Chapter 8

The Si surface and the Metal-Oxide-Semiconductor Structure

Outline

This chapter deals with what is perhaps the most important material system in modern microelectronics: the metal-oxide-semiconductor structure or MOS. The MOS structure is at the heart of integrated microelectronics. A basic CMOS inverter is made out of two metal-oxide-semiconductor field-effect transistors (MOSFETs) in which carrier transport is controlled by the field-effect action of a MOS structure. Hundreds of millions of them are now integrated together on a single chip. MOS capacitors are also the storage element of Dynamic Random-Access Memories (DRAM). Billions of them are routinely integrated on a memory chip. MOS structures also appear when interconnect metal lines run over an oxide-covered semiconductor surface. In a typical modern microprocessor, there are tens of meters worth of them.

In spite of its seemingly restrictive name, the MOS structure is to be viewed here as a generic sandwich of a highly-conducting material on top of an insulator, all on top of a semiconductor (a better term might be MIS, for Metal-Insulator-Semiconductor). The treatment of the MOS structure presented in this chapter can be easily extended to heavily-doped polysilicon gates which are the preferred choice in modern CMOS. Similarly, insulators other than silicon dioxide are widely used in modern microelectronics. A prominent example is Silicon Nitride (Si_3N_4). With some precautions, the MOS theory can be used to treat MIS structures based on semiconductors other than Si and in which the "insulator" is actually a wide bandgap semiconductor. A prominent example is the metal/AlGaAs/InGaAs heterostructure at the heart of modern High Electron Mobility Transistors (HEMTs).

The goal of this chapter is to understand the changes brought to the electrostatics of the MOS structure upon the application of a voltage to the metal with respect to the semiconductor. Even though current cannot flow through the insulator, the charge distribution in the semiconductor close to the insulator/semiconductor interface is affected in a major way. One can induce a depletion region, accumulate majority carriers, or all together "invert" the semiconductor type and create a thin layer of minority carriers. This "inversion layer" constitutes the conducting channel in most MOSFETs.

An additional goal of this chapter is to continue building up the tool kit for studying semiconductor devices by introducing the Poisson-Boltzmann formulation. This is a powerful formalism to rigorously analyze the electrostatics of the MOS structure. The usefulness of the Poisson-Boltzmann formulation transcends this structure as it can be applied to other junctions that appear in semiconductor devices.

This chapter places special emphasis in the understanding of the capacitance-voltage characteristics of the MOS structure. As in the metal-semiconductor and PN junctions, the C-V characteristics summarize in a compact way the electrostatics of the structure. C-V characteristics form also the basis for a variety of techniques to characterize and diagnose semiconductor surfaces, interfaces, dielectrics and devices.

This is one of the most important chapters in this book. Solid command of the physics of the MOS structure is essential to understand MOSFETs and CMOS.

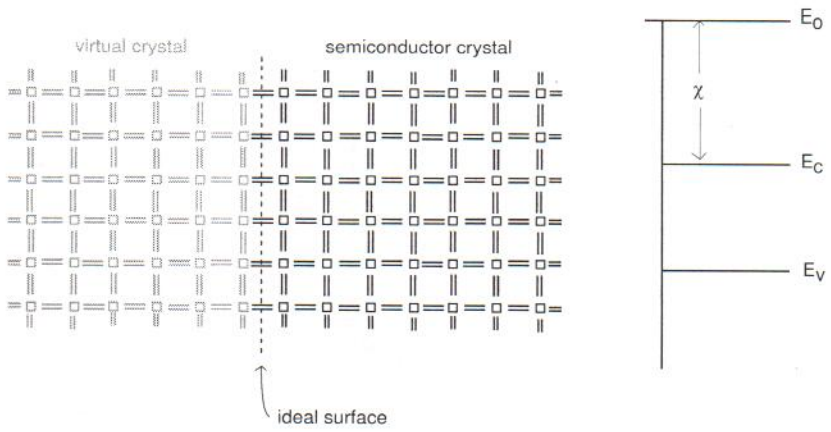


Figure 8.1: Left: Schematic diagram of an ideal surface with unperturbed semiconductor bonds. This is as if the real crystal continued into a "virtual crystal" where carriers are forbidden. Right: energy band diagram in the vicinity of an ideal surface. The bands are "bulk-like" all the way to the surface.

8.1 The semiconductor surface

At a semiconductor surface, the perfect crystalline periodicity of the solid comes to an abrupt end. This affects the transport properties of electrons and holes in its vicinity in a major way. Providing a simple and intuitive picture of what happens at a surface is the goal of this section. Before we do that, let us introduce the concept of *ideal semiconductor surface.*

An "ideal surface" can be thought of as one in which the semiconductor lattice comes to an end but the bulk electronic properties are unmodified by this disruption. Schematically, we can visualize an ideal surface if we introduce an imaginary plane inside the bulk of a semiconductor that carriers cannot traverse. This is illustrated on the left of Fig. 8.1. On one side is the real crystal, on the other side is a "virtual" crystal. Along the ideal surface plane, the semiconductor bonding arrangement remains undisturbed. Carriers can come up to that plane without feeling its presence in any way. They just simply cannot cross it because there are no allowed states on the other side of the interface for the carriers to go to. Another view of this situation is to consider that a large energy barrier exists right at the surface, as sketched on the right of Fig. 8.1. The barrier height for electrons in the conduction band is equal to the electron affinity of the semiconductor. It is even higher for the holes. As a consequence, the hole and electron current components normal to that plane are both zero. This concept of an ideal surface, while naive, is useful. We already encountered it in our discussion on the metal-semiconductor junction. We will invoke it many more times in this book.

In a "real" surface the four-fold coordination of the semiconductor atoms cannot be preserved. Even if nothing else happens, unsatisfied bonds will be present. This is a rather unstable situation as the atoms are eager to lower their energy by attempting to emulate four-fold coordination. As a consequence, a surface with unsatisfied bonds is very reactive. The surface atoms can find various ways to lower their total energy, as schematically illustrated in Fig. 8.2. One of them is to bond with foreign atoms or molecules, such as O or C or organic contaminants, that might

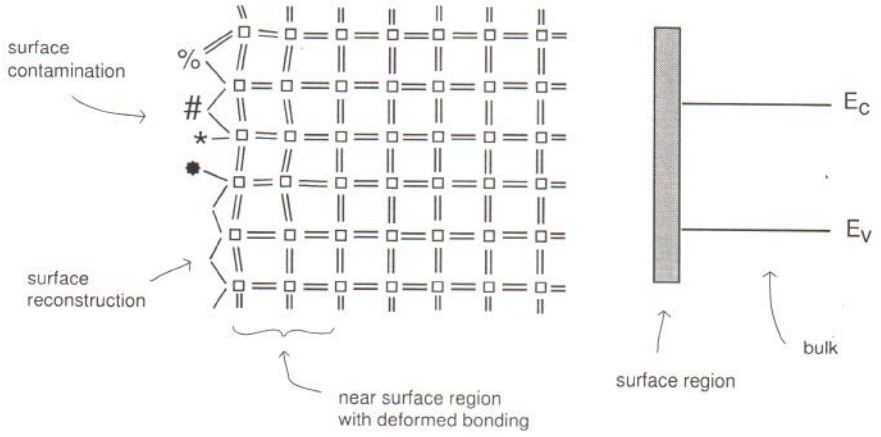


Figure 8.2: Left: Schematic illustration of a "real" surface with surface contaminants, surface reconstruction, and a subsurface region with deformed bonding. Right: Schematic energy band diagram in the vicinity of a "real" surface" showing a different electron energy distribution near the surface.

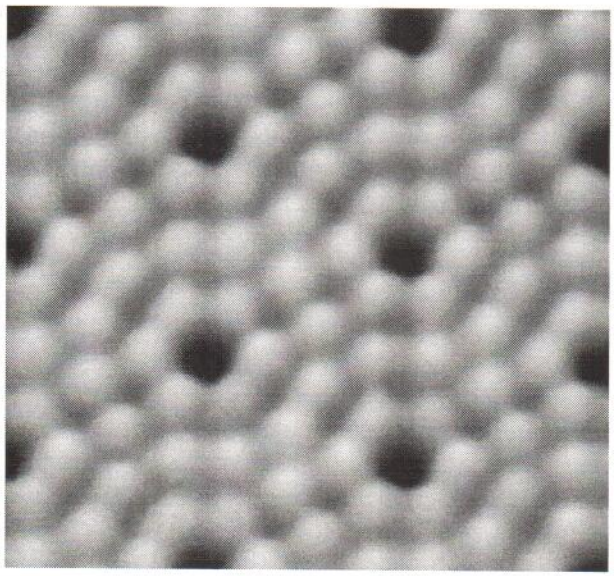


Figure 8.3: Scanning Tunneling Microscope picture of the 7x7 reconstruction of the (111) Si surface. Each round gray feature is associated with one distinct Si atom. The peculiar atomic arrangement pattern that is observed arises from minimization of the surface energy [*courtesy of R. Martel and Ph. Avouris, IBM*].

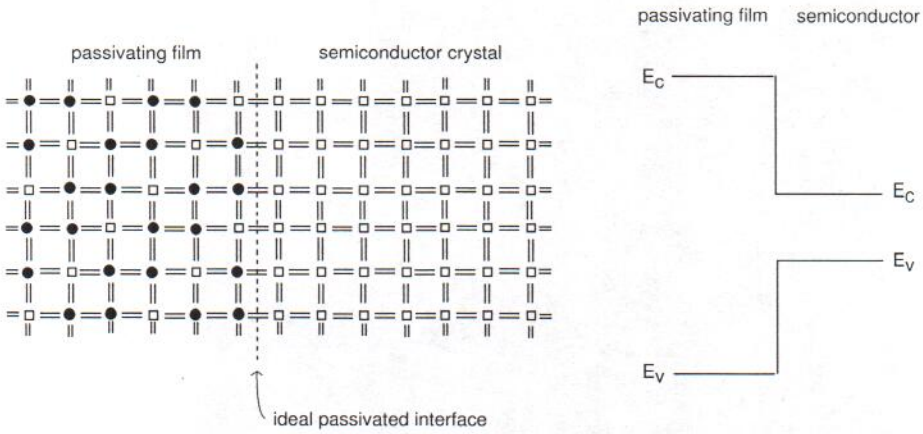


Figure 8.4: Left: Schematic illustration of surface passivation. A dielectric film has been placed on top of the semiconductor surface so that bulk-like bonding exists at the dielectric/semiconductor interface. Right: Energy band diagram in the vicinity of a passivated interface. The bands in the semiconductor are "bulk-like" all the way to the interface.

be present in the atmosphere surrounding the surface. A semiconductor surface, if exposed to air, oxidizes and gets contaminated very easily. A second avenue for surface energy lowering is *surface reconstruction* in which the surface atoms bond among themselves in an arrangement that is rather different from the one in the bulk. A (111) Si surface, for example, reconstructs itself into a spectacular diamond-shape 7x7 arrangement, pictured in Fig. 8.3.

When considering the impact that a "real" surface has on semiconductor device behavior, we might be tempted to believe that whatever it is, it should be confined to the immediate vicinity of the surface, that is the region where the semiconductor bonding is upset from its bulk arrangement. This deformed region is actually very thin, perhaps two or three monolayers. In consequence, we might be lead to believe that the influence of the surface does not extend more than a few nanometers below the surface. From a crystalline point of view, this argument is true, but from an electrostatic and transport one, it is not. We will learn in this chapter that a surface can modify the electrostatics of the bulk underneath to distances that can be as high as a few microns. Furthermore, we already saw in Ch. 5 that in minority carrier type situations the influence of surface recombination reaches several diffusion lengths into the bulk. Depending on doping level and processing details, this can be as far as hundreds of microns.

The high reactivity of an exposed semiconductor surface makes it unsuitable for semiconductor device applications. The microelectronics revolution would have not taken place without the discovery of *surface passivation*. This consists of the deposition of a dielectric material or the formation of a thin-film on top of a semiconductor surface such that bulk-like bonding at the interface is preserved. This is schematically illustrated in Fig. 8.4.

The most important passivating film is SiO_2. When Si is oxidized in an ultraclean environment, Si reacts with O to form SiO_2, which grows on the Si surface. SiO_2 formed this way is an amorphous dielectric material with remarkable properties. One of them is that at the Si-SiO_2 interface, nearly all Si bonds are satisfied. If the black circles are taken as the O atoms, Fig. 8.4

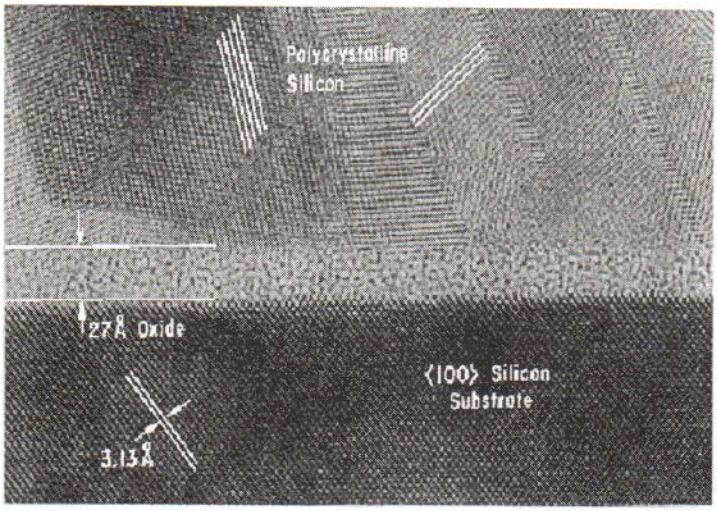


Figure 8.5: Transmission electron micrograph of modern MOS structure showing 1-2 monolayer surface roughness at the Si-SiO_2 interface. The gate, made out of heavily-doped polysilicon, shows several grains with different orientations. The extraordinary resolution of this picture permits the direct observation of the lattice planes that provides an absolute length calibration [*courtesy of D. Buchanan, IBM*].

can be interpreted as an ideal SiO_2 film on a Si surface. Notice that there are twice as many O atoms as Si atoms (the open squares) in the dielectric film. Note also that all atoms in the dielectric as well as at the dielectric/semiconductor interface share a total of eight valence electrons, the ideal bonding configuration.

This simple picture is however misleading in at least two important ways. First, SiO_2 films used in microelectronics are amorphous, that is, they do not have the long-range crystalline order that is suggested in Fig. 8.4. There is some short-range order, to be sure, with a bonding configuration similar to the one implied in this figure, but the film is really made out of multiple tiny crystallites with all kinds of orientations. A second problem is that Fig. 8.4 suggests that the Si/SiO_2 interface is perfectly flat. This is also not the case. Even carefully prepared interfaces have surface roughness on the scale of one to two monolayers. This is seen in the transmission-electron microscope picture of Fig. 8.5 that shows a cross section of a MOS structure used in modern CMOS applications. The MOS structure of the picture has a polysilicon gate on top of a 2.7 nm thick SiO_2 layer grown on a (100) Si substrate. If one looks in detail, one can see an Si/SiO_2 interface that is not flat but exhibits a roughness of about two monolayers. Surface roughness has very important implications for carrier transport along the Si/SiO_2 interface, as it happens in a MOSFET. This will be explored in Ch. 10. Interestingly, the extraordinary resolution of the picture in Fig. 8.5 allows the direct observation of the lattice planes in the Si substrate. This provides an absolute length calibration for the picture. Also, the polycrystalline nature of the gate is apparent.

In an ideal surface, the energy band diagram is bulk-like all the way to the surface, as sketched on the right of Fig. 8.1. The electrons do not "fall off" the semiconductor at the surface because they face an energy barrier equal to the electron affinity of the semiconductor. Electrons need to have an energy equal to the vacuum energy to escape from the semiconductor.

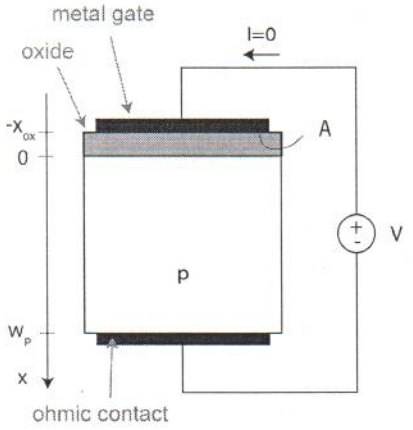


Figure 8.6: Sketch of ideal MOS structure on p-type substrate.

The energy band diagram of a perfectly passivated semiconductor surface is similar to an ideal surface, *i.e.*, the semiconductor bands remain bulk-like all the way to the interface. As we discussed in Ch. 1, the band diagram of an insulator is qualitatively identical to that of a semiconductor, except that the bandgap is much wider. The bandgap of SiO_2, for example, is about 9 eV. In consequence, at the insulator/semiconductor interface, there is a large discontinuity in the conduction band and the valence bands that prevents carriers from escaping. This is shown on the right of Fig. 8.4.

The energy band description of the bulk semiconductor is inappropriate in the vicinity of an unpassivated surface. The particular bonding arrangement of the surface and the presence of contamination results in an electron state configuration that is rather different from the one that prevails in the bulk. More importantly, the existence of a bandgap devoid of electron states can no longer be guaranteed, and in fact, a general continuum of states is more appropriate. This is pictured on the right of Fig. 8.2.

8.2 The ideal Metal-Oxide-Semiconductor structure

As in the case of the PN diode and the Schottky diode, it is useful to define the concept of an "ideal MOS structure". This captures the essence of the MOS physics while hiding some of the second-order effects. Later in this chapter, we will relax some of the assumptions made in this section.

A sketch of an ideal MOS structure is shown in Fig. 8.6. It consists of a p-type semiconductor (one can equally define it on an n-type substrate), an oxide of thickness x_{ox}, and a metal film. the metal is referred to this as the "gate". The semiconductor is contacted at the bottom by an ideal ohmic contact.

The following assumptions characterize the ideal MOS structure:

- This is a one dimensional situation. There are no 2D or 3D effects.
- We assume an ideal semiconductor surface at the oxide-semiconductor interface, i.e., smooth and defect free.
- The doping level in the semiconductor is uniform throughout.
- We assume that the metal work function and the doping level in the semiconductor are such that under zero bias, there is a depletion region in the semiconductor.
- In situations in which there is a depletion region in the semiconductor, we treat it under the depletion approximation.
- Under the application of a DC voltage, we assume that no DC current flows.
- The semiconductor is contacted through an ideal ohmic contact, as defined earlier in this book.
- We neglect any sidewall effects associated with the edges of the device.
- Maxwell-Boltzmann statistics apply to both types of carriers under all conditions.

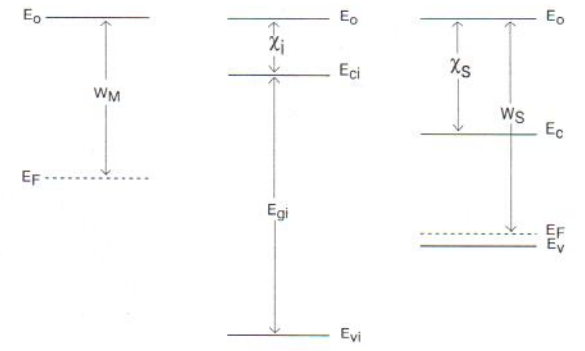
Fig. 8.6 also defines the axis that we will use in our analysis. The origin is placed at the oxide/semiconductor interface. The area of the MOS structure is A. The voltage convention that we will follow is also shown in the figure. For a MOS structure on a p-type substrate, V is defined as the potential difference between the gate and the semiconductor.

8.3 The ideal Metal-Oxide-Semiconductor structure at zero bias

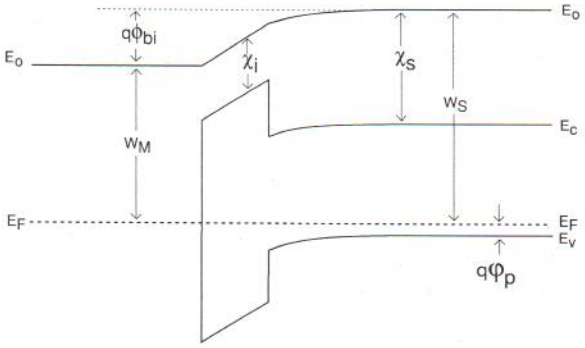
We start by analyzing the MOS structure at zero bias. The first question that one might ask is why do we use this notation, "zero bias," as opposed to "thermal equilibrium" used earlier in this book. Is it the same?

In an ideal MOS structure with a perfectly insulating oxide, no electron flow can take place between the metal and the semiconductor and the system cannot really attain thermal equilibrium. We can deal with this problem by shorting together the metal and the semiconductor through an external wire so that electron redistribution can take place through it. This is equivalent to applying a zero bias across the structure.

Consider a metal/oxide/p-type semiconductor structure with the three materials far apart, as sketched in an energy band diagram in Fig. 8.7a (dual situations occur with an n-type semiconductor substrate, as we will summarize later). The common reference energy for the three band diagrams is the vacuum level. The electronic structure of the metal is characterized by its work function. The semiconductor and the insulator are characterized by their electron affinity and their bandgaps, as well as the location of the Fermi level. In the case of the semiconductor, the Fermi level is set by doping. For the insulator, in principle one could establish the location



a) metal, insulator and semiconductor far apart



b) metal, insulator and semiconductor in intimate contact

Figure 8.7: Energy band diagrams a) for isolated metal, insulator and semiconductor, and b) after bringing them in intimate contact at zero bias.

of the Fermi level as if it was an intrinsic semiconductor. In practice, the bandgap of the insulator is so large, that the Fermi level can swing widely through most of it without changing in any significant way the occupation of the bands. For all practical purposes, the valence band is always completely full of electrons and the conduction band is perfectly empty. Because of this, the Fermi level is not usually drawn inside an insulator.

The situation shown in Fig. 8.7a is one in which the Fermi level in the semiconductor is lower than in the metal. In consequence, if these three materials are brought in intimate contact through an external wire, electrons will tend to flow from the metal to the semiconductor. When this happens, the metal becomes positively charged and the semiconductor acquires negative charge. This sets up an electric field that eventually brings carrier exchange to a halt establishing thermal equilibrium. The bottom of Fig. 8.7 shows the resulting energy band diagram at zero bias.

There are several interesting features in this energy band diagram. First of all, far away from their interfaces with the insulator, bulk thermal equilibrium conditions prevail in the metal and the semiconductor. In the vicinity of the oxide/semiconductor interface, the charge exchange

that has taken place between the metal and the semiconductor implies that the band diagram of the semiconductor bends *down*. In addition to getting more electrons close to the insulator, thermal equilibrium demands that fewer holes are present there. The semiconductor has become negatively charged. In fact, from a charge point of view, since holes are the majority carriers, the reduction of the hole concentration is far more significant than the increase of the electron concentration. The net effect, in any case, is that the semiconductor is negatively charged in the neighborhood of the interface with the insulator. This negative charge is imaged at the metal-insulator interface where there is a deficit of electrons. In the length scale of interest, this does not bend the band diagram of the metal in a significant way and it is not shown (refer to a related discussion in Section 7.2).

The dipole of charge that has appeared across the MOS structure results in two features in the band diagram. First, it bends the bands in the insulator. The fact that there is not bulk charge inside an ideal insulator implies that the bands are inclined but straight. Second, there is a built-in potential difference across the entire structure. As in previous junctions of dissimilar materials, the built-in potential is entirely set by the work functions of the end materials. In this case, it is given by:

$$q\phi_{bi} = W_S - W_M \tag{8.1}$$

A final observation to be made about the energy band diagram of Fig. 8.7 is the appearance of large conduction band and valence band discontinuities at the semiconductor/insulator interface. They constitute huge energy barriers to carrier flow and ideally prevent the injection of electrons and holes from the semiconductor to the insulator.

Before we proceed to analyze in detail the electrostatics of this situation, let us first derive some general relationships for the electrostatics of the MOS structure. We will derive them at zero bias, but with suitable modification they can also be used under bias.

8.3.1 General relations for the electrostatics of the ideal MOS structure

Fig. 8.8 shows a general MOS structure at zero bias. The coordinate system is selected to point into the wafer, with the origin placed at the semiconductor/insulator interface, as sketched on the left of Fig. 8.8. The right of the figure shows the volume charge density, the electric field, and the electrostatic potential along the space coordinate x.

The top of Fig. 8.8 shows the charge distribution in space. In the semiconductor we depict a generic charge distribution that extends to some depth from the Si/SiO_2 interface and that integrates to a total charge areal density Q_s (in units of C/cm^2). The insulator is assumed to be ideal and devoid of any charge inside. The metal also rejects volume charge but allows it at its surfaces. We can think of this charge as a "sheet" located at the metal-oxide interface with an areal density Q_g (in units of C/cm^2). Overall charge neutrality demands that:

$$Q_g = -Q_s \tag{8.2}$$

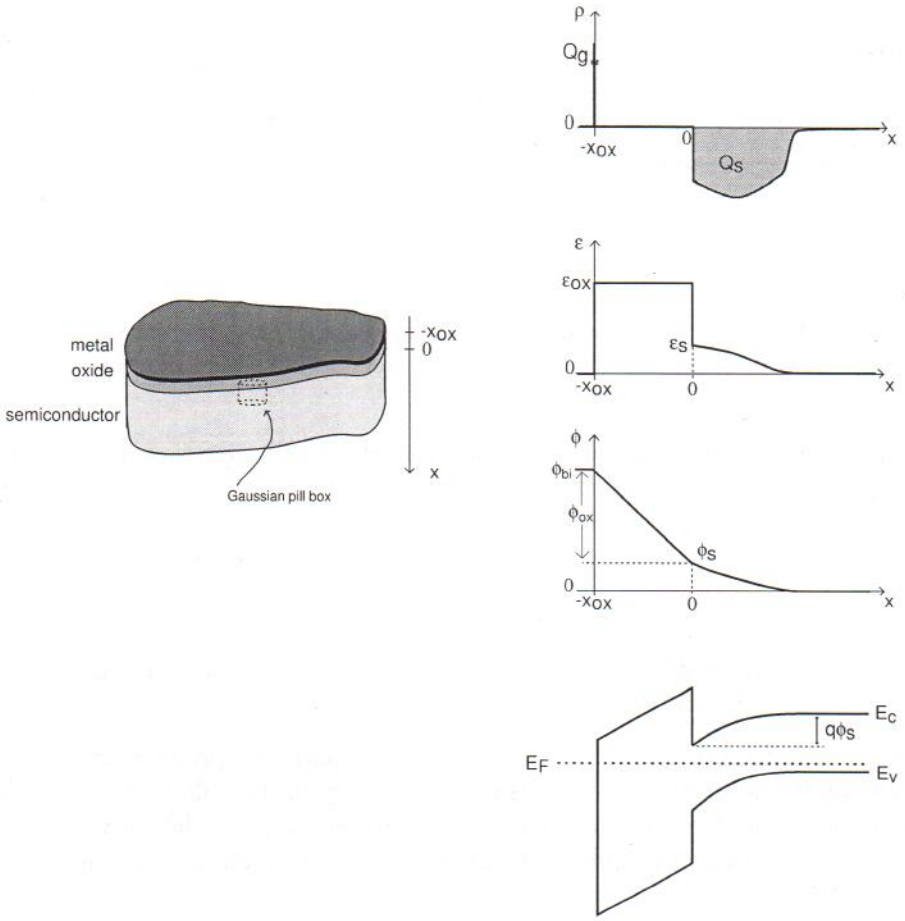


Figure 8.8: Schematic of MOS structure defining the coordinate axis and a Gaussian pill box (left). Generic sketches of, from top to bottom, volume charge density, electric field, electrostatic potential end energy band diagram for the MOS structure (right).

The field lines produced by this dipole of charge originate in the sheet charge at the metal/insulator interface and terminate in the semiconductor. As a result, the field inside the metal as well as sufficiently deep inside the semiconductor are zero. Application of Gauss law to this situation is straightforward. If we construct a "pill box" that includes all Q_s and penetrates partially into the oxide, as shown on the left of Fig. 8.8, we find the electric field in the oxide to be:

$$\mathcal{E}_{ox} = -\frac{Q_s}{\epsilon_{ox}} \tag{8.3}$$

where ϵ_{ox} is the permittivity of the oxide [1]. $\mathcal{E}_{ox}$ is uniform in space inside the oxide since there

[1] It is important to keep in mind the distinction between *permittivity* (ϵ, in F/cm^2) and *dielectric constant* (κ, no units). They are related through $\epsilon = \kappa\epsilon_o$ where ϵ_o is the permittivity of vacuum (value given in Appendix A).

are no charges there.

The permittivities of the insulator and the semiconductor are generally different. This implies that at the Si/insulator interface, the electric field is discontinuous. The relationship between the normal components of the electric field on both sides of the interface is obtained from the continuity of the electric displacement vector, that is:

$$\epsilon_s \mathcal{E}_s = \epsilon_{ox} \mathcal{E}_{ox} \tag{8.4}$$

The permittivity of SiO_2 is 3.9. This is exactly three times smaller than the permittivity of Si (see Appendix B). Therefore, the field inside the oxide is three times the field at the semiconductor surface.

Substituting Eq. 8.4 into 8.3 provides a relationship between the electric field at the semiconductor/insulator interface on the semiconductor side and Q_s:

$$\mathcal{E}_s = -\frac{Q_s}{\epsilon_s} \tag{8.5}$$

We could have obtained this result by selecting a different pillbox that includes all of Q_s and has one of its surfaces placed exactly at $x = 0$.

