2 / MOS Fundamentals

2.1 IDEAL STRUCTURE DEFINITION

As pictured in Fig. 2.1, the MOS-capacitor is a simple two-terminal device composed of a thin (0.01 μm–1.0 μm) SiO_2 layer sandwiched between a silicon substrate and a metallic field plate. The most common field plate materials are aluminum and heavily doped polycrystalline silicon.† A second metallic layer present along the back or bottom side of the

† Heavily doped Si is metallic in nature. Polysilicon gates, used extensively in complex MOS device structures, are deposited by a chemical-vapor process and then heavily *n*- or *p*-doped by either diffusion or ion implantation.

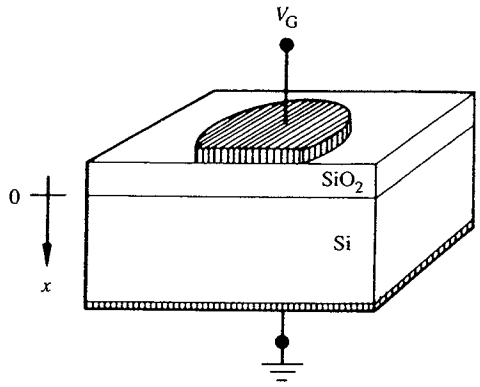


Figure 2.1 The metal–oxide–semiconductor capacitor.

semiconductor provides an electrical contact to the silicon substrate. The terminal connected to the field plate and the field plate itself are referred to as the gate; the silicon terminal, which is normally grounded, is simply called the back or substrate contact.

The ideal MOS structure has the following explicit properties: (1) the metallic gate is sufficiently thick so that it can be considered an equipotential region under a.c. as well as d.c. biasing conditions; (2) the oxide is a *perfect insulator* with *zero current* flowing through the oxide layer under *all* static biasing conditions; (3) there are no charge centers located in the oxide or at the oxide–semiconductor interface; (4) the semiconductor is uniformly doped; (5) the semiconductor is sufficiently thick so that, regardless of the applied gate potential, a field-free region (the so-called Si "bulk") is encountered before reaching the back contact; (6) an *ohmic* contact has been established between the semiconductor and the metal on the back side of the device; (7) the MOS-C is a one-dimensional structure with all variables taken to be a function only of the x-coordinate (see Fig. 2.1); and (8) $\Phi_M = \Phi_S = \chi + (E_c - E_F)_{FB}$. The material parameters appearing in idealization 8 were previously introduced in Section 14.1 and will be reviewed in the next section.

All of the listed idealizations can be approached in practice and the ideal MOS structure is fairly realistic. For example, the resistivity of SiO_2 can be as high as 10^{18} ohm-cm, and the d.c. leakage current through the layer is indeed negligible for typical oxide thicknesses and applied voltages. Moreover, even very thin gates can be considered equipotential regions and ohmic back contacts are quite easy to achieve in practice. Similar statements can be made concerning most of the other idealizations. Special note, however, should be made of idealization 8. The $\Phi_M = \Phi_S$ requirement could be omitted and will in fact be eliminated in Chapter 18. The requirement has only been included at this point to avoid unnecessary complications in the initial description of the static behavior.

2.2 ELECTROSTATICS—MOSTLY QUALITATIVE

2.2.1 Visualization Aids

Energy Band Diagram

The energy band diagram is an indispensable aid in visualizing the internal status of the MOS structure under static biasing conditions. The task at hand is to construct the diagram appropriate for the ideal MOS structure under equilibrium (zero-bias) conditions.

Figure 2.2 shows the surface-included energy band diagrams for the individual components of the MOS structure. In each case the abrupt termination of the diagram in a vertical line designates a surface. The ledge at the top of the vertical line, known as the vacuum level, denotes the minimum energy (E_0) an electron must possess to completely free itself from the material. The energy difference between the vacuum level and the Fermi energy in a metal is known as the metal workfunction, Φ_M. In the semiconductor the height of the surface energy barrier is specified in terms of the electron affinity, χ, the energy difference between the vacuum level and the conduction band edge at the surface. χ is used instead of $E_0 - E_F$ because the latter quantity is not a constant in semiconductors, but varies as a function of doping and band bending near the surface. Note that $(E_c - E_F)_{FB}$ is the energy difference between E_c and E_F in the flat band (FB) or field-free portion of the semiconductor. The remaining component, the insulator, is in essence modeled as an intrinsic wide-gap semiconductor where the surface barrier is again specified in terms of the electron affinity.

The conceptual formation of the MOS zero-bias band diagram from the individual components involves a two-step process. First the metal and semiconductor are brought together until they are a distance x_o apart and the two-component system is allowed to equilibrate. Once the system is in equilibrium the metal and semiconductor Fermi levels

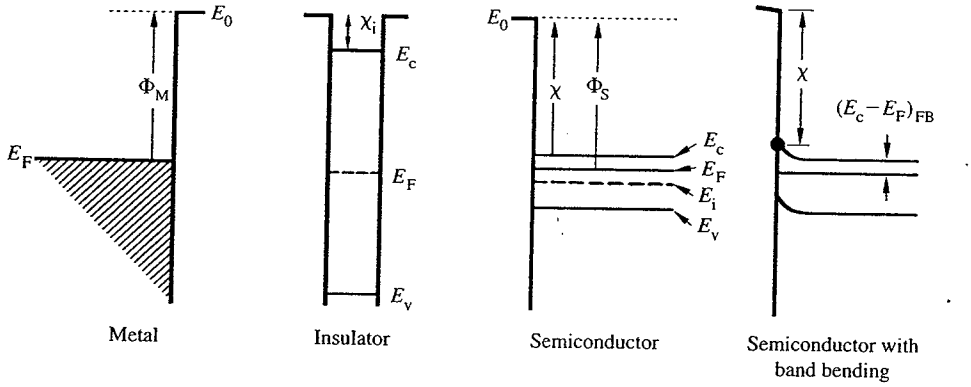


Figure 2.2 Individual energy band diagrams for the metal, insulator, and semiconductor components of the MOS structure. The diagram labeled "semiconductor with band bending" defines $(E_c - E_F)_{FB}$ and shows χ to be invariant with band bending. The value of χ, it should be emphasized, is measured relative to E_c at the semiconductor surface.

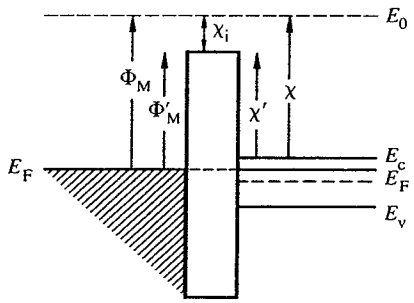


Figure 2.3 Equilibrium energy band diagram for an ideal MOS structure.

must be at the same energy. Moreover, the vacuum levels of the M and S components must also be in alignment because we have specified $\Phi_M = \Phi_S$. The foregoing implies that there are no charges or electric fields anywhere in the metal–gap–semiconductor system. Next the insulator of thickness x_o is inserted into the empty space between the metal and semiconductor components. Given the zero electric field in the x_o gap, the only effect of inserting the insulator is to slightly lower the barrier between the M and S components. Thus the equilibrium energy band diagram for the ideal MOS structure is concluded to be of the form pictured in Fig. 2.3.

Block Charge Diagrams

Complementary in nature to the energy band diagram, block charge diagrams provide information about the approximate charge distribution inside the MOS structure. As just noted in the energy band diagram discussion, there are no charges anywhere inside the ideal MOS structure under equilibrium conditions. However, when a bias is applied to the MOS-C, charge appears within the metal and semiconductor near the metal–oxide and oxide–semiconductor interfaces. A sample block charge diagram is shown in Fig. 2.4.

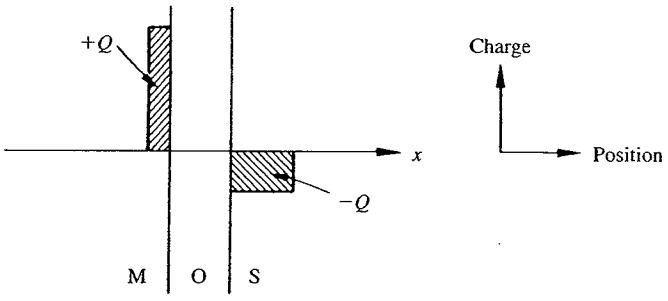


Figure 2.4 Sample block charge diagram.

