115

Semiconductor Silicon Crystal Technology

Leave the beaten track occasionally and dive into the woods. You will be certain to find something that you have never seen before.

Alexander Graham Bell

Semiconductor Silicon Crystal Technology

Fumio Shimura

Department of Materials Science and Engineering
North Carolina State University
Raleigh, North Carolina

Academic Press, Inc.
Harcourt Brace Jovanovich, Publishers
San Diego New York Berkeley Boston
London Sydney Tokyo Toronto

ACADEMIC PRESS, INC.
San Diego, Caliifornia 92101

United Kingdom Edition published by
ACADEMIC PRESS LIMITED
24-28 Oval Road, London NW1 7DX

Library of Congress Cataloging-in-Publication Data

Shimura, Fumio.
Semiconductor silicon crystal technology.

Includes bibliographical references and index.
1. Semiconductors. 2. Silicon crystals. I. Title.
TK7871.85.S523 1988 621.3815'2 88-6279
ISBN 0-12-640045-8 (alk. paper)

PRINTED IN THE UNITED STATES OF AMERICA
88 89 90 91 9 8 7 6 5 4 3 2 1

EA5455

Contents

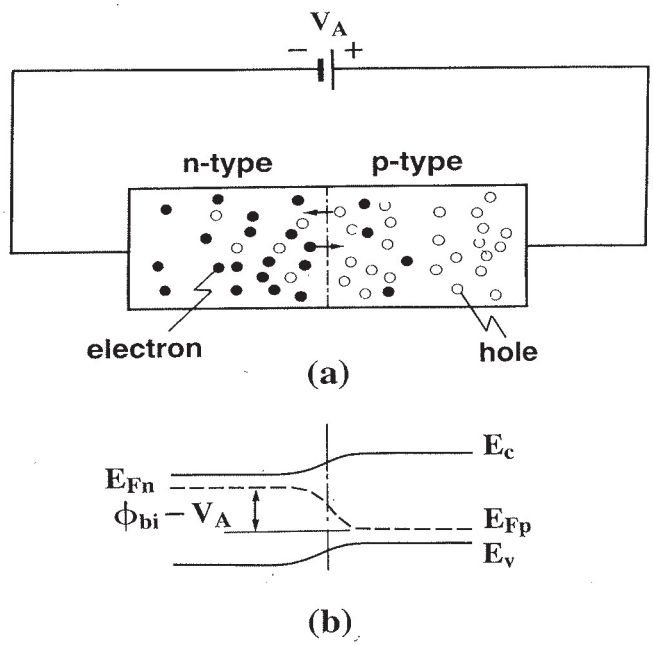

Fig. 4.18. Effect of an external forward bias potential on a *p-n* junction: (a) behavior of carriers and (b) energy-band structure.

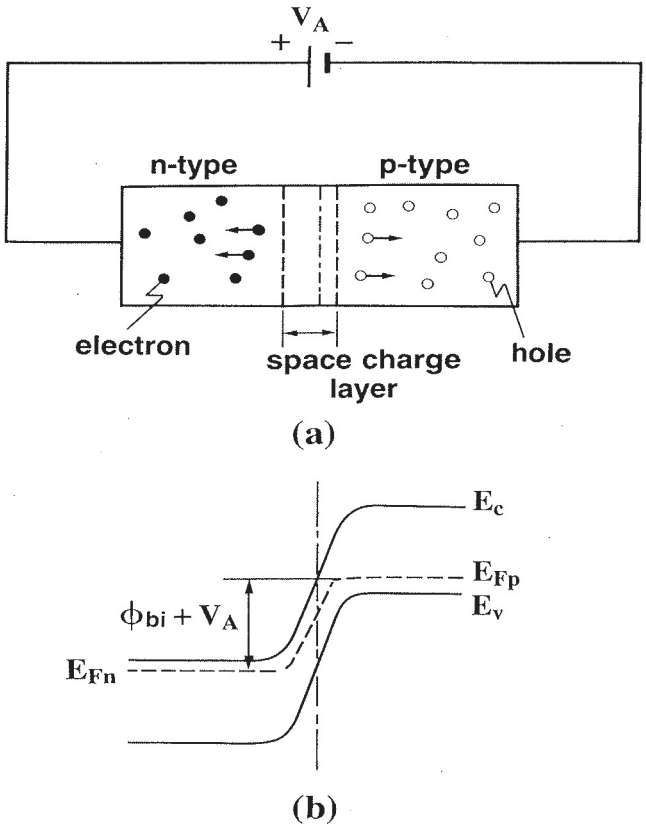

Fig. 4.19. Effect of an external reverse bias potential on a *p-n* junction: (a) behavior of carriers and (b) energy band structure.