The electric field distribution results in an electrostatic potential profile that is sketched in Fig. 8.8. In this graph, the origin of potentials is selected deep in the bulk of the semiconductor. The total potential difference across the structure at zero bias is ϕ_{bi}. The total potential drop across the structure is equal to the sum of the drop in the semiconductor, ϕ_s, plus the drop in the insulator, ϕ_{ox}, that is:

$$\phi_{bi} = \phi_s + \phi_{ox} \tag{8.6}$$

ϕ_s is called the *surface potential.* It is related to the detailed charge distribution in space inside the semiconductor. ϕ_{ox} is the potential build-up across the insulator. Since the field in the oxide is uniform, ϕ_{ox} is given by:

$$\phi_{ox} = x_{ox} \mathcal{E}_{ox} \tag{8.7}$$

Eq. 8.6 is the basis for a general relationship between ϕ_s and Q_s. Towards deriving this, let us first define the capacitance per unit area of the insulator:

$$C_{ox} = \frac{\epsilon_{ox}}{x_{ox}} \tag{8.8}$$

If we now plug Eq. 8.7 into Eq. 8.6 and use Eqs. 8.4 and 8.5, we can write:

$$\phi_{bi} = \phi_s - \frac{Q_s}{C_{ox}} \tag{8.9}$$

The relationships just derived apply to any ideal MOS structure under any circumstances. They have been obtained from fundamental statements about the electrostatics of the structure. In particular, they all apply to situations under bias if we substitute ϕ_{bi} by the appropriate expression for the total potential build-up across the structure.

The bottom sketch in Fig. 8.8 is the energy band diagram. A significant point to be made is that with the electrostatic potential reference that was selected above, the total band bending inside the semiconductor is given by $q\phi_s$.

We can now proceed to study in some detail the electrostatics of the zero-bias situation outlined above.

Exercise 8.1: *A certain MOS structure with $x_{ox} = 4.5$ nm exhibits $\phi_{bi} = 1$ V and $\phi_{ox} = 0.42$ V at zero bias. Calculate ϕ_s, Q_s, $\mathcal{E}_s$, and $\mathcal{E}_{ox}$.*

There are multiple ways to go about solving this problem. We can use Eq. 8.7 to get $\mathcal{E}_{ox}$:

$$\mathcal{E}_{ox} = \frac{\phi_{ox}}{x_{ox}} = \frac{0.42\ V}{4.5 \times 10^{-7}\ cm} = 9.3 \times 10^5\ V/cm$$

From here, we can calculate $\mathcal{E}_s$ using Eq. 8.4:

$$\mathcal{E}_s = \frac{\epsilon_{ox}}{\epsilon_s}\mathcal{E}_{ox} = 0.33 \times 9.3 \times 10^5\ V/cm = 3.1 \times 10^5\ V/cm$$

We can get Q_s from either Eq. 8.3 or 8.5. Using the later one:

$$Q_s = \epsilon_s \mathcal{E}_s = 11.7 \times 8.85 \times 10^{-14} F/cm \times 3.1 \times 10^5\ V/cm = 3.2 \times 10^{-7}\ C/cm^2$$

Finally, ϕ_s can be obtained from Eq. 8.6:

$$\phi_s = \phi_{bi} - \phi_{ox} = 0.58\ V$$

8.3.2 Electrostatic of the MOS structure under zero bias

For a uniformly-doped semiconductor, deriving an analytical formulation of the electrostatics of the MOS structure under zero bias is relatively straightforward under the *depletion approximation*. If one focuses on the semiconductor in Fig. 8.7, one sees that close to the interface with the insulator there is a region with a space charge region, while far away from it, the bulk is quasi-neutral. This is similar to the situations that we observed in the PN diode and the Schottky diode. Consistent with the treatment that was followed there, we can perform the depletion approximation and assume that up to a certain depth, x_d, the semiconductor is devoid of majority carriers (holes in this case) and beyond x_d, the semiconductor is perfectly neutral. Under this approximation, the volume charge density, electric field, electrostatic potential and equilibrium carrier concentrations and energy band diagram of Fig. 8.9 result.

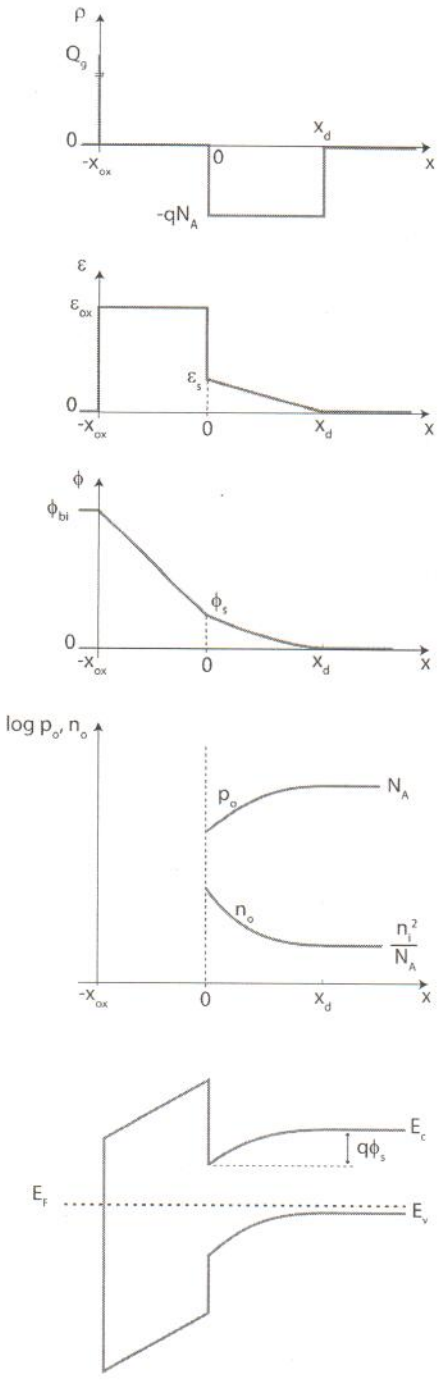


Figure 8.9: From top to bottom: sketches of volume charge density, electric field, electrostatic potential, equilibrium carrier concentration and energy band diagram in a MOS structure under zero bias under the depletion approximation (p-substrate).

Within the depletion approximation, the volume charge density inside the depletion region is uniform and equal to $\rho = -qN_A$. Hence, the integrated semiconductor charge is:

$$Q_s \simeq -qN_A x_d \tag{8.10}$$

The electric field distribution in the semiconductor has a triangular shape with a maximum at the semiconductor interface with the oxide. Plugging Eq. 8.10 into Eq. 8.5, the electric field at the semiconductor surface is given by:

$$\mathcal{E}_s \simeq \frac{qN_A x_d}{\epsilon_s} \tag{8.11}$$

Eq. 8.3 gives an expression that allows the calculation of the electric field inside the oxide. We can then write:

$$\mathcal{E}_{ox} \simeq \frac{qN_A x_d}{\epsilon_{ox}} \tag{8.12}$$

The electrostatic potential in the semiconductor has a parabolic shape. The surface potential is obtained by integrating the triangularly shaped electric field. This yields:

$$\phi_s \simeq \frac{qN_A x_d^2}{2\epsilon_s} \tag{8.13}$$

The only unknown at this time is the depth of the depletion region x_d. This can be obtained by plugging Eqs. 8.13 and 8.10 into the fundamental electrostatics relation 8.9 to yield:

$$\phi_{bi} \simeq \frac{qN_A x_d^2}{2\epsilon_s} + \frac{qN_A x_d}{C_{ox}} \tag{8.14}$$

Solving this quadratic equation we obtain:

$$x_d \simeq \frac{\epsilon_s}{C_{ox}}\left(\sqrt{1 + \frac{4\phi_{bi}}{\gamma^2}} - 1\right) \tag{8.15}$$

where γ is the so-called *body factor coefficient*:

$$\gamma = \frac{1}{C_{ox}}\sqrt{2\epsilon_s qN_A} \tag{8.16}$$

which has units of $V^{1/2}$. Eq. 8.15 states that the larger the difference in work functions between the metal and the semiconductor, the larger x_d is. This makes sense since the higher ϕ_{bi}, the larger the charge exchange that needs to take place.

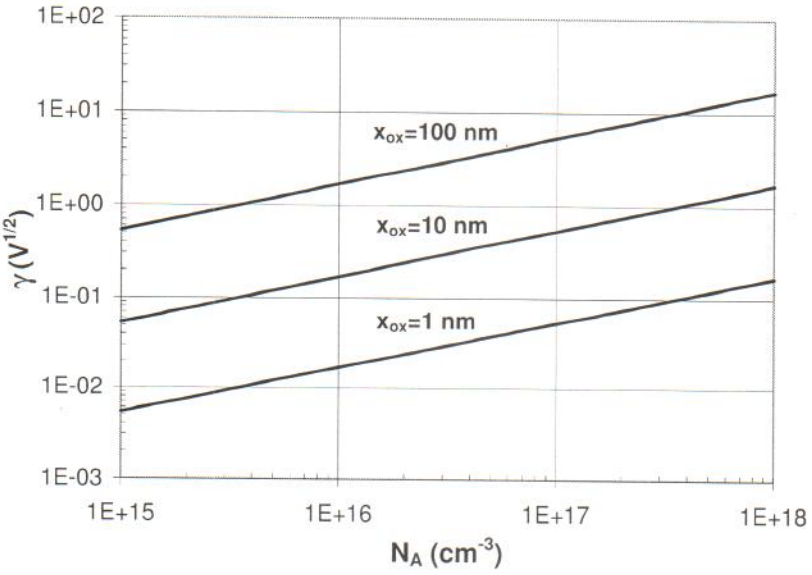


Figure 8.10: Body factor coefficient, γ, vs. substrate doping level for different values of the SiO_2 thicknesses.

It is also of interest to focus on the total charge in the depletion layer. This is given by:

$$Q_d \simeq -qN_A x_d \simeq -\sqrt{2\epsilon_s q N_A \phi_s} \tag{8.17}$$

where we have used Eq. 8.13. As in a PN junction and a Schottky junction, there is a square-root relationship between depletion charge and potential build up across a depletion region.

γ is an important parameter of a MOS structure. As Eq. 8.16 indicates, γ depends on the doping level of the substrate and the capacitance per unit area of the insulator. As it will be better understood later on in this chapter, γ makes a statement about the relative magnitude of the depletion region capacitance to the oxide capacitance. The higher the doping level, the thinner the depletion region in the semiconductor and the higher the depletion capacitance, hence γ goes up. Similarly, the thinner the oxide, the higher the oxide capacitance and the lower γ gets. Both dependencies are seen in Fig. 8.10 which graphs γ vs. N_A and the SiO_2 thickness.

In a well designed MOSFETs, a proper balance between the oxide and depletion region thicknesses is required. Hence, γ is typically on the order of 0.1 to 1 $V^{1/2}$. In time, γ has tended to decrease as MOSFET size has shrunk.

Once the electrostatic potential in the semiconductor is known, it is straightforward to obtain first-order expressions for the zero bias carrier densities in the semiconductor. The procedure is

identical to that one followed in the case of the metal-semiconductor junction in Section 7.2.3.

Exercise 8.2: *Consider a MOS structure made out of a W gate ($W_M = 4.54\ eV$), a 4.5 nm SiO_2 insulator, and a p-type doped Si substrate with $N_A = 6 \times 10^{17}\ cm^{-3}$. Estimate the electric field in the oxide and the surface potential at zero bias.*

Since W_M is about $\chi + E_g/2$, this structure is likely to be in depletion. In order to compute $\mathcal{E}_{ox}$, we first need to calculate x_d. Before that we need to compute ϕ_{bi}, C_{ox} and γ, which can be obtained, respectively, from Eqs. 8.1, 8.8 and Eq. 8.16. In order to get ϕ_{bi}, we need to first compute W_S:

$$W_S = \chi + E_g + kT \ln \frac{N_A}{N_v} = 4.04 + 1.12 + 0.026 \times \ln \frac{6 \times 10^{17}}{3.1 \times 10^{19}} = 5.06\ eV$$

Now:

$$\phi_{bi} = \frac{1}{q}(W_S - W_M) = 5.06 - 4.54 = 0.52\ V$$

$$C_{ox} = \frac{\epsilon_{ox}}{x_{ox}} = \frac{3.45 \times 10^{-13}\ F/cm}{4.5 \times 10^{-7}\ cm} = 7.7 \times 10^{-7}\ F/cm^2$$

$$\gamma = \frac{1}{C_{ox}}\sqrt{2\epsilon_s q N_A} = \frac{\sqrt{2 \times 1.04 \times 10^{-12}\ F/cm \times 1.6 \times 10^{-19}\ C \times 6 \times 10^{17}\ cm^{-3}}}{7.7 \times 10^{-7}\ F/cm^2} = 0.58\ V^{1/2}$$

We can now use Eq. 8.15 to obtain x_d:

$$x_d = \frac{\epsilon_s}{C_{ox}}\left(\sqrt{1 + \frac{4\phi_{bi}}{\gamma^2}} - 1\right) = \frac{1.04 \times 10^{-12}\ F/cm}{7.7 \times 10^{-7}\ F/cm^2}\left(\sqrt{1 + \frac{4 \times 0.52\ V}{0.58^2\ V}} - 1\right) = 2.3 \times 10^{-6}\ cm = 23\ nm$$

$\mathcal{E}_{ox}$ can now be obtained using Eq. 8.12:

$$\mathcal{E}_{ox} = \frac{qN_A x_d}{\epsilon_{ox}} = \frac{1.6 \times 10^{-19}\ C \times 6 \times 10^{17}\ cm^{-3} \times 2.3 \times 10^{-6}\ cm}{3.45 \times 10^{-13}\ F/cm} = 6.4 \times 10^5\ V/cm$$

The surface potential can be obtained from Eq. 8.13:

$$\phi_s = \frac{qN_A x_d^2}{2\epsilon_s} = \frac{1.6 \times 10^{-19}\ C \times 6 \times 10^{17}\ cm^{-3} \times (2.3 \times 10^{-6}\ cm)^2}{2 \times 1.04 \times 10^{-12}\ F/cm} = 0.24\ V$$

In the next section we will learn how to confirm that indeed this structure is in depletion.

8.4 The ideal Metal-Oxide Semiconductor structure under bias

The application of a voltage to the gate with respect to the body of a MOS structure affects the potential and charge distributions throughout. For the case of a p-type substrate, if we define as positive the voltage in the metal with respect to the semiconductor, the total potential difference across the structure goes from a zero bias value of ϕ_{bi} to a new value equal to $\phi_{bi} + V$, where V

is the applied voltage. The charge dipole across the MOS structure has to adapt itself so as to produce such a potential build up. In what way is it modified and what are the consequences of this? These are the topics that we study in this section.

Before discussing the details, we need to recognize that the presence of the oxide in the MOS structure blocks the flow of current even when a voltage is applied. Current continuity then implies that no current flows across the semiconductor. So, even though a bias has been applied, the semiconductor remains in a quasi-equilibrium condition. We can therefore conclude that similar to the PN diode or the Schottky diode, the electrostatics of a MOS structure with a built-in potential ϕ_{bi} under a voltage V are identical to the electrostatics of a MOS structure with a built-in potential equivalent to $\phi_{bi} + V$ under zero bias[2]. As a consequence, the general relation for the electrostatics of the MOS structure derived in Eq. 8.9 applies here for a total potential difference that is given by $\phi_{bi} + V$:

$$\phi_{bi} + V = \phi_s(V) - \frac{Q_s(V)}{C_{ox}} \tag{8.18}$$

where, in general, we note that ϕ_s and Q_s now depend on V.

Let us now qualitatively discuss the changes that take place in the electrostatics of a MOS structure as a result of the application of a voltage. A pictorial view is sketched in Fig. 8.11. The center diagram shows the structure at zero bias in which a depletion region exists in the semiconductor. An identical amount of positive charge resides at the metal-oxide interface. If a negative voltage is applied, the potential difference across the structure is reduced. In consequence, the magnitude of the charge dipole must decrease. This implies that the depletion layer shrinks. This voltage regime in which the depletion region thickness is being modulated by the applied bias is referred to as **depletion**.

For sufficient negative bias, the potential difference across the structure is wiped out. This calls for the charge dipole to disappear and the depletion region to completely collapse. This condition is called **flatband**. If a more negative voltage than this is applied, the potential difference across the structure reverses sign. This now demands the appearance of positive charge in the semiconductor and a negative sheet of charge at the metal-oxide interface. This can only be accomplished by accumulating holes against the oxide-semiconductor interface, a condition that is referred to as **accumulation**. This is shown in the top diagram of Fig. 8.11.

Moving downwards from $V = 0$ now, if a positive voltage is applied to the gate with respect to the semiconductor, the total potential difference across the structure increases and the depletion region has to widen to satisfy the electrostatics. If the voltage is made sufficiently large, eventually electrons start piling up at the insulator/semiconductor interface. The onset of this phenomenon is called **threshold**. For voltages higher than this, the electron concentration at the oxide-semiconductor interface becomes significant, a condition that is known as **inversion**. This is the regime that is exploited in MOSFETs.

[2]In the Schottky junction and the PN junction, we assumed that the electrostatics under bias were identical to equilibrium even though current was flowing by. In most cases, this assumption gives reasonable results. In the case of the MOS structure, this assumption is particularly well justified since no current flows.

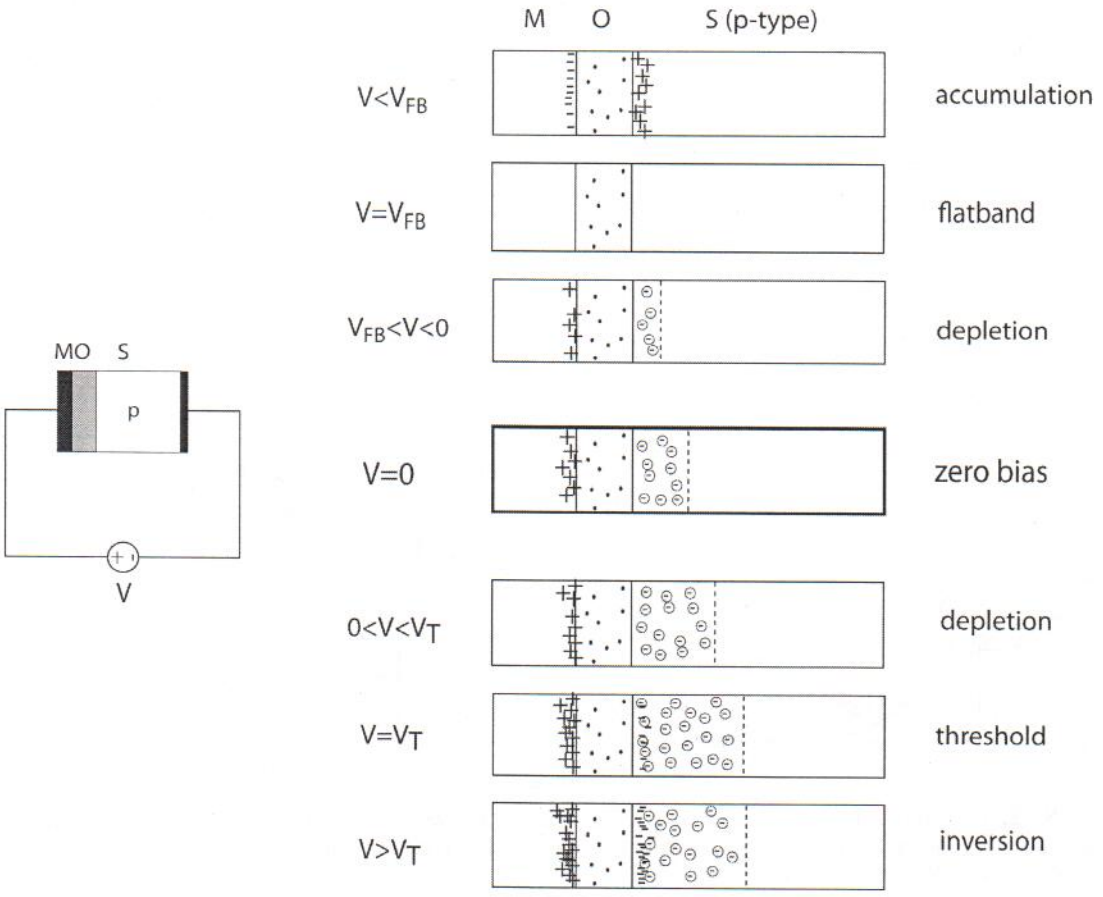


Figure 8.11: Impact of the application of a voltage to the charge distribution of a MOS structure. For negative voltages of the metal with respect to the semiconductor, the depletion region shrinks. For a large enough negative voltage, the structure is driven into accumulation. For positive voltages, the depletion region widens. For large enough voltage, the structure is driven into inversion.

We proceed to analyze these different situations in more detail in the following subsections.

8.4.1 Depletion

A detailed view of the MOS electrostatics in the depletion regime is shown in Fig. 8.12. Starting with the potential distribution, the application of a voltage to the gate with respect to the body changes the potential difference across the MOS structure. This requires a modulation of the charge distribution. For $V > 0$, the potential difference increases from the zero bias value of ϕ_{bi} to $\phi_{bi} + V$. In consequence, the depletion region widens and the electric field everywhere increases. For $V < 0$, the contrary happens and the potential difference is reduced, the depletion region shrinks and the electric field is lowered throughout.

The situation is well reflected in the evolution of the energy band diagram, which is also shown

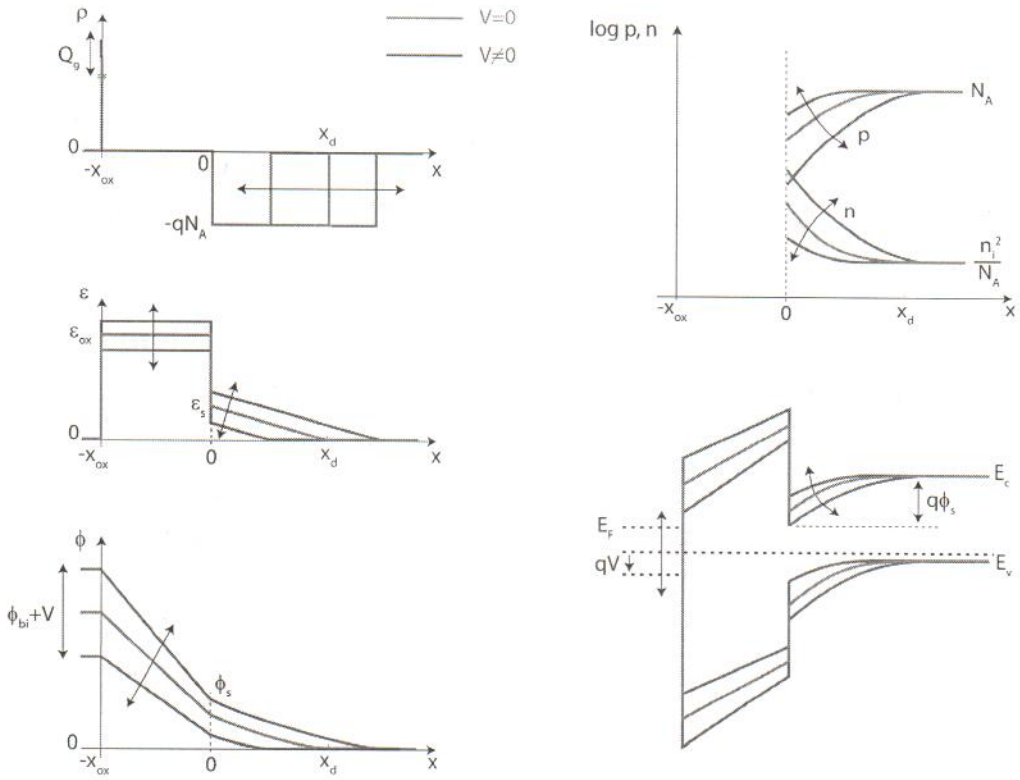


Figure 8.12: From top to bottom and left to right: sketches of volume charge density, electric field, electrostatic potential, carrier concentration and energy band diagram of a MOS structure under zero bias and for bias negative and positive but close to zero (depletion regime).

in Fig. 8.12. For $V > 0$, the Fermi level in the metal is pulled down with respect to the Fermi level in the semiconductor. As a result, the band bending across the oxide and the semiconductor increases. The contrary happens for $V < 0$. Note how the Fermi level in the semiconductor is drawn flat and that the hole and electron quasi-Fermi levels overlap. This reflects the absence of current through the structure and the consideration that the semiconductor remains in a quasi-equilibrium state.

It is interesting to reflect on what happens to the carrier concentrations. This is sketched in the top right of Fig. 8.12. For $V > 0$, the hole concentration drops close to the semiconductor/oxide interface as the depletion region widens. The contrary happens when $V < 0$. Since the pn product has to remain constant (consistent with zero current), for $V > 0$, the electron concentration close to the oxide/semiconductor interface rises while it drops for $V < 0$.

In the depletion regime, the change in the potential build up across the structure modulates the depletion region thickness. Nevertheless, the essential character of the electrostatics remains unchanged from the situation that we discussed under zero bias. In consequence, an analytical description of the depletion regime can be obtained simply by going to our zero bias results and substituting ϕ_{bi} by $\phi_{bi} + V$.

In particular, going back to the results obtained in Sec. 8.3.2, the depletion region thickness as a function of V is given by:

$$x_d(V) \simeq \frac{\epsilon_s}{C_{ox}}\left(\sqrt{1+4\frac{\phi_{bi}+V}{\gamma^2}}-1\right) \tag{8.19}$$

The total charge in the depletion region, from Eq. 8.10, can be written as:

$$Q_s(V) = Q_d(V) \simeq -\frac{1}{2}\gamma^2 C_{ox}\left(\sqrt{1+4\frac{\phi_{bi}+V}{\gamma^2}}-1\right) \tag{8.20}$$

where we have used the definition of γ in Eq. 8.16.

The evolution of the surface potential with V can be easily obtained from Eq. 8.13:

$$\phi_s(V) \simeq \frac{1}{4}\gamma^2\left(\sqrt{1+4\frac{\phi_{bi}+V}{\gamma^2}}-1\right)^2 \tag{8.21}$$

Similar relations can be obtained for any other variable of interest by starting from the zero bias equations derived in Sec. 8.3.2.

Exercise 8.3: *Consider an MOS structure identical to that of Exercise 8.2 except for the use of a poly-Si gate ($W_M = 4.04\ eV$). Calculate Q_s and its extent into the semiconductor for $V = -0.5\ V$.*

Let us assume that for $V = -0.5\ V$ this structure is in depletion (we will verify this in Exercise 8.4). The first step is to compute ϕ_{bi} as in Exercise 8.2. This yields $\phi_{bi} = 1.0\ V$. Q_s can be obtained from Eq. 8.20 (the values of $C_{ox} = 7.7 \times 10^{-7}\ F/cm^2$, and $\gamma = 0.58\ V^{1/2}$ were also obtained in Exercise 8.2):

$$\begin{aligned} Q_s &= -\frac{1}{2}\gamma^2 C_{ox}\left(\sqrt{1+4\frac{\phi_{bi}+V}{\gamma^2}}-1\right) \\ &= -\frac{1}{2}(0.58)^2\ V \times 7.7\times 10^{-7}\ F/cm^2\left[\sqrt{1+4\frac{(1.0-0.5)\ V}{(0.58\ V)^2}}-1\right] = -2.1\times 10^{-7}\ C/cm^2 \end{aligned}$$

The extent of the depletion region is simply:

$$x_d = \frac{Q_s}{-qN_A} = \frac{-2.1\times 10^{-6}\ C/cm^2}{-1.6\times 10^{-19}\ C \times 6\times 10^{17}\ cm^{-3}} = 22\ nm$$

8.4.2 Flatband

For large enough negative voltage applied to the gate with respect to the body of the MOS structure, the depletion region is wiped out. This is referred to as flatband. In this instance, the potential build up across the structure is precisely zero. As a result, there is no charge dipole,

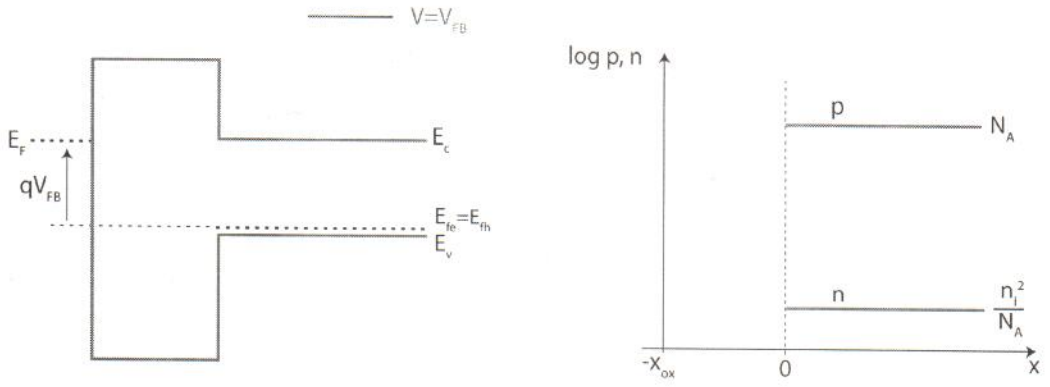


Figure 8.13: Energy band diagram and carrier concentration distribution in the MOS structure at flatband.

and there is no electric field anywhere. As sketched in Fig. 8.13, the energy bands are perfectly flat, hence the name, and the carrier concentrations are constant throughout the semiconductor which is then perfectly charge neutral.

The voltage required to drive the structure to flatband is referred to as the *flatband voltage*. Since there is no charge in the semiconductor, $Q_s = 0$. As a result, the band bending is zero and the surface potential $\phi_s = 0$. Then, from the general relation for electrostatics in Eq. 8.18, the flatband voltage is given by:

$$V_{FB} = -\phi_{bi} \tag{8.22}$$

Looking back at the expression of ϕ_{bi} given in Eq. 8.1, it is clear that the flatband voltage is entirely set by the bulk properties of the metal and the semiconductor.

The flatband voltage is an important reference voltage in the operation of MOS devices. We will extensively use it in the treatment of the MOS structure and in the analysis of MOSFETs.

