is needed but the current (per unit of capacitance) is relatively small. There are two major types of electrolytic capacitors: aluminum and tantalum. Tantalum capacitors have more capacitance per volume, but they are more expensive because of the material's higher cost.

Aluminum Electrolytic Capacitors The first step in producing aluminum electrolytic capacitors is to chemically etch an aluminum foil, making its surface very rough. This process increases the surface area of the foil (which is connected to the positive terminal of the capacitor) by a factor of as much as 100. A very thin layer of aluminum oxide, Al_2O_3 , is then formed on the foil's surface. This oxide, which has a dielectric constant of $K = 8.4$, is the dielectric in the capacitor. A liquid electrolyte forms the other "plate" of the capacitor as it comes in contact with the oxide. A second foil of aluminum makes electrical contact to this electrolyte, and is the negative terminal of the capacitor. As the electrolyte is not a very good conductor compared to a metal, the resistance of an aluminum electrolytic capacitor is relatively high.

The behavior of the electrolyte is the reason that the life of electrolytic capacitors is short compared to that of other components used in power circuits. The electrolyte slowly escapes through the seal of the package until enough has left to change the electric characteristics of the capacitor dramatically. The rate at which this occurs is temperature dependent; the higher the temperature, the shorter the life. The loss of electrolyte occurs during storage as well as during operation of the power circuit, although usually at a much slower rate because the capacitor is cooler. Before the electrolyte runs out, the most common cause of failure of an aluminum electrolytic capacitor is a short circuit through the oxide.

Tantalum Electrolytic Capacitors Tantalum oxide, Ta_2O_5 , has a dielectric constant approximately three times that of Al_2O_3 , giving tantalum capacitors a significant volumetric specific density advantage over aluminum capacitors. Three different methods are used to construct tantalum electrolytics. The *tantalum foil* construction is identical to that used for making aluminum electrolytics. The *wet slug* and *solid* constructions each start with a porous tantalum pellet made by sintering powdered tantalum. A wire, which forms the positive terminal, is

attached to this pellet. In the wet slug construction, the negative electrode is a liquid electrolyte, which is absorbed by the porous pellet. The negative electrode of the solid tantalum electrolytic is manganese dioxide, created on the Ta_2O_5 surface by another chemical reaction.

The temperature dependency of tantalum capacitors is fairly linear over their operating range, varying from -5% at -55°C to $+5\%$ at 125°C . They are also stable with respect to applied voltage.

Electric Double-Layer Capacitors The electric double-layer capacitor (EDLC, DLC, or “super-capacitor”) has an extremely high capacitance, on the order of 100s of farads, but very low voltage ratings, approximately 3 V. They are capable of very high rates of charge and discharge, and useful in pulsed power applications, such as absorbing regenerated braking energy in electric vehicles. While their volumetric specific energy density is relatively low compared to batteries, their volumetric specific power density is much higher. These capacitors can be used for energy storage, but not filtering applications, as their high-frequency performance is poor.

15.4.3 Multilayer Ceramic Capacitors

Multilayer ceramic capacitors are made by stacking many layers of thin, unfired (“green”) ceramic sheets on which a conductor has been screen printed using conductive inks. This sandwich is then sintered to form a solid, monolithic capacitor.

There are two types of ceramic dielectrics, designated Class I and Class II. Class I ceramics have a dielectric constant (normalized permittivity) of between 10 and 500 that is very stable (typically $\pm 3\%$) over wide ranges of voltage, time, and temperature. The dissipation factor of these materials is also low, typically 0.1% at 1 MHz. Because of their low permittivity, however, Class I ceramics have a low volumetric specific density compared to Class II ceramics. Capacitors designated NPO (or COG) are made of Class I ceramic.

Class II ceramics have much higher dielectric constants than Class I materials and come in two subgroups. The first has a dielectric constant that ranges from 500 to 3000 and is relatively stable over voltage, time, and temperature. The second has a relatively unstable dielectric constant ranging from 3000 to 20 000.

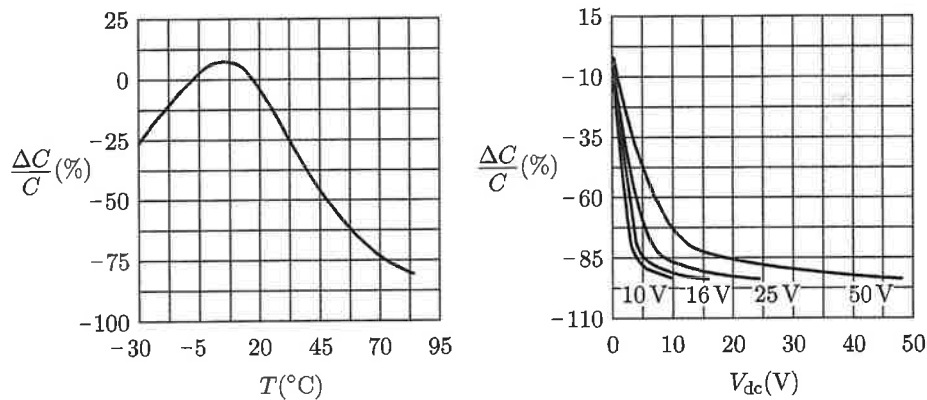
A three-symbol code is used to specify the performance of Class II ceramics. The first and second symbols specify the low and high limits of an operating temperature range. The third symbol specifies by how much the capacitance value at the temperature extremes will change from its value at 25°C . Table 15.1 defines this code. Two common types are X7R ($\pm 15\%$ from -55°C to $+125^\circ\text{C}$) and Z5U ($+22\%$, -56% from $+10^\circ\text{C}$ to $+85^\circ\text{C}$).

The Class II code designation provides information only about the temperature stability of the capacitance. However, Class II ceramics have a strong voltage dependency about which the code is silent. This dependency can result in an 80% decrease in capacitance at rated voltage. Class II ceramics also exhibit a logarithmic degradation of permittivity over time, losing on the order of 10% of their capacitance at one year, and 20% at 10 years. Representative voltage and temperature dependencies for typical Class II ceramic capacitors are shown in Fig. 15.13.

Multilayer ceramic capacitors can have very low series resistance and inductance and, in addition to their use in high-frequency resonant circuits, are often used as high-frequency bypass

Table 15.1 Class II ceramic capacitor codes.

	Low T (°C)	High T (°C)	Change from 25 °C value (%)
X	-55	4 +65	A ±1.0
Y	-30	5 +85	B ±1.5
Z	+10	6 +105	C ±2.2
		7 +125	D ±3.3
		8 +150	E ±4.7
			F ±7.5
			P ±10
			R ±15
			S ±22
			T +22, -33
			U +22, -56
			V +22, -82

**Figure 15.13** Representative temperature and voltage dependencies of Class II ceramic capacitors.

capacitors in parallel with electrolytics in filtering applications. In some cases they can provide energy densities comparable to electrolytic capacitors, but are relatively expensive.

15.5 Inductors and Transformers

Inductors and transformers used in power electronic circuits are often custom designed for the specific application. Seldom can these components be obtained from a manufacturer as a standard product. Even where off-the-shelf components are available, custom-designed components can often yield higher system performance. As a result, you must become familiar with the materials and manufacturing techniques used in producing magnetic components, subjects covered in Chapter 20. As a foundation for that presentation, in Chapter 2 we discussed the behavior of transformers and their circuit models, and here we provide an overview of loss mechanisms, materials, and the common physical forms of inductors and transformers.

15.5.1 Magnetic Materials

Most magnetic components used in power circuits contain a core of high-permeability material. Use of a core results in a smaller structure than otherwise, and the core prevents magnetic fields from extending great distances from the component. Such stray fields could cause interference or induction heating of other components or structures. Two broad classes of materials are used to make cores: most commonly ferrous alloys, and otherwise magnetic ceramics known as *ferrites*.

A great variety of ferrous alloys is available, and in several forms. These alloys differ in permeabilities, loss characteristics, and mechanical workability. For low-frequency applications (below 1 kHz), iron with a low silicon content, often known as *transformer steel*, is generally used. At higher frequencies, other elements are added to the iron to increase its permeability and lower its loss. These more exotic magnetic steels – containing various amounts of cobalt, nickel, molybdenum, and chromium – can be used to frequencies of about 20 kHz. They are very expensive relative to the silicon–iron materials. The ferrous alloys are available in either sheet or powdered form. When powdered, they can be used at frequencies higher than those cited, with nanocrystalline materials useful up to around 100 kHz. However, both their permeability and saturation flux density are considerably less than those of the bulk material.

Ferrites are ceramics made of ferrous oxide, zinc, and either manganese or nickel. Many proprietary types exist, and each is identified by a unique designation – for example, 4C4 or 3C6. Their maximum flux densities are about one third those of the ferrous alloys, and at temperatures above about 150°C their magnetic characteristics deteriorate markedly. Cores made of this material are limited in size because of their brittleness. Ferrites exhibit much lower losses at high frequencies than do ferrous alloys and thus can be used at much higher frequencies. The Mn materials are useful to approximately 1 MHz, while the Ni materials are acceptable to 10s of megahertz.

15.5.2 Magnetic Core Geometries

Figure 15.14 shows some of the commonly used core shapes. The “C,” “E,” and “pot” cores are the easiest to utilize, because their windings can be made on an open bobbin. The bobbin is then slipped over the legs of the “C” or “E” cores or the center post of the “pot” core. A toroidal core, on the other hand, must be wound by threading the wire through the core, a process that can be automated, but which is generally more expensive than winding on a bobbin.

15.5.3 Losses in Magnetic Components

Loss in a magnetic component arises from two sources: the resistance of the copper windings, and hysteresis and eddy currents in the core. At high frequencies, the copper loss is aggravated by skin and proximity effects, but in general we can calculate it in a relatively straightforward manner. Core loss, however, is a nonlinear function of both frequency and flux level. For a particular material, core manufacturers generally provide data giving core loss in W/kg as a graph parametric in frequency and flux level, as shown in Fig. 15.15 for ordinary transformer steel. For ferrites, the loss data is normally provided in W/cm³.

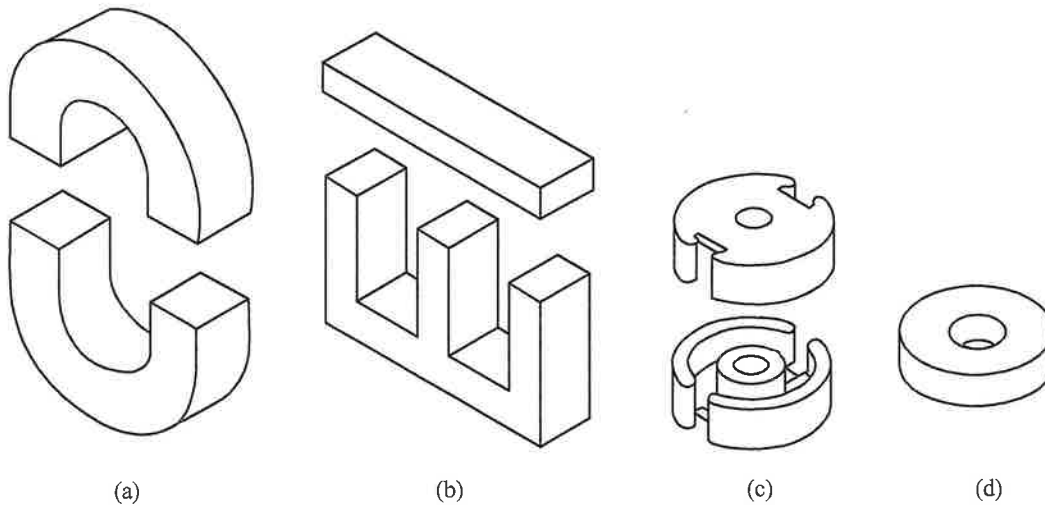


Figure 15.14 Core geometries used in power circuits. (a) Double “C” core. (b) “E” core with an “I” closing piece. (c) “Pot” core. (d) Toroidal core.

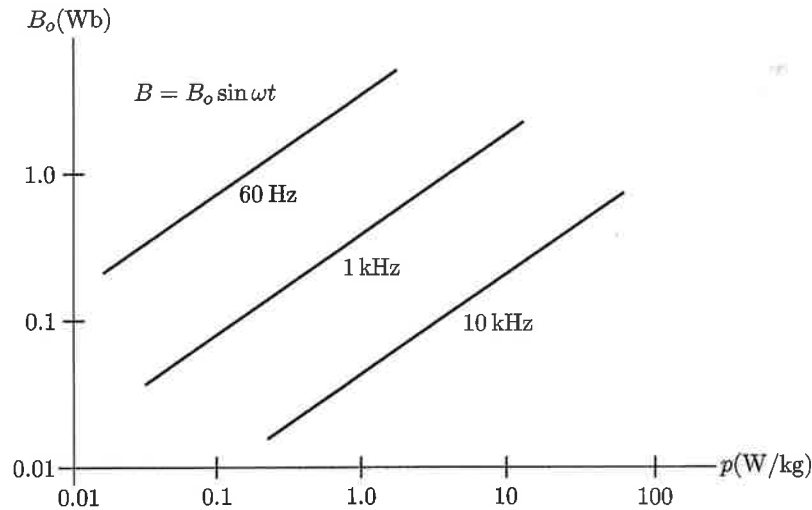


Figure 15.15 A core-loss graph for ordinary transformer steel.

PROBLEMS

15.1 A simple light dimmer circuit is shown in Fig. 15.16. When the TRIAC, Q_T , is off, C charges through R_1 and the lamp. The device Q_S is a *silicon bilateral switch* (SBS), one of a family of *trigger devices* designed to be used in thyristor gate-drive circuits. Think of it as a gateless TRIAC having symmetrical and low breakover voltages, $\pm V_{BO}$. If the magnitude of the voltage across Q_S exceeds V_{BO} , the SBS will turn on and thereafter behave as a TRIAC.

In this circuit, the dimming level is controlled by varying R_1 . Note that the dimmer is connected between the hot (H) and dimmed hot (DH) nodes and does not require access to the neutral wire (N), which is often not available at a wall switch being retrofitted with

a dimmer. The SBS has specified breakover voltages of $V_{BO} = \pm 27$ V. Determine and sketch the lamp voltage v_L . Assume that you can model the lamp as a $100\ \Omega$ resistor.

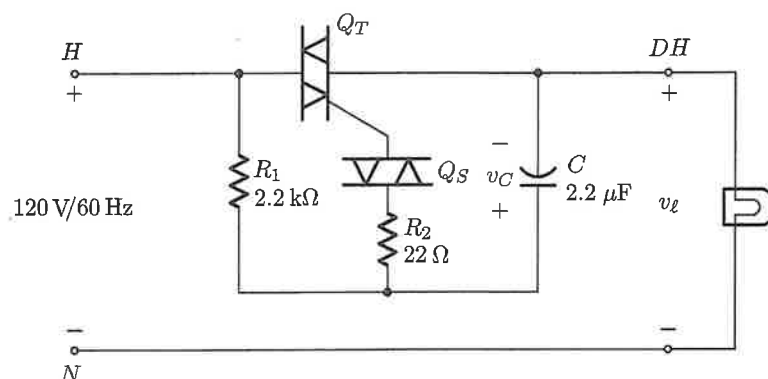


Figure 15.16 The light dimmer circuit of Problem 15.1. The TRIAC is fired by the SBS, Q_S .

- 15.2** Sketch the transistor voltage and current waveforms v_Q and i_Q for the down converter of Fig. 15.3. In what ways does the diode affect the voltage and current ratings of the transistor?
- 15.3** The converter of Fig. 15.17(a) is operating at 100 kHz, and the base drive signal to the transistor has a duty ratio of 50%. If the transistor voltage and current waveforms during switching are as shown in Fig. 15.17(b), what is the output voltage of the converter?

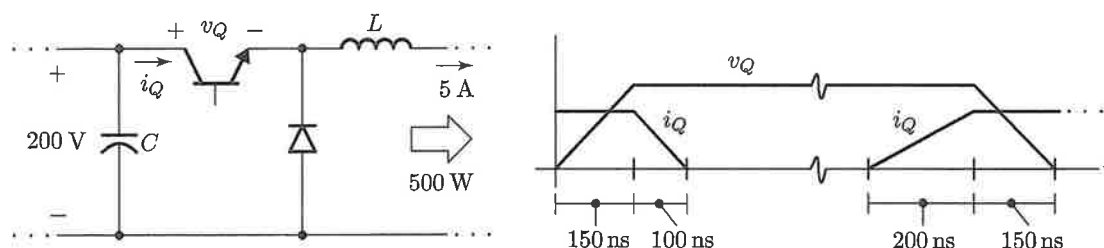


Figure 15.17 (a) The buck converter of Problems 15.3 and 15.4. (b) Voltage and current waveforms during switching for the converter of (a).

- 15.4** The buck converter of Fig. 15.17(a) is operating at a frequency of 20 kHz and delivering 500 W to the load. The diode is ideal, but the bipolar transistor has an on-state drop of 1.5 V, a leakage current of $2\ \mu\text{A}$, and nonzero switching times whose values are shown in Fig. 15.17(b). What is the efficiency of the converter? What is the efficiency if the switching frequency is increased to 50 kHz?
- 15.5** What is the maximum effective duty ratio that we can achieve in the 100 kHz converter of Problem 15.3 if we require the diode to turn on for an arbitrarily short time before the transistor turns on again?
- 15.6** In discussing the SCR, we said that the opto-SCR has a low dv_{AK}/dt rating. Explain why it is unsuitable for use as a bridge switch.

- 15.7** We stated that the current gain of a Darlington pair is *approximately* $h_{FE1}h_{FE2}$. If $h_{FE1} = h_{FE2} = 20$, what is the error in h_{FE} that results from using this approximation?
- 15.8** The diode in the buck converter of Fig. 15.18 is in series with 500 nH of parasitic lead inductance, and the diode current i_D is as shown. Calculate and plot the diode and transistor voltages v_D and v_Q and the transistor current i_Q . If the converter is operating at 200 kHz, what are the switching losses in the diode and transistor?

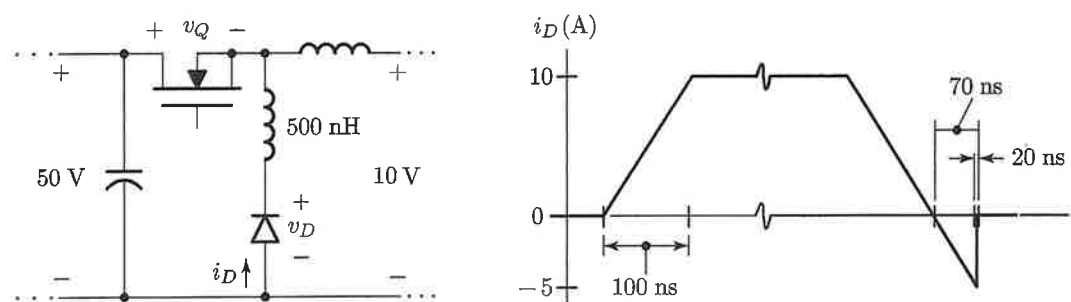


Figure 15.18 The buck converter of Problem 15.8, and its diode current waveform.

- 15.9** An inverter-grade SCR, Q , is used to excite a tank circuit, as shown in Fig. 15.19. Q has a reverse recovery time of $t_{rr} = 10 \mu\text{s}$ and a recovery characteristic so snappy that you can assume $t_f \approx 0$. The diode D and the 20Ω resistor form a snubber to maintain continuity of energy in L when Q turns off. The SCR turns on at $t = 0$. Calculate and sketch i_a and v_a .

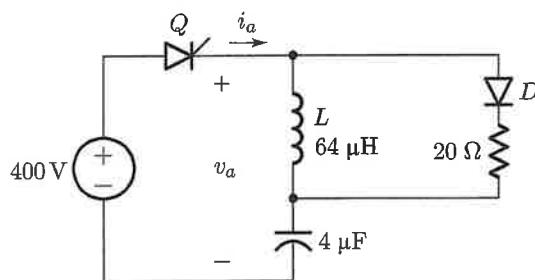


Figure 15.19 The circuit of Problem 15.9, in which v_a and i_a are calculated and sketched.

- 15.10** The circuit of Fig. 15.20 is a 5 kHz PWM inverter generating a sinusoidal short-term average voltage $\bar{v}_a$, so that $i_a = 200 \sin \omega_a t \approx \bar{v}_a/R$; that is, at the output frequency ω_a , the LR load appears resistive. The switching sequence for $\bar{v}_a > 0$ is as shown. The body diodes in Q_1 and Q_3 have reverse recovery times of $t_{rr} = 300 \text{ ns}$, and t_f is a negligible part of this time. Because of this nonzero recovery time, an inductor, $L_s = 1 \mu\text{H}$, is placed in series with the source V_d to limit the reverse recovery current when Q_1 and Q_3 toggle.

The diode D_s and resistor R_s dissipate the excess energy trapped in L_s at the end of the Q_1 or Q_3 recovery period. What is the required power rating of R_s ?

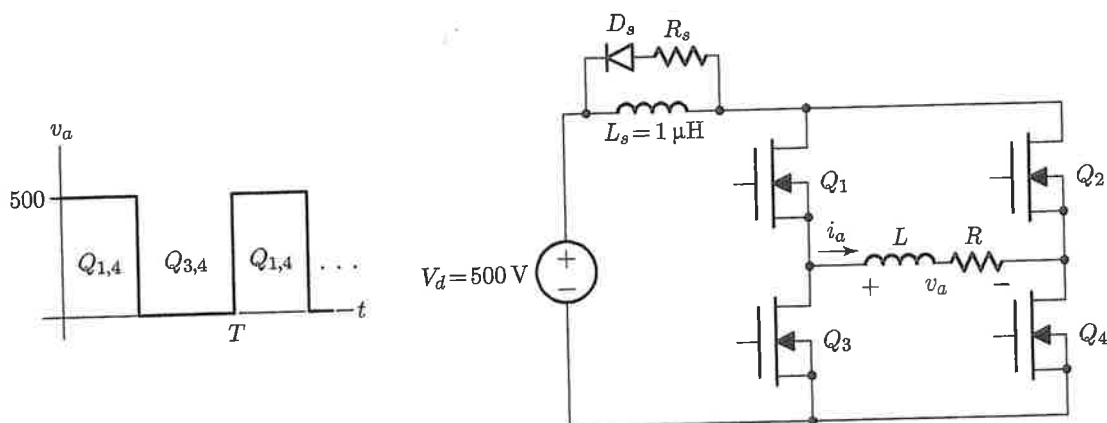


Figure 15.20 A PWM inverter with source inductance introduced to limit the peak reverse recovery currents of the Q_1 and Q_3 body diodes. Problem 15.10 considers the rating of R_s .

