

THE OXFORD SERIES IN ELECTRICAL AND COMPUTER
ENGINEERING

Adel S. Sedra, Series Editor

Allen and Holberg, *CMOS Analog Circuit Design*, 2nd Edition
Bobrow, *Elementary Linear Circuit Analysis*, 2nd Edition
Bobrow, *Fundamentals of Electrical Engineering*, 2nd Edition
Burns and Roberts, *An Introduction to Mixed-Signal IC Test and Measurement*
Campbell, *The Science and Engineering of Microelectronic Fabrication*, 2nd Edition
Chen, *Digital Signal Processing*
Chen, *Linear System Theory and Design*, 3rd Edition
Chen, *Signals and Systems*, 3rd Edition
Comer, *Digital Logic and State Machine Design*, 3rd Edition
Comer, *Microprocessor-based System Design*
Cooper and McGillem, *Probabilistic Methods of Signal and System Analysis*, 3rd Edition
DeCarlo and Lin, *Linear Circuit Analysis*, 2nd Edition
Dimitrijević, *Understanding Semiconductor Devices*
Fortney, *Principles of Electronics: Analog & Digital*
Franco, *Electric Circuits Fundamentals*
Ghausi, *Electronic Devices and Circuits: Discrete and Integrated*
Guru and Hiziroglu, *Electric Machinery and Transformers*, 3rd Edition
Houts, *Signal Analysis in Linear Systems*
Jones, *Introduction to Optical Fiber Communication Systems*
Krein, *Elements of Power Electronics*
Kuo, *Digital Control Systems*, 3rd Edition
Lathi, *Linear Systems and Signals*, 2nd Edition
Lathi, *Modern Digital and Analog Communications Systems*, 3rd Edition
Lathi, *Signal Processing and Linear Systems*
Martin, *Digital Integrated Circuit Design*
Miner, *Lines and Electromagnetic Fields for Engineers*
Parhami, *Computer Arithmetic*
Roberts and Sedra, *SPICE*, 2nd Edition
Roulston, *An Introduction to the Physics of Semiconductor Devices*
Sadiku, *Elements of Electromagnetics*, 3rd Edition
Santina, Stubbcrud, and Hostetter, *Digital Control System Design*, 2nd Edition
Sarna, *Introduction to Electrical Engineering*
Schaumann and Van Valkenburg, *Design of Analog Filters*
Schwarz and Oldham, *Electrical Engineering: An Introduction*, 2nd Edition
Sedra and Smith, *Microelectronic Circuits*, 5th Edition
Stefani, Savant, Shahian, and Hostetter, *Design of Feedback Control Systems*, 4th Edition
Tsitvidis, *Operation and Modeling of the MOS Transistor*, 2nd Edition
Van Valkenburg, *Analog Filter Design*
Warner and Grung, *Semiconductor Device Electronics*
Wolovich, *Automatic Control Systems*
Yariv, *Optical Electronics in Modern Communications*, 5th Edition
Žak, *Systems and Control*

FIFTH EDITION

MICROELECTRONIC CIRCUITS

Adel S. Sedra
University of Waterloo

Kenneth C. Smith
University of Toronto

New York Oxford
OXFORD UNIVERSITY PRESS
2004



region is a straight line rather than the exponentials shown in Fig. 3.50.

(a) For the short-base diode, sketch a figure corresponding to Fig. 3.50, and assume, as in Fig. 3.50, that $N_A \gg N_D$.

(b) Following a derivation similar to that given on page 205–206, show that if the widths of the p and n regions are denoted W_p and W_n then

$$I = Aq n_i^2 \left[\frac{D_p}{(W_n - x_n)N_D} + \frac{D_n}{(W_p - x_p)N_A} \right] (e^{V/V_T} - 1)$$

and

$$Q_p = \frac{1}{2} \frac{(W_n - x_n)^2}{D_p} I_p$$

$$\approx \frac{1}{2} \frac{W_n^2}{D_p} I_p, \quad \text{for } W_n \gg x_n$$

(c) Also, assuming $Q = Q_p$, $I \approx I_p$, show that

$$C_d = \frac{\tau_T I}{V_T}$$

where

$$\tau_T = \frac{1}{2} \frac{W_n^2}{D_p}$$

(d) If a designer wishes to limit C_d to 8 pF at $I = 1$ mA, what should W_n be? Assume $D_p = 10$ cm²/s.

