5

MOS Capacitor

CHAPTER OBJECTIVES

This chapter builds a deep understanding of the modern MOS (metal–oxide–semiconductor) structures. The key topics are the concepts of surface depletion, threshold, and inversion; MOS capacitor C–V; gate depletion; inversion-layer thickness; and two imaging devices—charge-coupled device and CMOS (complementary MOS) imager. This chapter builds the foundation for understanding the MOSFETs (MOS Field-Effect Transistors).

The acronym **MOS** stands for **metal–oxide–semiconductor**. An MOS capacitor (Fig. 5–1) is made of a semiconductor body or substrate, an insulator film, such as SiO_2, and a metal electrode called a **gate**. The oxide film can be as thin as 1.5 nm. One nanometer is equal to 10 Å, or the size of a few oxide molecules.

Before 1970, the gate was typically made of metals such as Al (hence the M in MOS). After 1970, heavily doped polycrystalline silicon (see the sidebar, Three Kinds of Solid, in Section 3.7) has been the standard gate material because of its ability to

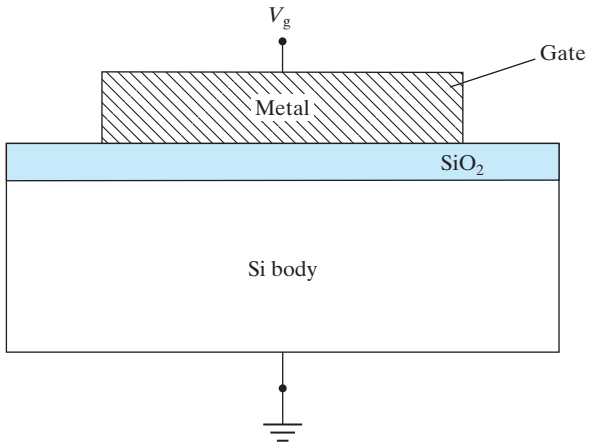

FIGURE 5–1 The MOS capacitor.

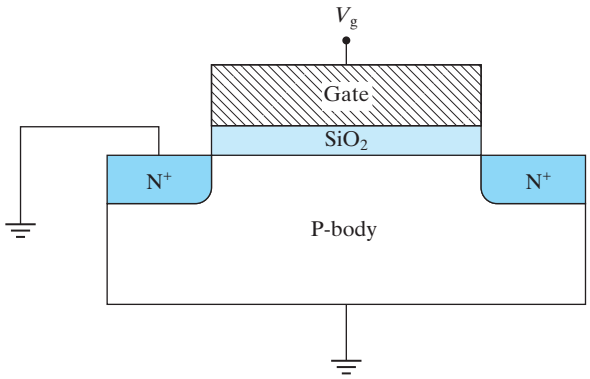

FIGURE 5–2 An MOS transistor is an MOS capacitor with PN junctions at two ends.

withstand high temperature without reacting with SiO_2. But the MOS name stuck. Unless specified otherwise, you may assume that the gate is made of heavily doped, highly conductive, polycrystalline silicon, or poly-Si for short. After 2008, the trend is to reintroduce metal gate and replace SiO_2 with more advanced dielectrics for the most advanced transistors (see Section 7.4).

The MOS capacitor is not a widely used device *in itself*. However, it is part of the MOS transistor—the topic of the next two chapters. The MOS transistor is by far the most widely used semiconductor device. An MOS transistor (Fig. 5–2) is an MOS capacitor with two PN junctions flanking the capacitor. This transistor structure is often a better structure for studying the MOS *capacitor* properties than the MOS capacitor itself as explained in Section 5.5.

5.1 • FLAT-BAND CONDITION AND FLAT-BAND VOLTAGE •

It is common to draw the energy band diagram with the oxide in the middle and the gate and the body on the left- and right-hand sides as shown in Fig. 5–3. The band diagram for $V_g = 0$ (Fig. 5–3b) is quite complex.

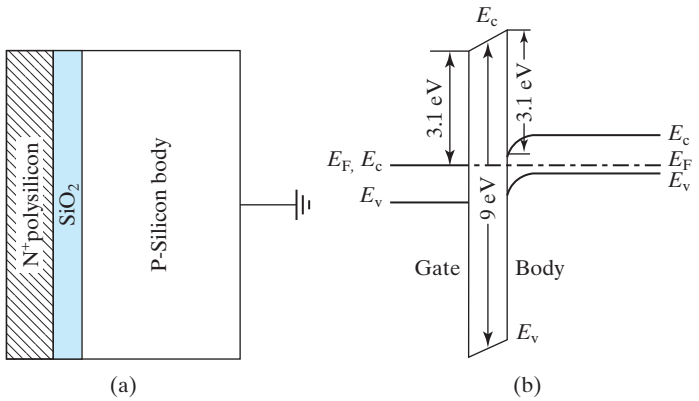

FIGURE 5–3 (a) Polysilicon-gate/oxide/semiconductor capacitor and (b) its energy band diagram with no applied voltage.

It is a good strategy to first study the energy band diagram for a special bias condition called the **flat-band condition**. Flat band is the condition where the energy band (E_c and E_v) of the substrate is flat at the Si–SiO_2 interface as shown in Fig. 5–4. This condition is achieved by applying a negative voltage to the gate in Fig. 5–3b, thus raising the band diagram on the left-hand side. (See Section 2.4 for the relation between voltage and the band diagram.) When the band is flat in the body as in Fig. 5–4, the surface electric field in the substrate is zero. Therefore the electric field in the oxide is also zero[1], i.e., E_c and E_v of SiO_2 are flat, too. E_c and E_v of SiO_2 are separated by 9 eV, the E_g of SiO_2. E_0, the **vacuum level**, is the energy state of electrons outside the material. E_0 of SiO_2 is above E_c by 0.95 eV. The difference between E_0 and E_c is called the **electron affinity**, another material parameter just as E_g is a material parameter. Si has an electron affinity equal to 4.05 eV. E_0 must be continuous at the Si–SiO_2 interface as shown in Fig. 5–4 (otherwise the electric field would be infinite). Therefore, E_c of SiO_2 is 3.1 eV higher than E_c of Si. This 3.1 eV is the Si–SiO_2 **electron energy barrier**. The **hole energy barrier** is 4.8 eV in Fig. 5–4. Because of these large energy barriers, electrons and holes normally cannot pass through the SiO_2 gate dielectric. E_c in the poly-silicon gate is also lower than the E_c of SiO_2 by 3.1 eV (the Si–SiO_2 energy barrier). Finally, E_F of the N^+poly-Si may be assumed to coincide with E_c for simplicity. In SiO_2, the exact position of E_F has no significance. If we place E_F anywhere around the middle of the SiO_2 band gap,

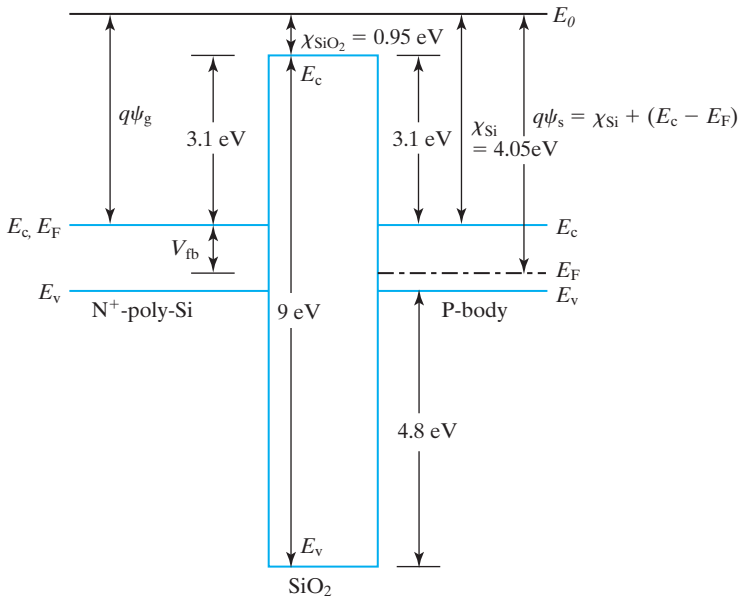

FIGURE 5–4 Energy band diagram of the MOS system at the flat-band condition. A voltage equal to V_{fb} is applied between the N^+-poly-Si gate and the P-silicon body to achieve this condition. ψ_g is the gate-material work function, and ψ_s is the semiconductor work function. E_0 is the vacuum level.