8.4.3 Accumulation

Nothing prevents us from applying a voltage across the MOS structure that is more negative than V_{FB}. When this is done, the gate is placed at a more negative potential than the body of the semiconductor. This forces a charge dipole with positive charge in the semiconductor and an identical amount of negative charge at the metal-oxide interface. Positive charge on the semiconductor can be obtained if the hole concentration is higher than the acceptor concentration. Since holes are mobile, they will tend to pile up, or accumulate, at the oxide-semiconductor interface. This is the accumulation regime.

A detailed view of the electrostatics in accumulation is shown in Fig. 8.14. The thin hole accumulation layer is imaged at the gate-oxide interface where there is a negative sheet of charge. This yields electric field and electrostatic potential distributions as sketched in the figure and a band structure that now bends upwards towards the metal.

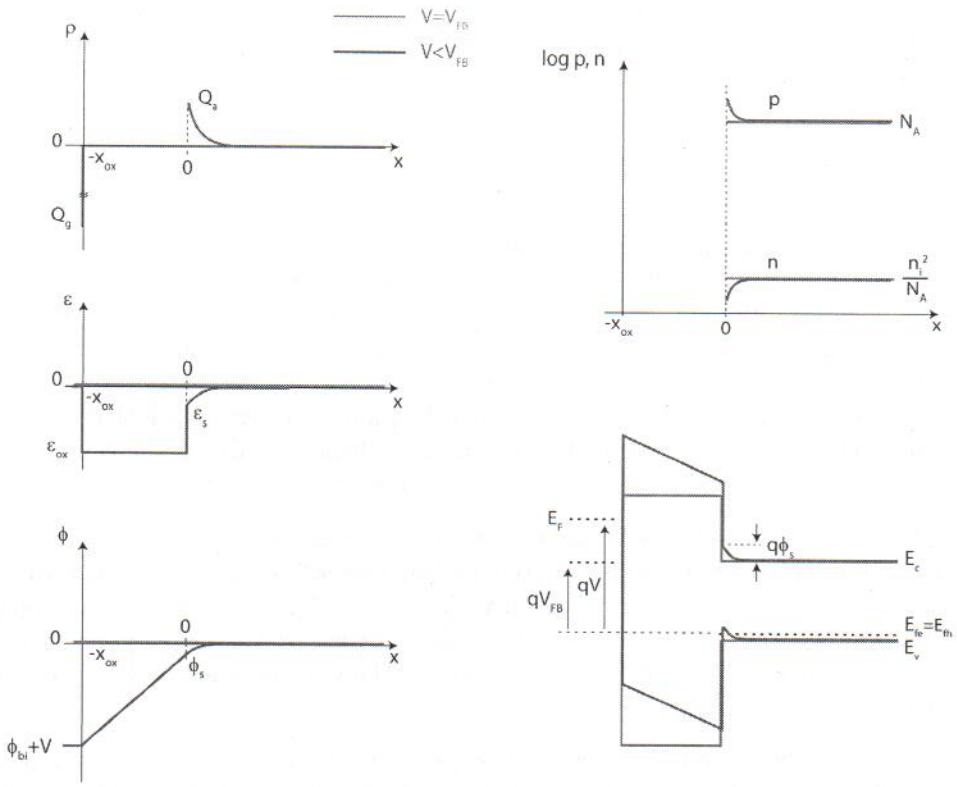


Figure 8.14: From top to bottom and left to right: sketches of volume charge density, electric field, electrostatic potential, carrier concentration and energy band diagram of a MOS structure at flatband and in accumulation.

It is relatively straightforward to derive first-order expressions for the hole accumulation charge and other electrostatic parameters under the *sheet-charge approximation*. This assumes that the accumulation layer is much thinner than any other dimensions in the structure. The hole distribution in the semiconductor reflects a balance between drift and diffusion. Hence, its extent is expected to be of the order of the Debye length, which can be quite small.

Under the sheet charge approximation, the surface potential in accumulation is:

$$\phi_s \simeq 0 \tag{8.23}$$

This is because a sheet of charge produces a sharp change in the electric field over zero distance and this does not accrue any potential buildup.

If we denote as Q_a the sheet hole charge density in the accumulation layer, then, using Eq. 8.18, we can easily obtain:

$$Q_a \simeq C_{ox}(V_{FB} - V) \tag{8.24}$$

where we have also used Eq. 8.22. This is an important result. Once in accumulation, the hole accumulation charge increases linearly with the voltage that is applied beyond the flatband voltage.

8.4.4 Threshold

We now come back to applying positive voltages to the gate with respect to the body. As we discussed above, for $V > 0$, the depletion region in the semiconductor widens. This is necessary to accommodate the additional negative charge required to satisfy the higher external potential difference between the gate and the body (see Fig. 8.12). A consequence of this can also be seen in this figure and that is that the electron concentration at the surface of the semiconductor increases. This is also reflected in the energy band diagram in which we see that as a result of the increased band bending associated with the enlarged depletion region, the conduction band edge at the surface of the semiconductor gets closer to the Fermi level.

It is clear that there is a voltage for which the electron concentration at the surface becomes equal to the hole concentration in the body of the semiconductor (see Fig. 8.15). This conditions is called threshold and it marks the onset of the inversion regime. "Threshold" is a very appropriate name as at this voltage, the semiconductor type changes at the surface from p-type to n-type. For gate voltages more positive than the threshold voltage, the electrostatics of the MOS structure drastically change in character.

It is relatively straightforward to derive an expression for the threshold voltage of a MOS structure. We start by finding an expression for the surface potential at threshold, ϕ_{sT}. For this, we can use the Boltzmann relation and relate the difference in electrostatic potential between the surface and the bulk, precisely ϕ_{sT}, to the ratio of electron concentration at these two locations. This yields:

$$\phi_{sT} = \frac{kT}{q} \ln \frac{n(x=0)|_T}{\frac{n_i^2}{N_A}} = 2\frac{kT}{q} \ln \frac{N_A}{n_i} = 2\phi_f \tag{8.25}$$

where we have taken $n(x=0)|_T = N_A$. We discuss the meaning of ϕ_f below.

We next need an expression for the charge in the semiconductor at threshold. If we can neglect the contribution from the electrons at the surface (not a bad assumption since the depletion region is much thicker than this thin electron layer), then the semiconductor charge at threshold is associated with the depletion region. For reasons that will become clear soon, we denote the thickness of the depletion region at threshold as x_{dmax} and the depletion charge as Q_{dmax}. With a potential build up across the depletion layer of ϕ_{sT}, using Eq. 8.17, x_{dmax} and Q_{dmax} are, respectively, given by:

$$x_{dmax} \simeq \sqrt{\frac{2\epsilon_s \phi_{sT}}{qN_A}} \tag{8.26}$$

and

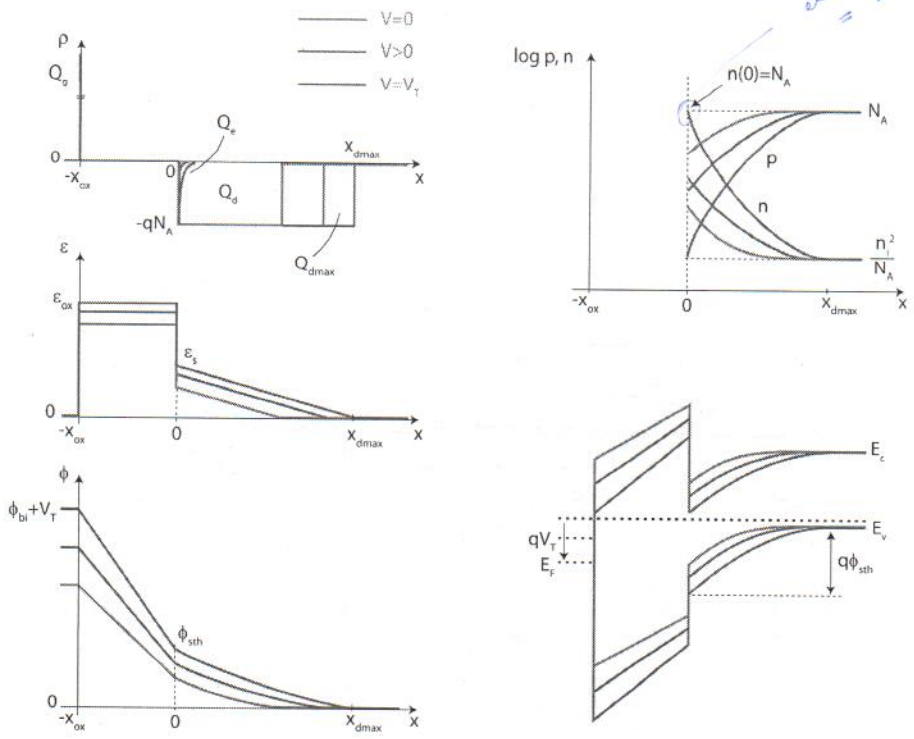


Figure 8.15: From top to bottom and left to right: sketches of volume charge density, electric field, electrostatic potential, carrier concentration and energy band diagram of a MOS structure at zero volts, in depletion regime and at threshold.

$$Q_{dmax} \simeq -\sqrt{2\epsilon_s q N_A \phi_{sT}} \tag{8.27}$$

We can now use the general relation for electrostatics of the MOS structure (Eq. 8.18) at threshold to obtain:

$$V_T = -\phi_{bi} + \phi_{sT} - \frac{Q_{dmax}}{C_{ox}} = V_{FB} + \phi_{sT} + \gamma\sqrt{\phi_{sT}} \tag{8.28}$$

where we have also substituted Eqs. 8.22, 8.27 and 8.16.

V_T is the threshold voltage of the MOS structure and it plays a key role in MOSFET operation. V_T depends on all the key parameters of the MOS structure: semiconductor doping level, oxide thickness and gate work function. Fig. 8.16 shows V_T as a function of N_A and x_{ox} for a n^+-polySi gate and SiO_2 dielectric. As N_A or x_{ox} increase, so does V_T. Notice that the dependence on doping is much weaker for thin oxides. This is because the doping dependence of V_T is mainly through the square root doping dependence of γ (Eq. 8.16). For a thin oxide, C_{ox} is high, and as Eq. 8.16 shows, the absolute magnitude of γ is very small.

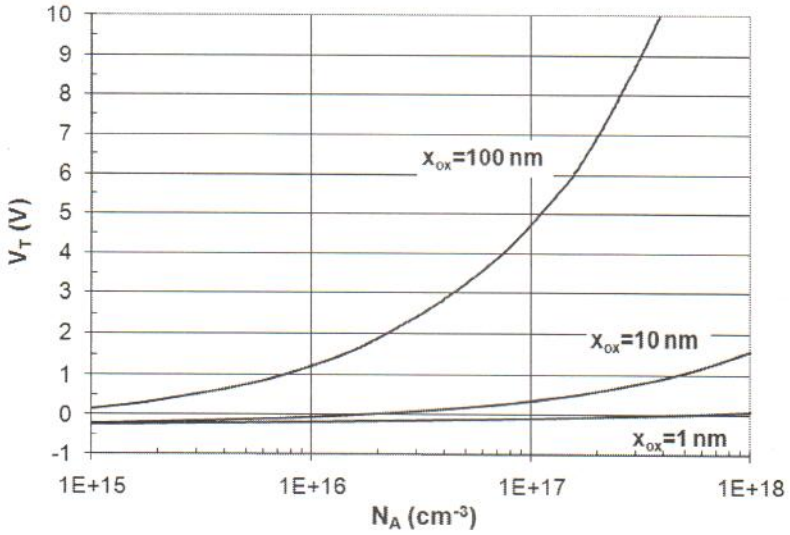


Figure 8.16: Threshold voltage, V_T, vs. substrate doping level for different values of the oxide thickness. The gate is made out of n^+-polySi ($W_M = 4.04\ eV$) and the oxide is SiO_2.

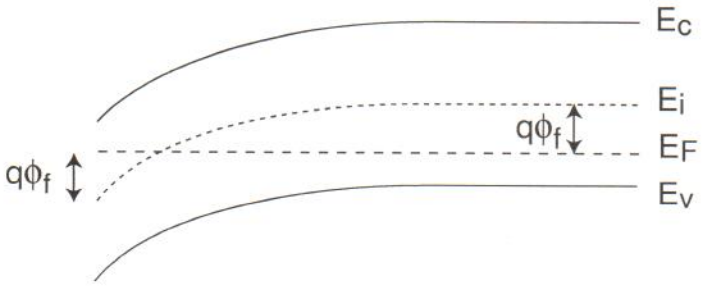


Figure 8.17: Energy band diagram of semiconductor at the onset of inversion (p-substrate). The definition of ϕ_f is indicated.

In Eq. 8.25, the surface potential at threshold was defined in terms of ϕ_f. Fig. 8.17 illustrates this definition. ϕ_f is the distance in energy between the intrinsic Fermi level and the Fermi level in the body of the semiconductor. When the band bending in the semiconductor is precisely equal to $2\phi_f$, then the electron concentration at the surface is identical to the hole concentration in the bulk.

8.4.5 Inversion

Threshold marks the onset of the creation of an inversion layer of electrons at the oxide-semiconductor interface of a MOS structure. This does not occur abruptly at the treshold voltage but for $V > V_T$ its existence can no longer be ignored. In fact, for $V > V_T$ the electrostatics of the MOS structure are profoundly affected by the presence of this inversion layer.

Applying a gate voltage that exceeds V_T demands an increase in the magnitude of the charge dipole across the MOS structure. To understand what this implies on the semiconductor side, we need to note the dependences of the inversion layer charge, Q_i, and depletion region charge, Q_d, on the surface potential. While Q_i is expected to increase exponentially with ϕ_s (more on this later), Q_d evolves as the square root of ϕ_s (Eq. 8.17). Hence, any increase in ϕ_s brought about by an increase in voltage will mostly grown the inversion layer while it will only produce a minor increase in the depletion region thickness beyond its value at threshold. It is because of this that we denote the thickness and charge associated with the depletion region at threshold as x_{dmax} and Q_{dmax}, respectively.

Fig. 8.18 illustrates the electrostatics of the MOS structure at threshold and in inversion. Starting from the top-right sketch, in inversion the electron concentration at the surface of the semiconductor increases in a prominent way. This is more clearly seen in the top-left sketch that shows the charge distribution across the structure. In inversion, the electron concentration at the surface exceeds the background acceptor concentration. As a result of the appearance of the inversion layer, the electric field exhibits a rather abrupt jump at the semiconductor surface. However, since the inversion layer is expected to be very thin (the electrons are "jammed" against the oxide), the electrostatic potential does not change very much. Due to the exponential dependence on Q_i on ϕ_s, not much of a change in ϕ_s is needed to induce the required amount of Q_i to satisfy the electrostatics. This is also reflected in the band diagram on the bottom right which shows a band bending across the semiconductor that has not changed much beyond threshold with most of the extra applied gate voltage used to bend the bands across the oxide.

It is relatively straightforward to obtain a first-order expression for the evolution of the inversion charge with the gate voltage. This can easily be done under the context of the *sheet-charge approximation*. As in accumulation, we assume that the inversion layer charge is very thin in the scale of the other dimensions of the problem (x_{ox} and x_{dmax}). Under this approximation, the surface potential in inversion is approximately equal to the value that it has at threshold:

$$\phi_s \simeq \phi_{sT} \tag{8.29}$$

We can now use the general relation for the electrostatics of the MOS structure (Eq. 8.18) in inversion:

$$\phi_{bi} + V = \phi_{sT} + \frac{Q_{dmax} + Q_i}{C_{ox}} \tag{8.30}$$

We now solve for Q_i and substitute the expressions of V_{FB} (Eq. 8.22) and V_T (Eq. 8.28) to get:

$$Q_i = -C_{ox}(V - V_T) \qquad \text{for } V \geq V_T \tag{8.31}$$

This is a very important equation. It is often called the *charge-control relation of the inversion layer*. It states that once reached threshold, the application of additional voltage across the MOS structure grows the inversion layer charge in a linear manner.

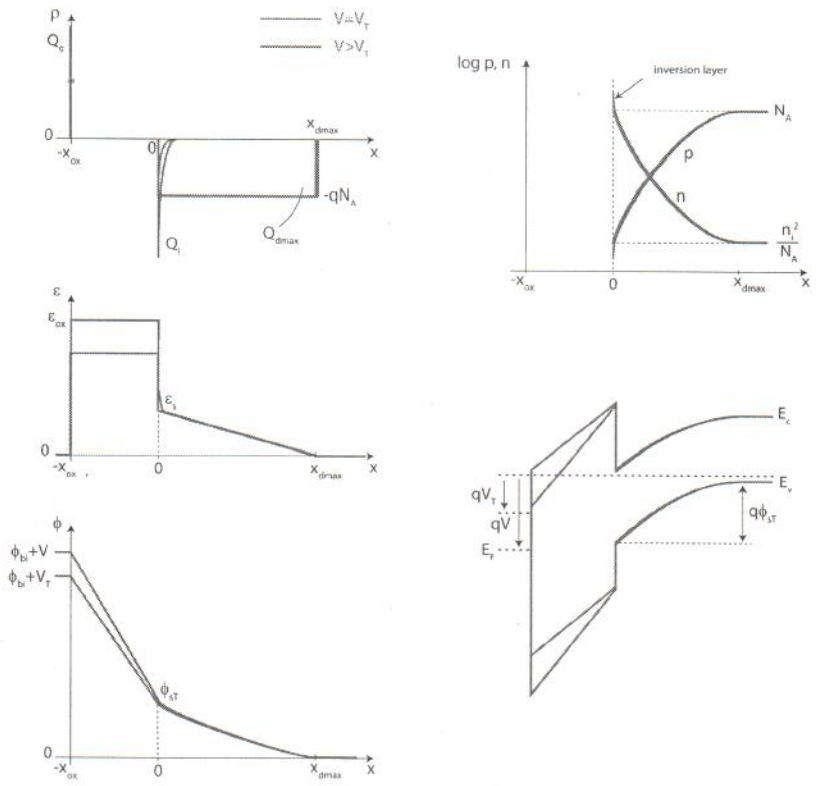


Figure 8.18: From top to bottom and left to right: sketches of volume charge density, electric field, electrostatic potential, carrier concentration and energy band diagram of a MOS structure at zero volts, at threshold and in inversion.

While we have obtained important and useful results using a simple view of the electrostatics of the MOS structure, a more rigorous analysis is often desirable. Advanced Topic AT8.1 at the end of this chapter presents the so-called Poisson-Boltzmann formulation of the MOS electrostatics. This approach yields more accurate analytical models for many important aspects of the MOS electrostatics.

Exercise 8.4: *Consider an MOS structure identical to the one studied in Exercise 8.3. Compute V_T and Q_i for $V = 2\ V$.*

We use results obtained for ϕ_{bi}, C_{ox}, and γ from Exercise 8.3. The first step is to compute ϕ_{sT} using Eq. 8.25:

$$\phi_{sT} = 2\frac{kT}{q}\ln\frac{N_A}{n_i} = 2 \times 0.026 \times \ln\frac{6 \times 10^{17}}{1.02 \times 10^{10}} = 0.93\ V$$

We then use Eq. 8.28 to obtain V_T:

$$V_T = V_{FB} + \phi_{sT} + \gamma\sqrt{\phi_{sT}} = -1.0\ V + 0.93\ V + 0.58\ V^{1/2} \times \sqrt{0.93\ V} = 0.49\ V$$

The calculation of Q_i is now straightforward through Eq. 8.31:

$$Q_i = -C_{ox}(V - V_T) = -7.7 \times 10^{-7}\ F/cm^2 \times (2 - 0.49\ V) = -1.2 \times 10^{-6}\ C/cm^2$$

8.4.6 Summary of charge-voltage characteristics

Figs. 8.19 summarizes the charge-voltage characteristics of the MOS structure discussed in this section. On the top, the absolute of the semiconductor charge is graphed against ϕ_s. In the accumulation regime, $Q_s = Q_a > 0$. Under the sheet-charge approximation, in accumulation $\phi_s \simeq 0$. In depletion, $Q_s = Q_d < 0$ and $|Q_d|$ evolves in a square-root form with ϕ_s. In inversion, $Q_s = Q_{dmax} + Q_i < 0$. Under the sheet-charge approximation, $\phi_s \simeq \phi_{sT}$.

The graph in the middle shows ϕ_s vs. V. For $V < V_{FB}$, we have the accumulation regime with $\phi_s \simeq 0$. In depletion, ϕ_s evolves with V according to Eq. 8.21. In inversion, for $V > V_T$, $\phi_s \simeq \phi_{sT}$.

At the bottom of Fig. 8.19 the absolute of the semiconductor charge is graphed against the voltage. For $V < V_{FB}$, the structure is in accumulation and Q_s is linear on $V_{FB} - V$. Between V_{FB} and V_T, the structure is in depletion and the dependence of Q_s with V follows the square-root form given by Eq. 8.20. For $V > V_T$, the structure is in inversion and Q_i is linear on $V - V_T$.

In the simple model developed in this section, the use of the sheet-charge approximation "pins" the surface potential in accumulation and inversion at 0 and ϕ_{sT}, respectively. In reality, carriers at the semiconductor surface are distributed in depth to some extent. This implies that ϕ_s becomes slightly negative in accumulation and it also increases somehow past ϕ_{sT} in inversion. This is indicated by the dashed lines in Fig. 8.19.

A more rigorous model that accounts for this is developed in Appendix AT8.1. This is referred to as the Poisson-Boltzmann formulation. This model does not make any assumptions about the spatial distribution of carriers at the semiconductor surface. It therefore yields a more accurate description of the electrostatics that yields, among other results, the evolution of carrier charge as a function of surface potential, and the surface potential vs. the applied voltage. These key results are summarized here.

In accumulation, the accumulation charge is found to depend on ϕ_s according to the following

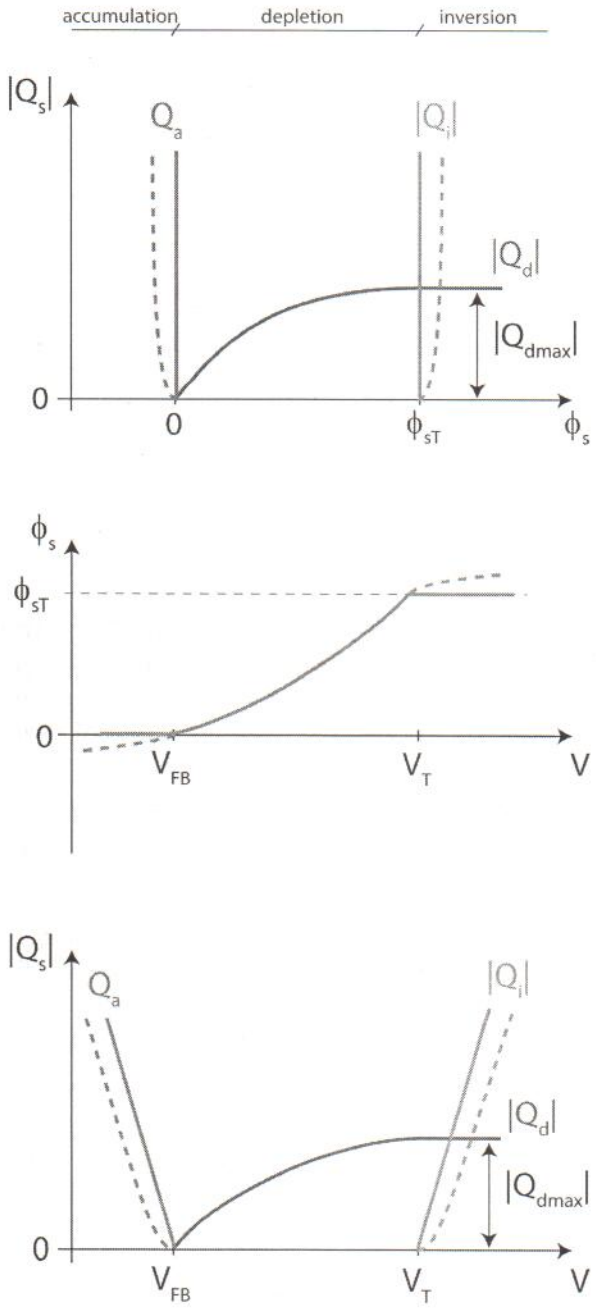


Figure 8.19: Summary of charge-voltage characteristics of MOS structure. Top: Semiconductor charge vs. surface potential. Middle: Surface potential vs. voltage. Bottom: Semiconductor charge vs. voltage. Dashed lines represent improved dependencies obtained from the Poisson-Boltzmann model of Appendix AT8.1.

expression (from Eq. 8.106):

$$Q_a \simeq \sqrt{2\epsilon_s kTN_A}\sqrt{\exp\frac{-q\phi_s}{kT}+\frac{q\phi_s}{kT}-1} \qquad \text{for } \phi_s < 0 \tag{8.32}$$

Notice how Q_a goes to zero, as ϕ_s goes to zero. The leading term of Q_a for strong accumulation ($|\phi_s| \gg kT/q$) depends exponentially on $-q\phi_s/2kT$.

Another important result from the Poisson-Boltzmann formulation is the relationship between ϕ_s and V. In accumulation, this is approximately (from Eq. 8.108):

$$\phi_s \simeq -\frac{kT}{q}\ln\{1+[\frac{C_{ox}(V_{FB}-V)}{\sqrt{2\epsilon_s kTN_A}}]^2\} \qquad \text{for } V < V_{FB} \tag{8.33}$$

At flatband, $\phi_s = 0$, but as $V < V_{FB}$, ϕ_s becomes negative but it does so quite slowly.

In inversion, the dependence of Q_i on ϕ_s is found to be (from Eq. 8.119):

$$Q_i \simeq -\sqrt{2\epsilon_s kTN_A}[\sqrt{\frac{q2\phi_f}{kT}+\exp\frac{q(\phi_s-2\phi_f)}{kT}-1}-\sqrt{\frac{q2\phi_f}{kT}}] \qquad \text{for } \phi_s > 2\phi_f \tag{8.34}$$

In this expression, Q_i goes to zero as ϕ_s approaches $2\phi_f$. In strong inversion, the leading term of Q_i goes as $\sim \exp(q\phi_s/2kT)$.

The dependence of ϕ_s on V is given by (from Eq. 8.122):

$$\phi_s \simeq 2\phi_f + \frac{kT}{q}\ln\{[\frac{C_{ox}(V-V_T)}{\sqrt{2\epsilon_s kTN_A}}-\sqrt{\frac{q2\phi_f}{kT}}]^2-\frac{q2\phi_f}{kT}+1\} \qquad \text{for } V > V_T \tag{8.35}$$

At $V = V_T$, $\phi_s = 2\phi_f$, consistent with our definition of threshold. In inversion, ϕ_s does increase somehow above $2\phi_f$ but it does so in a logarithmic way.

Other interesting results for the MOS electrostatics under the Poisson-Boltzmann formulation are derived in Appendix AT8.1.

8.5 Dynamics of the MOS structure

A MOS structure looks in many ways like a capacitor. On one side there is a metal plate; on the other side there is the semiconductor which is in effect another conducting "plate". In the middle, there is a dielectric, the oxide. Depending on the voltage that is applied to the MOS structure, net electrical charge exists at the metal/dielectric interface and in the semiconductor.

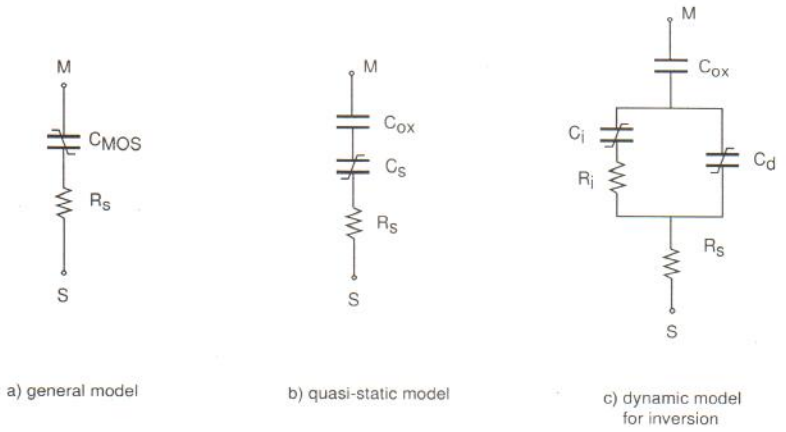


Figure 8.20: Equivalent circuit models for the MOS structure: a) general model; b) quasi-static model; c) dynamic model for inversion.

The total charge at these two plates is of equal magnitude but opposite sign. If the voltage is changed, the charge on both plates is modified by the same amount in absolute terms. At all times, the magnitude of the charge on each plate is identical. The MOS structure behaves as a true capacitor. A simple equivalent model description of the MOS structure consists then of a capacitor, capturing the MOS electrostatics, in series with a resistor that accounts for the finite conductance of the substrate. This is sketched on the left of Fig. 8.20. This simple equivalent circuit provides a reasonable description of the dynamics of the MOS structure.

For the simple *RC* circuit of the left of Fig. 8.20 to be useful, all the dependencies of the MOS capacitance need to be properly described. There are two critical dependences to understand and represent. The first one is the impact of voltage. The second one, perhaps less obvious, is time or the dynamic behavior of the electrostatics themselves. We study these two factors separately in the next two subsections.

8.5.1 Quasi-static C-V characteristics

The capacitance-voltage characteristics of an MOS structure have tremendous practical importance as a diagnostic tool. From the C-V characteristics, one can extract the flat-band voltage, the threshold voltage, the oxide thickness and the semiconductor doping level. Also a detailed study of the C-V characteristics can reveal non-idealities in the structure such as mobile oxide charge, interface states, or gate doping depletion (in the case of a polySi gate). The capacitance-voltage characteristics of a MOS structure are also of capital importance in just about any of its applications. The C-V characteristics summarize the complex behavior of the MOS structure that was described in detail in the preceding sections. Studying in detail the C-V characteristics is an opportunity to review and solidify our understanding of the physics of the MOS structure.