17

Power Semiconductor Devices

Power semiconductor devices are distinguished by the high current densities at which they operate while on, and the high voltages they must withstand when off. These requirements have serious consequences for both their physical structure and electrical behavior. In this chapter we consider their structural and behavioral departures from the ideal devices studied in Chapter 16. A significant departure from the analyses in that chapter common to almost all bipolar power diodes and transistors is that when conducting they are operating in a regime known as *high-level injection*, where the excess minority concentration becomes comparable to or greater than the equilibrium majority concentration, whereas in Chapter 16 we considered only low-level injection (excess minority concentration $\ll$ equilibrium majority concentration). High-level injection is also referred to as *conductivity modulation* for reasons to be explained.

We first consider the bipolar power diode, which contains a wide space-charge layer to support high voltage, and exhibits current and voltage transients when switching between states. We also consider the Schottky diode, which, while it cannot support the high voltages of the bipolar diode, has a lower forward drop at comparable currents and much better switching characteristics. We then look at the power BJT, whose structure, which is required to support high voltage and high current simultaneously, reduces its usable switching frequency. Next we consider the power MOSFET, including the *superjunction* MOSFET, through which current travels vertically instead of laterally as discussed in Chapter 16. We then review the IGBT as a combination of the power BJT and MOSFET, which provides a very high voltage and current capability at the expense of switching speed. We also discuss the favorable characteristics of devices fabricated in the wide-bandgap materials SiC and GaN. Following this we consider the SCR as the regenerative connection of an npn and a pnp transistor, and the TRIAC. We conclude with a description of the device datasheet and its interpretation.

17.1 Bipolar Diode

The characteristics limiting the reverse voltage rating of a bipolar diode are leakage current and avalanche breakdown. Although we discussed both phenomena in Chapter 16, a number of simplifying assumptions are not consistent with the behavior of a device designed specifically for power applications. To block a high reverse voltage, a power bipolar diode requires a wide space-charge layer at the junction. This has implications for its behavior in both forward and reverse bias. Under reverse bias the leakage current of the device is substantially larger than

that given by (16.44). And in forward bias the additional resistance caused by the increased device dimension necessitated by the wide SCL produces a resistive forward voltage drop in addition to the junction drop. This drop is mitigated by conductivity modulation of the region, but results in a forward voltage transient before the region is fully conductivity modulated. In addition, at high currents the excess charge in the neutral regions on either side of the junction is substantial and requires time (known as the *reverse recovery time*) to be removed when switching from conducting to blocking.

17.1.1 Leakage Current

In Chapter 16 we showed that a forward-biased diode whose neutral regions are in low-level injection – and are long compared to the minority carrier diffusion lengths – have the excess carrier profile shown in Fig. 17.1 and an i - v relationship of

$$J = \left[\frac{qD_h n_i^2}{N_D L_h} + \frac{qD_e n_i^2}{N_A L_e} \right] (e^{qv_A/kT} - 1) = J_s (e^{qv_A/kT} - 1). \quad (17.1)$$

In reverse bias, that is, $v_A \ll 0$, $J = -J_s$, where J_s is the leakage current density.

For Si at room temperature and under forward bias, J_s as defined by (17.1) is typically very small compared to J . For instance, if we assume that $v_A = 0.6$ V, J is 2.6×10^{10} times greater than J_s . Such a small leakage current results in negligible reverse-bias power dissipation, even when the diode is blocking several thousand volts.

However, J_s is proportional to n_i^2 , which is a strong function of temperature. That is,

$$n_i^2 \propto T^3 e^{-E_g/kT}. \quad (17.2)$$

At 25°C, $n_i^2 \approx 2 \times 10^{20}/\text{cm}^6$, but at 175°C (a typical maximum operating temperature for a power diode) $n_i^2 = 3.8 \times 10^{27}/\text{cm}^6$. This seven orders of magnitude increase gives a leakage current whose effects are no longer negligible.

Carrier Generation in the SCL and Diffusion Regions We know that not all the current in a reverse-biased diode is accounted for in (17.1). As discussed in Chapter 16, electron-hole

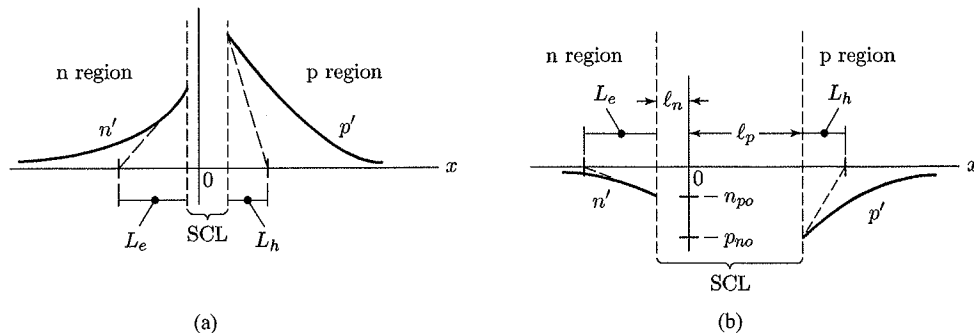


Figure 17.1 Excess minority carrier concentrations in a long diode. (a) Forward bias. (b) Reverse bias.

pairs are thermally generated throughout the device at a rate that depends on the intrinsic concentration n_i , which in turn depends on temperature. In thermal equilibrium, this generation is exactly canceled by recombination, which occurs at a rate proportional to the product of the carrier concentrations. The result is that $n_o p_o = n_i^2$ everywhere in the device. When the diode is reverse biased, however, it is no longer in thermal equilibrium. The excess carrier concentrations within the space-charge layer and the adjacent diffusion regions are substantially below their equilibrium values n_o and p_o . In these regions, then, the thermal generation rate, which is independent of n and p , exceeds the recombination rate. The carrier concentrations do not build up because the thermally generated carriers are swept out of the SCL by the intense electric field that exists there. Holes flow to the p region's contact and electrons flow to the n region's contact. These carrier flows give a component of the measured reverse-bias leakage current not accounted for by (17.1).

The generation rate in the depleted and diffusion regions is proportional to n_i , which is a function only of material and temperature. If we define a characteristic lifetime τ_{SCL} so that the generation rate is $G = n_i/\tau_{\text{SCL}}$, the total current density caused by thermal generation in these regions is

$$J_{\text{SCL}} \approx q(L_e + L_h + \ell_p + \ell_n)G = \frac{q(L_e + L_h + \ell_p + \ell_n)n_i}{\tau_{\text{SCL}}}, \quad (17.3)$$

where $(L_e + L_h + \ell_p + \ell_n)$ is the width of the SCL and adjacent diffusion regions.

Because the width of the SCL and the diffusion regions are so long in a power device, we make τ_{SCL} as large as possible. This lifetime is related to, but is usually longer than, the low-injection-level minority-carrier lifetimes τ_e and τ_h . In general, J_{SCL} is much larger than J_s at room temperature, but at high operating temperatures J_s is larger.

Example 17.1 Calculation of Leakage Current Components

For the diode described by (17.1), assume that $\tau_e = \tau_h = 0.2 \mu\text{s}$, $\tau_{\text{SCL}} = 10 \mu\text{s}$, and $N_A = N_D = 3 \times 10^{15}/\text{cm}^3$. These doping concentrations are such that for a reverse voltage of $v_A = -100 \text{ V}$, the SCL of this symmetric diode extends approximately $4.6 \mu\text{m}$ into both the n and p regions, that is, $\ell_p + \ell_n = 9.2 \mu\text{m}$.

We calculate J_s at 25°C using (17.1) and parameters from Table 16.1, obtaining $2.4 \times 10^{-10} \text{ A/cm}^2$. For this value of J_s , a forward-biased current density of 3 A/cm^2 requires $v_A = 0.58 \text{ V}$. This current density is very low compared to the typical current density for a power diode but is about the level at which this diode will no longer be in low-level injection. Although we do not yet have an accurate model for the diode in high-level injection, we can expect it to carry a forward current density of 50 A/cm^2 , resulting in $v_A \approx 1 \text{ V}$. Operating at this current density, the diode's dissipation is 50 W/cm^2 .

The minority carrier diffusion lengths ($\sqrt{D\tau}$) in this diode are $L_e = 26 \mu\text{m}$ and $L_h = 15 \mu\text{m}$. The total width of the SCL and diffusion regions is therefore approximately $50 \mu\text{m}$ at -100 V . At room temperature, this width yields a thermally generated leakage current $J_{\text{SCL}} = 1.12 \times 10^{-6} \text{ A/cm}^2$. Although this leakage current is much larger than J_s , it causes a reverse-bias power dissipation of only 0.112 mW/cm^2 , which is negligible compared to the forward-bias dissipation.

At 175°C , $J_s = 4.6 \times 10^{-3} \text{ A/cm}^2$ and $J_{\text{SCL}} = 5.4 \times 10^{-3} \text{ A/cm}^2$. With these currents, the reverse-bias dissipation is 1 W/cm^2 , which is still small but is quickly becoming important. At 200°C , $J_s = 28 \times 10^{-3} \text{ A/cm}^2$, $J_{\text{SCL}} = 13 \times 10^{-3} \text{ A/cm}^2$, and the reverse dissipation is 30 W/cm^2 .

A precise quantitative prediction of leakage current requires consideration of effects in addition to those discussed here. For instance, temperature affects lifetimes, and generation

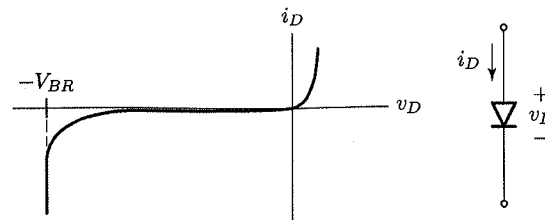


Figure 17.2 The i - v characteristic of a power diode, showing the fast rise in reverse leakage current when the avalanche breakdown voltage $-V_{BR}$ is reached.

within the SCL is not uniform. However, we can draw some important conclusions from the preceding analysis:

- Generation in the wide SCL and diffusion regions of a power device will yield substantially more leakage current at 25°C than is predicted by the classic diffusion model, (17.1).
- To keep the leakage current acceptably small, the minority carrier lifetime in the SCL should be as long as possible.
- The component of leakage current resulting from thermal generation in the SCL, J_{SCL} , grows as n_i , which approximately doubles for every 11°C increase in T between -50°C and 200°C . But J_s grows as n_i^2 , so it eventually dominates at high temperatures. This increase in reverse leakage current limits the maximum operating junction temperature of a power device.

17.1.2 Avalanche Breakdown

We assumed in Chapter 16 that applying an electric field to a semiconductor causes the holes and electrons to flow with a velocity proportional to the electric field. Actually, the particles frequently collide with the lattice imperfections (impurity atom sites, physical defects, or phonons). What we really meant was that the *average* velocity of the particles is proportional to the electric field. If the electric field is strong enough, the velocity of a particle becomes large enough that, during a collision, energy is imparted to the lattice to create an electron-hole pair. This process is called *impact ionization*. It is a multiplicative process: the newly generated particles are accelerated by the field, collide with the lattice, and create additional electron-hole pairs, and so on.

As the reverse bias voltage across a diode increases, so does the peak electric field in its SCL. In Si, when the field reaches a magnitude of about $3 \times 10^5 \text{ V/cm}$, the impact ionization process begins. This field is up to an order of magnitude larger for wide-bandgap materials like SiC or GaN. The number of electron-hole pairs generated in the SCL by impact ionization grows very quickly with increasing electric field above this value. Because these generated carriers add to the thermally generated ones to give more reverse leakage current, the diode's reverse bias characteristic looks like that shown in Fig. 17.2. When impact ionization reaches an infinite rate, we say that the diode is in *avalanche breakdown*.

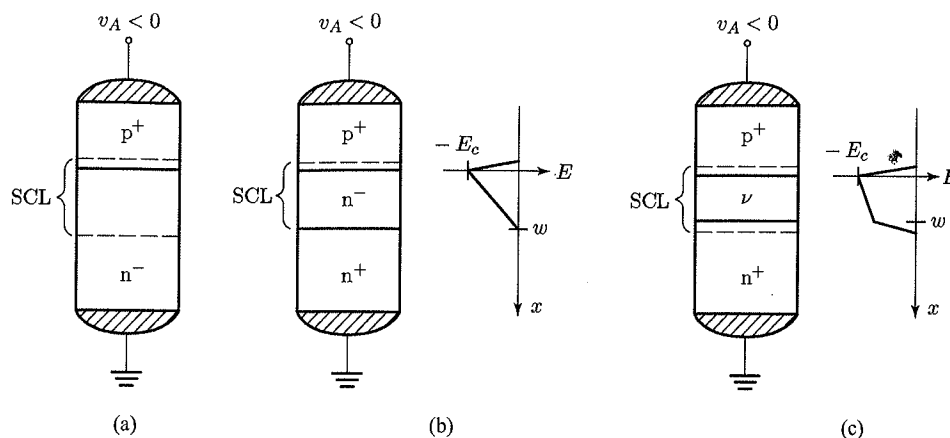


Figure 17.3 Evolution of the pin diode structure. (a) Conventional asymmetric diode with a large n^- substrate resistance. (b) The diode of (a) with the neutral part of the n^- region replaced by a low-resistance n^+ region. (c) The diode of (b) with the n^- region replaced by a very lightly n-doped ν region, producing a pin (or $p\nu n$) diode. The resulting field distributions for (b) and (c) are also shown.

The onset of avalanche breakdown occurs approximately when the peak electric field in the SCL reaches a critical value, E_c . This statement is not strictly accurate, because the avalanche process requires that impact ionization occur over some minimum distance so the multiplication process can sustain itself. Because the gradient of the electric field is much greater in the SCL of a heavily doped diode than in the SCL of a lightly doped diode, the distance over which ionization occurs is shorter. The peak electric field can therefore be greater in a heavily doped diode before avalanche begins. For a doping level of $10^{16}/\text{cm}^3$ (typical of a 40–50 V device), $E_c \approx 3.7 \times 10^5 \text{ V/cm}$, but for a doping level of $10^{14}/\text{cm}^3$ (typical of a 800–1000 V device), $E_c \approx 2.1 \times 10^5 \text{ V/cm}$.

17.1.3 pin Diode

A reverse-biased, asymmetrically doped diode supports nearly all the voltage across the SCL in the lightly doped n^- region. Making a vertical power diode by simply diffusing the p^+ region into an n^- substrate, as shown in Fig. 17.3(a), would result in a large n^- substrate resistance and a large resistive component of the forward drop. The reason is that the substrate must be about 500 μm thick to maintain mechanical integrity, which introduces a large series resistance. However, because the SCL has a width of between 10 and 200 μm , sandwiching a relatively thin n^- layer between p^+ and n^+ layers, as shown in Fig. 17.3(b), can lower the resistive component of the forward drop. The figure shows the field distribution when the voltage is such that the critical field, E_c , occurs at the junction.

Doping the n^- region so lightly that it is nearly intrinsic (a doping level well below n^- sometimes referred to as ν) further improves performance. The resulting electric field magnitude is shown in Fig. 17.3(c), with v_A such that the field is again E_c at the junction. Because the field

distribution is almost rectangular, the i region width is half that of the n^- region of Fig. 17.3(b) for the same voltage rating. The diode of Fig. 17.3(c) is called a *pin diode*. In a practical device, the intrinsic region is really very slightly n -type (ν) or p -type (π), and the device is known as a $p\nu n$ or $p\pi n$ diode. Almost all power diodes have this type of structure, which we refer to generically as a pin diode. The highly doped n^+ region quickly terminates the field and is sometimes known as a *field stop* layer.

Example 17.2 Electric Field and Breakdown Voltage of a pin Diode

If the doping concentration in the n^- region of the diode shown in Fig. 17.3(b) is $N_D = 2 \times 10^{14}/\text{cm}^3$, what is the minimum width, w , of this region resulting in the maximum breakdown voltage? We assume that the SCL ends at the $n^- - n^+$ junction, so all the voltage is dropped across the n^- region. When the device just reaches avalanche breakdown, the field is $-E_c$ at the $p^+ - n^-$ junction and 0 at the $n^- - n^+$ junction, as shown in the figure. Using Gauss's law and $\epsilon_{Si} = 11.8\epsilon_0$, we determine $E_{n^-}(x)$, the field in this region:

$$\frac{dE_{n^-}}{dx} = + \frac{qN_D}{\epsilon_{Si}} = -3.08 \times 10^7 \text{ V/cm}^2, \quad (17.4)$$

$$E_{n^-}(x) = -E_c + 3.08 \times 10^7 w. \quad (17.5)$$

The width w necessary to create this condition is determined by setting (17.5) to zero. For a doping concentration of $N_D = 2 \times 10^{14}$, the critical field $E_c = 2 \times 10^5 \text{ V/cm}$, giving $w = 65 \mu\text{m}$. Since the E field is triangular, V_{BR} is easily calculated:

$$V_{BR} = \left(\frac{1}{2}\right) E_c w = \left(\frac{1}{2}\right) (2 \times 10^5 \text{ V/cm})(66 \mu\text{m}) = 660 \text{ V}. \quad (17.6)$$

If we now reduce N_D to $10^{12}/\text{cm}^3$ to create the ν region shown in Fig. 17.3(c), what is the resulting width w required to support the same 660 V? The field in the ν region is now trapezoidal, and if we ignore the voltage contributions of the narrow SCLs in the p^+ and n^+ regions, the resulting field in the ν region is

$$E_\nu(x) = -E_c + 1.546 \times 10^5 x. \quad (17.7)$$

We calculate the required width by integrating E_ν and solving for $x = w$ at $V = 660 \text{ V}$. The resulting width is $w = 33 \mu\text{m}$, which is a substantial reduction.

17.1.4 Forward Bias

We based the analysis leading to the pn diode $i-v$ characteristic, (17.1), on the assumption that both sides of the junction were in low-level injection. This assumption is not valid for the intrinsic region of a pin diode because in this region $n \approx n' \gg n_0$ and $p \approx p' \gg p_0$ at any reasonable injection level, and since neutrality requires $n' = p'$, $n \approx p$ in the neutral regions. When the total carrier concentration in a region is dominated by the excess concentration, as in this case, we say that the region is in *high-level injection*, also called conductivity modulation, because the conductivity is no longer determined by the majority doping concentration – it is now a function of injection level.

The diode we choose for analysis is shown in Fig. 17.4, along with its minority carrier concentrations. We assume that the width of the intrinsic region W_i is comparable to the characteristic diffusion length. The intrinsic region is in high-level injection everywhere, but the

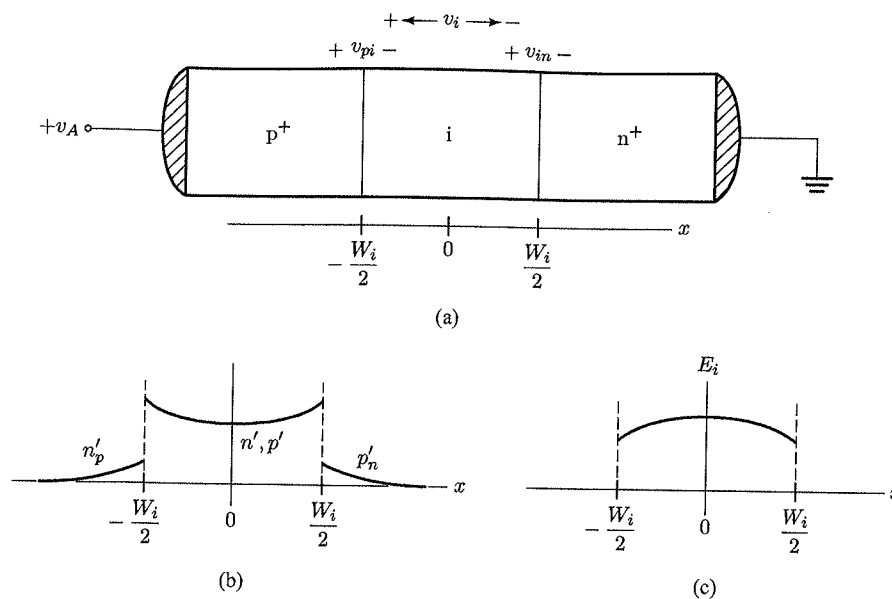


Figure 17.4 (a) A one-dimensional pin diode structure. (b) Carrier concentrations with the intrinsic region in high-level injection. (c) The electric field in the intrinsic region.

heavily doped p⁺ and n⁺ regions remain in low-level injection. Because the end regions are so heavily doped, the diffusion of minority carriers in these regions is negligible, as it was in the heavily doped side of the asymmetric diode analyzed in Chapter 16. Only hole current flowing by drift is significant in the p⁺ region, and only electron current flowing by drift is significant in the n⁺ region. These holes and electrons flow into the intrinsic region where they recombine. It is this recombination that determines the terminal current.

