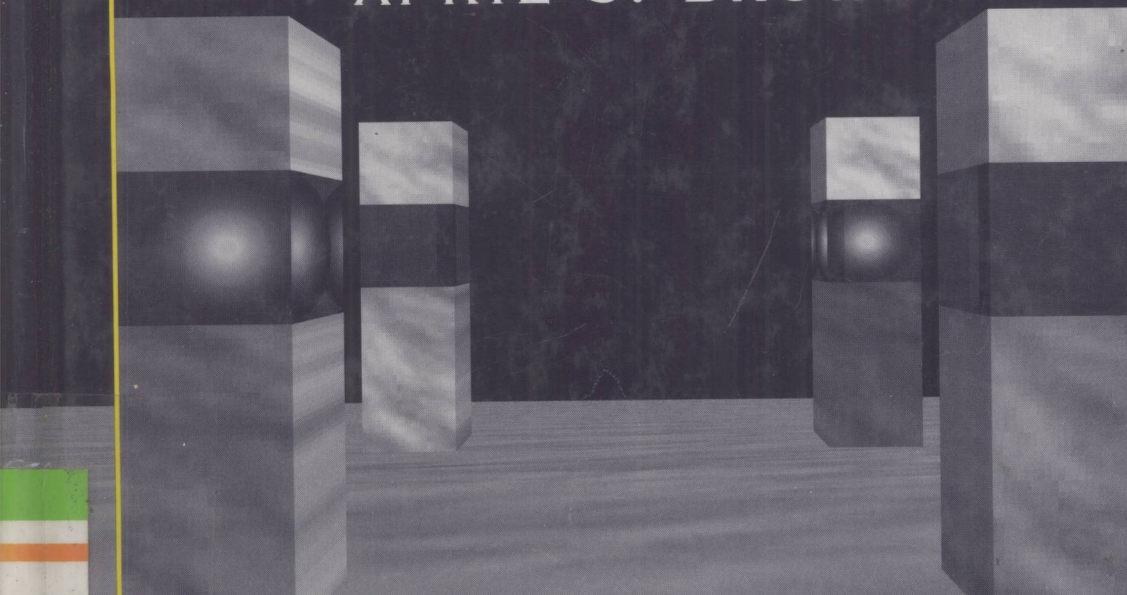


# THEORY OF MODERN ELECTRONIC SEMICONDUCTOR DEVICES

---

KEVIN F. BRENNAN  
APRIL S. BROWN



TN 303  
B 838

---

# THEORY OF MODERN ELECTRONIC SEMICONDUCTOR DEVICES

---

**KEVIN F. BRENNAN**

**APRIL S. BROWN**

Georgia Institute of Technology



E200201490

A Wiley-Interscience Publication  
**JOHN WILEY & SONS, INC.**

This book is printed on acid-free paper. ☺

Copyright © 2002 by John Wiley & Sons, Inc., New York. All rights reserved.

Published simultaneously in Canada.

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except as permitted under Sections 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 750-4744. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 605 Third Avenue, New York, NY 10158-0012, (212) 850-6011, fax (212) 850-6008, E-Mail: PERMREQ@WILEY.COM.

For ordering and customer service, call 1-800-CALL-WILEY.

*Library of Congress Cataloging-in-Publication Data Is Available*

ISBN 0-471-41541-3

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

THEORY OF  
MODERN ELECTRONIC  
SEMICONDUCTOR  
DEVICES

To our families,  
Lea and Casper  
and  
Bob, Alex, and John

# PREFACE

---

The rapid advancement of the microelectronics industry has continued in nearly exponential fashion for the past 30 years. Continuous progress has been made in miniaturizing integrated circuits, thus increasing circuit density and complexity at reduced cost. These circumstances have fomented the continuous expansion of computing capability that has driven the modern information age. Explosive growth is occurring in computing technology and communications, driven mainly by the advancements in semiconductor hardware. Continued growth in these areas depends on continued progress in microelectronics.

At this writing, critical device dimensions for commercial products are already approaching  $0.1\text{ }\mu\text{m}$ . Continued miniaturization much beyond  $0.1\text{-}\mu\text{m}$  feature sizes presents myriad problems in device performance, fabrication, and reliability. The question is, then, will microelectronics technology continue in the same manner as in the past? Can continued miniaturization and its concomitant increase in circuit speed and complexity be maintained using current CMOS technology, or will new, radically different device structures need to be invented?