As we have encountered before in this book, a quasi-static situation is one without significant memory effects. Effectively, at all times, the MOS structure behaves as if it was in DC. The term quasi-static C-V characteristics refers to the capacitance that is measured at a certain bias

voltage in a small-signal configuration when the frequency of the driving signal is low enough for the electrostatics to be independent of frequency (how small the frequency needs to be will become clear later). The formalism of the electrostatics of the MOS structure described so far in this chapter completely applies to a quasi-static situation and is therefore the basis for a derivation of the quasi-static C-V characteristics.

A general definition of capacitance of the MOS structure is the change of charge on the gate as a result of a small change in voltage across the structure:

$$C = \frac{dQ_g}{dV} \tag{8.36}$$

Since $Q_g = -Q_s$, using the chain rule for derivatives, the capacitance can be expressed as:

$$C = -\frac{dQ_s}{dV} = -\frac{dQ_s}{d\phi_s}\frac{1}{\frac{dV}{d\phi_s}} \tag{8.37}$$

From the general relationship 8.18, we have:

$$\frac{dV}{d\phi_s} = 1 - \frac{1}{C_{ox}}\frac{dQ_s}{d\phi_s} \tag{8.38}$$

Let us define the semiconductor capacitance, C_s, as:

$$C_s = -\frac{dQ_s}{d\phi_s} \tag{8.39}$$

As we will see below, C_s defined this way is always positive. Armed with this definition, Eq. 8.38 can be rewritten as:

$$\frac{dV}{d\phi_s} = 1 + \frac{C_s}{C_{ox}} \tag{8.40}$$

Substituting this in Eq. 8.37, we get:

$$C = \frac{1}{\frac{1}{C_{ox}} + \frac{1}{C_s}} \tag{8.41}$$

There is a simple physical interpretation to the result embodied in Eqs. 8.39 and 8.41. The capacitance of the MOS structure can be thought of as two capacitances in series, one associated with the oxide, C_{ox}, and another one associated with the semiconductor C_s. The definition of C_s of Eq. 8.39 is consistent with this interpretation since it represents the change of the charge in the semiconductor as a result of a change in the electrostatic potential build-up across the semiconductor. This simple picture of two capacitors in series is schematically illustrated at the center of Fig. 8.20.

We now derive first-order expressions for C_s in the various regimes. In accumulation, $Q_s = Q_a$. Eq. 8.32 gives Q_a vs. ϕ_s. From this, we can easily derive C_s in accumulation. For strong accumulation ($|\phi_s| \gg kT/q$), we find:

$$C_s = -\frac{dQ_a}{d\phi_s} \simeq \frac{q}{2kT} C_{ox}(V_{FB} - V) \qquad \text{for } V < V_{FB} \tag{8.42}$$

where we have used Eq. 8.24. In strong accumulation, C_s increases linearly with the gate voltage below flatband.

Close to flatband, when $\phi_s \simeq 0$, the semiconductor capacitance is:

$$C_s = -\frac{dQ_a}{d\phi_s} \simeq \frac{\epsilon}{L_D} \qquad \text{for } V \simeq V_{FB} \tag{8.43}$$

where L_D is the Debye length. This is a particularly intuitive result since we know that the Debye length gives a sense of the distance over which small perturbations to charge neutrality in an extrinsic semiconductor are screened out.

In depletion, $Q_s = Q_d$. The dependence of the depletion charge with ϕ_s is given by Eq. 8.17. Using also the dependence of ϕ_s on V given by Eq. 8.21 and the definition of γ in Eq. 8.16, we can obtain:

$$C_s = -\frac{dQ_d}{d\phi_s} \simeq \frac{C_{ox}}{\sqrt{1 + 4\frac{V - V_{FB}}{\gamma^2}} - 1} \qquad \text{for } V_{FB} < V < V_T \tag{8.44}$$

This expression for the semiconductor capacitance in depletion can also be obtained if we realize that:

$$C_s \simeq \frac{\epsilon_s}{x_d} \qquad \text{for } V_{FB} < V < V_T \tag{8.45}$$

and we use the expression of $x_d(V)$ given in Eq. 8.19. In the depletion region, there is a depletion region underneath the oxide. When the surface potential changes, the edge of the depletion region moves. This is at a distance x_d away from the semiconductor/oxide interface.

Looking at the Poisson-Boltzmann formulation in Appendix AT8.1, Eq. 8.44 is only valid for $\phi_s \gg kT/q$ or with V not to close to V_{FB}. As V goes to zero, this equation cleearly diverges. Using the more rigorous model of Appendix AT8.1, we can easily find that in the depletion regime for V close to zero, the semiconductor capacitance approaches the value given by Eq. 8.43.

In inversion, $Q_s = Q_i$. Eq. 8.34 gives the dependence of Q_i on ϕ_s. We also use Eq. 8.31 to get:

$$C_s = -\frac{dQ_i}{d\phi_s} \simeq \frac{q}{2kT} Q_i = \frac{q}{2kT} C_{ox}(V - V_T) \qquad \text{for } V > V_T \tag{8.46}$$

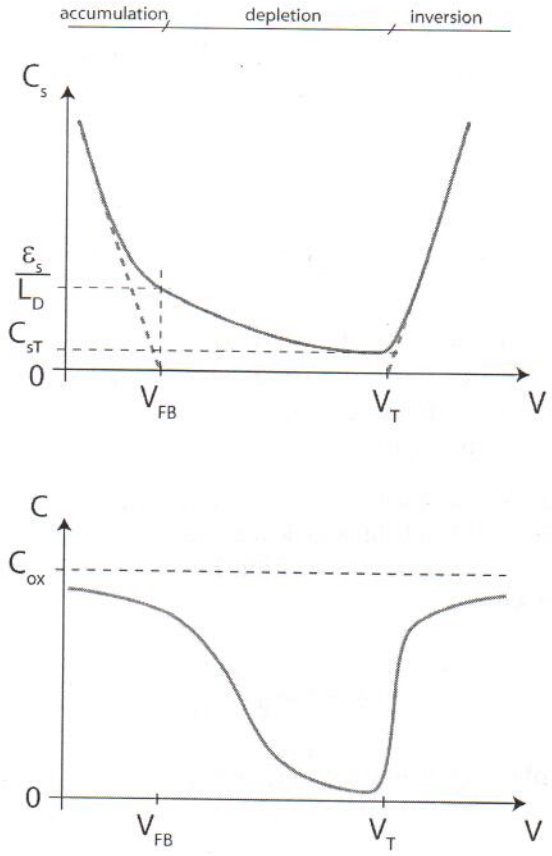


Figure 8.21: Top: Sketch of semiconductor capacitance of MOS structures in different regimes. Bottom: Overall MOS capacitance vs. voltage.

The evolution of the semiconductor capacitance with V is show at the top of Fig. 8.21. In accumulation and inversion, C_s increases in a linear fashion with V. In depletion, it evolves in a square-root form. The minimum capacitance is obtained around threshold.

The overall MOS capacitance is obtained by inserting these expressions for C_s into Eq. 8.41. In accumulation, we have:

$$C \simeq C_{ox} \frac{1}{1 + \frac{2kT}{q(V_{FB}-V)}} \qquad \text{for } V < V_{FB} \tag{8.47}$$

Around flatband, we get:

$$C \simeq C_{ox} \frac{1}{1 + C_{ox}\frac{\epsilon_s}{L_D}} \qquad \text{for } V \simeq V_{FB} \tag{8.48}$$

In depletion, we obtain:

$$C \simeq \frac{C_{ox}}{\sqrt{1+4\frac{V-V_{FB}}{\gamma^2}}} \qquad \text{for } V_{FB} < V < V_T \tag{8.49}$$

In inversion, we get:

$$C \simeq C_{ox}\frac{1}{1+\frac{2kT}{q(V-V_T)}} \qquad \text{for } V > V_T \tag{8.50}$$

The bottom of Fig. 8.21 sketches C vs. V. Since we are looking at two capacitances in series, the smaller one tends to dominate. Hence, under strong accumulation and inversion, the overall MOS capacitance approaches C_{ox}. In depletion, C_s decreases as V increases beyond V_{FB}, and the overall capacitance can become quite small.

The capacitance of the semiconductor in the depletion regime at threshold is of special significance as it enters in models for the subthreshold behavior of the MOS structure which we study below. An expression for this can be obtained from Eq. 8.44 by setting $V = V_T$. After some algebra, this can be written as:

$$C_{sT} \simeq C_{ox}\frac{\gamma}{2\sqrt{2\phi_f}} \tag{8.51}$$

This result can also be obtained if we realize that $C_{sT} = \epsilon_s/x_{dmax}$ and use the expression for x_{dmax} in Eq. 8.26.

8.5.2 High-frequency C-V characteristics

The quasi-static C-V characteristics described above are obtained for a driving signal with low-enough frequency. If a high frequency is employed, the physics change. This is described in this section. Let us examine the three voltage regimes separately.

The application of an AC voltage modulates the charge in the semiconductor. In the accumulation regime, this charge is made out of holes in the accumulation layer. Since holes are majority carriers in the semiconductor, they are in good communication with the outside world through the ohmic contact. The proper time constant to get holes in and out of the accumulation layer is the dielectric relaxation time of the substrate or the RC time constant of the whole structure, whichever is largest. Either case, it is a rather fast phenomenon. In consequence, the C-V characteristics in accumulation derived in the previous section apply up to fairly high frequencies. The situation is identical in depletion, where the hole concentration is modulated at the edge of the depletion region in response to the application of an AC voltage. The C-V characteristics in depletion obtained above also apply up to rather high frequencies.

The situation is quite different in inversion. As we discussed above, in strong inversion, the electron concentration in the inversion layer is modulated in response to the AC voltage. The thickness of the depletion region is basically unchanged. The extra electrons required to satisfy

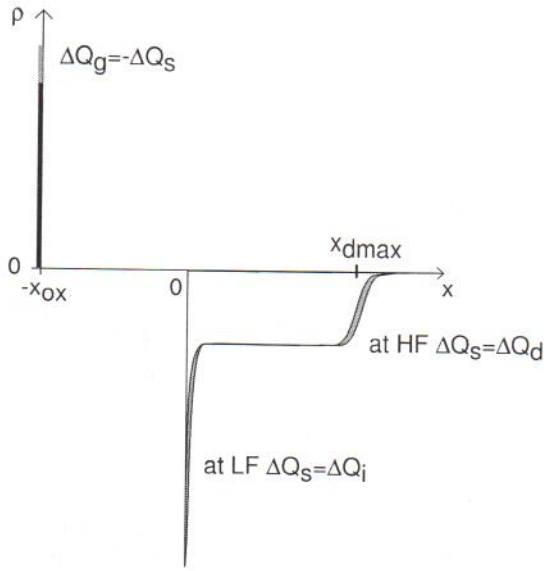


Figure 8.22: Illustration of change in charge in semiconductor in response to an AC voltage signal at low frequencies and high frequencies.

the electrostatics have to be generated in the semiconductor. The proper time constant for this is the generation lifetime, typically in the $\mu s - ms$ range. In consequence, the C-V characteristics derived above only apply if the AC signal changes slowly in the scale of the inverse of τ_g, this typically means frequencies in the kHz range or below. For higher frequencies, the inversion capacitance picture ought to look quite different from what has been discussed so far.

In fact, if the frequency is high enough, the electron concentration in the inversion layer is unable to respond to the AC voltage signal. Since the voltage is nevertheless changing, charge modulation must take place somewhere to satisfy the electrostatics. The only other place where charge can be modulated is at the edge of the depletion region with the substrate. This can be accomplished by getting holes in and out of the substrate through the ohmic contact, a process that as we discussed above, is limited by the RC time constant of the structure. This is sketched in Fig. 8.22.

This understanding allows us to conclude that the high-frequency inversion semiconductor capacitance is, to the first order, equal to the depletion capacitance at threshold (derived in Eq. 8.51):

$$C_{s,HF} \simeq C_{sT} \tag{8.52}$$

This is because in inversion $\phi_s \simeq \phi_{sT}$ and $x_d \simeq x_{dmax}$. C_{sT} can be a substantially lower value than the quasi-static capacitance of the structure for the same bias point (given by Eq. 8.46).

A comparison of the theoretical C-V characteristics of a typical MOS structure at low-frequency and high frequency is shown in Fig. 8.23. In inversion, C remains at the lowest value

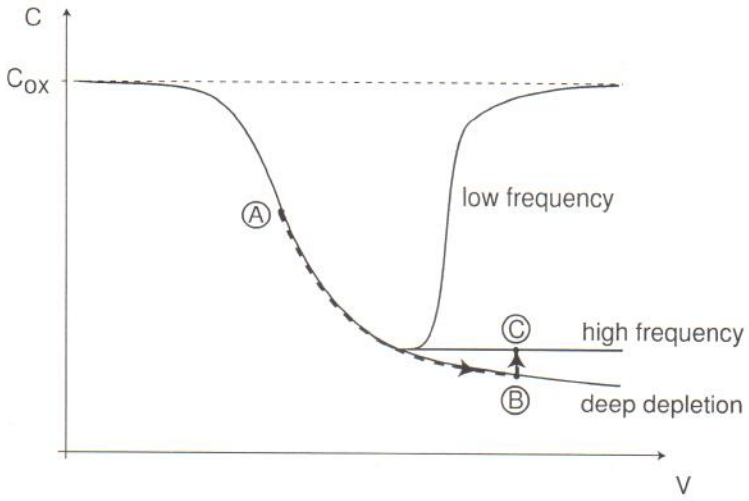


Figure 8.23: C-V characteristics of typical MOS structure at low-frequency (quasi-static), high frequency, and under a fast ramp or after a step (deep depletion). The marked points refer to the transient described in Fig. 8.25.

attained at threshold. Experimental results showing the evolution of the C-V characteristics of a typical MOS with frequency are depicted in Fig. 8.24.

An equivalent circuit description of the high frequency behavior of the MOS structure is shown on the right of Fig. 8.20. In this model, we have broken the semiconductor capacitance in two capacitors in parallel: one associated with the depletion region C_d, and another one associated with the inversion layer C_i. In series with the inversion layer capacitance is a resistor R_i. This resistor captures the difficulty of providing electrons to the inversion layer. The $R_i - C_i$ branch behaves as a low-pass filter. Its time constant, $R_i C_i$, is the inverse of the generation lifetime τ_g that controls electron generation in the depletion region. For low frequencies, the impedance of this branch is dominated by C_i. Since it is in parallel but it is much larger than C_d, the capacitance of the MOS structure is approximately C_i. At high frequencies, on the other hand, the impedance is dominated by R_i. In consequence, the capacitance of the MOS structure is approximately C_d.

8.5.3 Deep depletion

The previous section has revealed the rich dynamics of the MOS structure in the inversion regime that originate from the "sluggishness" of response of the inversion layer to changes in applied voltage. The condition just studied was one in which a small high-frequency signal was applied on top of a DC bias. Another dynamic situation of great interest is one that results when a MOS structure is suddenly switched into inversion by the application of a step or a fast voltage ramp. This condition is called deep depletion.

Consider a MOS structure at zero bias. Let us assume that at $V = 0$ it is in depletion. Let us consider what happens if at $t = 0$, we apply a voltage step, as sketched in Fig. 8.25. If the step

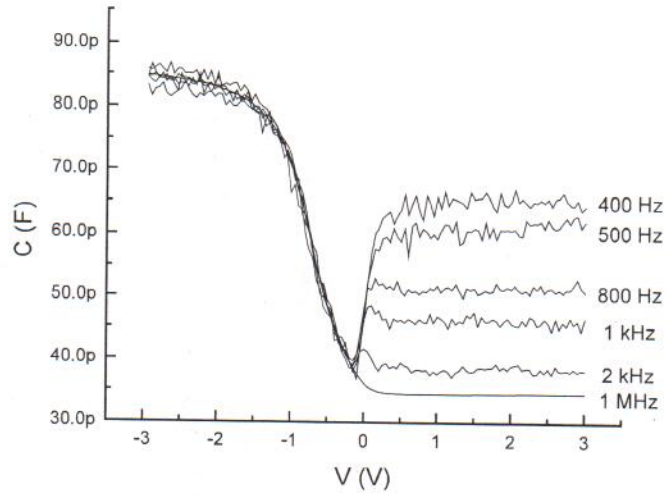


Figure 8.24: Experimental C-V characteristics of an MOS structure as a function of frequency. At low frequency, the measurement becomes noisy and in general it is not possible to obtain the ideal quasi-static behaviour sketched in Fig. 8.21 [*courtesy of T. Payakapan, MIT*].

is negative and of a magnitude that exceeds the flat-band voltage, the structure will be driven into accumulation. After an RC time constant, a relatively short time, the depletion region will be wiped out and holes will accumulate at the semiconductor/oxide interface. If the step is negative but of a magnitude smaller than V_{FB} or positive but smaller than V_T, the depletion region thickness will change. This also happens in a time scale of the corresponding RC time constant.

If the voltage step is positive and exceeds the threshold voltage, the situation gets complicated. A short time after the application of the step (following the inevitable RC delay), the only way to satisfy the electrostatics of the structure is for the depletion region to widen beyond x_{dmax}. This is because even though the voltage is such that an inversion layer should be formed, there has not been enough time for the electrons to be generated. For this to happen, we must wait a time of order of the generation lifetime, τ_g, a fairly long time in the scale of many useful electronic functions. The MOS structure has gone into what is called "deep depletion" because for a while the depletion region thickness widens beyond the value that corresponds to the inversion regime. If the step has enough duration, eventually the required electrons will be generated, the inversion layer will be formed and the depletion region will collapse down to x_{dmax}. This is sketched in Fig. 8.25.

The dynamics of the deep depletion regime can be observed experimentally by measuring the capacitance of the MOS structure during the transient. This can be accomplished by overlapping a high-frequency sinusoidal signal on top of the voltage step. The result is sketched in Fig. 8.25. After the RC transitory has died off but before a significant amount of electrons has been generated, the capacitance in deep depletion will be lower than the high-frequency capacitance in the inversion regime. This is because $x_d > x_{dmax}$. Given enough time, the inversion layer will eventually grow, the depletion region will collapse back to x_{dmax} and the capacitance will revert

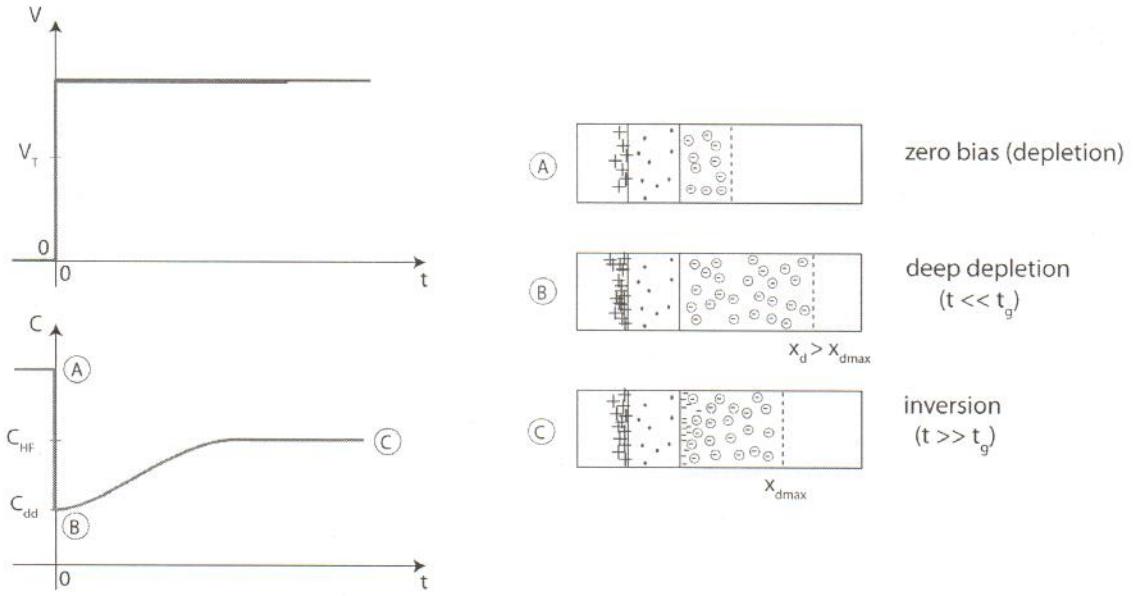


Figure 8.25: Evolution of the electrostatics of the MOS structure after the application of a voltage step exceeding V_T. The marked points correspond to the locations indicated in Fig. 8.23. Immediately following the application of the step, the inversion layer does not have enough time to form and the depletion region widens beyond x_{dmax}. This is called *deep depletion*. Given enough time, electrons are generated, the inversion layer fully forms and the depletion region collapses to x_{dmax}. If the capacitance is being measured as a function of time by means of a high-frequency AC voltage overlapping the step, we will observe that immediately upon the application of the step, the capacitance drops below its inversion value. As time goes on, the capacitance relaxes to the value corresponding to inversion.

to the high-frequency inversion value discussed in the previous section.

Deriving a model for the electrostatics of the MOS structure in deep depletion is quite simple because it is a fairly straightforward extension of the model for the electrostatics of the depletion regime for $V > V_T$ that we have derived earlier in this chapter. At $t = 0^+$, right after the voltage step is applied, the depletion region widens to a thickness that we can denote as $x_{d,dd}$. In analogy with Eq. 8.19, $x_{d,dd}$ is given by:

$$x_{d,dd} = \frac{\epsilon_s}{C_{ox}}\left(\sqrt{1 + 4\frac{V - V_{FB}}{\gamma^2}} - 1\right) \tag{8.53}$$

Since at $t = 0^+$ the charge in the semiconductor is entirely given by the depletion charge, the semiconductor capacitance at this point in time, is simply given by:

$$C_{s,dd} \simeq \frac{\epsilon_s}{x_{d,dd}} \tag{8.54}$$

The overall capacitance of the MOS structure is then:

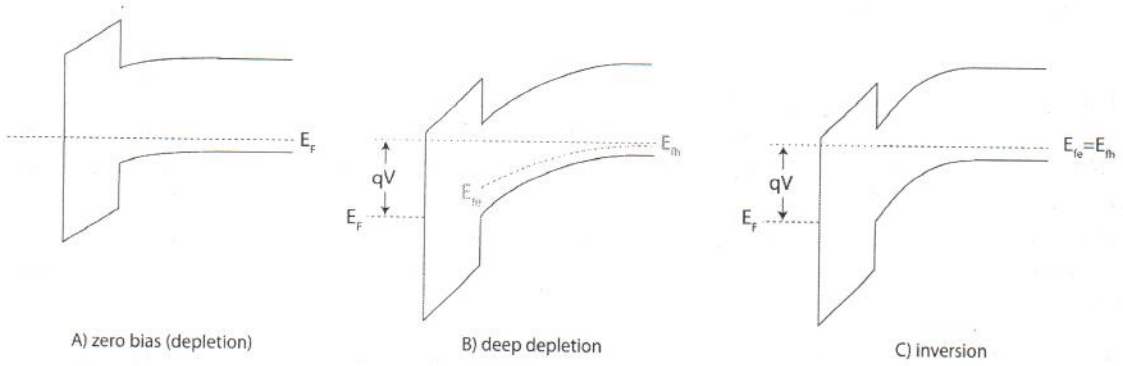


Figure 8.26: Evolution of the energy band diagrams in the MOS structure of Fig. 8.25 at key points in time.

$$C_{dd} = \frac{1}{\frac{1}{C_{ox}} + \frac{1}{C_{s,dd}}} \simeq \frac{C_{ox}}{\sqrt{1 + 4\frac{V - V_{FB}}{\gamma^2}}} \tag{8.55}$$

For $V > V_T$, $x_{d,dd} > x_{dmax}$, $C_{s,dd} < C_{sT}$ and $C_{dd} < C_{HF}$. In deep depletion, the higher the voltage the structure is stepped into, the more the depletion region widens and the lower the capacitance. This is often sketched in the C-V characteristics as in Fig. 8.23.

To further the understanding of the deep depletion regime, Fig. 8.26 illustrates the evolution of the energy band diagrams of the MOS structure at key points in the transient of Fig. 8.25. The band diagrams at points A and C correspond to depletion and inversion and have already been studied before. The interesting new band diagram is the one corresponding to point B or deep depletion. The extended depletion region produces a large band bending in the semiconductor. In spite of this, there is no inversion layer because there has not been enough time for the electrons to be generated. How could these two facts be reconciled in the energy band diagram? Since the electron concentration is given by the distance between the quasi-Fermi level for electrons and the edge of the conduction band, this situation requires that E_{fe} be substantially below E_c at the surface of the semiconductor. At the same time, the hole quasi-Fermi level must remains flat across the structure as these are majority carriers and they respond quickly to the change of voltage across the structure. Interestingly, satisfying these constrains implies that the electron quasi-Fermi level ends up below the hole quasi-Fermi level across the depletion region, as indicated in Fig. 8.26. This is consistent with the occurrence of electron-hole pair generation which we have argued needs to take place for the structure to eventually reach the steady-state situation of point C.

In deep depletion, the semiconductor is out of equilibrium, as in a reverse bias PN junction. The resulting electron generation contributes to growing the inversion layer and the hole generation shrinks the depletion layer. If the body of the semiconductor extends to enough depth, in deep depletion, E_{fe} remain below E_{fh} over a length in the scale of the electron diffusion length.

For deep depletion to occur, there must not be a supply of electrons readily available to form the inversion layer. That is the situation studied here where the electrons need to be generated. As we will see later in this Chapter, in a three-terminal MOS structure in which there is an

n^+-region right next to the channel that can supply the required electrons, deep depletion cannot be produced.

The dynamic behavior of the MOS structure under pulsed conditions that has been discussed in this section is at the heart of Charge-Coupled Devices (CCDs) and CCD-based image sensors (see Prob. 8.28). It is also exploited as a characterization technique. The recovery time from deep depletion in a MOS structure is related to the generation lifetime which itself depends on the purity and defect concentration in the vicinity of the oxide-semiconductor interface. A large concentration of defects leads to a fast recovery while the contrary happens in the MOS structure is of high quality.

Exercise 8.5: *Consider an MOS structure identical to that of Exercise 8.4. Calculate the quasi-static capacitance and the high-frequency capacitance at a bias $V = 2\ V$. Calculate also the capacitance immediately after pulsing the structure from $V = 0$ to $V = 2\ V$.*

Since for this structure $V = 2\ V$ corresponds to inversion, tot he first order, the quasi-static capacitance is simply the oxide capacitance:

$$C \simeq C_{ox} = 7.7 \times 10^{-7}\ F/cm^2$$

The high-frequency capacitance is the parallel of C_{sT} and C_{ox}. C_{sT} is given by Eq. 8.51:

$$C_{sT} = C_{ox}\frac{\gamma}{2\sqrt{2\phi_f}} = \frac{0.58\ V^{1/2}}{2\sqrt{0.93\ V}}\, 7.7 \times 10^{-7}\ F/cm^2 = 2.3 \times 10^{-7}\ F/cm^2$$

Then:

$$C_{HF} = \frac{1}{\frac{1}{C_{ox}} + \frac{1}{C_s}} = \frac{1}{\frac{1}{7.7\times10^{-7}} + \frac{1}{2.3\times10^{-7}}} = 1.8 \times 10^{-7}\ F/cm^2$$

When the structure is pulsed from $V = 0$ to $V = 2\ V$, it goes into deep depletion. In order to compute the capacitance, we must first calculate the extent of the depletion region right after the step. This is given by Eq. 8.53:

$$x_{d,dd} = \frac{\epsilon_s}{C_{ox}}\left(\sqrt{1 + 4\frac{V - V_{FB}}{\gamma^2}} - 1\right) = \frac{1.04 \times 10^{-12}\ F/cm}{7.7 \times 10^{-7}\ F/cm^2}\left(\sqrt{1 + 4\frac{2+1}{0.58^2}} - 1\right) = 6.9 \times 10^{-6}\ cm$$

$C_{s,dd}$ is, then, from Eq. 8.54:

$$C_{s,dd} = \frac{\epsilon_s}{x_{d,dd}} = \frac{1.04 \times 10^{-12}\ F/cm}{6.9 \times 10^{-6}\ cm} = 1.5 \times 10^{-7}\ F/cm^2$$

When put in parallel with C_{ox}, we obtain the capacitance of the whole structure in deep depletion:

$$C_{dd} = \frac{1}{\frac{1}{C_{ox}} + \frac{1}{C_{s,dd}}} = \frac{1}{\frac{1}{7.7\times10^{-7}} + \frac{1}{1.5\times10^{-7}}} = 1.3 \times 10^{-7}\ F/cm^2$$

8.6 Weak inversion and the subthreshold regime

The subthreshold regime, also sometimes called *weak inversion*, refers to the regime of operation of a MOS structure for gate voltages just below threshold. In this range of voltages, the electron concentration at the oxide-semiconductor interface is lower than the acceptor doping level in the bulk. From an electrostatic point of view, the dominant charge originates in the uncompensated acceptors in the depletion region and that is why up to this point, we have labeled this mode of operation the depletion regime. However, immediately below threshold, the electron concentration is of some significance and in a MOSFET, it can result in non-negligible current that is known as the *subthreshold current.*

In modern microelectronics, there is strong emphasis on low-power operation. One of its implications is the need to tightly control the trickle current that inevitably flows through a MOSFET when it is supposed to be OFF. This is often called the *OFF current*, I_{off}, and is a key consideration in device design. In addition, in order to predict power dissipation in a circuit, it is important to model this current correctly. Furthermore, in dynamic circuits, the time that a certain amount of charge can get stored in a dynamic node depends inversely on the leakage current that flows to neighboring nodes. In MOSFET circuits, this is often set by the subthreshold current of the transistors. In order to design devices with minimum OFF current, good physical understanding of the subthreshold regime is required. In this subsection, we develop a first-order model for the total electron concentration in the semiconductor in the subthreshold regime. This will be the basis for the derivation of a model for the subthreshold current of a MOSFET in Ch. 9.