4.4 Transistors

The term *transistor*[17] is from "transfer resistor," and transistors are unquestionably the most important semiconductor devices. The transistors discussed in this section include two types: (1) *bipolar* or *carrier injection transistors*, and (2) *unipolar* or *field-effect transistors*. First the structure of bipolar transistors is briefly discussed. In a bipolar transistor, the internal currents are obtained by both majority and minority carriers, hence the name bipolar transistors. Then field-effect transistors, which have become more and more important in the VLSI/ULSI technology, are discussed.

4.4.1. Bipolar Transistor

A bipolar transistor consists of a three-zone structure—*emitter* E, *collector* C, and *base* B—with two parallel *p-n* junctions very near each other built in the same semiconductor crystal. Two distinct types of the three-zone structure are possible: a *p-n-p* transistor (Fig. 4.20) and an *n-p-n* transistor (Fig. 4.21). In these figures, the basic conceptional structure (a), the cross section of a practical silicon planar bipolar transistor (b), and the circuit symbol (c) are shown. The operational principles of the two types are identical except for the

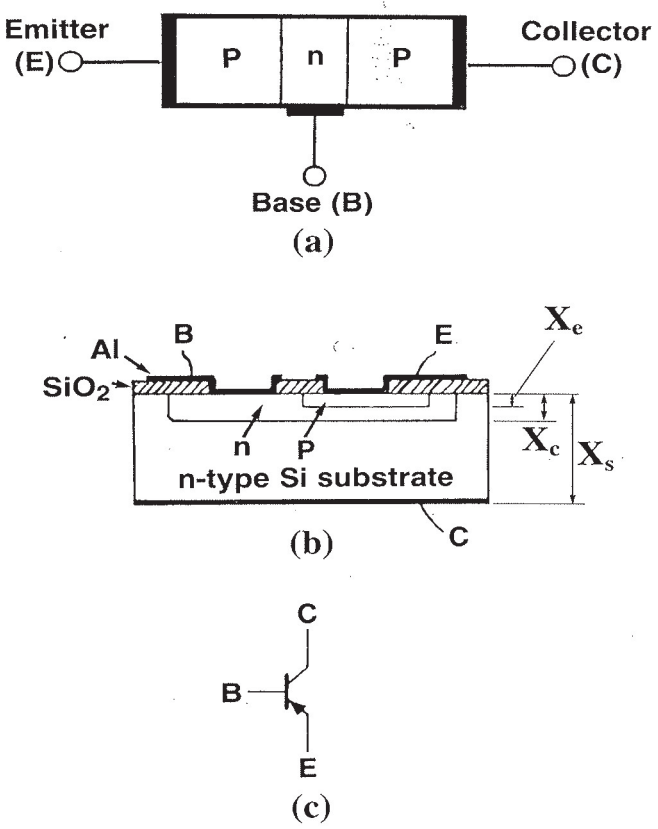

Fig. 4.20. Structure of *p-n-p* transistor: (a) basic conceptional structure, (b) cross section of a silicon planar transistor, and (c) circuit symbol.

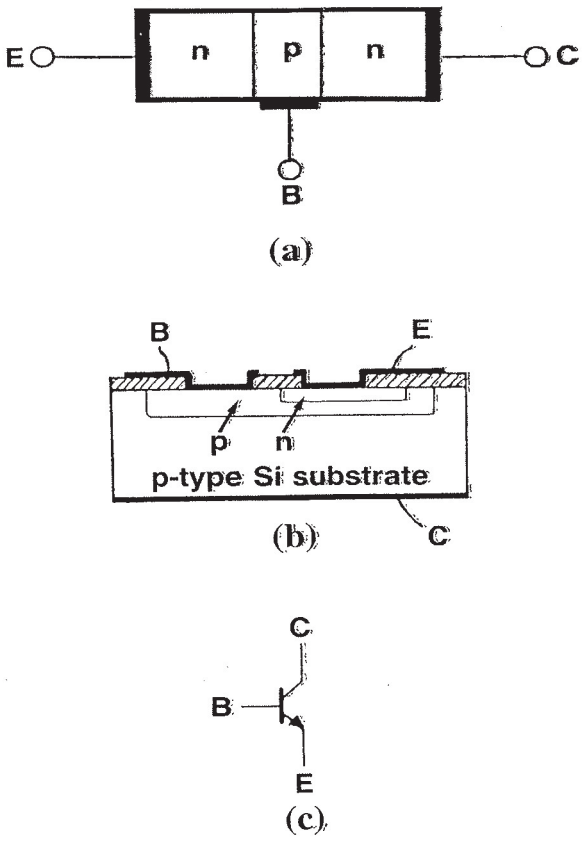

Fig. 4.21. Structure of *n-p-n* transistor: (a) basic conceptional structure, (b) cross section of a silicon planar transistor, and (c) circuit symbol.

interchange of minority and majority carrier types, and the polarity of the applied bias voltages.

