Lecture 14: MOSFETs

Contents

1 Introduction **1**

2 MOS structure **1**

3 Depletion and inversion width in a MOS **4**
3.1 Depletion width calculation 6

4 Role of oxide layer **6**
4.1 High k dielectrics 9

1 Introduction

Metal oxide semiconductor field effect transistor (MOSFET) is an example of a voltage controlled device where the current between the source and drain is determined by the channel width, which in turn is controlled by the applied voltage at the gate. MOSFETs are widely used in the semiconductor industry, especially due to the ease of fabrication of the MOS structure, and the use of silicon dioxide (SiO_2) as the insulator. The channel width in the semiconductor controls the conductivity. The important parameters are the threshold voltage required for inversion, width of the inversion channel that is created, and the total width of the depletion region. It is important to understand the relation of these parameters to the dopant concentration, so that transistors with the desired I-V characteristics can be developed.

2 MOS structure

Consider an ideal metal insulator semiconductor (MIS) structure shown in figure 1. When the insulator is an oxide layer (typically thermal oxide) then this becomes a MOS structure. It is possible to draw the band diagram for

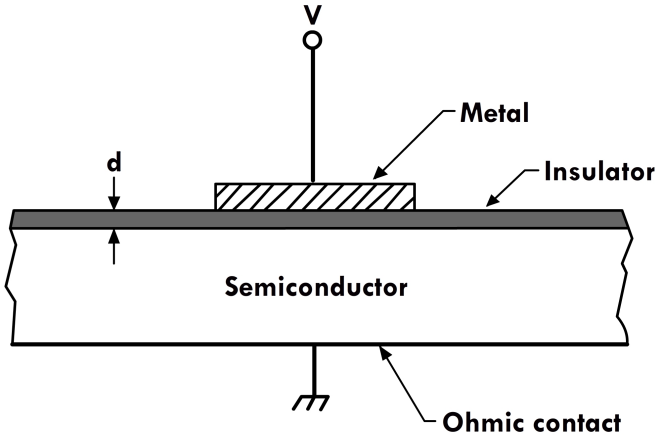

Figure 1: Metal insulator semiconductor structure. The semiconductor can be of p or n type. The insulator layer is usually SiO_x, though other insulators are possible. The gate can be made of metal or heavily doped poly silicon. Adapted from *Physics of semiconductor devices - S.M. Sze.*

this structure, depending on whether it is a n or p type structure. Both these structures are shown in figure 2. For an ideal MOS device the work functions of the metal and semiconductor are equal. So the *Fermi levels line up at equilibrium*, without any band bending or charge accumulation or depletion at the semiconductor interface. In a more general case, this band bending will have to be taken into account. The MOS structure can then be biased in different ways. This is shown in figure 3 for both p and n type semiconductors. One way to bias the structure is to create an accumulation region at the interface. For a p type semiconductor this would involve driving excess holes to the semiconductor - oxide interface, in some ways similar to an Ohmic junction.
This is not useful in a MOSFET since the source and drain are made of heavily doped n^+ semiconductors and we need to create an n channel. So the structure is usually biased, first to create a depletion region, and then an inversion region, with increasing bias. A depletion region is created when the concentration of holes (in a p type semiconductor) is smaller than N_A (the bulk acceptor concentration) but with $p > n$. In inversion, $n > p$ at the interface. The width of the n channel is usually defined from the surface to the depth at which $n = p = n_i$, i.e. the intrinsic carrier concentration.