[1] According to Gauss's Law, with no interface charge, $\varepsilon_s \mathscr{E}_s = \varepsilon_{ox}\mathscr{E}_{ox}$ where $\mathscr{E}_s$ and $\mathscr{E}_{ox}$ are the body surface field and the oxide field.

$n = N_c\exp[(E_c - E_F)/kT]$ would be a meaninglessly small number such as 10^{-60} cm^{-3}. Therefore, the position of E_F in SiO_2 is immaterial.

The applied voltage at the flat-band condition, called V_{fb}, the **flat-band voltage**, is the difference between the Fermi levels at the two terminals.

$$V_{fb} = \psi_g - \psi_s \tag{5.1.1}$$

ψ_g and ψ_s are the gate work function and the semiconductor work function, respectively, in volts. The work function is the difference between E_0 and E_F. For an N^+-poly-Si gate, $\psi_g = 4.05$ V.[2] For the P-Si body, $\psi_s = 4.05\text{ V} + (E_c - E_F)/q$. For the example at hand, Eq. (5.1.1) and Fig. 5–4 indicate a negative V_{fb}, about –0.7 V.

5.2 • SURFACE ACCUMULATION •

How would Fig. 5–4 change if a more negative V_g than V_{fb} is applied? The band diagram on the gate side would be pushed upward (see Section 2.4). The result is shown in Fig. 5–5. Note that Fig. 5–5 is not drawn to scale (e.g., 3.1 eV is not about three times the silicon band gap) for the economy of page space. Such not-to-scale drawings are the norm. When $V_g \neq V_{fb}$, ϕ_s (surface voltage) and V_{ox} (oxide voltage) will be non-zero in general. $q\phi_s$ is the band bending in the substrate. Because the substrate is the voltage reference, ϕ_s is negative if E_c bends upward toward the surface as shown in Fig. 5–5 and positive if E_c bends downward. If this discussion of the sign of ϕ_s sounds strange, please review Sec. 2.4. V_{ox} is the voltage across the oxide. Again, V_{ox} is negative if the SiO_2 energy band tilts up toward the gate as it does in Fig. 5–5, and positive if it tilts downward toward the gate.

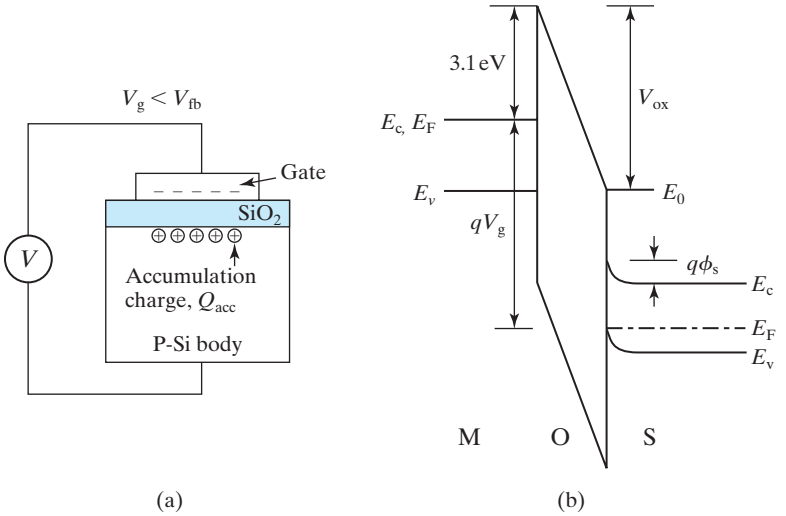

FIGURE 5–5 This MOS capacitor is biased into surface accumulation ($p_s > p_0 = N_a$). (a) Types of charge present. ⊕ represents holes and – represents negative charge. (b) Energy band diagram.

[2] In this case, ψ_g happens to be equal to χ_{Si}. In general, ψ_g is defined as the difference between E_0 and E_F.

Because E_v is closer to E_F at the surface than in the bulk, the surface hole concentration, p_s, is larger than the bulk hole concentration, $p_0 = N_a$. Specifically,

$$p_s = N_a e^{-q\phi_s/kT} \tag{5.2.1}$$

Since ϕ_s may be –100 or –200 mV, $p_s >> N_a$. That is to say, there are a large number of holes at or near the surface. They form an **accumulation layer** and these holes are called the **accumulation-layer holes**, and their charge the **accumulation charge, Q_{acc}**. This condition is known as **surface accumulation**. If the substrate were N type, the accumulation layer would hold electrons.

A relationship that we will use again and again is

$$\boxed{V_g = V_{fb} + \phi_s + V_{ox}} \tag{5.2.2}$$

At flat band, $V_g = V_{fb}$, $\phi_s = V_{ox} = 0$ and Eq. (5.2.2) is satisfied. If $V_g \neq V_{fb}$, the difference must be picked up by ϕ_s and V_{ox}. In the case of surface accumulation, ϕ_s may be ignored in a first-order model since it is quite small and Eq. (5.2.2) becomes

$$V_{ox} = V_g - V_{fb} \tag{5.2.3}$$

Using Gauss's Law, $\mathscr{E}_{ox} = -\dfrac{Q_{acc}}{\varepsilon_{ox}}$

$$V_{ox} = \mathscr{E}_{ox} T_{ox} = -\frac{Q_{acc}}{C_{ox}} \tag{5.2.4}$$

where C_{ox} is the oxide capacitance per unit area (F/cm^2) and Q_{acc} is the accumulation charge (C/cm^2). Equation (5.2.4) is the usual capacitor relationship, $V = Q/C$ (or $Q = C$–V) except for the negative sign. In $V = Q/C$, the capacitor voltage and charge are both taken from the same electrode. In the MOS capacitor theory, the voltage is the gate voltage, but the charge is the substrate charge because interesting things happen in the substrate. This unusual choice leads to the negative sign in Eq. (5.2.4). Equations (5.2.4) and (5.2.3) tell us

$$Q_{acc} = -C_{ox}(V_g - V_{fb}) \tag{5.2.5}$$

Therefore, the MOS capacitor in accumulation behaves like a capacitor with $Q = C$–V (or $-C$–V as explained earlier) but with a shift in V by V_{fb}. The shift is easily understandable because $Q_{acc} = 0$ when $V_g = V_{fb}$. In general, Eq. (5.2.4) should read

$$\boxed{V_{ox} = -Q_{sub}/C_{ox}} \tag{5.2.6}$$

where Q_{sub} is all the charge that may be present in the substrate, including Q_{acc}.