The majority of discrete transistors made today are of the silicon planar type, as shown in Figs. 4.20b and 4.21b. A *p-n-p* planar silicon transistor shown in Fig. 4.20b. for example, is made by diffusing donor impurities such as phosphorus into a *p*-type Si substrate through an open window of an SiO_2 mask on the substrate, that is, base diffusion. Then acceptor impurities such as boron are diffused through a small window of a second mask for forming the emitter region. Note that the thicknesses of the emitter (X_e), collector (X_c), and substrate (X_s) as illustrated in Figs. 4.20b and 4.21b do not represent their realistic proportions with each other. That is, X_e and X_c are practically on the order of a few micrometers while X_s is several hundreds of micrometers or even closer to a millimeter in the case of a recent large-diameter Si substrate as described in Chapter 5. State-of-the-art bipolar ICs are commonly fabricated by planar epitaxial technology.[18] The fundamental advantage of epitaxial substrates over bulk wafers is that the structures such as high-resistivity layer/low-resistivity wafer and thin layer/wafer with opposite doping types are easily formed by epitaxial growth. Figure 4.22 shows a schematic cross section of an *n-p-n* bipolar planar transistor built on an

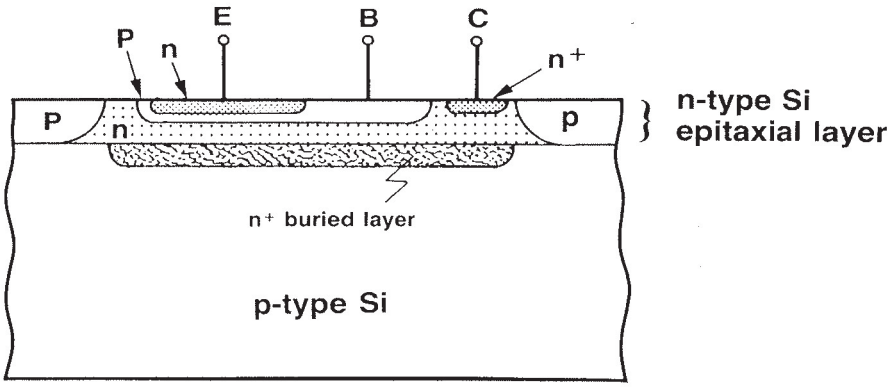

Fig. 4.22. Cross section of an *n-p-n* bipolar planar transistor built on an epitaxial substrate.

epitaxial Si substrate. The heavily doped n^+ buried layer functions as a subcollector, which reduces the series resistance of the devices and serves for alignment in subsequent wafer processing.

4.4.2 Physics of MOS

MOS Structure The metal oxide semiconductor field-effect transistor (MOSFET) has been realized as the most important device for VLSI/ULSI circuits such as memories and microprocessors because of its low fabrication cost, small size, and low power consumption. Figure 4.23 shows the structure of MOS or MIS (metal insulator semiconductor) and an MOS capacitor, which consists of a parallel-plate capacitor with one metallic plate, called the *gate*, and the other ohmic electrode.

C–V Characteristics The cross sections of an MOS capacitor fabricated on *p*-type Si are schematically shown with corresponding circuit symbols in Fig. 4.24. When the gate bias voltage $V_G \ll 0$ (Fig.4.24a), the MOS capacitor has the static capacitance C_0 per unit area, which is given by

$$C_0 = K_{ox}\varepsilon_0/t_{ox} \tag{4.36}$$

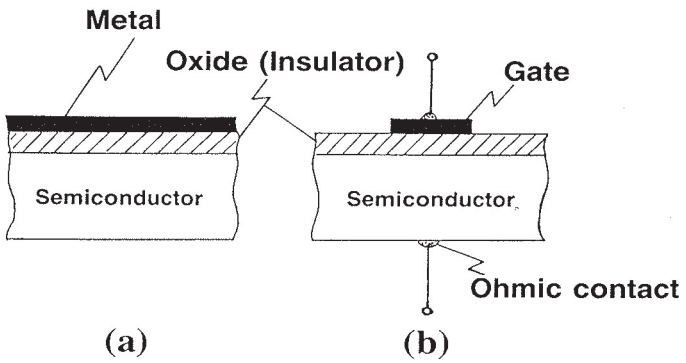

Fig. 4.23. Cross-sectional structure: (a) MOS and (b) MOS capacitor.