The growth in wireless and optical communications systems has closely followed the exponential growth in computing technology. The need not only to process but also to transfer large packets of electronic data rapidly via the Internet, wireless systems, and telephony is growing at a brisk rate, placing ever increasing demands on the bandwidth of these systems. Hardware used in these systems must thus be able to operate at ever higher frequencies and output power levels. Owing to the inherently higher mobility of many compound semiconductor materials compared to silicon, currently most high-frequency electronics incorporate compound semiconductors such as GaAs

and InP. Record-setting frequency performance at high power levels is invariably accomplished using either heterostructure field-effect or heterostructure bipolar transistors. What, though, are the physical features that limit the performance of these devices? What are their limits of performance? What alternatives can be utilized for high-frequency-device operation?

Device dimensions are now well within the range in which quantum mechanical effects become apparent and even in some instances dominant. What quantum mechanical phenomena are important in current and future semiconductor devices? How do these effects alter device performance? Can nanoelectronic devices be constructed that function principally according to quantum mechanical physics that can provide important functionality? How will these devices behave?

The purpose of this book is to examine many of the questions raised above. Specifically, we discuss the behavior of heterostructure devices for communications systems (Chapters 2 to 4), quantum phenomena that appear in miniaturized structures and new nanoelectronic device types that exploit these effects (Chapters 5, 6, and 9), and finally, the challenges faced by continued miniaturization of CMOS devices and futuristic alternatives (Chapters 7 and 8). We believe that this is the first textbook to address these issues in a comprehensive manner. Our aim is to provide an up-to-date and extended discussion of some of the most important emerging devices and trends in semiconductor devices. The book can be used as a textbook for a graduate-level course in electrical engineering, physics, or materials science. Nevertheless, the content will appeal to practicing professionals. It is suggested that the reader be familiar with semiconductor devices at the level of the books by Streetman or Pierret. In addition, much of the basic science that underlies the workings of the devices treated in this text is discussed in detail in the book by Brennan, *The Physics of Semiconductors with Applications to Optoelectronic Devices*, Cambridge University Press, 1999. The reader will find it useful to refer to this book for background material that can supplement his or her knowledge aiding in the comprehension of the current book.

The book contains nine chapters in total. The first chapter provides an overview of emerging trends in compound semiconductors and computing technology. We have tried to focus the book on the three emerging areas discussed above: telecommunications, quantum structures, and challenges and alternatives to CMOS technology. The balance of the book examines these three issues in detail. There are sections throughout that can be omitted without loss of continuity. These sections are marked with a dagger. We end the book with a chapter on magnetic field effects in semiconductors. It is our belief that although few devices currently exploit magnetic field effects, the unusual physical properties of reduced dimensional systems when exposed to magnetic fields are of keen interest and may point out new directions in semiconductor device technology. Again, the instructor may elect to skip Chapter 9 completely without compromising the main focus of the book.

From a pedagogic point of view, we have developed the book from class notes we have written for a one-semester graduate-level course given in the School of Electrical and Computer Engineering at the Georgia Institute of Technology. This course is generally taught in the spring semester following a preparatory course taught in the fall. Most students first study the fall semester course, which is based on the first nine chapters of the book by Brennan, *The Physics of Semiconductors with Applications to Optoelectronic Devices*. Nevertheless, the present book can be used independent of a preparatory course, using the book by Brennan as supplemental reference material. The present book is fully self-contained and refers the reader to Brennan's book only when needed for background material. Typically, we teach Chapters 2 to 8 in the current book, omitting the optional (Sections 2.5, 5.7, 6.3, and 7.1). The students are asked to write a term paper in the course following up in detail on one topic. In addition, homework problems and a midterm and final examinations are given. The reader is invited to visit the book Web site at [www.ece.gatech.edu/research/labs/comp\\_elec](http://www.ece.gatech.edu/research/labs/comp_elec) for updates and supplemental information. At the book Web site a password-protected solutions manual is available for instructors, along with sample examinations and their solutions.

