

Chapter 2

The MOS Structure

2.1 Introduction

The metal-oxide-semiconductor diode or MOS capacitor is an important structure, which is incorporated in the surface of most semiconductor devices. It forms an essential part of a MOSFET which in turn is an important device used in large-scale integration. Therefore, all the studies related to any kind of MOS device needs at first the basic understanding of the MOS structure. In order to achieve this objective, the present chapter is devoted to the study of MOS structure. A simple physical approach applied to MOS structure and a behavior of ideal MOS capacitor [1–5] that are necessary for understanding the analyses that will follow subsequently, are described. At first, all the basic concepts and quantities are introduced. Then, the charge distribution that sets in a MOS structure when the latter is biased in either one of the three biasing modes (accumulation, depletion, and inversion) is analyzed. This charge distribution is used to obtain the value of the capacitance and its dependence on the magnitude and the frequency of the applied small signal using a phenomenological approach. A real MOS structure always contains so-called “oxide charges” located in the bulk of the oxide or at the oxide-silicon interface. The impact of these charges on the behavior of real MOS structure and in particular on the flat-band voltage is also examined.

2.2 A Simple Physical Approach Applied to MOS Structure

The MOS capacitor consists of an oxide film sandwiched between a P- or N-type silicon substrate and a metal plate called gate as shown in Fig. 2.1. The study of the behavior of this capacitor under a varying bias applied between substrate and gate is a powerful way to investigate the quality of the oxide layer and the quality of the oxide-silicon interface.

Fig. 2.1 Cross-sectional view of a MOS structure

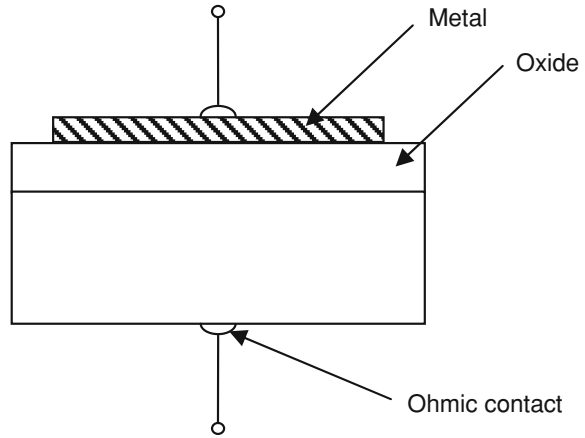
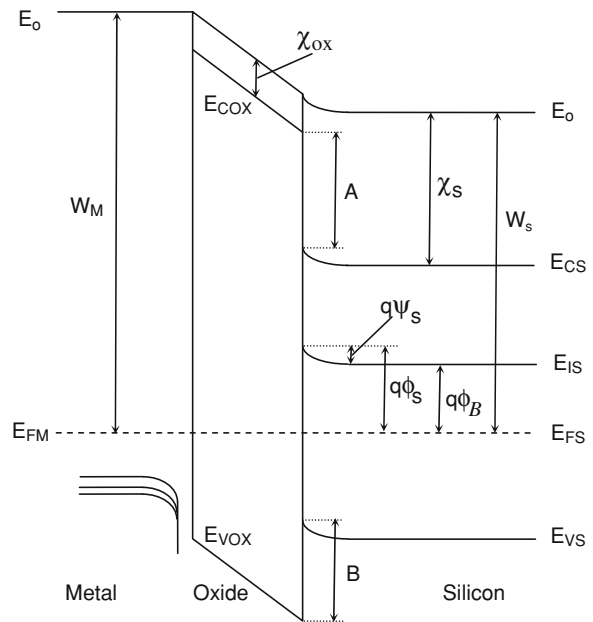


Fig. 2.2 Energy band diagram of unbiased real MOS structure ($W_M \neq W_S$)