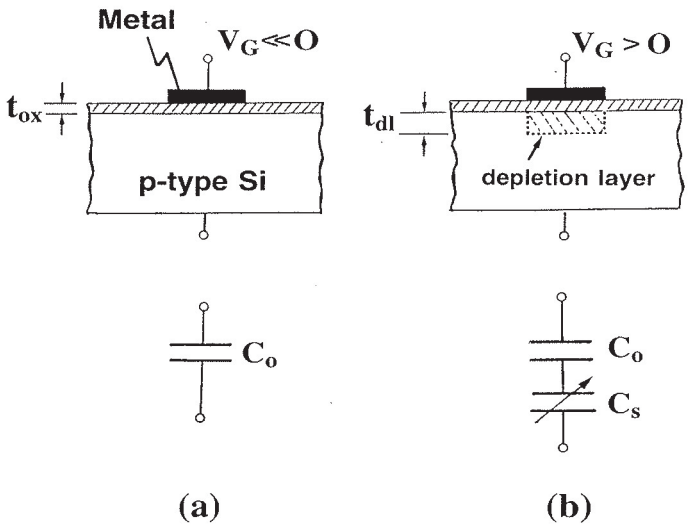

Fig. 4.24. Capacitance in an MOS capacitor fabricated on p-type silicon: (a) $V_G \ll 0$, and (b) $V_G > 0$.

where K_{ox} the dielectric constant of silicon oxide and ε_0 the dielectric permittivity of vacuum, and t_{ox} the thickness of oxide layer. When $V_G > 0$, holes are repelled and a depletion layer of t_{dl} in thickness is formed under the gate, as shown in Fig. 4.24b. The thickness t_{dl} depends predominantly on V_G and the doping level. Since there are no free carriers in the depletion layer, the depletion layer can be regarded as an insulator, and then the capacitance C_s, which is nonlinearly variable with V_G, is given by

$$C_s = K_{si}\varepsilon_0/t_{dl} \tag{4.37}$$

where K_{si} is the dielectric constant of silicon. The circuit model can be assumed as two series capacitors; the total capacitance C is then given by

$$1/C = 1/C_0 + 1/C_s \tag{4.38}$$

Figure 4.25 illustrates the relationship between the MOS capacitance (C/C_0) and gate bias voltage (V_G) for a p-type semiconductor in an ideal

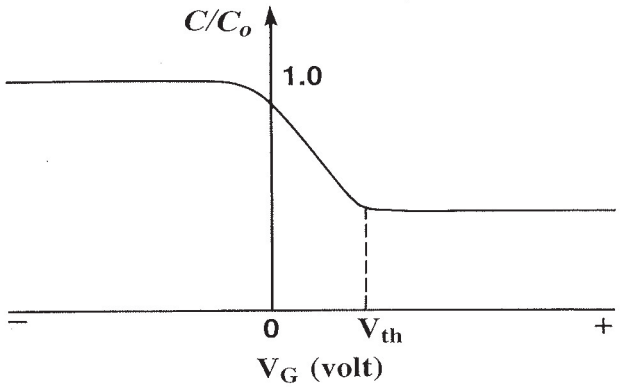

Fig. 4.25. Ideal C–V relation in an MOS capacitor fabricated on a p-type semiconductor.

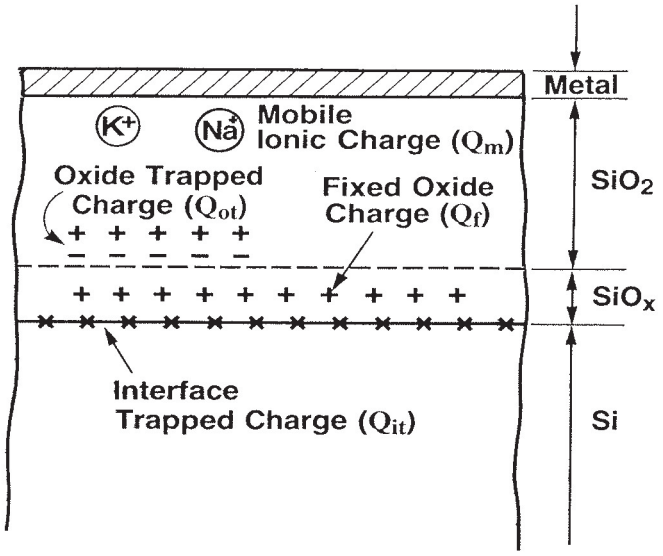

Fig. 4.26. Charges associated with thermally oxidized silicon. (After Deal.[20] ©1980 IEEE.)