We would like to thank our many colleagues and students at Georgia Tech for their interest and helpful insight. Specifically, we are deeply grateful to Dr. Joe Haralson II, who assisted greatly in the design of the cover and in revising many of the figures used throughout. We are also grateful to Tsung-Hsing Yu, Dr. Maziar Farahmand, Louis Tirino, Mike Weber, and Changhyun Yi for their help on technical and mechanical aspects of manuscript preparation. Additionally, we thank Mike Weber and Louis Tirino for setting up the book Web site. Finally, we thank Dr. Dan Tsui of Princeton University, Dr. Wolfgang Porod of Notre Dame University, Dr. Mark Kastner of MIT, Dr. Stan Williams of Hewlett-Packard Laboratories, and Dr. Paul Ruden of the University of Minnesota at Minneapolis for granting permission to reproduce some of their work in this book and for helpful comments in its construction.

Finally, both of us would like to thank our families and friends for their enduring support and patience.

Atlanta  
November 2000

KEVIN F. BRENNAN  
APRIL S. BROWN



# CONTENTS

---

<b>PREFACE</b>	<b>xi</b>
<b>1 OVERVIEW OF SEMICONDUCTOR DEVICE TRENDS</b>	<b>1</b>
1.1 Moore’s Law and Its Implications	1
1.2 Semiconductor Devices for Telecommunications	7
1.3 Digital Communications	11
<b>2 SEMICONDUCTOR HETEROSTRUCTURES</b>	<b>14</b>
2.1 Formation of Heterostructures	14
2.2 Modulation Doping	20
2.3 Two-Dimensional Subband Transport at Heterointerfaces	25
2.4 Strain and Stress at Heterointerfaces	45
2.5 Perpendicular Transport in Heterostructures and Superlattices	57
2.6 Heterojunction Materials Systems: Intrinsic and Extrinsic Properties	66
Problems	81
<b>3 HETEROSTRUCTURE FIELD-EFFECT TRANSISTORS</b>	<b>84</b>
3.1 Motivation	84
3.2 Basics of Heterostructure Field-Effect Transistors	88
3.3 Simplified Long-Channel Model of a MODFET	92
3.4 Physical Features of Advanced State-of-the-Art MODFETs	104
	<b>vii</b>

3.5	High-Frequency Performance of MODFETs	115
3.6	Materials Properties and Structure Optimization for HFETs	123
	Problems	127
<b>4</b>	<b>HETEROSTRUCTURE BIPOLAR TRANSISTORS</b>	<b>130</b>
4.1	Review of Bipolar Junction Transistors	130
4.2	Emitter–Base Heterojunction Bipolar Transistors	141
4.3	Base Transport Dynamics	152
4.4	Nonstationary Transport Effects and Breakdown	158
4.5	High-Frequency Performance of HBTs	170
4.6	Materials Properties and Structure Optimization for HBTs	183
	Problems	192
<b>5</b>	<b>TRANSFERRED ELECTRON EFFECTS, NEGATIVE DIFFERENTIAL RESISTANCE, AND DEVICES</b>	<b>195</b>
5.1	Introduction	195
5.2	$k$ -Space Transfer	196
5.3	Real-Space Transfer	206
5.4	Consequences of NDR in a Semiconductor	213
5.5	Transferred Electron-Effect Oscillators: Gunn Diodes	217
5.6	Negative Differential Resistance Transistors	220
†5.7	IMPATT Diodes	222
	Problems	232
<b>6</b>	<b>RESONANT TUNNELING AND DEVICES</b>	<b>234</b>
6.1	Physics of Resonant Tunneling: Qualitative Approach	234
6.2	Physics of Resonant Tunneling: Envelope Approximation	239
†6.3	Inelastic Phonon Scattering Assisted Tunneling: Hopping Conduction	249
6.4	Resonant Tunneling Diodes: High-Frequency Applications	258
6.5	Resonant Tunneling Diodes: Digital Applications	265
6.6	Resonant Tunneling Transistors	273
	Problems	276
<b>7</b>	<b>CMOS: DEVICES AND FUTURE CHALLENGES</b>	<b>279</b>
†7.1	Why CMOS?	279
7.2	Basics of Long-Channel MOSFET Operation	288

†Optional material.