2.2.1 Basic Concepts and Quantities

Figure 2.2 shows the energy band diagram of an unbiased MOS structure when the work function of the metal W_M and work function of silicon W_S are different. The diagram shows the position of the different energy levels such as Fermi level in the gate (E_{FM}) and in the silicon (E_{FS}). In this figure, χ_s represents the electron affinity for the silicon and χ_{ox} for the oxide. Figure 2.2 also shows that certain energy barriers exist between the metal and the oxide as well as between the silicon and

Table 2.1 Some typical values in the energy bands of a MOS structure

Metal	Oxide	Silicon
$W_M = 4.8 \text{ eV (Au)}$	$\chi_{\text{ox}} = 0.9 \text{ eV}$	$\chi_S = 4.1 \text{ eV}$
$W_M = 4.1 \text{ eV (Al)}$	$E_{\text{COX}} - E_{\text{VOX}} = 8.1 \text{ eV}$	$E_{\text{CS}} - E_{\text{VS}} = 8.1 \text{ eV}$
	$A = 3.2 \text{ eV}$	$4.1 \text{ eV} < W_S^* < 5.2 \text{ eV}$
	$B = 3.8 \text{ eV}$	

* W_S varies with doping concentration and temperature

the oxide. For example, an energy ($W_M - q\chi_{\text{ox}}$) would be needed to move an electron from the Fermi level of the metal E_{FM} to the lowest unoccupied states in the oxide, and $A + (E_{\text{CS}} - E_{\text{VS}})$ would be needed to move an electron from the silicon valence band to the lowest unoccupied states in the oxide, where W_M is the work function of metal, E_{CS} and E_{VS} the bottom of conduction band and top of valence band of silicon respectively. “A” difference between the bottom level of the conduction bands of oxide and silicon at the Si-SiO₂ interface and q the electron charge. The importance of these energy barriers is that they prevent the free flow of carriers from the metal to the silicon or vice versa. Some typical values for such a structure are shown in Table 2.1 [6–8].

2.2.2 Definition of Potentials

Figure 2.2 shows the various potentials. The potential may be defined by the following equation,

$$q\phi = E_F - E_i(x). \quad (2.1)$$

where E_F is the extrinsic Fermi level and E_i is the intrinsic energy level in the silicon. The potential $\phi(x)$ is called the *bulk potential* ϕ_B in the bulk ($x \rightarrow \infty$) and the *surface potential* ϕ_S at the surface ($x = 0$).

Location of any other energy level e.g. an interface trap level within the silicon band gap may be specified by stating its distance in electron volt from the intrinsic level. The *band bending* $\psi(x)$ is defined as:

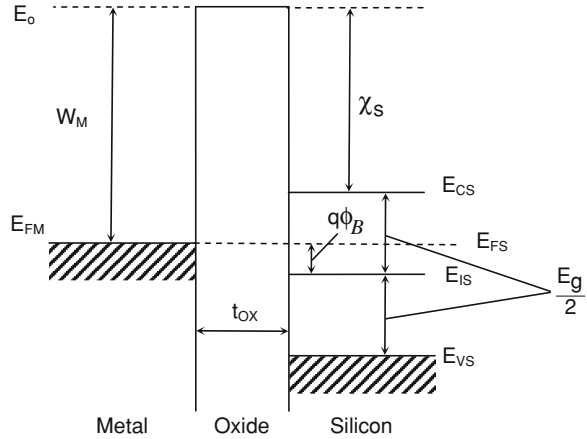
$$\psi(x) = \phi(x) - \phi_B. \quad (2.2)$$

where $\psi(x)$ represents the potential at any point x in the depletion layer with respect to its value in the bulk. In particular, the barrier height $\psi_S = \phi_S - \phi_B$ is the *total band bending*.

2.3 Ideal MOS Capacitor

Before characterizing electrically the real MOS device by taking into consideration the defects contained in the SiO₂, at first, the ideal MOS structure will be studied. The MOS structure is called ideal if the following two conditions are met:

Fig. 2.3 Energy band diagram of unbiased ideal MOS structure



1. The work function of metal W_M and work function of silicon W_S are equal, $W_M = W_S$, which implies that in the three materials, all energy levels are flat, when no voltage applied to the structure. This case is illustrated in Fig. 2.3.
2. There exists no charge in the oxide and at the Si-SiO₂ interface, which implies that the electric field is zero everywhere in the absence of any applied voltage.