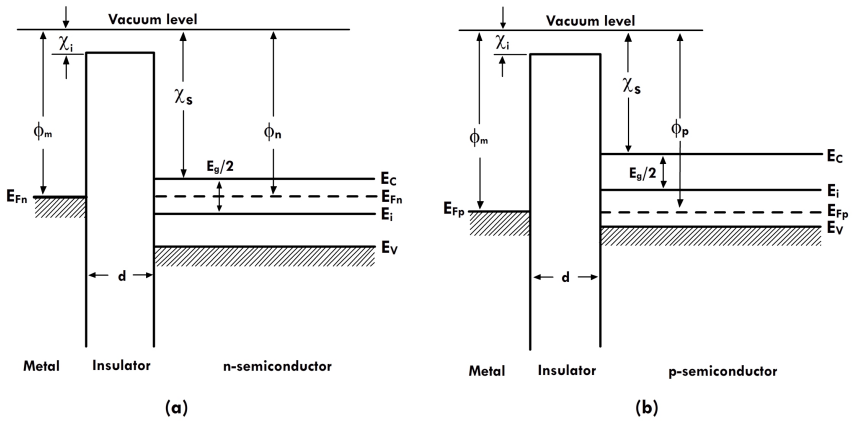

Figure 2: Metal insulator semiconductor band alignment for (a) n and (b) p type semiconductor. The diagram shows the flat band alignment in an ideal MIS where the work functions of the metal and semiconductor taken to be equal. The work function and electron affinity for the semiconductor is marked. Adapted from *Physics of semiconductor devices - S.M. Sze.*

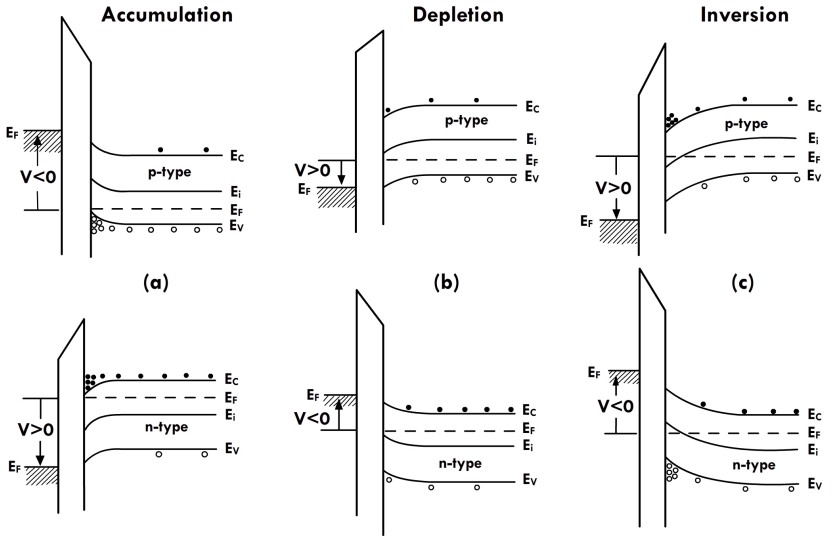

Figure 3: Metal insulator semiconductor band alignment for n and p type semiconductor under (a) accumulation, (b) depletion, and (c) inversion. The biases and hence the band bending are reversed for n and p type. Adapted from *Physics of semiconductor devices - S.M. Sze.*

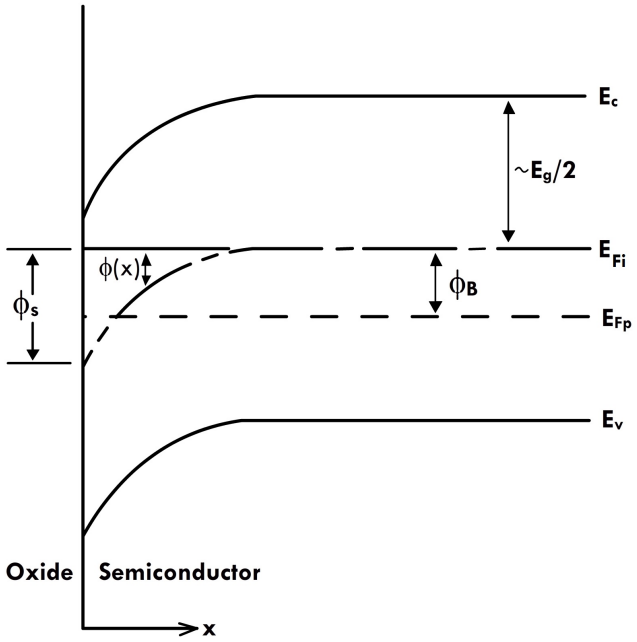

Figure 4: Oxide semiconductor band alignment for a p type semiconductor, under inversion. The bulk potential, ϕ_B, represents the acceptor concentration, while the surface potential, ϕ_s, represents the amount of band bending due to the external potential. The bulk is p type while the surface has inverted to n type. Adapted from *Solid state electronic devices - Streetman and Banerjee.*