Because $n' = p'$, in high-level injection the recombination rate for holes and electrons is the same and is characterized by the *ambipolar lifetime*, τ_a . Furthermore, as the majority and minority concentrations are the same, we can no longer make the approximation that the minority current flows only by diffusion. The continuity equations in the intrinsic region must now contain the drift term that we eliminated in our analysis of the diode in low-level injection. Including the drift term, the continuity equations are

$$\begin{aligned} \frac{\partial n}{\partial t} &= \frac{1}{q} \frac{\partial J_e}{\partial x} - \frac{n'}{\tau_a} \\ &= \mu_e \left(n \frac{\partial E}{\partial x} + E \frac{\partial n}{\partial x} \right) + D_e \frac{\partial^2 n}{\partial x^2} - \frac{n'}{\tau_a}, \end{aligned} \quad (17.8)$$

$$\begin{aligned} \frac{\partial p}{\partial t} &= -\frac{1}{q} \frac{\partial J_h}{\partial x} - \frac{p'}{\tau_a} \\ &= -\mu_h \left(p \frac{\partial E}{\partial x} + E \frac{\partial p}{\partial x} \right) + D_h \frac{\partial^2 p}{\partial x^2} - \frac{p'}{\tau_a}. \end{aligned} \quad (17.9)$$

The concentrations are equal to each other and dominated by their excess in high-level injection, that is, $n \approx n' = p \approx p'$. Recognizing this, (17.8) and (17.9) can be rewritten as

$$\frac{\partial n'}{\partial t} = \mu_e \left(n' \frac{\partial E}{\partial x} + E \frac{\partial n'}{\partial x} \right) + D_e \frac{\partial^2 n'}{\partial x^2} - \frac{n'}{\tau_a}, \quad (17.10)$$

$$\frac{\partial n'}{\partial t} = -\mu_h \left(n' \frac{\partial E}{\partial x} + E \frac{\partial n'}{\partial x} \right) + D_h \frac{\partial^2 n'}{\partial x^2} - \frac{n'}{\tau_a}. \quad (17.11)$$

Multiplying (17.10) by D_h and (17.11) by D_e and adding the two equations gives

$$\begin{aligned} (D_e + D_h) \frac{\partial n'}{\partial t} &= (\mu_e D_h - \mu_h D_e) \left(n' \frac{\partial E}{\partial x} + E \frac{\partial n'}{\partial x} \right) \\ &\quad + 2D_e D_h \frac{\partial^2 n'}{\partial x^2} - (D_e + D_h) \frac{n'}{\tau_a}. \end{aligned} \quad (17.12)$$

Substituting for μ_e and μ_h in (17.12) using the Einstein relation, $\mu = (q/kT)D$, and simplifying yields

$$(D_e + D_h) \frac{\partial n'}{\partial t} = 2D_e D_h \frac{\partial^2 n'}{\partial x^2} - (D_e + D_h) \frac{n'}{\tau_a}. \quad (17.13)$$

The continuity equations in high-level injection are

$$\frac{\partial n'}{\partial t} = D_a \frac{\partial^2 n'}{\partial x^2} - \frac{n'}{\tau_a}, \quad (17.14)$$

$$\frac{\partial p'}{\partial t} = D_a \frac{\partial^2 p'}{\partial x^2} - \frac{p'}{\tau_a}, \quad (17.15)$$

where D_a is the *ambipolar diffusion constant*,

$$D_a = \frac{2D_e D_h}{D_e + D_h}. \quad (17.16)$$

Because $n' = p'$, (17.14) and (17.15) are identical. We solve them to determine the spatial dependencies of the carrier concentrations and the diffusion components of the currents in the intrinsic region. As in our previous analyses, the concentrations vary exponentially with x . But the characteristic diffusion length, called the *ambipolar diffusion length*, $L_a = \sqrt{D_a \tau_a}$, is now the same for both holes and electrons.

There is also a drift component of current in the intrinsic region, and the drift field gives rise to a voltage drop, v_i , which adds to the on-state voltage of the diode. Using a straightforward calculation, we can obtain the value of this drift field, E_i :

$$J_e = q\mu_e \left(\frac{kT}{q} \frac{dn'}{dx} + n'E_i \right), \quad (17.17)$$

$$J_h = q\mu_h \left(-\frac{kT}{q} \frac{dp'}{dx} + p'E_i \right). \quad (17.18)$$

Because $J_e + J_h = J$ (the total diode current), which is constant throughout the device, and $n' = p'$,

$$E_i = -\frac{kT}{q} \left(\frac{\mu_e - \mu_h}{\mu_e + \mu_h} \right) \frac{1}{n'} \frac{dn'}{dx} + \frac{J}{q(\mu_e + \mu_h)n'}. \quad (17.19)$$

The first term in (17.19) is the field necessary to ensure quasi-neutrality when $\mu_e \neq \mu_h$ (and, therefore, $D_e \neq D_h$). The second term is simply the ohmic drop in the conductivity modulated intrinsic region. The conductivity of this region now depends on injection level, and is $\sigma_i = q(\mu_e + \mu_h)n'$.

An interesting consequence of high-level injection is that the electric field E_i is independent of injection level and therefore of current. We can show this condition as follows. The solution to the diffusion equation, (17.14), is a pair of exponentials with characteristic length L_a . Thus, the first term in (17.19), in which the spatial derivative of n' is divided by n' , does not depend on n' . And as we assumed that all the current flowing through the diode results from recombination within the intrinsic region – and as this recombination is proportional to the excess carrier concentration – we know that $J \propto n'$. The second term in (17.19) is therefore also independent of n' . Thus, the electric field E_i , and therefore the voltage v_i , are dependent only on physical parameters and the length of the intrinsic region.

The i - v characteristic of the pin diode of Fig. 17.4(a) can be determined by assuming that $D_e = D_h = D_a$ and therefore $\mu_e = \mu_h$ in the intrinsic region. This is a reasonably accurate assumption in regions of very high-level injection. Because the p^+ and n^+ end regions of the diode are heavily doped, their conductivity is high and we can ignore the voltage drop across these regions. With these assumptions the terminal voltage v_A is the sum of the junction drop and the drop v_i across the wide intrinsic region. Although we do not give the details of the calculation here, the intrinsic region drop is

$$v_i = \int_{-W_i/2}^{W_i/2} E_i dx = \frac{2kT}{q} \sinh\left(\frac{W_i}{2L_a}\right) \tan^{-1} \left[\sinh\left(\frac{W_i}{2L_a}\right) \right]. \quad (17.20)$$

For $W_i/2 \leq L_a$,

$$v_i \approx \frac{2kT}{q} \left(\frac{W_i}{2L_a} \right)^2, \quad (17.21)$$

and for $W_i/2 \gg L_a$,

$$v_i \approx \frac{\pi kT}{2q} e^{(W_i/2L_a)}. \quad (17.22)$$

If we assume v_i is small enough to ignore, the v - i relationship of a pin diode is

$$J = \frac{2qL_a n_i}{\tau_a} \tanh\left(\frac{W_i}{2L_a}\right) (e^{qv_A/2kT} - 1). \quad (17.23)$$

Note that the current is now proportional to $e^{v_A/2}$. Using $L_a = \sqrt{D_a \tau_a}$ to substitute for τ_a results in a relationship similar to that for the short diode of Example 16.3, the major difference being that for a given current the pin diode produces twice the forward drop of the short diode. We have also ignored any ohmic drop in the intrinsic region, which, depending on temperature and the width of the region, would produce an additional drop not present in the short diode.

Example 17.3 Forward Drop in the Intrinsic Region

We consider a 600 V, 100 A pin diode with an intrinsic region width W_i of 240 μm and an ambipolar carrier lifetime of 1 μs . The ambipolar diffusion constant for Si at 300 K is approximately 15 cm^2/s , yielding a diffusion length $L_a = \sqrt{D_a \tau_a}$ of 40 μm . Since $W_i/2 > L_a$, we calculate the drop using (17.22), which gives a value of 0.79 V. This is a significant additional forward drop contributed by the field in the intrinsic region.

17.1.5 Additional Effects of High Current Densities

Two phenomena reduce the ambipolar diffusion length L_a at the high carrier concentrations reached when the diode's current density gets very large. One is *carrier-carrier scattering*, which reduces the mobility of the carriers and therefore the ambipolar diffusion constant D_a . The other is *Auger recombination*, which reduces the ambipolar lifetime τ_a . As the diode current increases, the corresponding reduction in L_a caused by these two effects makes the voltage across the intrinsic region v_i grow quickly, as expressed by (17.21) or (17.22).

The injection-level dependence of L_a , the increased significance of end-region diffusion currents, and the increase of the $n_o p_o$ product cause the i - v characteristic of a power diode, (17.23), to bend at high currents. Figure 17.5 shows the curve that results from these effects.

17.1.6 Temperature Dependence of J

We show in (17.23) that the diode current is proportional to n_i if the injection level is not too high. The temperature dependence of n_i is

$$n_i \propto T^{3/2} e^{-E_g/2kT}. \quad (17.24)$$

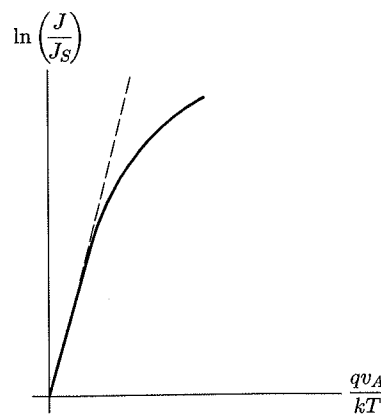


Figure 17.5 The i - v characteristic of a typical power diode, showing the curve bending at high currents because of high-level injection effects.

Because n_i increases more rapidly than $e^{q\psi_j/kT}$ decreases with increasing temperature, J increases with temperature for a given L_a and τ_a . Conversely, the forward drop decreases with temperature if the forward current is fixed. Although we do not discuss here how L_a and τ_a vary with temperature, the effect of these additional temperature dependencies is to increase the current flowing through a pin power diode with temperature for a given applied voltage.

17.2 Switch Transitions in a pin Diode

We now turn to the transitions of a diode between its static on and static off states. For simplicity, we ignore the capacitance of the SCL and focus on the dynamics of the excess stored charge in the neutral regions outside the SCL. We can easily add the effects of the SCL to those we describe here. We also ignore the change in the lifetime in the intrinsic region as the device goes from no excess carriers to high-level injection, and the effect of the doping concentration in the intrinsic region on the nature of the transitions.

The two salient phenomena during switching are the *forward recovery* voltage and the *reverse recovery current*. While all bipolar diodes exhibit these phenomena, they are greatly aggravated in power diodes because of their wide SCL.

17.2.1 Forward Recovery

We consider first the turn-on transition, in which the diode current steps instantly from zero to I_F . As Fig. 17.6(a) shows, the terminal voltage v_A first rises to a value V_{fp} that is much higher than the static forward drop. It then decays to the steady-state value v_F , determined by the current I_F . This process is called *forward recovery*, and V_{fp} is known as the *forward recovery voltage*. The duration of the forward recovery process is called the forward recovery time t_{fr} .

We can explain the overshoot in the diode voltage by considering how the carrier concentrations in the intrinsic region change during the turn-on transition. Initially, there are no excess carriers. However, as the forward current flows, holes drift through the p^+ region and are injected into the intrinsic region. Similarly, the electrons drift through the n^+ region and are injected into the intrinsic region. Over time, these injected carriers accumulate, as indicated by the profiles shown in Fig. 17.6(b).

Consider the carrier profiles at t_1 , a short time after the forward current starts flowing. As Fig. 17.6(b) shows, the carriers build up near the junctions but not in the center of the intrinsic region. Some of the holes injected into the intrinsic region at the p^+ -i junction add to the accumulation of carriers at this junction, and the rest flow across the intrinsic region to add to the accumulation of carriers at the i- n^+ junction. Similarly, some of the electrons injected at the i- n^+ junction flow across the intrinsic region to add to the accumulation at the p^+ -i junction.

How do these injected carriers flow across the intrinsic region? Not by diffusion, because at this time there is no gradient to the carrier profiles in most of the i region. Therefore, flow in the middle of the i region must be by drift. The voltage that results from this drift is very large. The reason is that the i region is not in high-level injection, so its conductivity, $\sigma_i = q\mu_e n + q\mu_h p$, is

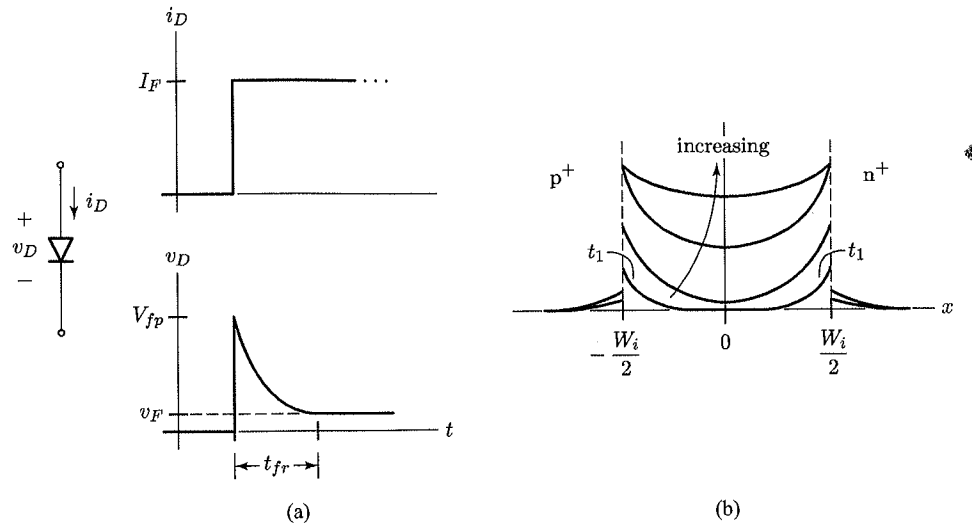


Figure 17.6 Forward recovery of a pin diode. (a) Typical diode voltage and current waveforms during the turn-on transition. (b) The time evolution of excess carrier concentrations in the neutral regions of the device.

very low and yet the region is carrying the full current I_F . Even if the region is not intrinsic, but is a lightly doped ν or π region, its resistance is high compared to its value upon reaching high-level injection. This resistance is the source of the peak transient voltage V_{fp} of the waveforms of Fig. 17.6(a).

As time passes, the growing carrier concentrations in the middle region modulate the region's conductivity and reduce its resistance. As a result, the voltage across the middle region drops. When the carrier concentrations reach their final value, so too does the diode's voltage. The charge in the middle region at steady state is proportional to the forward current ($I_F \propto Q/\tau_a$). Hence, the time required to supply this charge – in other words, the duration of the forward recovery transition – is approximately τ_a .

Example 17.4 Magnitude of the Forward Recovery Voltage

Assume that we have a pvn diode with the center region doped at $N_D = 10^{14}/\text{cm}^3$ and designed to withstand 400 V. This voltage requires a ν -region width of $W_\nu \approx 20 \mu\text{m}$. The cross-sectional area of the device is $A = 0.5 \text{ cm}^2$, and its forward current steps from zero to 50 A at $t = 0$. What is V_{fp} ?

Initially, the resistance of the ν region is

$$R_\nu = \frac{W_\nu}{\sigma A} = \frac{W_\nu}{q(\mu_e n + \mu_h p)A} = 185 \text{ m}\Omega, \quad (17.25)$$

where $n = N_D$, $p = N_A = n_i^2/N_D$ is negligible because it is several orders of magnitude less than N_D , and $\mu_e \approx 1350 \text{ cm}^2/\text{Vs}$ at this concentration. With $I_F = 50 \text{ A}$, the drop across this region is 9.25 V, which is the diode's peak forward recovery voltage V_{fp} .

17.2.2 Reverse Recovery

Consider the diode in the circuit of Fig. 17.7(a), where, for $t < 0$ and ignoring the forward diode drop, it is carrying a forward current of $I_F = V_F/R_F$. At $t = 0$, the switch connects the diode to the negative voltage V_R through a second resistor, R_R . Because the excess charge in the i region and diffusion regions of the diode cannot change instantaneously, the p^+i and $i-n^+$ junctions remain forward biased for some time after $t = 0$. As the diode voltage is approximately zero during this time relative to V_R , the diode current is negative and equals $-V_R/R_R$. This reverse current aids the removal of excess charge, until the concentrations at the SCL edges become negative and the junction can begin to support a reverse voltage. This process is called *reverse recovery*.

Figure 17.7(b) shows how the carrier profiles in the intrinsic region change as the reverse recovery current flows. Note that, just after $t = 0$, the excess carrier concentrations at the junction edges are still positive, and therefore the junction voltages are also positive. To support the negative current, the excess carrier distributions develop a negative slope near the junction edges. Holes diffuse out of the intrinsic region, are injected into the p^+ region, and drift to the left ohmic contact. Similarly, electrons leave the intrinsic region at the $i-n^+$ junction and drift to the right contact. So long as the excess concentrations at the SCL edges are greater than zero, and $|V_R| \gg v_F$, the total diode current equals $-V_R/R_R$. Continuing recovery current flow eliminates the excess carrier concentrations in the intrinsic region, as the profiles in Fig. 17.7(c) show.

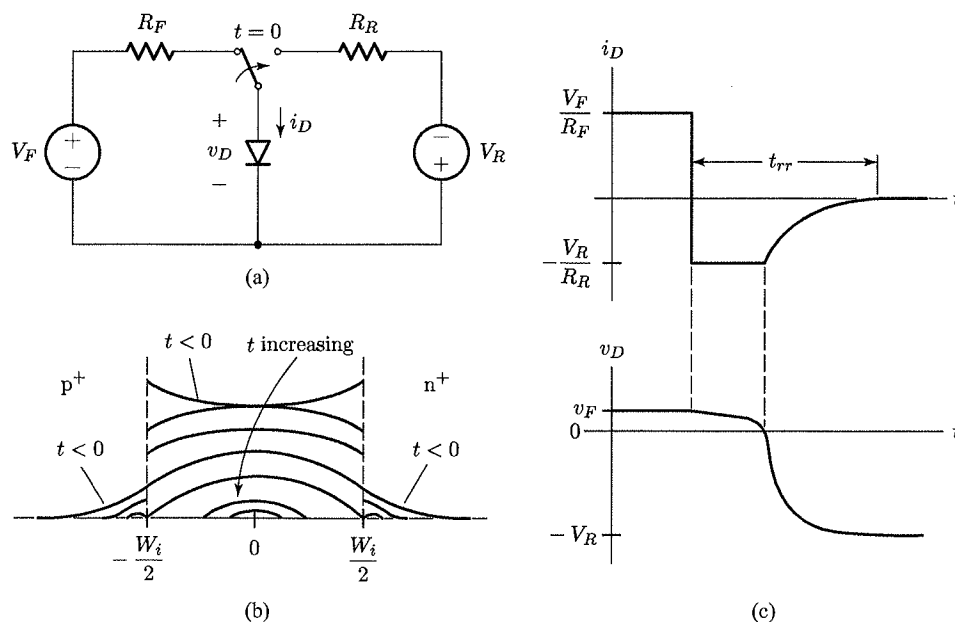


Figure 17.7 (a) A hypothetical circuit to illustrate pin diode reverse recovery. (b) Carrier profiles in the intrinsic region during the turn-off transition. (c) Waveforms showing reverse diode current during the reverse recovery time t_{rr} .

After the excess carrier concentrations at the junction edges reach zero, the concentration gradients decrease and the diode current can no longer be maintained at $i_D = -V_R/R_R$. As $|i_D|$ decreases, the drop across R_R also decreases. The difference between V_R and the voltage across R_R appears across the diode SCLs, which begin to widen. The dynamics of this process produce an approximately exponential rise of i_D to zero and fall of v_D to $-V_R$, as Fig. 17.7(c) shows. During the initial phase of the reverse recovery process, the diode voltage changes only slightly from its static forward value v_F (the small decrease results from the change in sign of v_i that occurs when the current changes direction).

The duration of reverse current flow in the diode is t_{rr} , the reverse recovery time. Based on our discussion so far, t_{rr} is directly proportional to the reverse current I_R and the charge Q that is initially stored in the intrinsic region. However, reverse current does not remove all of this stored charge; some of it recombines. How much depends on the magnitude of the negative current relative to the forward current. The initial stored charge is $Q = I_F \tau_a$. If $I_R \gg I_F$, almost all the stored charge is swept from the intrinsic region in a time much less than τ_a , and only a small amount will be removed by recombination. But if $I_R \leq I_F$, a substantial fraction of the charge will recombine. In the limit of $I_R = 0$, when all the charge is removed through recombination, the reverse recovery time is $t_{rr} \approx \tau_a$.

From the preceding qualitative discussion of the turn-on and turn-off transitions of a pin diode, we can conclude that the less charge the diode stores for a given forward current, the faster the diode will switch. As $Q = I_F \tau_a$, the way to store less charge is to shorten the lifetime. However, doing so reduces L_a , resulting in less conductivity modulation of the intrinsic region and a higher on-state drop, v_i , across it, as given by (17.21) and (17.22). For this reason, the forward drop of a fast bipolar diode is typically greater than that of a slower bipolar diode.

17.3 The Schottky Barrier Diode

When metal leads are attached to a semiconductor device, the interface between the metal and semiconductor is usually ohmic. An *ohmic contact* is one in which the resistances of the contacting materials govern current flow. Obtaining a good ohmic contact between a semiconductor and a metal is difficult. A great deal of research has been directed at developing an understanding of this interface and the metallurgy necessary to produce ohmic contacts. When we fail to pay attention to these processes, the result is a contact with a poor rectifying characteristic, which was the basis for early “cat whisker” crystal receivers and point-contact transistors. However, making a good, large-area, rectifying contact is even more difficult than making a good ohmic contact. A diode made with a metal–semiconductor junction is called a *Schottky barrier diode*, or SBD. The symbol for a Schottky diode is shown in Fig. 17.8.

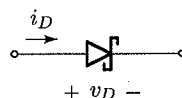


Figure 17.8 The circuit symbol for a Schottky barrier diode.

In principle, either n-type or p-type semiconductor material can be used to make a Schottky diode, although n-type material is used because in Si electron mobility is greater than that for holes. SBDs can be fabricated in Si, SiC, GaN, or GaAs, but we address here only Si and SiC devices. Silicon SBDs typically have low reverse blocking voltage ratings – with maximums on the order of 100 V – while SiC SBDs are available with ratings near 2 kV. As we demonstrate in this section, Schottky diodes offer two advantages over bipolar diodes. The first is that, for the same current density, the forward voltage drop of a Si SBD is several tenths of a volt less than that of a Si p–n diode. Although, because of its larger bandgap, the forward drop of a SiC p–n diode ($\sim 3\text{--}3.5\text{ V}$) is much higher than that of a Si p–n diode, the forward drop of a SiC SBD is similar to that of a Si p–n diode, making the SBD an attractive application of SiC. The second is that the current in a Schottky flows only by drift, so there is no need to accumulate or remove excess carriers, thereby eliminating forward and reverse recovery phenomena.

The disadvantage of Schottky diodes is that they have more leakage current when reverse biased than do bipolar diodes. This is especially true for Si Schottkys. Schottky diodes using SiC, however, can have leakage currents in the μA range at reverse blocking voltages in excess of 1 kV. The parameters of SiC Schottkys are also less temperature dependent than those for Si devices, especially the forward voltage drop. However, the forward drop for a SiC device is usually higher than for a Si device.