Note that no attempt is made to represent the exact charge distributions inside the structure. Rather, a squared-off or block approximation is employed and hence the resulting figure is called a block charge diagram. Block charge diagrams are intended to be qualitative in nature; the magnitude and spatial extent of the charges should be interpreted with this fact in mind. Nevertheless, because the electric field is zero in the interior of both the metal and the semiconductor (see idealization 5), the charges within the structure must sum to zero according to Gauss's law. Consequently, in constructing block charge diagrams, the area representing positive charges is always drawn equal to the area representing negative charges.

2.2.2 Effect of an Applied Bias

General Observations

Before examining specific-case situations, it is useful to establish general ground rules as to how one modifies the MOS energy band diagram in response to an applied bias. Assume normal operating conditions where the back side of the MOS-C is grounded and let V_G' be the d.c. bias applied to the gate.†

With $V_G' \neq 0$ we note first of all that the *semiconductor Fermi energy is unaffected by the bias and remains invariant (level on the diagram) as a function of position.* This is a direct consequence of the assumed zero current flow through the structure under all static biasing conditions. In essence, the semiconductor always remains in equilibrium independent of the bias applied to the MOS-C gate. Second, as in a *pn* junction, the applied bias separates the Fermi energies at the two ends of the structure by an amount equal to qV_G'; that is,

$$E_F(\text{metal}) - E_F(\text{semiconductor}) = -qV_G' \qquad (2.1)$$

Conceptually, the metal and semiconductor Fermi levels may be thought of as "handles" connected to the outside world. In applying a bias, one grabs onto the handles and rearranges the relative up-and-down positioning of the Fermi levels. The back contact is grounded and the semiconductor-side handle therefore remains fixed in position. The metal-side handle, on the other hand, is moved downward if $V_G' > 0$ and upward if $V_G' < 0$.

Since the barrier heights are fixed quantities, the movement of the metal Fermi level obviously leads in turn to a distortion in other features of the band diagram. The situation is akin to bending a rubber doll out of shape. Viewed another way, $V_G' \neq 0$ causes potential drops and E_c (E_v) band bending interior to the structure. No band bending occurs, of course, in the metal because it is an equipotential region. In the oxide and semiconductor, however, the energy bands must exhibit an upward slope (increasing E going from the gate toward the back contact) when $V_G' > 0$ and a downward slope when $V_G' < 0$. Moreover, the application of Poisson's equation to the oxide, taken to be an ideal insulator with no carriers or charge centers, yields $d\mathscr{E}_{oxide}/dx = 0$ and therefore $\mathscr{E}_{oxide}$ = constant. Hence, the slope

---

† The prime in V_G', it should be interjected, indicates at a glance that one is referring to the ideal structure; the unprimed symbol V_G is specifically reserved for the gate voltage applied to an actual MOS-C.

of the energy bands in the oxide is a constant—E_c and E_v are linear functions of position. Naturally, band bending in the semiconductor is expected to be somewhat more complex in its functional form, but per idealization 5, must always vanish ($\mathscr{E} \rightarrow 0$) before reaching the back contact.

Specific Biasing Regions

Given the general principles just discussed, it is now a relatively simple matter to describe the internal status of the ideal MOS structure under various static biasing conditions. Taking the Si substrate to be n-type, consider first the application of a positive bias. The application of $V_G' > 0$ lowers E_F in the metal relative to E_F in the semiconductor and causes a positive sloping of the energy bands in both the insulator and semiconductor. The resulting energy band diagram is shown in Fig. 2.5(a). The major conclusion to be derived from Fig. 2.5(a) is that the electron concentration inside the semiconductor, $n = n_i \exp[(E_F - E_i)/kT]$, increases as one approaches the oxide–semiconductor interface. This particular situation, where the majority carrier concentration is greater near the oxide–semiconductor interface than in the bulk of the semiconductor, is known as *accumulation.*

When viewed from a charge standpoint, the application of $V_G' > 0$ places positive charges on the MOS-C gate. To maintain a balance of charge, negatively charged electrons must be drawn toward the semiconductor–insulator interface—the same conclusion established previously by using the energy band diagram. Thus the charge inside the device as a function of position can be approximated as shown in Fig. 2.5(b).

Consider next the application of a *small* negative potential to the MOS-C gate. The application of a small $V_G' < 0$ slightly raises E_F in the metal relative to E_F in the semiconductor and causes a small negative sloping of the energy bands in both the insulator and semiconductor, as displayed in Fig. 2.5(c). From the diagram it is clear that the concentration of majority carrier electrons has been decreased, depleted, in the vicinity of the oxide–semiconductor interface. A similar conclusion results from charge considerations. Setting $V_G' < 0$ places a minus charge on the gate, which in turn repels electrons from the oxide–semiconductor interface and exposes the positively charged donor sites. The approximate charge distribution is therefore as shown in Fig. 2.5(d). This situation, where the electron and hole concentrations at the oxide–semiconductor interface are less than the background doping concentration (N_A or N_D), is known for obvious reasons as *depletion.*

Finally, suppose a larger and larger negative bias is applied to the MOS-C gate. As V_G' is increased negatively from the situation pictured in Fig. 2.5(c), the bands at the semiconductor surface will bend up more and more. The hole concentration at the surface (p_s) will likewise increase systematically from less than n_i when $E_i(\text{surface}) < E_F$, to n_i when $E_i(\text{surface}) = E_F$, to greater than n_i when $E_i(\text{surface})$ exceeds E_F. Eventually, the hole concentration increases to the point shown in Fig. 2.5(e) and (f), where

$$\boxed{E_i(\text{surface}) - E_i(\text{bulk}) = 2[E_F - E_i(\text{bulk})]} \qquad (2.2)$$

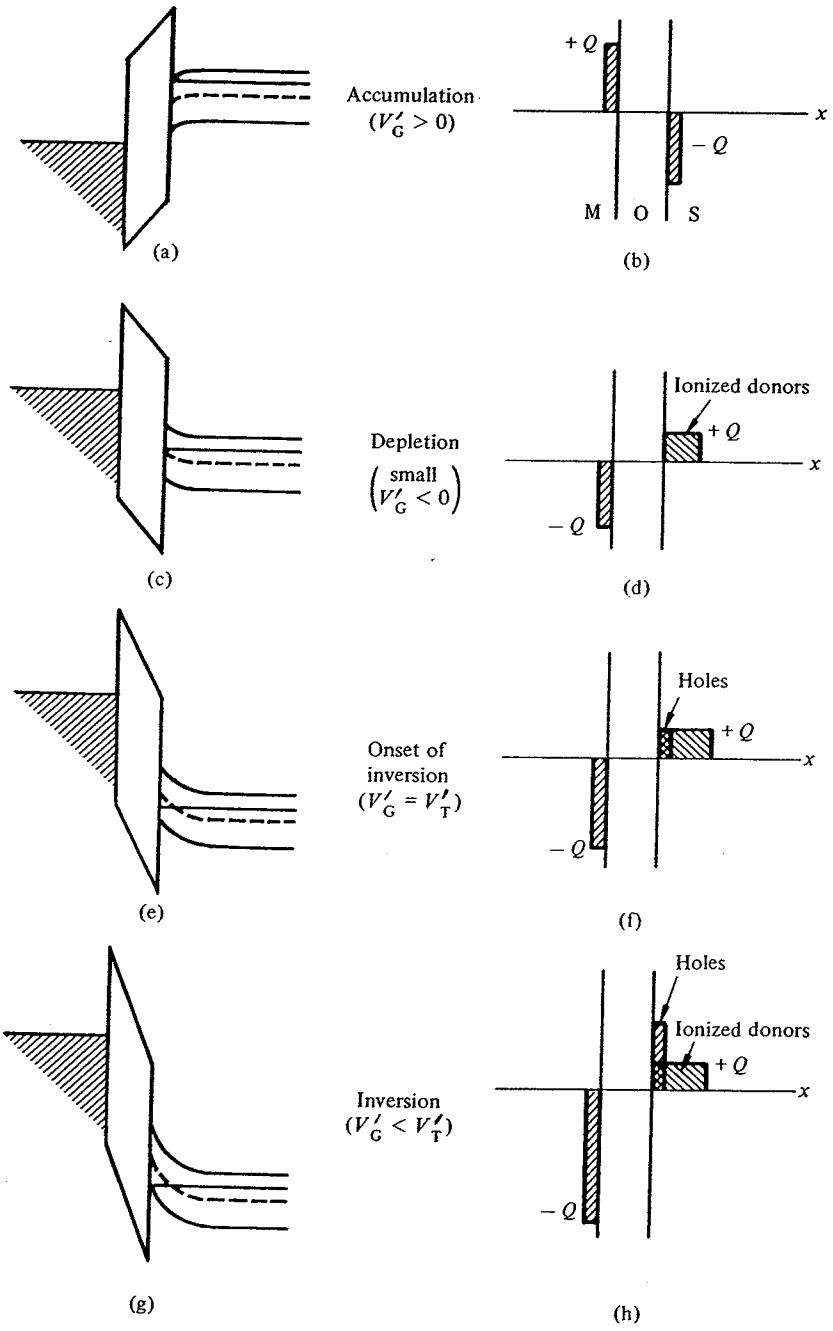


Figure 2.5 Energy band diagrams and corresponding block charge diagrams describing the static state in an ideal *n*-type MOS-capacitor.