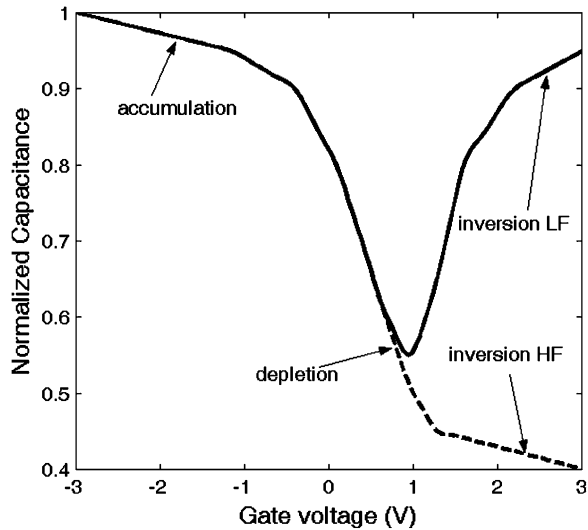
MOS capacitance will vary with the applied gate to substrate voltage. The capacitance versus voltage characteristics of MOS capacitors that result from the modulation of the width of the surface space charge layer (SCL) by the gate field have been found to be extremely useful in the evaluation of the electrical properties of oxide-silicon interfaces. There are three regions of interest, namely, accumulation, depletion and inversion in the C-V characteristics of the MOS capacitor as shown in Fig. 2.4. A MOS capacitor fabricated on a P-type substrate is the case treated here.

2.3.1 Accumulation

When an external voltage V_G is applied to the silicon surface in MOS capacitor, the carrier densities change accordingly in its surface region. With large negative bias applied to the gate, holes are attracted by the negative charges to form an accumulation layer (Fig. 2.5). The high concentration of these holes will form the second electrode of a parallel plate capacitor with first electrode at the gate. Since the accumulation layer is an indirect ohmic contact with the P-type substrate, the capacitance of the structure under accumulation conditions must be approximately equal to the capacitance of the oxide [1],

$$C_{ox} = \frac{\epsilon_o \epsilon_{ox}}{t_{ox}} \quad (2.3)$$

Fig. 2.4 Typical ideal C–V curves showing the three modes: accumulation, depletion, inversion for both high and low frequency in a P-MOS structure



where ϵ_o is the permittivity of the free space, ϵ_{ox} the relative permittivity of oxide, and t_{ox} the oxide thickness. This capacitance is always expressed per unit gate area [F cm^{-2}]. It does not vary with bias V_G as long as the structure is maintained in accumulation mode (Fig. 2.4). It is also independent of the frequency as long as the motion of the majority carriers, which contribute to substrate charge ΔQ_S , can keep pace with the incremental speed of gate charge ΔQ_M . This is true if the frequency of the applied small signal is smaller than the reciprocal of the dielectric time constant of silicon, i.e. 10^{11} Hz. Under this condition, the Fermi level near the silicon surface will move to a position closer to the valance band edge as shown in Fig. 2.5c.

2.3.2 Depletion

When negative charges are removed from the gate, holes leave the accumulation layer until the silicon will be neutral everywhere. This applied gate bias is called the flat band voltage. As the bias on the gate is made more positive with respect to flat band, holes are repelled and a region is formed at the surface which is depleted of carriers (Fig. 2.6b). Under depletion conditions, the Fermi level near the silicon surface will move to a position closer to the center of the forbidden region as illustrated in Fig. 2.6c. Increasing the positive voltage V_G will tend to increase the width of the surface depletion region X_D , the capacitance from the gate to the substrate associated with MOS structure will decrease, because the capacitance associated with the surface depletion region will add in series to the capacitance across the oxide. Thus the total capacitance per unit area from the gate to substrate under depletion conditions is given by

Fig. 2.5 Schematic representation of P-MOS structure under bias resulting in accumulation mode, **a** biasing condition, **b** charge distribution, **c** energy band diagram