To start, we need to revisit our discussion of the depletion regime in Sec. 8.4.1. Our interest here is in the total electron charge in the semiconductor. This can be obtained using the Boltzmann relationship for electrons as follows [3]:

$$Q_e = -q\int_0^\infty n(x)dx = -q\frac{n_i^2}{N_A}\int_0^\infty \exp\frac{q\phi}{kT}dx = -q\frac{n_i^2}{N_A}\int_{\phi_s}^0 \exp\frac{q\phi}{kT}(\frac{dx}{d\phi})d\phi \tag{8.56}$$

In the last step, we have changed variables from x to ϕ. The fact that the upper limit of the integral goes to $x = \infty$ or $\phi = 0$ is not a problem since its most significant contributions are right next to the oxide-semiconductor interface.

$d\phi/dx$ is the minus electric field in the depletion region. In the depletion regime this can be written in terms of ϕ as:

$$\frac{d\phi}{dx} = -\mathcal{E} = -\sqrt{\frac{2qN_A\phi}{\epsilon_s}} \tag{8.57}$$

We plug this into 8.57 to obtain:

[3] We use here Q_e rather than Q_i to emphasize the important point that the MOS structure is not in inversion.

$$Q_e = -q\frac{n_i^2}{N_A}\sqrt{\frac{\epsilon_s}{2qN_A}}\int_0^{\phi_s}\frac{1}{\sqrt{\phi}}\exp\frac{q\phi}{kT}d\phi \tag{8.58}$$

The major contributions to this integral come around $\phi = \phi_s$. Therefore, we can approximately write:

$$Q_e \simeq -q\frac{n_i^2}{N_A}\sqrt{\frac{\epsilon_s}{2qN_A\phi_s}}\int_0^{\phi_s}\exp\frac{q\phi}{kT}d\phi \simeq -\frac{kT}{q}\sqrt{\frac{q\epsilon_s N_A}{2\phi_s}}\exp\frac{q(\phi_s - 2\phi_f)}{kT} \tag{8.59}$$

where we have assumed that $\phi_s \gg kT/q$ and we have used the definition of ϕ_f in Eq. 8.25.

We obtain here the important result that in the subthreshold regime, $|Q_e| \sim \exp\frac{q\phi_s}{kT}$, which makes good sense. The rigorous numerical calculation presented in Appendix AT8.1 confirms the validity of this result.

We now need to relate ϕ_s to V. We could use Eq. 8.21 which is valid in the depletion regime. However, since we are interested in a relatively small range of voltages immediately below threshold, we can obtain a simpler expression by linearizing the voltage dependence of ϕ_s around threshold in the following way:

$$\phi_s - \phi_{sT} \simeq \phi_s - 2\phi_f \simeq \frac{d\phi_s}{dV}|_T(V - V_T) \tag{8.60}$$

The derivative $d\phi_s/dV$ around threshold can be obtained from Eq. 8.40:

$$\frac{d\phi_s}{dV}|_T = \frac{1}{\frac{dV}{d\phi_s}|_T} = \frac{1}{1 + \frac{C_{sT}}{C_{ox}}} \tag{8.61}$$

Substituting this into Eq. 8.59, approximating $\phi_s \simeq 2\phi_f$ inside the square root and using the definition of C_{sT} in Eq. 8.51, we finally get:

$$Q_e \simeq -\frac{kT}{q}C_{sT}\exp\frac{q(V - V_T)}{(1 + \frac{C_{sT}}{C_{ox}})kT} \qquad \text{for } V \leq V_T \tag{8.62}$$

Eq. 8.62 is a very important result. Below threshold, $|Q_e|$ drops exponentially with V. The sharpness of this drop is determined by the ratio of C_{sT} to C_{ox}. The lower this ratio, the sharper the drop. We can understand this result if we realize that, differentially, we are in front of a capacitive divider with C_{ox} and C_{sT} in series, and ϕ_s being the potential at the common node (see inset in Fig. 8.27). The higher C_{ox} is with respect to C_{sT}, the more closely ϕ_s follows changes in V. Or, in other words, the tighter the electrostatic control of the gate over the surface potential, the sharper the inversion layer can be turned on and off. The factor multiplying kT in the denominator is often called the *ideality factor* of the subthreshold regime. It is clear that it is always larger than unity.

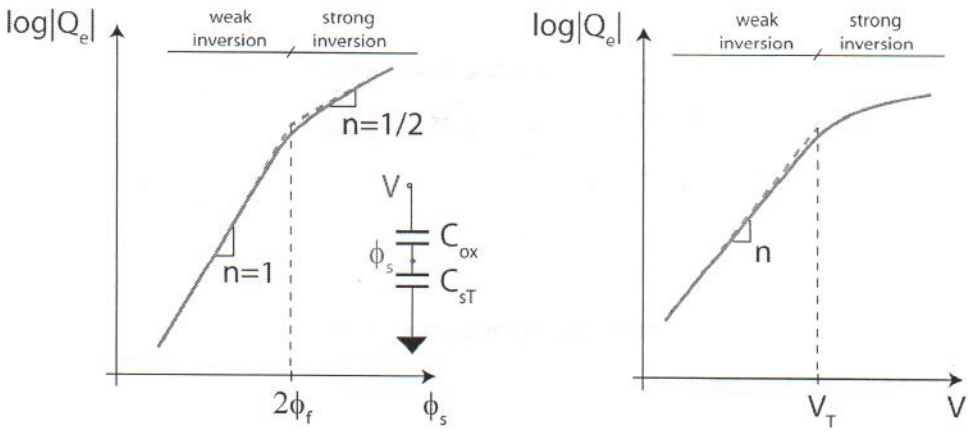


Figure 8.27: Sketches of logarithm of electron charge at MOS oxide-semiconductor interface from weak inversion to strong inversion. Top: vs. surface potential. Bottom: vs. voltage. Indicated is the ideality factor of the exponential regions.

Fig. 8.27 sketches the log of $|Q_e|$ vs. ϕ_s on the left and vs. V on the right from weak inversion to strong inversion. In weak inversion, the integrated electron charge increases as $\sim \exp(q\phi_s/kT)$. In contrast, in strong inversion, the dependence is $\sim \exp(q\phi_s/2kT)$. In weak inversion, $|Q_e|$ increases exponentially with V with an ideality factor that is greater than 1, while in strong inversion, $|Q_e|$ increases linearly with $V - V_T$.

The sharpness of the drop of $|Q_e|$ with V below threshold is characterized by the so-called *inverse subthreshold slope* or more commonly referred to as the *subthreshold swing*, S. This is defined as the voltage required to increase the magnitude of Q_e by a decade. From Eq. 8.62, we have:

$$S = n\frac{kT}{q}\ln 10 = (1 + \frac{C_{sT}}{C_{ox}})\frac{kT}{q}\ln 10 \tag{8.63}$$

S is usually given in mV/dec. At best, if $n = 1$, $S = 60\ mV/dec$ at room temperature. Typically, S is higher than that, meaning, it is less sharp than the ideal value.

Exercise 8.6: *Consider a MOS structure such as the one of Exercises 8.4. Calculate the subthreshold swing S at room temperature.*

First, we compute the ideality factor n:

$$n = 1 + \frac{C_{sT}}{C_{ox}} = 1 + \frac{\gamma}{2\sqrt{2\phi_f}} = 1 + \frac{0.58}{2\sqrt{0.93}} = 1.3$$

The subthreshold swing is then obtained from Eq. 8.63. At room temperature:

$$S = n\frac{kT}{q}\ln 10 = 1.3 \times 0.026\ V \times \ln 10 = 0.078\ V = 78\ mV/dec$$

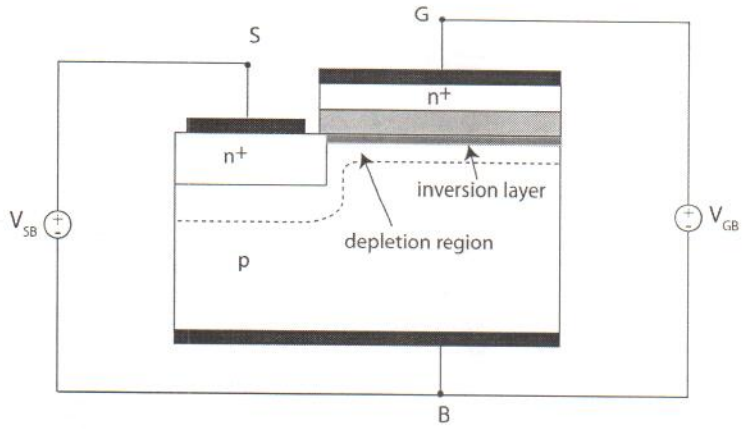


Figure 8.28: Schematic diagram of a MOS structure with a contact to the inversion layer.

8.7 Three-terminal MOS structure

The great usefulness of the MOS structure in modern electronics comes from the possibility of contacting the inversion layer *independently* from the rest of the structure. This allows its use as a conducting "channel" with a conductivity that can be controlled through the gate voltage. This is the essence of the MOSFET. Making a contact to the inversion layer is not difficult. One only has to place an n^+-doped region right next to the edge of the gate, as sketched in Fig. 8.28. This n^+-region communicates with the outside world through an ohmic contact. It is important that there be a small overlap between the n^+-region and the actual MOS structure so that a reliable, low-resistance connection to the inversion layer is obtained. A feature of this three-terminal MOS structure is that the depletion regions of the n^+-p region and the MOS structure overlap.

A contact to the inversion layer allows the application of a voltage between the inversion layer and the semiconductor. This enables the manipulation of the surface potential in the semiconductor and the inversion layer charge independently from the gate. Before we examine the consequences of doing this, let us label the voltages. The n^+-region that provides access to the inversion region is called the *source*, since it will play the role of source of electrons for the MOSFET. The semiconductor substrate is often referred to also as the *body*. The metal continues to be denoted as *gate*. Therefore, the source to body voltage is denoted as V_{SB} and the gate-to-body voltage is labeled as V_{GB}.

The range of voltages that in practice we might be interested in applying to the inversion layer is restricted by the presence of the n^+-p source-body junction diode. Under forward bias, substantial forward diode current can flow, a largely undesirable condition in most applications. Hence, we will discuss the physics of the situation for reverse bias applied to the n^+-p junction, that is, for $V_{SB} \geq 0$. Our models, however, also apply for $V_{SB} \leq 0$. In this discussion, we will keep V_{GB} constant and at a level such that an invesion layer is present under all conditions.

The best way to discuss the effect of $V_{SB} \geq 0$ is to sketch appropriate energy band diagrams, as done in Fig. 8.29. The left of this figure shows a diagram that corresponds to a MOS structure

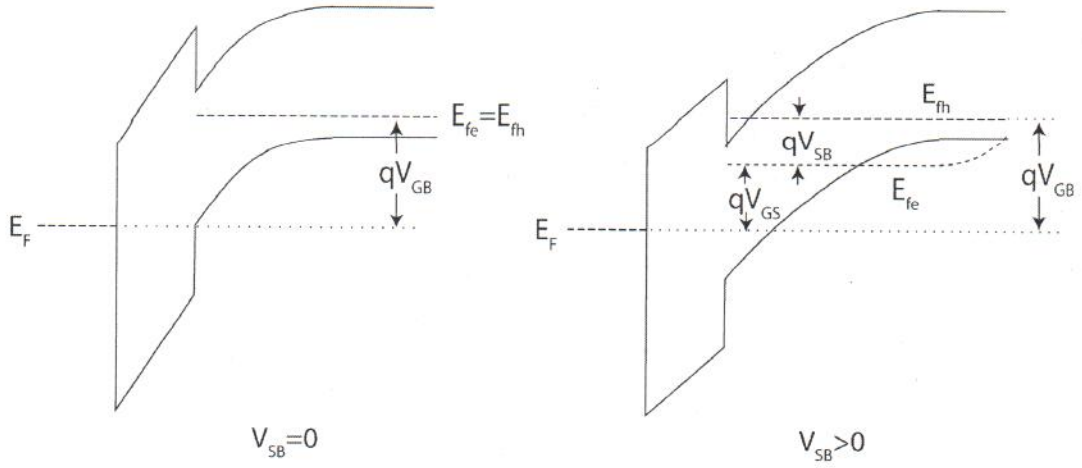


Figure 8.29: Energy-band diagram of three-terminal MOS structure with $V_{SB} = 0$, (left), and $V_{SB} > 0$ (right).

in inversion with $V_{SB} = 0$. The right diagram shows the same structure under an identical value of V_{GB}, but with $V_{SB} > 0$. On the left, $E_{fe} = E_{fh}$ throughout the semiconductor body which is in a quasi-equilibrium estate. On the right, this is not the case. Through the n^+-region, the ohmic contact "grabs" on the majority carrier quasi-Fermi level of the inversion layer, that is, E_{fe}. Since the substrate contact holds onto the hole quasi-Fermi level in the body, the Fermi level is split across the semiconductor by qV_{SB} in the manner that is sketched in the diagram. It is clear that a consequence of this is that the surface potential, the total potential difference across the semiconductor, increases by about the same amount, that is:

$$\phi_{sT}(V_{SB}) = \phi_{sT}(V_{SB} = 0) + V_{SB} \tag{8.64}$$

The most interesting question to try to answer is what happens to the absolute magnitude of the inversion layer charge, *i.e.* does it change, and if so, does it increase or decrease? It is clear from Fig. 8.29 that the charge in the depletion layer has to increase in absolute terms in order to accrue a higher surface potential. This happens through a widening of the depletion region. It is less evident how the inversion charge evolves. In order to answer this, it is convenient to look in some detail at the changes that occur in the electrostatics. These are sketched in Fig. 8.30.

It is easiest to start with the third diagram from the top in Fig. 8.30 which sketches the potential distribution across the structure. Since V_{GB} is unchanged, the total potential build up from gate to substrate contact remains unchanged and equal to $\phi_{bi} + V_{GB}$. The fact that ϕ_s has increased upon the application of $V_{SB} > 0$ implies that ϕ_{ox} must have decreased. This in turn means a reduction of $\mathcal{E}_{ox}$ and as a consequence of $\mathcal{E}_s$. Gauss's law requires that the magnitude of the total charge in the semiconductor, $|Q_s|$, must also decrease. Since $Q_s = Q_i + Q_d$, and we argued above that $|Q_d|$ was increased, the only possible way to reconcile the picture is for $|Q_i|$ to be reduced. In summary, the application of $V_{SB} > 0$ results in a reduction of the electron concentration in the inversion layer.

An important consequence of the application of a voltage between the inversion layer and

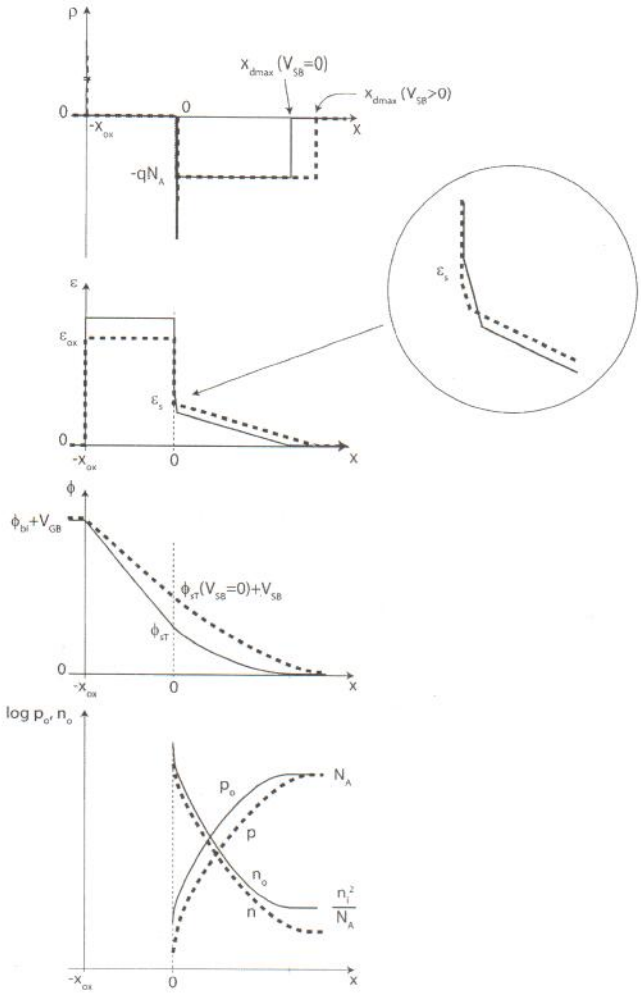


Figure 8.30: Evolution of the electrostatics of the three-terminal MOS structure upon the application of $V_{SB} > 0$.

the body is the split of quasi-Fermi levels in the semiconductor. In particular, for $V_{SB} > 0$, $E_{fe} < E_{fh}$. This is a situation in which the carrier concentrations are driven *below* the thermal equilibrium values, as sketched in the bottom diagram of Fig. 8.30. As we learned in Ch. 4, the semiconductor reacts by trying to reestablish equilibrium, that is, by generating electron-hole pairs. A consequence of this is the presence of a small leakage current flowing through the source and body terminals. This adds up to the reverse current of the n^+-p source-body junction. The magnitude of this current depends on the nature and concentration of traps inside the depletion region. As in a reverse-biased PN junction, the minority carrier concentration gets reduced in the body below equilibrium over a distance of the order of the diffusion length.

The impact of V_{SB} on the electrostatics of the MOS structure can be viewed also as producing a shift in the threshold voltage. Since the application of $V_{SB} > 0$ results in a reduction of the magnitude of the inversion charge that corresponds to a certain V_{GB}, this can be captured by an

increase in the threshold voltage. Through equation 8.31, we see that this produces the desired result. It is straightforward to derive a new expression for V_T that accounts for the impact of V_{SB}. Since the application of V_{SB} increases ϕ_s by V_{SB}, at threshold to the first order, the surface potential goes from ϕ_{sT} to $\phi_{sT} + V_{SB}$. Hence, the expression of the threshold voltage of Eq. 8.28 becomes:

$$V_T^{GB} = V_{FB} + \phi_{sT} + V_{SB} + \gamma\sqrt{\phi_{sT} + V_{SB}} \tag{8.65}$$

This can also be written as:

$$V_T^{GB}(V_{SB}) = V_T^{GB}(V_{SB} = 0) + V_{SB} + \gamma(\sqrt{\phi_{sT} + V_{SB}} - \sqrt{\phi_{sT}}) \tag{8.66}$$

Since we are now in front of a device with three terminals, it is important to specify the two terminals that the threshold voltage refers to. This is reflected in the new notation used in the equations above. For MOSFET operation, we are not as much interested in the gate-to-body threshold voltage, V_T^{GB}, as we have been considering until now, but in the gate-to-source threshold voltage V_T^{GS}. Since $V_{GB} = V_{GS} + V_{SB}$, the two threshold voltages are related by:

$$V_T^{GS}(V_{SB}) = V_T^{GB}(V_{SB}) - V_{SB} = V_{FB} + \phi_{sT} + \gamma\sqrt{\phi_{sT} + V_{SB}} \tag{8.67}$$

This can be easily rewritten as:

$$V_T^{GS}(V_{SB}) = V_T^{GS}(V_{SB} = 0) + \gamma(\sqrt{\phi_{sT} + V_{SB}} - \sqrt{\phi_{sT}}) \tag{8.68}$$

which increases as V_{SB} increases.

As V_T shifts with the application of V_{SB}, the charge in the inversion layer is affected. The charge control relation, Eq. 8.31, must be modifed to account for this. It can be written in two ways, in terms of the gate-to-body voltage:

$$Q_i = -C_{ox}[V_{GB} - V_T^{GB}(V_{SB})] \tag{8.69}$$

or, as more commonly done, in terms of the gate-to-source voltage:

$$Q_i = -C_{ox}[V_{GS} - V_T^{GS}(V_{SB})] \tag{8.70}$$

Both formulations give, of course, the same result.

The phenomenon that we have just discussed, the *back-bias effect* as it is sometimes referred to, has important implications for MOSFET and CMOS circuit operation and is a major concern to device and circuit designers. Three examples follow.

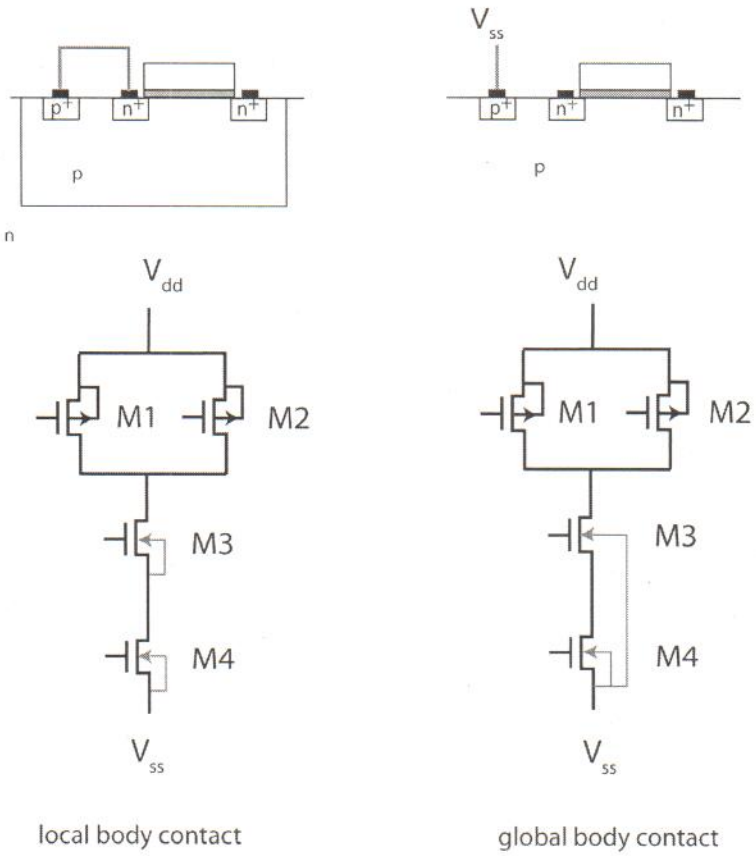


Figure 8.31: Illustration of two approaches to contact bodies of transistors. The left diagrams depict a local body contact where the body of every n-MOS is tied up to its own source. On the right is depicted a global body contact with all bodies of n-MOSFETs tied up to V_{ss}. The bottom diagrams show the schematic of a NAND gate making explicit the body connections.

- The optimum MOSFET technology to be used for a given application depends on the desired circuit topologies and their sensitivity to the back-bias effect. This is illustrated in Fig. 8.31 which shows a typical CMOS NAND logic gate [4]. Let us focus on the n-MOS devices at the bottom. The technology of the left of the figure features a p-well process on an n-type substrate. In consequence, a local body contact is available for every transistor (the p^+-region). The technology on the right has a p-type substrate that is shared among all n-MOSFETs. In this technology, only a global contact is available. Ideal circuit operation calls for each transistor to have its body shorted to its own source. This can be accomplished in the technology on the left, as is made explicit in the diagram, but for the technology on the right, this is not possible. In this case, transistor M4 has its body connected to its source but M3 does not because the source sits at a higher voltage than the bottom rail. There are two deleterious consequences of this. First, the logic gate on the left switches

[4] In logic circuits, the term "gate" is used to refer to elemental circuit configurations that perform simple logic functions. The use of the word "gate" here should not be confused with the metal gate of a transistor.

faster than the logic gate on the right. This is because as the source voltage of transistor M3 on the right circuit rises, its current driving capability is degraded. Second, the operation of the logic gate on the right highly depends on the relative timing of the signals arriving to the gates of transistors M3 and M4 and hence displays a lot more "jitter" or uncertainty in its switching.

- The *body effect* refers to the shift in the threshold voltage along the inversion layer in a MOSFET biased in a current flowing state. This originates in the lateral ohmic drop that occurs in the inversion layer with respect to the body that is produced by the current. The details of this effect will be studied in the next chapter. Suffice it to say here that it is a deleterious effect that results in lower transistor current that one would otherwise obtain.

- The shift in V_T that results from the application of a back-bias can be exploited to turn off a transistor, a logic gate, or entire circuits in a system. In spite of the extra circuit complexity required to implement this concept, this feature is becoming increasingly useful in the development of low-power systems in which certain functions can be disabled in a selective way when they are not needed.

There are several additional consequences of applying a bias to the inversion layer with respect to the source in a MOS structure. For $V_{SB} > 0$, the depletion region widens. This implies that the capacitance associated with the substrate for a given V_{GB} is reduced. In particular, at threshold, C_{sT} is smaller the higher V_{SB} is. An interesting consequence of this is the reduction of the subthreshold swing, that is, the device exhibits sharper subthreshold characteristics.

One final point about the energy band diagram on the right of Fig. 8.29. In this diagram, it is clear that the quasi-Fermi level for electrons is below the quasi-Fermi level for holes in the top region of the semiconductor. This will extend an electron diffusion length into the substrate. A consequence of this is that the carrier concentrations are below the thermal equilibrium values and the semiconductor responds by generating electron-hole pairs. The generated holes are swept towards the body and the generated electrons are extracted by the inversion layer. Hence, in a three-terminal MOS structure biased in this way, there is a small current that enters the source and flows out of the body. This adds to the current associated with the reverse-biased source-body PN junction.

Exercise 8.7: *Consider a MOS structure such as the one of Exercises 8.4 with an additional contact to the source that allows the application of V_{SB}. Calculate ϕ_{sT}, V_T^{GB}, V_T^{GS} and C_{sT} for $V_{SB} = 3\ V$.*

Let us first calculate ϕ_{sT}. From Exercise 8.4, we know that for $V_{SB} = 0$, $\phi_{sT} = 0.93\ V$. With $V_{SB} = 3\ V$, Eq. 8.64 implies:

$$\phi_{sT}(V_{SB} = 3\ V) = \phi_{sT}(V_{SB} = 0) + V_{SB} = 0.93 + 3 = 3.93\ V$$

V_T^{GB} is obtained from Eq. 8.66:

$$\begin{aligned} V_T^{GB}(V_{SB} = 3\ V) &= V_T^{GB}(V_{SB} = 0) + V_{SB} + \gamma(\sqrt{\phi_{sT} + V_{SB}} - \sqrt{\phi_{sT}}) \\ &= 0.5 + 3 + 0.58(\sqrt{0.93 + 3} - \sqrt{0.93}) = 4.1\ V \end{aligned}$$

V_T^{GS} is obtained simply from:

$$V_T^{GS}(V_{SB}) = V_T^{GB}(V_{SB}) - V_{SB} = 4.1 - 3 = 1.1\ V$$

C_{sT} is lowered by the application of V_{SB} because the depletion region under the inversion layer widens. It can be obtained by adapting Eq. 8.51 in the following way:

$$C_{sT}(V_{SB}) = \frac{\gamma C_{ox}}{2\sqrt{2\phi_f + V_{SB}}} = \frac{0.58\ V^{1/2} \times 7.7 \times 10^{-7}\ F/cm^2}{2\sqrt{0.93 + 3\ V}} = 1.1 \times 10^{-7}\ F/cm^2$$

8.8 Summary

- The Si/SiO_2 interface is nearly ideal. Most Si bonds are satisfied at the interface. However, the interface exhibits a roughness of the order of two monolayers.
- At the Si/SiO_2 interface, the semiconductor can swing all the way from accumulation to inversion by selecting a proper metal or through the application of a voltage to the metal with respect to the semiconductor.
- Beyond threshold, the absolute inversion charge increases linearly with voltage.
- To the first order, the surface potential in inversion and accumulation does not change with V.
- In the subthreshold regime, the magnitude of the minority carrier charge at the oxide-semiconductor interface drops exponentially with voltage below threshold.
- The inversion capacitance depends on the frequency of the measurement due to the relatively slow generation lifetime.
- If the MOS structure is driven quickly from below threshold to above threshold, a *deep depletion* condition is established until the carriers that form the inversion layer have had enough time to be generated.

- Application of a body bias with respect to the inversion layer increases the threshold voltage.
- The Poisson-Boltzmann formulation is a powerful tool to examine electrostatics of complex structures in a rigorous way.

8.9 Further reading

Operation and Modeling of the MOS Transistor (Third Edition) by Y. P. Tsividis and C. McAndrew, Oxford, 2011, ISBN 978-0-19-517015-3, TK7871.99.M44T77) is a classic textbook on MOSFET physics and modeling that is now in its third edition. The level of treatment is similar to the one adopted here though it goes in greater depth in some areas. A drawback of earlier editions in the absence of energy band diagrams has now been corrected. Chapters relevant here are Ch. 2 that deals with the two-terminal MOS structure and Ch. 3 that addresses the three-terminal structure. In these chapters, there is a detailed treatment of threshold. It also discusses in some depth the various regimes of inversion. This is an important reference book.

AT8.1 Poisson-Boltzmann formulation of MOS electrostatics

This section presents a general formulation for the electrostatics of the MOS structure under a variety of conditions. The Poisson-Boltzmann formulation, as it is known, allows us to get detailed descriptions of ϕ, $\mathcal{E}$, and ρ, in the oxide and the semiconductor. It also forms the basis for a detailed discussion of the C-V characteristics. A similar formulation can be constructed to treat the electrostatics of the PN junction and metal-semiconductor junction.