and

$$p_s = n_i e^{[E_i(\text{surface}) - E_F]/kT} = n_i e^{[E_F - E_i(\text{bulk})]/kT} = n_{\text{bulk}} = N_D \qquad (2.3)$$

Clearly, when $p_s = N_D$ for the special applied bias $V_G' = V_T'$ the surface is no longer depleted. Moreover, for further increases in negative bias ($V_G' < V_T'$), p_s exceeds $n_{\text{bulk}} = N_D$ and the surface region appears to change in character from n-type to p-type. In accordance with the change in character observation, the $V_G' < V_T'$ situation where the minority carrier concentration at the surface exceeds the bulk majority carrier concentration is referred to as *inversion.* Energy band and block charge diagrams depicting the inversion condition are displayed in Fig. 2.5(g) and (h).

If analogous biasing considerations are performed for an ideal p-type device, the results will be as shown in Fig. 2.5′. It is important to note from this figure that biasing regions in a p-type device are reversed in polarity relative to the voltage regions in an n-type device; that is, accumulation in a p-type device occurs when $V_G' < 0$, and so forth.

In summary, then, one can distinguish three physically distinct biasing regions—accumulation, depletion, and inversion. For an ideal n-type device, accumulation occurs when $V_G' > 0$, depletion when $V_T' < V_G' < 0$, and inversion when $V_G' < V_T'$. The cited voltage polarities are simply reversed for an ideal p-type device. No band bending in the semiconductor or *flat band* at $V_G' = 0$ marks the dividing line between accumulation and depletion. The dividing line at $V_G' = V_T'$ is simply called the depletion–inversion transition

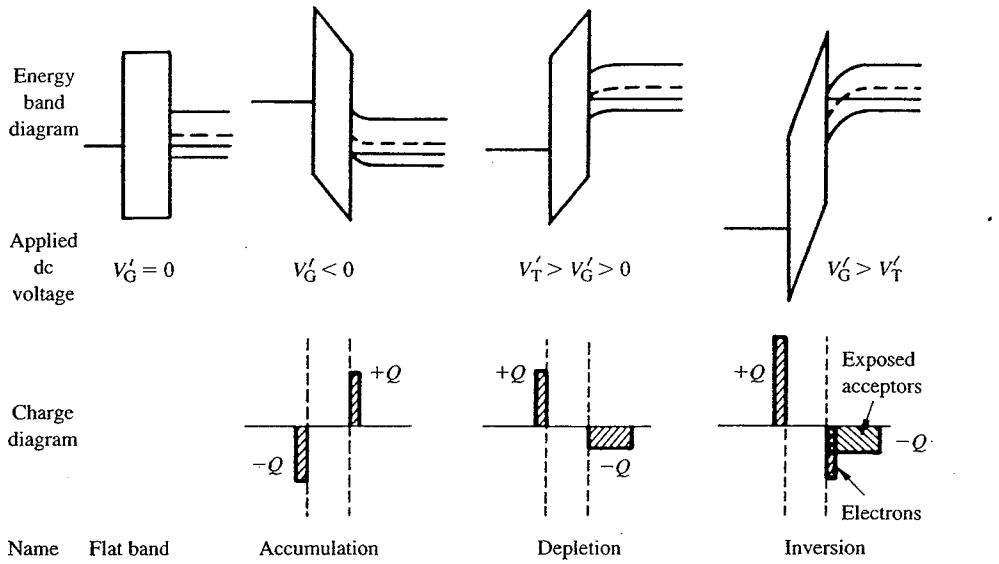


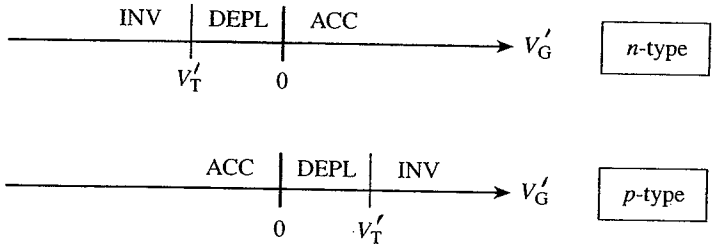
Figure 2.5′ Energy band and block charge diagrams for a p-type device under flat band, accumulation, depletion, and inversion conditions.

point, with Eq. (2.2) quantitatively specifying the onset of inversion for both n- and p-type devices.

Exercise 2.1

P: Construct line plots (with V_G' plotted along the x-axis) that visually identify the voltage ranges corresponding to accumulation, depletion, and inversion in ideal n- and p-type MOS devices.

S: The "plots" shown below are in essence a graphical representation of the word summary given at the end of the preceding section. Note that *acc, depl,* and *inv* are standard abbreviations for accumulation, depletion, and inversion, respectively.



2.3 SEMICONDUCTOR ELECTROSTATICS

2.3.1 Definition of Parameters

The purpose of this section is to establish analytical relationships for the charge density, electric field, and electrostatic potential existing inside the semiconductor component of an MOS-C under static conditions. Like the *pn* junction analysis (Chapter 5, SDF), obtaining a mathematical description of the dc state within the semiconductor is tactically a matter of solving Poisson's equation. Although an approximate solution paralleling that presented in the *pn* junction analysis will be included herein, our initial efforts are directed toward obtaining a solution which is "exact" within the ideal structure framework. An exact solution is possible in the MOS-C case because the semiconductor is always in equilibrium regardless of the applied dc bias. As the reader might suspect, however, the exact formulation is somewhat more involved and therefore requires a certain amount of preparatory development.

We begin by letting x be the depth into the semiconductor as measured from the oxide–semiconductor interface (see Fig. 2.6). Note that, under the assumption the semiconductor is sufficiently thick so that the electric field vanishes in the bulk of the material (idealization #5), it is permissible to treat the semiconductor mathematically as if it extended from $x = 0$ to $x = \infty$. Furthermore, since the electrostatic potential is arbitrary to within a constant, we can choose the potential to be zero in the semiconductor bulk; that is, let $V = 0$ at $x = \infty$.

In solving Poisson's equation one could work with the standard system parameters and variables such as the semiconductor doping (N_A, N_D) and the electrostatic potential V. However, in performing mathematical manipulations and in interpreting results, it is far more convenient to deal in terms of normalized parameters. It is therefore customary to introduce the dimensionless quantities,

$$U_F = [E_i(\text{bulk}) - E_F]/kT \tag{2.4}$$

and

$$U = [E_i(\text{bulk}) - E_i(x)]/kT = V/(kT/q) \tag{2.5}$$

U_F and U are also defined graphically in Fig. 2.6. U_F is called the doping parameter and is directly related to the semiconductor doping concentration. $U(x)$ is the electrostatic potential normalized to kT/q and is usually referred to as simply "the potential" if no

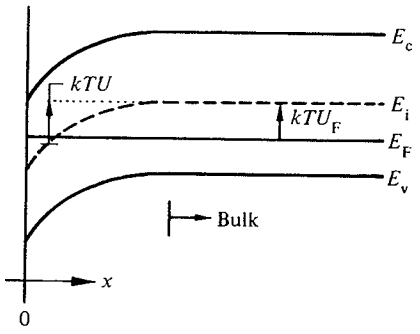


Fig. 2.6 Graphical definition of U and U_F.

ambiguity exists. U evaluated at the oxide-semiconductor interface is given the special symbol, U_S, and is known as the "surface potential." Also note that $U(x \to \infty) = 0$ in agreement with the choice of $V = 0$ at $x = \infty$.