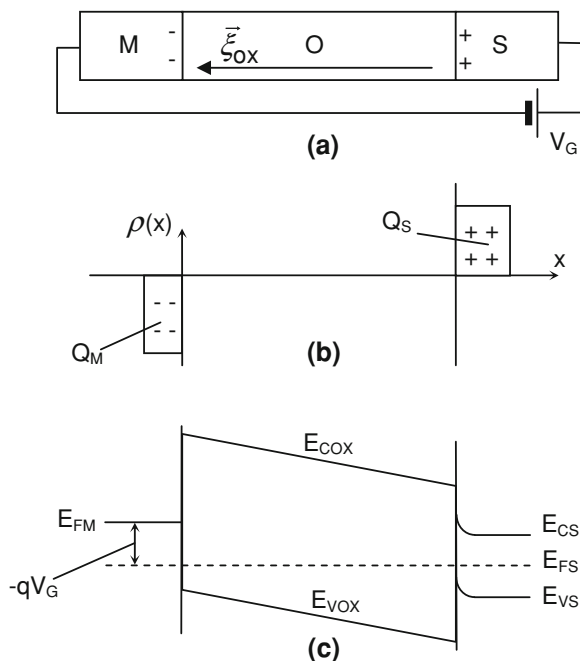
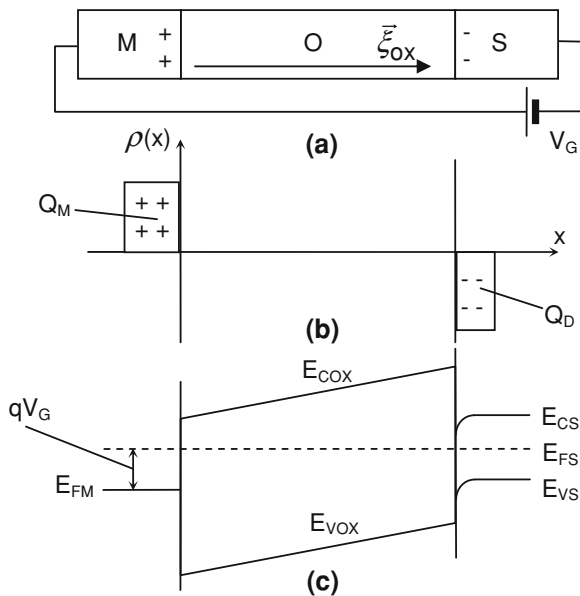


Fig. 2.6 Schematic representation of P-MOS structure under bias resulting in depletion mode, **a** biasing condition, **b** charge distribution, **c** energy band diagram



$$C(V_G) = \left(\frac{1}{C_{ox}} - \frac{1}{C_S(V_G)} \right)^{-1}, \quad (2.4)$$

where C_S is the silicon capacitance per unit area, is given by

$$C_S(V_G) = \frac{\epsilon_o \epsilon_S}{X_D}, \quad (2.5)$$

and,

$$X_D = \sqrt{\frac{2\epsilon_o \epsilon_S \psi_S}{qN_A}}. \quad (2.6)$$

Where the relation between the applied gate voltage V_G and the total band bending ψ_S can be written as

$$V_G = \psi_S + \frac{\sqrt{2\epsilon_o \epsilon_S q N_A \psi_S}}{C_{ox}} \quad (2.7)$$

Since only majority carriers contribute to the substrate charge ΔQ_D , the capacitance is independent of frequency.