5.3 • SURFACE DEPLETION •

How would Fig. 5–4 change if a more positive V_g than V_{fb} is applied? The band diagram on the gate side will be pulled downward as shown in Fig. 5–6b. Clearly, *there is now a depletion region at the surface* because E_F is far from both E_c and E_v

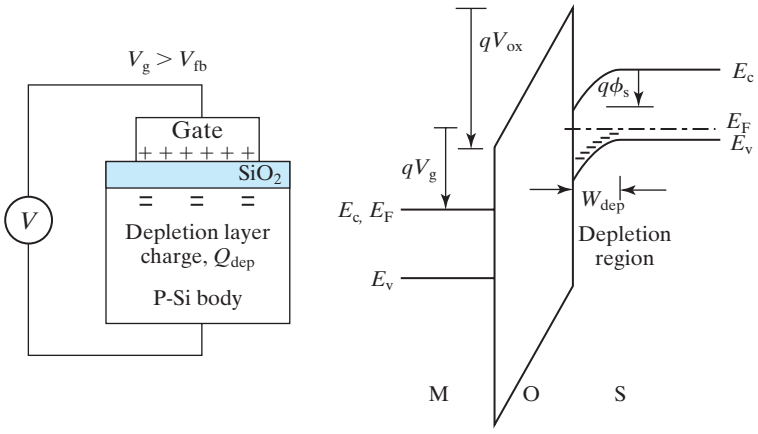

FIGURE 5–6 This MOS capacitor is biased into surface depletion. (a) Types of charge present; (b) energy band diagram.

and electron and hole densities are both small. This condition is called **surface depletion**. The depletion region has a width, W_{dep}. Equation (5.2.6) becomes

$$V_{ox} = -\frac{Q_{sub}}{C_{ox}} = -\frac{Q_{dep}}{C_{ox}} = \frac{qN_aW_{dep}}{C_{ox}} = \frac{\sqrt{qN_a2\varepsilon_s\phi_s}}{C_{ox}} \tag{5.3.1}$$

$$\phi_s = \frac{qN_aW_{dep}^2}{2\varepsilon_s} \tag{5.3.2}$$

Q_{dep} is negative because the acceptor ions (after accepting the extra electrons) are negatively charged. In Eqs. (5.3.1) and (5.3.2), we used $W_{dep} = \sqrt{(2\varepsilon_s\phi_s)/(qN_a)}$ [Eq. (4.2.10)]. Combining Eqs. (5.3.1), (5.3.2), and (5.2.2),

$$V_g = V_{fb} + \phi_s + V_{ox} = V_{fb} + \frac{qN_aW_{dep}^2}{2\varepsilon_s} + \frac{qN_aW_{dep}}{C_{ox}} \tag{5.3.3}$$

This equation can be solved to yield W_{dep} as a function of V_g. With W_{dep} determined, V_{ox} [Eq. (5.3.1)] and ϕ_s [Eq. (5.3.2)] become known.

5.4 • THRESHOLD CONDITION AND THRESHOLD VOLTAGE •

Let's make V_g in Fig. 5–6 increasingly more positive. This bends the energy band down further. At some V_g, E_F will be close enough to E_c at the Si–SiO_2 interface that the surface is no longer in depletion but at the **threshold** of **inversion**. The term inversion means that the surface is inverted from P type to N type, or electron rich. Threshold is often defined as the condition when the surface electron concentration, n_s, is equal to the bulk doping concentration, N_a. That means $(E_c - E_F)_{surface} = (E_F - E_v)_{bulk}$, or $A = B$ in Fig. 5–7.[3] That, in turn, means

[3] Assuming $N_c = N_v$, we conclude that $A = B$ when $n_s = N_a$.

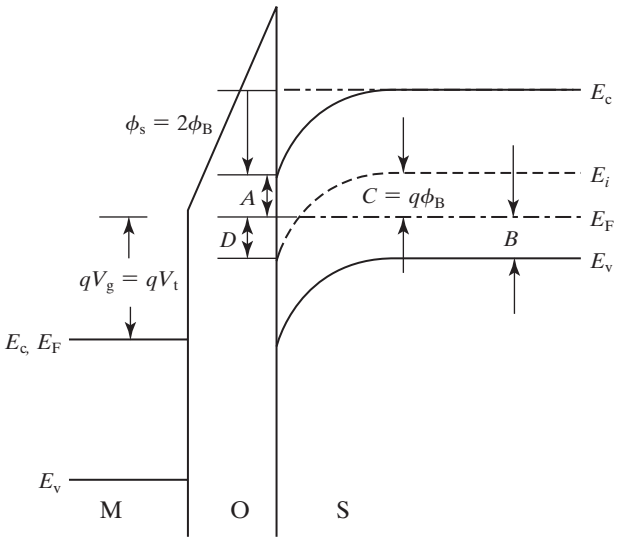

FIGURE 5–7 The threshold condition is reached when $n_s = N_a$, or equivalently, $A = B$, or $\phi_s = \phi_{st} = 2\phi_B$. Note that positive ϕ_{st} corresponds to downward band bending.

$C = D$. E_i is a curve drawn at **midgap**, which is half way between E_c and E_v. Let the surface potential (band bending) at the threshold condition be ϕ_{st}. It is equal to $(C + D)/q = 2C/q = 2\phi_B$.

Using Eqs. (1.8.12) and (1.8.8) and assuming $N_c = N_v$,

$$q\phi_B \equiv \frac{E_g}{2} - (E_F - E_v)\Big|_{bulk}$$
$$= kT\ln\frac{N_v}{n_i} - kT\ln\frac{N_v}{N_a} = kT\ln\frac{N_a}{n_i} \tag{5.4.1}$$

ϕ_s at the threshold condition is

$$\boxed{\phi_{st} = 2\phi_B = 2\frac{kT}{q}\ln\frac{N_a}{n_i}} \tag{5.4.2}$$

The V_g at the threshold condition is called the **threshold voltage**, V_t. Substituting Eqs. (5.4.2) and (5.3.1) into Eq. (5.2.2),

$$\boxed{V_t = V_{fb} + 2\phi_B + \frac{\sqrt{qN_a 2\varepsilon_s 2\phi_B}}{C_{ox}}} \tag{5.4.3}$$

The threshold voltage as a function of T_{ox} and body doping using Eq. (5.4.3) is plotted in Fig. 5–8. In this figure, the gate dielectric is assumed to be SiO_2 with dielectric constant $\varepsilon_{ox} = 3.9$.

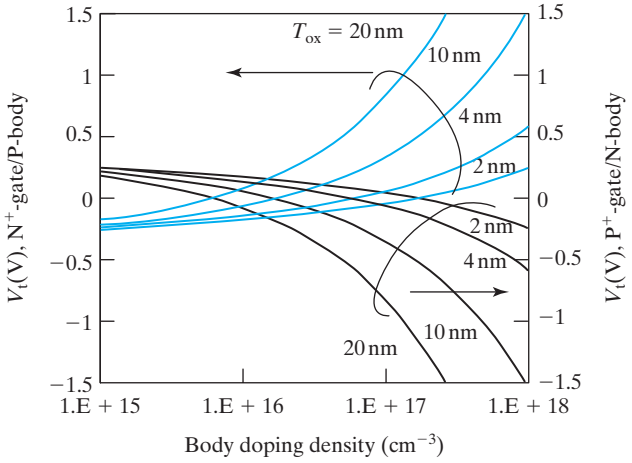

FIGURE 5–8 Theoretical threshold voltage vs. body doping concentration using Eq. (5.4.3). See Section 5.5.1 for a discussion of the gate doping type.