17.3.1 The Schottky Barrier

Current flow across a rectifying metal–semiconductor junction is the result of *thermionic emission* of electrons from the metal into the semiconductor and from the semiconductor into the metal. In thermal equilibrium, these two flows are equal and there is no net current. When a voltage is applied to the junction, one direction of electron emission dominates and a net electron current results.

Because free electrons in the semiconductor are more energetic than those in the appropriately chosen metal, more electrons initially flow into the metal. They leave behind a positive space charge layer of donor atoms and create a compensating negative charge on the metal's surface.

17.3.2 Forward Bias

We can now determine the effect of an applied junction voltage, v_j , on the flow of electrons across the Schottky barrier junction. As for a bipolar diode, v_j will change the SCL width. The junction is highly asymmetric, so nearly all of v_j appears across the semiconductor side of the SCL. If $v_j > 0$ (that is, the diode is forward biased), the electric field of SCL origin is reduced, reducing the potential barrier to electron emission from the semiconductor. However, the barrier to emission from the metal has been changed only slightly. The result is a forward current that is a strong function of v_j and consists of electrons flowing from the semiconductor to the metal.

Although we do not show the derivation here, the i – v characteristic of a *forward-biased* Schottky diode is

$$J_F = J_o \left(e^{qv_j/kT} - 1 \right), \quad (17.26)$$

which is identical in form to the characteristic of a bipolar diode. However, the current coefficient J_o is of a substantially different form from J_s , the leakage current as defined in (17.1) for a bipolar diode. The current coefficient J_o has different values in forward and reverse bias, as we show later.

The current coefficient J_o for a forward-biased silicon Schottky diode is typically four to six orders of magnitude larger than J_s for a silicon bipolar diode. For this reason, its forward drop is 0.23–0.34 V lower than that of a bipolar diode operating at the same current density. This lower on-state voltage is a Schottky's major advantage.

Forward Voltage Drop The forward drop of a Schottky diode contains a component in addition to v_j that is equal to the voltage across the semiconductor. The resistance of this region has two parts: (i) the lightly doped material into which the SCL grows in reverse bias; and (ii) the heavily doped substrate. Figure 17.9 illustrates this structure. For low-voltage Schottkys (< 40 V), these two resistances are comparable because, even though the lightly doped region has a higher resistivity, the heavily doped region is longer (typically 500 μm versus 5 μm). The total resistance of a 1 cm^2 silicon die might be 0.5–1 $\text{m}\Omega$, giving a drift region drop of 50–100 mV at a current of 100 A.

The resistance of the lightly doped region quickly dominates the resistance of the drift region as the Schottky voltage rating increases from approximately 40 V. The reason is that a higher voltage rating requires a longer lightly doped region and a lower doping concentration, both of which increase the resistance of this region. Unlike in the bipolar diode, conductivity modulation does not reduce this resistance when the junction is forward biased. The large resistance is one of the reasons that silicon Schottky diodes are not generally available at voltage ratings above approximately 100 V.

Temperature Dependence of the Forward Drop The forward voltage drop of a Schottky depends on temperature for two reasons: (i) the i - v characteristic, (17.26), has an explicit temperature dependence; and (ii) the resistance of the drift region in the semiconductor increases with temperature because electron mobility decreases as temperature increases. The mechanism responsible for this decrease is the increase in the amplitude of lattice vibrations. The first

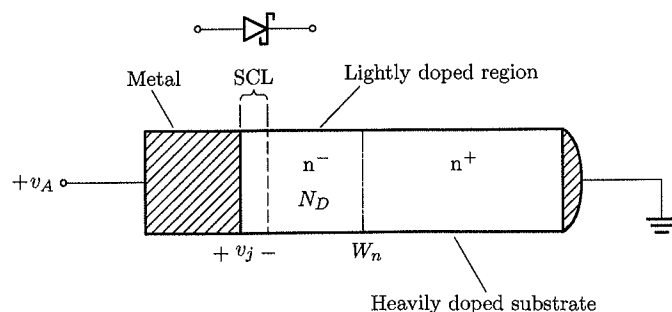


Figure 17.9 The Schottky diode structure showing the metal contact, the lightly doped region which supports the reverse voltage, and the heavily doped substrate.

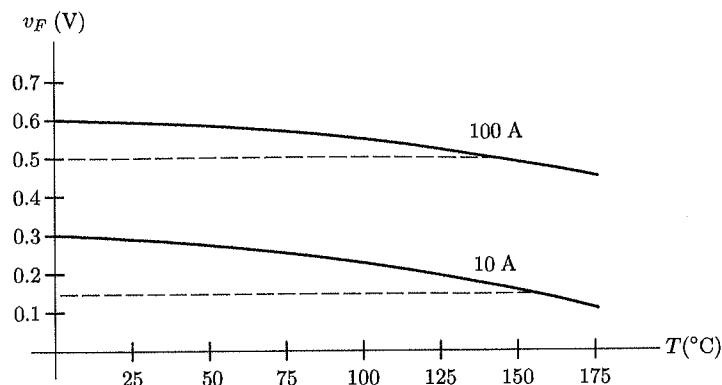


Figure 17.10 The forward voltage drop v_F as a function of temperature for a 50 V, 100 A Schottky diode operating at two different currents.

mechanism dominates the temperature dependence of the forward drop in almost all practical Schottky diodes.

The forward voltage as a function of T is shown in Fig. 17.10 for a 50 V, 100 A Schottky diode. The influence of the drift region resistance at high currents is apparent from the different slopes of the two curves in the figure.

17.3.3 Reverse Bias

Under reverse bias, the barrier to the flow of electrons from the semiconductor is greatly increased, virtually preventing electron flow from the semiconductor to the metal. Reverse current consists primarily of electrons flowing from the metal to the semiconductor. Because of the voltage dependence of the metal barrier height, this current is a strong function of reverse voltage, even though it is still much smaller than the forward current. This reverse voltage dependency makes the leakage current much greater than $-J_0$. The dissipation resulting from this leakage current is the principle determinant of the reverse voltage rating of a Schottky diode.

17.3.4 Switch Transitions in a Schottky Diode

The ideal Schottky diode is a *majority carrier* device, meaning that, generally, negligible excess charge is stored during any operating condition. However, minority carriers do constitute a normally negligible component of total device current. But in high-voltage Schottkys there is minority charge storage in their lightly doped regions, and because in high-voltage devices this region is wide, high-voltage Schottkys do exhibit reverse recovery behavior. In addition, the Schottky diode contains SCL capacitance, which must be charged and discharged during switching.

17.3.5 Selection of Barrier Metal

Because the potential of the Schottky barrier depends on the work function of the metal contact, we can change the forward and reverse characteristics of the diode by using different metals. Chrome, molybdenum, platinum, and tungsten are typical choices. Chrome gives the lowest junction drop but the highest reverse leakage current, so it is usually reserved for low-temperature applications. Tungsten is commonly used because, although it has the highest forward drop of the metals mentioned, it also exhibits the lowest leakage current and so can be used at high temperatures.

17.4 Power BJT

As we noted in Chapter 16, although the BJT has been largely displaced by the power MOSFET and IGBT, familiarity with its operation and characteristics is necessary for understanding the behavior of the SCR and IGBT. The power BJT differs from the BJT discussed in Chapter 16 primarily in its structure required to support high voltages, which results in corresponding behavioral differences.

17.4.1 Structure and Operating Regions of the Power BJT

Figure 17.11(a) shows a cutaway view of a typical npn bipolar power transistor. It is a vertical device, with the collector connection on the bottom (substrate) of the die and the base and emitter connections on the top. Note from the top view in Fig. 17.11(a) that the base and emitter metal traces are *interdigitated*. This pattern is used to keep the distance between the base contact and the centerline of the emitter (line $A-A'$) short, the reasons for which we discuss later.

Looking at the profile of the doping concentration along line $A-A'$, you can see the one-dimensional representation of the transistor shown in Fig. 17.11(b). Note that the collector of this transistor has two regions: a lightly doped ν region not present in a conventional (signal-level) transistor, and a heavily doped substrate. The purpose of the ν collector region, which is more lightly doped than the p-type base, is identical to that for the pin diode – to make the collector–base junction SCL grow into it, rather than into the base. The SCL can then be allowed to punch through to the substrate region, as shown in the plot of E in Fig. 17.11(b). As with the similar pin diode structure, we can nearly double the breakdown voltage for a given ν -region width by allowing this punch-through to occur.

When this transistor is in its on state, the entire ν region is in high-level injection so that its contribution to the on-state drop is small, and current flows by drift. For this reason, we sometimes refer to the ν region as the *drift region*. The operating condition during the drift region's transition to high-level injection results in four, rather than the usual three, distinct regions of operation.

The first is the off-state region, also known as *cutoff*. In this region both the collector–base and emitter–base junctions are reverse biased, the ν region is depleted, and the collector current

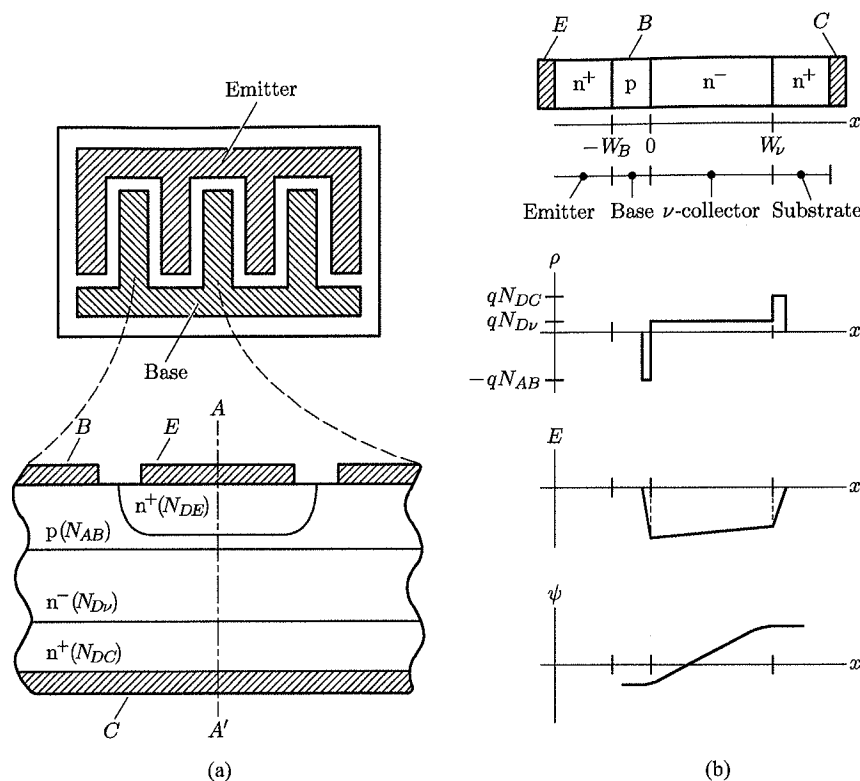


Figure 17.11 The npn power bipolar junction transistor. (a) Top and side-section views of the transistor. (b) One-dimensional representation of the power transistor doping profile along the line $A-A'$ bisecting the emitter, and the charge, electric field, and potential profiles of the reverse-biased collector-base junction.

is approximately zero. The second is the *forward-active* region, where the collector-base junction is strongly reverse biased, while the emitter-base junction is forward biased. Operation in this region resembles that of a conventional transistor. At low values of v_{CB} , corresponding to low values of v_{CE} , when the ν region is not yet in high-level injection (conductivity modulated) the transistor enters the third region, called *quasi-saturation*, corresponding to a forward-biased $p-\nu$ junction that injects excess carriers into the ν region, so that *part* of this region is now in high-level injection. The fourth region of operation, *hard saturation*, occurs when the entire ν region is conductivity modulated, and v_{CE} assumes its lowest on-state value for a given current. The transistor characteristics in each of the four operating regions in the i_C-v_{CE} plane are shown in Fig. 17.12.

The Cutoff Region One might reasonably expect the voltage rating of the transistor in cutoff to be determined by avalanche at the collector base junction. However, this is only true if the emitter is open-circuited. In this case the breakdown voltage is V_{CBO} . But the emitter is infrequently open-circuited, resulting in transistor action which produces a multiple (β_F) of the

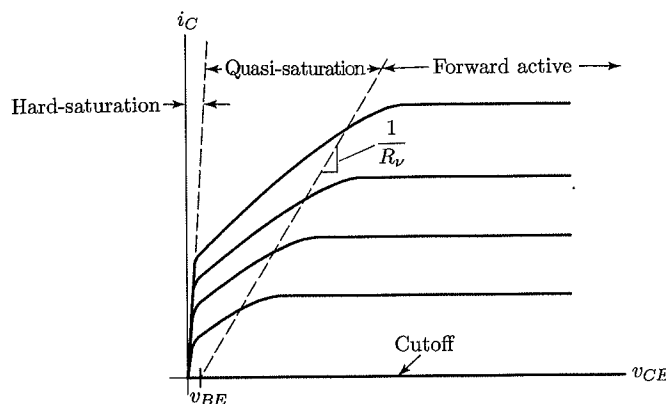


Figure 17.12 The i_C - v_{CE} curves for an npn power transistor with a ν collector region.

base current at the collector junction. For this reason the maximum collector-emitter voltage rating of a BJT is a function of the conditions existing at the base terminal.

If the base terminal is left open, the only source of majority carriers (holes) to support excess charge in the base is the leakage current of the collector junction. But this leakage current has the same effect as the base current in determining excess base charge and therefore the emitter current. Because of transistor action, this emitter current is larger than the “base” (collector leakage) current by a factor β_F , and as the base is open, this increased emitter current will increase the collector current by the same amount. The multiplicative avalanche process further increases the collector current as v_{CE} is increased, creating a positive feedback process, which, at some value of collector leakage current, becomes unstable, resulting in unconstrained collector current. The collector-emitter voltage at which this occurs is known as the *sustaining voltage*, $V_{CEO(sus)}$, which is lower than V_{CBO} . If the base terminal is not open, V_{CEX} (where X refers to the base terminal variable) is a function of i_B , decreasing for $i_B > 0$ and increasing for $i_B < 0$, as shown in Fig. 17.13.

The Forward-Active Region From its off state, the transistor enters the forward-active region when the emitter junction is forward biased and minority carriers are injected into the base region. In most of the forward-active region, the collector is strongly reverse biased and the entire ν collector is depleted. At small values of reverse bias, however, the ν collector is *not* depleted but presents a region of high resistance to the flow of collector current. The value of this resistance, R_ν , is

$$R_\nu = \frac{W_\nu}{q\mu_e N_{D\nu} A_\nu}, \quad (17.27)$$

where A_ν is the cross-sectional area of the ν region.

Figure 17.14(a) shows a simple model for the transistor when it is operating in the forward-active region but at low collector-base voltages where the p - ν junction SCL is very narrow. Figure 17.14(b) shows the corresponding minority carrier distributions. The p - ν junction SCL is not shown because it is so narrow at the low v_{CE} . The transistor will continue to operate in

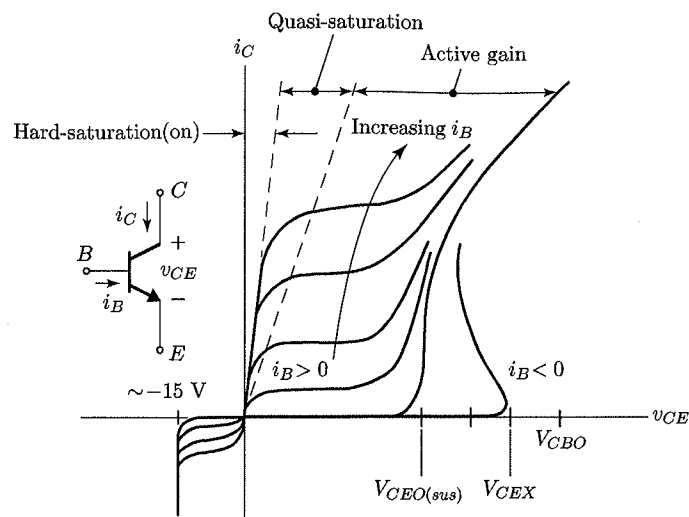


Figure 17.13 Typical i_C - v_{CE} curves for an npn power transistor showing the sustaining breakdown voltage as a function of i_B and its relationship to V_{CBO} .

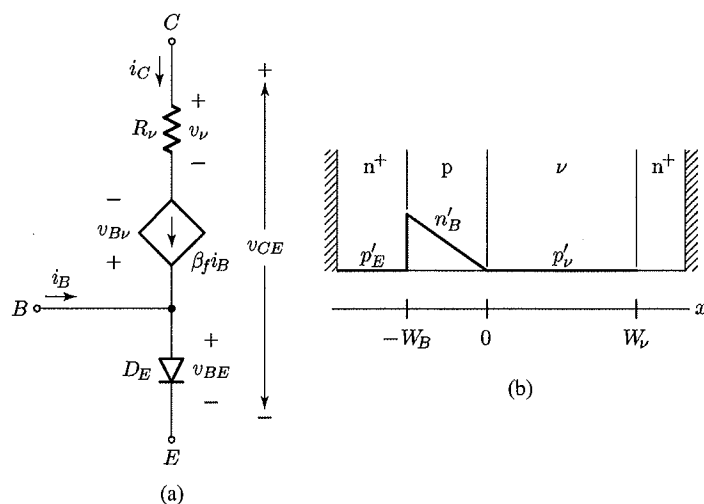


Figure 17.14 (a) A simple model for the power BJT operating in the forward-active region. (b) Excess minority carrier distributions in the power BJT operating in the forward-active region.

this region so long as the collector-base junction (p - v) remains reverse biased. If we assume that $v_{BE} \approx 0.8$ V, this junction is reverse biased if $v_{CE} > i_C R_v + 0.8$ V. On the v_{CE} - i_C plane, this boundary between the forward-active and quasi-saturation regions is a straight line with a slope $1/R_v$, intersecting the v_{CE} axis at $v_{BE} = 0.8$ V, as shown in Fig. 17.12. The region between this boundary and the $+v_{CE}$ axis defines forward-active operation.

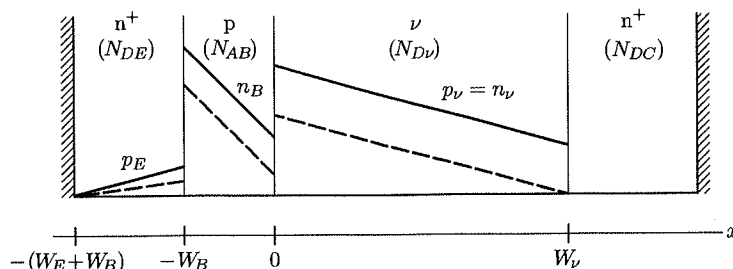


Figure 17.15 Carrier distributions for a transistor in hard saturation. The dashed profiles are those at the boundary of quasi- and hard saturation.

The Quasi-Saturation Region At the quasi-saturation region boundary with the forward-active region the p - ν junction voltage is zero and v_{CE} is the sum of $i_C R_\nu$ and v_{BE} . As v_{CE} continues to decrease, the p - ν junction becomes forward biased and the ν region begins to enter high-level injection and conductivity modulation, reducing R_ν . To match the decrease in v_{CE} , i_C must decrease, which is what we see in Fig 17.12.

The Hard-Saturation Region If, for a given collector current, we drive the transistor with a high enough base current i_B , we will push it into the hard-saturation region. Figure 17.15 shows the carrier profiles for a transistor in this condition. The slopes of the profiles in the base and ν regions are the same as they were at the onset of saturation, but the levels have shifted upwards.

Although the collector current has remained constant, the hole current injected into the emitter from the base increases by the same factor that the excess electron concentration at $x = -W_B$ has been increased. Furthermore, because the ν - n^+ junction is now forward biased, holes originating as base current will be injected into the n^+ collector (similar to reverse injection of holes into the emitter). For these reasons, the required base current quickly increases as we try to push the transistor further into the hard-saturation region, and the gain falls rapidly.

The common-emitter current gain $\beta_F = I_C/I_B$ is a very important parameter for users of power transistors (this gain is often labeled h_{FE} in specification sheets). But unlike users of signal-processing transistors, who only want to know β_F in the forward-active region, we want to know β_F at the onset of saturation in order to determine the minimum base current needed to ensure the on state. As noted above, β_F falls rapidly at high current levels as the transistor is pushed into hard saturation. While signal level BJTs have active region β_F s in the 100s, power BJT's gains in saturation are in the low double or even single digits. But the gain also falls at very low current levels because of the base current needed to support recombination in the emitter-base SCL. This also corresponds to a decrease in the emitter efficiency, γ_E , the ratio of electron current to hole current at the emitter junction discussed in Section 16.3.2. We will see that this current dependence of γ_E is critical to the operation of the SCR. A typical specification of β_F as a function of collector current at constant v_{CE} is shown in Fig. 17.16. Note that the gain is specified at $v_{CE} = 4$ V, which is not very deep into hard saturation. Driving the transistor harder to reduce v_{CE} will further reduce the gain. Note that the gain is also temperature dependent.

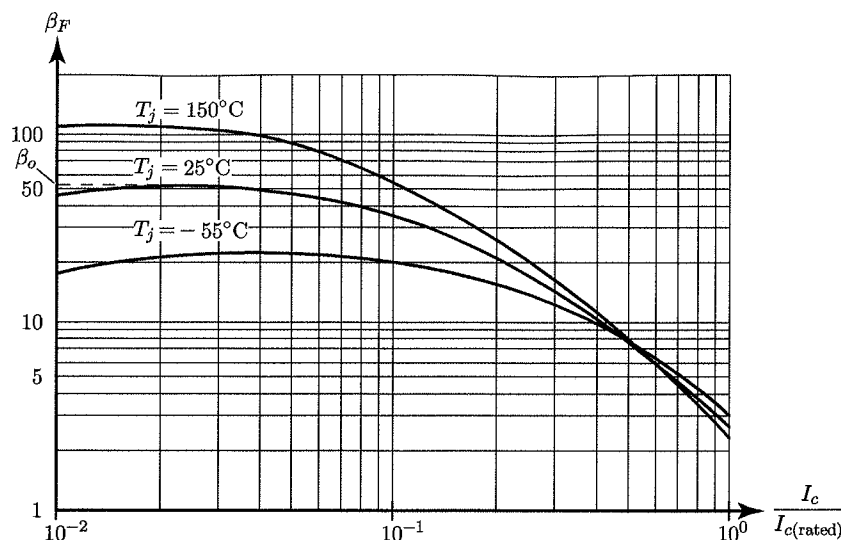


Figure 17.16 A typical specification of common-emitter current gain (β_F) as a function of collector current while holding $V_{CE} = 4$ V. The parameter β_o is the value of β_F specified in the datasheet.