2.3.3 Inversion

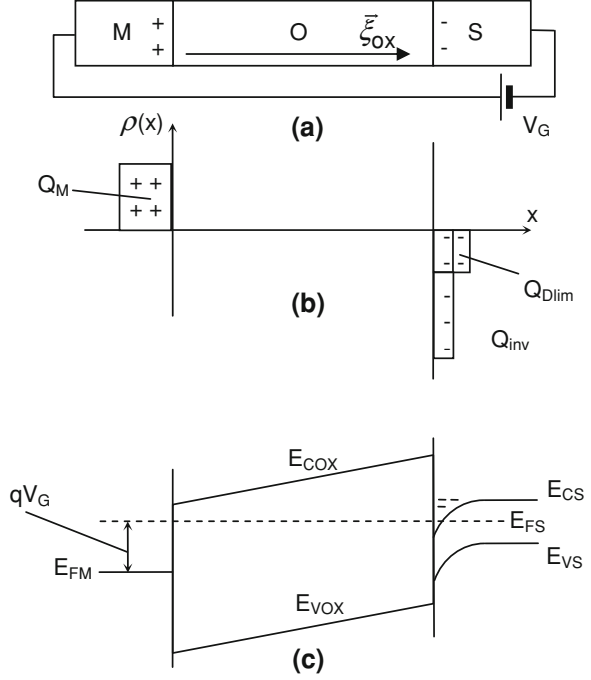
With increasingly applying positive voltage, the surface depletion region will continue to widen until the onset of surface inversion is observed (n-type), an inversion layer is formed, the Fermi level near the silicon surface will now lie close to the bottom of conduction band (Fig. 2.7). This inversion layer is very thin (1–10 nm) and separated from the bulk of silicon by the depletion layer. The build-up of inversion layer is a threshold phenomenon. The threshold condition marks the equality of the concentration of minority carriers to the doping concentration. At the onset of inversion, the depletion layer width reaches a limit, X_{DLim} as shown in Fig. 2.7b. Since the charge density in the inversion layer may or may not be able to follow the ac variation of the applied gate voltage, it follows that the capacitance under inversion conditions will be a function of frequency.

Low frequency Capacitance

This case, illustrated in Fig. 2.4, corresponds to the thermal equilibrium in which the increase in the gate charge δQ_M is balanced by the substrate charge δQ_{inv} . It arises when the frequency of the small signal is sufficiently low (typically less than 10 Hz). The low frequency capacitance of the structure, C_{LF} , is equivalent to that of the oxide layer, just as in accumulation mode,

$$C_{LF} = C_{ox}. \quad (2.8)$$

Fig. 2.7 Schematic representation of P-MOS structure under bias resulting in inversion mode, **a** biasing condition, **b** charge distribution, **c** energy band diagram



High Frequency capacitance

The case illustrated in Fig. 2.4, corresponds to the higher frequencies of the applied small signal (typically above 10^5 Hz). The increase of charge in the metal side δQ_M is now balanced by the substrate charge δQ_D , since the minority carriers can no longer adjust their concentrations. The charge modulation δQ_D occurs at distance X_{Dlim} of the Si-SiO₂ interface. It follows that the high frequency capacitance of the MOS structure, C_{HF} , is given,

$$\frac{1}{C_{HF}} = \frac{1}{C_{ox}} + \frac{1}{C_{Dlim}}. \quad (2.9)$$

where

$$C_{Dlim} = \frac{\epsilon_o \epsilon_S}{X_{Dlim}}, \quad (2.10)$$

and,

$$X_{Dlim} = \sqrt{\frac{4\epsilon_o \epsilon_S k T L n \left(\frac{N_A}{n_i} \right)}{q^2 N_A}}. \quad (2.11)$$

As shown in Fig. 2.4, the capacitance is practically independent of positive or negative bias for both high frequency inversion and low frequency inversion.

2.4 The Actual (Non-ideal) MOS Structure

An ideal MOS device does not agree with experimental results, and this difference is due to the presence of the oxide charges and the work function difference that exists in practice but was not taken into account in the theoretical treatment of an ideal MOS capacitor. Early studies of the MOS devices showed that the threshold voltage V_{Th} and the flat band voltage V_{FB} could strongly be affected by these charges. The understanding of the origin and nature of these charges is very important if they are to be controlled or minimized during device processing [2, 9]. The net result of the presence of any charge in the oxide is to induce a charge of opposite polarity in the underlying silicon. The amount of charge induced will be inversely proportional to the distance of the charge from the silicon surface. Thus, an ion residing in the oxide very near the Si-SiO₂ interface will reflect all of its charge in the silicon, while an ion near the oxide outer surface will cause little or no effect in the silicon. The charge is measured in terms of the net charge per unit area at the silicon surface. Most oxide charge evaluations can be made using the capacitance voltage (C-V) method. This method is simple and rapid [10, 11] and in most cases provide a quantitative or at least a semiquantitative measure of the surface charge.

