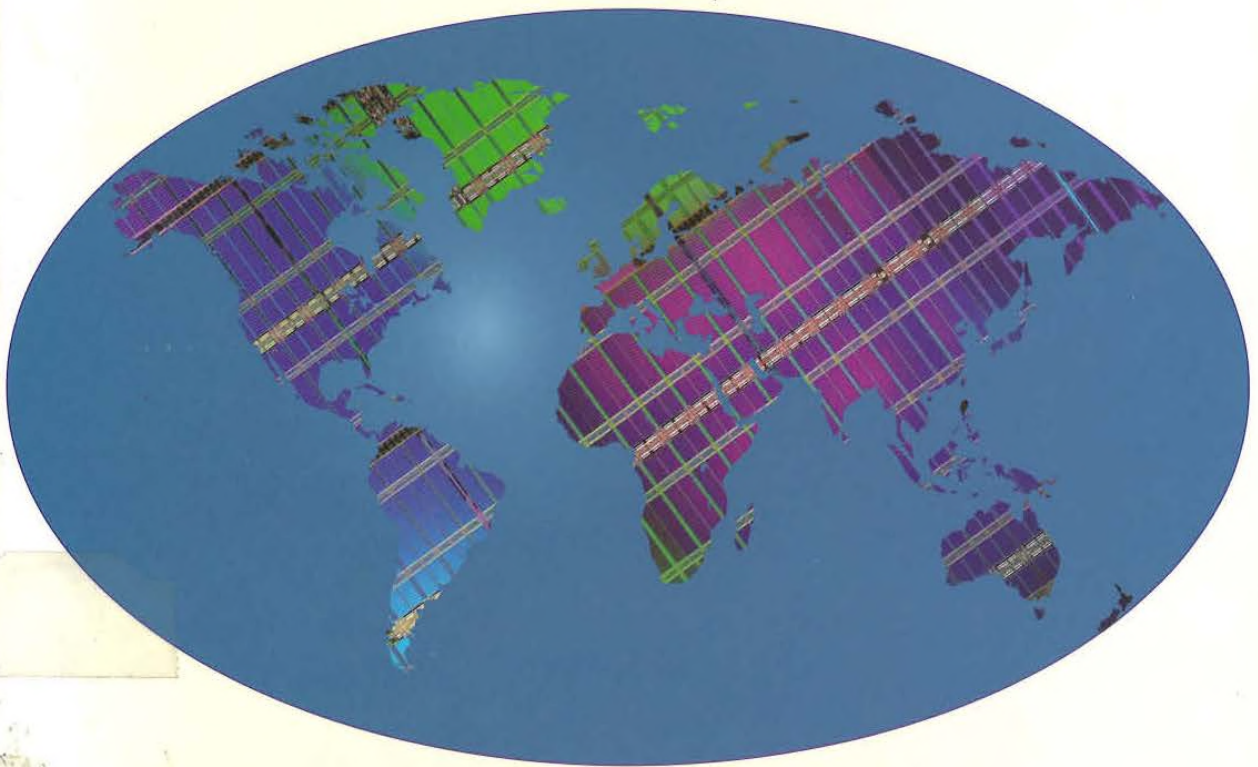


Microelectronics

AN INTEGRATED APPROACH



ROGER T. HOWE

CHARLES G. SODINI

Prentice Hall Electronics and VLSI Series—Charles G. Sodini, Series Editor

015276247
EN6I
REPL

Library of Congress Cataloging-in-Publication Data

Howe, Roger Thomas.

Microelectronics : an integrated approach / Roger T. Howe,

Charles G. Sodini.

p. cm.

Includes bibliographical references and index.

ISBN 0-13-588518-3

1. Microelectronics. 2. Solid state electronics. I. Sodini,
Charles G. II. Title.

TK7874.H688 1997

621.3815--dc20

96-17867

CIP

Acquisitions editor: Eric Svendsen

Production service: TKM Productions

Editorial/production supervision: Sharyn Vitrano

Copy editor: TKM Productions

Interior design: Lisa Delgado

Cover design: Joseph Sengotta

Art director: Amy Rosen

Managing editor: Bayani Mendoza DeLeon

Editor-in-chief: Marcia Horton

Director of production and manufacturing: David W. Riccardi

Manufacturing buyer: Julia Meehan

Editorial assistant: Kathryn Cassino



© 1997 by Prentice-Hall, Inc.

Simon & Schuster / A Viacom Company

Upper Saddle River, New Jersey 07458

All rights reserved. No part of this book may
be reproduced, in any form or by any means,
without permission in writing from the publisher.

The author and publisher of this book have used their best efforts in preparing this book. These efforts include the development, research, and testing of the theories and programs to determine their effectiveness. The author and publisher make no warranty of any kind, expressed or implied, with regard to these programs or the documentation contained in this book. The author and publisher shall not be liable in any event for incidental or consequential damages in connection with, or arising out of, the furnishing, performance, or use of these programs.

Printed in the United States of America

10 9 8 7 6 5

ISBN 0-13-588518-3

Prentice-Hall International (UK) Limited, London

Prentice-Hall of Australia Pty. Limited, Sydney

Prentice-Hall Canada Inc., Toronto

Prentice-Hall Hispanoamericana, S.A., Mexico

Prentice-Hall of India Private Limited, New Delhi

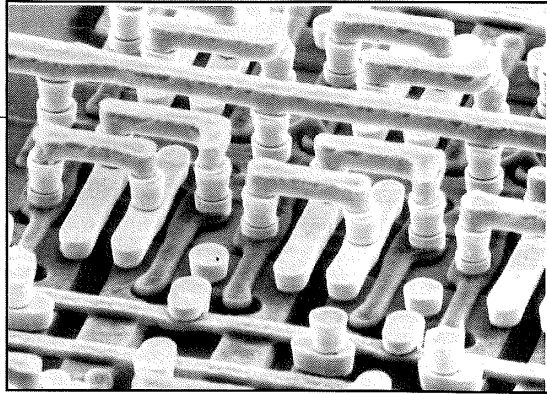
Prentice-Hall of Japan, Inc., Tokyo

Simon & Schuster Asia Pte. Ltd., Singapore

Editora Prentice-Hall do Brasil, Ltda., Rio de Janeiro

chapter 2

Semiconductor Physics and IC Technology



This scanning electron micrograph (SEM) photograph shows the complexity of modern integrated circuit interconnect technology. The inter-level dielectrics have been etched away to show the metallization patterns. (Courtesy of International Business Machines Corporation. Unauthorized use not permitted.)

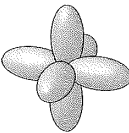
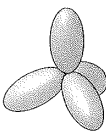
Microelectronic engineers need a working knowledge of the physical processes underlying the behavior of electronic devices. This chapter provides a short introduction to semiconductor physics, with the goal of developing the basic concepts and terminology needed to model electronic devices described in Chapters 3, 4, 6, and 7. In addition a basic familiarity with IC fabrication is essential to understand the device structures in later chapters. This chapter includes a short overview of the essential features of this powerful technology.

Chapter Objectives

- ◆ The bond model for silicon—the electron and hole as mobile charges.
- ◆ The control of electron and hole concentrations through doping.
- ◆ Drift and diffusion of electrons and holes.
- ◆ IC processes and the connection between mask layout and device cross section.
- ◆ Sheet resistance and the integrated-circuit resistor.

► 2.1 PURE SEMICONDUCTORS

To begin our discussion of semiconductor physics, we consider the dominant semiconductor used in modern microelectronics—crystalline silicon. Silicon is in Group IV of the periodic table, with atomic number 14. Recalling basic atomic chemistry, its 14 electrons occupy the first three energy levels. The first three orbitals, $1s$, $2s$, and $2p$, are filled by 10 electrons, leaving four electrons to occupy the $3s$ and $3p$ orbitals. In order to minimize the overall

(1s)² (filled) (2s)² (filled) (2p)⁶ (filled) (3s)² 

4 tetrahedral hybrid orbitals, half-filled with 4 valence electrons

electronic orbitals
in silicon

energy, the 3s and 3p orbitals hybridize to form four, tetrahedral 3sp orbitals. The electronic configuration of silicon is

$$(1s)^2 (2s)^2 (2p)^6 (3s)^2 (3p)^2$$

Each of the four outer hybrid orbitals have one electron and are capable of forming a bond with a neighboring atom. These four electrons that participate in bonding are called **valence electrons**. The inset on the left shows the electronic configuration of silicon.

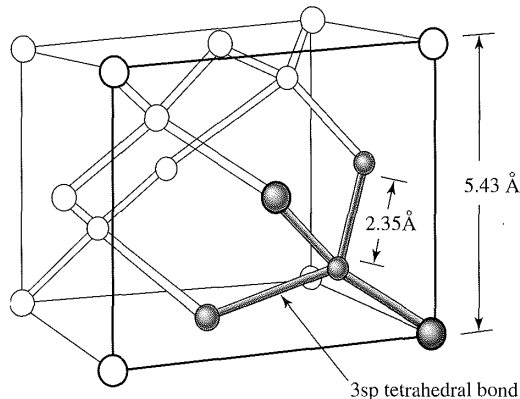
The crystalline form of silicon has the diamond crystal structure. Each silicon ion in the crystal has four nearest neighbors that are aligned with the tetrahedral 3sp hybrid orbitals. Each ion contributes an electron to fill the hybrid orbital. Figure 2.1 shows the unit cell of silicon, with the tetrahedral bonds from one atom highlighted. The edge of the unit cell is $a_o = 5.43 \text{ \AA}$.

How many atoms should be assigned to the unit cell? The eight corner atoms are shared among the eight cells touching the corner and therefore, contribute one atom to the unit cell. The six atoms on the faces of the cube are shared with adjacent cells and thus, contribute three atoms. Finally, there are four atoms that lie entirely within the cube, for a total of eight atoms in the unit cell.

The atomic density of crystalline silicon, in atoms per cubic centimeter, is

$$N_{Si} = 8 \cdot n_{cell} = 8 \cdot \frac{1}{a_o^3} = \frac{8}{(5.43 \times 10^{-8} \text{ cm})^3} = 5.00 \times 10^{22} \text{ cm}^{-3} \quad (2.1)$$

where n_{cell} is the number of unit cells per cm^3 . Since each atom has four tetrahedral bonds and contributes one electron to each bond, the concentration of valence electrons is $4 N_{Si} = 2 \times 10^{23} \text{ cm}^{-3}$.



► **Figure 2.1** Unit cell of crystalline silicon (After R. S. Muller and T.I. Kamins, *Device Electronics for Integrated Circuits*, 2nd ed., Wiley, 1986.)

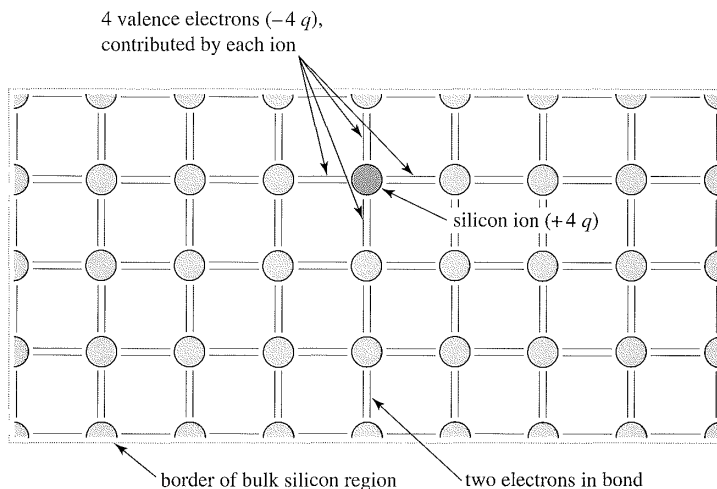
2.1.1 Electrons, Holes, and the Bond Model

We now investigate the electronic properties of the bonded network of silicon atoms in Fig. 2.1, which we refer to as the silicon **lattice**. In order to sketch the silicon lattice conveniently, we use the two-dimensional collapsed view shown in Fig. 2.2. The portion of the lattice shown in Fig. 2.2 is in the interior, or **bulk**, of the silicon crystal away from any surfaces. Each valence electron in a bond is symbolized by a connecting line, which means that two lines are used to represent each tetrahedral bond between silicon atoms. Note that this convention differs from the one used in chemistry, where a single line represents both electrons (in a single bond).

We refer to this abstract picture of the silicon crystal as the **bond model**. The silicon ions (nucleus and 10 inner electrons) each have a charge of $(+4q)$, where $q = 1.6 \times 10^{-19} \text{ C}$ is the magnitude of the charge of an electron. The four valence electrons per ion in the surrounding four bonds contribute a charge of $(-4q)$, making the crystal electrically neutral. In Fig. 2.2, all of the bonds are shown to be complete, which is true only for a perfect crystal at a temperature of absolute zero (0 degrees Kelvin).

An **intrinsic semiconductor** is one in which there are no impurities. A perfect crystal of intrinsic silicon at absolute zero has all of its valence electrons tied up in bonds since this is the lowest energy configuration. Since there are no “loose” or “free” electrons to carry current, intrinsic silicon is an insulator at absolute zero.

At temperatures above absolute zero, lattice vibrations occur that can impart enough energy to the valence electrons to break a very small percentage of the bonds. Each broken bond results in a mobile electron that can move around the crystal. This process is known as **thermal generation**. Note that the electron is still confined to the



► **Figure 2.2** Collapsed two-dimensional silicon lattice at absolute zero, with all bonds complete. This portion of the crystal is in the interior, or bulk, away from surfaces. It is electrically neutral since the four valence electrons balance the net charge on the silicon ion.

crystal since much more energy is required for it to escape into the surrounding ambient.

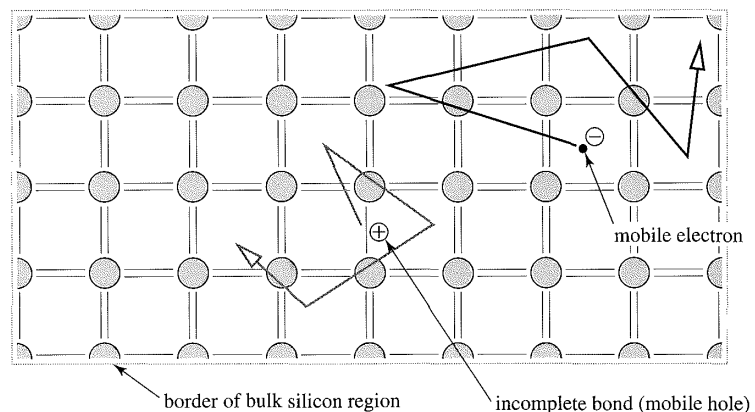
Figure 2.3 shows the silicon lattice after one bond has been broken due to thermal generation. The mobile electron has a negative charge $-q = -1.6 \times 10^{-19}$ C and leaves behind an incomplete bond that takes on a charge of $+q$. At room temperature, the electron moves quite rapidly and undergoes frequent collisions as suggested by the trajectory sketched in Fig. 2.3.

Since electrons have particle-like behavior in a vacuum, it seems plausible that the mobile electron resulting from the broken bond is capable of moving through the crystal lattice. For example, a cathode-ray tube used in a computer monitor works by steering electron trajectories with electric fields. What is surprising is that the positive charge arising from the broken bond is also mobile with similar properties to the electron, except that it has an equal and opposite charge. Figure 2.3 depicts that the positive charge is also free to move around the lattice. This mobile positive charge is called a **hole**.

2.1.2 Limitations of the Bond Model

The bond model is useful in visualizing the creation of electron-hole pairs. However, it has some shortcomings. Electrons and holes in crystals have both particle-like and wave-like behavior that can only be understood from the perspective of quantum mechanics. The bond model necessarily depicts both the electron and the hole as localized in space, which is not correct.

Fortunately, it is not necessary to take a solid-state physics course prior to studying semiconductor devices and circuits. This text considers that electrons and holes can effectively be considered as *classical* particles that obey Newton's Laws and have unit negative and positive charge. Quantum mechanics has a negligible effect on the



► **Figure 2.3** Electron and hole resulting from breaking one bond in intrinsic silicon.

accuracy of models for microelectronic devices that have dimensions of 0.1 μm or larger.¹

2.2 GENERATION, RECOMBINATION, AND THERMAL EQUILIBRIUM

We have seen that mobile electrons and holes can be created in pure, also known as **intrinsic**, silicon by the breaking of bonds due to thermal generation. The thermal generation rate $G_{th}(T)$, with units of $\text{cm}^{-3}\text{s}^{-1}$, increases exponentially with temperature. Another type of generation process is **optical generation**, with rate G_{op} . Optical generation occurs when a bond absorbs a photon with sufficient energy to break it. The minimum energy for bond-breaking in silicon is around 1.1 eV, corresponding to a wavelength of around 1 μm (in the near-infrared region of the optical spectrum.) In contrast to thermal generation, the optical generation rate is essentially independent of temperature. The total generation rate G is the sum of the thermal and optical rates

$$G = G_{th}(T) + G_{op} \quad (2.2)$$

If mobile electrons and holes are continuously generated, a balancing process is necessary that continuously removes them. **Recombination** is the elimination of an electron and hole from the crystal by completing one of the broken bonds. Recombination processes are classified according to what happens to the energy released when the bond forms.

Thermal recombination—the energy goes to vibrations in the crystal lattice. This is equivalent to saying the silicon crystal is heated.

Optical recombination—the energy goes to photon emission. Optical recombination is of negligible importance in silicon. In other semiconductors such as gallium arsenide, optical recombination is a very significant component of the total recombination rate. In these materials, light emission due to optical recombination is the basic principle behind light-emitting diodes and laser diodes.

For recombination to occur, both an electron and a hole are necessary. The recombination rate will be small if there are few electrons despite an abundance of holes or vice versa. Therefore, it is reasonable that the recombination rate is proportional to the *product* of the electron and hole concentrations

$$R = k(n \cdot p) \quad (2.3)$$

where n is the electron concentration (cm^{-3}), p is the hole concentration (cm^{-3}), and k is the recombination rate constant (cm^3s^{-1}). In contrast, the generation rate G is independent of the electron and hole concentrations since the number of broken bonds is a tiny fraction of the total number of bonds in the crystal.

If the environment of the silicon crystal (optical radiation flux, electric field intensity, temperature, etc.) is *constant*, then the electron and hole concentrations must eventually reach stable values that are set by the balance between the recomb-

steady-state

1. Current research is exploring **nanoelectronic** devices with critical dimensions on the order of 10 nm or less. Quantum effects are crucial to understanding their basic operation.

nation and generation rates. We define this situation as **steady-state**. The total recombination rate and total generation rates must be equal for steady-state conditions

$$R = k(n \cdot p) = G = G_{th}(T) + G_{op} \quad (2.4)$$

Given the particular environmental conditions that determine $G_{th}(T)$ and G_{op} , the product of the electron and hole concentrations is

$$n \cdot p = \frac{G_{th}(T) + G_{op}}{k} \quad (2.5)$$

thermal equilibrium

An idealized, special case of the steady-state condition called **thermal equilibrium** is of great importance in semiconductor physics. Thermal equilibrium is defined as the prolonged absence of external energy sources. The silicon sample is completely isolated from optical or electrical sources (e.g., power supplies and light sources) by placing it in a “black box” at a constant temperature. Some optical generation occurs in addition to thermal generation, due to black-body radiation at the sample temperature. For the case of thermal equilibrium, Eq. (2.5) reduces to

$$n_o \cdot p_o = \frac{G_{th}^o(T) + G_{op}^o(T)}{k_o} \quad (2.6)$$

where all quantities are subscripted or superscripted with “o” to denote thermal equilibrium. The right-hand side of this equation is a function of temperature only (the single variable in thermal equilibrium). Defining this function as $n_i^2(T)$, we can re-express Eq. (2.6) as the **mass-action law**

mass-action law

$$n_o \cdot p_o = n_i^2(T) \quad (2.7)$$

This important result relates the product of the thermal equilibrium electron and hole concentrations to a constant that depends only on temperature. The unusual name for this equation is borrowed from chemistry, since Eq. (2.7) is analogous to an equilibrium equation for a chemical reaction. For example, the equilibrium between hydrogen ions (concentration $[H^+]$) and hydroxyl ions (concentration $[OH^-]$) reacting to form water is expressed as

$$[H^+] \cdot [OH^-] = K_{eq}(T) \quad (2.8)$$

where $K_{eq}(T)$ is the equilibrium constant.

In pure silicon in thermal equilibrium, the electron and hole concentrations are equal since the electrons and holes are created in pairs when the bonds are broken. Therefore, it follows from substitution in Eq. (2.7)

$$n_o \cdot p_o = n_o^2 = p_o^2 = n_i^2(T) \quad (2.9)$$

and consequently, the electron and hole concentrations in pure silicon are both equal to the **intrinsic concentration** $n_i(T)$ in thermal equilibrium. Before proceeding fur-

ther, it is worthwhile to consider the order of magnitude of the electron and hole concentrations in pure silicon at room temperature

$$n_i \cong 10^{10} \text{ cm}^{-3} \text{ at } 300 \text{ K} \quad (2.10)$$

intrinsic concentration

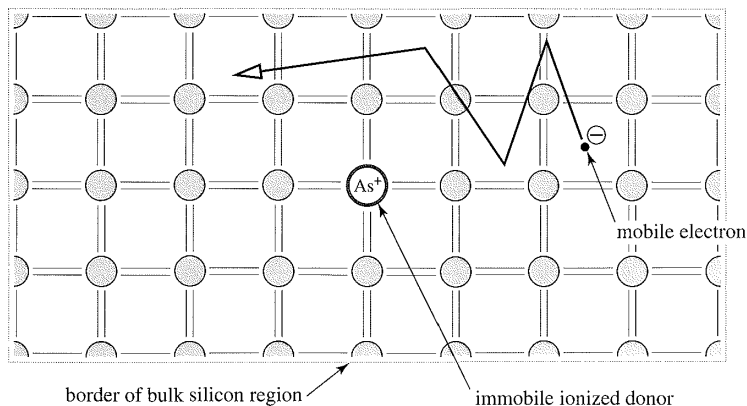
Given that the number of bonds is $4 \times N_{Si} = 2 \times 10^{23} \text{ cm}^{-3}$, at room temperature only an extremely small fraction (1 in 20 trillion) of the bonds are broken. As temperature increases, the intrinsic concentration increases exponentially—approximately doubling for every 10°C rise over room temperature.

► 2.3 DOPING

Silicon would be of little use if the hole and electron concentrations were both equal to 10^{10} cm^{-3} . Fortunately, the hole and electron concentrations can be controlled by adding minute quantities of specific elements to the silicon crystal. There are a variety of ways to incorporate these impurities into the silicon lattice using a process called **doping**. We will first consider the case of **donors**, which are elements in Group V of the periodic table.

2.3.1 Donors

The bond model is useful for visualizing how donor atoms control the concentration of electrons in silicon. Typical donors for silicon are phosphorus (P), arsenic (As), and antimony (Sb), all of which are in Group V of the periodic table. Figure 2.4 shows the bond model of silicon with the addition of an arsenic atom on a substitutional lattice site. Since arsenic has five valence electrons, it is able to form bonds with all four neighboring silicon atoms. The fifth valence electron is left after the bonds are



► **Figure 2.4** Bond model with arsenic incorporated into a substitutional site in a bulk region of the silicon crystal. Due to thermal excitation, arsenic loses (donates) its fifth valence electron as a mobile electron resulting in an immobile positive charge.

formed and is very loosely bound to the arsenic atom. For temperatures $T \geq 100$ K = -173° C, there is enough thermal energy for this fifth electron to leave and move through the lattice, leaving behind an immobile, positively charged arsenic ion.

The charges in the bulk silicon region in Fig. 2.4 are the key to quantifying the electron and hole concentrations for arsenic-doped silicon. The region was electrically neutral before the arsenic was incorporated into the lattice. After the arsenic is added and ionizes, which occurs for $T \geq 100$ K = -173° C, the region is still neutral since the negative electron charge is balanced by the positive charge on the arsenic ion. Adding up all the charge densities (units: C/cm³) with their proper signs, we require that the net charge density ρ is zero:

$$\rho = -qn_o + qp_o + qN_d = q(p_o + N_d - n_o) = 0 \quad (2.11)$$

The concentration of dopants is extremely low and their presence in the lattice does not significantly alter the mass-action law in thermal equilibrium. Solving Eq. (2.7) for the hole concentration and substituting into Eq. (2.11), we find that

$$\frac{n_i^2}{n_o} + N_d - n_o = 0 \quad (2.12)$$

By multiplying both sides by the electron concentration n_o and rearranging, the result is a quadratic equation

$$n_o^2 - N_d n_o - n_i^2 = 0 \quad (2.13)$$

Using the quadratic formula, we find that the electron concentration in silicon doped with donors is

$$n_o = \frac{N_d + \sqrt{N_d^2 + 4n_i^2}}{2} = \frac{N_d}{2} + \frac{N_d}{2} \sqrt{1 + \frac{4n_i^2}{N_d^2}} \quad (2.14)$$

where we have taken the positive root since the electron concentration must be positive. The hole concentration p_o can be found from the electron concentration using the mass-action law in Eq. (2.7).