situation.* Under a reverse bias condition ($V_G < 0$), C is equal to C_0 since no depletion layer exists. In the region $0 < V_G < V_{th}$ (V_{th} is referred to as the *threshold voltage*), the total capacitance C decreases with increasing t_{dl}, in turn decreasing C_s. When V_G is larger than V_{th}, C/C_0 becomes constant. However, the voltage due to the difference between the work functions of the metal and the semiconductor makes the C–V curve deviate from the ideal situation. In addition, the charges such as shown in Fig. 4.26[20] in the MOS capacitor also result in deviation of the C–V curve from the ideal situation. As shown in Fig. 4.27, when a positive charge exists at the SiO_2/Si interface, the C–V curve shifts to the left (a), while a negative charge shifts the C–V curve to the right (b). The voltage that is required on the gate of a MOS capacitor in order to achieve the ideal situation—that is, flat-band condition

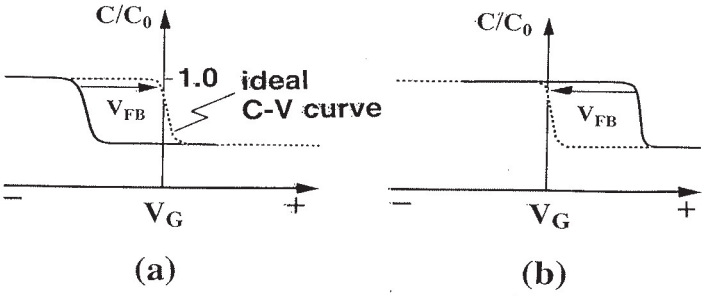

Fig. 4.27. Deviation of C–V curve from the ideal situation for a p-type semiconductor due to (a) positive charge ($V_{FB} > 0$) and (b) negative charge ($V_{FB} < 0$).

* Figure 4.25 shows the high-frequency C–V characteristics. At low frequencies, the capacitance goes through a minimum at V_{min} ($\sim V_{th}$) and then increases again as the inversion layer forms at the surface and C/C_0 reaches unity.[19] In the following, the high-frequency C–V plot, which is most commonly used to characterize MOS capacitors, is concerned.

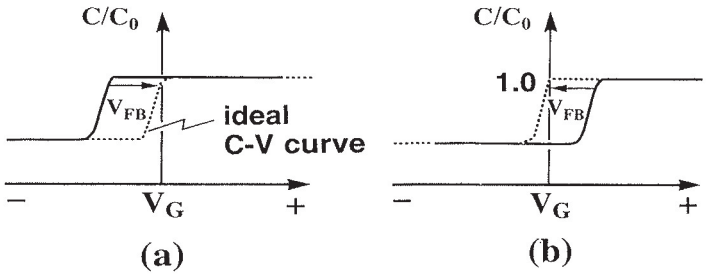

Fig. 4.28. Deviation of C–V curve from the ideal situation for an n-type semiconductor due to (a) positive charge ($V_{FB} > 0$) and (b) negative charge ($V_{FB} < 0$).

—is defined as the *flat-band voltage* and is represented with V_{FB}. For an n-type semiconductor, the C–V curve shifts similarly due to the existing states as shown in Fig. 4.28.

The flat-band voltage directly affects the threshold voltage V_{th}, which is defined as the gate bias voltage required to form an inversion layer in the surface of a semiconductor and is one of the most important factors for the operation of MOS devices. Constant V_{FB} is required to ensure high performance of MOS device operation.

4.4.3 MOS Transistor

Structure of MOSFET Devices in which the conduction involves only one polarity are referred to as *unipolar devices*, in contrast to bipolar devices. Among the unipolar devices, the MOSFET is the most important for VLSI/ULSI circuits. Many acronyms, such as MOST (MOS transistor), IGFET (insulated-gate FET), and MISFET (metal insulator semiconductor FET), represent the same device. Figure 4.29 depicts the basic structures of MOSFETs of two different types: (a) n-channel MOSFET (NMOS) and (b) p-channel MOSFET (PMOS). Note that the vertical and horizontal scales illustrated in Fig. 4.29, as in Figs. 4.20 and 4.21, do not represent a practical relation with each other. A MOSFET consists of four essential parts: *source* (S), *gate* (G), *channel*, and *drain* D. Heavily doped polysilicon or a combination of silicide and polysilicon as well as metals such as aluminum are used as