CHAPTER 4

MOS Field-Effect Transistors (MOSFETs)

Introduction	235	4.8 The MOSFET Internal Capacitances and High-Frequency Model	320
4.1 Device Structure and Physical Operation	236	4.9 Frequency Response of the CS Amplifier	326
4.2 Current–Voltage Characteristics	248	4.10 The CMOS Digital Logic Inverter	336
4.3 MOSFET Circuits at DC	262	4.11 The Depletion-Type MOSFET	346
4.4 The MOSFET as an Amplifier and as a Switch	270	4.12 The SPICE MOSFET Model and Simulation Example	351
4.5 Biasing in MOS Amplifier Circuits	280	Summary	359
4.6 Small-Signal Operation and Models	287	Problems	360
4.7 Single-Stage MOS Amplifiers	299		

INTRODUCTION

Having studied the junction diode, which is the most basic two-terminal semiconductor device, we now turn our attention to three-terminal semiconductor devices. Three-terminal devices are far more useful than two-terminal ones because they can be used in a multitude of applications, ranging from signal amplification to digital logic and memory. The basic principle involved is the use of the voltage between two terminals to control the current flowing in the third terminal. In this way a three-terminal device can be used to realize a controlled source, which as we have learned in Chapter 1 is the basis for amplifier design. Also, in the extreme, the control signal can be used to cause the current in the third terminal to change from zero to a large value, thus allowing the device to act as a switch. As we also

learned in Chapter 1, the switch is the basis for the realization of the logic inverter, the basic element of digital circuits.

There are two major types of three-terminal semiconductor device: the metal-oxide-semiconductor field-effect transistor (MOSFET), which is studied in this chapter, and the bipolar junction transistor (BJT), which we shall study in Chapter 5. Although each of the two transistor types offers unique features and areas of application, the MOSFET has become by far the most widely used electronic device, especially in the design of integrated circuits (ICs), which are circuits fabricated on a single silicon chip.

Compared to BJTs, MOSFETs can be made quite small (i.e., requiring a small area on the silicon IC chip), and their manufacturing process is relatively simple (see Appendix A). Also, their operation requires comparatively little power. Furthermore, circuit designers have found ingenious ways to implement digital and analog functions utilizing MOSFETs almost exclusively (i.e., with very few or no resistors). All of these properties have made it possible to pack large numbers of MOSFETs (>200 million!) on a single IC chip to implement very sophisticated, very-large-scale-integrated (VLSI) circuits such as those for memory and microprocessors. Analog circuits such as amplifiers and filters are also implemented in MOS technology, albeit in smaller less-dense chips. Also, both analog and digital functions are increasingly being implemented on the same IC chip, in what is known as mixed-signal design.

The objective of this chapter is to develop in the reader a high degree of familiarity with the MOSFET: its physical structure and operation, terminal characteristics, circuit models, and basic circuit applications, both as an amplifier and a digital logic inverter. Although discrete MOS transistors exist, and the material studied in this chapter will enable the reader to design discrete MOS circuits, our study of the MOSFET is strongly influenced by the fact that most of its applications are in integrated-circuit design. The design of IC analog and digital MOS circuits occupies a large proportion of the remainder of this book.



4.1 DEVICE STRUCTURE AND PHYSICAL OPERATION

The enhancement-type MOSFET is the most widely used field-effect transistor. In this section, we shall study its structure and physical operation. This will lead to the current-voltage characteristics of the device, studied in the next section.

4.1.1 Device Structure

Figure 4.1, shows the physical structure of the n -channel enhancement-type MOSFET. The meaning of the names “enhancement” and “ n -channel” will become apparent shortly. The transistor is fabricated on a p -type substrate, which is a single-crystal silicon wafer that provides physical support for the device (and for the entire circuit in the case of an integrated circuit). Two heavily doped n -type regions, indicated in the figure as the n^+ source¹ and the n^+ drain regions, are created in the substrate. A thin layer of silicon dioxide (SiO_2) of thickness t_{ox} (typically 2–50 nm),² which is an excellent electrical insulator, is grown on the surface of the substrate, covering the area between the source and drain regions. Metal is deposited on top of the oxide layer to form the gate electrode of the device. Metal contacts are also made to the source region, the drain region, and the substrate, also known as the

¹ The notation n^+ indicates heavily doped n -type silicon. Conversely, n^- is used to denote lightly doped n -type silicon. Similar notation applies for p -type silicon.