In practice, the donor concentration is much greater than the intrinsic concentration $N_d \gg n_i$. The square root in Eq. (2.14) is approximately one and the electron concentration n_o is

n-type electron
concentration

$$n_o \cong \frac{N_d}{2} + \frac{N_d}{2} = N_d \quad (2.15)$$

Silicon doped with $N_d \gg n_i$ is called **extrinsic**. The electron concentration n_o in extrinsic silicon is determined by the donor concentration and is not a function of temperature-dependent thermal generation. The result in Eq. (2.15) can be easily remembered since it corresponds to “one electron per donor,” which can be seen from the bond model in Fig. 2.4. The relatively large donor concentration means the thermally generated electrons can be neglected, as quantified by Eq. (2.14).

The thermal equilibrium hole concentration in silicon doped with donors can be found from the mass-action law

$$p_o = \frac{n_i^2}{n_o} \cong \frac{n_i^2}{N_d} \quad p_o = \frac{10^{20}}{N_d} \text{ cm}^{-3} \quad (T = 300\text{K}) \quad (2.16)$$

n-type hole
concentration

For example, for a donor concentration of $N_d = 10^{14} \text{ cm}^{-3}$, the hole concentration at room temperature is only $p_o = 10^6 \text{ cm}^{-3}$ —a factor of 100 million lower than the electron concentration! By adding donor impurities to the silicon sample, the mass-action law in Eq. (2.7) depresses the hole concentration. Few holes remain in equilibrium since recombination with the large number of electrons removes most of the them. Silicon doped with donors is referred to as **n-type**, in which electrons are the **majority carriers** (by far the most numerous) and holes are the **minority carriers**.

The lower limit of useful donor concentration is set by the requirement that $N_d \gg n_i$ in Eq. (2.14), so $N_{d\min} \cong 10^{12} \text{ cm}^{-3}$. On the opposite extreme, there is a maximum amount of substitutional impurities that the silicon lattice can absorb without forming a new material phase with a different crystal structure and electronic properties. This maximum varies among the donor elements, but a typical value is around $N_{d\max} \cong 10^{19} \text{ cm}^{-3}$. This corresponds to about 1 in 5000 lattice sites occupied by a donor since the atomic density of silicon is $N_{Si} = 5 \times 10^{22} \text{ cm}^{-3}$.

► EXAMPLE 2.1 Electron and Hole Concentrations in n-type Silicon

Find the concentration of holes and electrons for a bulk silicon sample in thermal equilibrium at room temperature that has a donor concentration $N_d = 10^{14} \text{ cm}^{-3}$.

SOLUTION

Using Eq. (2.14), the electron concentration is

$$n_o = \frac{10^{14}}{2} + \frac{10^{14}}{2} \sqrt{1 + \frac{4 \cdot 10^{20}}{10^{28}}} = 1.00000001 \times 10^{14} \text{ cm}^{-3}$$

The hole concentration p_o can be found from the mass-action law in Eq. (2.7) since the sample is in thermal equilibrium

$$p_o = \frac{n_i^2}{n_o} = \frac{(10^{10})^2}{1.00000001 \times 10^{14}} = 9.9999999 \times 10^5 \text{ cm}^{-3}$$

Verify that the bulk silicon region is electrically neutral by adding the charge densities of the holes, ionized donors, and electrons

$$\rho = q(p_o + N_d - n_o) = q(9.9999999 \times 10^5 + 1 \times 10^{14} - 1.00000001 \times 10^{14}) = 0$$

Note that since N_d exceeds n_i by four orders of magnitude, we could have used Eq. (2.15) to find $n_o \cong N_d = 10^{14} \text{ cm}^{-3}$, and then used the mass-action law to find the hole concentration to be $p_o \cong 10^6 \text{ cm}^{-3}$.

2.3.2 Acceptors

A substitutional impurity from Group III of the periodic table that has only three valence electrons is called an **acceptor**. The usual acceptor atom for silicon is boron (B). Figure 2.5 is the bond model's picture of doping by acceptors.

Since the Group III boron atom only has three valence electrons, it can complete only three of the four bonds with its silicon neighbors. However, an electron in an adjacent bond requires only a small amount thermal energy to jump over and complete the fourth bond. This process of “accepting” an electron accounts for the name of this class of dopants. The resulting hole is free to move around the crystal, whereas the immobile acceptor takes on a negative charge due to the extra electron as shown in Fig. 2.5.

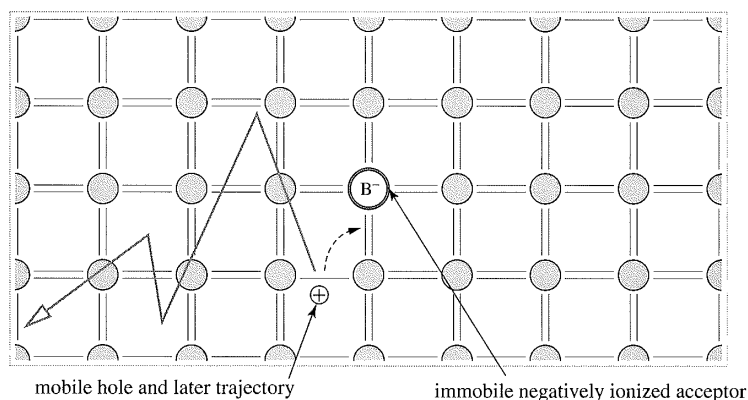
Given that the acceptor concentration is N_a (cm^{-3}), the equilibrium hole and electron concentrations can be found using the same approach used for donors. The bulk silicon region is electrically neutral and the hole, electron, and ionized acceptor charge densities must therefore sum to zero:

$$\rho = -qn_o + qp_o - qN_a = q(p_o - n_o - N_a) = 0 \quad (2.17)$$

The electron concentration can be eliminated by substituting $n_o = n_i^2 / p_o$ from the mass-action law in Eq. (2.7)

$$p_o - \frac{n_i^2}{p_o} - N_a = 0 \quad (2.18)$$

Using the quadratic formula to solve this equation yields an expression for the hole concentration in silicon doped with acceptors



► **Figure 2.5** Bond model showing boron on a substitutional lattice site in a bulk region of the silicon crystal. A hole is created when the incomplete fourth bond is filled by acceptance of an electron from a neighboring bond.

$$p_o = \frac{N_a}{2} + \frac{N_a}{2} \sqrt{1 + \frac{4n_i^2}{N_a^2}} \quad (2.19)$$

For the typical case where N_a is orders of magnitude larger than n_i , the intrinsic hole concentration, the square root term is approximately one and the hole concentration is

$$p_o \cong N_a \quad (2.20)$$

p-type hole
concentration

By applying the mass-action law, we find that the equilibrium electron concentration at room temperature is given by

$$n_o = \frac{n_i^2}{p_o} \cong \frac{n_i^2}{N_a} = \frac{10^{20}}{N_a} \text{ cm}^{-3} \text{ at } 300 \text{ K} \quad (2.21)$$

p-type electron
concentration

Silicon doped with acceptors is called **p-type**, in which holes are the majority carriers and electrons are the minority carriers. P-type silicon is **complementary** to n-type silicon, in which the relationship between holes and electrons is reversed.

Due to the mass-action law, doping with $N_a = 10^{15} \text{ cm}^{-3}$ of boron results in a hole concentration $p_o = 10^{15} \text{ cm}^{-3}$ and electron concentration $n_o = 10^5 \text{ cm}^{-3}$. As with donors, the minimum useful acceptor concentration to have extrinsic, p-type silicon is determined by the intrinsic concentration n_i . At room temperature $N_{a_{\min}} \approx 10^{12} \text{ cm}^{-3}$ for $N_a \gg n_i$. The maximum concentration of substitutional boron acceptors in silicon is on the order of $N_{a_{\max}} \approx 10^{19} \text{ cm}^{-3}$ since a boron-silicon compound begins to precipitate at higher concentrations.

2.3.3 Donors and Acceptors: Compensation

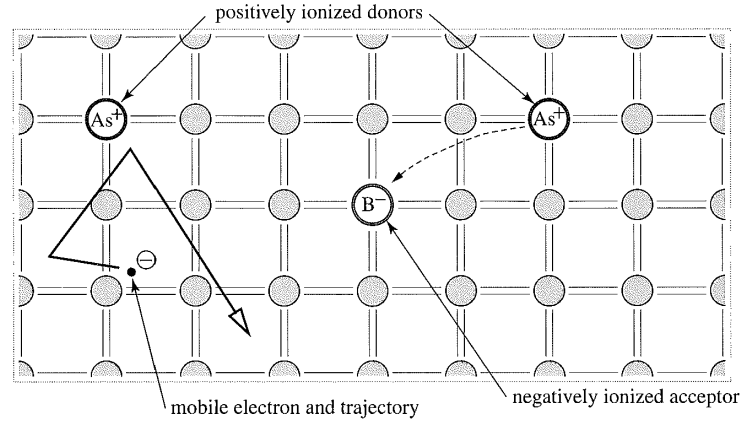
Silicon is often doped with both donors *and* acceptors. To find the equilibrium electron and hole concentrations, first consider the bond model in Fig. 2.6. Using the same reasoning used for the donor-only and acceptor-only cases, it can be seen that charge neutrality applies to the bulk silicon region in Fig. 2.6. The charge densities from electrons, holes, ionized donors, and ionized acceptors must therefore sum to zero

$$\rho = -qn_o + qp_o + qN_d - qN_a = q(p_o - n_o + N_d - N_a) = 0 \quad (2.22)$$

When $N_d > N_a$, we can substitute for the hole concentration in thermal equilibrium $p_o = n_i^2 / n_o$ from the mass-action law and find

$$n_o^2 - (N_d - N_a) n_o - n_i^2 = 0 \quad (2.23)$$

This quadratic equation is identical to Eq. (2.13) with the **net** donor concentration $N_d - N_a$ replacing the donor concentration. The solution can be found by substitution into Eq. (2.14)



► **Figure 2.6** Bond model showing a bulk section of silicon that is doped with both arsenic (donors) and boron (acceptors) with $N_d > N_a$. Donated electrons ionize the acceptors by completing their half-filled bonds, with the excess $N_d - N_a$ becoming mobile electrons. This process is called **compensation**.

$$n_o = \frac{(N_d - N_a)}{2} + \frac{(N_d - N_a)}{2} \sqrt{1 + \frac{4n_i^2}{(N_d - N_a)^2}} \quad (2.24)$$

where we assume that the donor concentration is greater than the acceptor concentration $N_d > N_a$. In practice, the excess of donors over acceptors $N_d - N_a$ will usually greatly exceed the intrinsic concentration since separately $N_d \gg n_i$ and $N_a \gg n_i$. For example, if the donor concentration exceeds the acceptor concentration by only 5% at a typical doping level of about 10^{15} cm^{-3} , then

$$N_d - N_a = 1.05 \times 10^{15} \text{ cm}^{-3} - 1.0 \times 10^{15} \text{ cm}^{-3} = 5 \times 10^{13} \text{ cm}^{-3} \gg n_i \quad (2.25)$$

Assuming $N_d - N_a \gg n_i$, we find that

compensated n-type
electron concentration

$$n_o \cong N_d - N_a \quad (N_d > N_a) \quad (2.26)$$

The physical interpretation of Eq. (2.26) is that electrons donated by the donors first ionize all the acceptors, with only the net donor concentration in excess of the acceptor concentration contributing mobile electrons to the crystal. This process is called **compensation**.

Since the sample is in thermal equilibrium, the minority hole concentration can be found by the mass-action law

compensated n-type
hole concentration

$$p_o \cong \frac{n_i^2}{N_d - N_a} \quad (N_d > N_a) \quad (2.27)$$

At first glance, the fact that there are negligible holes in a sample doped with acceptors seems counter-intuitive. However, the mass-action law depresses the hole concentration since the (majority) electron concentration far exceeds n_i .

For the case where the acceptor concentration is greater than the donor concentration, similar reasoning shows that the sample is p-type with

$$p_o \cong N_a - N_d \quad n_o \cong \frac{n_i^2}{N_a - N_d} \quad (N_a > N_d) \quad (2.28)$$

compensated p-type
carrier concentration

In summary, a silicon crystal that is doped with one type of impurities may be compensated by doping with an excess of impurities of the opposite type. In section 2.5, we will see that compensation is a cornerstone of integrated device fabrication.

► EXAMPLE 2.2 Electron and Hole Concentrations for Compensated Silicon

a) Find the thermal equilibrium electron and hole concentrations at room temperature for a bulk silicon region doped with donors at $N_d = 10^{14} \text{ cm}^{-3}$ and with acceptors at $N_a = 5 \times 10^{13} \text{ cm}^{-3}$.

SOLUTION

The sample is n-type since the donor concentration exceeds the acceptor concentration. Since the net donor concentration $N_d - N_a = 10^{14} \text{ cm}^{-3} - 5 \times 10^{13} \text{ cm}^{-3} = 5 \times 10^{13} \text{ cm}^{-3} \gg n_i$, we can equate the electron concentration to the net donor concentration (Eq. 2.26)

$$n_o \cong N_d - N_a = 5 \times 10^{13} \text{ cm}^{-3}$$

The mass-action law in Eq. (2.7) determines that the minority hole concentration is

$$p_o \cong \frac{n_i^2}{N_d - N_a} = 2 \times 10^6 \text{ cm}^{-3}$$

To verify that the region is electrically neutral, we add the contributions of the positive holes and ionized donors with the negative electrons and negative ionized acceptors

$$\rho = q(p_o + N_d - n_o - N_a) = q(2 \times 10^6 + 10^{14} - 5 \times 10^{13} - 5 \times 10^{13}) \approx 0$$

Note: The charge density would have been exactly zero if the exact expression for the compensated electron concentration in Eq. (2.24) had been used.

(b) Now consider a bulk n-type silicon sample with $N_d = 10^{14} \text{ cm}^{-3}$. We would like to add acceptors to make the hole concentration $p_o = 5 \times 10^{13} \text{ cm}^{-3}$. What acceptor doping concentration N_a is needed? What is the electron concentration?

SOLUTION

The desired hole concentration is much greater than the intrinsic concentration; therefore, the compensated sample is p-type. Using Eq. (2.28), we can solve for the acceptor concentration

$$p_o \equiv N_a - N_d \quad \rightarrow \quad N_a \equiv p_o + N_d = 5 \times 10^{13} \text{ cm}^{-3} + 10^{14} \text{ cm}^{-3} = 1.5 \times 10^{14} \text{ cm}^{-3}$$

The electron concentration follows from the mass-action law

$$n_o = \frac{n_i^2}{p_o} = \frac{10^{20}}{5 \times 10^{13}} = 2 \times 10^6 \text{ cm}^{-3}$$

► 2.4 CARRIER TRANSPORT

At room temperature, the holes and electrons in silicon are in agitated, random motion. Their average velocity is called the **thermal velocity** v_{th} , and is about 10^7 cm/s—only a factor of 3000 less than the speed of light! Holes and electrons collide frequently by scattering off ionized impurities, crystal defects, surfaces, and each other. A typical value for the average **collision time** τ_c at room temperature in bulk silicon is about 1×10^{-13} s = 0.1 picoseconds. In between collisions, carriers on average will travel a distance

$$\lambda = v_{th} \tau_c = 10 \text{ nm} = 0.01 \text{ } \mu\text{m} \quad (2.29)$$

where λ is defined as the **mean free path**. State-of-the-art silicon microelectronic devices have dimensions that are at least $0.1 \text{ } \mu\text{m}$, but are typically much larger. Therefore, holes and electrons undergo many collisions while traversing a device structure. In the next section, this important fact will allow us to find simple equations that describe the average transport of charge carriers due to electric fields or to concentration gradients.

2.4.1 Drift Velocity

In thermal equilibrium, the electrons and holes in a region of the silicon lattice have no net motion. In other words, the average position of all the electrons (or separately, all of the holes) does not change with time. Figure 2.7(a) illustrates this situation for electrons. The random trajectories of three electrons along the x axis are shown in Fig. 2.7(a) for a time interval $\Delta t = 7 \tau_c$, during which seven collisions occur. To show clearly the electron motion, the x -axis trajectories are spread out and all trajectories are shown originating at $x = x_i$. The average change in position after the interval Δt is

$$\overline{\Delta x} = \frac{(x_{f,1} - x_i) + (x_{f,2} - x_i) + (x_{f,3} - x_i)}{3} \approx 0 \quad (2.30)$$

using the approximate values for the position changes in Fig. 2.7(a). Of course, the average position change for electrons in a region of the silicon lattice will be taken

over a huge number of electrons. Since the trajectories are random, $\overline{\Delta x}$ will be zero after any time interval $\Delta t \gg \tau_c$.

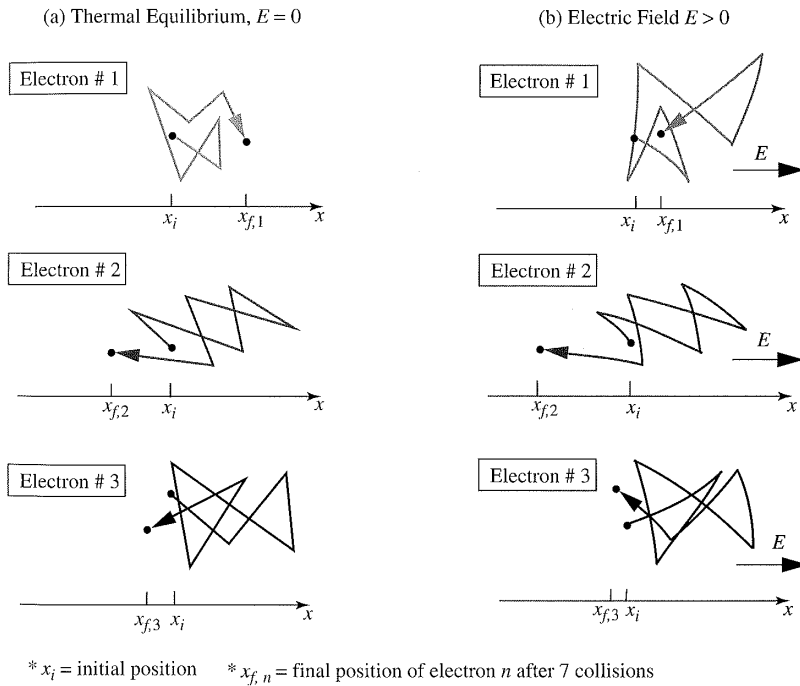
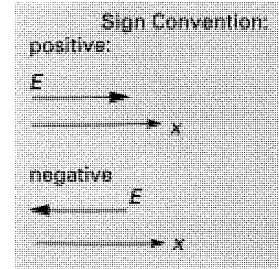
The situation with an applied electric field E is different, as shown in Fig. 2.7(b). It is important to first be sure of the sign for the electric field and other directional variables. By convention, the electric field is positive when it points in the same direction as the reference direction for the x axis. Thus, the electric field E in Fig. 2.7(b) is positive. The electrons are subject to an electrostatic force

$$F_{el} = -qE \quad (2.31)$$

that points in the $-x$ direction. During their “free flight” in between collisions, the electrons with mass m_n experience a constant acceleration $a = F_{el}/m_n = -qE/m_n$ that deflects their trajectories toward the $-x$ direction as shown in Fig. 2.7(b).

Averaging the position changes over the interval Δt , we find

$$\overline{\Delta x} = \frac{(x_{f,1} - x_i) + (x_{f,2} - x_i) + (x_{f,3} - x_i)}{3} < 0 \quad (2.32)$$



► **Figure 2.7** (a) Trajectories for three electrons in thermal equilibrium (zero applied electric field) for seven collisions. (b) Trajectories for the same three electrons for an electric field $E > 0$. All of the trajectories are drawn with the same initial position x_i .

using the three sample trajectories in Fig. 2.7(b). Dividing the average position change by the time interval, we define the **drift velocity** of the electrons as

$$v_{dn} = \frac{\overline{\Delta x}}{\Delta t} \quad (2.33)$$

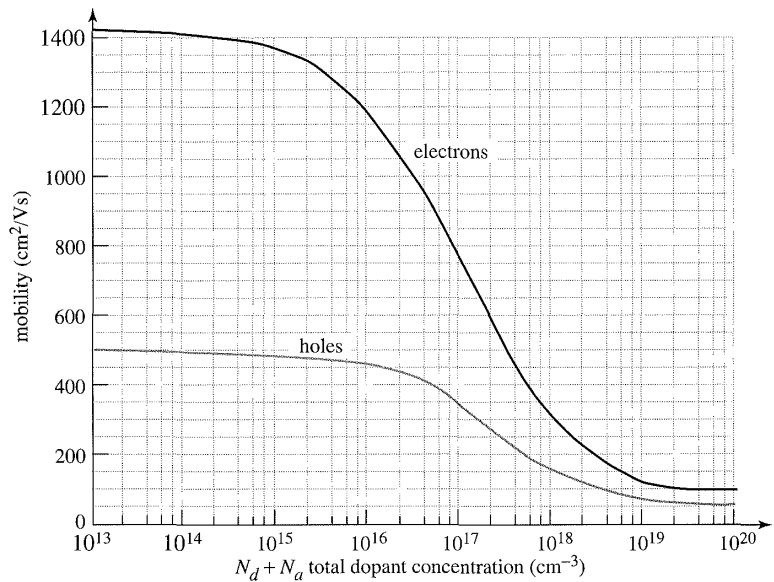
which is negative for $E > 0$. Note that the drift velocity is the average net velocity for a huge number of electrons in a region of the silicon lattice. Experimentally, the drift velocity v_{dn} is found to be proportional to the electric field

electron drift velocity

$$v_{dn} = -\mu_n E \quad (2.34)$$

where μ_n is the **electron mobility** with units of $\text{cm}^2/(\text{Vs})$.

The electron mobility is a function of both the temperature and total doping concentration. Mobility decreases with increasing temperature since thermal agitation causes lattice vibrations to scatter electrons. Ionized dopants also scatter electrons and, therefore, reduce the collision time. This means the mobility is a decreasing function of $N_d + N_a$. Figure 2.8 gives experimental measurements of electron mobility for $N_d + N_a = 10^{13}$ to 10^{20} cm^{-3} at room temperature. It is useful to have a rough number for electron mobility in silicon for approximate calculations. Substituting a typical total dopant concentration of $2.5 \times 10^{16} \text{ cm}^{-3}$, Fig. 2.8 gives an electron mobility of $\mu_n \approx 1000 \text{ cm}^2/(\text{Vs})$.



► **Figure 2.8** Linear-log plot of electron and hole mobilities at room temperature, as functions of the total doping concentration $N_d + N_a$. After R. S. Muller and T. I. Kamins, *Device Electronics for Integrated Circuits*, 2nd ed., Wiley, 1986.

Holes also drift in an electric field—the holes drifting *with* the field on average, because the positive charge causes them to accelerate in the same direction as E between collisions. The drift velocity for holes is also proportional to the electric field

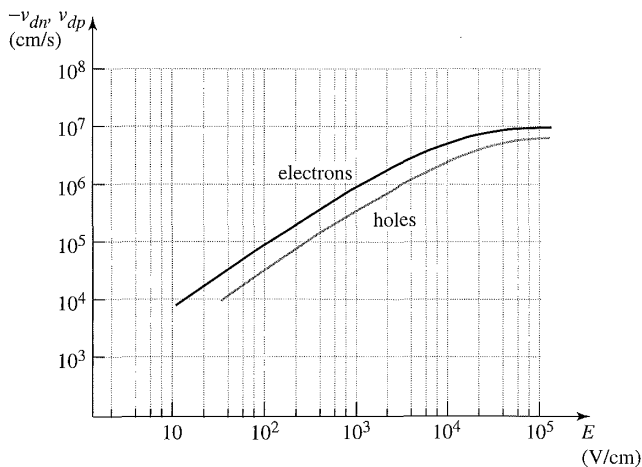
$$v_{dp} = \mu_p E \quad (2.35)$$

hole drift velocity

where μ_p is the **hole mobility**. In silicon, hole mobilities are lower than electron mobilities. They range from 100–500 $\text{cm}^2/(\text{Vs})$ at room temperature as shown in Fig. 2.8. For a typical dopant concentration of $2.5 \times 10^{16} \text{ cm}^{-3}$, the “round number” for hole mobility is $\mu_p \approx 400 \text{ cm}^2/(\text{Vs})$.

The linear relationship between drift velocity and electric field breaks down when the electric field is high enough that the drift velocity approaches $v_{sat} = 10^7 \text{ cm/s}$. Figure 2.9 shows $|v_d|$ as a function of electric field for electrons and holes and illustrates that both carriers have a maximum drift velocity of about 10^7 cm/s . This phenomenon is known as **velocity saturation**.

Advances in integrated-circuit technology are enabling ever-smaller electronic devices, a trend that is increasing the internal electric fields. Velocity saturation is an important phenomenon for accurate modeling of VLSI device characteristics. In this book, we will assume that the linear relationships between electric field and drift velocity in Eqs. (2.34) and (2.35) are valid unless velocity saturation is specifically mentioned.



► **Figure 2.9** Log-log plot of the drift velocity as a function of electric field for electrons and holes, showing the linear region where the velocity is proportional to the field and velocity saturation at high fields. After R. S. Muller and T. I. Kamins, *Device Electronics for Integrated Circuits*, 2nd ed., Wiley, 1986.

2.4.2 Drift Current Density

Drift of charged particles results in the transport of charge through the silicon lattice. We quantify the transport of charge by the **current density** J , which is defined as the rate at which charge passes through a reference plane in the $+x$ direction per unit area. The units for current density are $C/(s \text{ cm}^2)$ or A/cm^2 . The sign convention for current density is illustrated in Fig. 2.10. J is positive when positive charge passes in the $+x$ direction. Since the charge carriers can be positive (holes) or negative (electrons), there are two ways that J can be positive.