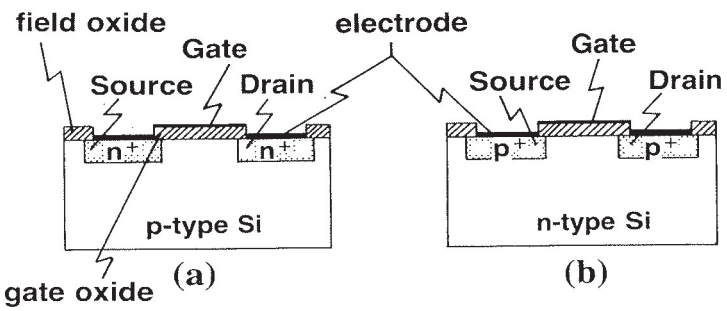

Fig. 4.29. Cross-sectional structure of enhancement-type MOSFETs: (a) n-channel MOSFET (NMOS) and (b) p-channel MOSFET (PMOS).

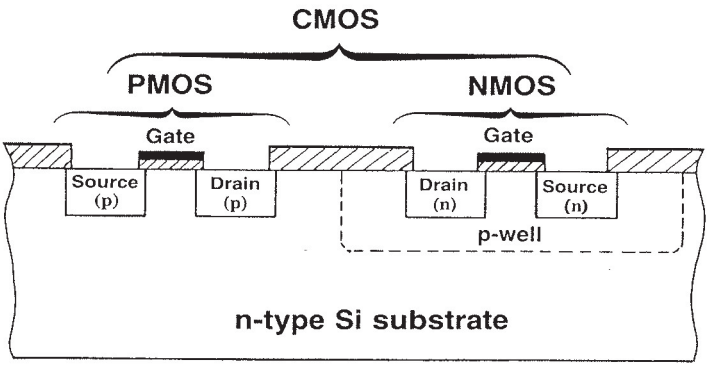

Fig. 4.30. Cross-sectional structure of CMOS.

the gate electrode. For NMOS, the source and drain regions are fabricated by either ion implantation or diffusion of donor impurities in the surface region of the *p*-type Si substrate. For PMOS, on the other hand, they are fabricated similarly with acceptor impurities in the *n*-type Si substrate. In an IC, a MOSFET is surrounded by the *field oxide*, which is thicker than the *gate oxide*, to isolate it from adjacent devices.

Although most of today's FETs, specifically random-access memories (RAMs), are NMOS designs, complementary MOS (CMOS) RAMs are the wave of the future VLSI/ULSI devices, including memories, microprocessors, and random logic. The basic structure of CMOS fabricated on an *n*-type Si substrate is diagrammatically shown in Fig. 4.30. It has a PMOS device fabricated with source-drain diffusion into the *n*-type substrate and an NMOS device fabricated with source-drain diffusion into the *p*-well (or *p*-tub), which was formed with *p*-type impurity diffusion into the substrate. Although the CMOS fabrication process is more complicated than that for simple NMOS or PMOS processes, CMOS technology[21] permits such advantages as reduction of power consumption, much simpler circuit design resulting in a much more efficient circuit layout, and thus smaller chips. There remain, however, a few CMOS problems that must be solved. The most notorious problem is *latchup*, which becomes more difficult to manage as the circuit geometry is reduced. For solutions, it is becoming common to fabricate CMOS circuits in high-resistivity epitaxial layers on low-resistivity substrates, or to utilize a trench structure in order to effectively separate each component.

Operating Principles Figure 4.31 diagrammatically explains the operating principle for NMOS. The n^+ source and drain regions, where the majority carrier electrons exist, are formed in the *p*-type Si substrate where the majority carrier holes dominate. When no voltage is applied to the gate (i.e., $V_G = 0$), drain current I_D does not flow even when a low voltage is applied to the drain, namely at $V_D - V_S > 0$, where V_D and V_S are the drain and source voltage, respectively, since no channel is formed between the two n^+ regions