• N-Type Body •

For an N-type body, Eq. (5.4.3) becomes

$$V_t = V_{fb} + \phi_{st} - \frac{\sqrt{2qN_d\varepsilon_s|\phi_{st}|}}{C_{ox}} \tag{5.4.4}$$

$$\phi_{st} = -2\phi_B \tag{5.4.5}$$

$$\phi_B = \frac{kT}{q}\ln\frac{N_d}{n_i} \tag{5.4.6}$$

Exercise: Draw the band diagram of an N-body MOS capacitor at threshold and show that the second term (ϕ_{st}) and the third term (V_{ox}) in Eq. (5.4.4) are negative.

5.5 • STRONG INVERSION BEYOND THRESHOLD •

Figure 5–9b shows the energy diagram at strong inversion, $V_g > V_t$. As shown in Fig. 5–9a, there is now an **inversion layer**, which is filled with **inversion electrons.** The **inversion charge density** is represented with Q_{inv} (C/cm^2). ϕ_s does not increase much further beyond $2\phi_B$ since even a 0.1 V further increase in ϕ_s would induce a much larger surface electron density and therefore a larger V_{ox} that would soak up the V_g in Eq. (5.2.2). If ϕ_s does not increase, neither will the depletion region width. Approximately speaking, W_{dep} has reached its maximum value

$$W_{dmax} = \sqrt{\frac{2\varepsilon_s 2\phi_B}{qN_a}} \tag{5.5.1}$$

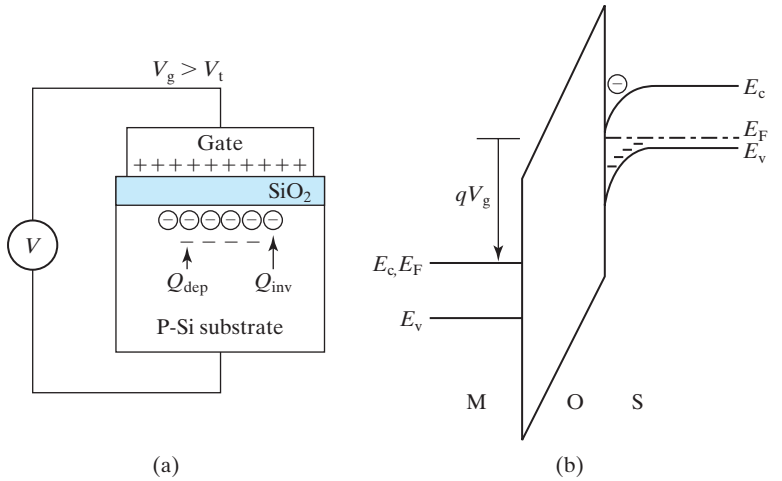

FIGURE 5–9 An MOS capacitor biased into inversion. (a) Types of charge present; (b) energy band diagram with arrow indicating the sense of positive V_g.

$$V_g = V_{fb} + 2\phi_B - \frac{Q_{dep}}{C_{ox}} - \frac{Q_{inv}}{C_{ox}} = V_{fb} + 2\phi_B + \frac{\sqrt{qN_a 2\varepsilon_s 2\phi_B}}{C_{ox}} - \frac{Q_{inv}}{C_{ox}}$$

$$= V_t - \frac{Q_{inv}}{C_{ox}} \tag{5.5.2}$$

Equations (5.2.2) and (5.2.6) are used in deriving Eq. (5.5.2).

$$\therefore \quad \boxed{Q_{inv} = -C_{ox}(V_g - V_t)} \tag{5.5.3}$$

Equation (5.5.3) confirms that the MOS capacitor in strong inversion behaves like a capacitor except for a voltage offset of V_t. At $V_g = V_t$, $Q_{inv} = 0$.

In this section, we have assumed that electrons will appear in the inversion layer whenever the closeness between E_c and E_F suggests their presence. However, there are few electrons in the P-type body, and it can take minutes for thermal generation to generate the necessary electrons to form the inversion layer. The MOS transistor structure shown in Fig. 5–2 solves this problem. The inversion electrons are supplied by the N^+ junctions, as shown in Fig. 5–10a. The inversion layer may be visualized as a very thin N layer (hence the term *inversion* of the surface conductivity type) as shown in Fig. 5–10b. The MOS transistor as shown in Figs. 5–2 and 5–10 is a more versatile structure for studying the MOS system than the MOS capacitor.

5.5.1 Choice of V_t and Gate Doping Type

The p-body transistor shown in Fig. 5–10 operates in an integrated circuit (IC) with V_g swinging between zero and a positive power supply voltage. To make circuit design easier, it is routine to set V_t at a small positive value, e.g., 0.4 V, so that, at $V_g = 0$, the transistor does not have an inversion layer and current does not flow between the two N^+ regions. A transistor that does not conduct current at $V_g = 0$ is called an **enhancement-type device**. This V_t value can be obtained with an N^+ gate and convenient body doping density as shown in Fig. 5–8. If the p-body device is paired with a P^+ gate,

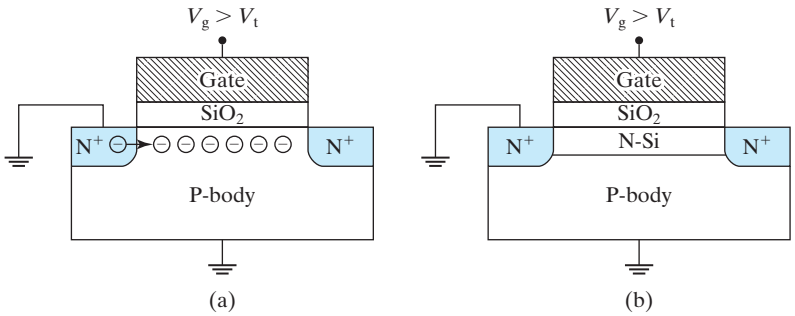

FIGURE 5–10 (a) The surface inversion behavior is best studied with a PN junction butting the MOS capacitor to supply the inversion charge. (b) The inversion layer may be thought of as a thin N-type layer.

V_t would be too large (over 1 V) and necessitate a larger power supply voltage. This would lead to larger power consumption and heat generation (see Section 6.7.3).

Similarly, an N-type body is routinely paired with a P^+ gate. In summary, P body is almost always paired with N^+ gate to achieve a small positive threshold voltage, and N body is normally paired with P^+ gate to achieve a small negative threshold voltage. The other body-gate combinations are almost never encountered.

● Review: Basic MOS Capacitor Theory ●

Let us review the concepts, nomenclatures, common approximations, and simple relationships associated with the MOS capacitor theory. We will do so using a series of figures, starting with Fig. 5–11. The surface potential, ϕ_s, is zero at V_{fb} and approximately zero in the accumulation region. As V_g increases from V_{fb} into the depletion regime, ϕ_s increases from zero toward $2\phi_B$. When ϕ_s reaches $2\phi_B$, the surface electron concentration becomes so large that the surface is considered **inverted**. The V_g at that point is called V_t, the **threshold voltage**.

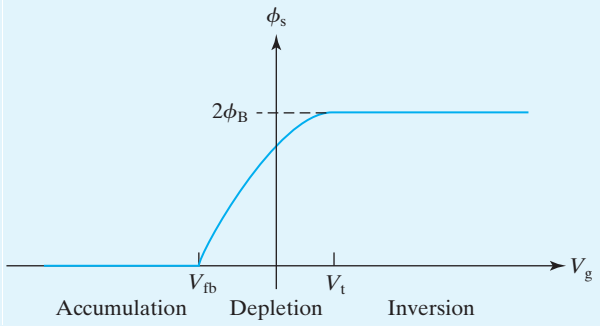

FIGURE 5–11 Surface potential saturates at $2\phi_B$ when V_g is larger than V_t.