1. Holes can pass through the reference plane in the $+x$ direction.
2. Electrons can pass through the reference plane in the $-x$ direction.

Note that either case results in a net transport of positive charge through the reference plane in the $+x$ direction making the current density J positive.

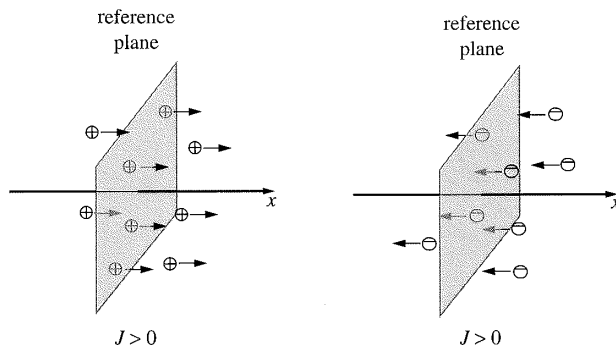
The hole current density due to drift in an electric field is expressed as J_p^{dr} —subscript p designates holes and superscript dr , drift. This can be expressed in terms of the electric field by directly applying the definition of current density, as shown in Fig. 2.11(a).

Over a time interval Δt , holes in the volume $\Delta V = A(v_{dp} \Delta t)$ will drift across the reference plane in the $+x$ direction. The number of holes ΔN_p that cross the plane in the interval Δt is given by the product of the concentration p of holes and the volume ΔV :

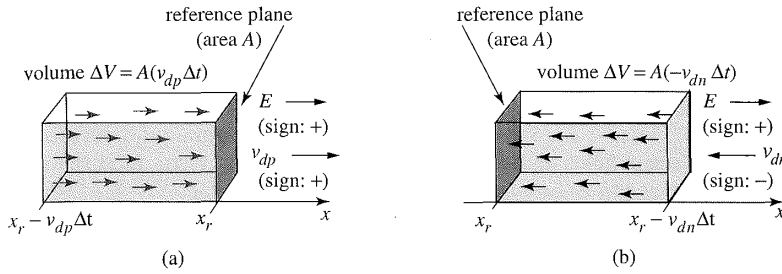
$$\Delta N_p = p \Delta V \quad (2.36)$$

The hole charge drifting across the reference plane ΔQ_p is the product of ΔN_p and the charge per hole q . According to the definition of current density

$$J_p^{dr} = \frac{\Delta Q_p}{A \Delta t} = \frac{qp \Delta V}{A \Delta t} = \frac{qp A v_{dp} \Delta t}{A \Delta t} = qp v_{dp} \quad (2.37)$$



► **Figure 2.10** Sign convention for current density J . Both cases represent positive values of J since there is a net transport of positive charge in the $+x$ direction.



► **Figure 2.11** (a) The definition of current density applied to drifting holes, and (b) to drifting electrons. The reference plane is located at $x = x_r$ for this example. For an electric field pointed in the $+x$ direction, holes drift through the reference plane from left to right, whereas electrons drift from left to right. The current density $J^{dr} > 0$ for both cases.

Substituting for the drift velocity from Eq. (2.35), we find that

$$J_p^{dr} = qp\mu_p E \text{ for } (v_{dp} \ll v_{sat}) \quad (2.38)$$

For electrons, the drift velocity is opposite in direction to the electric field as shown in Fig. 2.11(b). On average, electrons in a volume $\Delta V = A(-v_{dn} \Delta t)$ drift across the reference plane in the time interval Δt . The number of electrons crossing the plane is therefore $\Delta N_n = n\Delta V$, and the electron charge is equal to

$$\Delta Q_n = (-q) \Delta N_n = -qnA(-v_{dn} \Delta t) \quad (2.39)$$

The electrons drift across the reference plane in the $-x$ direction as shown in Fig. 2.11(b). Therefore, the drift current density is given by

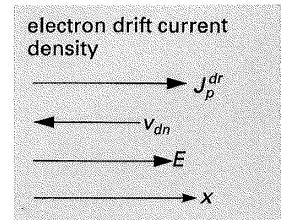
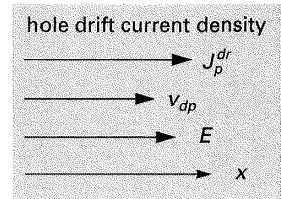
$$J_n^{dr} = -\left[\frac{\Delta Q_n}{A\Delta t}\right] \quad (2.40)$$

Substituting the electron charge from Eq. (2.39) and the drift velocity from Eq. (2.34), the electron drift current density is given by

$$J_n^{dr} = -\left[\frac{(qnAv_{dn}\Delta t)}{A\Delta t}\right] = -qn v_{dn} = -qn(-\mu_n E) = qn\mu_n E \quad (2.41)$$

Therefore, the electron drift current points in the same direction as the electric field. Restating Eq. (2.41)

$$J_n^{dr} = qn\mu_n E, \text{ for } |v_{dn}| \ll v_{sat}, \quad (2.42)$$



► EXAMPLE 2.3 Drift Current Density

Figure Ex2.3A shows a slab of p-type silicon ($N_a = 10^{16} \text{ cm}^{-3}$) under the influence of a uniform electric field E . Plot the drift current density as a function of E over the range $-5 \times 10^5 \text{ V/cm} < E < 5 \times 10^5 \text{ V/cm}$. Note: Velocity saturation occurs for holes drifting in electric fields with magnitudes in excess of 10^5 V/cm . Find the voltage across the slab at which the current reaches its maximum value.

SOLUTION

For small electric fields, we can use Eq. (2.38) to find the relationship between E and the drift current density

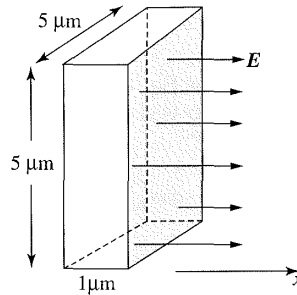
$$J_p^{dr} = q p \mu_p E = q N_a \mu_p E \quad v_{dp} \ll v_{sat}$$

where we have substituted $p = N_a$ for the hole concentration. Even though the holes are drifting with the electric field and the slab is not in thermal equilibrium, their concentration remains approximately equal to the acceptor concentration. As we will see later in this chapter, metal coatings on the left and right faces of the silicon slab can supply and remove the drifting carriers.

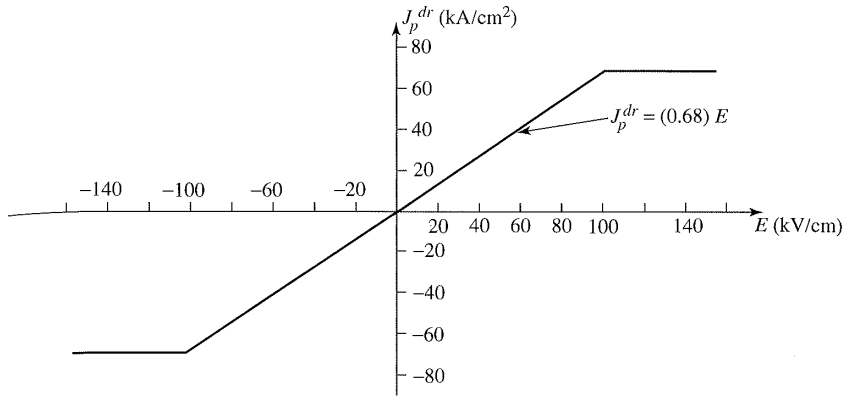
From Fig. 2.8, the hole mobility is $\mu_p = 425 \text{ cm}^2/(\text{Vs})$ and the low-field drift current density is $J_p^{dr} = (0.68) E$ (units: J : A/cm^2 , E : V/cm).

According to Fig. 2.9, this linear relationship is valid until the hole drift velocity saturates, which occurs at electric fields greater than $E_{crit} \approx 10^5 \text{ V/cm} = 100 \text{ kV/cm}$. Since we don't have a model for velocity saturation, we limit the drift current density at $J_{p,max}^{dr} = (0.68) \cdot E_{crit} = 68 \text{ kA/cm}^2$ as shown in Fig. Ex2.3B

Substituting for the thickness of the slab, the voltage needed to saturate the drifting holes is $E_{crit} \times 1 \mu\text{m} = (10^5 \text{ V/cm}) (10^{-4} \text{ cm}) = 10 \text{ V}$. The maximum current is given by the product of $J_{p,max}^{dr}$ and the area of the slab: $I_{max} = (68 \text{ kA/cm}^2) (5 \times 5 \times 10^{-8} \text{ cm}^2) = 17 \text{ mA}$.



► **Figure Ex2.3A** P-type silicon slab with uniform applied electric field E .



► **Figure Ex2.3B** Drift current density through slab as a function of the applied electric field.

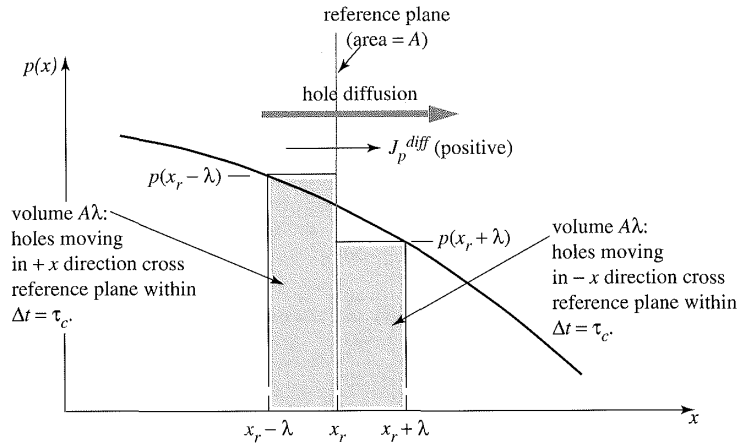
2.4.3 Diffusion Current Density

Diffusion is the process where gradients in concentration cause a net transport of particles. For example, suppose an ammonia bottle is opened in one corner of a room in which the air is at room temperature. Even if the air in the room is perfectly still, the random motions (with frequent collisions) of the air and ammonia molecules enable the transport of ammonia throughout the room. The net transport will be from regions of high concentration in the corner with the open bottle, to regions of low concentration in the opposite corner. If the bottle is capped, the concentration of ammonia will equalize throughout the room after a period of time. Diffusion reduces and eventually eliminates all gradients in ammonia concentration throughout the room.

Since electrons and holes also undergo random motion with frequent collisions, gradients in the electron or hole concentrations result in diffusion transport. In order to model diffusion, we consider holes having a concentration $p(x)$, which decreases with x as shown in Fig. 2.12. On the average, holes crossing the reference plane at $x = x_r$ started within one mean free path λ of x_r . The diffusion current density J_p^{diff} of holes across the reference plane has its origin in the difference in the number of holes in the two regions of volume $\Delta V = A\lambda$ to the left and right of the reference plane. Only half of the holes starting in either region will cross the reference plane in the time interval τ_c , while the other half are moving away from the reference plane. Therefore, the hole diffusion current density at $x = x_r$ is given by

$$J_p^{diff}(x=x_r) = q \left[\frac{\frac{1}{2} p(x=x_r-\lambda) A \lambda - \frac{1}{2} p(x=x_r+\lambda) A \lambda}{A \tau_c} \right] \quad (2.43)$$

where $A\lambda$ is the volume. Note that holes crossing from the right to the left, opposite to the $+x$ direction, enter into Eq. (2.43) with a negative sign.



► **Figure 2.12** One-dimensional picture for quantifying the hole diffusion current density as a function of the gradient in hole concentration.

For Eq. (2.43) to be a good approximation, the mean free path (about $0.01 \mu\text{m}$ at room temperature) must be much smaller than the distance over which $p(x)$ is varying. This requirement is usually satisfied for microelectronic device structures. Note that we have highly exaggerated the gradient in the hole concentration in Fig. 2.12 to make the difference in the hole population across the reference plane clearly visible. Using first-order Taylor expansions for $p(x = x_r + \lambda)$ and $p(x = x_r - \lambda)$ and simplifying, we find that

$$J_p^{\text{diff}}(x = x_r) = \frac{q\lambda}{\tau_c} \left(\frac{\left(p(x_r) - \left. \frac{dp}{dx} \right|_{x_r} \lambda \right) - \left(p(x_r) + \left. \frac{dp}{dx} \right|_{x_r} \lambda \right)}{2} \right) \quad (2.44)$$

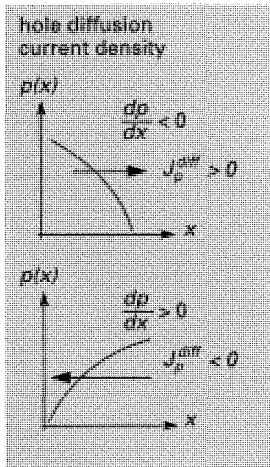
Canceling the hole concentration and collecting terms, Eq. (2.44) reduces to

$$J_p^{\text{diff}}(x = x_r) = -q \left(\frac{\lambda^2}{\tau_c} \right) \left. \frac{dp}{dx} \right|_{x=x_r} = -qD_p \left. \frac{dp}{dx} \right|_{x=x_r} \quad (2.45)$$

In Eq. (2.45), the term λ^2/τ_c is identified as the **hole diffusion coefficient** D_p , with units of cm^2/s . The diffusion current density for holes is

$$J_p^{\text{diff}} = -qD_p \frac{dp}{dx} \quad (2.46)$$

The proportionality of the diffusion current density to the concentration gradient is known as **Fick's Law** and is a fundamental result for diffusion processes. The negative sign in Eq. (2.46) reflects the fact that particles diffuse from regions of high con-



centration toward regions of low concentration. The concentration gradient dp/dx in Fig. 2.12 is negative, so the holes diffuse in the $+x$ direction and the diffusion current density is positive.

For the case of the electron concentration gradient shown in Fig. 2.13, we find the net transport of charge across the reference plane in the $+x$ direction as

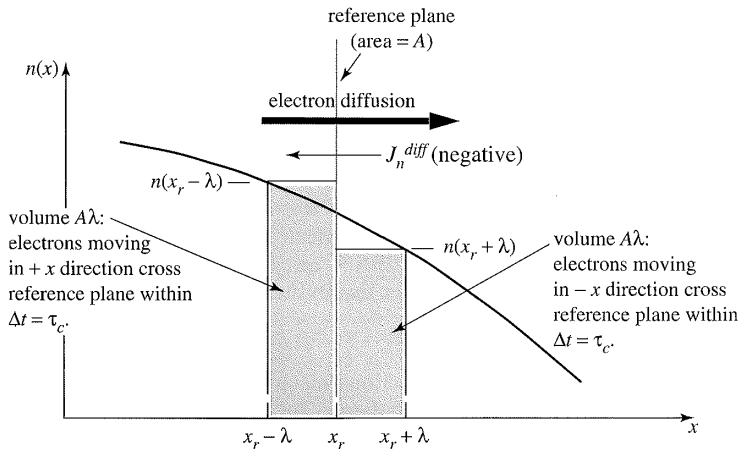
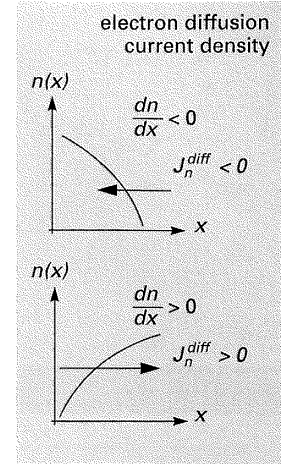
$$J_n^{diff}(x=x_r) = (-q) \left[\frac{\frac{1}{2}n(x=x_r-\lambda)A\lambda - \frac{1}{2}n(x=x_r+\lambda)A\lambda}{A\tau_c} \right] \quad (2.47)$$

This is identical in form to Eq. (2.43) for holes, except we have accounted for the negative charge on the electrons crossing the reference plane. Again, use first-order Taylor expansions for the concentrations at $x_r - \lambda$ and $x_r + \lambda$

$$J_n^{diff}(x=x_r) = \frac{(-q)\lambda}{\tau_c} \left(\frac{n(x_r) - \left. \frac{dn}{dx} \right|_{x_r} \lambda}{2} - \frac{n(x_r) + \left. \frac{dn}{dx} \right|_{x_r} \lambda}{2} \right) \quad (2.48)$$

Collecting terms, the electron diffusion current density can be expressed by

$$J_n^{diff}(x=x_r) = q \left(\frac{\lambda^2}{\tau_c} \right) \left. \frac{dn}{dx} \right|_{x=x_r} = qD_n \left. \frac{dn}{dx} \right|_{x=x_r} \quad (2.49)$$



► **Figure 2.13** One-dimensional picture for quantifying the electron diffusion current density as a function of the gradient in electron concentration. Note that the electrons diffuse left-to-right (in the $+x$ direction), but that J_n^{diff} points in the $-x$ direction and is negative.

where the **electron diffusion coefficient** is identified as D_n . The negative gradient in electron concentration ($dn/dx < 0$) shown in Fig. 2.12 means the electron diffusion current density is negative (i.e., it points in the $-x$ direction). The physical reason behind this result is that the electrons diffuse from regions of high concentration to regions of low concentration in the $+x$ direction in Fig. 2.13. However, each electron carries a charge $-q$, which results in J_n^{diff} pointing in the $-x$ direction.

The diffusion coefficients for holes and electrons are functions of doping concentration and temperature since these are determined by the mean free path and collision time. The mobility μ and diffusion coefficient D were derived from physical models based on large numbers of carriers undergoing random thermal motion with frequent collisions. In a solid-state physics course, it is shown that the two constants are proportional and their ratio is equal to

$$\frac{D}{\mu} = \frac{kT}{q} \quad (2.50)$$

where $k = 1.38 \times 10^{-23}$ J/K is Boltzmann's Constant. This important result is called the **Einstein Relation**—named after Albert Einstein, who was active in solid-state physics in addition to his other interests. Eq. (2.50) holds separately for holes and for electrons

Einstein Relation

$$\frac{D_p}{\mu_p} = \frac{kT}{q} \quad \text{and} \quad \frac{D_n}{\mu_n} = \frac{kT}{q} \quad (2.51)$$

One implication of the Einstein Relation is that only the mobility need be measured. The diffusion coefficient can be found directly from Eq. (2.51). Using room-temperature values for the electron and hole mobilities in silicon (1000 and 400 cm²/Vs), typical round numbers for the electron and hole diffusion coefficients are

$$D_n = 25 \text{ cm}^2/\text{s} \text{ and } D_p = 10 \text{ cm}^2/\text{s} \quad (2.52)$$

The **thermal voltage** is defined as the quantity kT/q :

thermal voltage

$$V_{th} = \frac{kT}{q} \quad (2.53)$$

The thermal voltage has units of Volts and will appear in many device physics equations. It ranges from 25 mV in a cool room in which the temperature is 17 °C or 62 °F to 26 mV in a somewhat warm room at 28.5 °C or 83 °F. In this book, we will choose either $V_{th} = 25$ mV or 26 mV—whichever results in “round numbers.”

2.4.4 Total Current Density

As we have seen in the previous two sections, electrons and holes drift in an electric field and diffuse due to concentration gradients. The total current density J is the sum of all four components

$$\begin{aligned} J &= J_n + J_p = J_n^{dr} + J_n^{diff} + J_p^{dr} + J_p^{diff} \\ J &= qn\mu_n E + qD_n \frac{dn}{dx} + qp\mu_p E - qD_p \frac{dp}{dx} \end{aligned} \quad (2.54)$$

The total electron and total hole current densities J_n and J_p are important results, therefore they are given separately

$$J_n = J_n^{dr} + J_n^{diff} = qn\mu_n E + qD_n \frac{dn}{dx} \quad (2.55)$$

$$J_p = J_p^{dr} + J_p^{diff} = qp\mu_p E - qD_p \frac{dp}{dx} \quad (2.56)$$

electron and hole total current densities

In most cases, we will be able to simplify Eqs. (2.55) and (2.56) for a particular device structure and focus on the dominant components of the current densities.

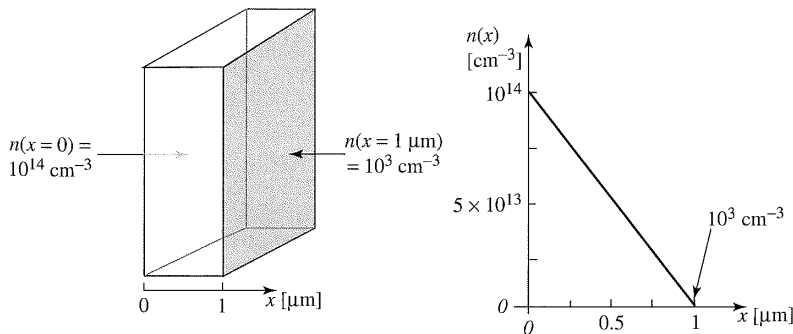
► EXAMPLE 2.4 Diffusion Current Density

Figure Ex2.4 shows a 1 μm -thick slab of p-type silicon having $N_a = 10^{17} \text{ cm}^{-3}$. By continuously injecting minority electrons on the left side at $x = 0$ and extracting them from the right side at $x = 1 \mu\text{m}$, we have set up a linear gradient in the electron concentration $n(x)$. What is the electron diffusion current density in the slab?

SOLUTION

On the linear plot of $n(x)$, we note that the electron concentration at $x = 1 \mu\text{m}$, 10^3 cm^{-3} , is indistinguishable from zero. For a total doping concentration of 10^{17} cm^{-3} , Fig. 2.8 indicates that the *electron* mobility is $\mu_n = 750 \text{ cm}^2/(\text{Vs})$. Substituting into the Einstein Relation from Eq. (2.51), we find that the electron diffusion coefficient is

$$D_n = V_{th} \cdot \mu_n = (25 \times 10^{-3} \text{ V}) \cdot (750 \text{ cm}^2/\text{Vs}) = 18.75 \text{ cm}^2/\text{s}$$



► **Figure Ex2.4** Electron concentration through a slab of p-type silicon.

We apply Eq. (2.49) to calculate the electron the diffusion current density in the silicon slab

$$J_n^{diff} = qD_n \frac{dn}{dx} = (1.6 \times 10^{-19}) (18.75) \left(\frac{10^3 - 10^{14}}{10^{-4}} \right) = -3 \text{ A/cm}^2$$

Note that the current density is negative since the electrons are diffusing from left to right and carry a negative charge. The magnitude of the diffusion current density in this example is also small in comparison with typical majority carrier drift current densities. A voltage of less than 1 mV between the left and right faces of the slab is all that is needed to cause a drift current of majority holes of 3 A/cm².

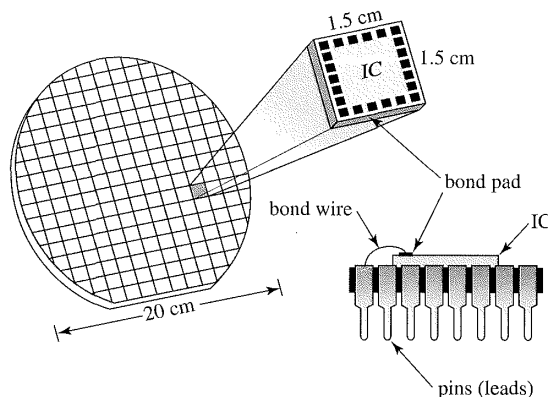
► 2.5 SILICON INTEGRATED CIRCUIT TECHNOLOGY

The focus of this text is silicon **integrated** devices and circuits, in which an electronic functional block, subsystem, or entire system is implemented on a single integrated circuit (IC). A basic appreciation of the manufacturing processes used to make ICs will be helpful in later chapters since it will give you an appreciation for the dimensional scales and physical structures of integrated devices. As you will see, the origin of the important “parasitic” circuit elements in ICs is obvious from the device structures.

This section provides a basic introduction to **IC fabrication technology**. This fascinating discipline draws from materials science and engineering, chemical engineering, and applied physics, as well as electrical engineering. By many measures, silicon IC technology is the crowning technological achievement of the late 20th century. Few areas of life have been unaffected by the penetration of silicon ICs such as microprocessors, microcontrollers, and modems. Since the fundamental physical limits of these products are at least two decades away, the early 21st century is certain to see further advances in the core fabrication processes that will have great impact on computers and other electronic systems. In addition, “spin-offs” from IC technology are catalyzing new fields such as micro electromechanical systems (MEMS). Silicon MEMS have display, sensing, or actuating functions. They may incorporate 10⁶ micromechanical elements on the silicon chip along with a complex electronic interface. In summary, although the first silicon IC dates to around 1960, the end of the IC revolution is by no means in sight.

Figure 2.14 shows the fundamental unit of IC manufacturing—a silicon **wafer**. The typical wafer diameter is 20 cm in the mid-1990’s. The wafer is a nearly perfect crystal that has been polished to be extremely flat. Wafers of this diameter cost about US\$50.00 in large quantities. A group of around 24 wafers, called a **lot**, undergoes the same sequence of fabrication steps together. In a complex IC such as a microprocessor or dynamic random-access memory (DRAM), the process has around 200 steps and takes about three weeks to complete for a typical manufacturing schedule.