Normalized parameters such as U_F and U are often accepted rather reluctantly. To surmount this problem it is helpful to learn as much as possible about the parameters. Relative to the doping parameter one should know (1) the significance of the sign (plus or minus) associated with U_F, (2) how to calculate U_F for a given impurity concentration, and (3) the range of U_F values normally encountered in practical problems. Since E_F lies above E_i in n-type materials and drops below E_i in p-type materials, an inspection of Eq. (2.4) rapidly reveals $U_F < 0$ given an n-type semiconductor and $U_F > 0$ given a p-type semiconductor. Consequently, the sign of U_F indicates the doping type. Turning next to the computation of U_F, we know

$$n_{bulk} = n_i e^{[E_F - E_i(bulk)]/kT} = N_D \qquad (\text{if } N_D \gg N_A) \tag{2.6a}$$

$$p_{bulk} = n_i e^{[E_i(bulk) - E_F]/kT} = N_A \qquad (\text{if } N_A \gg N_D) \tag{2.6b}$$

Thus, combining Eqs. (2.4) and (2.6) yields

$$U_F = -\ln(N_D/n_i) \qquad \text{for } n\text{-type semiconductor} \tag{2.7a}$$

$$U_F = \ln(N_A/n_i) \qquad \text{for } p\text{-type semiconductor} \tag{2.7b}$$

Finally, semiconductor doping concentrations in MOS devices typically lie somewhere between $10^{14}/cm^3$ and $10^{17}/cm^3$. Employing Eqs. (2.7) with $n_i \simeq 10^{10}/cm^3$ appropriate for Si at room temperature, we therefore find the range of U_F values normally encountered in practical problems is $\boxed{9 \le |U_F| \le 16}$.

The second parameter U is of course a function of both position inside the semiconductor and the voltage applied to the MOS-C gate. Of prime importance is the connection between the U_S value at the oxide–semiconductor interface and the biasing states described in the previous section. Clearly, under flat band conditions $U_S = 0$. Moreover, combining Eqs. (2.2), (2.4), and (2.5), *one concludes* $\boxed{U_S = 2U_F}$ *at the depletion-inversion transition point*. It therefore follows that, in a p-type semiconductor, $U_S < 0$ if the semicon-

ductor is accumulated, $0 < U_S < 2U_F$ if the semiconductor is depleted, and $U_S > 2U_F$ if the semiconductor is inverted. For an n-type semiconductor the inequalities are merely reversed. In other words, knowledge of U_S completely specifies the biasing state inside the semiconductor.

Another item of interest is the range of U_S values normally encountered in practical problems. Although greater excursions are possible, the band bending inside an MOS-C is routinely such that the Fermi level at the oxide-semiconductor interface is confined to band gap energies between E_v and E_c. Thus, given E_v (surface) $\leq E_F \leq E_c$ (surface), and assuming a Si substrate maintained at room temperature, the range of U_S values normally encountered in practice is $\boxed{U_F - 21 \lesssim U_S \lesssim U_F + 21}$. It is left to the reader as an exercise to verify this result. The cited range of U_S values, it should be noted, is also of interest from a theoretical standpoint. When E_F crosses into either the valence band or conduction band at the surface, the surface region becomes decidedly degenerate and degenerate relationships must be employed in calculating the carrier concentrations. Standard quantitative analyses, including the one presented herein, employ nondegenerate relationships and are therefore restricted in validity to the indicated range of U_S values.

In addition to U and U_F, quantitative expressions for the band bending inside of a semiconductor are normally formulated in terms of a special length parameter known as the *intrinsic Debye length*. The Debye length is a characteristic length which was originally introduced in the study of plasmas. (A plasma is a highly ionized gas containing an equal number of positive gas ions and negative electrons.) Whenever a plasma is perturbed by placing a charge in or near it, the mobile species always rearrange so as to shield the plasma proper from the perturbing charge. The Debye length is the shielding distance, or roughly the distance where the electric field emanating from the perturbing charge falls off by a factor of $1/e$. A semiconductor devoid of band bending can be viewed as a type of plasma with its equal number of ionized impurity sites and mobile electrons or holes. The placement of charge near the semiconductor, on the MOS-C gate for example, also causes the mobile species inside the semiconductor to rearrange so as to shield the semiconductor proper from the perturbing charge. The shielding distance or band bending region is again on the order of a Debye length, L_B, where

$$L_B = \left[\frac{K_S \varepsilon_0 kT}{q^2(n_{bulk} + p_{bulk})}\right]^{1/2} \tag{2.8}$$

Although the Debye length characterization applies only to small deviations from flat band, it is convenient to employ the Debye length appropriate for an *intrinsic* material as a normalizing factor in theoretical expressions. The *intrinsic* Debye length, L_D, is obtained from the more general L_B relationship by setting $n_{bulk} = p_{bulk} = n_i$; that is,

$$L_D = \left[\frac{K_S \varepsilon_0 kT}{2q^2 n_i}\right]^{1/2} \tag{2.9}$$

Exercise 2.2

P: (a) Construct line plots (with U_S plotted along the x-axis) that visually identify the surface potential ranges corresponding to accumulation, depletion, and inversion in ideal n-type and p-type MOS devices.

(b) For each of the U_F, U_S parameter sets listed below, indicate the doping type and the specified biasing condition. Also draw the corresponding energy band diagram and block charge diagram that characterize the static state of the ideal MOS system.

(i) $U_F = 12, \quad U_S = 12$

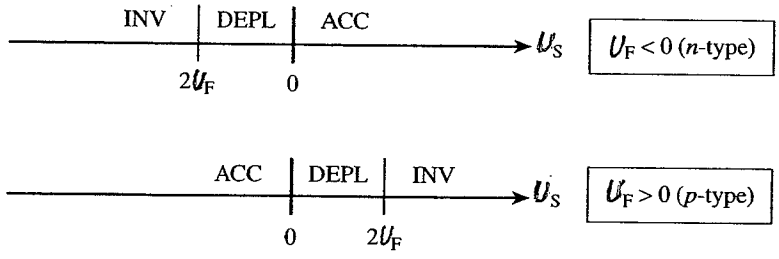
(ii) $U_F = -9, \quad U_S = 3$

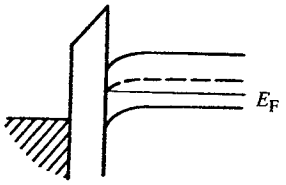
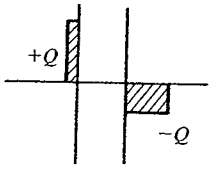
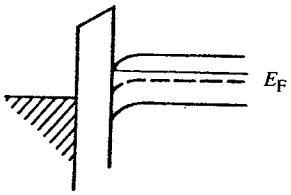
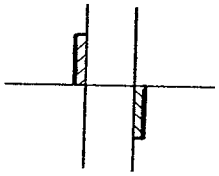
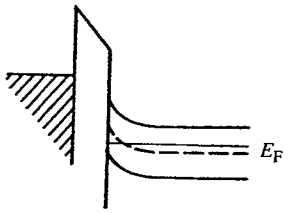
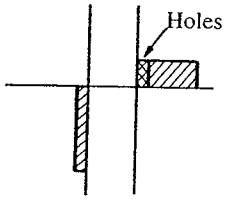
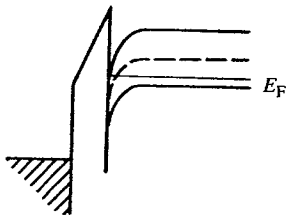
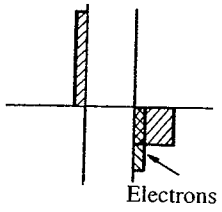
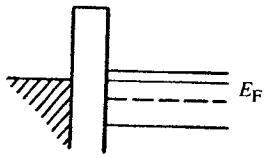
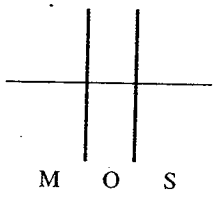
(iii) $U_F = -9, \quad U_S = -18$