Switch Transitions Figure 17.17 illustrates typical switching waveforms for a bipolar transistor. Consider the turn-on transition first. At $t = 0$ the device is cut off and the base current makes a step change from 0 to $+I_{B1}$. The transistor must first pick up the load current, as it commutates from the diode, before v_{CE} can fall, since the diode cannot support reverse voltage while it is conducting forward current. The collector voltage falls quickly at first but then “tails off” for the last few volts, reaching its final value only after the *voltage tail time*, t_{vt} . This behavior is similar to the forward recovery of a bipolar diode, and is caused by the gradual drop in the resistance of the transistor’s ν -collector region as it becomes conductivity modulated.

During turn-off, when the base current switches from $+I_{B1}$ to $-I_{B2}$, the excess charge in the base and ν regions begins to decrease, but their gradients remain unchanged so i_C remains constant. When the excess concentration at the collector ν - n junction reaches zero, we enter the quasi-saturation region and the fraction of the ν -collector that is conductivity modulated decreases, increasing its resistance and hence v_{CE} . This accounts for the slow rise of v_{CE} before i_C begins to decrease. The period over which this occurs is known as the *storage delay time*, t_s . As i_C begins to fall, the diode turns on and I_d begins to commutate from the transistor to the diode. The current i_C decreases as the base charge declines via the negative base current and recombination. The transistor enters cut-off when $i_C = 0$. During turn-off $-i_{B2}$ is pulling charge out of the base region, and until the excess concentration at the emitter-base junction reaches zero, the junction remains forward biased but conducting negative current. Once $i_C = 0$, v_{BE} can fall to the negative drive voltage $-V_x$.

During both the turn-on and turn-off transients, the transistor is simultaneously supporting nonzero i_C and v_{CE} and dissipating energy. This switching loss and the duration of the turn-on and turn-off processes determine the maximum switching frequency at which a BJT can operate efficiently.

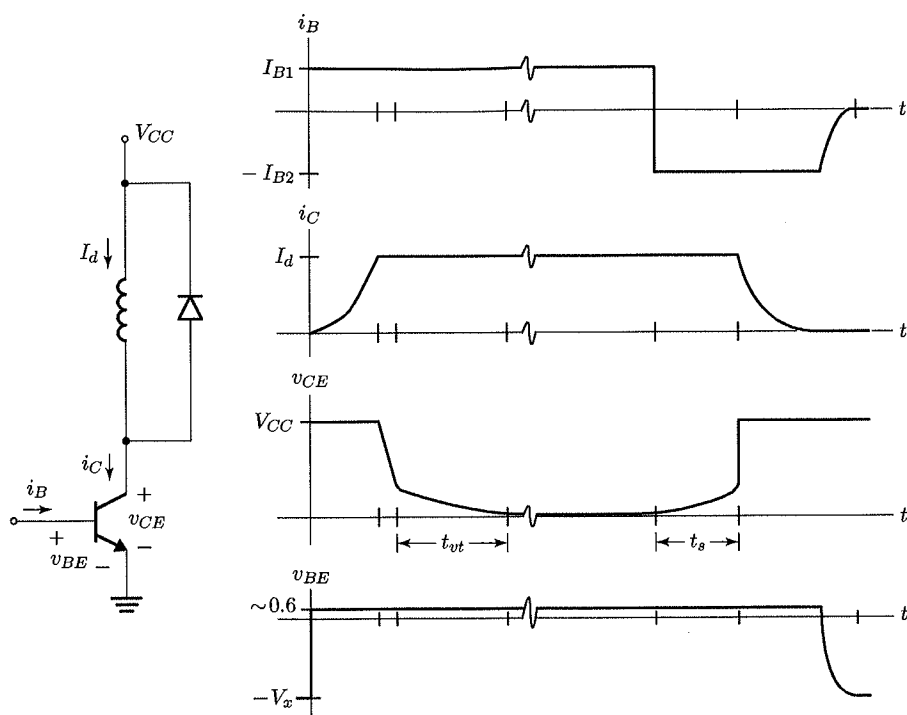


Figure 17.17 Typical switching waveforms exhibited by a power BJT.

17.5 Power MOSFET

Figure 17.18 shows the structure of an n-channel power MOSFET. Note that the drain contact is on the bottom of the die, rather than on the top as in a lateral signal-processing MOSFET, resulting in a device in which the current flows vertically. This structure provides for a thick lightly doped n-region, called the *extended drain*, to block the off-state voltage without compromising surface area. It also gives maximum area to both drain and source contacts in order to produce a low-resistance connection to the package terminals. The polysilicon (or metal) electrode gate is insulated from the source metal that covers it by a layer of SiO_2 . The gate electrode is connected by metal fingers that make contact through windows etched in the SiO_2 . The fabrication of power MOSFETs uses a double-diffusion process to create the n^+ and p regions, and the device is often referred to as a *vertical double-diffused MOSFET* (VDMOSFET or DMOS). The structure of Fig. 17.18 has a lateral channel. While lateral-channel designs currently dominate SiC MOSFETs, more modern Si power MOSFETs have a vertical channel which we discuss shortly. But we discuss the lateral-channel structure here because it is somewhat easier to visualize the various aspects of its operation.

Between the n^+ source and the n^- drain regions are *wells* of p-type material. These wells are known as the *body* region of the device. The channel of the MOSFET is formed on the surface of these p wells just beneath the gate oxide. Note that the p wells are shorted to the source by the source electrode. We described the reason for this source-body short in Section 16.4.2.

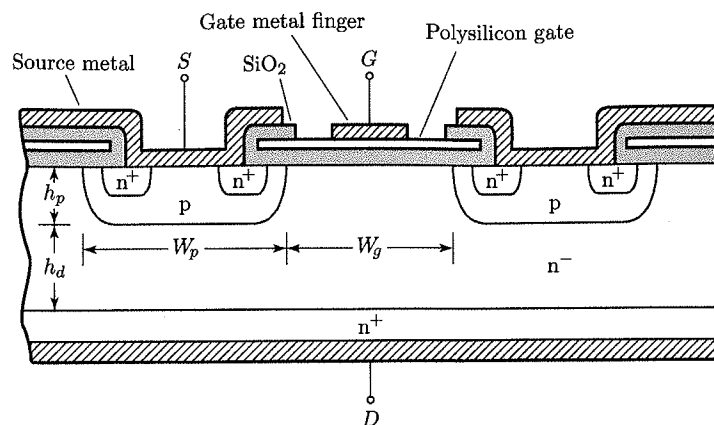


Figure 17.18 Structure of a vertical n-channel power MOSFET.

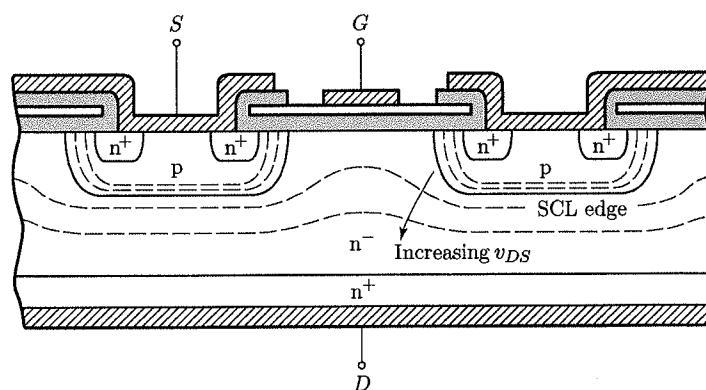


Figure 17.19 An illustration of how the SCL grows in a power MOSFET with increasing v_{DS} , pinching off the n^- region between the p wells.

The lightly doped n-type drain region is unique to power MOSFETs, and is provided to allow growth of a long SCL, permitting the device to block a high voltage when it is off. In this respect, the region functions like the v region in a power BJT. This lightly doped drain region is frequently referred to as the *extended drain* or *drift region*. As the SCL grows, it *pinches off* (depletes) the region between the p wells, as shown in Fig. 17.19 (the gate electrode acts as a field plate to promote this pinch-off). This feature of a power MOSFET's structure is important because it keeps the gate oxide from being subjected to the full drain voltage. In fact, the voltage just underneath the gate oxide typically reaches only 5–10 V with respect to the gate electrode, even though the drain voltage may be 200–400 V. As a result, we can make the gate oxide relatively thin, which keeps both the gate threshold voltage V_T and the gate energy loss low.

A further distinction between the signal-level and power MOSFETs is the relationship between i_D and v_{GS} . That for the signal-level MOSFET was given in (16.64), which shows the quadratic relationship resulting in the non-uniform spacing of the curves in Fig. 16.14. At the high gate voltages characteristic of power MOSFETs, the carrier velocity saturates at v_{sat} , which

is on the order of 10^7 cm/s for Si, and is no longer given by $\mu_e E$. Effectively, μ decreases as E (and V_{DS}) increases. An approximation to the i_D - v_{GS} relation better than (16.64) for the power MOSFET is

$$i_D = k v_{\text{sat}} (v_{GS} - V_T). \quad (17.28)$$

This is not particularly important for applying power MOSFETs, since they are seldom operated in their linear region. But it is important if one calculates transconductance as discussed with respect to the Miller effect in Chapter 16.

DMOS On-State Resistance Figure 17.20(a) shows how the drain current flows through the DMOS of Fig. 17.18 when it is on. As the current flows through the extended drain, it focuses in the area between the p wells, called the *neck* region. From there it must further focus in the thin entrance to the p-well channels on either side of the neck region. This latter focusing would result in a large resistance if the current had to flow through a progressively smaller cross-sectional area of the n^- region to get to the channel entrance. Fortunately, the region just below the gate oxide (and between the p wells) is *accumulated*, meaning that it has a very high concentration of mobile electrons, which makes it much more conductive than the bulk n^- region. The drain current between the p wells therefore tends to first flow directly upward to this layer (taking advantage of the full width between the p wells) and then flows horizontally (with little resistance) to the channel entrances.

Figure 17.20(b) shows how the MOSFET's total drain-source resistance, R_{DS} , can be broken down into four components: R_{xd} , R_{neck} , R_{accum} , and R_{ch} . For a DMOS designed for 400 V or above, R_{xd} is typically much larger than R_{neck} , R_{accum} , and R_{ch} , and thus dominates R_{DS} . For a DMOS designed for 100 V and below, however, $R_{\text{ch}} + R_{\text{neck}}$ is typically one third to one half of R_{DS} . Power MOSFETs are characterized by a normalized parameter called *specific on-state resistance*. It is defined as

$$R_{DS(\text{sp})} = (R_{DS})(A),$$

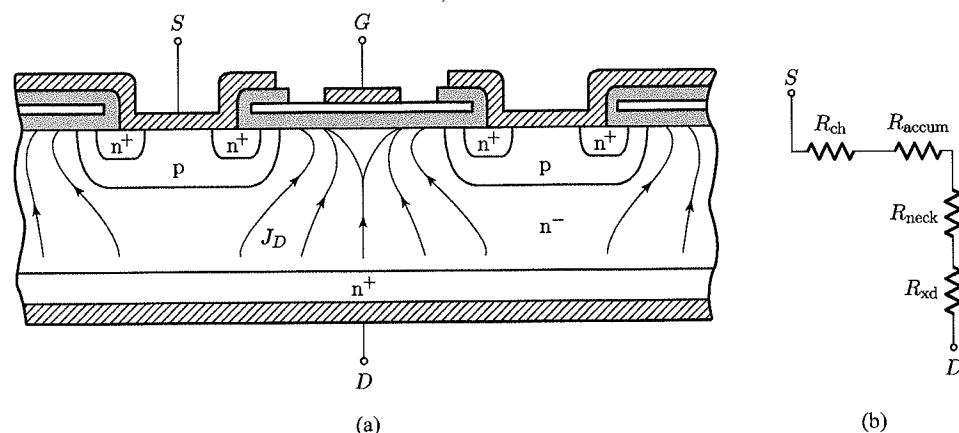


Figure 17.20 (a) Current flow in a power MOSFET. (b) A model of R_{DS} , showing its components.

where A is the active area of the device. This parameter is voltage specific, and illustrates what minimum R_{DS} is possible for a given device area and therefore current rating. For an ideal device when $R_{DS(on)}$ is dominated by the extended drain, this relationship is

$$R_{DS(on)} \propto V_{BR}^{2.5} / \text{Area}. \quad (17.29)$$

For example, a device designed to withstand 400 V would have $4\sqrt{2}$ times the resistance of a device designed to withstand only 200 V, if they both have the same die area. Actual devices obey this relationship only at higher voltage levels (above approximately 100 V). For devices rated below 100 V, $R_{DS(on)}$ is higher than (17.29) predicts because the channel and neck resistances become significant and because less of the total die area is actually used.

Note that current in the MOSFET is carried by majority carriers, which is why it is called a majority carrier device. Unlike the pin diode or BJT, the extended drain cannot be conductivity modulated to reduce its on-state resistance.

17.5.1 Trench or UMOS

Figure 17.21 illustrates the structure of a vertical-trench power MOSFET, sometimes referred to as a UMOS. Its operation is similar to that of the DMOS of Fig. 17.18, except that the channel is now vertical at the oxide walls of the trench. The benefits of this structure are that there is no neck region when the device is on, and the cells can be fabricated at a higher density than those of the DMOS design, that is, there are a greater number of channels per unit chip area. In addition there exists an accumulation of carriers at the gate oxide where the n-drift region is penetrated by the gate trench, reducing the resistance in this region. The result is a lower $R_{DS(on)}$ relative to the planar DMOS device, particularly for lower-voltage devices where the channel and accumulation region resistances are significant relative to that of the n-drift region.

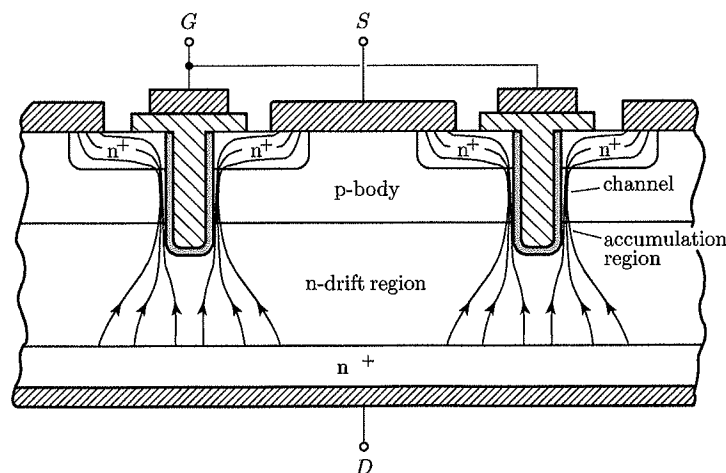


Figure 17.21 The trench or UMOS structure, also illustrating the drain current paths.

UMOS On-State Resistance Figure 17.21 shows the current paths in a trench-gate (UMOS) device. The principle contributions to $R_{DS(on)}$ are the drift region and channel. In low-voltage devices the substrate (the n^+ drain) and the source metal contact also contribute. In high-voltage devices (above about 100 V), the resistance of the drift region dominates because of its required depth to support the high voltage.

17.5.2 Temperature Dependence of $R_{DS(on)}$

Because the high-voltage vertical Si MOSFET's resistance is dominated by the extended drain, the temperature dependence of $R_{DS(on)}$ is the result of the temperature dependence of μ_e . For the range of interest (0–200 °C), mobility decreases with temperature as $\mu_e \propto T^{-2.2}$, which means that $R_{DS(on)}$ increases with temperature. A 100 °C rise produces an increase of approximately 90%.

17.5.3 DMOS Dynamic Behavior

Because the MOSFET is a majority carrier device, it does not have the buildup of excess carrier concentrations that controls the dynamic behavior of the BJT. The transient performance of the MOSFET is governed only by the oxide and SCL capacitances, and by the impedances that limit our ability to charge and discharge these capacitances. (Some of these impedances are internal to the device but most are external.) We discussed these capacitances briefly in Chapter 16, and described how they were characterized on datasheets. Figure 17.22 shows their physical origins in a DMOS. Comparable capacitances exist between the different elements of the UMOS structure.

Between the gate and the source are two parallel capacitances that comprise C_{GS} . One, C_{GS1} , is the result of source metalization actually covering the polysilicon gate, but being insulated from

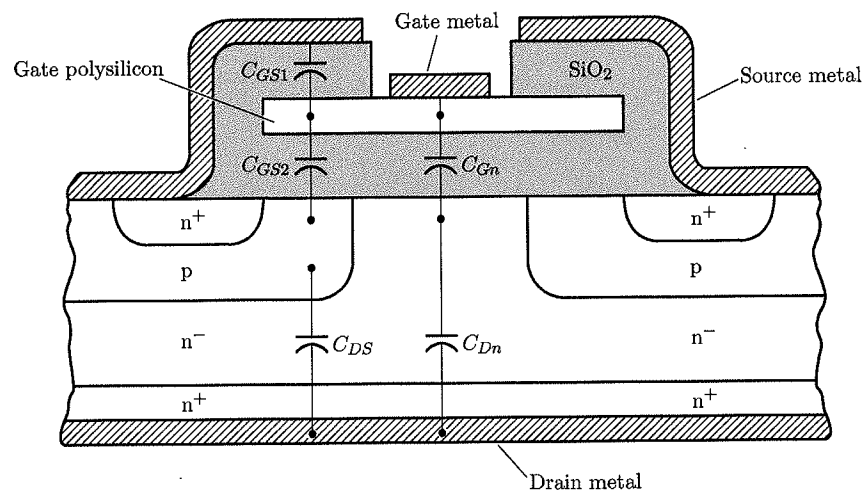


Figure 17.22 Origins of capacitances in a vertical MOSFET.

it by an oxide layer. Part of this capacitance is also the result of the overlap of the polysilicon and the n^+ source diffusion. Thus C_{GS1} is an oxide capacitance, which is not a function of the voltage across it. The second capacitance making up C_{GS} is the gate-to-p-well capacitance of the channel C_{GS2} . Although there is a resistance, R_p , between the p-well side of this capacitor and the source contact, it is small enough to ignore for typical rates of change of gate voltage. As C_{GS2} is the series combination of the oxide capacitance and the p-well's depletion capacitance, it is a strong function of the gate voltage when $v_{GS} < V_T$ (the capacitance gets smaller as the voltage increases). Once the gate voltage is above threshold, however, C_{GS2} is dominated by the oxide capacitance, which is constant.

Between the drain and the source is the capacitance of the drift region's space-charge layer, C_{DS} . Its value varies inversely with the square root of the drain-source voltage.

Finally, between the gate and drain is C_{GD} , a series combination of two capacitors. One, C_{Gn} , is the oxide capacitance, which is independent of voltage. The other, C_{Dn} , is the SCL capacitance between the drain and the neck-oxide interface. It gets smaller as the drain voltage rises. Note that these two capacitances are "connected" at the plane in which the accumulation layer grows.

At low values of v_{DS} , the SCL does not extend into the neck region, so $C_{DG} = C_{Gn}$. At higher voltages, the neck region becomes depleted, so C_{Dn} appears in series with C_{Gn} and reduces the net capacitance C_{DG} . As the drain voltage rises even higher, however, pinching-off of the neck region becomes complete, and very little incremental increase in drain voltage appears at the neck-oxide interface. Therefore C_{GD} becomes very small under this condition.

Note that as the channel and the accumulation layer form, the node connecting C_{Dn} and C_{Gn} is shorted to the source by the low resistances of these regions. Under this condition, C_{GD} becomes zero, C_{Gn} appears in parallel with the other gate-source capacitances, and C_{Dn} appears in parallel with C_{DS} .

The values of C_{GD} , C_{GS} , or C_{DS} are not provided in datasheets. Instead, they can be calculated from the measured incremental capacitances C_{iss} , C_{oss} , and C_{rss} , which are functions of V_{DS} and provided by the manufacturer in graphical form as shown in Fig. 17.23 for a typical DMOS. The relationship between these two sets of capacitances was illustrated in Fig. 16.15.

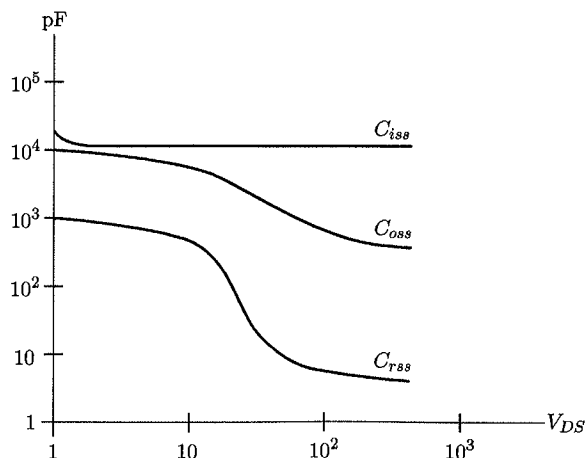


Figure 17.23 Dependence of DMOS datasheet capacitances C_{iss} , C_{oss} , and C_{rss} on V_{DS} .