3 Depletion and inversion width in a MOS

Consider a close up of band structure of the oxide semiconductor interface under bias, as shown in figure 4. Based on the band diagram it is possible to define a few potential terms

1. ϕ_b - bulk potential. $\phi_b = (E_{Fi})_{bulk} - (E_{Fp})_{bulk}$
2. ϕ_s - surface potential. $\phi_s = (E_{Fi})_{surface} - (E_{Fi})_{bulk}$

The bulk potential depends on the acceptor concentration in the semiconductor. The surface potential depends on the amount of band bending, based on the applied external potential. This can be understood by writing the

carrier concentration in terms of the Fermi energy positions

$$\begin{aligned} p &= n_i \exp[-\frac{(E_{Fp} - E_{Fi})}{k_B T}] \\ n &= n_i \exp[\frac{(E_{Fp} - E_{Fi})}{k_B T}] \end{aligned} \tag{1}$$

In the bulk of a p type semiconductor E_{Fp} is located below E_{Fi} and hence $p > n$. This can be rewritten in terms of ϕ_b and ϕ_s, so that electron concentration at the surface can be linked to n_i.

$$n_s = n_i \exp[\frac{(\phi_s - \phi_b)}{k_B T}] \tag{2}$$

$\phi_s - \phi_b$ is a measure of the band bending at the surface, due to the applied potential. ϕ_s also depends on the acceptor concentration (N_A) and to the depletion width (w_D) by

$$\boxed{\phi_s = \frac{e^2 N_A w_D^2}{2\epsilon_0 \epsilon_r}} \tag{3}$$

In general, it is possible to define a parabolic dependence of the potential within the depth as a function of the surface potential.

$$\boxed{\phi(x) = \phi_s (1 - \frac{x}{w_D})^2} \tag{4}$$

The surface of a MOS structure is said to be inverted when n is greater than p. When this happens, the bands bend such that the *Fermi level at the surface is closer to the conduction band than to the valence band.* This situation is depicted in figure 4. When the concentration of electrons at the surface (n_s) equals the acceptor concentration in the bulk, i.e.N_A, then it is defined as **strong inversion**. At strong inversion, *the surface is as much n type as the bulk is p type.* Then, the Fermi level at the surface is located as much above the intrinsic level as it is located below the intrinsic level in the bulk. When $n_s = N_A$

$$\boxed{\phi_s = 2\phi_B = 2k_B T \ln(\frac{N_A}{n_i})} \text{strong inversion only} \tag{5}$$

This can be substituted in equation 3 to calculate the depletion width at strong inversion (w_m).

$$\begin{aligned} 2k_B T \ln(\frac{N_A}{n_i}) &= \frac{e^2 N_A w_m^2}{2\epsilon_0 \epsilon_r} \\ \boxed{w_m = 2\sqrt{\frac{\epsilon_0 \epsilon_r k_B T}{e^2 N_A} \ln(\frac{N_A}{n_i})}} & \end{aligned} \tag{6}$$

To calculate the width of the inversion region (n channel) we can use equation 4 and calculate the depth (x_m) at which $\phi(x) = \phi_B$. This gives a n channel width related to depletion width (w_m) by

$$\boxed{x_m = w_m \left(\frac{\sqrt{2} - 1}{\sqrt{2}}\right)} \text{ strong inversion only} \quad (7)$$

3.1 Depletion width calculation

Consider a p type Si with N_A of 10^{16} cm^{-3} forming an interface with the oxide layer and metal. Since the semiconductor is Si, at room temperature n_i is 10^{10} cm^{-3} and band gap is 1.10 eV. The Fermi level position in the bulk (bulk potential, ϕ_b) is equal to 0.36 eV (below intrinsic Fermi level). To create a depletion and inversion region we need to apply an external potential so that electrons are driven to the surface or holes are driven away from the surface, see figure 3. At strong inversion, the electron concentration at the surface becomes equal to N_A. This happens when $\phi_s = 2\phi_B$ and the applied potential for this 0.72 V. Using equation 6, the depletion width at strong inversion (w_m) is 307 nm. For this depletion width, the inversion channel width (x_m), given by equation 7, is approximately 90 nm. The band bending and carrier concentration, as a function of depth, is shown in figure 5.