(iv) $U_F = 15, \quad U_S = 36$

(v) $U_F = -15, \quad U_S = 0$

S: (a) Converting the discussion in the preceding subsection into a graphical representation yields



(b) Set	Doping	Biasing Condition	Energy Band Diagram	Block Charge Diagram
(i)	*p*	Depletion		
(ii)	*n*	Accumulation		
(iii)	*n*	Depl / Inv Transition		
(iv)	*p*	Inversion		
(v)	*n*	Flat Band		

2.3.2 Exact Solution

Expressions for the charge density, electric field, and potential as a function of position inside the semiconductor are obtained by solving Poisson's equation. Since the MOS-C is assumed to be a one-dimensional structure (idealization #7), Poisson's equation simplifies to

$$\frac{d\mathscr{E}}{dx} = \frac{\rho}{K_S\varepsilon_0} = \frac{q}{K_S\varepsilon_0}(p - n + N_D - N_A) \tag{2.10}$$

Maneuvering to recast the equation in a form more amenable to solution, we note

$$\mathscr{E} = \frac{1}{q}\frac{dE_i(x)}{dx} = -\frac{kT}{q}\frac{dU}{dx} \tag{2.11}$$

The first equality in Eq. (2.11) is a restatement of Eq. (3.15) in SDF. The second equality follows from the Eq. (2.5) definition of U and the fact that $dE_i(\text{bulk})/dx = 0$. In a similar vein we can write

$$p = n_i e^{[E_i(x)-E_F]/kT} = n_i e^{U_F - U(x)} \tag{2.12a}$$

$$n = n_i e^{[E_F - E_i(x)]/kT} = n_i e^{U(x) - U_F} \tag{2.12b}$$

Moreover, since $\rho = 0$ and $U = 0$ in the semiconductor bulk,

$$0 = p_{\text{bulk}} - n_{\text{bulk}} + N_D - N_A = n_i e^{U_F} - n_i e^{-U_F} + N_D - N_A \tag{2.13}$$

or

$$N_D - N_A = n_i(e^{-U_F} - e^{U_F}) \tag{2.14}$$

Substituting the foregoing $\mathscr{E}$, p, n, and $N_D - N_A$ expressions into Eq. (2.10) yields

$$\boxed{\rho = qn_i(e^{U_F - U} - e^{U - U_F} + e^{-U_F} - e^{U_F})} \tag{2.15}$$

and

$$\frac{d^2U}{dx^2} = \left(\frac{q^2 n_i}{K_S\varepsilon_0 kT}\right)(e^{U-U_F} - e^{U_F-U} + e^{U_F} - e^{-U_F}) \tag{2.16}$$

or, in terms of the intrinsic Debye length,

$$\frac{d^2U}{dx^2} = \frac{1}{2L_D^2}(e^{U-U_F} - e^{U_F-U} + e^{U_F} - e^{-U_F}) \tag{2.17}$$

We turn next to the main task at hand. Poisson's equation, Eq. (2.17), is to be solved subject to the boundary conditions:

$$\mathscr{E} = 0 \quad \text{or} \quad \frac{dU}{dx} = 0, \qquad \text{at } x = \infty \tag{2.18a}$$

and

$$U = U_S, \qquad \text{at } x = 0 \tag{2.18b}$$

Multiplying both sides of Eq. (2.17) by dU/dx, integrating from $x = \infty$ to an arbitrary

point x, and making use of the Eq. (2.18a) boundary condition, we quickly obtain

$$\mathscr{E}^2 = \left(\frac{kT/q}{L_D}\right)^2 [e^{U_F}(e^{-U} + U - 1) + e^{-U_F}(e^{U} - U - 1)] \tag{2.19}$$

Equation (2.19) is of the form $y^2 = a^2$, which has two roots, $y = a$ and $y = -a$. As can be deduced by inspection from the energy band diagram, we must have $\mathscr{E} > 0$ when $U > 0$ and $\mathscr{E} < 0$ when $U < 0$. Since the right-hand side of Eq. (2.19) is always positive ($a \geq 0$), the proper polarity for the electric field is obviously obtained by choosing the positive root when $U > 0$ and the negative root when $U < 0$. We can therefore write

$$\mathscr{E} = -\frac{kT}{q}\frac{dU}{dx} = \hat{U}_S \frac{kT}{q}\frac{F(U, U_F)}{L_D} \tag{2.20}$$

where

$$F(U, U_F) \equiv [e^{U_F}(e^{-U} + U - 1) + e^{-U_F}(e^{U} - U - 1)]^{1/2} \tag{2.21}$$

and

$$\hat{U}_S = \begin{cases} +1 & \text{if } U_S > 0 \\ -1 & \text{if } U_S < 0 \end{cases} \tag{2.22}$$

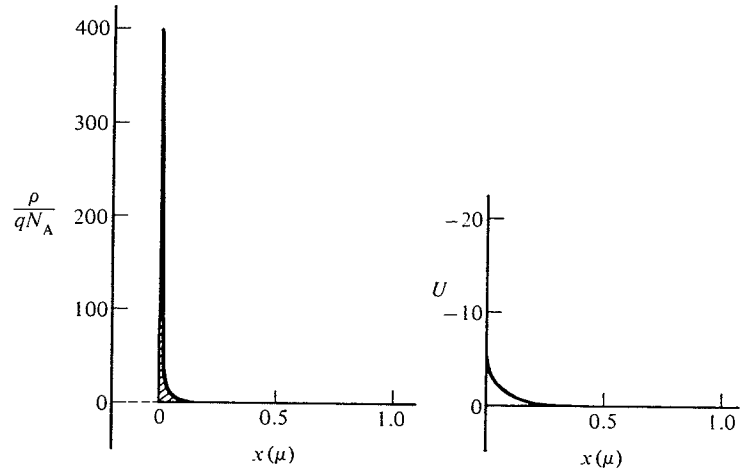
To complete the solution, one separates the U and x variables in Eq. (2.20) and, making use of the Eq. (2.18b) boundary condition, integrates from $x = 0$ to an arbitrary point x. The end result is Eq. (2.23),

$$\hat{U}_S \int_U^{U_S} \frac{dU'}{F(U', U_F)} = \frac{x}{L_D} \tag{2.23}$$

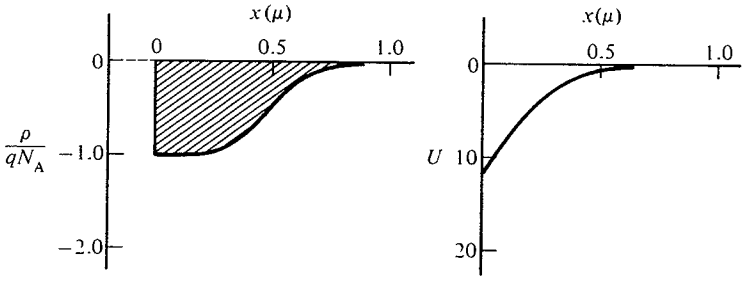
Although not in a totally explicit form, Eqs. (2.15), (2.20), and (2.23) collectively constitute an exact solution for the electrostatic variables. For a given U_S, numerical techniques can be used to compute U as a function of x from Eq. (2.23). Once U as a function of x is established, direct substitution into Eqs. (2.15) and (2.20) yields ρ and $\mathscr{E}$ as a function of x. Sample plots of U versus x and ρ versus x obtained in the manner just described are presented in Fig. 2.7.