Figure 5–12 uses W_{dep} to review the MOS capacitor. There is no depletion region when the MOS interface is in accumulation. W_{dep} in the PN junction and in the MOS capacitor is proportional to the square root of the band bending (ϕ_s in the MOS case). W_{dep} saturates at W_{dmax} when $V_g \geq V_t$, because ϕ_s saturates at $2\phi_B$.

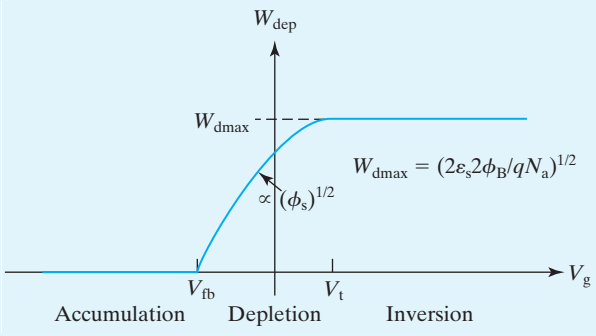

FIGURE 5–12 Depletion-region width in the body of an MOS capacitor.

Figure 5–13 reviews the three charge components in the substrate. The depletion charge Q_{dep} is constant in the inversion region because W_{dep} is a constant there. $Q_{inv} = -C_{ox}(V_g - V_t)$ appears in the inversion region. Q_{acc} shows up in the accumulation

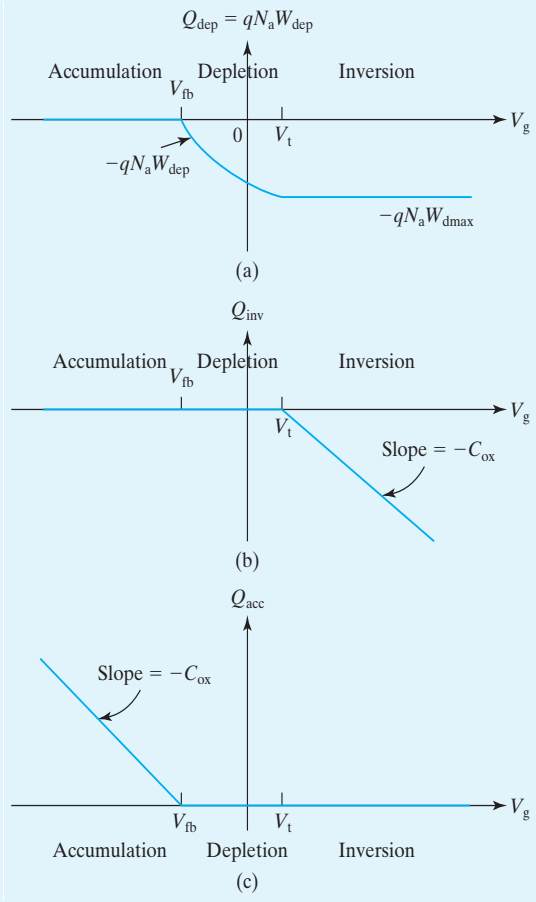

FIGURE 5–13 Components of charge (C/cm^2) in the MOS capacitor substrate: (a) depletion-layer charge; (b) inversion-layer charge; and (c) accumulation-layer charge.

region. In both (b) and (c), the slope is $-C_{ox}$. Figure 5–14 shows the total substrate charge, Q_{sub}. Q_{sub} in the accumulation region is made of accumulation charge. Q_{sub} is made of Q_{dep} in the depletion region. In the inversion region, there are two components, Q_{dep} that is a constant and Q_{inv} that is equal to $-C_{ox}(V_g - V_t)$.

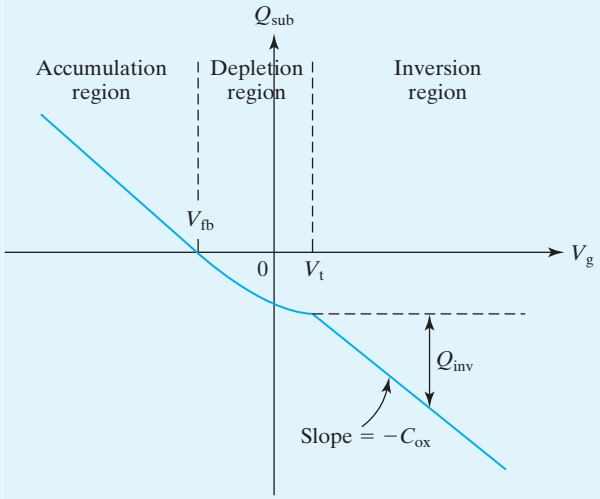

FIGURE 5–14 The total substrate charge, Q_{sub} (C/cm^2), is the sum of Q_{acc}, Q_{dep}, and Q_{inv}.

5.6 • MOS C–V CHARACTERISTICS •

The capacitance–voltage (C–V) measurement is a powerful and commonly used method of determining the gate oxide thickness, substrate doping concentration, threshold voltage, and flat-band voltage. The *C–V* curve is usually measured with a C–V meter (Fig. 5–15), which applies a DC bias voltage, V_g, and a small sinusoidal signal (1 kHz–10 MHz) to the MOS capacitor and measures the capacitive current with an AC ammeter. The capacitance is calculated from $i_{cap}/v_{ac} = \omega C$.

The capacitance in the MOS theory is always the **small-signal capacitance**

$$C \equiv \frac{dQ_g}{dV_g} = -\frac{dQ_{sub}}{dV_g} \tag{5.6.1}$$

The negative sign in Eq. (5.6.1) arises from the fact that V_g is taken at the top capacitor plate but Q_{sub} is taken at the bottom capacitor plate (the body). Q_{sub} is given in Fig. 5–14 and its derivative is shown in Fig. 5–16.

In the accumulation region, the MOS capacitor is just a simple capacitor with capacitance C_{ox} as shown in Fig. 5–17a. Figure 5–17b shows that in the depletion region, the MOS capacitor consists of two capacitors in series: the oxide capacitor, C_{ox}, and the depletion-layer capacitor, C_{dep}. Under the AC small-signal voltage, W_{dep} expands and contracts slightly at the AC frequency. Therefore, the AC charge appears at the bottom of the depletion layer as shown in Fig. 5–17b.

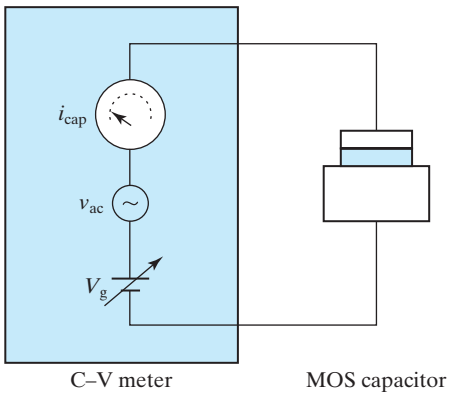

FIGURE 5–15 Setup for the C–V measurement.