We assume here a MOS structure with a geometry identical to that studied in the main body of this chapter. We also assume that the semiconductor is in a quasi-equilibrium condition. We work with Boltzmann statistics. The starting point to solve the electrostatics of this situation is the expression of the volume charge density. For an uncompensated uniformly-doped p-type semiconductor, this can be written as:

$$\rho = q(p - n - N_A) \tag{8.71}$$

This allows to write Poisson's equation in the form:

$$\frac{d^2\phi}{dx^2} = -\frac{q}{\epsilon_s}(p - n - N_A) \tag{8.72}$$

With the semiconductor under quasi-equilibrium, we can use the Boltzmann relations to express the carrier concentrations in terms of the electrostatic potential. Selecting as origin of potentials deep in the bulk, $\phi(bulk) = 0$, we have:

$$n = n_{oB} \exp \frac{q\phi}{kT} \tag{8.73}$$

$$p = p_{oB} \exp \frac{-q\phi}{kT} \tag{8.74}$$

where n_{oB} and p_{oB} are, respectively, the equilibrium electron and hole concentrations in the bulk.

In the bulk, charge neutrality prevails. Mathematically:

$$p_{oB} - n_{oB} - N_A = 0 \tag{8.75}$$

Plugging Eqs. 8.73, 8.74, and 8.75 into 8.72, we get:

$$\frac{d^2\phi}{dx^2} = -\frac{qN_A}{\epsilon_s}[(\exp \frac{-q\phi}{kT} - 1) - \frac{n_i^2}{N_A^2}(\exp \frac{q\phi}{kT} - 1)] \tag{8.76}$$

where we have also assumed that $p_{oB} \simeq N_A$ and $n_{oB} \simeq n_i^2/N_A$.

Eq. 8.76 is called the *Poisson-Boltzmann equation.* It contains a single unknown, the electrostatic potential ϕ. A double integration of this equation entirely solves the problem.

The first integration is performed using the following mathematical equality:

$$\frac{d}{dx}(\frac{d\phi}{dx})^2 = 2\frac{d\phi}{dx}\frac{d^2\phi}{dx^2} \tag{8.77}$$

Multiplying both sides of the differential equation 8.76 by $2\frac{d\phi}{dx}$ and using the relationship just presented, we get:

$$\frac{d}{dx}(\frac{d\phi}{dx})^2 = -\frac{2qN_A}{\epsilon_s}[(\exp\frac{-q\phi}{kT} - 1) - \frac{n_i^2}{N_A^2}(\exp\frac{q\phi}{kT} - 1)]\frac{d\phi}{dx} \tag{8.78}$$

We can now integrate this equation across the semiconductor. We start from the bulk, $x = \infty$ and integrate towards the surface. On the left hand side:

$$\int_\infty^x \frac{d}{dx}(\frac{d\phi}{dx})^2 dx = (\frac{d\phi}{dx})^2|_x - (\frac{d\phi}{dx})^2|_\infty = (\frac{d\phi}{dx})^2 \tag{8.79}$$

$\frac{d\phi}{dx}|_\infty = 0$ because sufficiently far away from the surface in the bulk of the semiconductor, the electric field is zero.

The right-hand side of Eq. 8.78 can also be easily integrated:

$$\begin{aligned}
&-\frac{2qN_A}{\epsilon_s}\int_\infty^x [(\exp\frac{-q\phi}{kT} - 1) - \frac{n_i^2}{N_A^2}(\exp\frac{q\phi}{kT} - 1)]\frac{d\phi}{dx}dx = \\
&-\frac{2qN_A}{\epsilon_s}\int_0^\phi [(\exp\frac{-q\phi}{kT} - 1) - \frac{n_i^2}{N_A^2}(\exp\frac{q\phi}{kT} - 1)]d\phi = \\
&\frac{2kTN_A}{\epsilon_s}[(\exp\frac{-q\phi}{kT} + \frac{q\phi}{kT} - 1) + \frac{n_i^2}{N_A^2}(\exp\frac{q\phi}{kT} - \frac{q\phi}{kT} - 1)]
\end{aligned} \tag{8.80}$$

Assembling now Eqs. 8.79 and 8.80, we get:

$$\frac{d\phi}{dx} = \pm\sqrt{\frac{2kTN_A}{\epsilon_s}}[(\exp\frac{-q\phi}{kT} + \frac{q\phi}{kT} - 1) + \frac{n_i^2}{N_A^2}(\exp\frac{q\phi}{kT} - \frac{q\phi}{kT} - 1)]^{1/2} \tag{8.81}$$

We now have to select the proper sign of this equation. This is best done with the help of Fig. 8.32. When ϕ is positive in the semiconductor, as in depletion and inversion, then the sketch of Fig. 8.32 indicates that $\frac{d\phi}{dx} < 0$. When ϕ is negative, as in accumulation, $\frac{d\phi}{dx} > 0$. In general, we can then write:

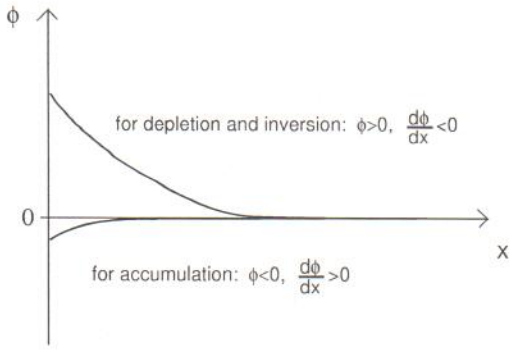


Figure 8.32: Sketch of electrostatic potential in the semiconductor. When ϕ is positive, $d\phi/dx$ is negative; when ϕ is negative, $d\phi/dx$ is positive.

$$sign(\frac{d\phi}{dx}) = -sign(\phi) = -\frac{\phi}{|\phi|} \tag{8.82}$$

We can define the function $F(\phi)$:

$$F(\phi) = \frac{\phi}{|\phi|}[(\exp\frac{-q\phi}{kT} + \frac{q\phi}{kT} - 1) + \frac{n_i^2}{N_A^2}(\exp\frac{q\phi}{kT} - \frac{q\phi}{kT} - 1)]^{1/2} \tag{8.83}$$

Using Eqs. 8.82 and 8.83, we can rewrite Eq. 8.81 in the following way:

$$\frac{d\phi}{dx} = -\sqrt{\frac{2kTN_A}{\epsilon_s}}F(\phi) \tag{8.84}$$

A second integration of Eq. 8.84 gives ϕ vs. x which constitutes a complete solution to the electrostatics problem:

$$\int_{\phi_s}^{\phi} \frac{d\phi}{F(\phi)} = -\sqrt{\frac{2kTN_A}{\epsilon_s}}x \tag{8.85}$$

Unfortunately, an analytical integration of this equation is not possible in general. If some simplifications can be made, analytical formulations for $\phi(x)$ can be derived. This is carried out in the three subsections below.

Even before the additional integration, Eq. 8.84 leads to several useful results. The electric field in the semiconductor is the negative derivative of ϕ in space, hence:

$$\mathcal{E} = \sqrt{\frac{2kTN_A}{\epsilon_s}}F(\phi) \tag{8.86}$$

In particular, at the insulator/semiconductor interface:

$$\mathcal{E}_s = \sqrt{\frac{2kTN_A}{\epsilon_s}}F(\phi_s) \tag{8.87}$$

Using this result in Eq. 8.5 gives us a general expression for the total charge in the semiconductor:

$$Q_s = -\sqrt{2\epsilon_s kTN_A}F(\phi_s) \tag{8.88}$$

We can also break the semiconductor charge into its two components, that is, those due to the electron and hole concentrations being different from their respective bulk values:

$$Q_e = -q\int_0^\infty (n - n_{oB})dx \simeq -q\frac{n_i^2}{N_A}\int_0^\infty (\exp\frac{q\phi}{kT} - 1)dx \tag{8.89}$$

$$Q_h = q\int_0^\infty (p - p_{oB})dx \simeq qN_A\int_0^\infty (\exp\frac{-q\phi}{kT} - 1)dx \tag{8.90}$$

Changing variables from x to ϕ in these two integrals and using Eq. 8.84, one gets:

$$Q_e = q\frac{n_i^2}{N_A}\sqrt{\frac{\epsilon_s}{2kTN_A}}\int_{\phi_s}^0 \frac{\exp\frac{q\phi}{kT} - 1}{F(\phi)}d\phi \tag{8.91}$$

$$Q_h = -\sqrt{\frac{\epsilon_s q^2 N_A}{2kT}}\int_{\phi_s}^0 \frac{\exp\frac{-q\phi}{kT} - 1}{F(\phi)}d\phi \tag{8.92}$$

Finally, we can also plug Eq. 8.88 into Eq. 8.18 and get a relationship between ϕ_s and V:

$$V = -\phi_{bi} + \phi_s + \frac{\sqrt{2\epsilon_s kTN_A}}{C_{ox}}F(\phi_s) = V_{FB} + \phi_s + \gamma\sqrt{\frac{kT}{q}}F(\phi_s) \tag{8.93}$$

If V is given, this equation can be easily solved in an iterative fashion to obtain ϕ_s. Once ϕ_s is known, Eq. 8.87 allows the calculation of the electric field at the oxide/semiconductor interface. Additionally, Eq. 8.88 can be used to get Q_s as a function of V. We can break Q_s on its two components Q_e and Q_h by numerically integrating Eqs. 8.91 and 8.92. A complete description of ϕ vs. x can only be obtained by integrating Eq. 8.85. This also yields n and p vs. x from Eqs. 8.73 and 8.74, and ρ vs. x from Eq. 8.71.

In order to better appreciate how the physics of the MOS structure evolves with applied voltage, it is worthwhile to first examine results obtained from the Poisson-Boltzmann formulation. To this end, Figs. 8.33-8.39 display a complete set of results for a typical case with $x_{ox} = 4.5\ nm$,

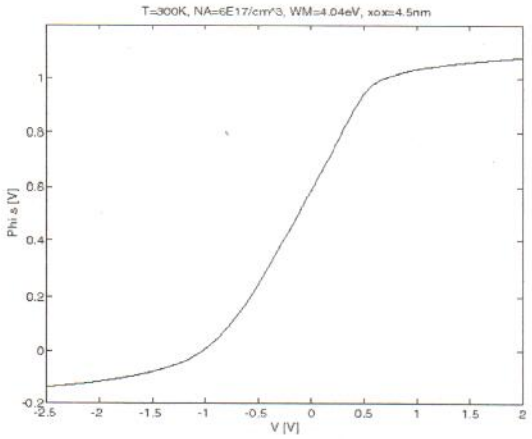


Figure 8.33: Surface potential vs. applied voltage for a typical MOS structure [*calculations courtesy of S. Mertens (MIT)*].

$N_A = 6 \times 10^{17}\ cm^{-3}$, $W_M = 4.04\ eV$, and $T = 300\ K$ as a function of applied voltage. For this case, $V_T = 0.5\ V$, $V_{FB} = -1\ V$, $\phi_{sT} = 0.93\ V$.

Fig. 8.33 graphs ϕ_s vs. V. As V increases, so does ϕ_s. The most prominent feature of this figure is its "S"-shape. For sufficiently negative values of V ($V < -1\ V$), the surface potential ϕ_s tends to saturate to a value that is not too different from zero. This is the accumulation regime. For intermediate values of V ($-1 < V < 0.5\ V$), ϕ_s increases rapidly with V. This is the depletion regime. For sufficiently positive values of V ($V > 0.5\ V$), ϕ_s tends to saturate again. This is the inversion regime.

The left of Fig. 8.34 graphs Q_s, Q_e and Q_h vs. V. Three regimes are apparent again. For negative and slightly positive voltages, $Q_s \simeq Q_h$. These are accumulation and depletion, respectively. For sufficiently positive voltages, Q_h saturates and Q_e starts rising in absolute terms, as we argued it should happen in inversion. Note that Q_e is always negative, while Q_h is both positive (in accumulation) and negative (in depletion and inversion).

The right of Fig. 8.34 shows $\mathcal{E}_s$ vs. V. Again the three regimes are apparent. For negative V, $\mathcal{E}_s$ is negative and depends linearly on V (accumulation). For positive but small V, $\mathcal{E}_s$ becomes positive and displays a sublinear dependence on V (depletion). For positive enough V, $\mathcal{E}_s$ increases linearly again with V (inversion).

The physics of the various regimes is illuminated by examining the evolution of the relevant variables as a function of spatial location in the semiconductor. The left of Fig. 8.35 displays the volume charge density in the semiconductor. For $V < 0$, there is a prominent positive peak at the interface. This is the hole accumulation region. As V is made positive, a negatively charged region grows from the interface into the semiconductor. This is the depletion region. Beyond $V = 0.5\ V$, threshold, the depletion region stops growing. Instead, a prominent negative peak due to the electron inversion layer appears at the oxide/semiconductor interface.

Consistent with this, the electric field in the semiconductor, right of Fig. 8.35, displays a

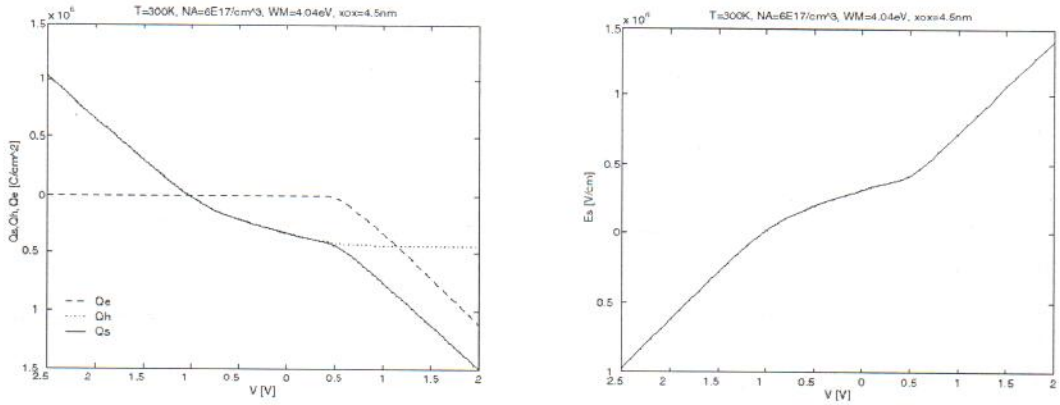


Figure 8.34: Left: Semiconductor charge vs. applied voltage for a typical MOS structure; graphed also are the two components of Q_s: Q_e and Q_h. Electric field at the insulator/semiconductor interface vs. applied voltage for the same MOS structure [*calculations courtesy of S. Mertens (MIT)*].

prominent peak at the interface for negative V, is zero everywhere for $V = -1.0\ V$, but increases and penetrates deeper into the semiconductor as V is made more positive. However, beyond $V = 0.5\ V$, the field in the bulk of the semiconductor does not change much, but increases substantially at the surface. This is consistent with the identification of the various regimes.

The spatial dependence of ϕ is shown on the left of Fig. 8.36 for six selected voltages. For $V = -2\ V$, ϕ is relatively small and negative and confined to the vicinity of the interface. This is accumulation. For $V = -1.0\ V$, $\phi = 0$ everywhere. This is flatband. As V becomes positive, ϕ increases (depletion) but beyond $V = 0.5\ V$ (threshold) it does not change much.

The right of Fig. 8.36 shows the energy band diagrams in the semiconductor for several bias points. At flatband, the bands are flat. In accumulation, the bands bend up right at the interface. In depletion, the bands bend down. In inversion, the conduction band closely approaches the Fermi level at the interface, indicative of the presence of an inversion layer.

Fig. 8.37 shows the electron (left) and hole (right) profiles across the semiconductor for the same voltage values. Deep into the semiconductor, p and n attain constant values that are unaffected by the applied voltage. These are the bulk values, p_{oB} and n_{oB} respectively, imposed by the doping level. In accumulation, for negative V, the electron concentration dips at the interface below the bulk value, but the hole concentration increases substantially there. As V is made more positive, n increases at the interface but also deeper into the semiconductor and p decreases accordingly. At $V = 0.5\ V$, the electron concentration at the interface equals the hole concentration in the bulk. This is threshold. For values of V larger than this threshold value, the electron concentration at the interface rises quickly in inversion. Note that at all points, $np = n_i^2$.

Fig. 8.34 (left) showing the evolution of charge with voltage is drawn on a linear scale. Using this scale, it is not possible to appreciate the evolution of the electron charge below threshold. The subthreshold regime is more clearly seen when we graph $|Q_e|$ vs. V in a semilog scale as done on the left of Fig. 8.38. This figure makes clear that below threshold ($V_T \simeq 0.5\ V$ in this case), Q_e drops in a nearly exponential fashion, as we discussed in Sec. 8.6.

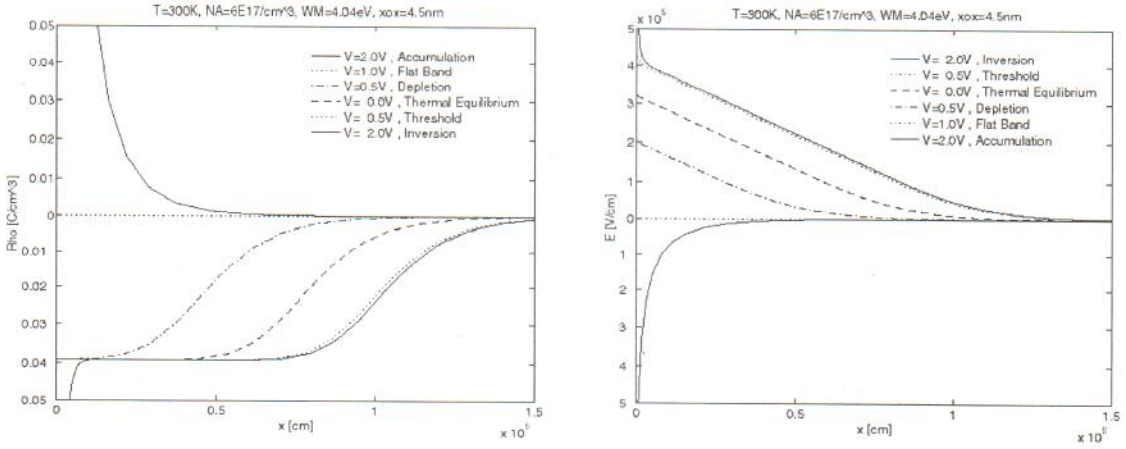


Figure 8.35: Volume charge density (left) and electric field (right) vs. space coordinate for several voltages for a typical MOS structure [*calculations courtesy of S. Mertens (MIT)*].

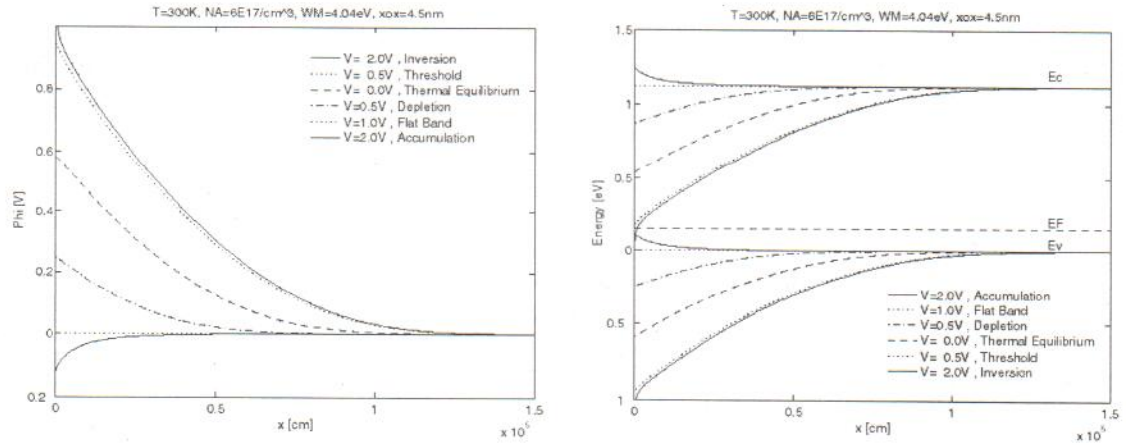


Figure 8.36: Electrostatic potential (left) and energy band diagrams (right) vs. space coordinate for several voltages for a typical MOS structure [*calculations courtesy of S. Mertens (MIT)*].

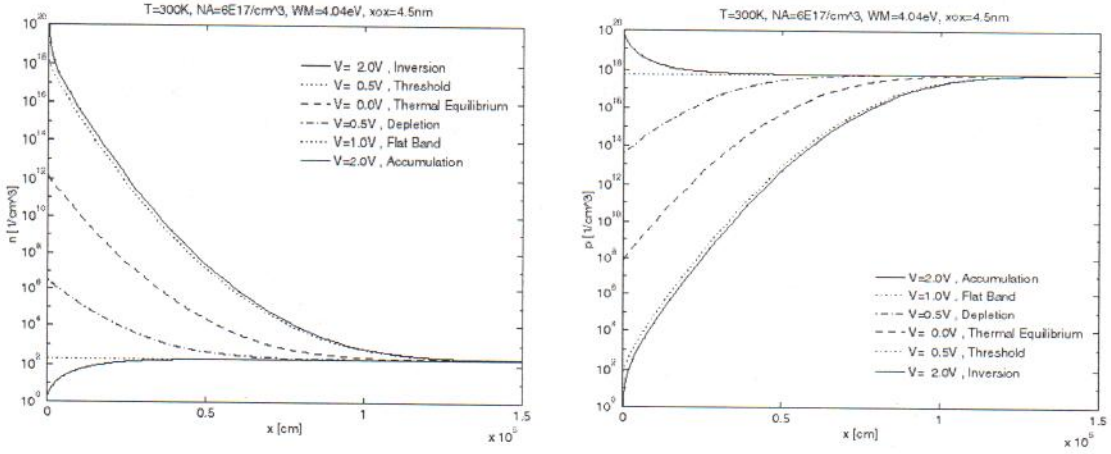


Figure 8.37: Electron and hole concentrations vs. space coordinate for several voltages for a typical MOS structure [*calculations courtesy of S. Mertens (MIT)*].

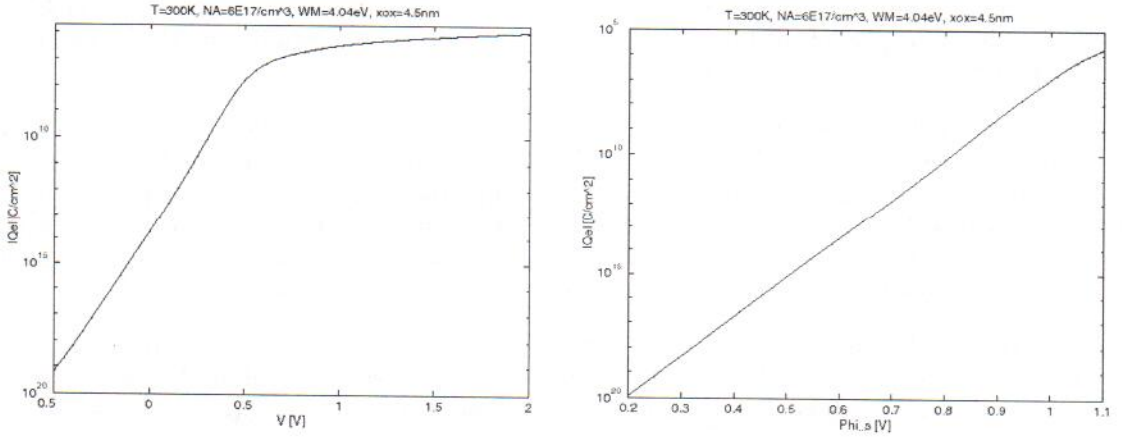


Figure 8.38: Q_e vs. V (left) and ϕ_s (right) in a semilogarithmic scale for a typical MOS structure. For small V, Q_e increases exponentially with V [*calculations courtesy of S. Mertens (MIT)*].

The evolution of $|Q_e|$ with ϕ_s in the subthreshold regime is shown on the right of Fig. 8.38. Below threshold, we observe a $|Q_e| \sim \exp \frac{q\phi_s}{kT}$ dependence, in agreement with Eq. 8.59. Above threshold, the dependence of $|Q_e|$ on ϕ_s softens and turns into a $\exp \frac{q\phi_s}{2kT}$ dependence, as predicted by Eq. 8.34.

The Poisson-Boltzmann formulation also yields a rigorous model for the capacitance-voltage characteristics of the MOS structure. Looking back at the model constructed in Sec. 8.5.1 and using the expression of Q_s in Eq. 8.88, the semiconductor capacitance can be expressed as:

$$C_s = \sqrt{2\epsilon_s kTN_A}\, \frac{dF(\phi_s)}{d\phi_s} \tag{8.94}$$

This is always positive, as a capacitance should be.

Using Eq. 8.83, we can calculate the derivative of the $F(\phi)$ function:

$$\frac{dF(\phi)}{d\phi} = \frac{q}{kT}\frac{1}{2F(\phi)}[(-\exp\frac{-q\phi}{kT}+1)+\frac{n_i^2}{N_A^2}(\exp\frac{q\phi}{kT}-1)] \tag{8.95}$$

Inserting this into Eq. 8.94, yields:

$$C_s = \frac{\epsilon_s}{\sqrt{2}L_D}\frac{1}{F(\phi_s)}[(-\exp\frac{-q\phi_s}{kT}+1)+\frac{n_i^2}{N_A^2}(\exp\frac{q\phi_s}{kT}-1)] \tag{8.96}$$

where L_D is the extrinsic Debye length.

An exact calculation of C_s *vs.* V using Eq. 8.96 for the same MOS structure as in Figs. 8.33-8.37 is shown on the left of Fig. 8.39. The overall capacitance, obtained by the series of C_s and C_{ox} as in Eq. 8.41, is shown on the right of Fig. 8.39.

These two figures display the three regimes of operation of the MOS structure that were discussed in Sec. 8.5.1. On the left of Fig. 8.39 we see that in accumulation, for $V < -1$ V, C_s increases in a roughly linear fashion as V_{GS} drops below the flatband voltage. This is what is predicted by Eq. 8.42. In depletion, C_s evolves in a rough square-root way, as given by Eq. 8.44. In inversion, C_s increases beyond threshold in an approximately linear way with a slope similar to that of the accumulation regime. This is what is predicted by Eq. 8.46.

The overall MOS capacitance is shown on the right of Fig. 8.39. In accumulation and inversion, C approaches C_{ox} the further we bias the structure below V_{FB} or above V_T, respectively. In depletion, the overall MOS capacitance evolves in a square root way consistent with Eq. 8.49.

The physical understanding gained in this section is essential to devise useful mathematical simplifications of the exact formalism. We study the three main regimes separately below.

To accomplish this, it is helpful to first examine the key dependences of the function $F(\phi)$ defined in Eq. 8.83. A sketch of $|F(\phi)|$ in a semilog scale is shown in Fig. 8.40 for different values of n_i^2/N_A^2. This graph illustrates several features of the function $F(\phi)$:

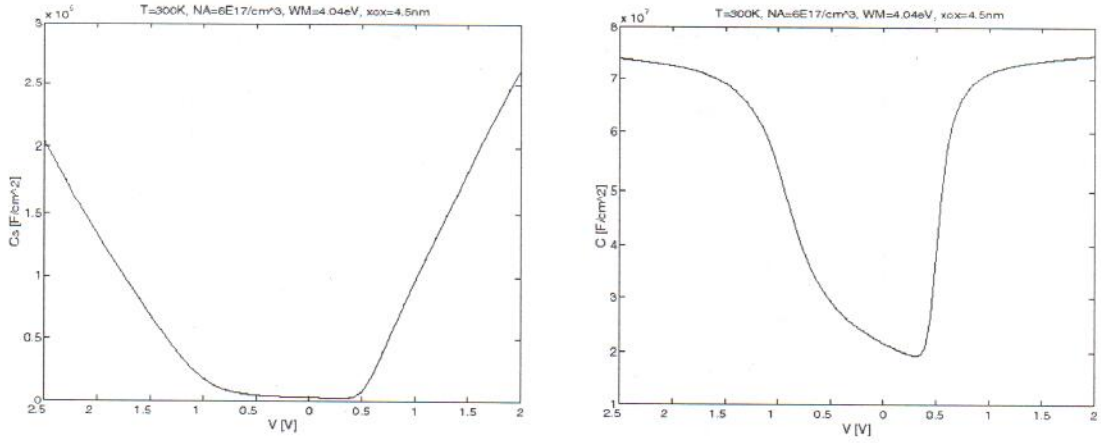


Figure 8.39: Semiconductor capacitance (left) and total capacitance of MOS structure (right) vs. applied voltage for a typical MOS structure under quasi-static conditions [*calculations courtesy of S. Mertens (MIT)*].