Figure 2.14 shows that when the process is complete, the wafer is tiled with an array of ICs interspersed with a few diagnostic test chips for monitoring the fabrication process. A typical complex IC has an area of about 1.5 cm × 1.5 cm, or about 2 cm², on which are several *million* electronic devices. For a 20 cm-diameter wafer, there will be around 150 chips of this area. The cost of the completed wafer is roughly US\$2,000, which is a function of the fabrication process. How many of the chips work



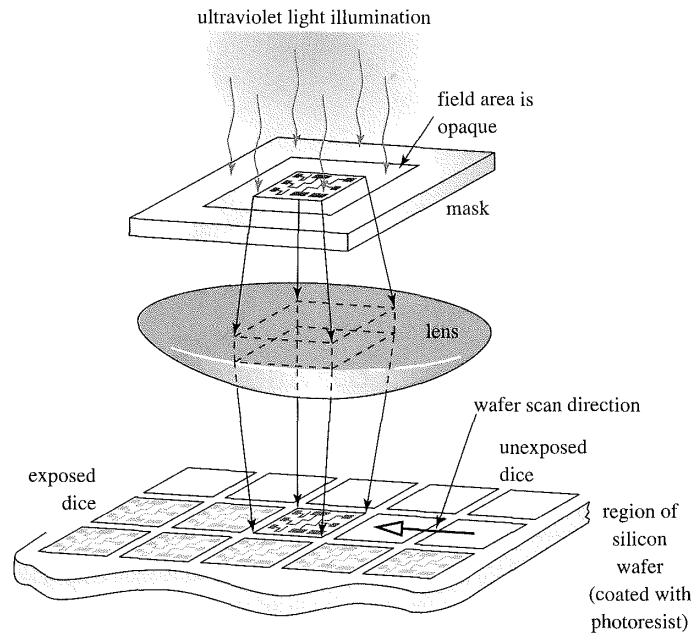
► **Figure 2.14** Silicon wafer and inset showing a single IC chip (die), with bond pads, and a side view of a dual in-line (DIP) package, shown with the lid removed and only a single bond wire. Wafer diameter and chip area are typical for mid-1990's large-volume, complex ICs.

(the **yield**) will affect the price at which the manufacturer sells each chip. The yield is a function of the quality control in the manufacturing plant, as well as the design margins built into the process and circuit. For example, a high-performance IC may have dimensions that approach the technological limits with the result that the yield is low. However, the high performance may justify a very high price and make the process profitable despite the low yield. As a particular IC design matures in production, the yield can be predicted to improve to around 90% or higher, which allows the selling price to drop steadily. The IC industry is extremely competitive. Companies aggressively track the **yield curve** of most products by dropping prices as yields increase. This maximizes volume and makes it more difficult for other companies to enter into direct competition on the product.

General-use ICs are packaged in plastic dual in-line (DIP) packages as shown in Fig. 2.14. In recent years, there has been great progress in economical packaging of ICs into high-density **multichip modules** for use in applications such as lap-top computers and cellular phones. Instead of the DIP shown in Fig. 2.14, chips are often encapsulated in low-profile surface mount packages or mounted directly on the circuit board using a “flip-chip” technique. Electrical connections between the board and inverted chip are made through electroplated solder bumps, rather than through bond wires, allowing for many more interconnects per chip.

2.5.1 Photolithography and Pattern Transfer

The essential process in IC fabrication is the imaging, alignment, and transfer of complex patterns onto the silicon wafer. A modern fabrication process uses around 15–20 separate patterning steps to define the transistors, diodes, and several levels of electrical interconnections (i.e., thin-film wires) that make up the IC. A **wafer stepper** is used to transfer patterns from an optical plate called a **mask**, to a photosensitive film coating the top surface of the wafer called **photoresist**. The process shown schematically in Fig. 2.15 depicts the exposure of a single die to the image of the mask.



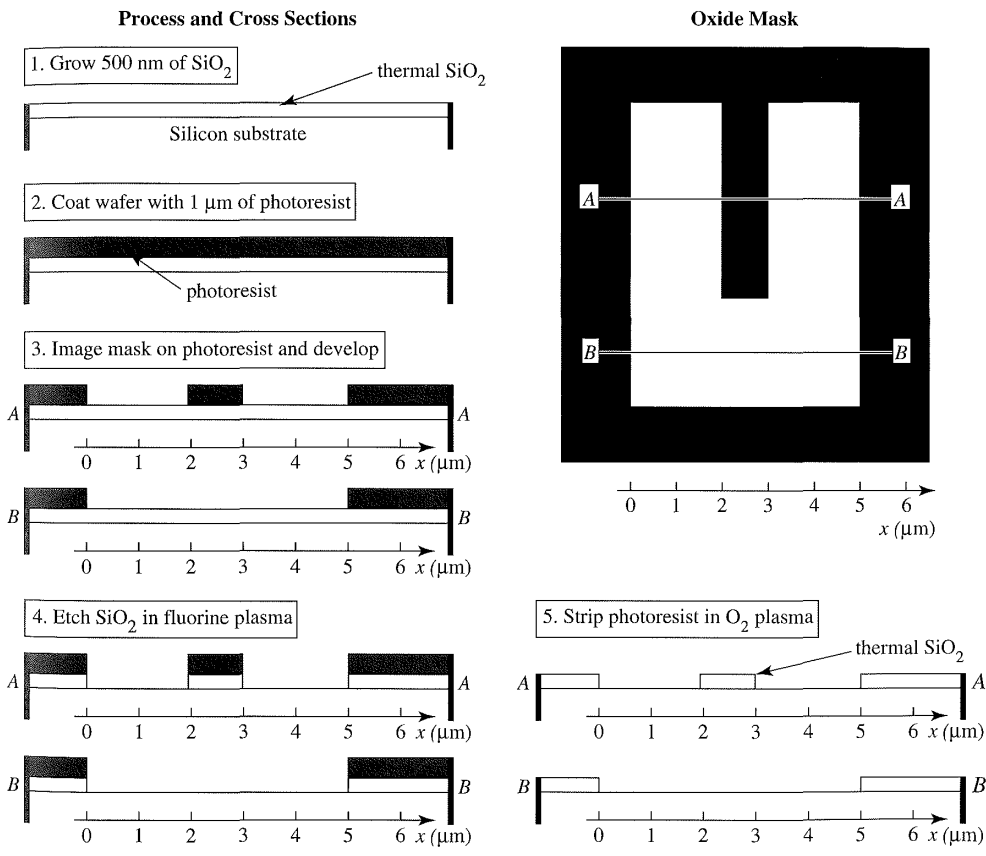
► **Figure 2.15** Schematic diagram showing imaging of the mask pattern on one die in a wafer stepper. The mask pattern is shown in gray on exposed dice. Unexposed dice are shown in outline only. In reality, the image in the photoresist is not visible until after development.

After completion of the exposure, the wafer stepper indexes to the next die, automatically aligns to special features previously etched on it, and repeats the exposure. Since the image of the mask pattern must be distortion-free and the error allowed in the alignment is less than $0.1\ \mu\text{m}$, it is not surprising that a high-volume wafer stepper costs well over US\$2 million in the mid-1990s.

Figure 2.16 shows the step-by-step process of how the transfer from the mask to the wafer is accomplished through the photolithography and pattern transfer process. Figure 2.16 also introduces the customary way of visualizing three-dimensional IC structures. The **layout** consists of the series of masks used to create the IC structure. Lines across the layout denote the locations of **cross sections** through the silicon wafer. Typically, only the top few μm of the wafer are needed to show the device structure. In order to draw the cross sections, we must know the sequence of masking and fabrication steps. This sequence is called the **fabrication process**, or **process** for short. In analogy with food preparation, the fabrication process is sometimes called the **recipe**.

The single-mask photolithography and pattern transfer process in Fig. 2.16 involves five steps.

1. A 500 nm-thick film of silicon dioxide, SiO_2 , is grown on the silicon wafer. In order to grow the thermal SiO_2 film, the wafer is placed in a furnace at an elevated temperature. Silicon surfaces react with oxygen to form a very high quality insulating film.
2. The wafer is uniformly coated with a 1 μm -thick film of photoresist by applying a mixture of photoresist and solvent, spinning the wafer at several thousand rpm for 30 seconds, and then baking to drive off the solvent.
3. Areas of this film that are exposed to ultraviolet light (UV) during the exposure step dissolve in an alkaline solution—a process known as **development**. Note carefully the differences in cross sections *A-A* and *B-B* after development of the mask image in the photoresist.



► **Figure 2.16** Cross sections for transferring the oxide mask pattern to a thermal oxide film. A shorthand description of this process is “grow 500 nm of thermal SiO_2 and pattern using the oxide mask.”

4. After baking the photoresist to drive off residual solvents, the wafer is immersed in a vacuum chamber containing a fluorine **plasma**. A typical example of a plasma is the glow discharge in a fluorescent light. The plasma reacts with SiO_2 and etches it much faster than it etches the photoresist film. Once the silicon dioxide is etched away from the exposed areas, the underlying silicon is also etched, but very slowly. This enables the etch to be terminated before significant silicon erosion occurs.
5. The final step is to remove the photoresist in an oxygen plasma. This plasma aggressively attacks organic materials, such as photoresist, but does not etch silicon or silicon dioxide. Comparing the cross sections for steps 3 and 5 in Fig. 2.16, we see that the mask pattern has been successfully transferred to the silicon dioxide film.

As we mentioned, an IC process involves many photolithography and pattern transfer steps on deposited films. In order to avoid repeating steps 1–5 each time we use this procedure, we will summarize Fig. 2.16 by the statement: “Grow 500 nm of SiO_2 and pattern using the oxide mask.” The photolithography and pattern transfer sequence can be thought of as a subroutine of the IC fabrication process.

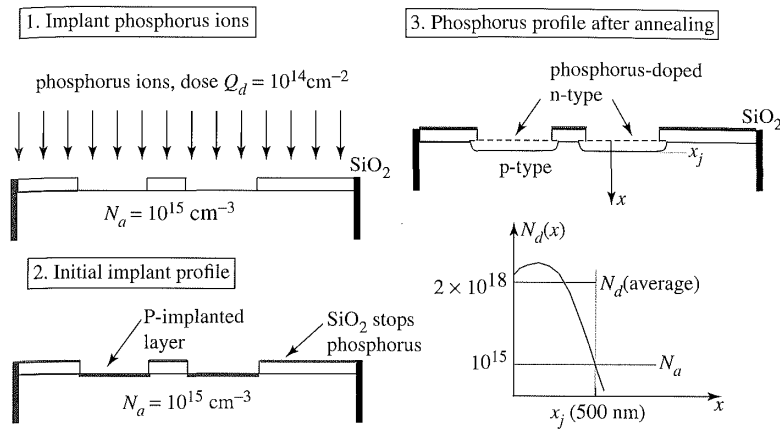
2.5.2 Doping by Ion Implantation

Acceptors and donors can be incorporated into selected regions of the top surface of a silicon wafer by means of **ion implantation**. Ions of boron, arsenic, or phosphorus are extracted from a plasma, accelerated to energies from 20 keV to 3 MeV (eV = electron volt), and formed into a tightly aligned ion beam. The ion beam is scanned over the surface of the silicon wafer. The beam is scanned until the desired **dose** of ions is implanted into the wafer. The ion dose has units of ions per unit area (cm^{-2}). After entering the silicon, the bombarding ions are slowed and then stopped by collisions with the silicon lattice.

Ion implantation creates a damaged region near the surface, since the collisions knock silicon atoms out of their positions in the lattice. The damaged region loses its crystal structure entirely and becomes **amorphous**. By heating the silicon wafer above 900°C , the damaged region recrystallizes, a process which is known as **annealing**. Remarkably, most of the dopant ions in the recrystallized region end up on substitutional lattice sites. During annealing, the implanted dopant ions also diffuse further into the silicon wafer.

Figure 2.17 illustrates how ion implantation can be used to dope specific regions of the wafer. We start with the patterned oxide film from Fig. 2.16, grown on a p-type wafer with a background donor concentration $N_a = 10^{15} \text{ cm}^{-3}$. A dose of $Q_d = 10^{14} \text{ cm}^{-2}$ phosphorous ions is implanted into the wafer. Their energy is such that the phosphorus ions do not penetrate the SiO_2 film as shown in Fig. 2.17. After a thermal cycle, the phosphorus concentration $N_d(x)$ is found to exceed the background donor concentration up to a depth $x_j = 500 \text{ nm}$, which is called the **junction depth**. Over the past 20 years, extensive research has investigated and modeled the details of dopant concentrations after implantation and annealing. Further discussion of this subject will be left to advanced courses on IC fabrication technology.

In this book, we will simply use the average donor concentration in the layer N_d as defined in the ion profile in Fig. 2.17. Since the dose Q_d of donors (units: cm^{-2}) is located in a layer of depth x_j after annealing, the average concentration is



► **Figure 2.17** Phosphorus implantation into the patterned oxide film shown in Fig. 2.16. The average concentration of phosphorus is defined in the plot of $N_d(x)$ after annealing.

$$N_d = \frac{Q_d}{x_j} \quad (2.57)$$

For the case shown in Fig. 2.17, $N_d = (10^{14} \text{ cm}^{-2}) / (5 \times 10^{-5} \text{ cm}) = 2 \times 10^{18} \text{ cm}^{-3}$.

2.5.3 Deposited Conducting and Insulating Films

In order to fabricate integrated circuits, we need conducting films that can be patterned into the microscopic wiring called **interconnects**. The interconnects enable electrical signals to be propagated around the IC. In addition, insulating films are needed to enable crossovers between two or more levels of interconnects. In the mid-1990s, three to five levels of interconnect are used for efficient packing and for minimum interconnect lengths in high-density digital ICs. As we will see in Chapter 5, short interconnects are important for achieving high switching speed.

Our first conductor is **polycrystalline silicon** or **polysilicon**. This material is an aggregate of crystallites or grains of single crystal silicon. Thin films of polysilicon are deposited on the surface of a silicon wafer at temperatures around 600 °C in a process known as **chemical vapor deposition** (CVD). Polysilicon is usually deposited with a maximum doping level of phosphorus. As we will see in the next section, polysilicon is a mediocre conductor; however, it has the virtues of being thermally stable and free of contaminants.

Aluminum is the most commonly used conducting film for interconnects. It is deposited by **sputtering**, a process in which material is deposited on the wafer from a nearby target that is bombarded by ions from a plasma discharge. One drawback of aluminum interconnects is that they cannot tolerate temperatures higher than about

450 °C due to degradation of the Al-Si junctions. Other metals used increasingly in ICs are tungsten and copper, both of which can be deposited by CVD processes.

There are several processes for depositing insulating films using CVD. The most common inter-metal insulator is silicon dioxide, SiO_2 . Phosphorus and boron are often added during growth in order to allow the film to soften and flow slightly at temperatures around 900° C. This process, called **reflow**, is important in smoothing (or planarizing) the surface of the wafer after several depositions and masking steps have created too rough a surface. Silicon nitride, Si_3N_4 , is another CVD insulating film that is able to mask a high-temperature thermal oxidation since it is impermeable to oxygen. In plasma-enhanced CVD processes, SiO_2 and Si_3N_4 can be grown at low temperatures (around 300° C) for use as a final protective layer over the IC.

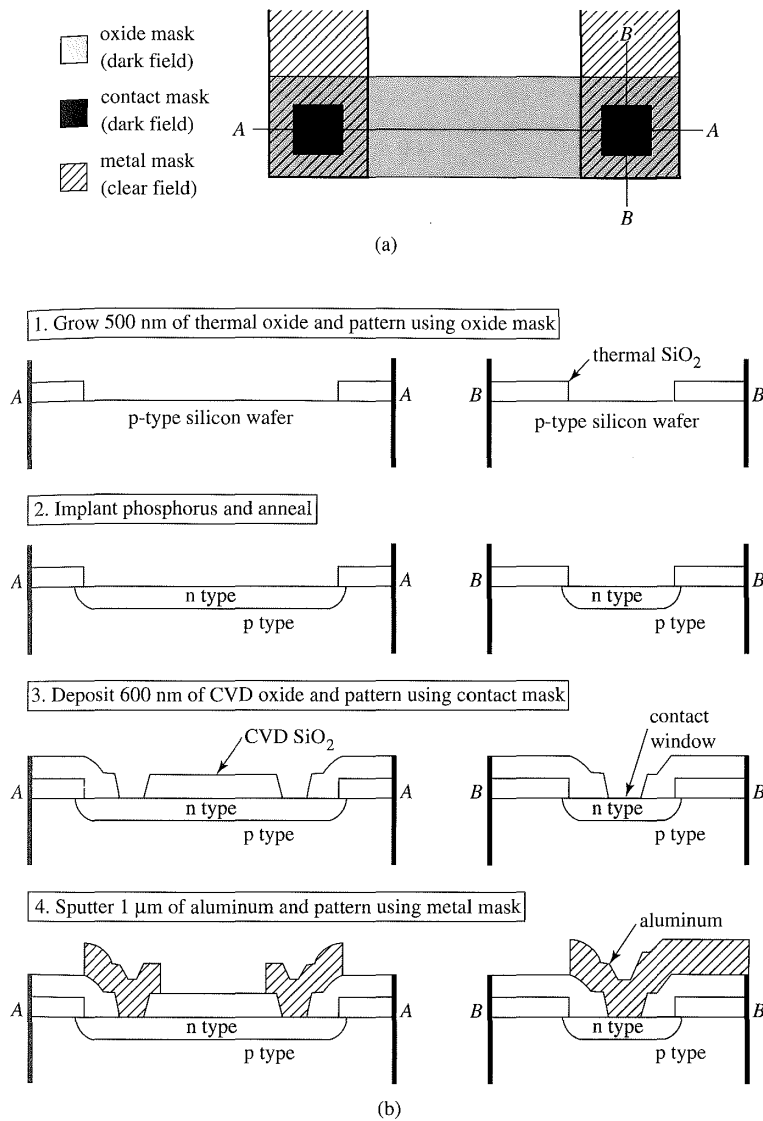
2.5.4 Layouts and Process Flows

You are now familiar with the lithography and etching processes and the basic materials that are the foundation of silicon IC technology. In this section, we will first describe a three-mask process for fabricating an ion-implanted resistor and then describe a more complex process to make an MOS field-effect transistor. You need not understand anything at this point about how these structures function as electronic devices. In following these sample processes, focus on the connection between the layout and the evolving device cross sections.

Figure 2.18(a) depicts the layout of the resistor, as it would be viewed in a computer-aided design (CAD) program. At first glance, you would expect that these patterns would correspond exactly to the actual masks used for photolithography. In fact, a mask that is nearly opaque (referred to as **dark field** mask) will obscure the other masks. Even with the use of color and a variety of “fill” and “border” designs, a nearly opaque mask will make it difficult to see features on the other masks in a layout. Fortunately, there is a simple solution: we display the negative, or complement, of the dark field mask in depicting the layout. **Clear field** masks have few opaque regions and are displayed directly. Of course, we must keep a careful record of which are clear field and which are dark field, so that the correct masks can be made. Many wafer lots have been ruined due to miscommunication on this minor, but critically important point.

In the resistor layout shown in Fig. 2.18(a), the oxide mask is identified as a dark field mask. Since the negative of the oxide mask is shown in the layout, the rectangle is clear on the actual mask. You should note carefully from cross sections *A-A* and *B-B* after step 1 that the oxide has been etched from the region of the wafer exposed by the clear rectangle. If this mask had been clear field, the oxide would have been removed everywhere *except* the rectangle, which would have been dark on the physical mask.

We now describe the process sequence in Fig. 2.18(b). The first step is to grow 500 nm of thermal oxide, which is done by heating the wafer in a furnace with oxygen flowing through it. This film is patterned using the oxide mask and the photolithography and etching process shown in Fig. 2.16. The wafer is then implanted with a dose of phosphorus ions that is sufficiently heavy to compensate the background acceptor concentration in the p-type silicon wafer. The thermal SiO_2 is thick enough to stop the phosphorus ions from penetrating into the underlying silicon substrate and is, therefore, effective at masking the implant.



► **Figure 2.18** (a) Layout and (b) process flow with cross sections A-A and B-B for fabrication of an ion-implanted resistor.

After annealing the wafer at high temperature, the phosphorus diffuses further into the substrate and creates a rectangular n-type region as shown in the cross sections after Step 2 in Fig. 2.18. Note that the phosphorus diffuses laterally under the oxide during the anneal, with the result that the dimensions of the n-type region are somewhat larger than the rectangle on the oxide mask. A chemical-vapor deposition (CVD) process is then used to deposit 600 nm of CVD SiO_2 . Contact windows are patterned in this film using a second photolithography and etching step. Note that the contact mask is also a dark field mask. A 1 μm -thick film of aluminum is sputtered and patterned using the metal mask, which is a clear field mask. In order for the aluminum to form good electrical contacts to the n-type silicon region, the wafer must be heated to around 425 °C. It is helpful to observe the differences in the *A-A* and the *B-B* cross sections in the completed structure and to connect these to the layout.

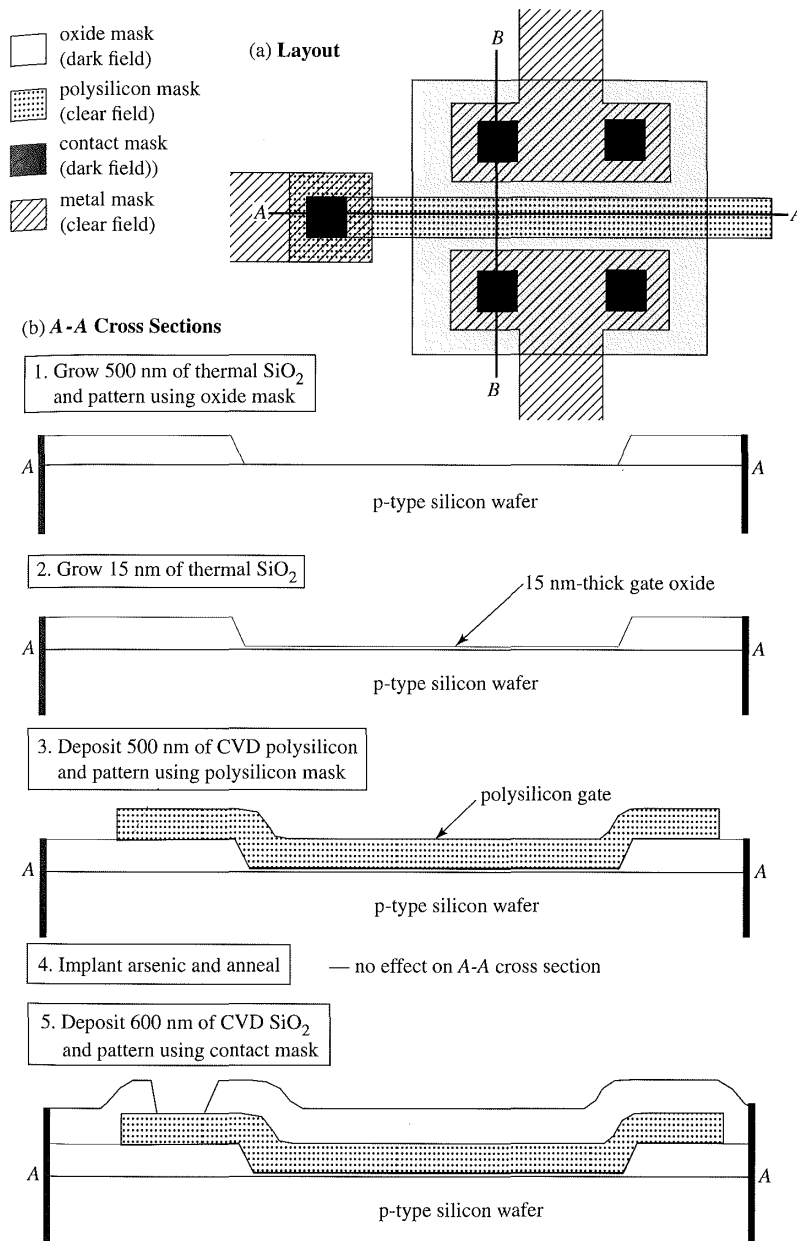
With only a slightly more complicated process, we can fabricate a “starter” version of an MOS field-effect transistor—the core device in digital ICs. Figure 2.19 gives the four mask patterns and the sequence of cross sectional views of the structure. After patterning a window in the initial 500 nm thick SiO_2 film in the first masking step, a 15 nm-thick SiO_2 “gate oxide” layer is grown in the bare silicon regions by a high-temperature oxidation process. Polysilicon is deposited by CVD with n-type doping and is patterned in the second masking step. The polysilicon film serves as the mask for a high-dose arsenic implant that penetrates the gate oxide. After annealing, n-type regions are formed on each side of the polysilicon “gate” that are seen only in the *B-B* cross section perpendicular to the gate. There is some lateral spreading of the arsenic implant during the anneal as seen in the *B-B* cross section. The next step involves the deposition of a CVD SiO_2 film, etching contact windows to the gate and to the two n-type regions adjacent to it. Finally, the MOSFET is completed by deposition and patterning of the aluminum interconnect layer.

An IC technology is often specified by λ , its minimum feature size. In an MOS technology, λ is the etched polysilicon gate length as shown in the final cross section in Fig. 2.19(b). The trend has been to ever-smaller dimensions, with state-of-the-art production processes at $\lambda < 0.5 \mu\text{m}$ in the mid-1990s.