CONTENTS	ix
7.3 Short-Channel Effects	297
7.4 Scaling Theory	310
7.5 Processing Limitations to Continued Miniaturization Problems	314
8 BEYOND CMOS: FUTURE APPROACHES TO COMPUTING HARDWARE	317
8.1 Alternative MOS Device Structures: SOI, Dual-Gate FETs, and SiGe	320
8.2 Quantum-Dot Devices and Cellular Automata	325
8.3 Molecular Computing	340
8.4 Field-Programmable Gate Arrays and Defect-Tolerant Computing	354
8.5 Coulomb Blockade and Single-Electron Transistors	358
8.6 Quantum Computing Problems	369
9 MAGNETIC FIELD EFFECTS IN SEMICONDUCTORS	379
9.1 Landau Levels	381
9.2 Classical Hall Effect	392
9.3 Integer Quantum Hall Effect	398
9.4 Fractional Quantum Hall Effect	407
9.5 Shubnikov–de Haas Oscillations Problems	413
REFERENCES	416
APPENDIX A: PHYSICAL CONSTANTS	419
APPENDIX B: BULK MATERIAL PARAMETERS	433
Table I: Silicon	435
Table II: Ge	436
Table III: GaAs	436
Table IV: InP	437
Table V: InAs	437
Table VI: InN	438
Table VII: GaN	438
Table VIII: SiC	439
Table IX: ZnS	439

Table X: ZnSe	440
Table XI: $\text{Al}_x\text{Ga}_{1-x}\text{As}$	440
Table XII: $\text{Ga}_{0.47}\text{In}_{0.53}\text{As}$	441
Table XIII: $\text{Al}_{0.48}\text{In}_{0.52}\text{As}$	441
Table XIV: $\text{Ga}_{0.5}\text{In}_{0.5}\text{P}$	442
Table XV: $\text{Hg}_{0.70}\text{Cd}_{0.30}\text{Te}$	442
<b>APPENDIX C: HETEROJUNCTION PROPERTIES</b>	<b>443</b>
<b>INDEX</b>	<b>445</b>

# CHAPTER 1

---

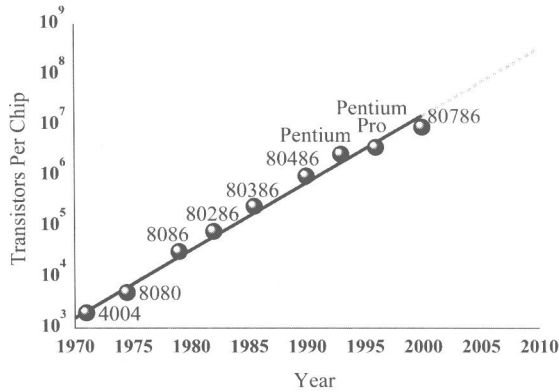
## Overview of Semiconductor Device Trends

---

The dawn of the third millennium coincides with what has often been referred to as the *information age*. The rapid exchange of information in its various formats has become one of the most important activities of our modern world. Shannon's early recognition that information in its most basic form can be reduced to a series of bits has led to a vast infrastructure devoted to the rapid and efficient transfer of information in bit form. The technical developments that underlie this infrastructure result from a blending of computing and telecommunications. Basic to these industries is semiconductor hardware, which provides the essential tools for information processing, transfer, and display.

### 1.1 MOORE'S LAW AND ITS IMPLICATIONS

The integrated circuit is the fundamental building block of modern digital electronics and computing. The rapid expansion of computing capability is derived mainly from successive improvements in device miniaturization and the concomitant increase in device density and circuit complexity on a single chip. Functionality per chip has grown in accordance with Moore's law, an historical observation made by Intel executive Gordon Moore. *Moore's law* states that functionality as measured by the number of transistors and bits doubles every 1.5 to 2 years. As can be seen from Figure 1.1.1, the number of transistors on a silicon chip has followed an exponential dependence since the late 1960s. This in turn has led to dramatic improvements in computing capability, leading the consumer to expect ever better products at reduced cost.