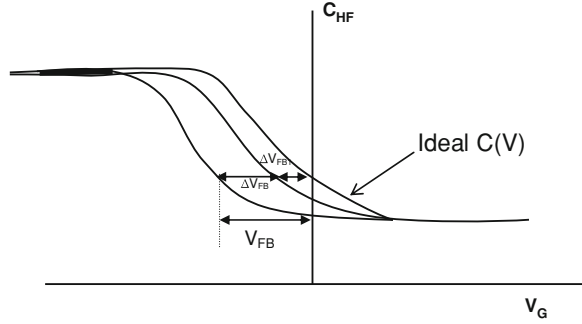
2.4.1 The Metal-Silicon Work Function Difference

In the real MOS structure, the work function of the metal and the work function of the silicon are different [6, 7]. For this reason, there exists an electric field in the oxide and in the top layer of the silicon even in the absence of an applied voltage (see band diagrams of Fig. 2.2). To obtain the flat band conditions, $\psi_s = 0$, a bias on the gate must be applied relative to the substrate, which can be written as

$$\Delta V_{FB1} = \frac{W_{MS}}{q}. \quad (2.12)$$

As an example, for Al-SiO₂-Si structure, a typical value of ΔV_{FB1} is 0.3 V for an n-type Si substrate and 0.8 V for a p-type [6]. The effect of a work function difference may cause a shift of the actual C-V curve with respect to the ideal one. The flat-band-voltage shift ΔV_{FB1} occurs along the voltage axis as illustrated in Fig. 2.8.

Fig. 2.8 Representation of the C–V curves of a P-type MOS structure showing the flat band voltage shifts introduced by work function difference and oxide charges



2.4.2 Effect of the Charge Distributed in the Oxide

Whether mobile ions or other types of oxide charges are distributed unevenly in the bulk, their density $\rho(x)$ varies with distance (and with time in case of time-dependent stress). To study the influence of oxide charges distribution on the properties of the MOS structure, at first, the effect of only those charges, which are located within a layer between x and $x + dx$, is calculated. The origin of the x -axis is taken at the metal-oxide interface as shown in Fig. 2.9. In a second step, the effect of the various layers from zero to t_{ox} is added. Using Gauss's law, the electric field in the oxide ζ_{ox} exhibits a discontinuity $\delta\zeta_{ox}$ when crossing this charge layer. This discontinuity is given by