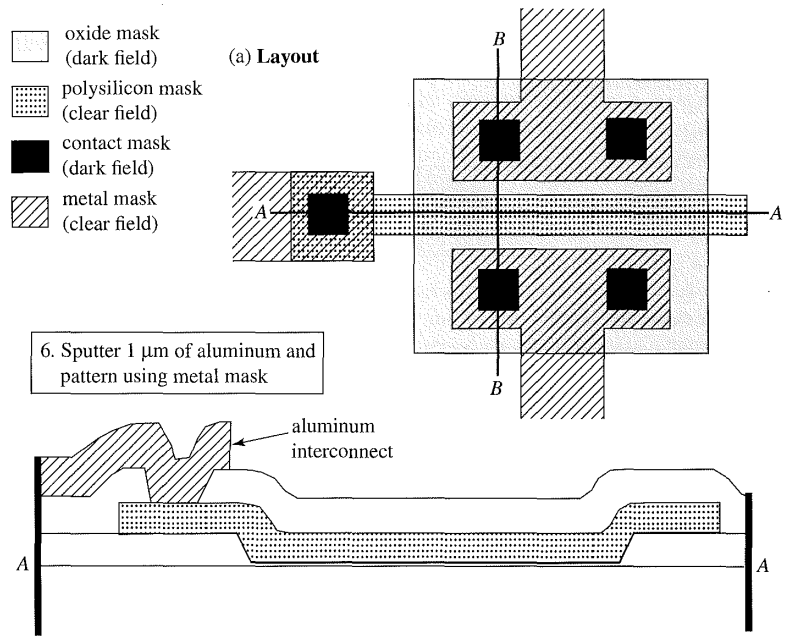
2.5.5 Geometric Design Rules

In Fig. 2.19(a), the contact nests inside both the oxide and metal patterns. The overlaps between patterns on different masks are specified in the **geometric design rules** for the process. These rules allow for systematic changes in the dimensions of the device structures, due to lithography, diffusion, or etching. They also incorporate knowledge of the random variations in these factors and the misalignment tolerances of the mask making and lithography equipment. There are “proximity” rules that specify the separation between the edges of patterns on different masks. Finally, there are rules that define the minimum line and space widths for each mask.

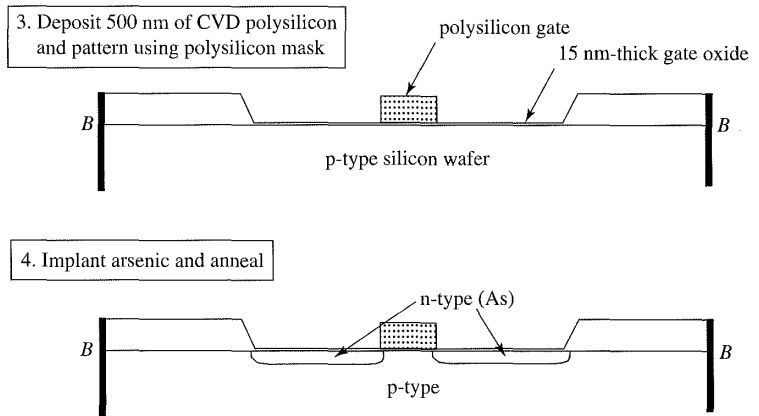
The yield of a particular product in a given process technology will be heavily influenced by the margins built into the design rules. With an IC containing 10^6 MOSFETs, each with at least three contacts, carefully chosen design rules for the contact mask nesting are essential for achieving a balance between high yield and a reasonable die size. Systematic experimentation and modeling of statistical variations are required to optimize the design rules for a particular technology and product family.



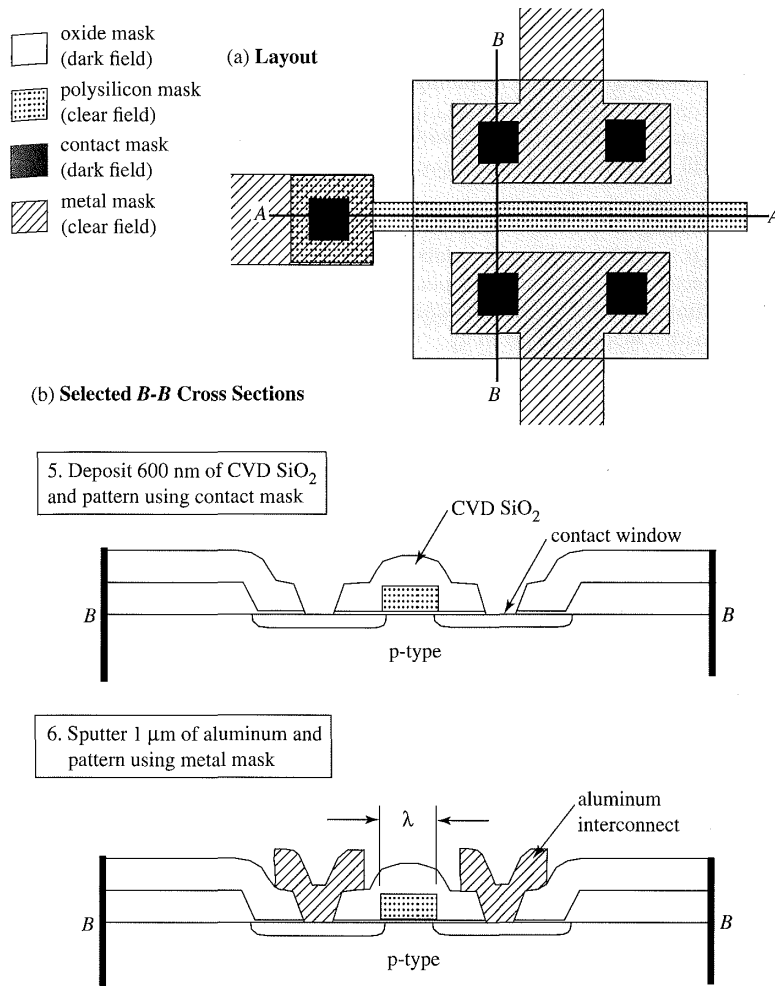
► **Figure 2.19** (a) Layout and (b) A-A cross sections for a simple MOSFET



(b) Selected *B-B* Cross Sections



► **Figure 2.19 (cont.)** (a) Layout and final *A-A* section and (b) selected *B-B* cross sections for a MOSFET.



► **Figure 2.19 (cont.)** (a) Layout and (b) selected *B-B* cross sections for the simple MOSFET.

► 2.6 IC RESISTORS

Now that you have a basic understanding of IC technology, we are prepared to model our first IC device—a silicon resistor. First we must derive Ohm's Law from the fundamental concept of drift current density. We will then re-express the resistance in a form that will be more useful for IC design.

2.6.1 Ohm's Law

Figure 2.20 is a schematic view of a slab of n-type silicon with two sides having metallized contacts. A DC voltage V is applied across the contacts. What is the magnitude of the resulting current I ? Our first observation is that there is an electric field E in the silicon with a magnitude given by

$$E = \frac{V}{L} \quad (2.58)$$

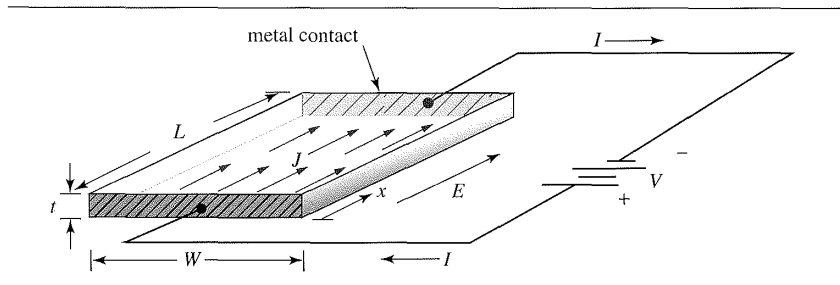
where L is the length of the slab in Fig. 2.20.

The electric field points in the direction of the x coordinate since it points from the contact with the higher potential (the nearer one in Fig. 2.20) to the contact with the lower potential. This electric field will cause electrons and holes in the silicon slab to drift, with the resulting drift current density being given by

$$J = J_n^{dr} + J_p^{dr} = qn\mu_n E + qp\mu_p E = (qn\mu_n + qp\mu_p) \left(\frac{V}{L} \right) \quad (2.59)$$

in which we have assumed that the electric field is low enough that velocity saturation is not an issue. There is no concentration gradient along the slab, since the mobile carriers are supplied and removed by the metal contacts. As a result, diffusion is ignored in Eq. (2.59). Since the current I [units: A] is the product of the current density J [units: A/cm²] and the cross sectional area of the slab $A = Wt$, we can find an equation relating current and voltage

$$I = J \cdot A = (qn\mu_n + qp\mu_p) \left(\frac{V}{L} \right) (Wt) = \left[(qn\mu_n + qp\mu_p) \left(\frac{Wt}{L} \right) \right] \cdot V \quad (2.60)$$



► **Figure 2.20** n-type silicon slab with length L , width W , and thickness t . A voltage V is applied across two metal contacts at opposite ends.

Ohm's Law states that $I = V/R$ and therefore, we can identify the resistance R of the slab from Eq. (2.60)

$$R = \frac{1}{(qn\mu_n + qp\mu_p)\left(\frac{Wt}{L}\right)} = \left(\frac{1}{qn\mu_n + qp\mu_p}\right)\left(\frac{L}{Wt}\right) \quad (2.61)$$

The first term on the right-hand side of Eq. (2.61) is defined as the **resistivity**, ρ [units: $\Omega \text{ cm}$]:

$$\rho = \frac{1}{qn\mu_n + qp\mu_p} \quad (2.62)$$

Substituting for the resistivity, Eq. (2.61) reduces to a simple expression

$$R = \rho \left(\frac{L}{Wt}\right) \quad (2.63)$$

In order to evaluate the resistivity, we must substitute the electron and hole concentrations n and p . The slab in Fig. 2.20 was assumed to be n-type with a donor concentration of N_d . Although the slab is not in thermal equilibrium, the carrier concentrations are not significantly perturbed from their equilibrium values by their constant drift across the slab. Substituting $n = N_d$ and $p = n_i^2/N_d$ for the n-type slab in Eq. (2.62), the resistivity ρ_n of n-type silicon is given by

$$\rho_n = \left(qN_d\mu_n + \frac{qn_i^2\mu_p}{N_d}\right)^{-1} \cong \frac{1}{qN_d\mu_n} \quad (2.64)$$

n-type resistivity

In other words, we can neglect the contribution of minority carriers in the drift current density since their concentration is orders of magnitude smaller than the majority carriers. For a p-type silicon region, holes are the majority carriers and the resistivity ρ_p is

$$\rho_p = \left(q\frac{n_i^2}{N_d}\mu_n + qN_a\mu_p\right)^{-1} \cong \frac{1}{qN_a\mu_p} \quad (2.65)$$

p-type resistivity

For doping concentrations of 10^{13} cm^{-3} to 10^{19} cm^{-3} , the resistivity of silicon ranges from about $500 \Omega\text{-cm}$ to $5 \text{ m}\Omega\text{-cm}$.

2.6.2 Sheet Resistance

Integrated-circuit resistors are fabricated from regions of one doping type in a substrate of opposite type. Figure 2.21 shows the layout and cross section of an n-type, ion-implanted IC resistor, that is fabricated using the process illustrated in Fig. 2.18. Metal interconnects make electrical contact at each end of the implanted region through contact windows etched through the deposited oxide film. The rectangular central portion of the implanted region has length L , width W , and thickness t . The

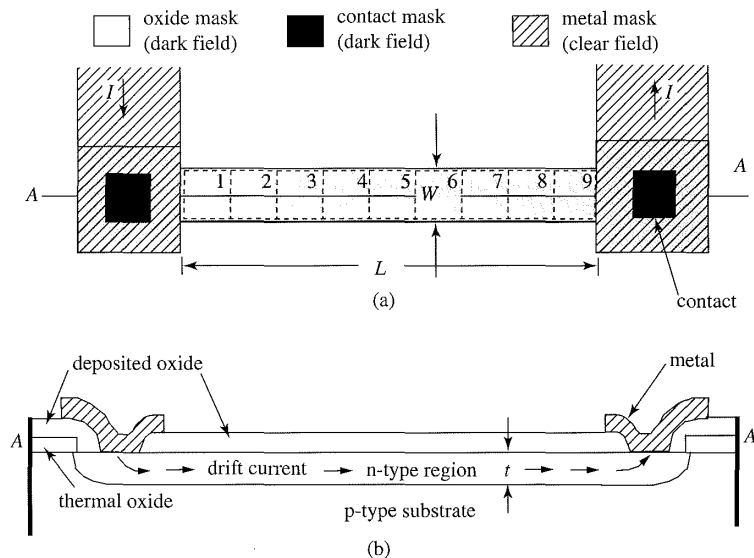
design rules require that the ends of the n-type region are widened to allow room for the contact windows.

If a voltage is applied across the contacts to the resistor, the resulting electric field will cause a lateral electron drift current density as shown in Fig. 2.21(b). As we will see in Chapter 3, if we apply positive voltages to both contacts with respect to the substrate, there is negligible current across the n-p junction between the n-type region and the p-type substrate. Hence, the n-type region is electrically isolated from the substrate. This simple structure hints at some of the difficulties in integrated circuit design, in that a p-n junction is the unwanted by-product of making an ion-implanted resistor. We will revisit the IC resistor in Example 3.6 in Chapter 3 and develop a more accurate model for it, after we have studied the reverse biased p-n junction.

From the previous section, the resistance of the rectangular region is given by

$$R = \rho_n \left(\frac{L}{Wt} \right) = \left(\frac{1}{qN_d\mu_n} \right) \left(\frac{L}{Wt} \right) \quad (2.66)$$

In IC design, it is very convenient to re-express Eq. (2.66) in a form that separates those parameters determined by the fabrication process (mobility μ_n , donor concentration N_d , and thickness t of the implanted layer) from those that are determined by



► **Figure 2.21** (a) Layout (b) cross section of an n-type resistor having $L/W = 9$ that is fabricated using the process sequence described in Fig. 2.18. Arrows show direction of electron drift current density for the case of a higher voltage on the left-hand contact.

the layout (length L and width W of the resistor). Collecting terms in Eq. (2.66) accordingly

$$R = \left(\frac{1}{q N_d \mu_n t} \right) \left(\frac{L}{W} \right) = R_{\square} \left(\frac{L}{W} \right) = R_{\square} N_{\square} \quad (2.67)$$

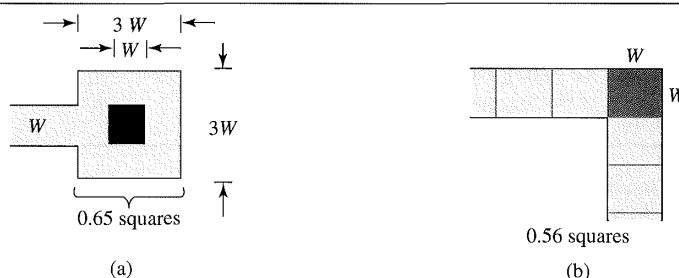
sheet resistance

in which we have defined the **sheet resistance** $R_{\square} = (q \mu_n N_d t)^{-1}$. The ratio of length to width is also defined as $N_{\square} = (L/W)$, the number of “squares.” The units for sheet resistance are by convention $\Omega/\square$, pronounced “ohms per square.” Of course, the number of “squares” is dimensionless. The layout in Fig. 2.21(a) illustrates the “square” concept by dividing the central portion into 9 segments of length W . In practical cases, the length L is often an integer multiple of the width W and $N_{\square}$ will be an integer.

For average doping levels of 10^{15} cm^{-3} to 10^{19} cm^{-3} and a typical layer thickness of $t = 1 \mu\text{m}$, the sheet resistance ranges from $50 \text{ k}\Omega/\square$ to $10 \Omega/\square$. For other IC materials, heavily phosphorus-doped (n^+) gate polysilicon films are typically 500 nm thick and have $R_{\square}$ approximately $20 \Omega/\square$. Aluminum interconnections are $1 \mu\text{m}$ in thickness and have a sheet resistance of around $30 \text{ m}\Omega/\square$ —far lower than for the polysilicon or implanted silicon layers.

Why is sheet resistance such a useful parameter to the IC designer? The answer is that the parameters of the fabrication process such as the resistor thickness, doping concentration, and mobility are all “givens” for an IC designer. The layout shown in Fig. 2.21(a) is all that is under the IC designer’s control in a particular process technology.

Thus far, we have neglected the contributions of the contact regions at the ends of the resistor to the total resistance. The details of how the current spreads out from the metal-silicon contact determine the resistance for a particular contact layout. We will neglect any additional resistance due to the metal-silicon junction not being an ideal ohmic contact. For the “dogbone” style contact shown in Fig. 2.22(a), each contact region adds 0.65 squares. It is convenient to express the additional resistances in units of squares, so that they can be counted into the total $N_{\square}$. In Fig. 2.22(b), the effect of the 2-D distribution of current density at a right angle bend is that the “corner square” contributes only 0.56 squares to the total $N_{\square}$. Accurate accounting for corners is important, since large-value resistors are typically folded in order to fit compactly into the layout. Example 2.5 illustrates the concept of sheet resistance for resistor layout.



► **Figure 2.22** Effective fractional square for (a) a “dogbone” contact area and (b) a corner. After R. C. Jaeger, *Introduction to Microelectronic Fabrication*, Addison-Wesley, 1988.

► EXAMPLE 2.5 IC Resistor Layout

(a) Given that a dose $Q_d = 10^{12} \text{ cm}^{-2}$ of arsenic is implanted into p-type silicon, with a junction depth of $x_j = 400 \text{ nm}$, estimate the sheet resistance of this n-type layer.

SOLUTION

The average arsenic concentration in the implanted layer is the thickness of the layer (in cm) divided into the dose (per cm^2)

$$N_d = \frac{Q_d}{x_j} = \frac{1 \times 10^{12}}{4 \times 10^{-5}} = 2.5 \times 10^{16} \text{ cm}^{-3}$$

From Fig. 2.8, the electron mobility for the average concentration is $\mu_n = 1000 \text{ cm}^2/\text{Vs}$. Substituting into the definition of sheet resistance in Eq. (2.67), we find that

$$R_{\square} = \frac{1}{q\mu_n N_d x_j} = \frac{1}{(1.6 \times 10^{-19})(1000)(2.5 \times 10^{16})(4 \times 10^{-5})} \cong 6 \text{ k}\Omega/\square$$

In letting $x_j = t$ in finding $R_{\square}$, we assume there is no penetration of the n-type region by the charged layer that forms at the p-n junction. This approximation is reasonable so long as the background acceptor concentration in the substrate, N_a , is much less than the donor concentration N_d in the implanted layer. We will quantify this approximation in Chapter 3. We have also made this same assumption in letting the electron concentration $n = N_d$, rather than $N_d - N_a$. For a typical substrate doping, these approximations lead to small errors.

(b) For this n-type layer, find the length L for a straight resistor (no corners) of value $R = 56 \text{ k}\Omega$. The minimum width $W = 2.5 \mu\text{m}$ and the two contact areas have the layout shown in Fig. 2.22(a)

SOLUTION

First, the total number of squares needed is

$$N_{\square} = R/R_{\square} = 56 \text{ k}\Omega / (6 \text{ k}\Omega/\square) = 9.3 \square$$

Subtracting off the contributions of the contacts, the rectangular portion of the resistor has

$$(L/W) = N_{\square} - 2(0.65) = 9.3 - 1.3 = 8$$

Therefore, the length $L = 8 \times 2.5 \mu\text{m} = 20 \mu\text{m}$.

2.6.3 Statistical Variations in IC Resistors

In order to use the IC resistor as a circuit element, its variation from lot to lot and over temperature must be understood. We will introduce a simple way to quantify the effect of manufacturing variations on the resistance, but will leave the subject of tem-

perature effects to a more advanced course. This section is applicable to other device structures since the sources of variation are fundamental to IC manufacturing.

From Eq. (2.66), the resistance is the product of the sheet resistance and the number of squares, each of which contains parameters or dimensions that vary from chip to chip:

$$R = \left(\frac{1}{qN_d\mu_n t} \right) \left(\frac{L}{W} \right) \quad (2.68)$$

Since a background in probability and statistics is *not* assumed, the variation of a particular quantity, such as the doping concentration N_d , will be expressed simply as

$$N_d = \bar{N}_d (1 \pm \epsilon_{N_d}), \quad (2.69)$$

where $\bar{N}_d$ is the average doping and ϵ_{N_d} is defined as the **normalized uncertainty**. It is convenient to use the normalized uncertainty rather than the absolute uncertainty in the particular quantity. We can determine the range of variation by multiplying the average by the normalized uncertainty. For example, a doping concentration defined by an average of $\bar{N}_d = 10^{17} \text{ cm}^{-3}$ and a normalized uncertainty of $\epsilon_{N_d} = 0.025$ means that

$$N_d = (10^{17} \text{ cm}^{-3}) (1 \pm 0.025) = 10^{17} \text{ cm}^{-3} \pm 2.5 \times 10^{15} \text{ cm}^{-3} \quad (2.70)$$

Returning to Eq. (2.68), we substitute for each term that is subject to manufacturing variations:

$$R = \left(\frac{1}{q\bar{N}_d (1 + \epsilon_{N_d}) \bar{\mu}_n (1 + \epsilon_{\mu_n}) \bar{t} (1 + \epsilon_t)} \right) \left(\frac{\bar{L} (1 + \epsilon_L)}{\bar{W} (1 + \epsilon_W)} \right) \quad (2.71)$$

It is reasonable that the average resistance should be

$$\bar{R} = \left(\frac{1}{q\bar{N}_d \bar{\mu}_n \bar{t}} \right) \left(\frac{\bar{L}}{\bar{W}} \right) \quad (2.72)$$

However, how do we predict the normalized uncertainty ϵ_R given the normalized uncertainties in the parameters and dimensions? Applying an important result from statistics, the normalized uncertainty in the resistance is the square root of the sum of the squares of the normalized uncertainties of each variable

$$\epsilon_R = \sqrt{\epsilon_{N_d}^2 + \epsilon_{\mu_n}^2 + \epsilon_t^2 + \epsilon_L^2 + \epsilon_W^2}. \quad (2.73)$$

This result is valid when the variation of one variable is independent of those of the other variables and when all variables follow the Gaussian, or normal, probability distribution. In practice, Eq. (2.73) is a good first-cut estimate for the case where the normalized uncertainties are relatively small (< 0.1). This “statistical” approach is convenient and we will use it to estimate the effects of process and layout variations—even though we are not going to verify the underlying assumptions.

We now turn to the origin of the variations in the terms in Eq. (2.71). The ion implanter will largely determine the run-to-run control on the doping concentration. Mobility variations are due to random local variations in crystal defects. Note that there is also coupling between the mobility and the doping concentration, as shown in Fig. 2.8. This coupling will be ignored in our first-cut estimate of uncertainty in Eq. (2.73). The thickness t of the resistor is subject to variations in the time and local wafer temperature during the furnace annealing process.

The lateral dimensions of the resistor are affected by random linewidth variations from the pattern transfer processes, as shown in Figure 2.23. Defining the uncertainty on one edge as $\delta/2$, the width and length of the resistor are

$$W = \bar{W} \pm \frac{\delta}{2} \pm \frac{\delta}{2} = \bar{W} \pm \delta \text{ and } L = \bar{L} \pm \frac{\delta}{2} \pm \frac{\delta}{2} = \bar{L} \pm \delta, \quad (2.74)$$

where we have added the uncertainties from each edge. The width and length can be rewritten in the form

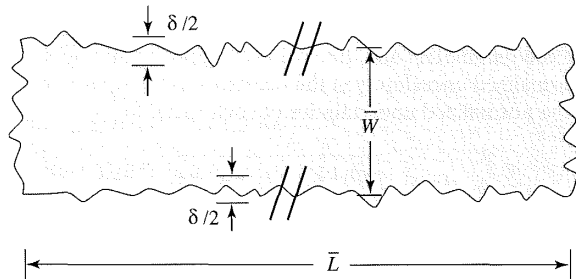
$$W = \bar{W} \left(1 \pm \frac{\delta}{\bar{W}} \right) = \bar{W} (1 \pm \epsilon_W) \text{ and} \quad (2.75)$$

$$L = \bar{L} \left(1 \pm \frac{\delta}{\bar{L}} \right) = \bar{L} (1 \pm \epsilon_L). \quad (2.76)$$

From Eqs. (2.75) and (2.76), the normalized uncertainties in the width and length are inversely proportional to their average values:

$$\epsilon_W = \delta / \bar{W} \text{ and } \epsilon_L = \delta / \bar{L}. \quad (2.77)$$

Therefore, the designer has control over the normalized uncertainties in width and length and can make trade-offs between the resistor area (the product of width and length) and its normalized uncertainty. It is common for the length to greatly



► **Figure 2.23** Highly magnified top view of the resistor showing the uncertainty $\delta/2$ in the edges due to random variations in the photolithography and etching processes. Note that the sketch only shows the ends of the relatively long and narrow resistor.

exceed the width, in which case the normalized uncertainty in the length becomes negligibly small, according to Eq. (2.77)

What are the magnitudes of the process uncertainties that enter into Eq. (2.73)? To answer this question, we must first specify the collection of samples used to evaluate the average and variation. The variation is smallest for structures that are nearby on the same chip, since they are subjected to nearly identical processes. If an entire wafer is considered, then the spread in parameters and dimensions increases. The uncertainties increase further when a wafer lot is evaluated and are largest when measurements are made on samples from many lots. Finally, the IC process equipment and the organization and discipline of the workforce are major factors in setting the baseline levels of process uncertainties.