² A nanometer (nm) is 10^{-9} m or 0.001 μm . A micrometer (μm), or micron, is 10^{-6} m. Sometimes the oxide thickness is expressed in angstroms. An angstrom ($\text{\AA}$) is 10^{-10} m, or 10^{-2} nm.

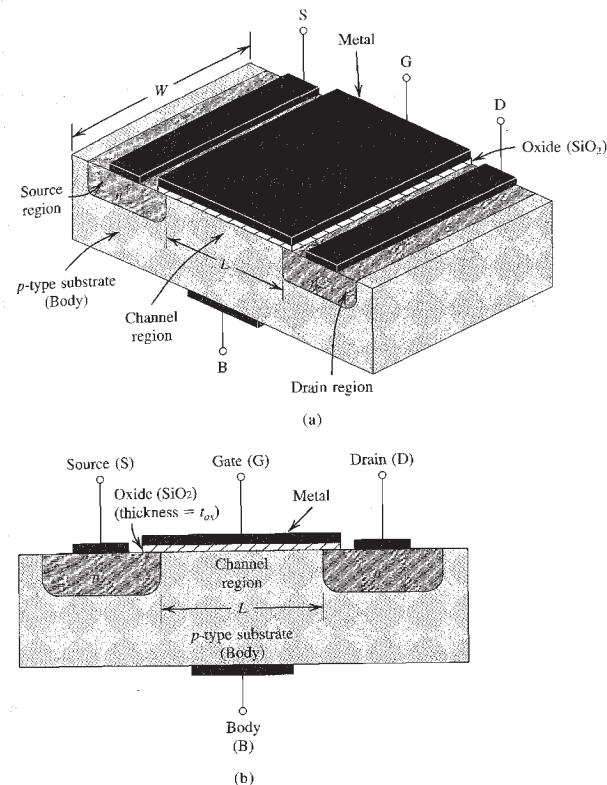


FIGURE 4.1 Physical structure of the enhancement-type NMOS transistor: (a) perspective view; (b) cross-section. Typically $L = 0.1$ to $3 \mu\text{m}$, $W = 0.2$ to $100 \mu\text{m}$, and the thickness of the oxide layer (t_{ox}) is in the range of 2 to 50 nm.

body.³ Thus four terminals are brought out: the gate terminal (G), the source terminal (S), the drain terminal (D), and the substrate or body terminal (B).

At this point it should be clear that the name of the device (metal-oxide-semiconductor FET) is derived from its physical structure. The name, however, has become a general one and is

³ In Fig. 4.1, the contact to the body is shown on the bottom of the device. This will prove helpful later in explaining a phenomenon known as the “body effect.” It is important to note, however, that in actual ICs, contact to the body is made at a location on the top of the device.

used also for FETs that do not use metal for the gate electrode. In fact, most modern MOSFETs are fabricated using a process known as silicon-gate technology, in which a certain type of silicon, called polysilicon, is used to form the gate electrode (see Appendix A). Our description of MOSFET operation and characteristics applies irrespective of the type of gate electrode.

Another name for the MOSFET is the **insulated-gate FET** or **IGFET**. This name also arises from the physical structure of the device, emphasizing the fact that the gate electrode is electrically insulated from the device body (by the oxide layer). It is this insulation that causes the current in the gate terminal to be extremely small (of the order of 10^{-15} A).

Observe that the substrate forms *pn* junctions with the source and drain regions. In normal operation these *pn* junctions are kept reverse-biased at all times. Since the drain will be at a positive voltage relative to the source, the two *pn* junctions can be effectively cut off by simply connecting the substrate terminal to the source terminal. We shall assume this to be the case in the following description of MOSFET operation. Thus, here, the substrate will be considered as having no effect on device operation, and the MOSFET will be treated as a three-terminal device, with the terminals being the gate (G), the source (S), and the drain (D). It will be shown that a voltage applied to the gate controls current flow between source and drain. This current will flow in the longitudinal direction from drain to source in the region labeled "channel region." Note that this region has a length L and a width W , two important parameters of the MOSFET. Typically, L is in the range of $0.1\text{ }\mu\text{m}$ to $3\text{ }\mu\text{m}$, and W is in the range of $0.2\text{ }\mu\text{m}$ to $100\text{ }\mu\text{m}$. Finally, note that the MOSFET is a symmetrical device; thus its source and drain can be interchanged with no change in device characteristics.