- For $\phi = 0$, $F(0) = 0$.
- For $\phi = 0$, $\frac{d^2F}{d\phi^2}|_0 = 0$.
- For $\phi < 0$, $F(\phi) < 0$. For $\phi > 0$, $F(\phi) > 0$.
- For all ϕ, $\frac{dF}{d\phi} > 0$.
- The F function in Eq. 8.83 contains six terms inside its square brackets. The last three are multiplied by a prefactor that represents the ratio of the minority carrier concentration to the majority carrier concentration in the bulk, a very small number (of order $\sim 10^{-10} - 10^{-14}$).
- For $\phi < 0$ and $|\phi|$ more than a few $\frac{kT}{q}$, the first term in Eq. 8.83 dominates and F displays an exponential behavior of the form:

$$F(\phi) \simeq -\exp\frac{-q\phi}{2kT} \tag{8.97}$$

- For $\phi > 0$ and ϕ more than a few $\frac{kT}{q}$, but not too large, the second term of Eq. 8.83 dominates and F displays a behavior of the form:

$$F(\phi) \simeq \sqrt{\frac{q\phi}{kT}} \tag{8.98}$$

- When $\phi > 0$ and sufficiently large, the fourth term eventually dominates and F follows a behavior:

$$F(\phi) \simeq \frac{n_i}{N_A}\exp\frac{q\phi}{2kT} \tag{8.99}$$

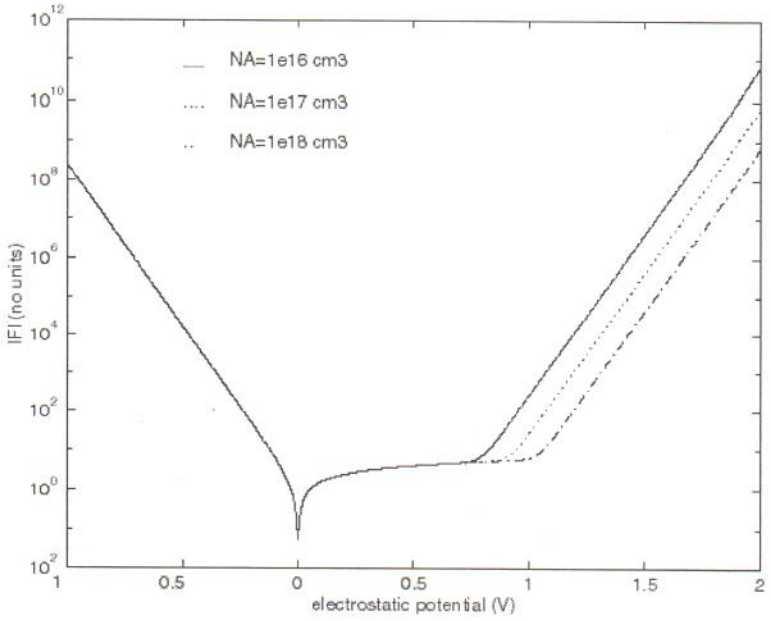


Figure 8.40: Graph of $|F|$ from Eq. 8.83 for several values of the minority to majority carrier concentration ratio, n_i^2/N_A^2, as a function of the normalized potential $q\phi/kT$.

As the prefactor n_i^2/N_A^2 becomes larger for smaller values of N_A, the onset of this last exponential behavior occurs earlier, as Fig. 8.40 shows.

We are now in a position to develop appropriate simplifications for the three key regimes in the MOS structure.

AT8.1.1 Approximations for depletion

In the depletion regime, the semiconductor charge is dominated by fixed mobile ions, the substrate dopants. In this regime, Fig. 8.33 shows that ϕ_s changes quickly with V and 8.34 shows that Q_s and $\mathcal{E}_s$ both have a square root dependence in V. It is easy to see why this should be the case.

As we saw in Eq. 8.98, in the depletion regime, the second term in $F(\phi)$ (Eq. 8.83) dominates. Inserting it into Eq. 8.85 we get:

$$\int_{\phi_s}^{\phi} \frac{d\phi}{\sqrt{\phi}} = -\sqrt{\frac{2qN_A}{\epsilon_s}}x \tag{8.100}$$

which can be easily integrated to yield:

$$\sqrt{\phi} = \sqrt{\phi_s} - \sqrt{\frac{qN_A}{2\epsilon_s}}x \tag{8.101}$$

This equation is only valid up to the point at which ϕ goes to zero, the potential in the bulk. This point is the end of the depletion region, x_d. Solving from 8.101, we find:

$$x_d = \sqrt{\frac{2\epsilon_s\phi_s}{qN_A}} \tag{8.102}$$

This is identical to Eq. 8.13. Inserting this expression back in Eq. 8.101, we get:

$$\phi = \phi_s(1 - \frac{x}{x_d})^2 \qquad \text{for } x \leq x_d \tag{8.103}$$

which is the expected result.

An expression for ϕ_s as a function of V can be obtained from Eq. 8.93. Inserting Eq. 8.98 into it, we get:

$$V = V_{FB} + \phi_s + \gamma\sqrt{\phi_s} \tag{8.104}$$

The solution of this quadratic equation is:

$$\phi_s = \frac{\gamma^2}{4}(\sqrt{1 + 4\frac{V - V_{FB}}{\gamma^2}} - 1)^2 \tag{8.105}$$

This is as obtained in Eq. 8.21. For small values of V in excess of V_{FB}, this equation is quadratic in V. For large values of V, it is linear. This is precisely the behavior observed in Fig. 8.33.

The rest of the results obtained in Sec. 8.4.1 follow easily.

AT8.1.2 Approximations for accumulation

For sufficiently negative bias, the depletion layer disappears and in its place a thin layer of holes is induced at the semiconductor interface. This is the accumulation regime. In this regime, Fig. 8.33 shows that ϕ_s saturates with V. Additionally, Fig. 8.34 shows that Q_h and $\mathcal{E}_s$ increase in magnitude linearly with V. Let us understand where these dependencies come from.

In accumulation, the first three terms of the $F(\phi)$ function in Eq. 8.83 are relevant. Plugging this into Eq. 8.88 gives us the hole accumulation sheet charge density as a function of ϕ_s:

$$Q_a = Q_s \simeq \sqrt{2\epsilon_s kTN_A}\sqrt{\exp\frac{-q\phi_s}{kT} + \frac{q\phi_s}{kT} - 1} \tag{8.106}$$

In strong accumulation, Q_a increases exponentially as $\exp \frac{-q\phi_s}{2kT}$ (remember that in accumulation ϕ_s is negative).

The relationship between the applied voltage and the surface potential can be obtained from Eq. 8.93:

$$V = V_{FB} + \phi_s - \frac{\sqrt{2\epsilon_s kTN_A}}{C_{ox}} \sqrt{\exp \frac{-q\phi_s}{kT} + \frac{q\phi_s}{kT} - 1} \tag{8.107}$$

This equation is not ameanable to an explicit solution since there are several terms that depend on ϕ_s. Of them, however, the exponential term is the dominant one. We can then obtain a first-order solution by setting the linear terms to zero since $\phi_s = 0$ marks the onset of the accumulation regime. After performing this substitution, we can solve for ϕ_s in the exponential term of Eq. 8.107 to get:

$$\phi_s \simeq -\frac{kT}{q} \ln\{1 + [\frac{C_{ox}(V_{FB} - V)}{\sqrt{2\epsilon_s kTN_A}}]^2\} \tag{8.108}$$

This equation exhibits the right dependence. For $V = V_{FB}$, $\phi_s = 0$. As V increases in absolute terms below flatband, ϕ_s increases in magnitude, although rather weakly.

An approximate expression for Q_a vs V can be obtained from the fundamental charge control relationship of Eq. 8.18 using Eq. 8.108 for ϕ_s. This yields:

$$Q_h \simeq C_{ox}(V_{FB} - V - \frac{kT}{q} \ln\{1 + [\frac{C_{ox}(V_{FB} - V)}{\sqrt{2\epsilon_s kTN_A}}]^2\}) \tag{8.109}$$

In strong accumulation, when $|V| \ll |V_{FB}$, the logarithmic term becomes negligible next to the linear term and we have the classic result:

$$Q_h \simeq C_{ox}(V_{FB} - V) \tag{8.110}$$

Eq. 8.110 explains the observation made in Fig. 8.34 that the accumulation charge in strong accumulation depends linearly in V. In fact, the slope of the dependence is precisely C_{ox} the capacitance per unit area of the oxide. Eq. 8.110 in only valid for $V < V_{FB}$. Only the voltage in excess of the flat-band voltage is used to grow the accumulation layer.

We can get a sense of the approximate distribution for the electrostatic potential in space and the thickness of the accumulation layer in strong accumulation by using the dominant term of $F(\phi)$ (the exponential term) and inserting this into Eq. 8.85. This yields:

$$\int_{\phi_s}^{\phi} \exp \frac{q\phi}{2kT} d\phi = \sqrt{\frac{2kTN_A}{\epsilon_s}} x \tag{8.111}$$

Integration of this equation leads to:

$$x = \sqrt{2}L_D(\exp\frac{q\phi}{2kT} - \exp\frac{q\phi_s}{2kT}) \tag{8.112}$$

This allows us to solve for the thickness of the accumulation layer which corresponds to $\phi = 0$:

$$x_{acc} \simeq \sqrt{2}L_D(1 - \exp\frac{q\phi_s}{2kT}) \simeq \sqrt{2}L_D \tag{8.113}$$

The simplification that has been made in Eq. 8.113 is possible because in accumulation ϕ_s is negative. Eq. 8.113 tells us that the accumulation layer is very thin, of the order of the Debye length. Also, interestingly, the thickness of the accumulation layer only depends on the doping level of the semiconductor and temperature, and is independent of other parameterers such as oxide thickness, or the magnitude of ϕ_s. For $N_A = 6 \times 10^{17}\ cm^{-3}$, for example, $x_{acc} \simeq 7.5\ nm$.

Exercise 8.8: *Consider an MOS structure such as the one of Exercise 8.3. Calculate Q_s and ϕ_s for $V = -2\ V$.*

In Exercise 8.3 we obtained for this structure $\phi_{bi} = 1.0\ V$. V_{FB} is then $-1\ V$. At $V = -2\ V$, the structure is then in accumulation. Hence $Q_s = Q_h$ which in a first pass can be computed in a straightforward way using Eq. 8.24:

$$Q_s = Q_h = C_{ox}(V_{FB} - V) \simeq 7.7 \times 10^{-7}\ F/cm^2(-1+2)\ V = 7.7 \times 10^{-7}\ C/cm^2$$

ϕ_s can be obtained from Eq. 8.108:

$$\begin{aligned}
\phi_s &\simeq -\frac{kT}{q}\ln\{1 + [\frac{C_{ox}(V_{FB} - V)}{\sqrt{2\epsilon_s kTN_A}}]^2\} \\
&= -0.026\ V \ln\{1 + [\frac{7.7 \times 10^{-7}\ F/cm^2(-1+2)\,V}{\sqrt{2 \times 1.04 \times 10^{-12}\ F/cm \times 0.026\ eV \times 1.6 \times 10^{-19}\ C \times 6 \times 10^{17}\ cm^{-3}}}]^2\} \\
&= -0.12\ V
\end{aligned}$$

As expected, this is much smaller than V.

With this estimate for ϕ_s we can now reevaluate Q_h more precisely using Eq. 8.109:

$$Q_s = Q_h \simeq C_{ox}(V_{FB} - V + \phi_s) = 7.7 \times 10^{-7}\ F/cm^2(-1+2-0.12)\ V = 6.8 \times 10^{-7}\ C/cm^2$$

This represents about a 10% correction over the result obtained above.

AT8.1.3 Approximations for inversion

In the inversion regime, a thin sheet of electrons appears at the semiconductor surface. In this regime, as Fig. 8.33 shows, ϕ_s tends to saturate with V. Additionally, Fig. 8.34 shows that Q_h saturates while Q_e and $\mathcal{E}_s$ increase linearly with V. Furthermore, the extent of the depletion region into the substrate stops growing with V. Let us understand more rigorously where these dependencies come from.

The leading term in the F function in inversion is indicated in Eq. 8.99. However, to smoothly connect with the depletion regime, we have to keep two additional terms, as in:

$$F(\phi_s) \simeq \sqrt{\frac{q\phi_s}{kT} - 1 + \frac{n_i^2}{N_A^2} \exp \frac{q\phi_s}{kT}} = \sqrt{\frac{q\phi_s}{kT} - 1 + \exp \frac{q(\phi_s - 2\phi_f)}{kT}} \tag{8.114}$$

where we have used the definition of ϕ_f in Eq. 8.25. This function smoothly connects with the F function in depletion when $\phi_s = 2\phi_f$.

The charge in the semiconductor in inversion is obtained from Eq. 8.88:

$$Q_s \simeq -\sqrt{2\epsilon_s kTN_A}\sqrt{\frac{q\phi_s}{kT} - 1 + \exp \frac{q(\phi_s - 2\phi_f)}{kT}} \tag{8.115}$$

This charge is the sum of the inversion charge plus the depletion charge. To separate the two terms, we need to consider the spatial distribution of ϕ.

In strong inversion, ϕ increases from zero in the bulk to a $\phi_s > 2\phi_f$ at the surface. While ϕ is not too large, the second term of $F(\phi)$ in Eq. 8.83 dominates. Close to the surface, ϕ becomes large enough that the fourth term of $F(\phi)$ becomes dominant. A suitable way to deal with this transition is through a piecewise approximation for $F(\phi)$. To the first order, we can select $2\phi_f$ as the value of ϕ at the boundary between these two regions. Hence, in the inversion regime:

$$F(\phi) \simeq \sqrt{\frac{q\phi}{kT}} \qquad \text{for } \phi < 2\phi_f \tag{8.116}$$

$$F(\phi) \simeq \sqrt{\frac{q\phi}{kT} - 1 + \exp \frac{q(\phi - 2\phi_f)}{kT}} \qquad \text{for } \phi > 2\phi_f \tag{8.117}$$

Eq. 8.117 applies to the electron inversion layer, while Eq. 8.116 applies to the depletion layer underneath. This piecewise description of $F(\phi)$ is reasonably simple and continuous and represents a good basis for obtaining a number of useful first-order results.

The electron charge in the inversion layer can be obtained by inserting Eq. 8.117 into Eq. 8.91 and integrating through the inversion layer:

$$Q_i = Q_e \simeq q\frac{n_i^2}{N_A}\sqrt{\frac{\epsilon_s}{2kTN_A}} \int_{\phi_s}^{2\phi_f} \frac{\exp \frac{q\phi}{kT}}{\sqrt{\frac{q2\phi_f}{kT} - 1 + \exp \frac{q(\phi - 2\phi_f)}{kT}}} d\phi \tag{8.118}$$

In the linear term inside the square root in the denominator, ϕ has been substituted by a constant value of $2\phi_f$ which is a reasonable approximation when compared with the exponential term on ϕ next to it.

It is not difficult to perform this integral and obtain:

$$Q_i \simeq -\sqrt{2\epsilon_s kTN_A}[\sqrt{\frac{q2\phi_f}{kT} + \exp\frac{q(\phi_s - 2\phi_f)}{kT} - 1} - \sqrt{\frac{q2\phi_f}{kT}}] \tag{8.119}$$

This result shows that in strong inversion, the inversion layer charge grows *exponentially* with the magnitude of ϕ_s in excess of $2\phi_f$. Also, the leading dependence is $\sim \exp(q\phi_s/2kT)$.

The depletion charge can be computed from Eq. 8.92 in a similar way. Alternatively, it can also be derived from the difference between Q_s in Eq. 8.115 and Q_i in Eq. 8.119, to obtain:

$$Q_d \simeq -\sqrt{2\epsilon_s qN_A 2\phi_f} = -C_{ox}\gamma\sqrt{2\phi_f} \tag{8.120}$$

This is a constant value, independent of ϕ_s or V, just as observed in Fig. 8.34.

The relationship between ϕ_s and V can be obtained by inserting Eq. 8.115 into Eq. 8.93:

$$V = V_{FB} + \phi_s + \frac{\sqrt{2\epsilon_s kTN_A}}{C_{ox}}\sqrt{\frac{q\phi_s}{kT} - 1 + \exp\frac{q(\phi_s - 2\phi_f)}{kT}} \tag{8.121}$$

The ϕ_s dependence on the right-hand side is complex and an explicit solution is not possible. However, the exponential term exhibits a much stronger dependence on ϕ_s than the other two terms. We can therefore obtain an approximate solution by equating $\phi_s \simeq 2\phi_f$ in the two other terms and solving for ϕ_s in the exponential. We obtain:

$$\phi_s \simeq 2\phi_f + \frac{kT}{q}\ln\{[\frac{C_{ox}(V - V_T)}{\sqrt{2\epsilon_s kTN_A}} - \sqrt{\frac{q2\phi_f}{kT}}]^2 - \frac{q2\phi_f}{kT} + 1\} \tag{8.122}$$

This equation exhibits the right limit. For $V = V_T$, with V_T as defined in Eq. 8.28, $\phi_s = 2\phi_f$. Also, as V increases above V_T, ϕ_s increases too but it does so in a logarithmic way. This is consistent with what we observed in Fig. 8.33.

In the inversion regime, the dependence of the inversion layer charge on voltage is of great importance. This can be obtained from the fundamental charge control relationship in Eq. 8.18. Again, noting that $Q_s = Q_i + Q_d$ and using the expressions for Q_d in Eq. 8.120 and for V_T in Eq. 8.28, we can write:

$$\begin{aligned} Q_i &\simeq -C_{ox}(V - V_T - \phi_s + \phi_{sT}) \\ &= -C_{ox}(V - V_T - \frac{kT}{q}\ln\{[\frac{C_{ox}(V - V_T)}{\sqrt{2\epsilon_s kTN_A}} - \sqrt{\frac{q2\phi_f}{kT}}]^2 - \frac{q2\phi_f}{kT} + 1\}) \end{aligned} \tag{8.123}$$

where we have assumed $\phi_{sT} = 2\phi_f$.

In strong inversion, the logarithmic term becomes less relevant and:

$$Q_i \simeq -C_{ox}(V - V_T) \tag{8.124}$$

This is the charge control relation for the inversion layer that was derived in Eq. 8.31. In the main body of this chapter, this equation was derived under two assumptions: that the inversion layer thickness is much thinner than any other vertical dimension in the problem (the *sheet-charge approximation*) and that beyond threshold, the depth of the depletion region stops increasing. Both assumptions have been mathematically justified in this section and their limits of applicability can be gauged from the more accurate formulation that is presented here.

The choice of threshold surface potential $\phi_{sT} \simeq 2\phi_f$ is a reasonable one and is widely used in the literature. You can see, however, that in the more exact charge control relationship of Eq. 8.123 there is an additional term with a negative sign that is increasing logarithmically with $V - V_T$. If we want to preserve the simple relationship of Eq. 8.124, then it would be advisable to define V_T at a slightly higher surface potential that $2\phi_f$. Often times, ϕ_{sT} is defined as:

$$\phi_{sT} \simeq 2\phi_f + m\frac{kT}{q} \tag{8.125}$$

with values for m between 1 and 6 often used.

The potential distribution in space can be obtained from integrating Eq. 8.85. For the inversion layer, we have to use Eq. 8.117. This is a rather messy integral. An acceptable description of the situation is obtained for ϕ_s sufficiently larger than $2\phi_f$, by just keeping the exponential term in Eq. 8.117. In this case, Eq. 8.85 becomes:

$$\int_{\phi_s}^{\phi} \exp\frac{-q(\phi - 2\phi_f)}{2kT} d\phi = -\sqrt{\frac{2kTN_A}{\epsilon_s}}x \qquad \text{with } \phi > 2\phi_f \tag{8.126}$$

Integrating, we obtain:

$$\exp\frac{-q(\phi - 2\phi_f)}{2kT} - \exp\frac{-q(\phi_s - 2\phi_f)}{2kT} = \frac{x}{\sqrt{2}L_D} \qquad \text{with } \phi > 2\phi_f \tag{8.127}$$

where we have used the definition of Debye length made in Eq. 4.68.

This equation allows us to estimate the inversion layer thickness. To be consistent with our piecewise model above, this is defined as the location at which $\phi = 2\phi_f$. If ϕ_s is larger than $2\phi_f$ by a few kT/q's at least, then Eq. 8.127 gives:

$$x_{inv} \simeq \sqrt{2}L_D[1 - \exp\frac{-q(\phi_s - 2\phi_f)}{2kT}] \simeq \sqrt{2}L_D \tag{8.128}$$

The last approximation applies for the case of strong inversion.

This is a simple and easy to remember result. The inversion layer thickness is of the order of the Debye length. For $N_A = 6 \times 10^{17}\ cm^{-3}$, for example, x_{inv} is about 7.5 nm.

Exercise 8.9: *Consider a MOS structure such as the one of Exercises 8.3. Estimate Q_s and ϕ_s for $V = 2\ V$.*

Since for this structure $V_T = 0.5\ V$, at $V = 2\ V$, it operates in inversion. Hence $Q_s = Q_i + Q_d$. Q_i, the inversion charge, can be evaluated using Eq. 8.124:

$$Q_i \simeq -C_{ox}(V - V_T) = -7.7 \times 10^{-7}\ F/cm^2(2 - 0.5)\ V = -1.2 \times 10^{-6}\ C/cm^2$$

Q_d is the depletion charge. It is obtained from Eq. 8.120:

$$Q_d \simeq -C_{ox}\gamma\sqrt{2\phi_f} = -7.7 \times 10^{-7}\ F/cm^2 \times 0.58\ V^{1/2}\sqrt{0.93} = -4.3 \times 10^{-7}\ Q/cm^2$$

Q_s is the sum of these two components:

$$Q_s = Q_i + Q_d = -1.2 \times 10^{-6}\ C/cm^2 - 4.3 \times 10^{-7}\ Q/cm^2 = -1.6 \times 10^{-6}\ Q/cm^2$$

ϕ_s is evaluated using Eq. 8.122:

$$\begin{aligned}\phi_s &\simeq 2\phi_f + \frac{kT}{q}\ln\{[\frac{C_{ox}(V - V_T)}{\sqrt{2\epsilon_s kTN_A}} - \sqrt{\frac{q2\phi_f}{kT}}]^2 - \frac{q2\phi_f}{kT} - 1\} \\ &= 0.93\ V + 0.026\ V\ln\{[\frac{1.2 \times 10^{-6}\ C/cm^2}{7.2 \times 10^{-8}\ C/cm^2} - \sqrt{\frac{0.93\ V}{0.026\ V}}]^2 - \frac{0.93\ V}{0.026\ V} - 1\} = 1.04\ V\end{aligned}$$

This is about 4 kT's over $2\phi_f$. In this calculation, we have used our evaluation of the square root in the denominator that was performed in Exercise 8.8.

Using this value of ϕ_s we can estimate Q_i more precisely using Eq. 8.123:

$$Q_i \simeq -C_{ox}(V - V_T - \phi_s + \phi_{sT}) = 7.7 \times 10^{-7}\ F/cm^2(2 - 0.5 - 0.11)\ V = -1.1 \times 10^{-6}\ C/cm^2$$

This is less than 10% different from our rough estimate above.

Problems

8.1 Estimate the conduction band and valence band discontinuities at the Si/SiO_2 interface.

8.2 Select the work function of the gate material of an MOS structure that is needed to obtain $V_{FB} = 0\ V$ on an n-type substrate with $N_D = 10^{17}\ cm^{-3}$ and $x_{ox} = 10\ nm$. Sketch the energy band diagram of this structure at zero bias.

• **8.3** Consider a MOS structure that consists of a n^+-poly-Si gate, a 9 nm SiO_2 insulator, on a p-Si substrate with a doping level $N_A = 3 \times 10^{17}\ cm^{-3}$. At room temperature and for $V_G = -3$, 0, 0.3 and 3 V, compute numerical values for:

a) the surface potential;

b) the total charge per unit area in the semiconductor and its breakdown into electron charge and hole charge;

c) the electric field in the oxide;

d) the extension of the depletion region in the semiconductor;

e) the low-frequency capacitance.

f) the high-frequency capacitance.

Do not solve the problem numerically. Use the analytical approximations described in class. Treat the n^+ poly-Si gate as a metal in which the Fermi level sits at the conduction band edge irrespective of bias.

Additionally, calculate:

g) the threshold voltage,

h) the accumulation layer charge per unit area at the oxide breakdown

i) the inversion layer charge per unit area at the oxide breakdown condition, condition.

The oxide breakdown field is 4 MV/cm.

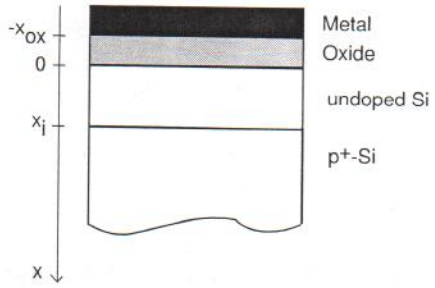
• **8.4** Consider an MOS structure using a p-Si substrate. At zero bias, this structure is in the depletion regime. Suppose we can now change gate materials and use a new one with a *higher* work function, W_M. Nothing else is changed. Indicate the impact that this would have on the parameters and figures of merit listed below.

Circle one: *increase* $=\uparrow$, *decrease* $=\downarrow$, or *no effect*. For each item, give the reason for your choice.

flatband voltage, V_{FB}:	$\uparrow$	$\downarrow$	*no effect*
surface potential at zero bias, $\phi_s(V = 0)$:	$\uparrow$	$\downarrow$	*no effect*
surface potential at threshold, ϕ_{sT}:	$\uparrow$	$\downarrow$	*no effect*
high-frequency capacitance in inversion, $C_{HF}(inv)$:	$\uparrow$	$\downarrow$	*no effect*
how does N_A have to change to maintain V_T unaffected?	$\uparrow$	$\downarrow$	*no effect*

• **8.5** Consider a metal-oxide-semiconductor structure as sketched below. In this structure, directly underneath the oxide over a depth x_i, the doping level in the semiconductor is negligible. Below x_i the semiconductor is heavily doped p-type. The work function of the metal is

smaller than the work function of the p^+ semiconductor, but the difference in work functions is smaller than the bandgap of the semiconductor.



a) Sketch the volume charge density distribution along x at zero bias.

b) Sketch the electric field distribution along x at zero bias.

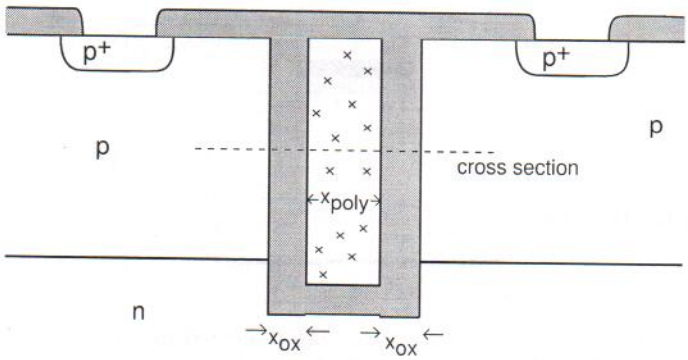
c) Sketch the electrostatic potential distribution along x at zero bias.

d) Sketch the energy band diagram along x at zero bias.

e) Derive an analytical expression for the *charge in the semiconductor* at zero bias as a function of relevant material-related parameters.

f) Derive an analytical expression for the *threshold voltage* of this MOS structure in terms of relevant material-related parameters. Define threshold as the condition that brings the conduction band edge in contact with the Fermi level at $x = 0$.

8.6 Trench isolation is a compact way to isolate devices. Typically, the trench has a lining of SiO_2 and is filled with poly-Silicon. There is no contact to the poly-Si filling. This is illustrated in the schematic below.



An important concern with trench isolation is the capacitance of the structure. This depends on the thickness of the SiO_2 lining, but also on the doping level and type of the poly-Si filling. Let's consider here the two limiting possibilities.

I. First assume that the trench is filled with p^+-polySi.

a) Sketch the energy band diagram at zero bias along the indicated cross section.

b) Calculate the low-frequency capacitance per unit area of trench that is seen between the ohmic contacts for $x_{ox} = 80\ nm$, $x_{polySi} = 100\ nm$, and $N_A = 2 \times 10^{16}\ cm^{-3}$. No DC bias is applied.

c) *Estimate* the electric field across the oxide lining.

II. Now consider the trench filled up with n^+-polySi.

d) Sketch the energy band diagram at zero bias along the indicated cross section.

e) Calculate the low-frequency capacitance per unit area of trench that is seen at the ohmic contacts for $x_{ox} = 80\ nm$, $x_{polySi} = 100\ nm$, and $N_A = 2 \times 10^{17}\ cm^{-3}$. No DC bias is applied.

f) Calculate the electric field across the oxide lining.

Regarding the poly-Si filling, assume that the Fermi level sits at E_c for the n^+ case and at E_v for the p^+ case. For all calculations, assume that no bias is applied between the two p^+ contacts.

• **8.7** Consider an MOS structure with a n^+-Si gate on a p-type Si substrate characterized by the following parameters: $W_M = \chi_s = 4.04\ eV$, $x_{ox} = 9\ nm$, $N_A = 1.2 \times 10^{17}\ cm^{-3}, T = 300\ K$.

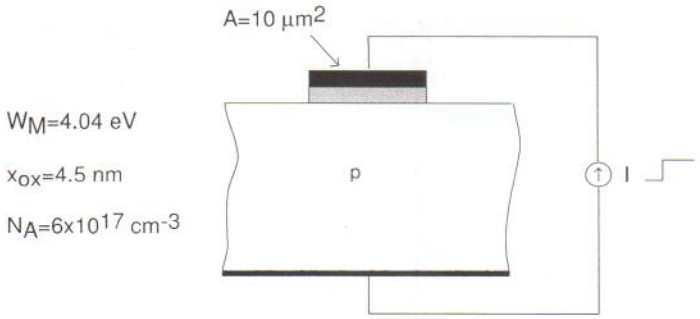
For $t < 0$, a voltage $V_1 = -2\ V$ has been applied from the gate to the bulk of the semiconductor for a long time. At $t = 0$, the applied voltage abruptly changes to $V_2 = 3\ V$.

a) Calculate the semiconductor charge and the electric field in the oxide at $t = 0^-$. What regime is the MOS structure in?

b) Calculate the semiconductor charge and the electric field in the oxide at $t = 0^+$. Consider any RC delays negligible. What regime is the MOS structure in?

c) Calculate the semiconductor charge and the electric field in the oxide for $t \to \infty$. What regime is the MOS structure in?

8.8 Consider an MOS structure as sketched in the figure below. For $t < 0$, the structure is at zero bias. In this condition, there is a depletion region in the semiconductor with a charge density of $Q_d = -3.4 \times 10^{-7}\ C/cm^2$.



At $t = 0$ a current source is switched on with the sign indicated in the figure. *Estimate* the time that it takes for the MOS structure to be brought to the onset of inversion for the three cases below. The generation lifetime in the depletion region is $\tau_g = 1\ \mu s$.