If the bulk acceptor concentration is increased by two orders of magnitude to 10^{18} cm^{-3}, then it becomes harder to create an inversion channel since now a greater potential has to be applied to drive the electrons to the surface. For this concentration, the depletion width (w_m) is reduced to 35.4 nm and correspondingly the inversion channel width (x_m) is also reduced to 10.4 nm. Since the channel width affects the conductivity (similar to the resistance of a thin film) increasing the bulk p concentration decreases the conductivity of the n channel and also decreases the source drain current (I_{DS}). The relation between depletion width and dopant concentration for Si and GaAs is shown in figure 6.

4 Role of oxide layer

The purity of the oxide layer at the interface between the metal and semiconductor is important. The externally applied potential is equal to the voltage drop across the oxide layer plus the voltage drop across the semiconductor. If there are defects in the oxide layer, especially charged defects, these can also migrate under the influence of the applied voltage and affect the threshold voltage for inversion. This is shown in figure 7. Defects in the oxide is

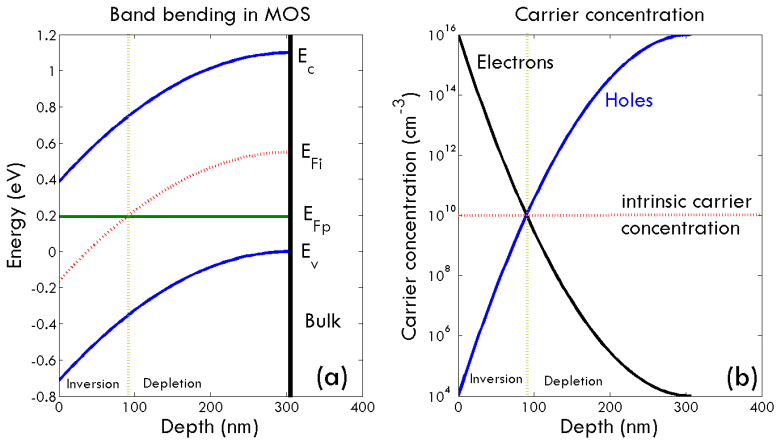

Figure 5: (a) Band bending, during strong inversion, for a p type semiconductor with N_A of 10^{16} cm^{-3}. The bending of the valence and conduction band and intrinsic Fermi levels are plotted. (b) The change in carrier concentration with depth. The surface is n type up to a depth of around 90 nm and then there is a depletion region up to 307 nm. The dotted vertical lines in both (a) and (b) mark the separation between the inversion and depletion regions. The plot was generated in MATLAB.

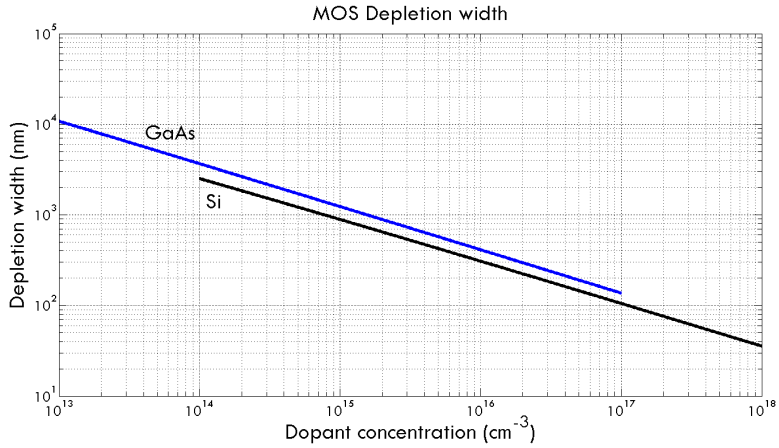

Figure 6: Depletion width, for strong inversion, vs. dopant concentration for Si and GaAs. This a logarithmic plot from equation 6. For the same dopant concentration, depletion width is slightly higher in GaAs, because the intrinsic carrier concentration is lower. The plot was generated in MATLAB.