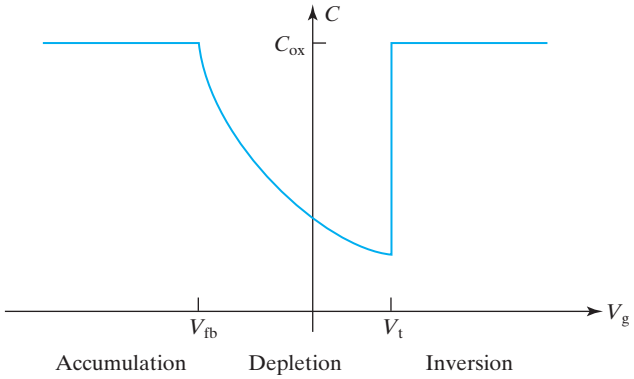

FIGURE 5–16 The quasi-static MOS C–V characteristics.

$$C_{dep} = \frac{\varepsilon_s}{W_{dep}} \tag{5.6.2}$$

$$\frac{1}{C} = \frac{1}{C_{ox}} + \frac{1}{C_{dep}} \tag{5.6.3}$$

$$\frac{1}{C} = \sqrt{\frac{1}{C_{ox}^2} + \frac{2(V_g - V_{fb})}{qN_a\varepsilon_s}} \tag{5.6.4}$$

To derive Eq. (5.6.4), one needs to solve Eq. (5.3.3) for W_{dep} as a function of V_g. The derivation is left as an exercise for the reader in the problems section at the end of the chapter. As V_g increases beyond V_{fb}, W_{dep} expands, and therefore C decreases as shown in Fig. 5–16.

Figure 5–17c shows that an inversion layer exists at the Si–SiO_2 interface. In response to the AC signal, Q_{inv} increases and decreases at the AC frequency. The inversion layer plays the role of the bottom electrode of the capacitor. Therefore, C reverts to C_{ox} in the inversion region as shown in Fig. 5–16. This C–V curve is called

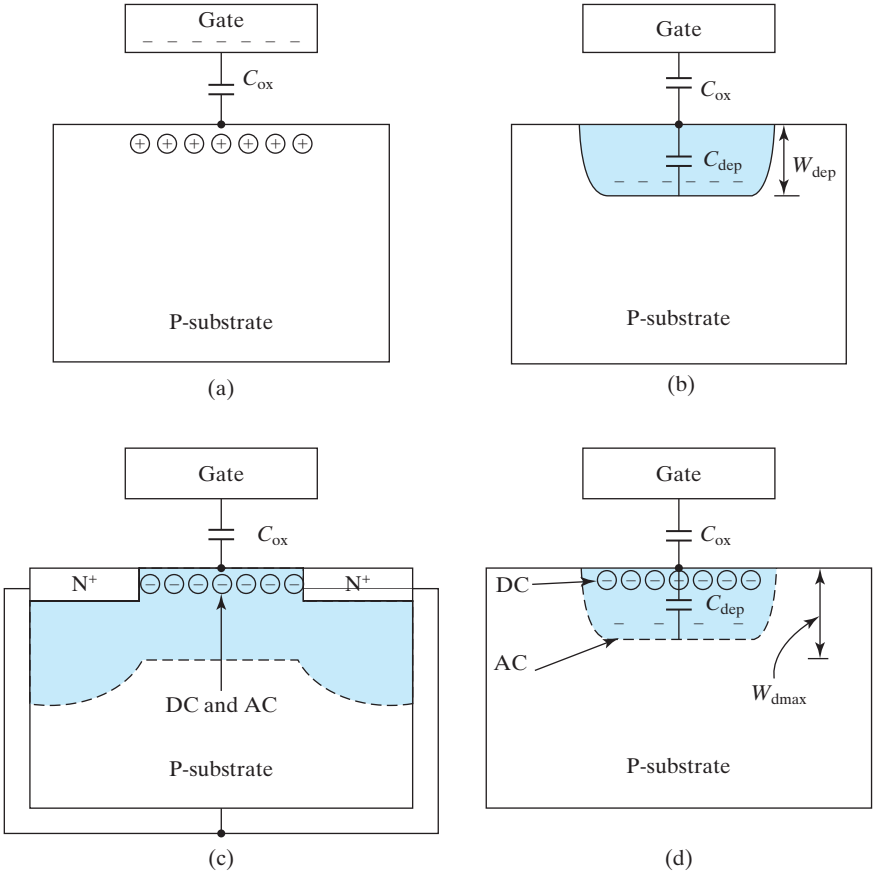

FIGURE 5–17 Illustration of the MOS capacitor in all bias regions with the depletion-layers shaded. (a) Accumulation region; (b) depletion region; (c) inversion region with efficient supply of inversion electrons from the N region corresponding to the transistor *C–V* or the quasi-static *C–V*; and (d) inversion region with no supply of inversion electrons (or weak supply by thermal generation) corresponding to the high-frequency capacitor *C–V* case.

the **quasi-static C–V** because Q_{inv} can respond to the AC signal as if the frequency were infinitely low (static case). That would require a ready source of electrons, which can be provided by the N region shown in Fig. 5–17c. PN junctions are always present in an MOS transistor. Therefore, the **MOS transistor C–V** characteristics at all frequencies follow the curve in Fig. 5–16, which is repeated as the upper curve in Fig. 5–18.

What if, as in Fig. 5–17d, the PN junctions are not present? The P-type substrate is an inefficient supplier of electrons. It produces electrons through thermal generation at a very slow rate (for the same reason the diode reverse leakage current is small.) Q_{inv} cannot respond to the AC signal and remains constant at its DC value. Instead, the AC signal causes ϕ_s to oscillate around $2\phi_B$

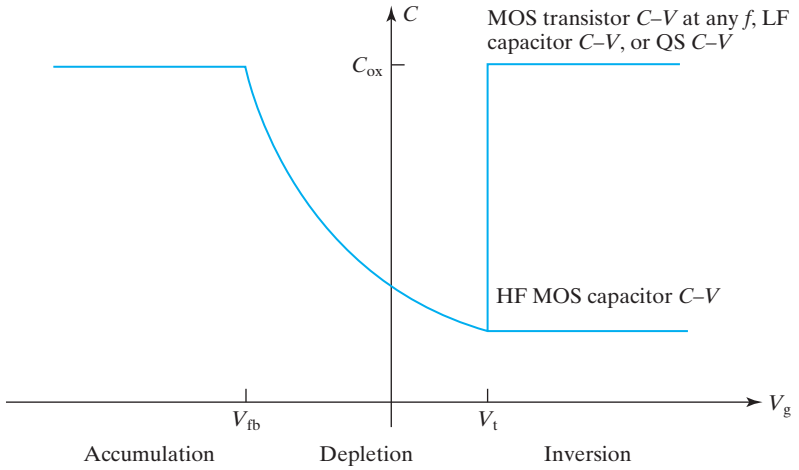

FIGURE 5–18 Two possible MOS C–V characteristics. The difference in the inversion region is explained in Fig. 5–17c and d.

and causes W_{dep} to expand and contract slightly around W_{dmax}. This change of W_{dep} can respond at very high frequencies because it only involves the movement of the abundant majority carriers. Consequently, the AC charge exists at the bottom of the depletion region. The result is a saturation of C at V_t as illustrated by the lower curve in Fig. 5–18. This curve is known as the capacitor C–V or the **high-frequency MOS capacitor *C–V*** (HF *C–V*). The name connotes that, in principle, at a sufficiently low frequency, even the MOS capacitor's C–V would follow the upper curve in Fig. 5–18. Following that reasoning, the upper curve is also known as the **low-frequency *C–V*** (LF *C–V*)**.** In reality, even at a low frequency such as 1 kHz, the C–V of modern high-quality MOS capacitors does not follow the LF *C–V* curve. At yet lower frequencies, the C–V meter is ineffective (the capacitative current is too low) for studying the MOS capacitor. The term *low-frequency C–V* has a historical significance and is still used, but it no longer has a practical significance.