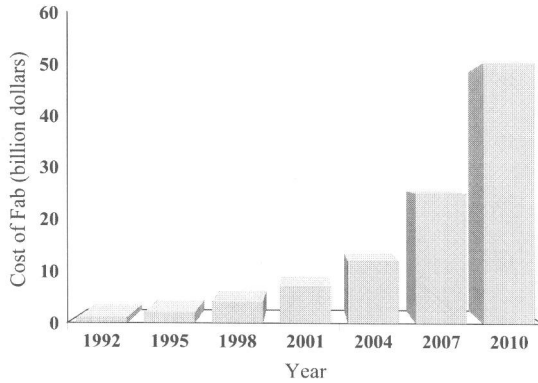


**FIGURE 1.1.1** Number of transistors per chip as a function of year, including the names of the processors. The dashed line shows projections to 2010. The exponential growth reflected by this graph is what is commonly referred to as Moore's law. (Data from Birnbaum and Williams, 2000.)

One of the questions that this book addresses is: Is there a limit to Moore's law? Can integrated-circuit complexity continue to grow exponentially into the twenty-first century, or are their insurmountable technical or economic challenges that will derail this progress?

The two prominent technical drivers of the semiconductor industry are dynamic random access memory (DRAM), and microprocessors. Historically, DRAM technology developed at a faster pace than microprocessor technology. However, from the late 1990s microprocessors have become at least an equal partner to DRAMs in driving semiconductor device refinement. In many instances, microprocessor units (MPUs), have become the major driver of semiconductor technology. There are different performance criteria for these two major product families. The major concerns for DRAMs are cost and memory capacity. The usual metric applied to DRAMs is *half-pitch*, which is defined as essentially the separation between adjacent memory cells on the chip. Consequently, minimization of the area of each memory cell to provide greater memory density is the primary development focus for DRAMs. Cost and performance also drive microprocessor development, but the key parameters in this case are gate length and the number of interconnect layers. The maintenance of Moore's law requires, then, aggressive reduction in gate length as well as in DRAM cell area.

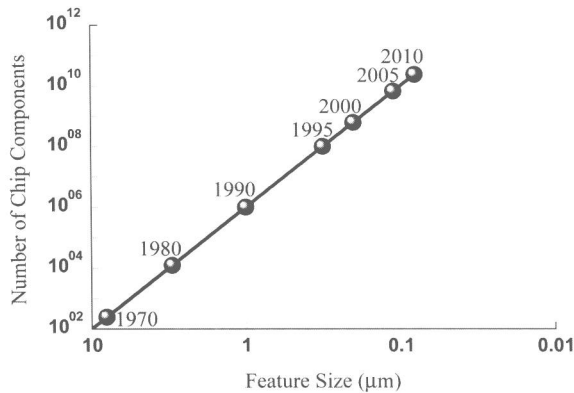
The semiconductor industry has tracked the technology trends of both DRAM and microprocessor technology and established technology road maps to project where the technology will be in subsequent years. Such road maps provide the industry with a rough guide to the technology trends expected. Below we discuss some of the implications of these trends and examine how they affect Moore's law.



**FIGURE 1.1.2** Fabrication plant cost as a function of year. Notice the tremendous cost projected by the year 2010. (Data from Birnbaum and Williams, 2000.)

Let us first consider economic challenges to the continuation of Moore's law. There is a second law attributed to Moore, Moore's second law, which examines the economic issues related to integrated-circuit chip manufacture. According to *Moore's second law*, the cost of fabrication facilities needed to manufacture each new generation of integrated circuits increases by a factor of 2 every three years. Extrapolating from fabrication plant costs of about \$1 billion in 1995, Figure 1.1.2 shows that the cost of a fabrication plant in 2010 could reach about \$50 billion. That a single company or even a consortium of companies could bear such an enormous cost is highly doubtful. Even if a worldwide consortium of semiconductor industries agreed to share such a cost, continuing to do so for future generations would certainly not be feasible. Therefore, it is likely that the economics of manufacturing will strongly influence the growth of the integrated-circuit industry in the near future.

Aside from the economic issues faced by continued miniaturization, several daunting technical challenges threaten continued progress in miniaturization of integrated circuits in accordance with Moore's law. As discussed in detail in Chapter 7, several technical issues threaten continued exponential growth of integrated-circuit complexity. Generally, these concerns can be classified as either physical or practical. By physical challenges we mean problems encountered in the physical operation of a device under continued miniaturization. Practical challenges arise from the actual fabrication and manufacture of these miniaturized devices. Among the practical challenges are lithography, gate oxide thickness reduction, and forming interconnects to each device. Physical operational problems encountered by continued miniaturization of devices include threshold voltage shifts, random fluctuation in the dopants, short-channel effects, and high-field effects. Any one or a combination of these effects could threaten continued progress of Moore's law. We examine these effects in detail in Chapter 7.