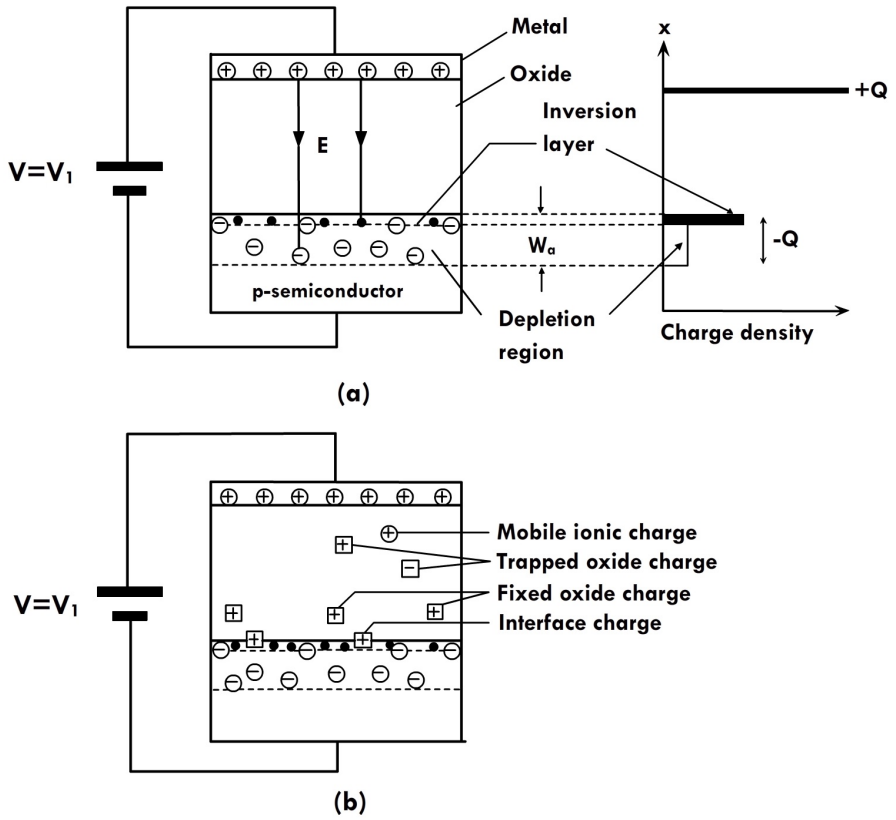

Figure 7: Defects in the oxide interface can affect threshold voltage for inversion in a MOS structure. (a) Shows an ideal oxide layer. (b) Commonly found defects in an grown oxide layer. These defects can be fixed or mobile. Adapted from *Principles of Electronic Materials - S.O. Kasap.*

related to the growth process. Thermally growth SiO_2 was originally used as the oxide layer for MOS due to the possibility to grow defect free oxides. This has been replaced now by a variety of other oxide materials.
The interface between the oxide and the metal also matters. Poly-Si gate can also be used instead of metal gate. To increase conductivity this is usually heavily doped. The advantage of poly-Si gate is that, during fabrication, the MOS structure can be subjected to high temperature processing while metal gate restricts processing temperature. Also, the threshold voltage for a poly Si gate can be controlled by doping the poly Si. This can be used to fabricate multiple MOSFETs with varying I-V characteristics in the same IC.

4.1 High k dielectrics

The oxide layer between the metal and semiconductor in a MOS acts as a capacitor. The capacitance of this layer is given by

$$C = \frac{\epsilon_0 \epsilon_r^{ox} A}{d} \tag{8}$$

where ϵ_r^{ox} is the relative permittivity of the oxide. For Si this is 3.9. d refers to the width of the oxide layer and A is the area. As device dimensions scale down, with transistor sizes shrinking in size and more of them packing in a smaller area, the thickness of the oxide layer decreases. But this can lead to leakage currents in device due to tunneling across the thin oxide interface. Thin interfaces also have higher defect densities which can affect the threshold voltage. One way to overcome this leakage current is to replace the oxide material with an insulator having a higher relative permittivity (ϵ_r^i). For the same capacitance, using equation 8, this gives an insulator thickness (d_i) by

$$d_i = d_{ox} \frac{\epsilon_r^i}{\epsilon_r^{ox}} \tag{9}$$

Another way of expressing this is that a higher capacitance can be retained by using a thicker insulator width with reduced leakage. Typical insulators that replaced the oxide layer include SiN_x ($\epsilon_r = 7$), Ta_2O_5 ($\epsilon_r = 25$), and TiO_2 ($\epsilon_r = 60 - 100$). Growing these insulators without interface defects are challenging. This adds another layer of complexity in integrated device manufacture.