• Measuring the Quasi-Static C–V Using an MOS Capacitor •

There *is* a practical way to obtain the "low frequency" or quasi-static *C–V* (upper branch of Fig. 5–18) using an MOS capacitor without the PN junction. It involves applying a very slow linear-ramp voltage (<0.1V/s) to the gate and measuring I_g with a very sensitive DC ammeter during the ramp. C is calculated from $I_g = C \cdot dV_g/dt$. This technique provides sufficient time for Q_{inv} to respond to the slowly changing V_g. Plotting $I_g/(dV_g/dt)$ vs. V_g produces the QS *C–V* curve shown in Fig. 5–18. This technique becomes impracticable if the gate dielectric has too large a leakage current.

EXAMPLE 5–1 C–V of MOS Capacitor and Transistor

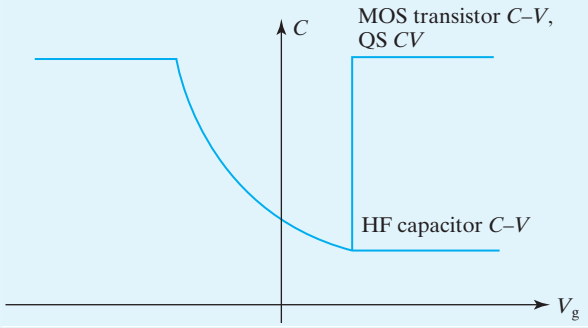

FIGURE 5–19 *C–V* curves of MOS capacitor and transistor.

For each of the following cases, does the QS *C–V* or the HF capacitor *C–V* apply?

(1) MOS transistor, 10 kHz. (Answer: QS *C–V*).
(2) MOS transistor, 100 MHz. (Answer: QS *C–V*).
(3) MOS capacitor, 100 MHz. (Answer: HF capacitor *C–V*).
(4) MOS capacitor, 10 kHz. (Answer: HF capacitor *C–V*).
(5) MOS capacitor, slow V_g ramp. (Answer: QS *C–V*).
(6) MOS transistor, slow V_g ramp. (Answer: QS *C–V*).

5.7 • OXIDE CHARGE—A MODIFICATION TO V_{fb} AND V_t[4] •

The basic MOS theory ignores the possible presence of electric charge in the gate dielectric. Assuming surface charge, Q_{ox} (C/cm^2), exists at the SiO_2–Si interface, the band diagram at the flat-band condition would be modified from Fig. 5–20a to 5–20b.

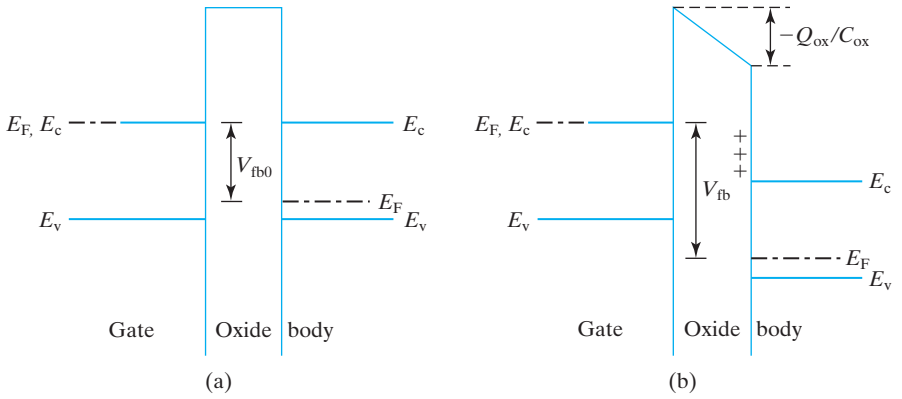

FIGURE 5–20 Flat-band condition (no band bending at body surface) (a) without any oxide charge; (b) with Q_{ox} at the oxide–substrate interface.

[4] This section may be omitted in an accelerated course.

The flat-band voltage in Fig. 5–20a is $\psi_g - \psi_s$ (Section 5.1). In Fig. 5–20b, the oxide charge (assumed to be located at the oxide–substrate interface for simplicity) induces an electric field in the oxide and an oxide voltage, $-Q_{ox}/C_{ox}$. Clearly, V_{fb} in part b is different from the V_{fb0} in part a. Specifically,

$$V_{fb} = V_{fb0} - Q_{ox}/C_{ox} = \psi_g - \psi_s - Q_{ox}/C_{ox} \tag{5.7.1}$$

Because Q_{ox} changes V_{fb}, it also changes V_t through Eq. (5.4.3).

There are several types of **oxide charge**. Positive **fixed oxide charge** is attributed to silicon ions present at the Si–SiO_2 interface. **Mobile oxide charge** is believed to be mostly sodium ions. Mobile ions can be detected by observing V_{fb} and V_t shift under a gate bias at an elevated temperature (e.g., at 200 °C) due to the movement of the ions in the oxide. **Sodium contamination** must be eliminated from the water, chemicals, and containers used in an MOS fabrication line in order to prevent instabilities in V_{fb} and V_t. In addition, significant **interface traps** or **interface states** may be present and they can trap and release electrons and generate noise (see Section 6.15.3) and degrade the subthreshold current of MOSFET (see Section 7.2).

● Reliability ●

More interface states and fixed oxide charge appear after the oxide is subjected to high electric field for some time due to the breaking or rearrangement of chemical bonds. This raises a reliability concern because the threshold voltage and transistor current would change with usage and can potentially cause sensitive circuits to fail. Engineers ensure device reliability by controlling the stress field and improving the MOS interface quality and verifying or projecting the reliability with careful long-term testing.

EXAMPLE 5–2 Interpret the measured V_{fb} dependence on oxide thickness in Fig. 5–21 using Eq. (5.7.1). It is known that the gate electrode is N^+ poly-Si. What can you tell about the capacitors?

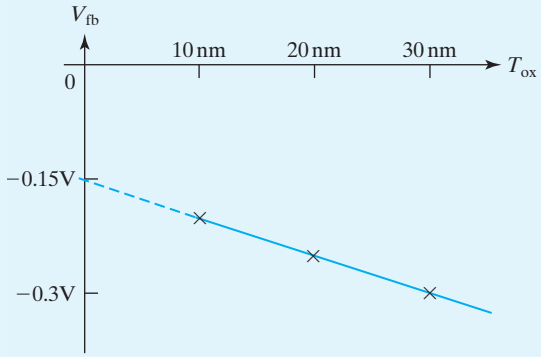

FIGURE 5–21 Measured V_{fb} of three capacitors with different oxide thicknesses.

SOLUTION:

$$V_{fb} = \psi_g - \psi_s - Q_{ox}T_{ox}/\varepsilon_{ox} \tag{5.7.1}$$

Equation (5.7.1) suggests that V_{fb} at $T_{ox} = 0$ is $\psi_g - \psi_s$. Therefore, $\psi_g - \psi_s = -0.15$ V. This is illustrated in Fig. 5–22.

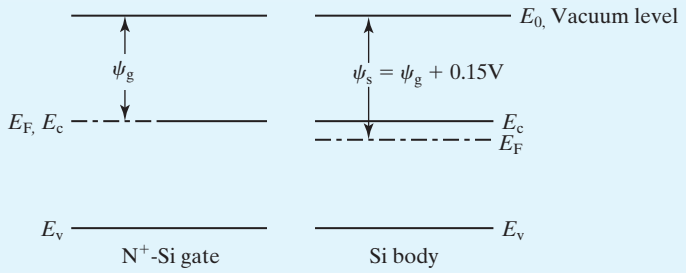

FIGURE 5–22 The relationship between ψ_g and ψ_s.