17.6 Superjunction MOSFET

The blocking voltage, V_{BR} , of a conventional MOSFET is governed by the avalanche breakdown of the drift region, the extended drain for high-voltage devices. Increasing the blocking voltage is achieved by increasing the length of the extended drain and/or reducing its majority carrier concentration, simultaneously increasing the specific on-state resistance, which, in the simplest case of uniform current flow, is given by (17.29). The relationship between the minimum R_{drift} and V_{BR} is complex and constrained by what is known as the *silicon limit*, which for an abrupt diode is

$$R_{\text{drift}(\text{min})} \propto V_{BR}^{2.5}. \quad (17.30)$$

Figure 17.24 shows the structure of a *superjunction* or SJFET. The p body is augmented with narrow columns of p^- regions extending vertically into the n^- drain region. The n^- drain is now also comprised of narrow columns. The result is two parallel pin diodes between the source and drain: n^+-p and n^+-p^- . There exists a $p^- - n^-$ junction between the columns which is reverse biased when the device is blocking (off). Since these columns are so lightly doped, they completely deplete at a low voltage. The resulting electric fields within the columns require a two-dimensional analysis, but insight into their behavior can be obtained by considering the stylized columns of Fig. 17.25.[†]

The equipotential lines shown in Fig. 17.25(a) are as they would be if the p-i-n columns were separated. However, since they share a boundary, the equipotentials as shown cannot remain since the potential must be continuous across a boundary unless there is an impulse of charge at the interface, which in this case does not exist. Consequently the equipotential

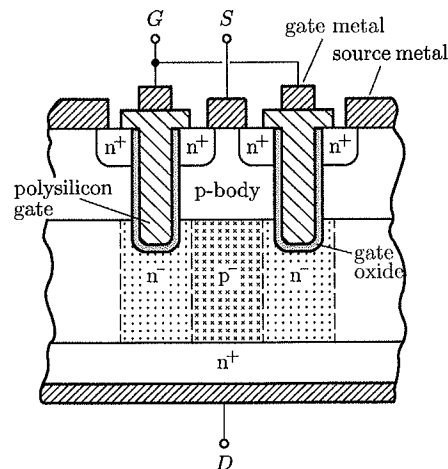


Figure 17.24 Basic structure of a superjunction MOSFET (SJFET).

[†] This representation is a derivative of that presented in R. Siemienieć, C. Braz, and O. Blank, "Design considerations for charge-compensated fast-switching power MOSFET in the medium-voltage range," *IET Power Electronics*, vol. 11, no. 4, pp. 638–645, 2018.

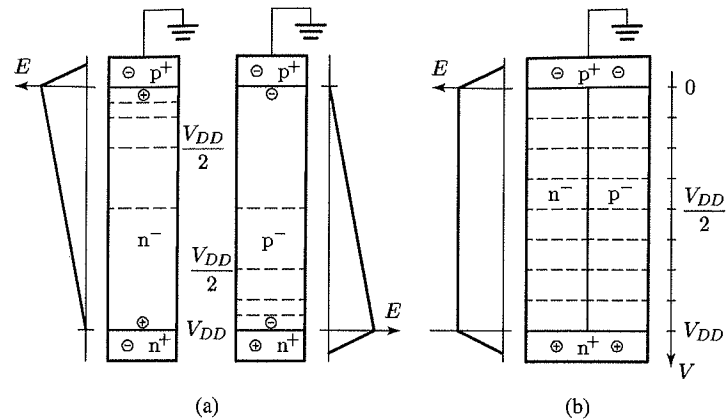


Figure 17.25 (a) A stylized representation of the n⁻ and p⁻ columns of the SJFET showing the E-field in each if they were separated. (b) The equipotential lines for an applied arbitrary voltage of V_{DD} .

lines must be distorted to achieve continuity. If there is a charge balance between the n⁻ and p⁻ columns at the junction between them, the result is equally spaced equipotentials (a result we will not prove here). If the columns are very narrow, the approximately equal spacing will penetrate the column widths, as shown in Fig. 17.25(b). Equal spacing of the equipotentials is the result of a constant E -field. Thus the combined columns behave as a single pin diode, with the near-intrinsic region forming the extended drain of the MOSFET and providing a higher breakdown voltage than a conventional device with the same extended drain thickness.

An additional benefit of the SJ structure is that because the columns are so narrow, they deplete at a low voltage. Therefore their doping concentrations can be increased, reducing the resistance of the extended drain and $R_{DS(on)}$. Compared with a conventional MOSFET with the same voltage rating, the thickness of the extended drain of the SJFET can be thinner, resulting in a further reduction of $R_{DS(on)}$. For the same $R_{DS(on)}$, the die size can be reduced by a factor of 2–5.

Because of the very different SCL shapes and behaviors as a function of V_{DS} , the dependencies of C_{iss} , C_{oss} , and C_{rss} on V_{DS} differ markedly from those shown in Fig. 17.23 for the DMOS structure. Representative behavior of these capacitances for the SJFET are shown in Fig. 17.26.

17.7 IGBT

The circuit symbol, equivalent circuit, and physical structure of a lateral channel IGBT are shown in Fig. 17.27. Although the physical architecture of the device incorporates a pnp BJT, its symbol showing an npn BJT is designed to represent its application behavior. The structure is labeled consistent with the circuit symbol, not the terminals of the pnp BJT. An IGBT structure using a trench channel is shown in Fig. 17.28. It is the design most commonly used because of its smaller contribution to the on-state drop of the device relative to a lateral design, as discussed in Section 17.5.1. Note that the n–p–n–p IGBT structure is similar to a thyristor. However, the

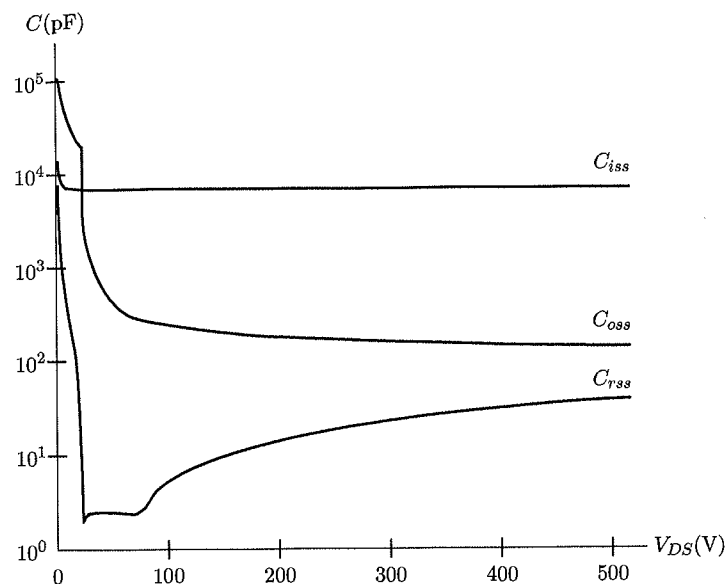


Figure 17.26 Capacitances vs. V_{DS} for a 600 V, 60 A SJMOSFET (at $T_C = 100^\circ\text{C}$).

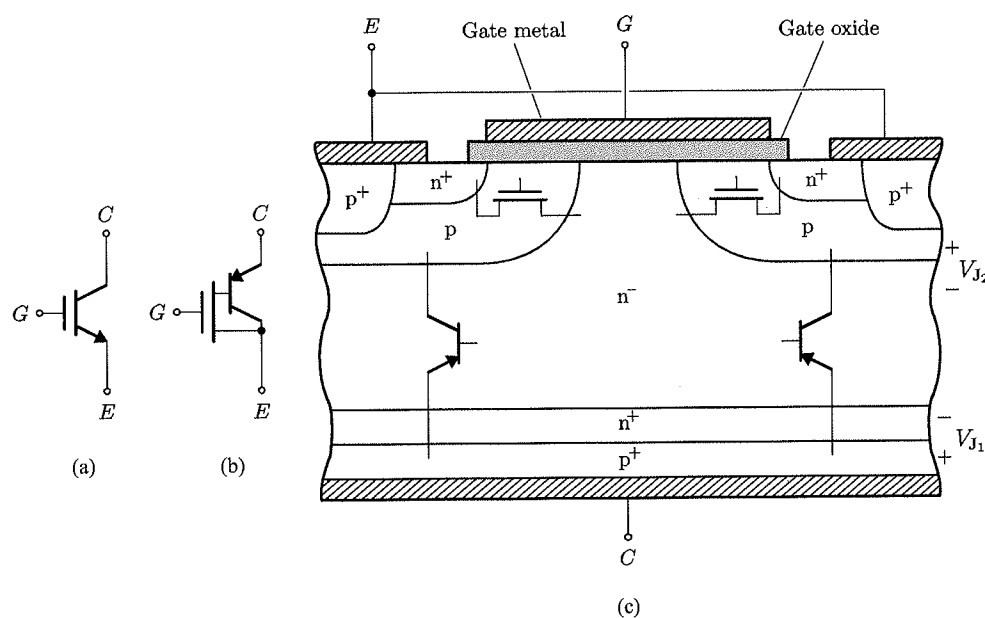


Figure 17.27 The lateral channel IGBT. (a) Schematic symbol. (b) Equivalent circuit. (c) Physical structure showing MOSFET and BJT regions.

n-p-n transistor is made inoperative by the short between its emitter and base created by the emitter metalization. Also shown in Fig. 17.28 is how a freewheeling diode can be integrated with the device by providing access to the collector contact by the n^+ field-stop layer.

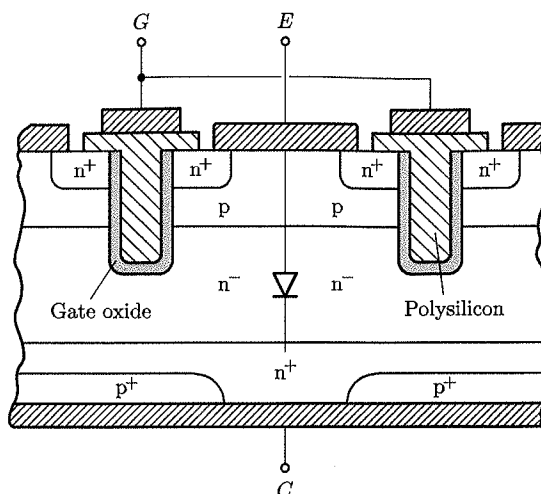


Figure 17.28 The IGBT structure designed with a trench channel and illustrating how a freewheeling diode can be integrated.

17.7.1 On State

With reference to Fig. 17.27(c), the n-channel MOSFET driving the BJT is comprised of the n^+ region as its source, the p region as its body (which is also the collector of the pnp BJT), and the n^- region as its drain. The n^- layer also functions as the base of the BJT, and the p^+ layer is its emitter. With $V_{CE} > 0$, the IGBT is turned on by applying a positive voltage to the gate (V_{GE}) which supplies electrons through the resulting channel to the n^- base region. Simultaneously, the forward-biased $p^+ - n^-$ junction floods the n^- base region, resulting in a large fraction of the region being conductivity modulated, reducing its contribution to the on-state drop $V_{CE(on)}$. Contrary to what one might infer from the IGBT symbol, the bipolar transistor conducts a relatively small fraction of the device current. This is due to its very small gain as a consequence of its very wide base region and poor emitter efficiency. The majority of the collector current is comprised of electrons from the FET channel.

While one might expect the forward drop of the IGBT to be that of the pnp transistor in saturation, the MOSFET channel is in parallel with the $p - n^-$ junction which is the collector junction of the pnp transistor. Therefore the $p - n^-$ junction remains slightly reverse biased, preventing full saturation of the transistor. The forward drop is then

$$V_F = V_{J_1} + V_{base} + R_{ch}I_{ch},$$

where V_{J_1} is the emitter–base voltage of the transistor, V_{base} is the drop across the n^- base region, and I_{ch} is frequently approximated as I_C . The total forward drop is typically on the order of 1.5 V for a 600 V/100 A device to 3.5 V for one rated at 6500 V/1000 A. The forward drop is a function of temperature and increases by about 35% between 25 °C and 175 °C.

17.7.2 Off State

In the off state the channel is not inverted (that is, not conducting) and the p-n⁻ junction J_2 is reverse biased and supports V_{DS} . The n⁺ layer at the collector junction is a buffer, or *field-stop* layer that prevents the SCL from punching through to the collector. As in the pin diode, this permits the lightly doped n⁻ region to be completely depleted and block a high V_{DS} . It is in this regard that the IGBT differs from a BJT in its voltage capability. If such a highly doped buffer layer were present at the emitter side of the base region (to which the BJT SCL spreads), the transistor's emitter efficiency would be very low, resulting in such a low gain that the device would be unusable.

In this state there is a collector-emitter leakage current across J_2 whose value at rated V_{CE} is the parameter I_{CES} known as the *collector cut-off current*, which ranges from 10s of μA to 10s of mA for devices with V_{CE} ratings of 100s of volts to 1000s of volts, respectively, at 25°C, and increases with temperature.

17.7.3 Turn-on and Turn-off Dynamics

Like the MOSFET, the dynamic behavior of the IGBT is determined in large measure by its internal capacitances. But unlike the MOSFET, the IGBT is controlled by minority carrier as well as majority carrier dynamics. Minority carrier lifetime in the n⁻ base region limits the turn-off time of the device. The datasheet specifications are shown in Fig. 17.29, along with their relationships to the capacitances seen from the device terminals. The measurements and relationships are similar to those for the DMOS of Fig. 17.23.

The IGBT turn-on and turn-off transients are shown in Fig. 17.30 assuming a clamped inductive load. That is, the inductor current is constant at I_L and circulating through the

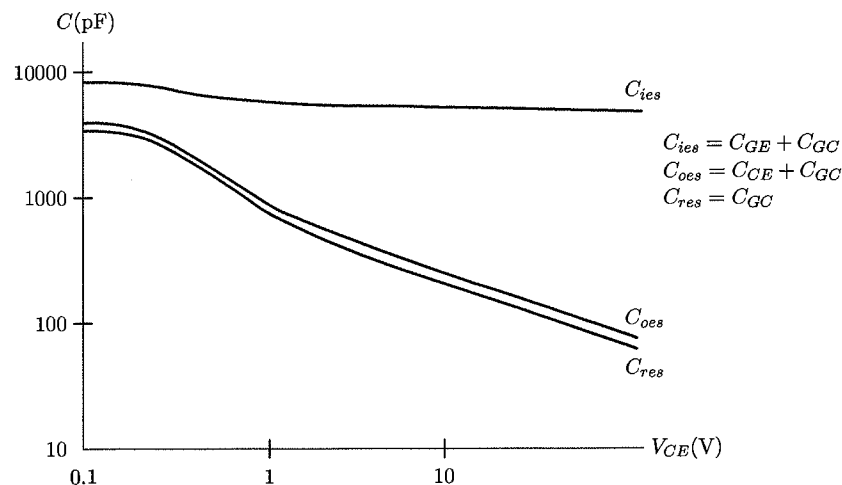


Figure 17.29 Typical variation of datasheet measurements of the input (C_{ies}), output (C_{oes}), and reverse transfer (C_{res}) capacitances with V_{CE} for a 600 V, 40 A trench gate IGBT.

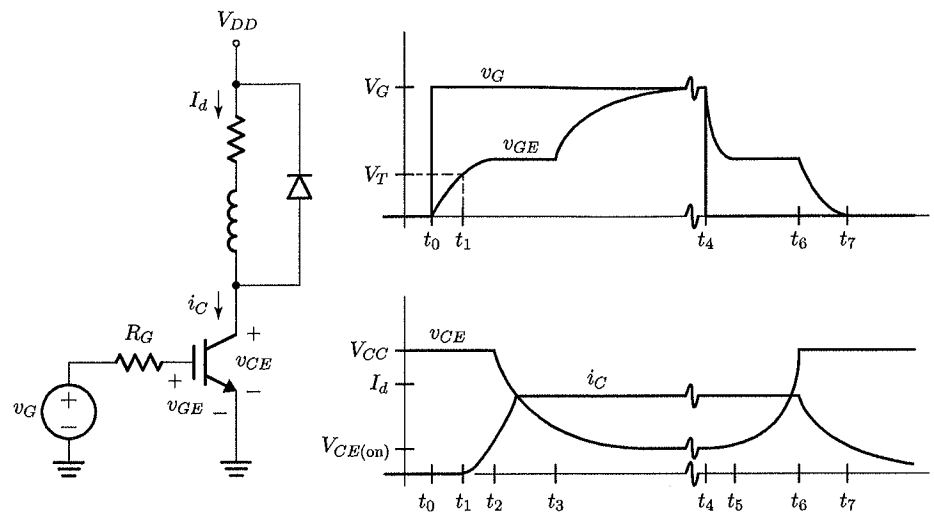


Figure 17.30 (a) The clamped inductive load circuit used to illustrate the IGBT switching behavior. (b) The IGBT turn-on and turn-off transients.

freewheeling diode and $v_{CE} = V_{CC}$. During turn-on between t_0 and t_1 , v_G rises exponentially to V_T at t_1 as the gate capacitance is charged. A further increase in v_G causes the collector current to rise as the diode current falls but v_{CE} remains at V_{CC} . When $i_C = I_L$ the diode turns off and v_{CE} falls as C_{CE} discharges. During the interval t_2 – t_3 , v_{GE} remains approximately constant resulting from the Miller effect as discussed in Chapter 23. At t_3 the Miller effect has diminished and v_{GE} continues its journey to V_G , the maximum gate voltage, while v_{CE} reaches $V_{CE(on)}$.

Because there is no external access to the n^- base region, the turn-off process cannot be accelerated by applying negative base current to extract the stored base charge. When $v_G = 0$ at t_4 , v_{GE} begins to decrease and again exhibits a Miller-effect plateau between t_5 and t_6 as v_{CE} begins to rapidly rise. The diode remains off, and $i_C = I_L$ until $v_{CE} = V_{CC}$ and the diode turns on. At this time there is still excess charge remaining in the base, which slowly recombines. Because the lifetime in the base region is long, the result is a persisting tail in the collector current as shown in the waveform of i_C . Since this tail occurs at high v_{DS} , it can account for an appreciable fraction of the turn-off loss. The length of this tail is a function of the lifetime in the n^- region, and for IGBTs designed for high frequency (50–100 kHz) the tail time is reduced by reducing lifetime at the expense of higher on-state losses.

17.8 Silicon Carbide and Gallium Nitride Devices

As mentioned in Chapter 16, the electronic and thermal properties of the wide-bandgap materials silicon carbide (SiC) and gallium nitride (GaN) provide improved power-device performance relative to Si. The primary reasons for this are the high values of their critical field and thermal conductivity. Although power devices are commercially available in both materials,

SiC dominates the high-voltage, high-current devices, while GaN is generally limited to voltages less than 650 V and moderate currents. The large bandgap for both devices is reflected in their large forward-junction drops, which are a disadvantage in bipolar but not field-effect devices. The wide bandgap would also permit these devices to operate at higher temperatures than Si, but packaging and bonding technologies at present limit maximum device temperature to less than 200 °C.

17.8.1 SiC Devices

Except for their increased threshold voltage (V_T), SiC MOSFETs are practically drop-in replacements for their Si counterparts. But the efficiency benefit of SiC is obtained when the on-state drop is dominated by the device drift region and the SiC drop is significantly lower than that for a Si device, which occurs around 900 V. The drift region, as illustrated in Figs. 17.20 and 17.21, must be thick enough to support the rated device voltage. Since SiC has a critical electric field E_B that is 10 times that of Si, the thickness of the drift region can be reduced significantly. Additionally, to support the same voltage in this thinner drift region, the E -field must be increased by increasing the doping density by the inverse of the square of the dimensional decrease. Combining these two changes, and accounting for the decreased electron mobility in SiC relative to Si, results in a (theoretical) reduction of the drift-region resistance by a factor of about 500 for devices of similar area. A further advantage of SiC MOSFETs is that the temperature dependence of $R_{DS(on)}$ is substantially less than that for Si devices, as illustrated in the next section on the datasheet.

On a per-unit basis, the device capacitances of SiC and Si MOSFETs are comparable because the dielectric constants of the two materials are approximately equal. However, for devices of equal voltage rating, and taking advantage of a thinner SiC die to achieve reduced $R_{DS(on)}$, for equal die area the SiC device will have a substantially higher C_{oss} due to the reduced width of its drain-source space-charge layer. Reducing the die area will reduce C_{oss} but increase $R_{DS(on)}$. There is, therefore, a trade-off between $R_{DS(on)}$ and C_{oss} depending upon the circuit design and whether switching or conduction dominates the device loss. But, for a given voltage rating, SiC devices have smaller $R_{DS(on)}C_{oss}$ products than Si, and therefore their use will result in lower losses.

The combination of a thinner SiC device and its high thermal conductivity allows the device to operate in ambient temperatures considerably higher than for Si. But other system components, particularly capacitors, can still limit the maximum ambient temperature, while bonding and packaging considerations limit the die temperature.

Unlike Si SBDs, which are limited to relatively low voltages, SiC SBDs are available with voltage ratings on the order of 10 kV. The wide bandgap reduces the rate of thermally generated carriers, reducing considerably the leakage current which limits the voltage rating of Si SBDs.

17.8.2 GaN Devices

Unlike SiC devices, those fabricated in GaN exhibit important structural and behavioral differences from their Si counterparts. Early GaN MOSFETs were depletion-mode devices,

meaning they were normally “on.” But manufacturers have developed the technology to produce enhancement-mode MOSFETs which are now in common use. Most GaN devices are not fabricated on a native substrate, that is, the starting wafer is not pure GaN. A thin layer of GaN is epitaxially deposited on a wafer of Al_2O_3 (sapphire), SiC, or Si. All three exhibit a lattice mismatch to GaN, but SiC comes closest at approximately 3.5%. But because of the significant cost differential, Si is commonly used for power GaN devices. Due to the relatively low maturity and high cost of native GaN substrates, the majority of current commercial GaN devices are lateral instead of vertical. While this reduces their achievable voltage capability, it provides an opportunity for their monolithic integration. Their favorable high-frequency behavior is due in part to a threshold voltage of less than 2 V and a very small $R_G C_{iss}$ product, allowing fast transitions of the gate voltage. Device processing incorporates a layer of AlGaIn on top of the GaN, resulting in what is known as a *heterojunction* between the two. For reasons not discussed here, this structure increases the mobility of electrons in the FET channel. The device is known as a *high-electron-mobility transistor* (HEMT).

A significant departure from the behavior of Si and SiC devices is a phenomenon known as *dynamic on-state resistance* exhibited by GaN MOSFETs. It manifests as an $R_{DS(on)}$ at turn-on after blocking a high voltage that is multiples of the datasheet static $R_{DS(on)}$. The increase is proportional to the magnitude of the blocking voltage. The cause is thought to be trapping of carriers in the buffer layer between the channel and oxide. The result is an on-state loss that is higher than anticipated based on datasheet parameters. One useful rule of thumb is to double the datasheet value of $R_{DS(on)}$ for design purposes.

Since GaN devices are most often used in very high-frequency converters, loss in the output capacitance, C_{oss} , can be significant. This loss is due to hysteresis in the capacitor dielectric's *D-H* trajectory, which is a function of frequency.