$$\delta\zeta_{ox} = \frac{\rho(x)dx}{\epsilon_o\epsilon_{ox}}. \quad (2.13)$$

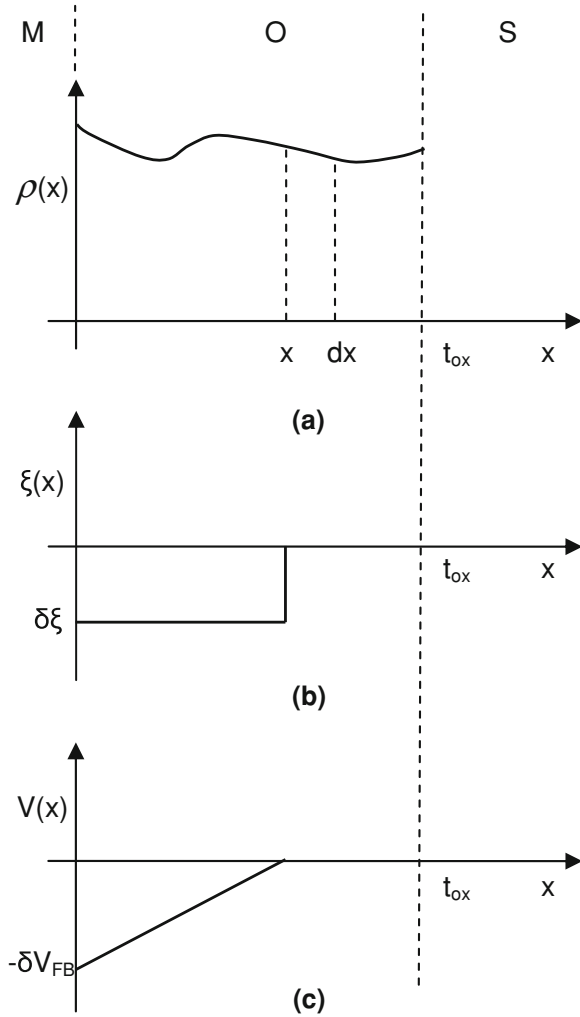
For ensuring flat band condition in the silicon, ζ_{ox} must be zero on the right hand side of the discontinuity. Thus, the profile of the electric field should be as shown in Fig. 2.9b and the corresponding gate voltage that ensures the flat band condition is given by:

$$\delta V_{FB} = \frac{-\rho(x)xdx}{\epsilon_o\epsilon_{ox}} \quad (2.14)$$

Using a classical result of electrostatics, namely the superposition theorem, the effects of all layers comprised between zero and t_{ox} are added and the gate voltage shift ΔV_{FB} , which is necessary to ensure a flat-band condition at the Si–SiO₂ interface, is found to be

$$\Delta V_{FB} = - \int_0^{t_{ox}} \frac{\rho(x)xdx}{\epsilon_o\epsilon_{ox}}. \quad (2.15)$$

Fig. 2.9 Distribution of **a** oxide charges, **b** electric field and **c** voltage within the oxide of MOS structure



The effect of each charge layer depends on its distance from the oxide-silicon interface as given in Eq. (2.15). A layer has no effect if it is located at the metal-oxide interface and has a maximum effect if it is located at the oxide-silicon interface.

References

1. Nicollian, E.H., Brews, J.R.: MOS Physics and Technology. Wiley, New York (1982)
2. Goetzberger, A., Sze, S.M.: Metal-insulator-semiconductor (MIS) physics. In: Wolfe, R. (ed.) Applied Solid State Science. Academic Press, New York (1969)

3. Grove, A.S.: *Physics and Technology of Semiconductor Devices*. Wiley, New York (1967)
4. Richman, P.: *MOS Field-Effect Transistors and Integrated Circuits*. Wiley, New York (1973)
5. Prezewlocki, H.M.: Work function difference in MOS structures: current understanding and new measurement methods (1982)
6. Maykusiak, B., Jakubowski, A.: A new method for the simultaneous determination of the surface-carrier mobility and the metal-semiconductor work-function difference in MOS transistors. *IEEE Trans. Elect. Dev.* **ED-35**, 439–443 (1988)
7. McNutt, M.J., Sah, C.T.: Determination of the MOS oxide capacitance. *J. Appl. Phys.* **46**, 3909–3913 (1975)
8. Deal, B.E.: Standardized terminology for oxide charge associated with thermally oxidized silicon. *IEEE Trans. Elect. Dev.* **ED-27**, 606–608 (1980)
9. Deal, B.E.: The current understanding of charges in the thermally oxidized silicon structure. *J. Electrochem. Soc.* **121**, 198–205 (1974)
10. Terman, M.: An investigation of surface states at silicon–silicon oxide interface employing metal oxide silicon diodes. *Solid St. Elect.* **5**, 285–299 (1962)
11. Snow, E.H., Grove, A.S., Deal, B.E., et al.: Ion transport phenomena in insulating films. *J. Appl. Phys.* **36**, 1664–1673 (1965)

Transport in Metal-Oxide-Semiconductor Structures

Mobile Ions Effects on the Oxide Properties

Bentarzi, H.

2011, XIV, 106 p., Hardcover

ISBN: 978-3-642-16303-6