The statistics on process variations are highly proprietary for each manufacturer, needless to say. However, some rough numbers are useful for gaining an appreciation for the limits of IC technology. Nominally identical, adjacent ion-implanted resistors can have a normalized uncertainty as low as 10^{-3} . With careful attention to the layout, resistor ratios can have a similar level of precision. A further benefit of adjacent resistors is that they are at nearly the same temperature. As a result, their resistances track with changing ambient temperature. When the set of resistors is expanded to include samples from a wafer, a lot, or a series of lots, the normalized uncertainty increases by a factor of 10 to 100—not including the significant variation over temperature. As a result, the precision is too poor to allow a designer to use an IC resistor for circuits where its absolute value must be tightly controlled. The absence of precision integrated resistors has been a major motivator of new approaches to circuit design.

► EXAMPLE 2.6 IC Resistor Process Variations

An ion-implanted resistor with the cross section shown in Fig. 2.21 is designed to have a resistance $R = 20 \text{ k}\Omega$. The process parameters and their normalized uncertainties over a series of lots are: $N_d = (4 \times 10^{17} \text{ cm}^{-3})(1 \pm 0.04)$, $\mu_n = (450 \text{ cm}^2/\text{Vs})(1 \pm 0.02)$, $t = 550 \text{ nm}(1 \pm 0.035)$.

(a) Find the average sheet resistance and its normalized uncertainty.

SOLUTION

Substituting the average values for the parameters into the definition of sheet resistance in Eq. (2.67), we find that

$$\bar{R}_{\square} = \frac{1}{(1.6 \times 10^{-19} \text{ C})(4 \times 10^{17} \text{ cm}^{-3})(450 \text{ cm}^2/\text{Vs})(0.55 \times 10^{-4} \text{ cm})} = 630 \Omega/\square.$$

The normalized uncertainty in the sheet resistance is given by

$$\epsilon_{R_{\square}} = \sqrt{\epsilon_{N_d}^2 + \epsilon_{\mu_n}^2 + \epsilon_t^2} = \sqrt{(0.04)^2 + (0.02)^2 + (0.035)^2} = 0.06$$

(b) Over several wafer lots, the linewidth edge variation is $\delta = 0.15 \mu\text{m}$. For nominal resistor widths of $W = 2 \mu\text{m}$ and $W = 4 \mu\text{m}$, find the normalized uncertainty for the resistor.

SOLUTION

The number of squares required is:

$$N_{\square} = \frac{R}{R_{\square}} = \frac{20 \text{ k}\Omega}{630 \text{ }\Omega/\text{sq.}} = 31.75$$

Due to the large number of squares, the uncertainty in the length can be neglected. According to Eq. (2.77), the normalized uncertainty in the resistor width is

$$\epsilon_w = \frac{0.15 \text{ }\mu\text{m}}{W} = 0.075 \text{ (} W = 2 \text{ }\mu\text{m)} \text{ and } \epsilon_w = 0.038 \text{ (} W = 4 \text{ }\mu\text{m)}$$

Therefore, the normalized uncertainties of the resistor for the two widths are

$$\epsilon_R = \sqrt{\epsilon_{R_{\square}}^2 + \epsilon_w^2} = \sqrt{(0.06)^2 + (0.075)^2} = 0.096 \text{ (} W = 2 \text{ }\mu\text{m)} \text{ and}$$

$$\epsilon_R = \sqrt{\epsilon_{R_{\square}}^2 + \epsilon_w^2} = \sqrt{(0.06)^2 + (0.038)^2} = 0.0625 \text{ (} W = 4 \text{ }\mu\text{m)}$$

The nominally 20 k Ω IC resistor has an absolute uncertainty of nearly 2 k Ω for the narrow width layout and about 1.25 k Ω for the wider layout. The penalty for the wider layout is an increase in area by a factor of 4 since the length must also be doubled to keep the number of squares at 31.75.

For large resistor values, it is desirable to fold the resistor in order to make a more compact IC structure. We conclude the chapter with a design example which explores the geometry involved in designing for a square layout.

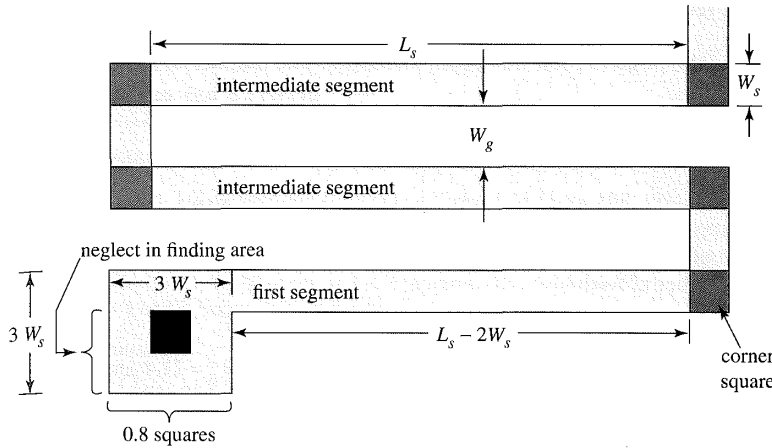
**DESIGN EXAMPLE 2.7** Folded IC Resistor Layout

Given an n-type layer with a sheet resistance $R_{\square} = 1 \text{ k}\Omega/\square$, we want to lay out a folded resistor with $R = 75 \text{ k}\Omega$. The geometric design rules for the n-type region specify that the minimum width is $W_s = 2 \text{ }\mu\text{m}$ and that the minimum gap between n-type regions is $W_g = 3 \text{ }\mu\text{m}$, as defined in the sample layout in Fig. Ex2.7. The contact regions each contribute $0.8\square$. In order to make the resistor easy to pack into the IC layout, we want to fold the resistor.

What is the segment length L_s (defined in the sample layout) that results in a nearly square resistor layout? How much area does the resistor require? You may neglect the portions of the contact areas outside the folded segments in finding the resistor area.

SOLUTION

Assuming that the resistor consists of N_s segments, the total number of squares $N_{\square}$ is



► **Figure Ex2.7** Layout of a folded IC resistor with offset contacts.

$$N_{\square} = 2(0.8) + 2\left(\frac{L_s - 2W_s}{W_s}\right) + (N_s - 2)\left(\frac{L_s}{W_s}\right) + (N_s - 1)\left(2(0.56) + \frac{W_g}{W_s}\right)$$

in which the first two terms are for the first and last segments (with the contacts), the third term is for the intermediate segments of length L_s , and the final term is for the folds (with the two corner squares contributing 0.56 squares each). Using the minimum width $W_s = 2 \mu\text{m}$ and the minimum gap $W_g = 3 \mu\text{m}$, we find

$$N_{\square} = 1.6 + L_s - 4 + \left(\frac{(N_s - 2)}{2}\right)L_s + (N_s - 1)(2.62) = \left(\frac{L_s}{2} + 2.62\right)N_s - 5.02$$

In order to achieve a square layout, Fig. Ex2.7 indicates that

$$L_s + 2W_s = (N_s - 1)(W_s + W_g) + W_s = (N_s - 1)(5 \mu\text{m}) + 2 \mu\text{m}$$

in which the left hand side is the total length of one segment and the right hand side is the length of the side with the folds. Solving for the number of segments, we find that

$$N_s = \frac{L_s + 2W_s + W_g}{W_s + W_g} = \frac{L_s + 7}{5}$$

The total number of squares needed is $N_{\square} = R/R_{\square} = 75 \text{ k}\Omega / (1 \text{ k}\Omega/\square) = 75$. Substituting for N_s in the expression for the total number of squares, we find a quadratic equation for the segment length

$$75 = \left[\frac{L_s}{2} + 2.62\right]\left(\frac{L_s + 7}{5}\right) - 5.02, \text{ which simplifies to}$$

$$L_s^2 + (12.24) L_s - 763.32 = 0$$

Taking the positive root, we find $L_s = 22.2 \mu\text{m}$. The number of segments needed is

$$N_s = \frac{L_s + 7}{5} = \frac{22.2 + 7}{5} = 5.84$$

One approach is to use an integer number of folds so we round up to $N_s = 6$. The length of each segment must be modified in order to maintain $N_\square = 75$ using the new number of folds

$$N_\square = 75 = \left(\frac{L_s}{2} + 2.62 \right) N_s - 5.02 = \left(\frac{L_s}{2} + 2.62 \right) 6 - 5.02 \rightarrow L_s = 21.4 \mu\text{m}$$

The final layout of the folded resistor with $L_s = 21.4 \mu\text{m}$ is not a perfect square, since the segments are $25.4 \mu\text{m}$ long (counting the folds) and the edge with the 6 folds is $27 \mu\text{m}$ long. An alternative to forcing N_s to be an integer would be to adjust the lengths of the first or the last segments, in order to make the overall layout closer to a square. In practice, the integer number of folds is close enough to a square layout. The area of the resistor is

$$A = (L_s + 2W_s) [(N_s - 1)(W_s + W_g) + W_s] = (21.4 + 4)(5 \cdot 5 + 2) = 686 \mu\text{m}^2$$

► SUMMARY

This chapter has provided a foundation in two subjects that are essential for integrated device and circuit modeling: the physics of holes and electrons in silicon and the technology of IC fabrication. Key concepts that you should understand are:

- ◆ The bond model and its description of an electron and a hole in silicon.
- ◆ The concept of thermal equilibrium and the mass-action law.
- ◆ Hole and electron concentrations for silicon with donors, with acceptors, or with both donors and acceptors.
- ◆ Drift velocity and drift current density.
- ◆ Diffusive transport and diffusion current density.
- ◆ The total electron and hole current densities.
- ◆ Lithography and pattern transfer.
- ◆ Masked ion implantation of impurities.
- ◆ Cross sections through device structures from the layout and the sequence of fabrication steps (process).
- ◆ Resistivity and sheet resistance.
- ◆ Ion-implanted IC resistors and their layout, including equivalent squares for contact regions and corners.

- ◆ Normalized uncertainty and its calculation for a function of several variables, given their uncertainties.

► FURTHER READING

Semiconductor Physics

1. C. G. Fonstad, *Microelectronic Devices and Circuits*. McGraw-Hill, 1994, Chapters 1 and 2. About the same level.
2. R. F. Pierret, *Semiconductor Fundamentals*, 2nd Ed. Modular Series on Solid State Devices, Vol. I, G. W. Neudeck and R. F. Pierret, eds., Addison-Wesley, 1988, Chapters 1 and 3. Somewhat more advanced.

IC Technology

1. W. Maly, *Atlas of VLSI Technologies*, Addison-Wesley, 1986. Several step-by-step process sequences provide a good intuitive understanding of the connection between layout, process, and device cross sections.
2. R. C. Jaeger, *Introduction to Microelectronic Fabrication*, Modular Series on Solid State Devices, Vol. V, G. W. Neudeck and R. F. Pierret, eds., Addison-Wesley, 1988, Chapters 1 and 2. Early chapters provide a good overview of IC fabrication.

► PROBLEMS

Note: Assume that the sample temperature is 27 °C , unless otherwise noted in the problem statement.

EXERCISES

- E2.1** A silicon wafer is doped with donors at a concentration of $N_d = 10^{15} \text{ cm}^{-3}$.
- (a) What is the electron concentration n_o (cm^{-3}) at room temperature?
 - (b) What is the hole concentration p_o (cm^{-3}) at room temperature?
- E2.2** A silicon wafer is doped with acceptors at a concentration of $N_a = 10^{14} \text{ cm}^{-3}$.
- (a) What is the electron concentration n_o (cm^{-3}) at room temperature?
 - (b) What is the hole concentration p_o (cm^{-3}) at room temperature?
- E2.3** Given that a silicon wafer is doped with acceptors with $N_a = 10^{14} \text{ cm}^{-3}$. We add a donor concentration of $N_d = 7.5 \times 10^{13} \text{ cm}^{-3}$ to a particular region.
- (a) What type is the region (n or p)?
 - (b) What is the electron concentration n_o (cm^{-3}) in the region?
 - (c) What is the hole concentration p_o (cm^{-3}) in the region?
- E2.4** Given that a silicon wafer is doped with donors with $N_d = 5 \times 10^{17} \text{ cm}^{-3}$. We add an acceptor concentration of $N_a = 5.5 \times 10^{17} \text{ cm}^{-3}$ to a particular region.
- (a) Which carrier is the majority carrier?
 - (b) What is the hole concentration p_o (cm^{-3}) in the region?
 - (c) What is the electron concentration n_o (cm^{-3}) in the region?
- E2.5** For a wafer doped with acceptors at $N_a = 10^{14} \text{ cm}^{-3}$, what concentration of donors must be added in order to reduce the hole concentration by 90%?

E2.6 In a subatomic particle detector, extremely low doped silicon substrates are needed. What donor concentration is required in order to have an electron concentration $n_o = 8 n_i$? You will need to account for thermal generation for this problem.

E2.7 In modern VLSI processes, the dimensions of device structures are approaching $1 \mu\text{m}^3$. For a donor concentration $N_d = 5 \times 10^{14} \text{ cm}^{-3}$, find the number of electrons and the number of holes in this volume. *Comment:* your answer will show one reason why there is a tendency for doping concentrations to increase as device dimensions are scaled into the submicron region.

E2.8 In a silicon region with donor concentration $N_d = 10^{16} \text{ cm}^{-3}$, we apply an electric field in the $+x$ direction with a magnitude of 10^3 V/cm .

- What is the electron drift velocity (magnitude and sign)?
- What is the electron drift current density (magnitude and sign)?
- How long is required for an electron to drift $1 \mu\text{m}$, on average?
- How many collisions occur while it is drifting? You can assume that the collision time for the drifting electron is $\tau_c = 0.1 \text{ ps}$.

E2.9 Repeat E2.8 with a donor concentration of $N_d = 10^{15} \text{ cm}^{-3}$.

E2.10 In a silicon region with acceptor concentration $N_a = 10^{18} \text{ cm}^{-3}$, we apply an electric field in the $+x$ direction with a magnitude of $2 \times 10^3 \text{ V/cm}$.

- What is the hole drift velocity (magnitude and sign)?
- What is the hole drift current density (magnitude and sign)?
- How long is required for a hole to drift $1 \mu\text{m}$, on average?
- How many collisions occur while it is drifting? You can assume that the collision time for the drifting hole is $\tau_c = 0.05 \text{ ps}$.

E2.11 A silicon region is doped with acceptors at $N_a = 10^{15} \text{ cm}^{-3}$ and with donors at $N_d = 8 \times 10^{14} \text{ cm}^{-3}$.

- What type is the silicon (n or p)?
- What electric field must be applied for the magnitude of the drift velocity of the majority carriers to be $4 \times 10^6 \text{ cm/s}$?
- What is the drift current density of the majority carriers for the case of the electric field in part (b)?

E2.12 By optical means, we set up a minority hole concentration gradient across a $2 \mu\text{m}$ -wide silicon region that is given by:

$$p(x) = (10^{18} \text{ cm}^{-4}) x$$

where x is the coordinate in the direction of the concentration gradient. The donor concentration in the region is $N_d = 10^{16} \text{ cm}^{-3}$.

- Find the hole diffusion current density in the region.
- How long does a hole take to diffuse across the $2 \mu\text{m}$ -wide region?

E2.13 The minority hole concentration gradient across a $2.5 \mu\text{m}$ -wide silicon region is given by:

$$p(x) = (5 \times 10^{18} \text{ cm}^{-4}) (2.5 \mu\text{m} - x)$$

where x is the coordinate in the direction of the concentration gradient. The donor concentration in the region is $N_d = 5 \times 10^{17} \text{ cm}^{-3}$.

- Find the hole diffusion current density in the region.
- How long does a hole take to diffuse across the $2.5 \text{ }\mu\text{m}$ -wide region?

E2.14 In a silicon sample, we establish a minority electron concentration gradient over the region $x \geq 0$ that is given by

$$n(x) = (10^{15} \text{ cm}^{-3}) e^{-x/(2 \text{ }\mu\text{m})}$$

The acceptor concentration is $N_a = 10^{17} \text{ cm}^{-3}$.

- Find the magnitude and sign of the electron diffusion current density at $x = 0$.
- Plot the electron diffusion current density over the interval $0 < x < 10 \text{ }\mu\text{m}$.

E2.15 The gradient in minority hole concentration in a silicon sample over the region $x \geq 0$ is given by

$$p(x) = (10^{13} \text{ cm}^{-3}) (1 - e^{-x/(1 \text{ }\mu\text{m})})$$

The donor concentration is $N_d = 5 \times 10^{15} \text{ cm}^{-3}$.

- What is the hole concentration at $x = 0$?
- What is the hole diffusion current density at $x = 0$? Provide an explanation for the apparent discrepancy between parts (a) and (b).
- Plot the hole diffusion current density over the interval $0 \leq x \leq 5 \text{ }\mu\text{m}$.

E2.16 By continuous improvements in fabrication technology, three generations of a microprocessor IC have areas that are given in the table below. The cost of processing a 20-cm-diameter wafer increases with each generation, due to the need for more expensive process equipment.

Technology	Chip Area	Wafer Cost
1 st Generation	$2 \times 2 \text{ cm}^2$	\$1750
2 nd Generation	$1.5 \times 1.5 \text{ cm}^2$	\$2000
3 rd Generation	$1 \times 1 \text{ cm}^2$	\$2250

The yield of good chips for each generation is 75%.

- Find the number of good chips on a 20-cm wafer for each generation. Do not consider the complication of dice that intersect the edge of the wafer—an estimate of the chip count per wafer is adequate.
- What is the manufacturing cost per good microprocessor chip for each generation?

E2.17 For an implant dose $Q_a = 10^{15} \text{ cm}^{-2}$ of acceptors and a post-anneal thickness of $t = 0.5 \text{ }\mu\text{m}$, what is the acceptor concentration in the layer?

E2.18 Given a silicon wafer with an acceptor concentration of $N_a = 10^{16} \text{ cm}^{-3}$. We would like to use ion implantation to dope a surface region so that the average acceptor concentration is $N_a = 10^{17} \text{ cm}^{-3}$. If the post-anneal thickness of the region is $t = 0.1 \text{ }\mu\text{m}$, what dose is needed?

E2.19 Starting with a silicon wafer with a donor concentration of $N_d = 10^{17} \text{ cm}^{-3}$, we would like to use ion implantation to convert a surface region to p-type, with an average acceptor concentration $N_a = 3 \times 10^{17} \text{ cm}^{-3}$. What implant dose is needed, if the post-anneal thickness of the implanted region is $0.25 \text{ }\mu\text{m}$.

E2.20 The five-step process for a two-layer structure made from silicon dioxide and polysilicon is as follows:

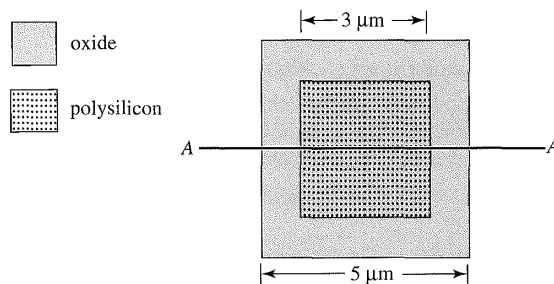
1. Grow $0.5 \text{ }\mu\text{m}$ of silicon dioxide (SiO_2)
2. Pattern using the oxide mask.
3. Grow $500 \text{ }\text{\AA}$ of SiO_2 .
4. Deposit $1 \text{ }\mu\text{m}$ of polysilicon.
5. Pattern using the polysilicon mask.

The CAD layout for this structure is shown in Fig. E2.20.

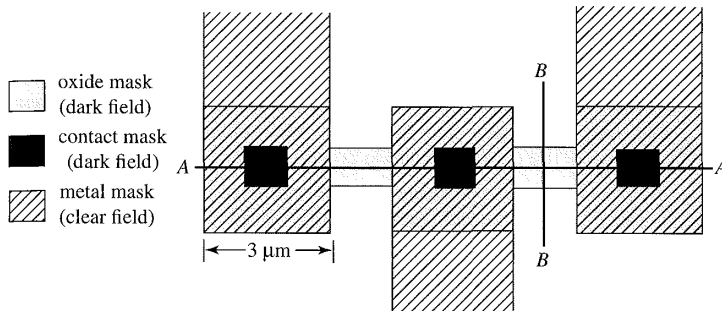
- (a) Sketch the cross section *A-A* from the process sequence, assuming that the oxide mask is dark field and the polysilicon mask is clear field.
- (b) Repeat part (a), with the oxide mask clear field and the polysilicon mask dark field.
- (c) Repeat part (a), with the oxide mask dark field and the polysilicon mask dark field.
- (d) Repeat part (a), with the oxide mask clear field and the polysilicon mask clear field.

E2.21 The layout in Fig. E2.21 is for the ion-implanted resistor process outlined in Fig. 2.18.

- (a) Sketch the cross section *A-A* from the layout in Fig. E2.21.
- (b) Sketch the cross section *B-B* from the layout in Fig. E2.21.



► **Figure E2.20**



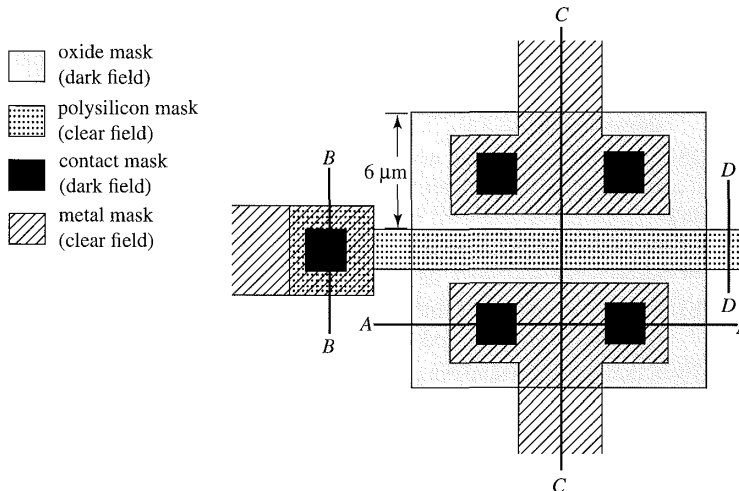
► **Figure E2.21**

E2.22 The layout in Fig. E.2.22 is for the simple MOS transistor process described in Fig. 2.19 (b).

- Sketch the cross section along $A-A$.
- Sketch the cross section along $B-B$.
- Sketch the cross section along $C-C$.
- Sketch the cross section along $D-D$.

E2.23 The parameters for an ion-implanted region are: arsenic dose $Q_d = 10^{12} \text{ cm}^{-2}$, thickness after anneal $t = 1.1 \text{ μm}$. The background doping of the p-type substrate is 10^{14} cm^{-2} .

- What is the average arsenic concentration in the ion-implanted layer?
- What is the sheet resistance in units of $\Omega/\square$.



► **Figure E2.22**

E2.24 After discovering that boron was implanted at a dose of $Q_a = 10^{11} \text{ cm}^{-2}$ instead of phosphorus, you try to salvage the lot by compensating the layer back to n-type by means of implanting a heavier dose of phosphorus.

- Assuming that the layer thicknesses for the phosphorus and boron are the same at $t = 1 \text{ } \mu\text{m}$ after annealing, what phosphorus dose is necessary to achieve the desired electron concentration?
- What is the sheet resistance for the compensated layer? Compare your answer to what the sheet resistance would have been if the original mistake had not been made.
- From your answer in part (b), what phosphorus dose is needed to achieve the original sheet resistance?

E2.25 Given a $1 \text{ } \mu\text{m}$ -thick implanted layer and a maximum average dopant concentration of $1 \times 10^{20} \text{ cm}^{-3}$ in this layer.

- What implant dose is needed to reach the maximum concentration? Neglect the background doping in the substrate.
- For an n-type layer, what is the minimum sheet resistance? Note that the mobility degrades with increasing doping concentration, as shown in Fig. 2.8.
- For a p-type layer, what is the minimum sheet resistance? Note that the mobility degrades with increasing doping concentration, as shown in Fig. 2.8.

E2.26 Given a sheet resistance of $100 \text{ } \Omega/\square$ in a $1 \text{ } \mu\text{m}$ -thick, n-type implanted layer.

- What is the average donor concentration in the layer? Neglect the background doping in the substrate.
- Sketch the layout for a $2.5 \text{ k}\Omega$ resistor, using the “dogbone” style contacts shown in Fig. 2.22(a).

E2.27 Given a sheet resistance of $250 \text{ } \Omega/\square$ in a $1 \text{ } \mu\text{m}$ -thick, p-type implanted layer.

- What is the average acceptor concentration in the layer? Neglect the background doping in the substrate.
- Sketch the layout for a $2 \text{ k}\Omega$ resistor, using the “dogbone” style contacts shown in Fig. 2.22(a).

E2.28 Given that the normalized uncertainties are: $\epsilon_{N_d} = 0.025$ in the donor concentration and $\epsilon_t = 0.02$ in the layer thickness. The uncertainty in the mobility can be neglected.

- For an average donor concentration of 10^{16} cm^{-3} , find the variation in N_d .
- For an average layer thickness of 500 nm , find the average sheet resistance and its normalized and absolute uncertainties.