17.9 Thyristor

As noted in Chapter 15, thyristor is the generic term for variants of a four-layer semiconductor device, the most common of which is the SCR. The structure is shown in Fig. 17.31(a). For the lower current ratings (10–100 A) the SCR is built on a small die of silicon, but for higher current ratings (100–6000 A), it is fabricated on an entire wafer. The heavily doped p region on the bottom of the device is the *anode*, and the heavily doped n region on the top of the device is the *cathode*. The p region under the cathode is the *gate*, and it is connected to the gate metalization on the top surface, as shown in Fig. 17.31(a). The interdigitation of the gate and cathode metalization is typically non-existent or sparse in an SCR. Figure 17.31(b) shows the top view of a device whose gate is restricted to the center of the wafer. Known as a *center-gate geometry*, it is used extensively for low-frequency SCRs. The gate pattern shown in Fig. 17.31(c) is called an *involute*, which has the important property of constant spacing between two gate fingers across the entire device.[‡] A typical spacing is 3–5 mm for a wafer-sized device. Interdigitation using this or other patterns is utilized in high-frequency thyristors.

[‡] Herb Storm of GE CR&D was the first to recognize the benefit of the involute structure and apply it as a thyristor gate geometry.

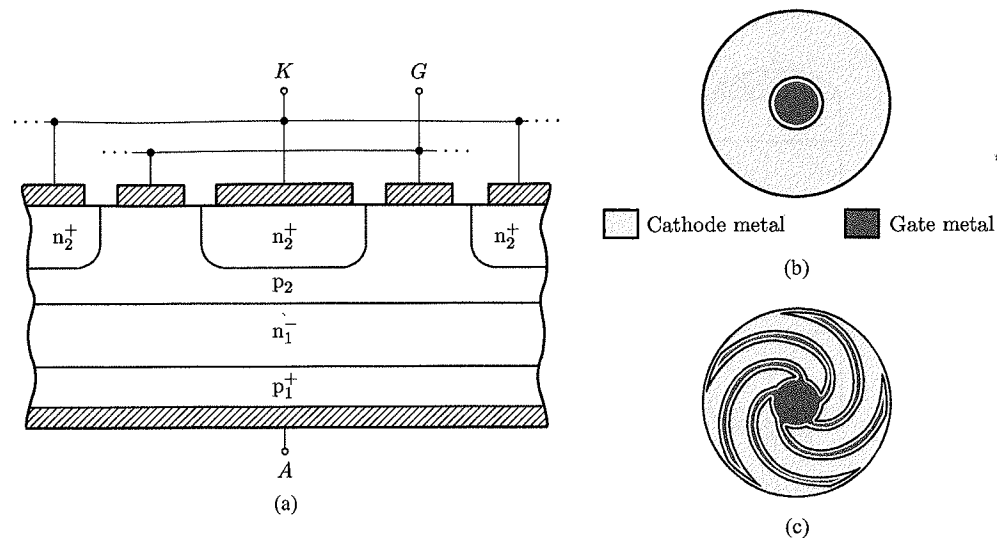


Figure 17.31 Structure of an SCR. (a) Cutaway view showing the four layers and terminal connections. (b) Top view of center-gate geometry. (c) Top view of involute gate geometry.

17.9.1 Overview of SCR Operation

The purpose of this section is to briefly describe the off-state characteristics, the regenerative turn-on process, and the on-state characteristics of the SCR. The sections that follow give greater detail on each aspect of the device's operation. Throughout these discussions, we use the one-dimensional model of the SCR shown in Fig. 17.32(a). (Note that the anode is now drawn on the top.) For future reference we have labeled each of the four layers (p_1 , n_1 , p_2 , and n_2) and each of the three junctions (J_1 , J_2 , and J_3).

Off State When off, the SCR can block either a reverse or a forward anode-cathode voltage, v_{AK} . When v_{AK} is negative, junctions J_1 and J_3 are reverse biased (assuming that the gate terminal is left open), and J_2 is forward biased. The doping levels on each side of junction J_3 are very high, so this junction breaks down at a relatively low voltage. However, the n_1 region is both lightly doped and long. Therefore J_1 is able to block a large reverse anode-cathode voltage. As the n_1 region is much more lightly doped than the p_1 region, the SCL at J_1 grows mostly into this lightly doped n_1 region rather than into the more heavily doped p_1 region, as shown in Fig. 17.32(b).

When v_{AK} is positive, junctions J_1 and J_3 are forward biased and junction J_2 , the only reverse-biased junction, withstands the voltage. As the n_1 region is also more lightly doped than the p_2 region, the SCL again grows mostly into the n_1 region, as shown in Fig. 17.32(c).

The n_1 region is therefore used to block both polarities of voltage when the SCR is off. Its doping level and length must be chosen to give the desired breakdown voltage. Note that breakdown can occur in two ways: *punch-through*, where the SCL extends far enough to reach the opposite edge of the n_1 region, and *avalanche*, where the peak electric field becomes high enough to cause impact ionization. Most SCRs are designed with the n_1 region long enough to cause avalanche to be the breakdown mechanism.

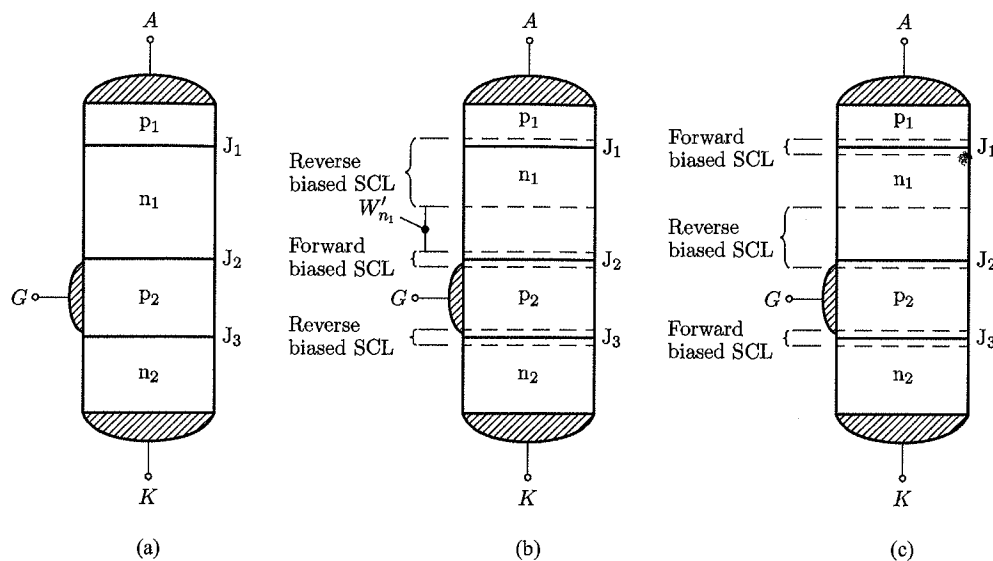


Figure 17.32 A one-dimensional model of the SCR. (a) The four layers and their junctions. (b) Space-charge layers when the SCR is blocking a reverse voltage. (c) Space-charge layers when the SCR is blocking a forward voltage.

Regenerative Turn-On Process To understand how the SCR can go from a forward-biased off state to its conducting on state, it is useful to view the device as a cross connection of two transistors. Figure 17.33(a) shows how the four-layer device can be considered as the interconnection of two three-layer devices – one a pnp and the other an npn transistor. The base of each transistor is connected to the collector of the other. Figure 17.33(b) illustrates this two-transistor SCR model.

Note that the base of each transistor is driven with a current that is β_F times its collector current. As long as the product $\beta_1\beta_2 > 1$, we can show that – once some current flows – the two transistors will drive each other harder and harder until they saturate. A short pulse of current into the gate terminal is one way to start this regenerative process. To desensitize the turn-on process to noise at the gate or leakage current at J₂ (which serves as base current for Q₂), Q₂ is designed to have a very low β_F at low emitter currents. The natural drop-off of β_F at low currents, illustrated in Fig. 17.16, is enhanced in Q₂ by fabrication techniques, one of which is to use the cathode metalization to provide a partial short between the gate and the cathode, as illustrated in Fig. 17.34. In simplest terms this can be modeled as a resistor between the base and emitter of Q₂, which shunts some of the gate current around J₃.

Note that the Q₁ collector leakage current functions as gate current and can turn the device on if it is too high. This is the reason that the SCR is a *silicon* device, and cannot be fabricated in germanium as the leakage current of a Ge transistor is orders of magnitude greater than that of Si.

The turn-on process is initiated by the injection of carriers into the n₂ region. A number of mechanisms can do this, some of which are more useful than others. Since collector leakage at

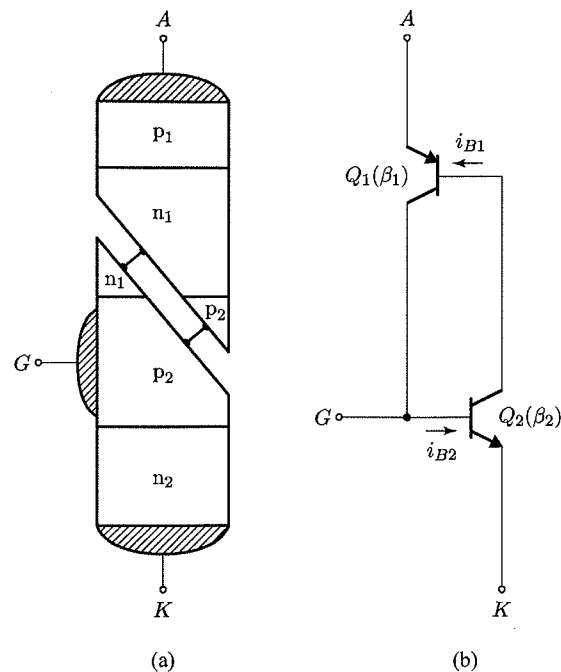


Figure 17.33 A two-transistor model of the SCR. (a) The four-layer device represented as the interconnection of two three-layer devices. (b) The two-transistor model of the interconnection of (a).

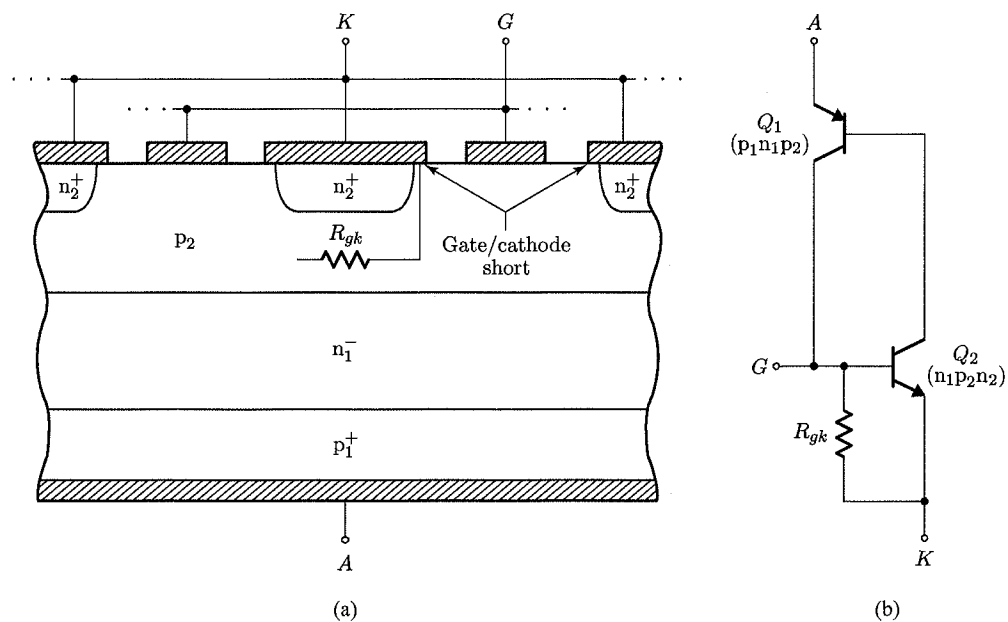


Figure 17.34 (a) A gate-cathode short produced on the die by making contact between the gate region and the cathode metalization. (b) The two-transistor SCR model showing the location of the gate-cathode short R_{gk} .

Q_1 is strongly temperature dependent, SCRs can be triggered if the device is operated above its temperature limit. If dv_{AK}/dt is high enough, the current at the Q_1 collector junction (J_2) capacitance will be sufficient to initiate turn-on, so all transistors have a maximum dv_{AK}/dt limit, typically between 200 and 2000 V/ μ s. (The time t_q as defined in Fig. 15.9 is usually specified under the assumption that the subsequent rise of the anode voltage is at this maximum rate, which is a worst-case condition.) As v_{AK} increases and J_2 approaches breakdown, its increasing current will also cause the device to turn on. And lastly, some devices are designed to allow optical access to the p_2 region and can be triggered by optical injection of carriers. This mechanism is very useful when a string of thyristors is used in very high-voltage applications, such as HVDC.

As an SCR turns on, it first begins to carry current in the regions of the die near the gate lead. Only after some time (the length of which depends on how finely interdigitated the gate and cathode metal contacts are) has the turned-on region spread sufficiently that the current is uniformly distributed over the entire device. During this *current-spreading* process, in which only a fraction of the SCR is conducting, it is important to keep the total anode current low enough to avoid a localized region of excessively high and destructive current density. For this reason, a maximum rate-of-rise of anode current is specified for all SCRs. Typically, di_A/dt must be kept below 1000 A/ μ s.

17.9.2 On State

As the two transistors drive each other into saturation, the excess carrier concentrations in their base regions reach high-level injection. At this point, the actual doping concentrations in the base regions are no longer relevant, and the SCR now behaves as a three-layer pin diode, with the two middle layers of the thyristor, n_1 and p_2 , corresponding to the intrinsic region of the diode. Figure 17.35 shows typical excess carrier concentrations in the SCR under this condition. Everything you learned in Section 17.2 about the pin diode's forward voltage drop as a function of current and design parameters also applies to the SCR in its on state.

Two parameters relevant to the on state are the *latching* and *holding* currents. The latching current is the minimum anode current necessary for the SCR to remain on when the gate drive is removed. At this time the SCR is not yet conducting over its entire area. Once the device is fully on, the holding current is the minimum anode current at which it will remain on without a gate current. The holding current is lower than the latching current because during turn-on the

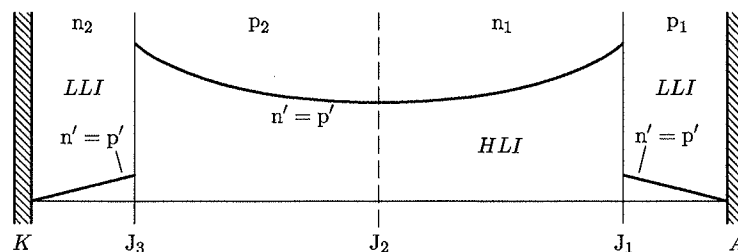


Figure 17.35 Excess carrier concentrations in an SCR in its on state. The two middle regions are in high-level injection, causing behavior like a pin diode.

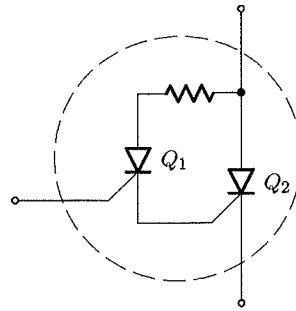


Figure 17.36 Darlington-type connection of two SCRs to produce an amplifying gate.

gate current aids the injection of electrons from the J_3 emitter as injection spreads from the J_3 region near the gate contact. When the device is fully on, the J_3 junction is uniformly forward biased (the device is conducting across its entire area) and providing adequate injection to the p_2 region to maintain the regenerative process without help from the gate.

We can connect two SCRs in a Darlington-like way, as shown in Fig. 17.36, in order to reduce the gate current requirement. In large devices, Q_1 is integrated on the same die as the main SCR, Q_2 , and is known as an *amplifying gate*. A special version of this device uses a photosensitive SCR for Q_1 . A photo-diode (LED) is then mounted in the same package and optically coupled to Q_1 . The combination provides a means of electrically isolating the gate drive circuit from the SCR. However, these *opto-SCRs* often have a very low dv_{AK}/dt rating.

17.9.3 Turn-Off Process

Once the SCR is on, it behaves as a pin diode because the two center regions are well into high-level injection for typical anode currents. When the power circuit tries to reverse the anode current, the SCR displays a reverse recovery characteristic similar to that of the pin diode. Current will flow in the negative direction until the excess charge stored in the center regions of the SCR is either removed or lost to recombination.

The same trade-off made between the speed of recovery and the on-state voltage of a pin diode can be made for an SCR. *Rectifier-grade* SCRs trade speed for a reduced forward drop. They generally have a simple center-gate electrode geometry, and are intended for line commutation applications of up to about 400 Hz. *Inverter-grade* SCRs are intended for higher-frequency applications. These devices are generally constructed with a highly interdigitated gate geometry. The lifetime for excess carriers in the base regions is made relatively short to give a fast reverse recovery at the expense of a higher forward voltage drop. Inverter-grade SCRs generally have a forward drop that is about 50% higher than a rectifier-grade device with comparable voltage and current ratings. As noted earlier, SCRs have largely been replaced by MOSFETs or IGBTs in high-frequency applications.

Turn-Off Time and Rate Effects When the SCR is off and blocking a forward voltage, it can be pushed into its regenerative turn-on mode if the current flowing through the device

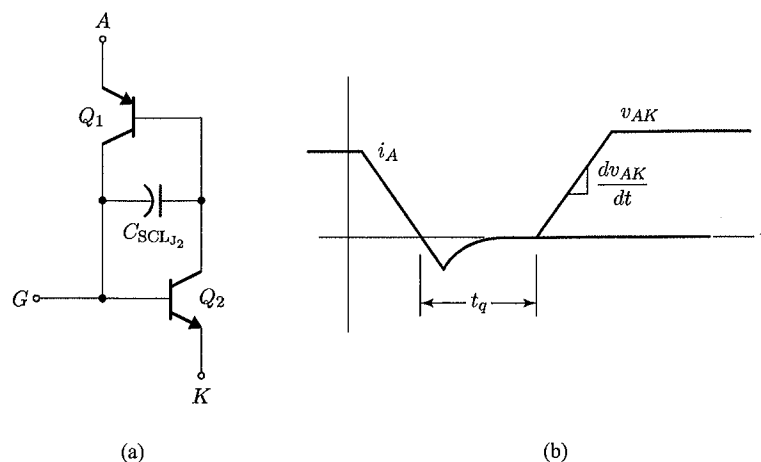


Figure 17.37 (a) Two-transistor model of an SCR showing the capacitance of the SCL at junction J_2 . (b) Reapplied forward voltage at the end of the minimum turn-off time t_q .

exceeds a critical value. One way to generate this current (besides driving the gate terminal) is to apply a positive change in voltage across the SCL of junction J_2 . As Fig. 17.37(a) shows, the junction can be modeled as having a parallel capacitance (because of the SCL), and the current flowing through this capacitance when dv_{AK}/dt is positive acts just like gate current. Therefore, if dv_{AK}/dt is large enough, the regenerative turn-on mechanism in the SCR can be triggered. For this reason, there is a maximum specified off-state dv_{AK}/dt above which the device is not guaranteed to remain off. Rates of 1000–2000 V/ μ s are typical for inverter-grade devices.

Note that the cathode short, which reduces the gain of Q_2 , also helps to make the SCR less sensitive to dv_{AK}/dt . In addition, applying a negative voltage across the gate–cathode terminals when the SCR is off also helps reduce the chance of false triggering.

Because the J_2 SCL capacitance is nonlinear, the maximum allowable dv_{AK}/dt is specified at the largest value of this capacitance. This condition occurs when $v_{AK} = 0$. Moreover, the SCR is most sensitive to triggering immediately after it has been turned off, as excess charge still exists in the middle base regions. For these reasons, the maximum allowable dv_{AK}/dt is specified at a fixed time after the anode current reverses and the SCR begins its reverse recovery phase. This time is called the *circuit commutated turn-off time* and is designated by t_q . A forward voltage can safely be reapplied (meaning that retriggering will not take place) at the specified rate so long as it is delayed for a time t_q past the time that the anode current reverses. The relationships among i_A , v_{AK} , and t_q during the turn-off process are shown in Fig. 17.37(b). Typical values for t_q are 100–200 μ s for rectifier-grade SCRs and 10–20 μ s for inverter-grade SCRs.

The circuit commutated turn-off time t_q depends on whether the circuit in which the SCR is connected permits the flow of significant reverse recovery current. In a situation in which an anti-parallel diode is connected across the SCR (a common configuration), the switch current will reverse through the diode, but little reverse current will be carried by the SCR. The result is that elimination of the charge stored in the base regions of the SCR is not aided by reverse current. Instead, the charge will decay through recombination at a rate determined by the excess

carrier lifetimes in the base regions. In this case, t_q will be longer than when an anti-parallel diode is absent. Datasheets generally give t_q under both conditions for inverter-grade SCRs, and their values differ by 50%–100%.

17.9.4 Gate Turn-Off Thyristor

So far we have assumed that a thyristor, once it is on, will not turn off until the power circuit drives its anode current to zero. But the two-transistor thyristor model of Fig. 17.33(b) implies that if we draw enough current out of the gate terminal of the device, we can reduce i_K to a value insufficient to maintain the regenerative process and the thyristor will turn off. We cannot do this in a conventional SCR, because the gate terminal affects only the region of the device in the immediate vicinity of the gate contact. The lateral voltage drop that would accompany the flow of current from regions remote from the gate contact effectively prevents the behavior of these regions being influenced by anything happening at the gate terminal. Therefore, while a negative gate current might stop the regenerative process in the immediate vicinity of the gate, the rest of the SCR would remain on. However, by using special design techniques, including dense gate interdigitation and reducing the gain of Q_1 , we can design a thyristor to be turned off by a large negative gate current. Such a device is known as a *gate turn-off* (GTO) thyristor. One consequence of the structure necessary to create a GTO is that it loses its reverse blocking capability, now sustaining a reverse voltage on the order of only 50 V.