4.1.2 Operation with No Gate Voltage

With no bias voltage applied to the gate, two back-to-back diodes exist in series between drain and source. One diode is formed by the *pn* junction between the n^+ drain region and the *p*-type substrate, and the other diode is formed by the *pn* junction between the *p*-type substrate and the n^+ source region. These back-to-back diodes prevent current conduction from drain to source when a voltage v_{DS} is applied. In fact, the path between drain and source has a very high resistance (of the order of $10^{12}\text{ }\Omega$).

4.1.3 Creating a Channel for Current Flow

Consider next the situation depicted in Fig. 4.2. Here we have grounded the source and the drain and applied a positive voltage to the gate. Since the source is grounded, the gate voltage appears in effect between gate and source and thus is denoted v_{GS} . The positive voltage on the gate causes, in the first instance, the free holes (which are positively charged) to be repelled from the region of the substrate under the gate (the channel region). These holes are pushed downward into the substrate, leaving behind a carrier-depletion region. The depletion region is populated by the bound negative charge associated with the acceptor atoms. These charges are "uncovered" because the neutralizing holes have been pushed downward into the substrate.

As well, the positive gate voltage attracts electrons from the n^+ source and drain regions (where they are in abundance) into the channel region. When a sufficient number of electrons accumulate near the surface of the substrate under the gate, an *n* region is in effect created, connecting the source and drain regions, as indicated in Fig. 4.2. Now if a voltage is applied between drain and source, current flows through this induced *n* region, carried by the mobile electrons. The induced *n* region thus forms a **channel** for current flow from drain to source and is aptly called so. Correspondingly, the MOSFET of Fig. 4.2 is called an ***n*-channel MOSFET** or, alternatively, an **NMOS transistor**. Note that an *n*-channel MOSFET is formed in a *p*-type substrate: The channel is created by *inverting* the substrate surface from *p* type to *n* type. Hence the induced channel is also called an **inversion layer**.

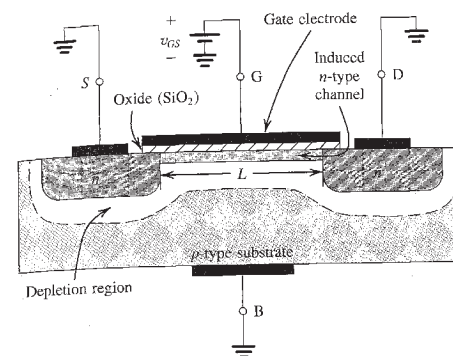


FIGURE 4.2 The enhancement-type NMOS transistor with a positive voltage applied to the gate. An *n* channel is induced at the top of the substrate beneath the gate.

The value of v_{GS} at which a sufficient number of mobile electrons accumulate in the channel region to form a conducting channel is called the **threshold voltage** and is denoted V_t .⁴ Obviously, V_t for an *n*-channel FET is positive. The value of V_t is controlled during device fabrication and typically lies in the range of 0.5 V to 1.0 V.

The gate and the channel region of the MOSFET form a parallel-plate capacitor, with the oxide layer acting as the capacitor dielectric. The positive gate voltage causes positive charge to accumulate on the top plate of the capacitor (the gate electrode). The corresponding negative charge on the bottom plate is formed by the electrons in the induced channel. An electric field thus develops in the vertical direction. It is this field that controls the amount of charge in the channel, and thus it determines the channel conductivity and, in turn, the current that will flow through the channel when a voltage v_{DS} is applied.

4.1.4 Applying a Small v_{DS}

Having induced a channel, we now apply a positive voltage v_{DS} between drain and source, as shown in Fig. 4.3. We first consider the case where v_{DS} is small (i.e., 50 mV or so). The voltage v_{DS} causes a current i_D to flow through the induced *n* channel. Current is carried by free electrons traveling from source to drain (hence the names source and drain). By convention, the direction of current flow is opposite to that of the flow of negative charge. Thus the current in the channel, i_D , will be from drain to source, as indicated in Fig. 4.3. The magnitude of i_D depends on the density of electrons in the channel, which in turn depends on the magnitude of v_{GS} . Specifically, for $v_{GS} = V_t$, the channel is just induced and the current conducted is still negligibly small. As v_{GS} exceeds V_t , more electrons are attracted into the channel. We may visualize the increase in charge carriers in the channel as an increase in the channel depth. The result is a channel of increased conductance or, equivalently, reduced resistance. In fact, the conductance of the channel is proportional to the **excess gate voltage** ($v_{GS} - V_t$), also

⁴ Some texts use V_T to denote the threshold voltage. We use V_t to avoid confusion with the thermal voltage V_T .