2.3.3 Delta-Depletion Solution

A closed-form though approximate solution for the charge density, electric field, and potential interior to the semiconductor can be established by utilizing the depletion approximation first introduced in the *pn* junction analysis. Because additional approximations based in large part on the nature of the exact solution are also employed in the



(a) Accumulation ($U_S = -6$)

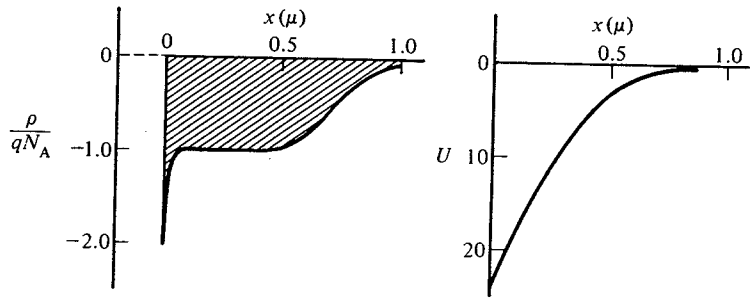


(b) Middle of depletion ($U_S = U_{i'} = 12$)

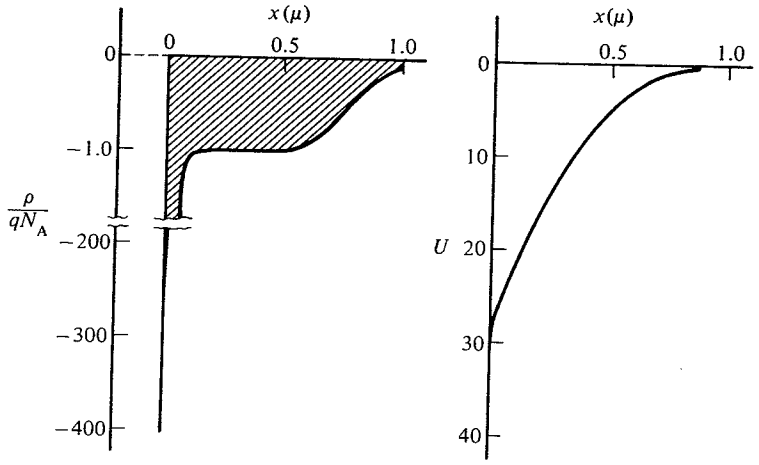
Fig. 2.7 Exact solution for the charge density and potential inside a semiconductor as a function of position assuming $U_F = 12$, $T = 23°C$ and an L_D appropriate for silicon ($L_D = 3.11 \times 10^{-3}$cm). (a) Accumulation ($U_S = 6$) and (b) middle of depletion ($U_S = U_F = 12$). (See next page for continuation of Fig. 2.7(c, d).)

formulation, the approximate relationships for the electrostatic variables are herein referred to collectively as the *delta-depletion solution*.

Since the nature of the exact solution plays a role in the approximate formulation, it is reasonable to spend a few moments examining the plots presented in Fig. 2.7. First of all, note the general correlation between the Fig. 2.7 plots and the semiconductor portion of the diagrams sketched in Fig. 2.5. Next, specifically note that *the charge associated with majority carrier accumulation and minority carrier inversion resides in an extremely narrow portion of the semiconductor immediately adjacent to the oxide–semiconductor interface*. By comparison, the depleted portion of the semicon-



(c) Onset of inversion ($U_S = 2U_F = 24$)



(d) Deep into inversion ($U_S = 2U_F + 6 = 30$)

Fig. 2.7 (continued). (c) Onset of inversion ($U_S = 2U_F = 24$), and (d) heavily inverted ($U_S = 2U_F + 6 = 30$). The ρ-diagrams were drawn on a linear scale and the $+ U$ axes oriented downward to enhance the correlation with the diagrams sketched in Fig. 2.5.

ductor under moderate depletion biasing extends much deeper into the semiconductor. Moreover, in comparing the depleted semiconductor regions when $U_S = U_F$ (middle of depletion, $U_S = 2U_F$ (onset of inversion), and $U_S = 2U_F + 6$ (inversion), we find *the depletion width increases substantially with increased depletion biasing, but increases only slightly once the semiconductor inverts*. This, we might interject, occurs because the highly peaked inversion charge near the oxide–semiconductor interface is nearly sufficient, in itself, to shield the interior of the semiconductor from any additional charge placed on the MOS-C gate.

In the delta-depletion formulation the very narrow extent of the accumulation and inversion layers is used as justification for approximately representing these layers by

δ-functions of charge located at $x = 0$, the oxide–semiconductor interface. Naturally, invoking the standard depletion approximation means the depleted region existing under depletion and inversion biases is also taken to be terminated abruptly a distance $x = W$ into the semiconductor. However, reflecting the observation that the depletion width increases only slightly once the semiconductor inverts, W is fixed at W_T for all inversion biases, where W_T is the depletion width at the onset of inversion ($U_S = 2U_F$).

The ramifications of the above approximations are as follows: Under accumulation biasing in the delta-depletion formulation the majority carriers pile up in a δ-function distribution at the oxide–semiconductor interface; $\rho = 0$, $\mathscr{E} = 0$, and $U = 0$ for all $x > 0$. In depletion W progressively increases with increased depletion biasing until $W = W_T$; ρ, $\mathscr{E}$, and V for a given depletion bias are computed using the standard depletion approximation. In inversion, minority carriers pile up in a δ-function distribution at the oxide–semiconductor interface, while W, ρ, $\mathscr{E}$, and V for all $x > 0$ remain fixed at their $U_S = 2U_F$ values.

From the foregoing discussion, then, to complete the delta-depletion solution we need only work out expressions for the electrostatic variables when the semiconductor is biased into depletion. Let us perform the required analysis. Invoking the depletion approximation (n and $p \ll N_A$ or N_D), we can write

$$\rho = q(N_D - N_A) \qquad (0 \le x \le W) \tag{2.24}$$

and

$$\frac{d\mathscr{E}}{dx} = -\frac{d^2V}{dx^2} = \frac{q(N_D - N_A)}{K_S\varepsilon_0} \qquad (0 \le x \le W) \tag{2.25}$$

The straightforward integration of Eq. (2.25) employing $\mathscr{E} = 0$ and $V = 0$ at $x = W$ yields

$$\mathscr{E}(x) = \frac{q(N_A - N_D)}{K_S\varepsilon_0}(W - x) \qquad (0 \le x \le W) \tag{2.26}$$

and

$$V(x) = \frac{q(N_A - N_D)}{2K_S\varepsilon_0}(W - x)^2 \qquad (0 \le x \le W) \tag{2.27}$$

The remaining unknown, W, is determined by applying the boundary condition $V = (kT/q)U_S$ at $x = 0$. Thus

$$\frac{kT}{q}U_S = \frac{q(N_A - N_D)}{2K_S\varepsilon_0}W^2 \tag{2.28}$$

and

$$W = \left[\frac{2K_S\varepsilon_0}{q(N_A - N_D)}\frac{kT}{q}U_S\right]^{1/2} \tag{2.29}$$

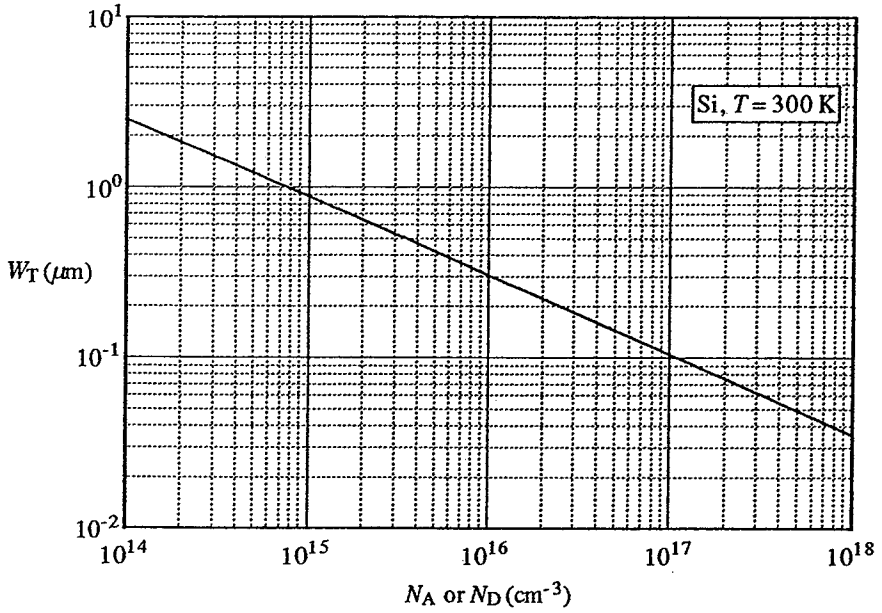


Figure 2.8 Doping dependence of the maximum equilibrium depletion width inside silicon devices maintained at 300 K.

Since $W = W_T$ when $U_S = 2U_F$, we also conclude

$$W_T = \left[\frac{2K_S\varepsilon_0}{q(N_A - N_D)}\frac{kT}{q}(2U_F)\right]^{1/2} \tag{2.30}$$

A plot of W_T versus doping covering the range $9 \lesssim |U_F| \lesssim 16$ is displayed in Fig. 2.8.