17.9.5 TRIAC

We described the operating characteristics of the TRIAC in Chapter 3. The device is a five-layer structure, as shown in Fig. 17.38. The MT_1 and MT_2 contacts short the n_2^+ region to the p_2

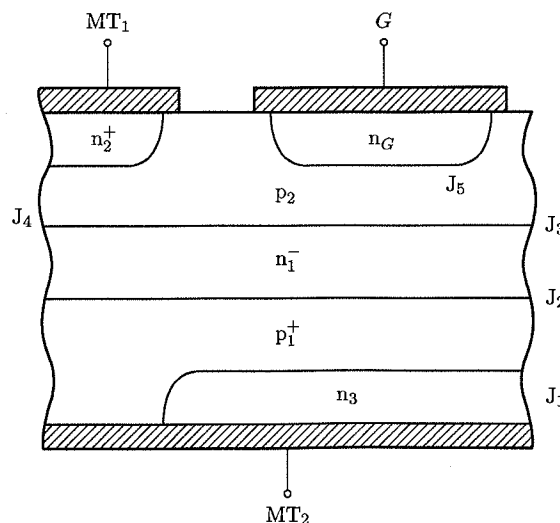


Figure 17.38 The TRIAC structure.

region and the n_3 region to the p_1^+ region. The result is the equivalent of two anti-parallel SCR structures: $p_1^+-n_1^--p_2-n_2^+$ and $n_3-p_1^+-n_1^+-p_2$. The n_G region is the TRIAC gate, and it is shorted to the p_2 region by the gate contact.

The possible modes of operation of the TRIAC are defined in Section 15.3.3 as I+, I−, III+, and III−. In mode I+, the device is operating as a conventional SCR, with the anode being the p_1 region. The gate is the p_2 region. It is connected directly to the gate terminal, because the gate metal extends beyond the edge of the n_G diffusion, shorting the terminal to the p_2 region. Therefore, triggering in this mode is effected in the normal manner, and this is the mode with the most sensitive triggering characteristic.

III− is the next most sensitive mode. In this mode the n_3 region is the cathode and the p_2 region the anode. Triggering is accomplished by pulling the gate negative with respect to MT_1 . This forward biases junction J_5 , which functions as the base-emitter junction of the npn transistor composed of n_1 , p_2 , and n_G . The injected current (electrons) is collected at J_3 , which is reverse biased in quadrant III. The resulting electron flow into the n_1 region is effectively base current for the $p_1n_1p_2$ transistor, and it begins to turn on, initiating regeneration and turning the TRIAC on.

17.10 Datasheet

All semiconductor devices are accompanied by a datasheet containing parameters that characterize the device. The meaning of some of the parameters is self-evident – $I_{\max}$, $V_{\max}$, T_J for example. Others are unique to the device type – V_T for the MOSFET or β_F for the BJT. Other parameters are more universal, but often complicated to interpret and apply. One such example is the safe operating area (SOA). Furthermore, many parameters are given at a specified temperature which is usually different from that in the application environment, requiring derating. And lastly, many parameters and operating characteristics are given in parametric graphs. In this section we discuss some of these relevant but somewhat confusing parameters, focusing on the MOSFET.

17.10.1 Switching Characteristics

Switching times and energy loss are based on the clamped inductive load circuit of Fig. 17.39, where the freewheeling diode D can be either an external diode or the body diode of the complementary MOSFET in a bridge leg. The inductor L is large enough for the device under test to be switched at a frequency and duty ratio which maintains the inductor current at the test value.

Although this circuit may not represent the application configuration, its use is nearly standard in the device industry, providing a common metric with which to compare the switching performance of different devices. The definition of switching times is shown in Fig. 17.39(b). These parameters and their associated losses E_{on} and E_{off} are given in the datasheet for one (typically a low) value of R_G , but graphs of these parameters are also given that are parametric in R_G . The junction temperatures at which the single data point and graphs are measured is also given, but may not be the same. For example, the single datasheet parameter may be at

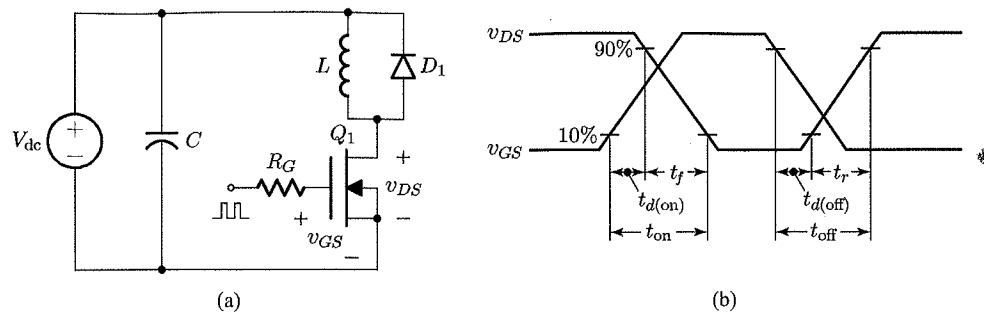


Figure 17.39 (a) Clamped inductive load test circuit used to determine MOSFET switching times and energy. (b) Waveforms of v_{GS} and V_{DS} showing how the different switching parameters are defined.

$T_J = 175^\circ\text{C}$ while the graphs give losses at $T_J = 25^\circ\text{C}$. There may be a graph included which provides the temperature dependence of loss versus T_J for one value of R_G from which the temperature dependence at other values of R_G may be inferred.

17.10.2 Thermal Parameters

Keeping T_J below its maximum allowed temperature, which is generally given on the first page of the datasheet, is critical to maintaining the integrity of the device. The primary parameter for the circuit designer is $R_{\theta JC}$, the *thermal resistance* between the device junction and its case. Thermal design and the use of this parameter is covered in Chapter 25, but in summary it relates the dissipated power to the difference between case and junction temperatures. It is a single number with units of $^\circ\text{C}/\text{W}$. Further thermal characterization is provided in a graph of *thermal impedance* versus pulse width with duty ratio as a parameter. The use of this graph is also covered in Chapter 26.

Another datasheet parameter related to the device's thermal behavior is maximum power dissipation, P_D . This number is generally surprisingly high because it is calculated based on $T_J = T_{J\max}$ and $T_C = 25^\circ\text{C}$, conditions difficult to replicate in practice. A graph is always included which provide a power derating curve for $T_J = T_{J\max}$ and $T_{J\min} < T_C < T_{J\max}$.

17.10.3 Device Capacitances

The device capacitances C_{iss} , C_{oss} , and C_{rss} are given in two places on the datasheet: a single numerical value with test conditions noted, and, because these capacitances are nonlinear, a graph of their incremental values versus V_{DS} . Since C_{oss} figures prominently in switching behavior, two additional linear parameters are provided: the *energy-related* capacitance $C_{o(er)}$ and the *time-related* capacitance $C_{o(tr)}$. The former is the capacitance which stores the same energy as C_{oss} as V_{DS} rises from 0 to a specified value. The latter is the capacitance which, when charged at the device's specified continuous drain current I_D , results in the same charging time to the specified V_{DS} as C_{oss} .

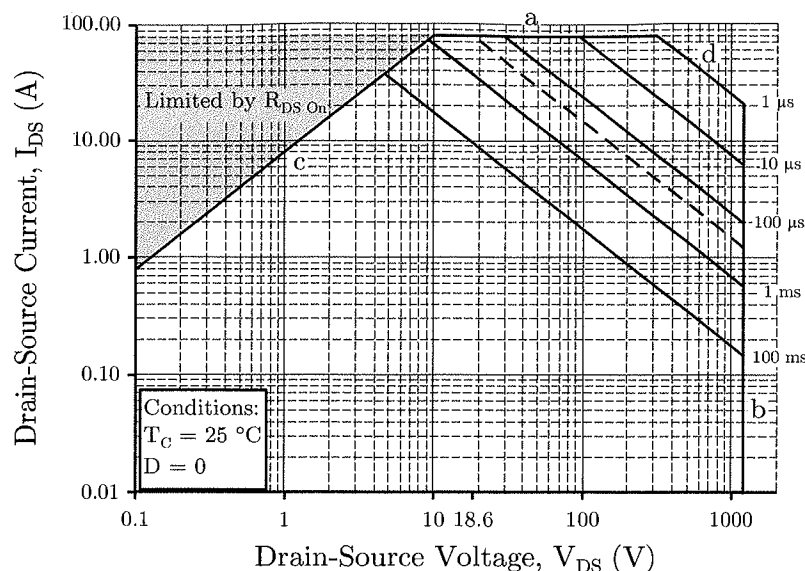


Figure 17.40 The safe operating area (SOA) graph for the Wolfspeed C3M0075120D SiC MOSFET. (Used with permission of Wolfspeed, Inc.)

17.10.4 Safe Operating Area

The SOA is presented in graphical form as shown in Fig. 17.40 and provides a region in the i_D - v_{DS} plane in which the device will operate safely at the specified T_C , which is usually 25 °C. It is most useful in applications where high I_D and V_{DS} occur simultaneously, such as non-switching (linear) audio amplifiers, or during the switching transient. Boundary (a) is defined by the datasheet parameter $I_{D(\text{pulse})}$, which is limited by the die bonding system and packaging constraints, and boundary (b) is constrained by the datasheet value of $V_{DS(\text{max})}$, sometimes identified as V_{DSS} . Boundary (c) shows the constraint on V_{DS} at a given I_d due to $R_{DS(\text{on})}$, that is, $V_{DS} = R_{DS(\text{on})}I_D$. The 45° lines (d) are the constant power loci for single pulses of different widths, each of which results in $T_J = T_{J(\text{max})}$. Since the test condition is $T_C = 25$ °C, for most applications the curves (d) have to be derated, reducing the SOA. The process of derating these curves is described in Chapter 25.

17.10.5 Diode Parameters

The body diode of the MOSFET is characterized in the datasheet with respect to forward voltage, continuous current, pulse current capability, and reverse parameters. The most useful of these is usually the reverse recovery time, t_{rr} .

17.10.6 Graphed Parameters

In addition to those graphed parameters already mentioned, other device parameters are also graphed as functions of T_C , T_J , V_{DS} , or V_{GS} . Their use is self-evident. The most useful of

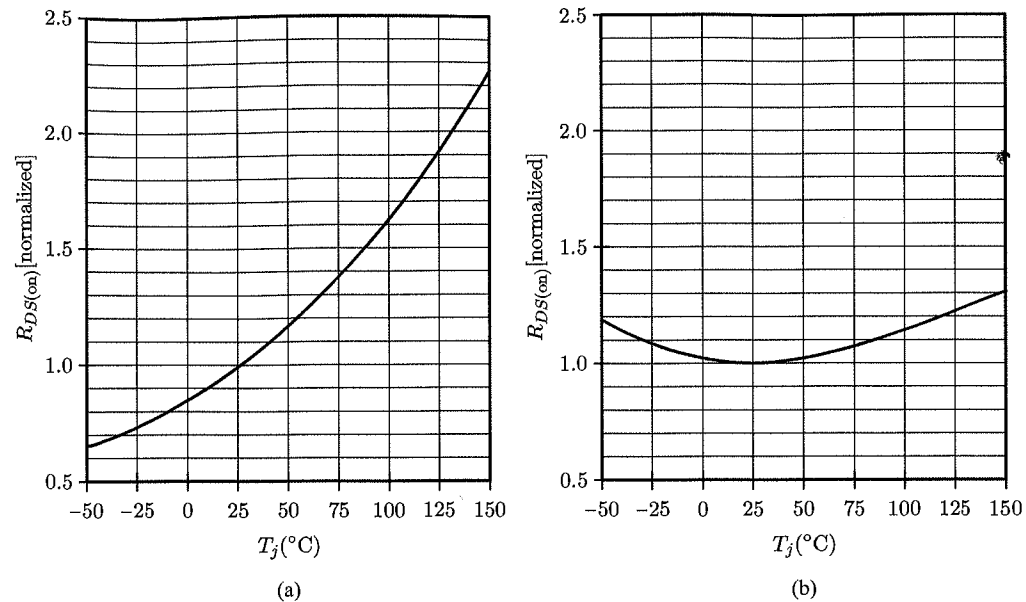


Figure 17.41 The temperature dependence of $R_{DS(on)}$ for (a) an IPW60R018CFD7 Si MOSFET, and (b) an IMW65R027M1H SiC MOSFET. (Used with permission of Infineon Technologies.)

these graphs are those which characterize $R_{DS(on)}$ as functions of T_J , I_D , and V_{GS} . Figure 17.41 illustrates the behavior of $R_{DS(on)}$ versus T for both a Si and SiC MOSFET which have similar voltage ratings. The datasheet generally gives a value of $R_{DS(on)}$ at $T_J = 25^\circ\text{C}$, and the graph provides its temperature dependence normalized to this value. The Si MOSFET exhibits substantial variation with T_J , increasing by approximately 125% over its operating temperature range. However, because of different mechanisms affecting channel and drift region mobilities in the SiC device, its $R_{DS(on)}$ variation with T_J is substantially less, increasing by only 30% over the same temperature range. There are frequently additional graphs which illustrate how V_{GS} and I_D affect $R_{DS(on)}$.

Notes and Bibliography

The best general references on the material presented in this chapter are [1] and [2]. Baliga's book has extensive coverage of all the devices discussed in this chapter, but focuses more on device physics than phenomenological behavior. Sze, while not devoted to power devices, is another source of how and why device parameters vary with temperature, dopant concentration, etc. A comprehensive treatment of the Schottky diode and the physics of thermionic emission is also provided.

Chapter 3 of [3] by J. D. Plummer and R. D. Blanchard provides an excellent discussion of the power MOSFET development and addresses the difference in the i_D - v_{GS} relationship between signal- and power-level devices. Udrea et al. [4] provides an excellent discussion of superjunction technology, with an extensive reference list for those wishing to explore the technology further.

A good appreciation of the benefits of SiC can be found in [5] and [6]. The Shinkansen application in [6] is especially interesting for the system implications of changing from Si to SiC devices. The problem

of device output capacitance loss explored in [7], though addressing the Schottky diode, is relevant to the similar phenomenon in Si SJFETs and GaN HEMTs.

A good discussion of the dynamic on-state resistance of GaN FETs can be found in [8]. The paper also presents methods of measuring the phenomenon. Although short on physics, [9] provides extensive information on the application of GaN devices.

1. B. J. Baliga, *Fundamentals of Power Semiconductor Devices*, 2nd ed., Cham: Springer, 2019.
2. S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed., New York: Wiley, 1981.
3. J. D. Plummer and R. A. Blanchard, in P. Antognetti, Ed., *Power Integrated Circuits: Physics, Design, and Applications*, New York: McGraw-Hill, 1986, pp. 3.1–3.58.
4. F. Udrea, G. Deboy, and T. Fujihara, “Superjunction Power Devices, History, Development, and Future Prospects,” *IEEE Trans. Electronic Devices*, vol. ED-64, pp. 713–727, Mar. 2017.
5. A. O. Adan, et al., “The Current Status and Trends of 1,200 V Commercial Silicon-Carbide MOS-FETs,” *IEEE Power Electronics Magazine*, pp. 36–47, June 2019.
6. K. Sato, H. Kato, and T. Fukushima, “Development of SiC Applied Traction System for Next Generation Shinkansen High-Speed Trains,” *IEEE J. Industry Applications*, vol. 9, pp. 453–459, 2020.
7. Z. Tong, G. Zulauf, J. Xu, J. Plummer, and J. Rivas-Davila, “Output Capacitance Loss Characterization of Silicon Carbide Schottky Diodes,” *IEEE J. Emerging and Selected Topics in Power Electronics*, vol. 7, pp. 865–878, June 2019.
8. R. Li, X. Wu, S. Yang, and K. Sheng, “Dynamic On-State Resistance Test and Evaluation of GaN Power Devices Under Hard- and Soft-Switching Conditions by Double and Multiple Pulses,” *IEEE Trans. Power Electronics*, vol. 34, pp. 1044–1053, Feb. 2019.
9. A. Lidow, J. Strydom, M. de Rooij, and D. Reusch, *GaN Transistors for Efficient Power Conversion*, 2nd ed., Chichester: Wiley, 2015.

PROBLEMS

- 17.1 In determining the voltage v_i across the intrinsic region of the pin diode we stated that “ $D_e = D_h = D_a$ and therefore $\mu_e = \mu_h$ in the intrinsic region.” Justify the “therefore” in this statement.
- 17.2 Reconsider Example 17.2 but with a doping density in the n^- region of $10^{13}/\text{cm}^3$ and a calculated n^- region width of $125\ \mu\text{m}$, which now creates a punch-through pin diode. Sketch the electric field in the diode and calculate its breakdown voltage. At this doping density the critical field is approximately $E_c = 10^5\ \text{V/cm}$.
- 17.3 In Section 17.4.1 we noted that the collector side of the collector junction is lightly doped to allow the SCL to expand into it. Why could we not instead create an intrinsic region in the base with a highly doped field-stop layer at the emitter junction to allow the SCL to grow into the base? (*Hint*: consider the issues discussed in Section 16.3.2.)
- 17.4 Figure 17.42(a) shows the vertical profile of a 77 mm-diameter fast-recovery rectifier rated at $3000\ \text{A}_{\text{rms}}$. Use the parameters as given in Fig. 17.42(a) for parts (a) and (b) of the problem.

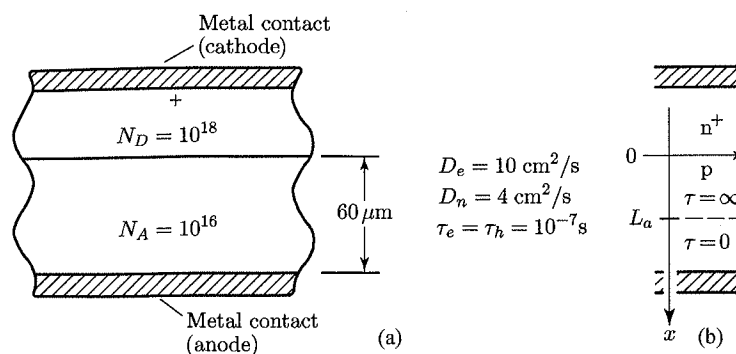


Figure 17.42 The 77 mm-diameter fast-recovery diode of Problem 17.4, and its lifetime profile.

- Determine the value of current density at which the device enters high-level injection on the p side of the SCL.
- Assuming no *bulk resistive drop*, that is, no drop outside the SCL, determine the terminal voltage of this diode at the rated current. Use the values for diffusion constants and lifetimes given in Fig. 17.42.
- What is the terminal voltage at the rated current if you do not ignore the drop across the p region outside the SCL? For the purposes of this calculation, assume that you can characterize the p region by an infinite lifetime within a diffusion length, L_a , of the junction and by a lifetime of zero throughout the rest of the p region, as shown in Fig. 17.42(b).

17.5 Figure 17.43 conceptually illustrates a MOSFET drain region whose physical dimensions and resistance you are to compare for fabrication in Si and SiC. Use the parameters for Si and SiC in Table 16.1 and assume that the drain in both cases will be required to block 1 kV.

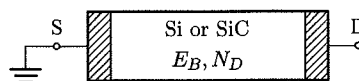


Figure 17.43 A stylized representation of the drain region of a MOSFET having length ℓ_d to be analyzed for fabrication in Si and SiC.

- Sketch the E -field and label its maximum value for both cases if the drain is fully depleted when the device is blocking 1 kV.
- Under the conditions of (a), determine the length ℓ_d for both materials.
- Determine the doping density, N_D , consistent with the answers to (a) and (b).
- Determine and compare $R_{DS(on)}$ for the Si and SiC drains.

17.6 This problem explores the effect of the nonlinearity of the device capacitances C_{iss} , C_{rss} , and C_{oss} on DMOS switching behavior. A DMOS switching a clamped inductive load, along with its graphically represented capacitances, are shown in Fig. 17.44. The device is carrying 50 A when it is turned off. Assume $v_{GS} = 0$ instantaneously at $t = 0$.

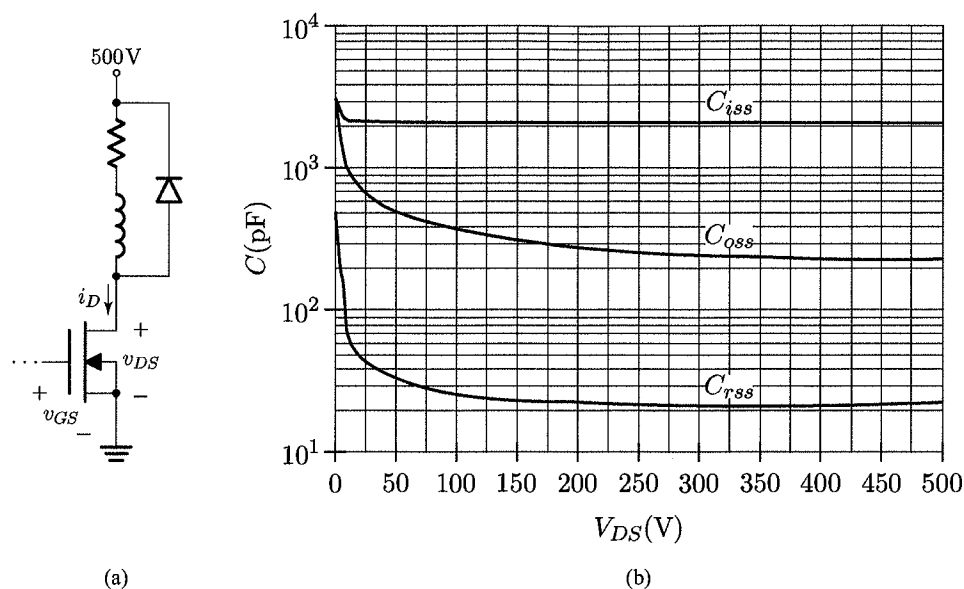


Figure 17.44 (a) A DMOS switching a clamped inductive load. (b) Capacitances as a function of V_{DS} for the DMOS of (a).

- Redraw Fig. 17.44(a) with its terminally viewed capacitances in terms of C_{iss} , C_{rss} , and C_{oss} . (You might want to refer to Fig. 16.15).
- Which capacitance(s) will dominate the turn-off transient?
- Draw the circuit relevant to determining the transient.
- Approximating the capacitances as constant at their average values during discrete intervals of $\Delta v_{DS} = 50$ V, plot $v_{DS}(t)$ during turn-off and estimate its rise time.
- The time-related output capacitance, $C_{o(tr)}$, for this device is 370 pF. Compare your estimated rise time to that calculated using $C_{o(tr)}$.