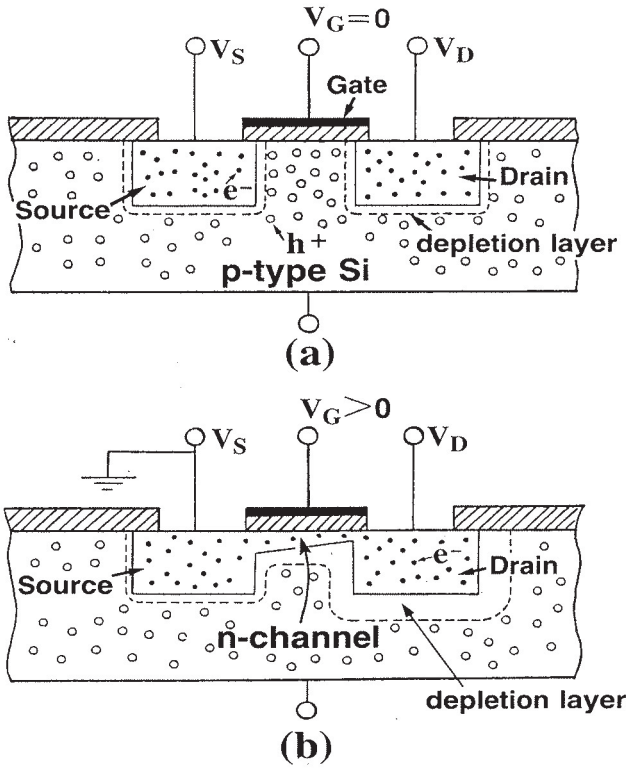

Fig. 4.31. Operating principle for NMOS: (a) $V_G = 0$ and (b) $V_G > 0$.

(see Fig. 4.31a). The only current that can flow from source to drain is the reverse leakage current. However, when a positive bias voltage V_G is applied to the gate and V_G exceeds the threshold voltage V_{th}, a surface inversion layer (i.e., *n*-channel) is formed between the source and drain, as shown in Fig. 4.31b. The two regions are then connected by a conducting *n*-channel through which a large drain current I_D can flow. The conductance of this channel can be modulated by varying the gate voltage. That is, the characteristics of the MOSFET are variable with the applied bias voltage to the gate. In the case of PMOS fabricated in *n*-type silicon, I_D is obtained similarly by applying a reverse bias voltage V_G, as shown in Fig. 4.32.

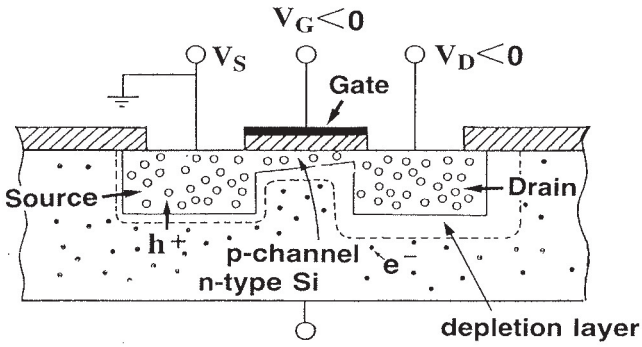

Fig. 4.32. Operating principle for PMOS at $V_G < 0$.

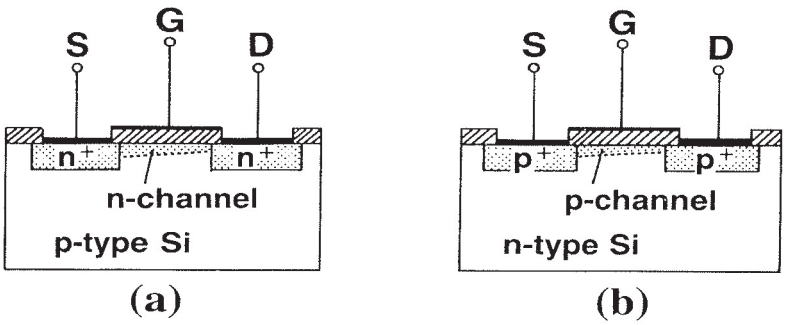

Fig. 4.33. Cross-sectional structure of depletion-type MOSFETs: (a) NMOS and (b) PMOS.

For the MOSFETs shown in Fig. 4.29, I_D does not flow as zero gate bias. These transistors are considered to be "normally off." This kind of transistor is called an *enhancement-type* device, because a gate bias voltage will enhance the conductivity of the channel. On the other hand, as shown in Fig. 4.33, MOSFETs that have a channel fabricated between the source and drain by *channel doping* are considered to be "normally on." A device that is normally on is called a *depletion-type* device, because a gate bias voltage is required to deplete the channel resulting in no I_D flow. In summary, there are basically four types of MOSFET depending on the types of channel and inversion layer: enhancement-type NMOS and PMOS, and depletion-type NMOS and PMOS. Although the operating principles of NMOS and PMOS are identical, NMOS is preferred to PMOS for devices that require high-speed switching. This is mainly because the mobility of electrons, μ_e, in an n-channel transistor is greater than that of holes, μ_h, in a p-channel transistor.