The delta-depletion solution just completed is clearly a relatively gross, first-order theory compared to the exact solution. However, because of its greater simplicity and the direct analogies possible with *pn* junction theory, the approximate formulation is widely employed in the analysis of MOS devices. Quite often the reader will encounter *both* solution approaches in a given problem. An initial analysis based on the delta-depletion formulation usually provides closed-form results that can be readily interpreted. Greater precision is obtained by performing a subsequent analysis based on the exact formulation.

2.4 GATE VOLTAGE RELATIONSHIP

Throughout the discussion of semiconductor electrostatics the biasing state was described in terms of the semiconductor surface potential, U_S. Results formulated in this manner are dependent only on the properties of the semiconductor. U_S, however, is an *internal* system constraint or boundary condition. It is the *externally* applied gate potential, V'_G, which is subject to direct control. Thus, if the results of the previous section are to be utilized in practical problems, an expression relating V'_G and U_S must be established. This section is devoted to deriving the required relationship.

We begin by noting that V'_G in the ideal structure is dropped partly across the oxide and partly across the semiconductor or, symbolically,

$$V'_G = \Delta V_{semi} + \Delta V_{ox} \tag{2.31}$$

Because $V = 0$ in the semiconductor bulk, however, the voltage drop across the semiconductor is simply

$$\Delta V_{semi} = V(x = 0) = \frac{kT}{q} U_S \tag{2.32}$$

The task of developing a relationship between V'_G and U_S is therefore reduced to the problem of expressing ΔV_{ox} in terms of U_S.

As stated previously (Section 2.2), in an ideal insulator with no carriers or charge centers

$$\frac{d\mathscr{E}_{ox}}{dx} = 0 \tag{2.33}$$

and

$$\mathscr{E}_{ox} = -\frac{dV_{ox}}{dx} = \text{constant} \tag{2.34}$$

Therefore

$$\Delta V_{ox} = \int_{-x_o}^{0} \mathscr{E}_{ox}\, dx = x_o \mathscr{E}_{ox} \tag{2.35}$$

where x_o is the oxide thickness. The next step is to relate $\mathscr{E}_{ox}$ to the electric field in the semiconductor. The well-known boundary condition on the fields normal to an interface between two dissimilar materials requires

$$(D_{semi} - D_{ox})|_{O\text{-}S\ interface} = Q_{O\text{-}S} \tag{2.36}$$

where $D = \varepsilon\mathscr{E}$ is the dielectric displacement and $Q_{O\text{-}S}$ is the surface center charge/unit area located at the interface. Since $Q_{O\text{-}S} = 0$ in the idealized structure (idealization #3),*

$$D_{ox} = D_{semi|x=0} \tag{2.37}$$

$$\mathscr{E}_{ox} = \frac{K_S}{K_O} \mathscr{E}_S \tag{2.38}$$

and

$$\Delta V_{ox} = \frac{K_S x_o}{K_O} \mathscr{E}_S = x'_o \mathscr{E}_S \tag{2.39}$$

---

*The development here is exact. If the delta-depletion formulation is invoked, the δ-function layers of carrier charge at the O-S interface constitute an effective "$Q_{O\text{-}S}$" under accumulation and inversion conditions.

where

$$x_o' = \frac{K_S x_o}{K_O} \tag{2.40}$$

K_S is the semiconductor dielectric constant; K_O, the oxide dielectric constant; and $\mathscr{E}_S$, the electric field in the semiconductor at the oxide–semiconductor interface.

Finally, if Eqs. (2.32) and (2.39) are substituted into Eq. (2.31), we obtain

$$V_G' = \frac{kT}{q} U_S + x_o' \mathscr{E}_S \tag{2.41}$$

or, making use of Eq. (2.20),

$$\boxed{V_G' = \frac{kT}{q}\left[U_S + \hat{U}_S \frac{x_o'}{L_D} F(U_S, U_F)\right]} \tag{2.42}$$

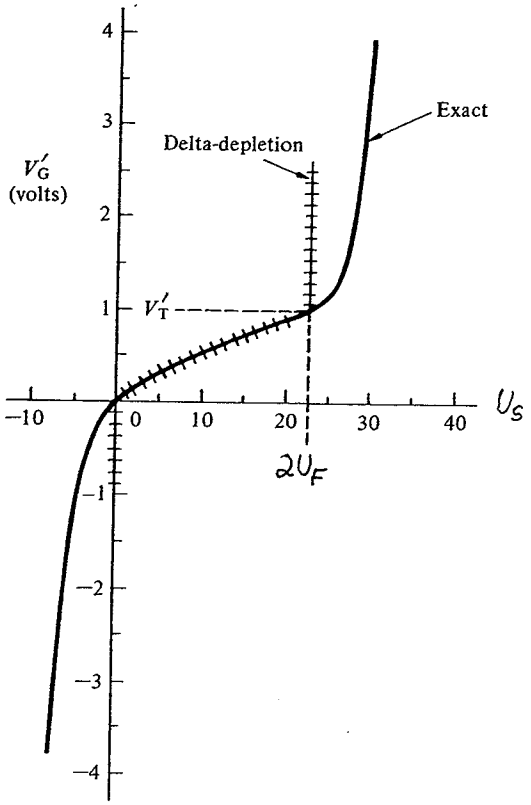


Figure 2.9 Typical interrelationship between the applied gate voltage and the semiconductor surface potential; +++++++ delta-depletion solution, ——— exact solution (x_o = 0.1 μm, $N_A = 10^{15}/\text{cm}^3$, T = 300 K).

The $V'_G - U_S$ dependence calculated from Eq. (2.42) employing a typical set of device parameters is displayed in Fig. 2.9. The figure nicely illustrates certain important features of the gate voltage relationship. For one, U_S is a rather rapidly varying function of V'_G when the device is depletion biased. However, when the semiconductor is accumulated ($U_S < 0$) or inverted ($U_S > 2U_F$), it takes a large change in gate voltage to produce a small change in U_S. This implies the gate voltage divides proportionally between the oxide and the semiconductor under depletion biasing. Under accumulation and inversion biasing, on the other hand, changes in the applied potential are dropped almost totally across the oxide. Also note that the depletion bias region extends only from $V'_G = 0$ to $V'_G \cong 1$ V. Since the character of the semiconductor changes drastically in progressing from one side of the depletion bias region to the other, we therefore anticipate a significant variation in the electrical characteristics over a rather narrow range of voltages.

2.5 SUMMARY AND CONCLUDING COMMENTS

The statics of the ideal MOS structure was the main concern of this chapter. Care was taken to clearly define what was envisioned as the ideal structure and, wherever possible, the specific use of a particular idealization was noted. In a subsequent chapter some of the idealizations will be removed and the ensuing perturbations on the device characteristics will be fully examined. The ideal structure, therefore, serves as a reference plane, a foundation for understanding and analyzing the more complex behavior of the real MOS structure.

Qualitatively, MOS statics has been envisioned in terms of the energy band diagram and the block charge diagrams. These visualization aids are, of course, not confined to this single application, but will be utilized again and again in later work. The terms accumulation, flat band, depletion, and inversion were given a physical interpretation using the aforementioned diagrams. Accumulation corresponds to the pile-up of majority carriers at the oxide–semiconductor interface; flat band, to no bending of the semiconductor bands, or equivalently, to no charge in the semiconductor; depletion, to the repulsion of majority carriers from the interface leaving behind an uncompensated impurity-ion charge; and inversion, to the pile-up of minority carriers at the oxide–semiconductor interface. The reader should be able to associate an energy band diagram and block charge diagram with each of these physical situations.

The quantitative formulation of MOS statics is a simple matter of solving Poisson's equation. Two solutions for the semiconductor part of the structure were presented herein: a first-order solution based on the depletion approximation and an exact solution based only on the assumptions of nondegeneracy, a constant doping profile, and a semiconductor of sufficient thickness so that $\mathscr{E} \to 0$ as $x \to$ back contact. At first glance the exact solution is complicated by the introduction of the normalized potential, U, and the normalized doping parameter, U_F. However, once these parameters are "digested," the mathematical convenience derived from their use becomes obvious. Knowledge of the parameters general properties greatly aids the "digestion" process. One should specifically

remember that, for typical silicon substrates at room temperature, $9 \lesssim |U_F| \lesssim 16$ and $U_F - 21 \lesssim U_S \lesssim U_F + 21$. Furthermore, $U_S = 0$ corresponds to the dividing line between accumulation and depletion while $U_S = 2U_F$ marks the end of depletion and the start of inversion. Finally, it should be remembered that a one-to-one relationship exists between U_S, an internal system constraint, and V'_G, the external gate voltage applied to the ideal structure.