Because E_F is 0.15 V below E_c, we conclude that the substrate is N-type with

$$N_d = n = N_c e^{-0.15\,eV/kT} \approx 10^{17}\,cm^{-3}$$

Further, Eq. (5.7.1) suggests that

$$Q_{ox} = -\varepsilon_{ox} \times \text{slope of line in Fig. 5–21}$$

$$= -\varepsilon_{ox} \times \frac{-0.15\ V}{30\ nm} = \frac{3.9 \times 8.85 \times 10^{-14} \times 0.15\ V}{300 \times 10^{-8}} = 1.7 \times 10^{-8}\ C/cm^2$$

This corresponds to 1.7×10^{-8} cm$^2 \div q = 9 \times 10^{10}$ cm^2 of positive charge at the interface. A high-quality MOS interface has about 10^{10} cm^2 of charge. Both numbers are small fractions of the number of silicon atoms on a (100) crystal plane, 7×10^{14} cm^{-2}. In this sense, the SiO_2–Si interface is remarkably well-behaved and charge-free.

5.8 • POLY-SI GATE DEPLETION—EFFECTIVE INCREASE IN T_{OX} •

Consider an MOS capacitor with P^+ poly-Si gate and N body. The capacitor is biased into surface inversion. Figure 5–23a shows that the continuity of electric flux requires that the band bends in the gate. This indicates the presence of a thin depletion layer in the gate. Depending on the gate doping concentration and the oxide field, the ***poly-Si gate depletion*** layer thickness, W_{dpoly}, may be 1–2 nm. According to Gauss's Law,

$$W_{dpoly} = \varepsilon_{ox}\mathscr{E}_{ox}/qN_{poly} \tag{5.8.1}$$

Because a depletion layer is present in the gate, one may say that a poly-silicon-gate capacitor is added in series with the oxide capacitor as shown in Fig. 5–23b. The MOS capacitance in the inversion region becomes

$$C = \left(\frac{1}{C_{ox}} + \frac{1}{C_{poly}}\right)^{-1} = \left(\frac{T_{ox}}{\varepsilon_{ox}} + \frac{W_{dpoly}}{\varepsilon_s}\right)^{-1} = \frac{\varepsilon_{ox}}{T_{ox} + W_{dpoly}/3} \tag{5.8.2}$$

This **poly-depletion effect** effectively increases T_{ox} by $W_{dpoly}\varepsilon_{ox}/\varepsilon_s$ or $W_{dpoly}/3$, and can have a significant impact on the C–V curve if T_{ox} is thin. The gate capacitance drops as the capacitor is biased deeper into the inversion region due to increasing poly-depletion as shown in Fig. 5–26. The poly-depletion effect is

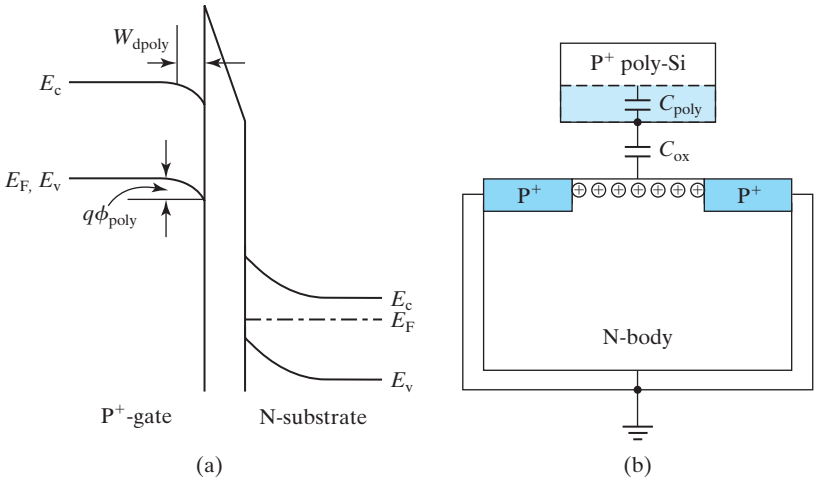

FIGURE 5–23 Poly-gate depletion effect illustrated with (a) the band diagram and (b) series capacitors representation. An N^+ poly-Si gate can also be depleted.

undesirable because a reduced C means reduced Q_{inv}, and reduced transistor current. The solution is to dope the poly-Si heavily. Unfortunately, very heavy doping may cause **dopant penetration** from the gate through the oxide into the substrate. **Poly-SiGe** gate can be doped to a higher concentration, thus improving gate depletion [1]. Poly-gate depletion is eliminated in advanced MOSFET technology by substitution of the poly-gate with a metal gate (see Section 7.4).

The effect of poly-gate depletion on Q_{inv} may be modeled in another way:

$$Q_{inv} = -C_{ox}(V_g - \phi_{poly} - V_t) \tag{5.8.3}$$

Poly-gate depletion effectively reduces V_g by ϕ_{poly}. Even 0.1 V ϕ_{poly} would be highly undesirable when the power-supply voltage (the maximum V_g) is only around 1 V.

EXAMPLE 5–3 Poly-Si Gate Depletion

Assume that V_{ox}, the voltage across a 2 nm thin oxide is –1 V. The P^+ poly-gate doping is $N_{poly} = 8 \times 10^{19}$ cm^3 and substrate N_d is 10^{17}cm^3. Estimate (a) W_{dpoly}, (b) ϕ_{poly}, and (c) V_g.

SOLUTION:

a. Using Eq. (5.8.1),

$$W_{dpoly} = \varepsilon_{ox}\mathscr{E}_{ox}/qN_{poly} = \varepsilon_{ox}V_{ox}/T_{ox}qN_{poly}$$

$$= \frac{3.9 \times 8.85 \times 10^{-14}(\text{F/cm}) \cdot 1\text{V}}{2 \times 10^{-7}\text{cm} \cdot 1.6 \times 10^{-19}\text{C} \cdot 8 \times 10^{19}\text{cm}^{-3}}$$

$$= \frac{34.5 \times 10^{-14}\text{cm}}{256 \times 10^{-8}} = 0.13 \times 10^{-6}\text{cm} = 1.3\,\text{nm}$$

b. W_{dpoly} is related to ϕ_{poly} by the depletion-region model

$$W_{dpoly} = \sqrt{\frac{2\varepsilon_s \phi_{poly}}{qN_{poly}}}$$

$$\phi_{poly} = qN_{poly}W_{dpoly}{}^2/2\varepsilon_s$$

$$= \frac{1.6\times10^{-19}\text{C}\cdot 8\times10^{19}\text{cm}^{-3}\cdot(1.3\times10^{-7}\text{cm})^2}{2\times12\times8.85\times10^{-14}\text{F/cm}}$$

$$= \frac{2.3\times10^{-13}\text{V}}{2.1\times10^{-12}} = 0.11\text{ V}$$

c. Equation (5.2.2) with a ϕ_{poly} term added is

$$V_g = V_{fb} + \phi_{st} + V_{ox} + \phi_{poly}$$

$$V_{fb} = \psi_g - \psi_s = \frac{E_g}{q} - \frac{kT}{q}\ln\frac{N_c}{N_d} = 1.1 - 0.15\text{ V} = 0.95\text{ V}$$

$$V_g = 0.95 - 0.8 - 1 - 0.11\text{ V} = -0.96\text{ V}$$

Using Eq. (5.4.5), $\phi_{st} = -2\phi_B = -2\frac{kT}{q}\ln\frac{N_d}{n_i} = -0.8$

Draw an energy band diagram to confirm the signs of terms in the last equation. The loss of 0.11 V to poly-depletion is a large loss relative to the 0.96 V applied voltage.