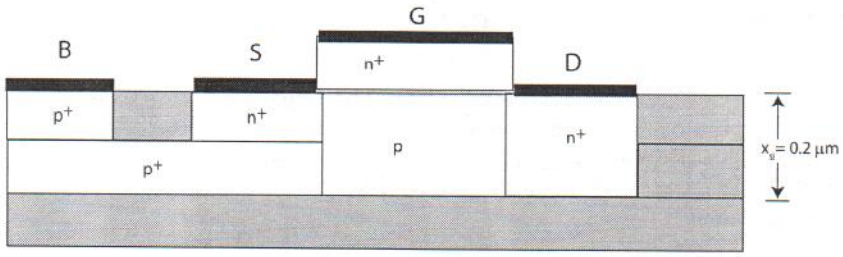
a) If $I = -100\ nA$.

b) If $I = 100\ nA$.

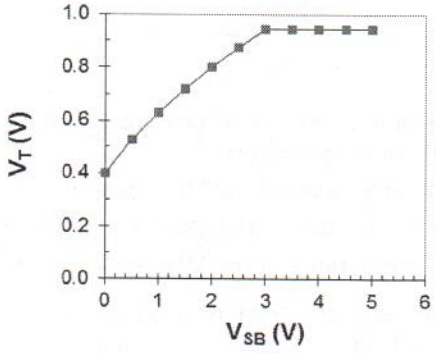
c) If $I = 1\ nA$.

8.9 One of your competitors is working on a new DRAM generation based on a body-contacted SOI (Silicon-on-insulator) MOSFET technology. In this approach, a MOSFET is fabricated

on a SOI substrate with a body contact on the side, as sketched below.



You have very little information about your competitor's design. You know that they use an n^+-polySi gate and a standard Si film thickness of $x_{si} = 0.2\ \mu m$. That is all you know. At a recent conference, your competitor presented a graph showing the dependence of the threshold voltage of the MOSFET as a function of the source to body voltage, V_{SB} (see below). Based on this new information, you can learn a bit more about their design.

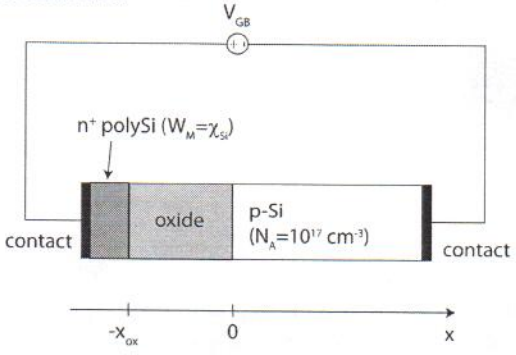


a) Explain the origin of the peculiar dependence of V_T on V_{SB}.

b) From the available information, extract the doping level, N_A, of the Si film.

c) From the available information, extract the gate oxide thickness, x_{ox}, of the MOSFET.

8.10 Consider the following MOS structure:



a) Calculate the flatband voltage.

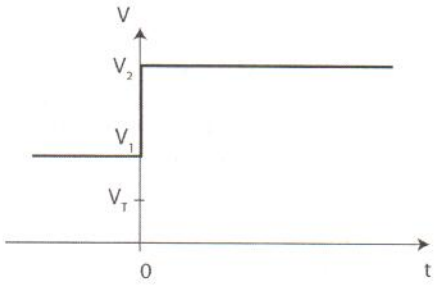
b) Calculate the extent of the depletion region in the semiconductor at threshold.

c) Calculate the electric field in the oxide at threshold.

d) Calculate the inversion layer sheet charge when the electric field in the oxide is $\mathcal{E}_{ox} = 10^6\ V/cm$.

e) Calculate the accumulation layer sheet charge when the electric field in the oxide is $\mathcal{E}_{ox} = -10^6\ V/cm$.

8.11 Consider an MOS capacitor on a p-type substrate. The gate material is n^+-polySilicon ($W_M = \chi_{Si} = 4.04\ eV$), the oxide thickness is 10 nm, the doping level of the substrate is $N_A = 1 \times 10^{17}\ cm^{-3}$. This structure has a $V_T = 0.33\ V$. For $t < 0$, the device is biased in inversion with $V_1 = 2\ V$. At $t = 0$, V is suddenly increased to $V_2 = 5\ V$ and remains at this value for a long time, as sketched below.



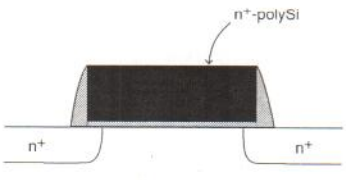
a) Sketch the evolution of the high-frequency capacitance of the MOS structure as a function of time. Explain the physics of the transitory.

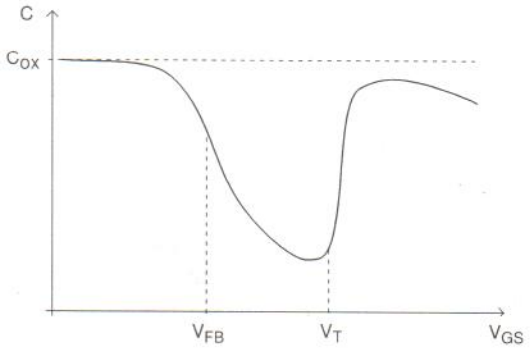
b) Calculate the high-frequency capacitance of the MOS structure at $t = 0^-$.

c) Calculate the high-frequency capacitance of the MOS structure at $t = 0^+$.

d) Calculate the high-frequency capacitance of the MOS structure at $t \to \infty$.

8.12 An emerging problem with deep-submicron CMOS is what is called *poly-depletion*. This problem manifests itself by the MOS capacitance in the strong-inversion regime, C_{inv}, failing to reach the oxide capacitance, C_{ox} (see sketch below for nMOS).





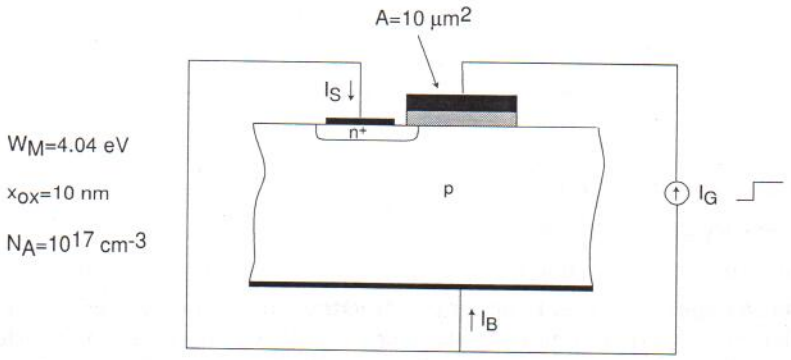
The poly-depletion problem arises at high V_{GS} as a result of a depletion region that is formed in the n^+-polySi gate due to its finite doping-level.

a) Carefully sketch an energy band diagram illustrating the gate-poly depletion problem.

b) Why doesn't this problem appear also for negatige V_{GS} when the MOS structure is in accumulation?

8.13 Consider a MOS structure with an n^+-Si gate on a p-type Si substrate characterized by the following parameters: $W_M = \chi_s = 4.04\ eV$, $x_{ox} = 10\ nm$, $N_A = 10^{17}\ cm^{-3}$. Calculate the electron and hole concentrations (in cm^{-3}) at the oxide-semiconductor interface at zero bias at $T = 300\ K$.

8.14 Consider a three-terminal MOS structure at $T = 300\ K$ as sketched in the figure below. Notice that this structure is characterized by the same physical parameters as that of the previous problem.



For $t < 0$, the structure is at zero bias. At $t = 0$, a current source of a magnitude $I_G = 10\ nA$ is switched on with the sign indicated in the figure.

Quantitatively sketch the time evolution of I_B and I_S with t for $t \geq 0$. Explain.

8.15 Consider an n^+-poly Si gate ($W_M = \chi_S = 4.04\ eV$) MOS structure with a gate oxide thickness of $x_{ox} = 15\ nm$ and a semiconductor doping level of $N_A = 10^{17}\ cm^{-3}$.

Answer the following questions at 300 K. For each question, state in what regime the MOS structure is operating. V refers to the voltage of the gate with respect to the body of the semiconductor.

To save you time, for this structure $C_{ox} = 2.3 \times 10^{-7}\ F/cm^2$ and $\gamma = 0.8\ V^{1/2}$.

a) Compute the electric field across the oxide for $V = -2\ V$.

b) Compute the hole concentration at the oxide/semiconductor interface when there is a depletion region in the semiconductor of $x_d = 50\ nm$.

c) Compute the electric field on the semiconductor side of the oxide semiconductor interface when, in steady state, there is a inversion layer charge of $|Q_i| = 5 \times 10^{-7}\ C/cm^2$.

d) Estimate the surface potential right after a pulse of $V = 4\ V$ is applied at $t = 0$ to the MOS structure in equilibrium.

e) Estimate the charge at the gate/oxide interface under the same conditions as d).

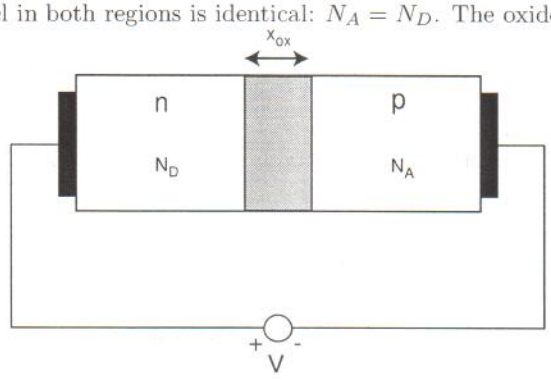
f) Estimate the high-frequency capacitance of the MOS structure under the same conditions as d).

8.16 You have learned that there is a simple technique of extracting the flat-band voltage from the high-frequency capacitance-voltage characteristics of an MOS structure. The only thing you know about this technique is that it is somehow based on graphing $1/C_{HF}^2$ vs. V.

In order to figure out how this technique might work, derive a complete analytical description of $1/C_{HF}^2$ vs. V for a MOS structure with a p-type body from accumulation to inversion. Sketch this relationship in a linear scale.

Discuss how one could extract V_{FB} from this graph.

8.17 Consider a n-Si/oxide/p-Si (Semiconductor-Oxide-Semiconductor or SOS) structure as sketched below. The doping level in both regions is identical: $N_A = N_D$. The oxide thickness is x_{ox}.



a) Sketch the charge distribution across the SOS structure at $V = 0$.

b) Sketch the energy band diagram at $V = 0$.

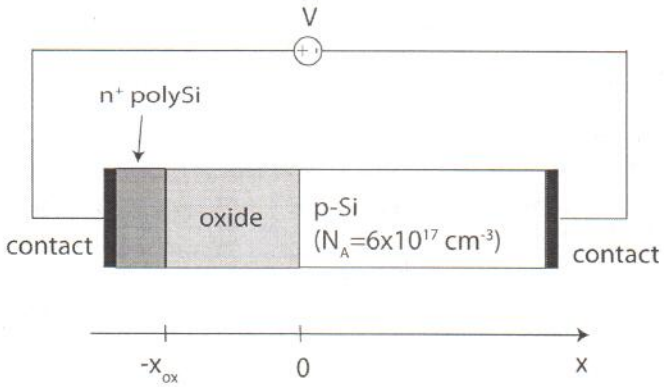
c) Derive an expression for the built-in potential of this structure at $V = 0$.

d) Now apply a negative voltage, V, across the structure. At a certain voltage, a characteristic situation occurs. How would you describe it? Derive an expression for the applied voltage required to produce this situation. Sketch the energy band diagram across the SOS structure at this bias.

e) Now apply a positive voltage across the structure. At a certain voltage, a characteristic situation occurs. How would you describe it? Derive an expression for the applied voltage required to produce this situation. Sketch the energy band diagram across the SOS structure at this bias.

f) Sketch the capacitance-voltage characteristics of this structure at low frequency. Place suitable tickmarks in the C and V axis.

8.18 Consider the following MOS structure at room temperature:



The oxide thickness is $x_{ox} = 5\ nm = 5 \times 10^{-7}\ cm$. The gate is made out of n^+-polySi with $W_M = \chi_S = 4.04\ eV$.

To save you time, for this structure $C_{ox} = 6.9 \times 10^{-7}\ F/cm^2$ and $\gamma = 0.65\ V^{1/2}$.

a) What is the magnitude of $\mathcal{E}_{ox}$ (electric field across the oxide) for a condition in which $Q_G = -2 \times 10^{-7}\ C/cm^2$?

b) What is the capacitance of the MOS structure at a bias point for which the total charge in the semiconductor is equal to $-2 \times 10^{-7}\ C/cm^2$?

8.19 Consider a two-terminal MOS structure identical to that of Exercises 8.5-8.11 in the notes. The key parameters of this structure are $N_A = 6 \times 10^{17}\ cm^{-3}$, $x_{ox} = 4.5\ nm$, and $W_M = 4.04\ eV$. For this problem, you can reuse as many results as you want from any of these exercises.

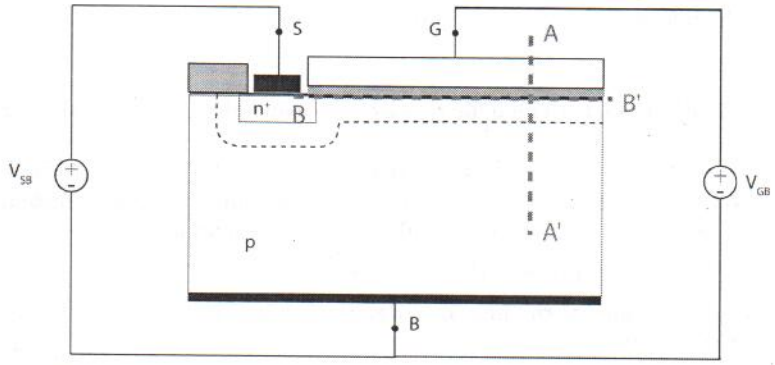
a) Estimate the surface potential at the onset of electron degeneracy, that is, at a bias point beyond which Fermi-Dirac statistics ought to be used to describe the statistics of electrons in the inversion layer. Sketch an energy band diagram for the semiconductor at this bias point.

b) At this same bias point, estimate the hole concentration at the surface of the semiconductor.

c) At this same bias point, estimate the inversion layer sheet-carrier density.

d) Estimate the gate voltage that leads to this condition.

8.20 Consider the three terminal MOS structure sketched below. The n^+ region has a doping level $N_D = 10^{19}\ cm^{-3}$. The MOS structure itself is identical to that of Exercises 8.5-8.11 in the notes. The key parameters of this structure are $N_A = 6 \times 10^{17}\ cm^{-3}$, $x_{ox} = 4.5\ nm$, and $W_M = 4.04\ eV$. For this problem, you can reuse as many results as you want from any of these exercises.



This problem is about studying the electrostatics of this structure across sections AA' and BB', indicated in the diagram, under several conditions. Section AA' cuts across the MOS structure vertically at a location far away from the n^+ region. $x = 0$ is located at the oxide-semiconductor interface. Section BB' runs along the oxide semiconductor interface (on the semiconductor side). It starts just inside the n^+ region and it runs until far away from it. $y = 0$ is the location of the n^+-p metallurgical junction.

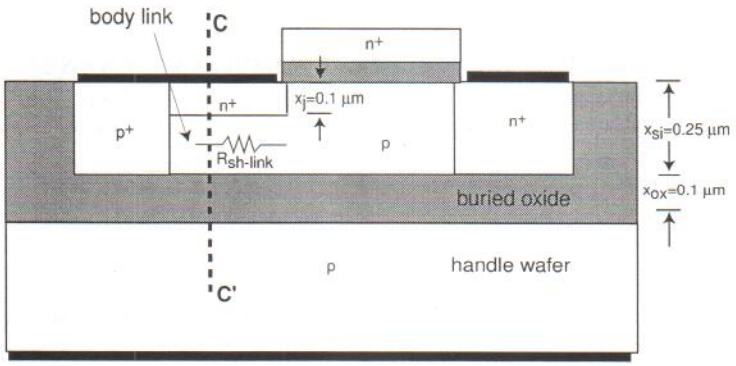
In each part that follows, *quantitatively* sketch the electrostatic potential, ϕ, along sections AA' and BB'. The sketches should include the values of the potential at key points and the spatial coordinate of the relevant point. Select $\phi = 0$ in the quasi-neutral p-type bulk of the semiconductor.

a) $V_{GB} = 0, V_{SB} = 0$.

b) $V_{GB} = 5\ V, V_{SB} = 0$.

c) $V_{GB} = 5\ V, V_{SB} = 3\ V$.

8.21 Consider an n-channel Silicon-on-Insulator MOSFET, as sketched below.



In this device design, there is a concern about the body resistance. This is because there is a relatively narrow link between the body and its body contact underneath the source. This problem is about estimating the sheet resistance of this link and the body resistance associated with it.

The doping levels in the various regions are:

-source-drain regions: $N_D^+ = 10^{19}\ cm^{-3}$
-body and body link region: $N_A = 10^{17}\ cm^{-3}$
-body contact: $N_A^+ = 10^{19}\ cm^{-3}$
-handle wafer: $N_A = 10^{17}\ cm^{-3}$

Note that the body contact, the handle wafer contact and the source are all shorted out.

a) Qualitatively sketch an energy band diagram along cross section CC' indicated in the figure. What is the regime of operation of the semiconductor-oxide-semiconductor (SOS) structure associated with the body link/buried oxide/handle wafer?

b) Estimate the sheet resistance of the p-type link.

c) Estimate the resistance of the link for a device that has a link length of 2 μm and a gate width of $W = 100\ \mu m$.

d) Suppose that the value of the link resistance that you obtain is too high. You wish to reduce it by applying a voltage to the handle wafer with respect to the body. What sign of the voltage should you apply: $V_{HB} > 0$? $V_{HB} < 0$? Explain and sketch an energy band diagram along CC' of the resulting situation.

8.22 Consider a two-terminal MOS structure identical to that of Exercises 8.5-8.11 in the notes. The key parameters of this structure are $N_A = 6 \times 10^{17}\ cm^{-3}$, $x_{ox} = 4.5\ nm$, and $W_M = 4.04\ eV$. For this problem, you can reuse as many results as you want from any of these exercises.

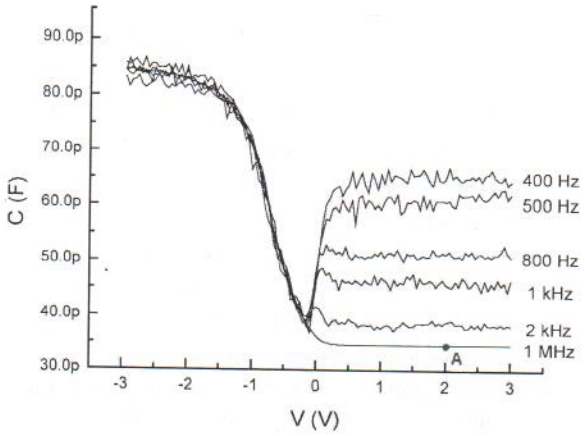
a) Estimate the surface potential at the bias that yields $n(x = 0) = p(x = 0)$ at the surface of the semiconductor. In what regime is the MOS structure operating?

b) At this same bias point, estimate the charge in the semiconductor.

c) At this same bias point, estimate the electric field across the oxide.

d) Estimate the gate voltage that leads to this condition.

8.23 The figure below graphs the experimental C-V characteristics obtained on a Si/SiO_2 MOS structure at room temperature at different frequencies. The metal is unknown. The area of the structure is $100 \times 100\ \mu m^2$.



Using the data shown in this graph and making reasonable assumptions, answer the following questions.

a) Estimate the thickness of the oxide.

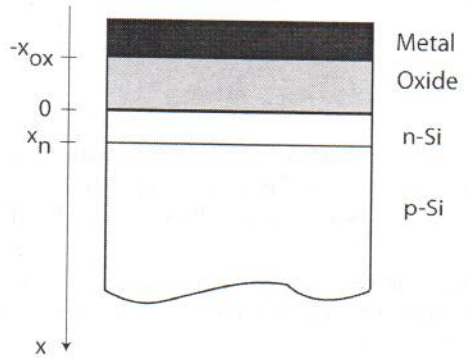
b) Estimate the extent of the depletion region in the semiconductor at threshold.

c) Estimate the doping level in the substrate.

d) Estimate the inversion layer sheet charge density when the structure is operated at point **A** in the figure.

8.24 In some applications, MOSFETs with different threshold voltages are desired. A way to accomplish this is to introduce a threshold voltage control implant. How this works is explored in this problem.

Consider a regular MOS structure on a p-type Si substrate. Assume that the structure is designed in such a way as to yield a positive threshold voltage. Now consider introducing a very thin n-type doped layer at the surface of the semiconductor, as sketched in the figure below. This is easily accomplished by implanting donor atoms under the right conditions. For simplicity, we will assume that the ion implantation results in a uniform doping distribution with a doping level of N_D and a thickness x_n.



Your assignment is to derive an expression for the shift in the threshold voltage that the introduction of the n-type layer causes. The n-type threshold control layer has been designed so that it is fully depleted at threshold. You can assume that $x_n \ll x_{dmax}$ (but not that $N_D x_n \ll N_A x_{dmax}$). Follow the steps below:

a) Qualitatively sketch the *volume charge density* in space at threshold.

b) Qualitatively sketch the *electric field* in space at threshold.

c) Qualitatively sketch the *electrostatic potential* distribution in space at threshold.

d) Derive an expression for the electric field at the surface of the semiconductor, $\mathcal{E}_s$ in terms of parameters that characterize the semiconductor, such as N_A, N_D, x_n, x_{dmax}, etc.

e) Is the surface potential at threshold, ϕ_{sT}, affected by the introduction of the n layer? If so, derive an expression for the new surface potential.

f) Is the extent of the depletion region in the p region at threshold, x_{dmax}, affected by the introduction of the n layer? If so, derive an expression for the new x_{dmax}.

g) Derive an expression for the threshold voltage of this structure. By comparing it with the threshold voltage of the conventional structure, give an expression for the shift in threshold voltage that the introduction of the n region produces.

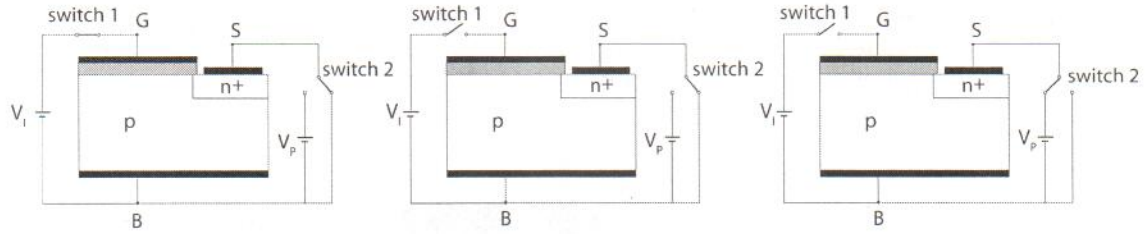
8.25 An ideal MOS structure on a Si p-type substrate is characterized by a threshold voltage $V_T = 0.6\ V$, a flat-band voltage $V_{FB} = -1\ V$, and a subthreshold swing $S = 86\ mV/dec$. All these values are given at room temperature. From this, you can reverse engineer the physical parameters of the structure. Answer the questions below. State any approximations that you need to make.

a) Estimate the doping level in the semiconductor.

b) Estimate the gate oxide thickness.

c) Estimate the work function of the gate metal.

8.26 The three-terminal MOS structure can be used to achieve what is called "parametric operation." Consider the sequence of events graphed below:



In *a)*, the MOS structure has been biased for a long time with $V_{GB} = V_I > V_T$ and the source is shorted to the body. At $t = 0$, the gate is open and left to float. This leads to situation *b)*. Some time later in *c)*, the source is applied a voltage $V_{SB} = V_P$ that is large enough so that the inversion layer is just turned off.

The reference for all voltages is the body voltage. V_T refers to the threshold voltage of the MOS structure measured from the gate to the body. Answer the questions below in terms of the usual MOS parameters: N_A, V_{FB}, V_T, x_{ox}, ϕ_{sT}, C_{ox}, etc.

a) At $t = 0^-$, provide expressions for the sheet charge density in the gate, the inversion layer and the depletion region of the MOS structure.

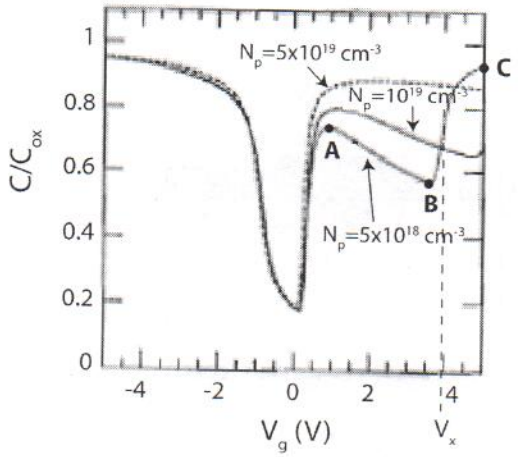
b) At $t = 0^+$, what is the voltage at the gate V_{GB}?

c) At $t = 0^+$, provide expressions for the sheet charge density in the gate, the inversion layer and the depletion region of the MOS structure.

d) Derive an expression for the voltage V_P that is required to *just* turn the inversion layer off at $t = T$.

e) Derive an expression for the voltage at the gate at $t = T$. Is this higher or lower than the gate-body voltage obtained in *b*? By how much?

8.27 The figure below shows the low-frequency C-V characteristics of an MOS structure on a p-type substrate at room temperature [5]. The gate is made out of n^+-polysilicon doped at different donor concentrations. The figure reveals prominent "poly depletion", as is known, that is more pronounced the lower the doping level in the polysilicon gate is.



In this problem, we focus on the trace that corresponds to a doping level in the polisilicon gate of $N_p = 5 \times 10^{18}\ cm^{-3}$.

a) First, consider point **A**. Carefully sketch an energy band diagram across the structure. In what regime (accumulation, depletion, or inversion) is the p-type body of the MOS structure? In what regime (accumulation, depletion, or inversion) is the n^+ polysilicon gate? Explain your answers.

b) Now consider point **B**. Carefully sketch an energy band diagram across the structure. In what regime (accumulation, depletion, or inversion) is the p-type body of the MOS structure? In what regime (accumulation, depletion, or inversion) is the n^+ polysilicon gate? Why is the overall capacitance of the structure so much lower than at point **A**? Explain your answers.

c) Derive an algebraic expression for the voltage V_x at which the capacitance abruptly rises right past point **B**. This should be in terms of the doping level in the substrate N_A, oxide thickness x_{ox}, and doping level in the polysilicon gate, N_p.

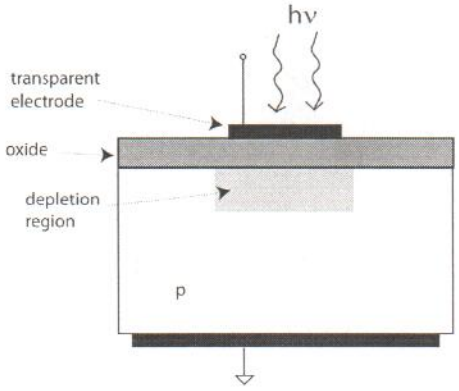
[5] from Rios and Arora, IEDM 1994

d) Now consider point **C**. Carefully sketch an energy band diagram across the structure. In what regime (accumulation, depletion, or inversion) is the p-type body of the MOS structure? In what regime (accumulation, depletion, or inversion) is the n^+ polysilicon gate? Why is the overall capacitance of the structure so much higher than at point **B**? Explain your answers.

8.28 The invention of the charge-coupled device (CCD) image sensor was honored in 2009 by the awarding of Nobel Price in physics to Willard Boyle and George Smith. They did their seminal work in the 70s at Bell Labs. CCD imagers are today widely used in all kinds of cameras and imaging systems.

The CCD imager relies on a MOS capacitor pulsed into deep depletion. If illuminated, the photogenerated electrons will collect at the oxide-semiconductor interface. The amount of electrons is proportional to the ligh intensity. There is a narrow window of time for this charge to be collected and to be sensed and that is the generation lifetime of the structure. This problem explores some of the issues involved.

Consider a pixel in a CCD imager that consists of a MOS capacitor with a transparent gate electrode, as pictured below:



The MOS structure is characterized by the following parameters: transparent metal gate with $W_M = \chi_S = 4.04\ eV$, $x_{ox} = 15\ nm$, uniform $N_A = 10^{17}\ cm^{-3}$.

To save you time, for this structure $C_{ox} = 2.3 \times 10^{-7}\ F/cm^2$, $\gamma = 0.8\ V^{1/2}$, $\phi_{sT} = 0.84\ V$, $V_{FB} = -1\ V$, $V_T = 0.6\ V$.

The active area of the pixel is $10 \times 10\ \mu m^2$.

a) Consider first this structure in the dark. For $t = 0^-$, the MOS structure is biased at $V = 0$ for a long time. At $t = 0$, a voltage pulse of $V_G = 5\ V$ is suddenly applied to the gate with respect to the body. Calculate the thickness of the depletion region at $t = 0^+$.

b) Now, holding $V_G = 5\ V$, a very short time after $t = 0$ the shutter opens and light illuminates the structure. The light intensity and wavelength are such that photogeneration takes place right at the surface of the semiconductor at a rate of $g_l = 10^{15}\ cm^{-2} \cdot s^{-1}$. How many photogenerated electrons are collected at the oxide-semiconductor interface after 1 ms of exposure? Assume that this span of time is much shorter than the generation lifetime for this structure.

c) After 1 ms of illumination, the shutter is closed while V_G remains at 5 V. What is the thickness of the depletion region right after the shutter closes?

8.29 An alternative threshold definition in a MOSFET is the condition that leads to the oxide capacitance being equal to the semiconductor capacitance, that is, $C_{ox} = C_s$.

a) Derive a simple expression for this alternative definition of the threshold voltage, V_T', in terms of the threshold voltage defined in class, V_T. State any assumptions that you need to make.

b) Derive an expression for the inversion layer charge at this new threshold voltage.