PROBLEMS

2.1 For the U_F, U_S parameter sets listed below first indicate the specified biasing condition and then draw the energy band diagram and block charge diagram which characterizes the static state of the system. Assume the MOS structure to be ideal.

(a) $U_F = 18$, $U_S = 9$;

(b) $U_F = -12$, $U_S = 0$;

(c) $U_F = 12$, $U_S = 24$;

(d) $U_F = -15$, $U_S = 3$;

(e) $U_F = 9$, $U_S = 21$.

2.2 Given Si maintained at $T = 300K$ with a donor doping of $N_D = 10^{15}/cm^3$, compute:

(a) L_D

(b) L_B

(c) U_F

(d) $\mathscr{E}_S$ when $U_S = U_F$ (exact value)

(e) $\mathscr{E}_S$ when $U_S = 2U_F$ (exact value)

(f) W at $U_S = U_F$

(g) W_T

2.3 In this problem we wish to verify the text statement that the range of U_S values is normally restricted to $U_F - 21 \lesssim U_S \lesssim U_F + 21$ for Si substrates maintained at room temperature.

(a) Referring to Fig. 2.6, draw the semiconductor energy band diagrams corresponding to the situations where (i) $E_F = E_c$ at $x = 0$ and (ii) $E_F = E_v$ at $x = 0$. Indicate kTU_S and kTU_F on your diagrams.

(b) Noting $E_c - E_i \simeq E_i - E_v \simeq E_G/2$, show that:

(i) $U_S \simeq U_F + E_G/2kT$, if $E_F = E_c$ at $x = 0$;

(ii) $U_S \simeq U_F - E_G/2kT$, if $E_F = E_v$ at $x = 0$;

(iii) $U_F - E_G/2kT \lesssim U_S \lesssim U_F + E_G/2kT$, if E_F is confined to band gap energies between E_v and E_c.

(c) At $T = 300K$ the Si band gap is 1.12 eV and $kT = 0.0259$ eV. Specialize the (iii) result in part (b) to room temperature ($T \simeq 300K$). Is your result compatible with the U_S limits given in the text? (*Note*: The text limits were chosen to be on the "safe" side to allow for variations in the meaning of "room temperature.")

2.4 Let us examine Fig. 2.7, particularly Fig. 2.7(c), more closely.

(a) Draw the block charge diagram describing the charge situation inside an ideal p-bulk MOS-C biased at the onset of inversion.

(b) Is your part (a) diagram in agreement with the plot of ρ/qN_A versus x in Fig. 2.7(c)? Explain why the ρ/qN_A plot has a spike-like nature near $x = 0$ and shows a value of $\rho/qN_A = -2$ at $x = 0$.

(c) Noting that $U_F = 12$ was assumed in constructing Fig. 2.7, determine the W_T predicted by the delta-depletion theory. Is the W_T obtained from the delta-depletion theory consistent with the charge density plot in Fig. 2.7(c)?

2.5 (a) Construct a computer program (possibly employing a hand-held calculator) which will automatically give V'_G versus U_S based on Eq. (2.42) for U_S stepped in single-unit values between the limits $U_F - 21 \leq U_S \leq U_F + 21$. Only U_F and x_o are to be considered input variables. Let $T = 300K$ ($kT/q = 0.0259$ V and $L_D = 2.91 \times 10^{-3}$ cm).

(b) Setting $x_o = 0.1\ \mu$, use your program to compute V'_G versus U_S for $N_A = 10^{15}/cm^3$. Check your results against Fig. 2.9.

2.6 The exact solution approach presented in Section 2.3 can also be applied to obtain the *exact* electric field and electrostatic potential variation in a *pn* step junction under *equilibrium* conditions.

(a) Assuming a nondegenerately doped step junction, develop a set of equations that constitutes an exact solution for the electrostatic variables ($\rho, \mathscr{E}, V$) inside the *pn* junction under equilibrium conditions. In the development you will find it convenient to introduce the normalized potentials: $U_{BI} \equiv \ln(N_A N_D/n_i^2)$; $U_{FP} \equiv \ln(N_A/n_i)$; and $U_{FN} \equiv -\ln(N_D/n_i)$. Set up $x = 0$ at the metallurgical junction, take $U = 0$ at $x = -\infty$ on the p-side of the junction, and let U at $x = 0$ be U_0.

(b) The exact solution approach of Section 2.3 *cannot* be employed to obtain the electrostatic variables inside a *pn* junction when $V_A \neq 0$. Why is it the solution approach is valid for any applied bias in an MOS-C and not in a *pn* junction?

2.7 Construct a W_T versus doping plot similar to Fig. 2.8 that is appropriate for GaAs. Assume $T = 300$ K; $K_S = 12.85$.

2.8 The energy band diagram for an ideal $x_o = 0.2\ \mu$m MOS-C operated at $T = 300$ K is sketched in Fig. P2.8. Note that the applied gate voltage causes band bending in the semiconductor such that $E_F = E_i$ at the Si–SiO_2 interface. Invoke the delta-depletion approximation as required in answering the questions that follow.

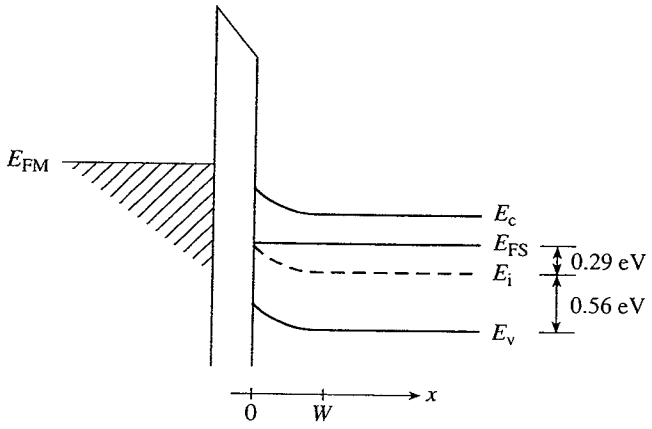


Figure P2.8

(a) Sketch the electrostatic potential (V) inside the semiconductor as a function of position.

(b) Roughly sketch the electric field ($\mathscr{E}$) inside the oxide and semiconductor as a function of position.

(c) Do equilibrium conditions prevail *inside the semiconductor*? Explain.

(d) Roughly sketch the electron concentration versus position inside the semiconductor.

(e) What is the electron concentration at the Si–SiO_2 interface?

(f) $N_D = ?$

(g) $U_S = ?$

(h) $V_G' = ?$

(i) What is the voltage drop (ΔV_{ox}) across the oxide?

2.9 Figure P2.9 is a dimensioned energy band diagram for an ideal MOS-C operated at $T = 300$ K with $V_G' \neq 0$. Note that $E_F = E_i$ at the Si–SiO_2 interface.

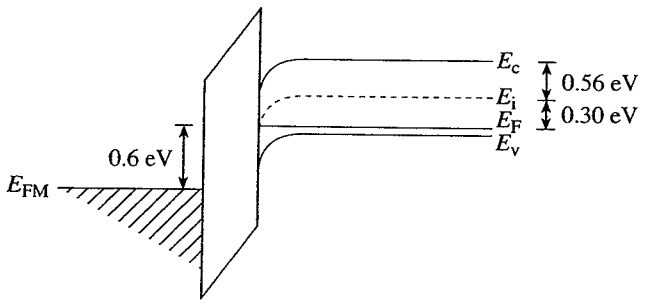


Figure P2.9

(a) Do equilibrium conditions prevail *inside the semiconductor*?

(b) $U_F = ?$

(c) $U_S = ?$

(d) $V_G' = ?$

(e) $x_o = ?$

(f) Draw the block charge diagram corresponding to the state pictured in the energy band diagram. For reference purposes, include the maximum equilibrium depletion width, W_T, on your diagram.