Contamination Effect on MOSFET Although the principle of a MOSFET device had been recognized since the late 1930s, the modern practical MOSFET was made in 1960s after the bipolar transistor was invented in 1947.[4,5] This is mainly due to the unsatisfactory technology of semiconductor surface preparation and oxide growth. As discussed, the field effect is highly sensitive to surface states, and in turn to surface contamination such as positive sodium ions. The presence of positive ions in the oxide at the SiO_2–Si interface attracts electrons in the silicon. This brings about a deleterious effect on an enhancement-type NMOS and a depletion-type PMOS (see Figs. 4.29a and 4.33b). For an IC structure, a current may flow between NMOSs, which must be isolated each other, because of the unfavorable channel formed by electrons attracted to the silicon surface. That is, only an enhancement-type PMOS may be influenced little by positive ion contamination. In fact, ICs using MOSFETs were originally based on PMOS devices,[22] although NMOS is preferred to PMOS because of the higher electron mobility with respect to hole mobility as described above. The development of material and process technologies have made NMOS dominate in the IC market since the

early 1970s. Again, it should be emphasized that the harmful effects of contamination on the device performance have to be minimized.

References

1. W. Schockley, "Electrons and Holes in Semiconductors." Van Nostrand-Reinhold, New York, 1950.
2. W. C. Dunlap, "An Introduction to Semiconductors." New York, 1962.
3. J. L. Moll, "Physics of Semiconductors." McGraw-Hill, New York, 1964.
4. S. M. Sze, "Physics of Semiconductor Devices," 2nd ed. Wiley, New York, 1981.
5. E. H. Nicollian and J. R. Brews, "MOS (Metal Oxide Semiconductor) Physics and Technology." Wiley, New York, 1982.
6. C. Kittel, "Elementary Solid State Physics: A Short Course." Wiley, New York, 1962.
7. J. M. Ziman, "Principles of the Theory of Solids." Cambridge Univ. Press, London and New York, 1964.
8. A. Bar-Lev, "Semiconductors and Electronic Devices," 2nd ed. Prentice-Hall, Englewood Cliffs, New Jersey, 1984.
9. Y. P. Vashni, Temperature dependence of the energy gap in semiconductors. *Physica (Amsterdam)* **34**, 149–154 (1967).
10. C. D. Thurmond, The standard thermodynamic function of the formation of electrons and holes in Ge, Si, GaAs and GaP. *J. Electrochem. Soc.* **122**, 1133–1141 (1975).
11. B. M. Welch, Advances in GaAs LSI/VLSI processing technology. *Solid State Technol.* Feb., pp. 95–101 (1980).
12. J. Bardeen and W. Schockley, Deformation potentials and mobilities in non-polar crystals. *Phys. Rev.* **80**, 72–80 (1950).
13. E. Conwell and V. F. Wesskopf, Theory of impurity scattering in semiconductors. *Phys. Rev.* **77**, 388–390 (1950).
14. L. A. Azaroff and J. J. Brophy, "Electronic Properties in Materials." McGraw-Hill, New York, 1963.
15. "Annual Book of ASTM Standards," Vol. 10.05 Electronics (II) F723–82, pp. 598–614. Am. Soc. Test. Mater. Philadelphia, Pennsylvania, 1984.
16. J. C. Irvin, Resistivity of bulk silicon and of diffused layers in silicon. *Bell Syst. Tech. J.* **41**, 387–410 (1962).
17. J. Bardeen and W. H. Brattain, The transistor, a semiconductor triode. *Phys. Rev.* **74**, 230–231 (1948).
18. H. M. Liaw, J. Rose, and P. L. Fejes, Epitaxial silicon for bipolar integrated circuits. *Solid State Technol.* May, pp. 135–143 (1984).
19. A. S. Grove, B. E. Deal, E. H. Snow, and C. T. Sah, Investigation of thermally oxidized silicon surface using metal-oxide-semiconductors. *Solid-State Electron.* 145–163 (1985).
20. B. E. Deal, Standard terminology for oxide charges associated with thermally oxidized silicon. *IEEE Trans. Electron Devices* **ED-27**, 606–608 (1980).
21. D. B. Scott, K.-L. Chen, and R. D. Davies, CMOS VLSI technology. *In* "VLSI Handbook" (N. G. Einspruch, ed.), pp. 121–150. Academic Press, New York, 1985.
22. L. C. Parrillo, VLSI process integration. *In* "VLSI Technology" (S. M. Sze, ed.). pp. 445–505. McGraw-Hill, New York, 1983.