E2.29 Given that the uncertainty in the lateral dimension of an ion-implanted resistor is $\delta = 500 \text{ } \text{\AA}$ and that the sheet resistance is $100 \text{ } \Omega/\square$.

- Neglecting the uncertainty in the sheet resistance, what are the normalized uncertainty and the variation in a $5 \text{ k}\Omega$ resistor for a resistor width $W = 2 \text{ } \mu\text{m}$.
- How wide must the resistor be made in order to have the normalized uncertainty less than 0.01 (1%)?

E2.30 For a doped polysilicon film, the sheet resistance is $25 \Omega/\square$ for a $5000 \text{ \AA}$ -thick film. If the thickness of the film is controlled to $\pm 2.5\%$ over the wafer, what is the variation in the resistance of a $20 \text{ k}\Omega$ resistor?

PROBLEMS

P2.1 A sample of silicon is doped with phosphorus with a concentration of 10^{15} cm^{-3} .

- What is the average volume occupied by an electron in this sample, in μm^3 ?
- What is the average volume occupied by a holes in this sample, in μm^3 ?

P2.2 A region of a silicon wafer is doped with three impurities: arsenic: 10^{16} cm^{-3} , boron: $1.15 \times 10^{16} \text{ cm}^{-3}$, phosphorus: $2.5 \times 10^{15} \text{ cm}^{-3}$. Find the electron and hole concentrations n_o and p_o .

P2.3 A sample of silicon is doped with $N_d = 1.1 \times 10^{13} \text{ cm}^{-3}$ and $N_a = 1 \times 10^{13} \text{ cm}^{-3}$.

- Find the electron and hole concentrations n_o and p_o at room temperature.
- Show that the charge density $\rho = 0$ (exactly).

P2.4 The intrinsic concentration in silicon n_i has a strong temperature dependence, as shown by these measurements over the temperature range for automotive electronic applications:

$T (^{\circ}\text{C})$	$n_i (\text{cm}^{-3})$
-55	2.24×10^6
0	1.42×10^9
70	2.82×10^{11}
125	5.06×10^{12}

A course in semiconductor physics shows that the functional dependence of the intrinsic concentration is

$$n_i = A \cdot T^{3/2} e^{-6608.7/T},$$

where T is in Kelvin.

- Find the constant A by fitting the function to the measurements above.
- How rapidly the intrinsic concentration varies near room temperature is important for device modeling. Find the temperature increase (in degrees Centigrade) needed to double n_i and the temperature decrease needed to halve n_i .

P2.5 A sample of silicon is doped with $N_d = 10^{13} \text{ cm}^{-3}$. As the temperature increases, the contribution of thermal generation to the electron and hole concentrations eventually becomes significant. Using the result from P2.4(a) for the intrinsic concentration as a function of temperature,

- Plot $n_o(T)$ over the temperature range $-55^{\circ}\text{C} < T < 125^{\circ}\text{C}$ on a log-linear plot.
- What is the maximum temperature for which $p_o/n_o < 10^{-3}$?

P2.6 A sample of silicon is doped with $N_a = 2 \times 10^{14} \text{ cm}^{-3}$ and $N_d = 9 \times 10^{13} \text{ cm}^{-3}$. Using the result from P2.4(a) for the intrinsic concentration as a function of temperature,

- Plot $p_o(T)$ over the temperature range $-55^\circ \text{ C} < T < 125^\circ \text{ C}$ on a log-linear plot.
- What is the maximum temperature for which $n_o/p_o < 5 \times 10^{-4}$?

P2.7 Figure P2.7 is a top view of a silicon region which has metal electrical contacts on two sides, across which a voltage $V_A = 2 \text{ V}$ is applied. The silicon layer is doped with arsenic at a concentration of 10^{13} cm^{-3} and is $2 \mu\text{m}$ thick and you can assume room temperature. You should use the appropriate mobility for the given dopant concentration.

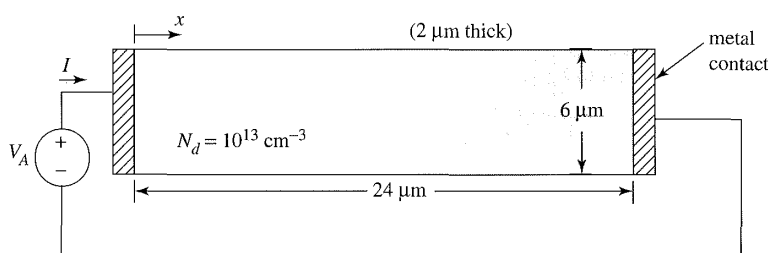
- Find the value of the electric field E_x in the silicon.
- Find the value of the drift velocity of electrons and of holes in the silicon. How long does it take an electron, on average, to drift across the silicon region.
- Find the drift current densities for electrons and for holes. What is the percentage error in the total current I if the minority carrier drift current density is neglected?
- What is the value of the resistance R in ohms?

P2.8 If the arsenic concentration is increased to 10^{18} cm^{-3} , repeat P2.7.

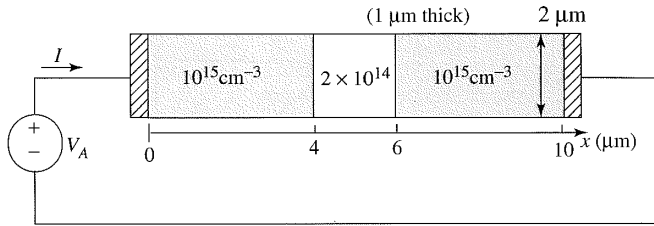
P2.9 Plot the maximum current $I_{\max}$ across the silicon region in Fig. P2.7 as a function of doping concentration from $N_d = 10^{13} \text{ cm}^{-3}$ to $N_d = 10^{18} \text{ cm}^{-3}$, using a log-log scale. You can assume that the saturation velocity of electrons is $v_{\text{sat}} = 10^7 \text{ cm/s}$ and should include the dependence of mobility on doping concentration. Plot the applied voltage $V_{A,\max}$ at which $I_{\max}$ is reached as a function of doping concentration.

P2.10 Figure P2.10 is a top view of a $1 \mu\text{m}$ -thick silicon region with variable acceptor concentration, across which a voltage V_A is applied through metal electrical contacts. Assume that the hole saturation velocity is $v_{\text{sat}} = 10^7 \text{ cm/s}$.

- For an applied voltage $V_A = 1 \text{ V}$, plot the electric field $E(x)$. *Hint:* you will find it helpful to consider the silicon region as three series resistors.
- What is the maximum current $I_{\max}$ through this structure?
- What voltage $V_{\max}$ must be applied to reach this maximum current?
- Plot the electric field $E(x)$ corresponding to $V_{\max}$.



► Figure P2.7



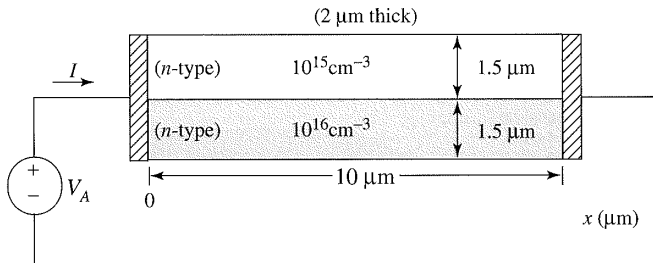
► Figure P2.10

P2.11 Figure P2.11 is a top view of a 1 μm -thick silicon region with variable donor concentration. A voltage V_A is applied through metal electrical contacts.

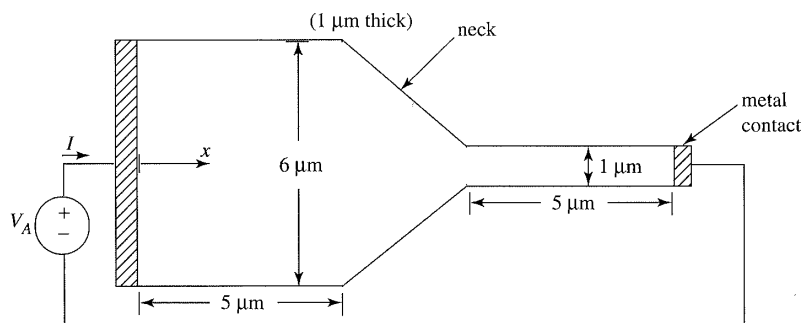
- For an applied voltage $V_A = 1\text{ V}$, what is the current I ? You may find it helpful to view the silicon region as two resistors in parallel.
- Plot the current as a function of voltage over the range $-20\text{ V} < V_A < 20\text{ V}$. You can assume that the saturation velocity of electrons $v_{sat} = 10^7\text{ cm/s}$ and should include the variation of mobility with donor concentration.

P2.12 The top view of a 2 μm -thick silicon region with a varying width is shown in Fig. P2.12. The donor concentration is $N_d = 10^{17}\text{ cm}^{-3}$ throughout. You can assume that 25% of the applied voltage is dropped across the “neck,” as long as the saturation velocity is not reached in any part of the region.

- Find the ratio of the current density in the wide region to that in the narrow region, for the case where $V_A = 2\text{ V}$.
- Sketch the electric field $E(x)$ along the center line of the structure, for the case when $V_A = 2\text{ V}$. The numerical value in the regions of constant width should be accurate. It will be helpful to view the silicon region as three resistors in series, with the voltage drop across the “neck” region given as 25% of V_A .
- What is the maximum current I_{max} through this structure? What is the applied voltage when I_{max} is reached?



► Figure P2.11



► Figure P2.12

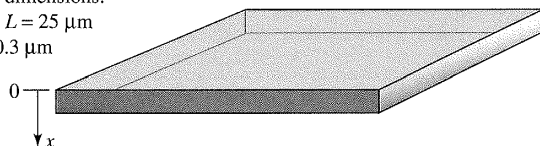
P2.13 In the slab of silicon region in Fig. P2.13, the minority carrier concentration is maintained at its thermal equilibrium value at the bottom surface ($x = 0.3\text{ }\mu\text{m}$). The silicon is doped with acceptors at $N_a = 10^{16}\text{ cm}^{-3}$.

- If the minority carrier concentration at the top surface is 1/10 the majority carrier concentration and we assume a linear gradient from top to bottom, what is the diffusion current density in the slab?
- How long does it take for a minority carrier to diffuse across the slab, on average?

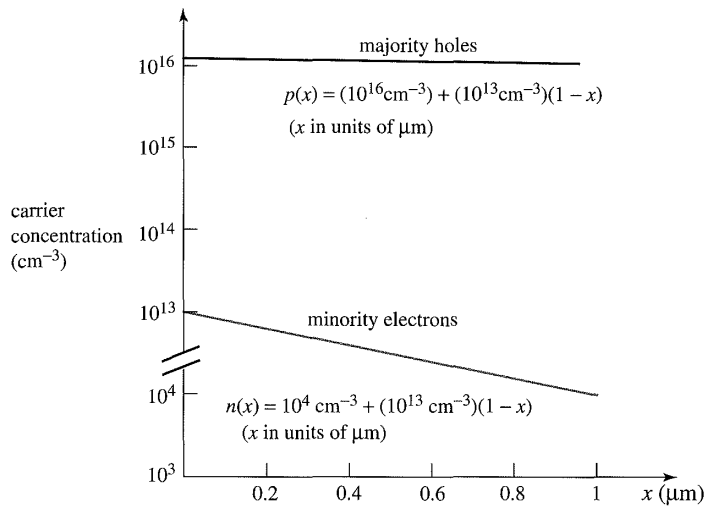
P2.14 When a gradient in minority carrier concentration is set up in a semiconductor region, an identical gradient in the majority carrier concentration is quickly established, in order to keep the silicon region electrically neutral. This situation is shown in the carrier concentration distributions in Fig. P2.14.

- Find the minority electron diffusion current density.
- Find the majority hole diffusion current density.
- Since the silicon region is open-circuited, the total current density must be zero. Find the required drift current density.
- Assuming that the contribution of the minority carriers to the drift current is negligible (which it is), what electric field is required in the sample to generate the drift current density needed from part (c).

slab dimensions:
 $W = L = 25\text{ }\mu\text{m}$
 $t = 0.3\text{ }\mu\text{m}$



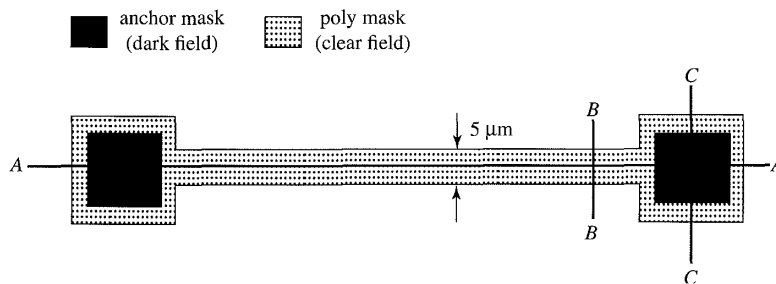
► Figure P2.13



► Figure P2.14

P2.15 The process sequence and layout for making a simple polysilicon micromechanical bridge are listed below and in Fig. P2.15. Using IC-based processes, we can make structures for sensing non-electrical inputs, such as acceleration.

1. Deposit $2 \mu\text{m}$ of CVD SiO_2 .
2. Pattern with the anchor mask.
3. Implant phosphorus and anneal.
3. Deposit $2 \mu\text{m}$ of phosphorus-doped polysilicon and anneal.
4. Pattern with the poly mask.
5. Etch for 4 minutes in hydrofluoric acid (SiO_2 etch rate: $1 \mu\text{m}$ per minute); rinse wafer in water and dry.



► Figure P2.15

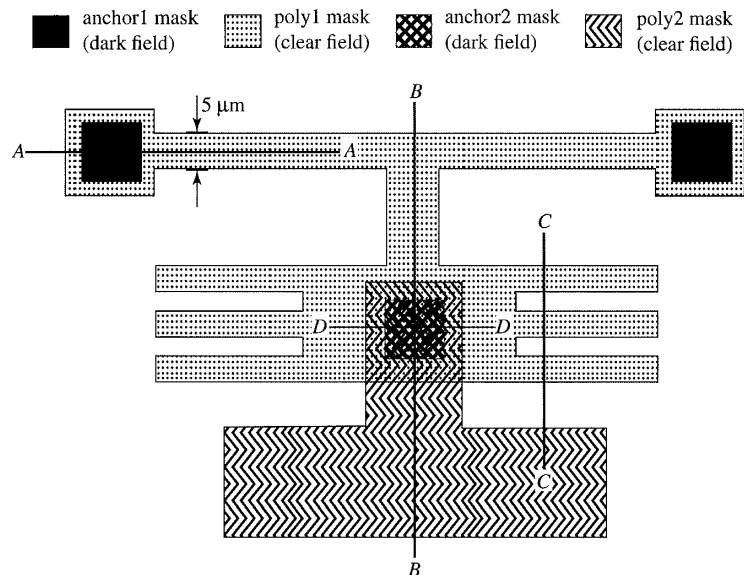
- (a) Draw the cross sections *A-A* after step 4 and after completion of the process.
- (b) Draw the cross section *B-B* after step 5 is half completed (2 minutes of HF etching.)
- (c) Draw the cross section *C-C* after completion of the process.

P2.16 This problem considers the fabrication of a two-level micromechanical structure, which is accomplished by extending the process in P2.15.

Steps 1-4: same as those in P2.15, with the masks renamed as *anchor1* and *poly1*

- 5. Deposit 1 μm of SiO_2 .
- 6. Pattern the oxide with the anchor2 mask.
- 7. Deposit 1 μm of phosphorus-doped polysilicon and anneal.
- 8. Pattern the second polysilicon layer with the poly2 mask.
- 9. Etch for 12 minutes in hydrofluoric acid (SiO_2 etch rate: 1 μm per minute); rinse wafer in water and dry.
- (a) Draw the cross section *A-A* after step 8.
- (b) Draw the cross section *B-B* after completion of the process.
- (c) Draw the cross section *C-C* after completion of the process.
- (d) Draw the cross section *D-D* after completion of the process.

P2.17 Texturing the silicon surface is a way to increase the surface area of a capacitor; the following process and layout are one way to achieve this goal.

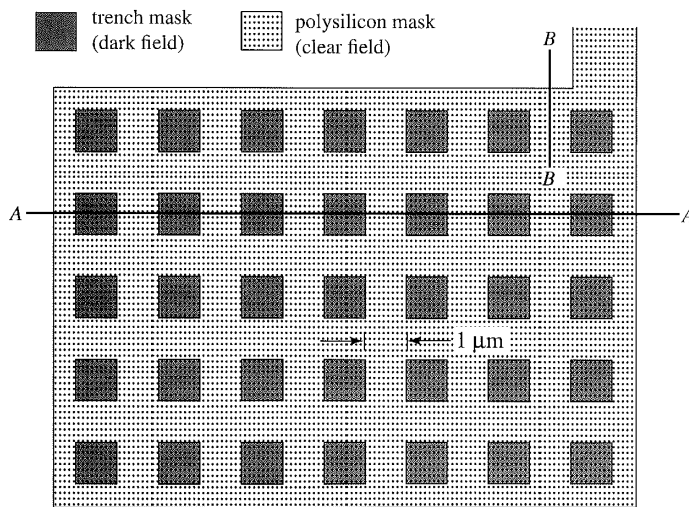


► **Figure P2.16**

1. Etch $5\text{ }\mu\text{m}$ deep into the silicon substrate using the trench mask*.
2. After stripping the photoresist and cleaning the wafer, grow $500\text{ }\text{\AA}$ of thermal SiO_2 .
3. Deposit $7500\text{ }\text{\AA}$ of phosphorus-doped polysilicon.
4. Pattern using the polysilicon mask.
 - (a) Draw the cross section $A-A$ after step 2.
 - (b) Draw the cross section $A-A$ after step 3.
 - (c) Draw the cross section $B-B$ after completion of the process.

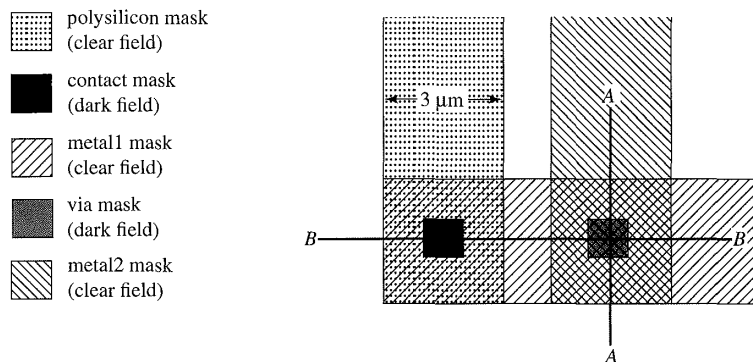
P2.18 A simple double metallization process is as follows; the mask layout is shown in Fig. P2.18.

1. Grow $0.5\text{ }\mu\text{m}$ of thermal SiO_2 .
2. Deposit $0.5\text{ }\mu\text{m}$ of phosphorus doped polysilicon and pattern using polysilicon mask.
3. Deposit $0.5\text{ }\mu\text{m}$ of CVD SiO_2 and pattern using contact mask.
4. Deposit $1\text{ }\mu\text{m}$ of sputtered Al and pattern using metal1 mask.
5. Deposit $0.5\text{ }\mu\text{m}$ of CVD SiO_2 and pattern using via mask.
6. Deposit $1\text{ }\mu\text{m}$ of sputtered Al and pattern using metal2 mask.
 - (a) Accurately sketch the cross section $A-A$, assuming that the deposited CVD films conformally coat the surface.



► **Figure P2.17**

*A reactive-ion etch is used for the trench etch into the silicon substrate, which is optimized to yield vertical walls in silicon with negligible mask undercutting.



► Figure P2.18

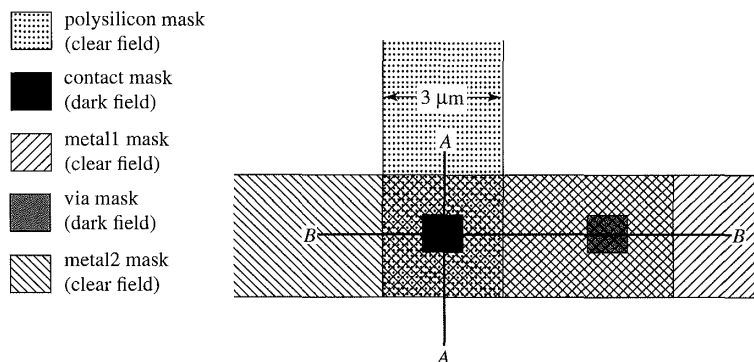
- (b) Accurately sketch the cross section *B-B*, assuming that the deposited CVD films conformally coat the surface.

P2.19 Using the double metal process from P2.18, consider the layout in Fig. P2.19.

- (a) Accurately sketch the cross section *A-A*, assuming that the deposited CVD films conformally coat the surface.
- (b) Accurately sketch the cross section *B-B*, assuming that the deposited CVD films conformally coat the surface. Note that the rough topography will make lithography difficult. The geometric design rules can prevent this situation by prohibiting the stacking of metal1 and metal2.

P2.20 The mask layout in Fig. 2.20 uses the simple MOSFET process in Fig. 2.19(b).

- (a) Sketch the cross section along *A-A*.

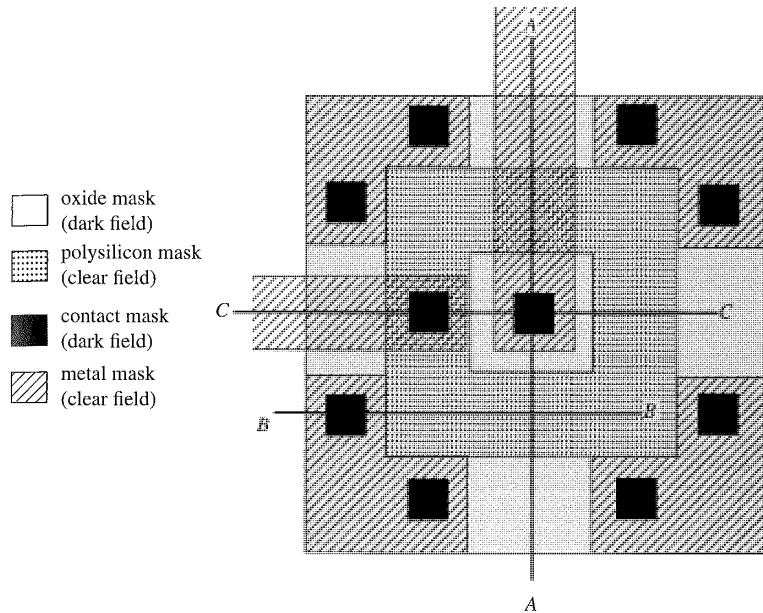


► Figure P2.19

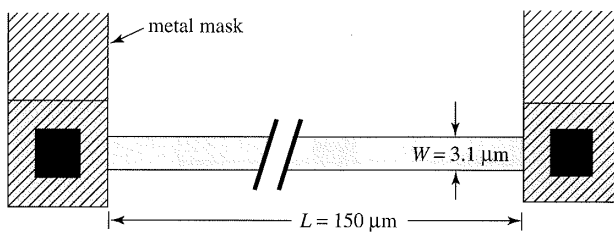
- (b) Sketch the cross section along $B-B$.
 (c) Sketch the cross section along $C-C$.

P2.21 The resistor whose layout is shown in Fig. P2.21 is a test structure on an integrated circuit.

The measured resistance is $R = 22.3 \text{ k}\Omega$. What is the sheet resistance of the ion-implanted layer in ohms per square? You can assume that the contact regions each contribute 0.65 squares.



► Figure P2.20



► Figure P2.21

P2.22 The two ion-implanted resistors in Fig. P2.22 have a different style of contact region compared to that shown in Fig. 2.22. The longer one has a measured resistance of $R_1 = 395 \, \Omega$, with the shorter one measuring $R_2 = 270 \, \Omega$.

From these measurements and the dimensions, what is the sheet resistance of this layer and the effective number of squares for this type of contact region?

P2.23 Given a p-type ion-implanted resistor with an average acceptor concentration of $N_{a1} = 10^{16} \, \text{cm}^{-3}$ over the depth $0 < x < 0.5 \, \mu\text{m}$ and $N_{a2} = 7.5 \times 10^{15} \, \text{cm}^{-3}$ over the depth $0.5 \, \mu\text{m} < x < 1.25 \, \mu\text{m}$. Find the sheet resistance in $\Omega/\square$.

P2.24 Given an n-type ion-implanted layer with thickness $t = 1 \, \mu\text{m}$ and average doping concentration $N_d = 10^{17} \, \text{cm}^{-3}$.

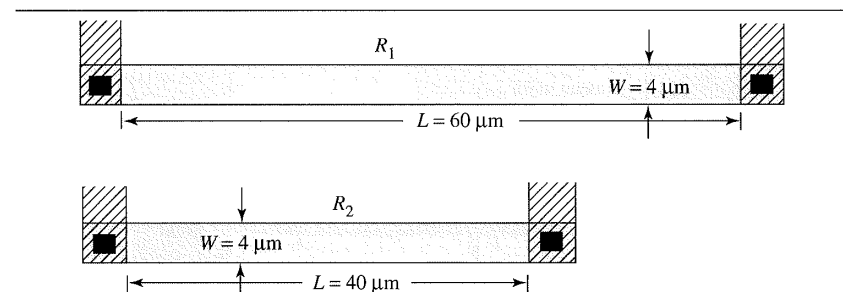
- What is the sheet resistance?
- In order to meet the target sheet resistance of $750 \, \Omega/\square$, we can perform a second implant anneal to add additional dopants to a layer of thickness $0.25 \, \mu\text{m}$. Neglecting the change in the overall layer thickness after annealing, what dose and dopant type (n or p) are needed?