**FIGURE 1.1.3** Number of chip components as a function of device feature size. Both the historical trend and projected values according to the Semiconductor Institute of America (SIA) road map are shown. (Data from Birnbaum and Williams, 2000.)

Figure 1.1.3 shows how the number of chip components scales with the on-chip feature size. As can be seen, to maintain the historical trend in the near future, device feature sizes will need to be scaled significantly below  $0.1 \mu\text{m}$ . However, doing so requires overcoming many of the challenges listed above. At the time of this writing, strategies for overcoming the physical and practical challenges to further miniaturization are not known.

In addition to the practical and physical challenges to continued miniaturization, there are other issues that may thwart further progress. Formidable challenges arise in design, testing, and packaging integrated circuits containing billions of transistors. For example, testing a chip containing billions of transistors that operates at gigabit speeds is presently not possible. It is not sufficient that the chip contain billions of devices and operates at gigabit speeds. Just as important is the ability to extract signals from the transistors within the chip at gigahertz frequencies. Therefore, new packaging schemes will need to be developed to ensure that chip output progresses along with the chip itself.

Although there are many practical and physical challenges that may thwart further reduction of integrated circuits, progress may still occur through either evolutionary or revolutionary advancements. By *evolutionary*, we mean continued progress in device reduction through progressive refinements in the main integrated-circuit technology itself, complementary metal-oxide semiconductor (CMOS), technology. In Chapter 7 we examine several evolutionary technologies that could potentially extend CMOS in accordance with Moore's law. Alternatively, *revolutionary* technologies that go well beyond conventional CMOS may be required to continue progress in miniaturization. In Chapter 8 we examine several leading candidate technologies that may form the basis of computing hardware in the future. However, before anyone begins to invest and develop these alternative technologies aggressively, it is first nec-



essary to examine just how much better computing hardware can be made to be. It would be very unwise to pursue an expensive revolutionary technology aggressively if only marginal performance improvements can be made.

To this end, it is important to ask: What are the fundamental limits to computation, if any? Can these limits be quantified? Knowledge of the ultimate limits that the laws of physics place on computation, coupled with where CMOS technology is at present, clearly enable us to assess what more can be done and whether such advancements warrant development. Rolf Landauer and Richard Feynman pondered these questions in the late 1950s. They considered independently what the thermodynamic limits are to computing after recognizing that information could be treated as a physical entity and could thus be quantified.

Work by Lloyd (2000) has examined in detail the physical limits to computation. To that end, it is useful to determine the maximum speed of a logical operation. To perform an elementary logic operation in a certain time, a minimum amount of energy is required. The minimum energy required to perform a logic operation in time  $\Delta t$  is given by Lloyd (2000) as

$$E \geq \frac{\pi \hbar}{2 \Delta t} \quad 1.1.1$$

Therefore, if a computer has energy  $E$  available to it for computation, it can perform a maximum number of logical operations  $N$  per second given as

$$N = \frac{2E}{\pi \hbar} \quad 1.1.2$$

It is important to recognize that the foregoing limit applies to both serial and parallel processing. The computational limit is independent of the computer architecture. The total number of computations per second is the same whether one chooses to spend all the available energy to do serial computation faster or spread the energy out to do many calculations at once but with each calculation taking more time.

How do present computers compare to this ultimate speed limit? If one had an ideal computer, one in which all its mass could be transformed into computational energy, the limits imposed by Eqs. 1.1.1 and 1.1.2 would lead to very high computational speeds. Lloyd (2000) has estimated that a 1-kg computer all of whose mass-energy could be used for computation could perform about  $10^{50}$  operations per second. A modern computer falls far short of this rate mainly because (1) very little of its mass-energy is used in performing logic operations, and (2) many electrons are used to encode 1 bit. In modern computers, anywhere from  $10^6$  to  $10^9$  electrons are used to encode a "1." In principle, only one electron is needed to encode a "1." As we will see in Chapter 8, single-electron transistors have been developed that require only one electron for charging and storage of a bit.