P2.25 The normalized uncertainty in the average acceptor concentration in an implanted layer is $\epsilon_{N_a} = 0.025$. The uncertainties in the mobility and in the layer thickness ($\bar{t} = 1.4 \, \mu\text{m}$) are negligible for this problem. Finally, the dimensions of the resistor pattern have an uncertainty of $\delta = 400 \, \text{\AA}$.

- For a 120 square resistor with $\bar{N}_a = 10^{17} \, \text{cm}^{-3}$ and $\bar{W} = 2.5 \, \mu\text{m}$, what is the average resistance and its uncertainty?
- If we increase the doping by a factor of two, what is the width necessary in order to keep the average resistance the same? What is the new uncertainty?

P2.26 It is an attractive idea, at first glance, to compensate a silicon region so that $N_d = N_a$ and $n_o = p_o = n_i$. Assume that the wafer is doped with acceptors at precisely $N_a = 10^{16} \, \text{cm}^{-3}$ and that we can introduce donors uniformly into a $1 \, \mu\text{m}$ -thick surface region with a normalized uncertainty of 0.005.

- If we wish to guarantee that the surface region stays p-type, what donor concentration should we add? Note that the normalized uncertainty is 0.005.



► Figure P2.22

- (b) What is the range of sheet resistances for the surface region, given the doping in part (a)?

DESIGN PROBLEMS

D2.1 There are often geometric constraints in implementing resistors on an integrated circuit. Figure D2.1 defines the outline of the area in which the $160\text{ k}\Omega$ resistor must fit, given the following geometric design rules. The locations for the “dogbone” style contacts are set by the other circuit elements and cannot be moved. The oxide mask pattern can abut the edges of the outline in Fig. D2.1.

Geometric design rules:

minimum width of oxide mask pattern = minimum resistor width =

$$W = 2\text{ }\mu\text{m}$$

minimum spacing between oxide patterns = $W_g = 2.5\text{ }\mu\text{m}$

Resistor layer characteristics:

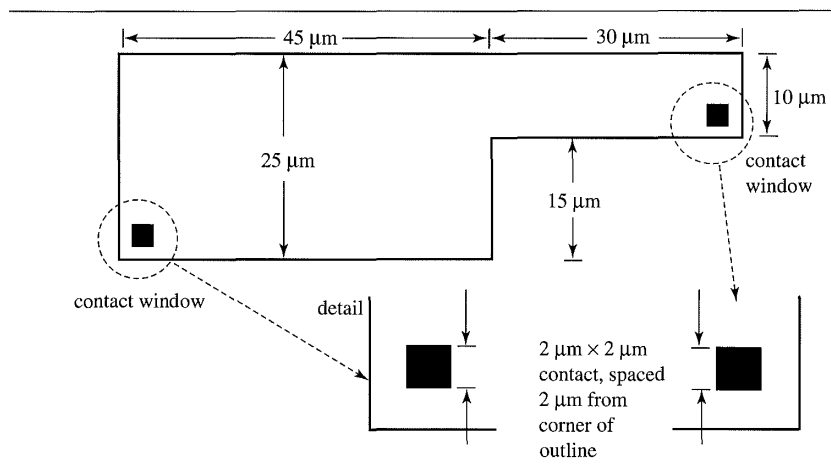
doping concentration $N_d = \bar{N}_d(1 \pm \epsilon_{N_d}) = 10^{17}\text{ cm}^{-3} (1 \pm 0.025)$

thickness $t = \bar{t}(1 \pm \epsilon_t) = 0.6\text{ }\mu\text{m} (1 \pm 0.02)$

variation in resistor lateral dimension: $\delta = 650\text{ }\text{\AA}$.

- Lay out the $160\text{ k}\Omega$ resistor to fit inside the outline in Fig. D2.1.
- What is the uncertainty in the resistance?
- Modify your layout in (a) to minimize the uncertainty, while keeping within the boundaries of the outline. By what factor can you reduce the uncertainty?

D2.2 In some cases, the IC designer aims for the maximum resistance possible in a particular area. In this problem, you will lay out a polysilicon resistor in an area of $100\text{ }\mu\text{m} \times 120\text{ }\mu\text{m}$. The contacts should use the “dogbone” style in Fig. 2.22(a) and are



► Figure D2.1

to be located as shown on Fig. D2.2. The polysilicon resistor can abut the edges of the outline. The geometric design rules and parameters for the polysilicon film (lot-to-lot uncertainties) are:

Geometric design rules:

minimum polysilicon line width $W = 1\ \mu\text{m}$

minimum spacing between polysilicon lines $= W_g = 1\ \mu\text{m}$

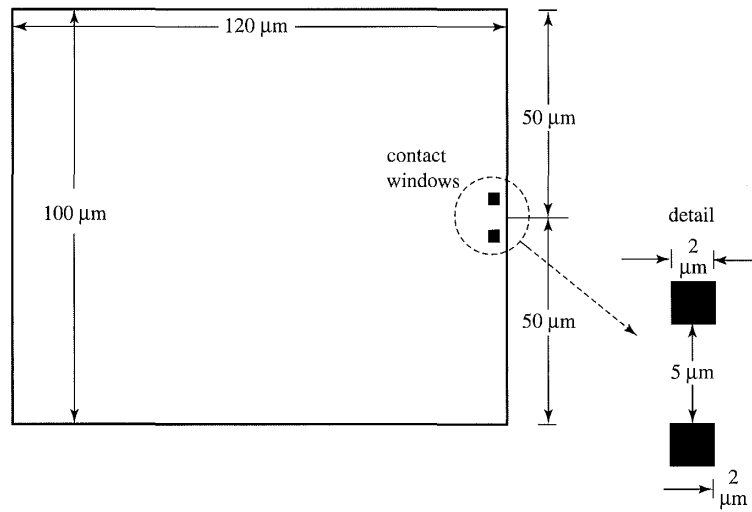
Polysilicon film parameters:

resistivity $\rho = \bar{\rho} (1 \pm \epsilon_\rho) = (1\text{ m}\Omega\text{-cm}) (1 \pm 0.05)$

thickness $t = \bar{t} (1 \pm \epsilon_t) = 0.5\ \mu\text{m} (1 \pm 0.05)$

variation in polysilicon line: $\delta = 500\ \text{\AA}$.

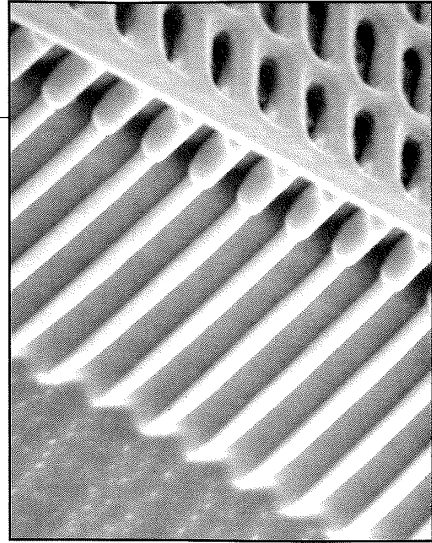
- Design the polysilicon resistor layout for the maximum number of squares in this area, given the design rules and fixed location of the contacts.
- What is the maximum resistance and its variation?



► Figure D2.2

chapter 3

P_N Junction and MOS Electrostatics



SEM of trenches used to form extremely dense MOS capacitors used in state-of-the-art dynamic random access memories. (Courtesy of International Business Machines Corporation. Unauthorized use not permitted.)

With a basic understanding of semiconductor physics and IC technology from Chapter 2, we are ready to tackle two important device structures: the pn junction and the metal-oxide-semiconductor (MOS) capacitor. A good command of electrostatics is needed to find the electrostatic potential, electric field, and charge density in these structures. The electrostatics of the MOS capacitor is the foundation for understanding the MOS field effect transistor in Chapter 4. The electrostatics of the pn junction plays a similar role for the pn diode in Chapter 6 and the bipolar transistor in Chapter 7.

Chapter Objectives

- ◆ The interrelationships between charge density, electric field, and potential.
- ◆ The boundary conditions for electric field and electrostatic potential.
- ◆ The connection between potential and carrier concentration in thermal equilibrium.
- ◆ The depletion approximation and its application in simplifying the analysis of the pn junction and MOS capacitor.
- ◆ The qualitative spatial variation of charge density, electric field, and potential through the pn junction and MOS capacitor.
- ◆ Quantitative models for the depletion widths, electric fields, and potential for both structures.
- ◆ The definition of capacitance and its application to the reverse-biased diode and the MOS capacitor.

► 3.1 APPLIED ELECTROSTATICS

In introductory physics courses, an overview of the classical theory of electromagnetism is presented. These courses focus on the underlying assumptions, the three dimensional mathematical structure, and the inherent beauty of this remarkably powerful theory. Problem solving is usually limited to highly symmetrical geometries where the various surface integrals can be evaluated exactly. As a result, students often master the mathematics of electromagnetism, but fail to gain an intuitive grasp of the subject. When presented with a simple problem in one dimension, students are often unsure which way the electric field points, let alone grasp the details of the potential variation!

In order to understand semiconductor devices, we must develop a practical, *working knowledge* of electrostatics. Semiconductor device structures have complicated boundaries that make it difficult, if not impossible, to find closed-form analytical solutions in two or three dimensions. Therefore, we will use the one-dimensional electrostatic equations to develop insight and a basic understanding of device operation. For detailed modeling, device physicists solve the three-dimensional equations numerically using simulation programs.

3.1.1 Gauss's Law

We will first collect the basic electrostatic equations that relate charge density, electric field, and electrostatic potential. Gauss's Law connects the electric field E [units: volts per centimeter, V/cm] with the charge density ρ [units: Coulombs per cubic centimeter, C/cm³]. Its one dimensional, differential form is

Gauss's Law in differential form

$$\frac{dE}{dx} = \frac{\rho}{\epsilon} \quad (3.1)$$

where ϵ is the electric permittivity [units: Farads per centimeter, F/cm]. If the electric field is known, we can differentiate it (i.e., find its slope) and use Eq. (3.1) to plot the charge density. In many cases, we will know the charge density ρ and need to find the electric field. Integrating Eq. (3.1) along an interval $x_a \leq x' \leq x$ we find the integral form of Gauss's Law:

Gauss's Law in integral form

$$\int_{x_a}^x d[\epsilon E(x')] = \epsilon E(x) - \epsilon_a E(x_a) = \int_{x_a}^x \rho(x') dx' = Q(x) \quad (3.2)$$

where $Q(x)$ is the charge in the interval [units: C/cm²]. The possibility of a change in permittivity due to a material interface in the interval has been accounted for by keeping the permittivity together with the field in the integral. This result is useful since we can often conclude from the distribution of charge in semiconductor device structures that $E(x_a) = 0$ for a particular boundary x_a .

3.1.2 Potential and Poisson's Equation

The electrostatic potential $\phi(x)$ is defined with respect to its value at a point x_0 as the integral of the negative of the electric field E from x_0 to x :

$$\phi(x) - \phi(x_o) = \int_{x_o}^x -E(x) dx \quad (3.3)$$

definition of electrostatic potential

Eq. (3.3) allows us to calculate the potential *difference* between two points given the electric field. To find a numerical value for $\phi(x)$, it is necessary to select an arbitrary reference point x_o where by definition $\phi(x_o) = 0$. We will consider the question of selecting references for $\phi(x)$ in semiconductors later in this chapter.

Differentiation of Eq. (3.3) allows us to find the electric field given the potential

$$E(x) = -\frac{d\phi(x)}{dx} \quad (3.4)$$

The potential can be related to the electric field, and vice versa, through Eqs. (3.3) and (3.4). Since the derivative of the electric field is proportional to the charge density through the differential form of Gauss's Law (Eq. (3.1)), differentiation of Eq. (3.4) and substitution leads to a second-order differential equation for the potential:

$$\frac{d^2\phi(x)}{dx^2} = -\frac{dE(x)}{dx} = -\frac{\rho(x)}{\epsilon} \quad (3.5)$$

Poisson's Equation

The result, called Poisson's Equation, can be useful since it directly links the potential with the charge distribution.

3.1.3 Intuition and Sign Conventions

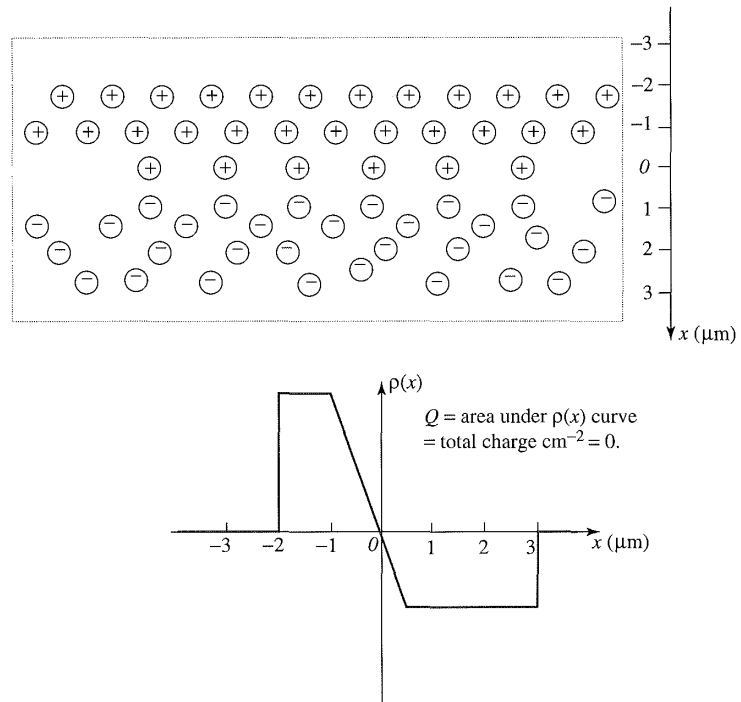
Even though Eqs. (3.1), (3.4), and (3.5) are ordinary differential equations with constant coefficients, the mathematical details of solving them can get in the way of finding the answer. For instance, a sign error will cause the electric field to point in the wrong direction. In this section we will introduce some techniques for sketching the answers *before* solving the equations.

As an example to motivate these techniques, Figure 3.1 sketches the charge density in a material with permittivity ϵ . Positive charges are located on the top (negative x) and negative charges are on the bottom (positive x). The *total* charge in the region is zero—which is always true for electronic device structures, as we will discover later in this chapter. From Eq. (3.2), the electric field $E(-2 \mu\text{m})$ at $x = -2 \mu\text{m}$ and the electric field $E(3 \mu\text{m})$ at $x = 3 \mu\text{m}$ are equal

$$E(3 \mu\text{m}) - E(-2 \mu\text{m}) = \int_{-2 \mu\text{m}}^{3 \mu\text{m}} \frac{\rho(x)}{\epsilon} dx = 0 \quad (3.6)$$

In device physics, we are interested in the internal electric field caused by the charge distribution inside the device. For our purposes, any external electric field at the boundaries of the charged region $E(-2 \mu\text{m}) = E(3 \mu\text{m})$ will be ignored and, for convenience, set to zero.

In this text, the x axis will be drawn horizontally with the positive direction to the right. The physical orientation of the x axis will be specified on the device cross section as in the top of Fig. 3.1. The charge distribution can be integrated graphically to find the electric field, using the integral form of Gauss's Law



► **Figure 3.1** Charge distribution and corresponding charge density $\rho(x)$. The area under the curve is the total charge Q (per cm^2) and is zero for this charge distribution—and for *all* charge distributions in silicon devices in this text.

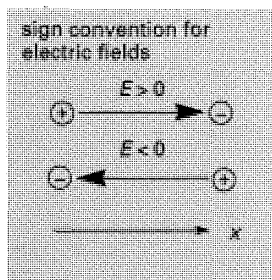
$$E(x) = E(-2\mu\text{m}) + \int_{-2\mu\text{m}}^x \frac{\rho(x')}{\epsilon} dx' = \int_{-2\mu\text{m}}^x \frac{\rho(x')}{\epsilon} dx' \quad (3.7)$$

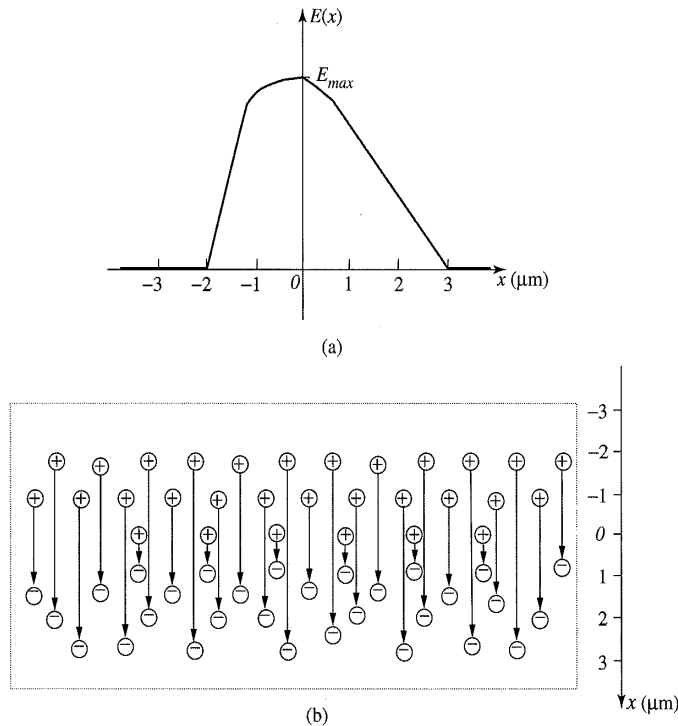
since $E(-2\mu\text{m}) = 0$. Figure 3.2(a) shows the electric field.

The electric field varies linearly with x for regions with constant charge density and varies quadratically in the interval $-1\mu\text{m} < x < 0.5\mu\text{m}$, where the charge density varies linearly. The electric field in Fig. 3.2(a) reaches its maximum value which is equal to

$$E_{\text{max}} = \int_{-2\mu\text{m}}^0 \frac{\rho(x')}{\epsilon} dx' \quad (3.8)$$

The electric field is positive over the entire charged region, meaning that it points in the same direction as the x axis. Note that $E(x)$ is positive in the interval $0 < x < 3\mu\text{m}$,



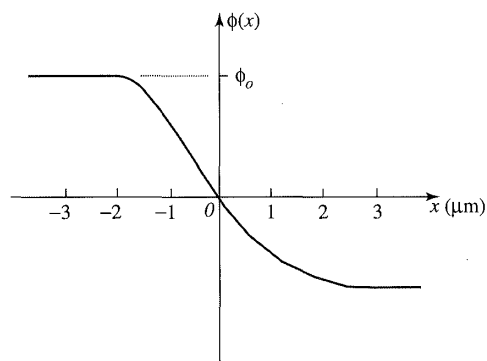


► **Figure 3.2** (a) Electric field and (b) qualitative picture of electric field lines connecting the positive and negative charges for the charge distribution in Fig. 3.1.

where the charge density $\rho(x)$ is negative. In order to clarify the connection between charge density and electric field, Figure 3.2(b) shows electric field lines connecting the positive and negative charges. From this figure, we can see that the electric field's magnitude increases (i.e., there are more field lines) as we move from $x = -2 \mu\text{m}$ to the origin, after which it decreases as the field lines terminate on the negative charges.

From the electric field $E(x)$, the electrostatic potential $\phi(x)$ can be found by integration from Eq. (3.3). Since we have not yet considered a potential reference in silicon, we will set the potential for $x < -2 \mu\text{m}$ to an arbitrary potential ϕ_0 in Fig. 3.3. Since the electric field is zero outside the charged region, the potential in this region is constant. To check the sketch of $\phi(x)$, the potential at $x = -2 \mu\text{m}$ must be higher than at $x = 3 \mu\text{m}$ since the electric field points from regions of high potential (positive charge) to regions of low potential (negative charge).

In summary, the basic features of the electric field and potential can be found from the charge density without excessive use of formal mathematics. In order to understand electronic devices, it is very important to *know* which way the electric field points or whether an answer for the potential “makes sense.” There will be



► **Figure 3.3** Electrostatic potential for the charge distribution in Fig. 3.1. The reference potential is ϕ_o at $x < -2 \mu\text{m}$.

many opportunities in this chapter, as well as Chapters 4 and 6, to practice the skills needed to master applied electrostatics.

3.1.4 Boundary Conditions for Potential and Electric Field

Electronic devices are made of layers of different materials. In order to solve for the potential or electric field through such an inhomogeneous structure, we need conditions for ϕ and E at the boundary between two materials. Starting with the potential, we observe that an abrupt jump in ϕ would lead to an infinite electric field at the boundary according to Eq. (3.4). Since infinite electric fields are not possible (they would tear the material apart), we conclude that $\phi(x)$ must be continuous

continuity of potential at a boundary

$$\phi(x_b^-) = \phi(x_b^+) \quad (3.9)$$

where the boundary is located at $x = x_b$.

On the other hand, the electric field usually jumps at a boundary. Figure 3.4(a) shows the case of a boundary between materials 1 and 2, where we have located the boundary at $x = 0$ for convenience. The electric field is indicated at two points located a very small distance Δ on either side of the boundary.

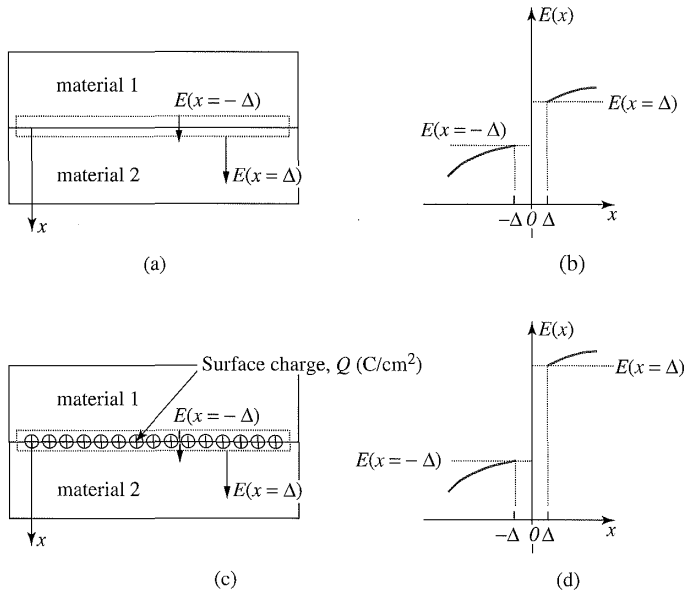
We can use the integral form to Gauss's Law to relate the electric field at $x = -\Delta$ to that at $x = +\Delta$

$$\int_{-\Delta}^{+\Delta} d\epsilon E(x) = \epsilon_2 E(x=\Delta) - \epsilon_1 E(x=-\Delta) = \int_{-\Delta}^{+\Delta} \rho(x) dx = 0 \quad (3.10)$$

The permittivity is left inside the integral on the left hand side since it varies with position through the interval $-\Delta \leq x \leq \Delta$. The boundary condition is found by letting $\Delta \rightarrow 0$

electric field jump for charge-free boundary

$$E(0^+) = \left(\frac{\epsilon_1}{\epsilon_2} \right) E(0^-) \quad (3.11)$$



► **Figure 3.4** (a) Boundary between materials 1 and 2 with permittivities $\epsilon_1 > \epsilon_2$ and (b) resulting jump in the electric field. (c) Boundary, with a surface charge Q located at the interface and (d) resulting jump in the electric field.

in which $E(0^+)$ is the electric field in material 2 at the boundary and $E(0^-)$ is the electric field in material 1 at the boundary. Figure 3.4(b) shows the jump in the electric field with Δ not quite zero, for the case where $\epsilon_1 > \epsilon_2$. When $\Delta \rightarrow 0$, the plot of electric field becomes discontinuous at the boundary.

A common interface in microelectronic devices is between silicon ($\epsilon_s = 11.7 \epsilon_0$ where $\epsilon_0 = 8.85 \times 10^{-14}$ F/cm is the permittivity of vacuum) and SiO_2 (silicon dioxide) which has $\epsilon_{ox} = 3.9 \epsilon_0$. For this case, the electric field in the silicon $E_s(0^+)$ is related to the electric field in the oxide $E_{ox}(0^-)$ by

$$E_s(0^+) = \left[\frac{3.9}{11.7} \right] E_{ox}(0^-) = \frac{E_{ox}(0^-)}{3.0} \quad (3.12)$$

The electric field therefore drops by a factor of three from the SiO_2 side to the silicon side of the boundary.

In some cases of practical importance, there can be a sheet of charge Q [units: C/cm²] at the boundary as shown in Fig. 3.4(c). The presence of this charge changes the integral in Eq. (3.10)

$$\int_{-\Delta}^{+\Delta} d\epsilon E(x) = \epsilon_2 E(x=\Delta) - \epsilon_1 E(x=-\Delta) = \int_{-\Delta}^{+\Delta} \rho(x) dx = Q \quad (3.13)$$