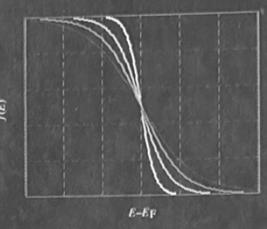
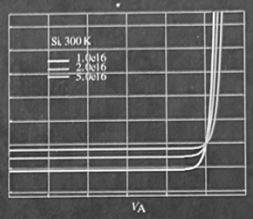
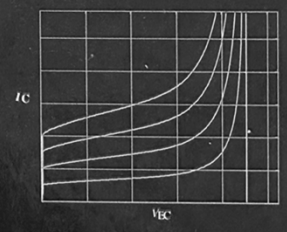
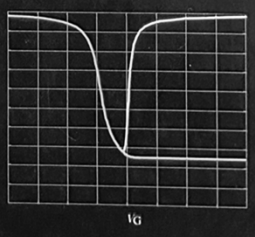
Robert F. Pierret





Semiconductor Device Fundamentals





with Computer-Based Exercises and Homework Problems

APPLE-1030

SEMICONDUCTOR DEVICE FUNDAMENTALS

Robert F. Pierret
School of Electrical and Computer Engineering
Purdue University



ADDISON-WESLEY PUBLISHING COMPANY
Reading, Massachusetts · Menlo Park, California · New York
Don Mills, Ontario · Wokingham, England · Amsterdam · Bonn
Sydney · Singapore · Tokyo · Madrid · San Juan · Milan · Paris

Katherine Harutunian *Associate Editor*
Helen Wythe *Senior Production Supervisor*
Hugh Crawford *Manufacturing Supervisor*
Barbara Atkinson *Associate Cover Design Supervisor*
Peter Blaiwas *Cover design*
Kenneth J. Wilson *Text design*
Joyce Grandy *Copyeditor*
Sandra Rigney *Production Packaging Services*
G & S Typesetters, Inc. *Composition*
Publishers' Design and Production Services, Inc. *Illustrations*

Library of Congress Cataloging-in-Publication Data
Pierret, Robert F.
Semiconductor device fundamentals / Robert F. Pierret.
p. cm.
Includes index.
ISBN 0-201-54393-1
1. Semiconductors. I. Title.
TK7871.85.P484 1996
621.3815'2–dc20 95-17387
CIP

Many of the designations used by manufacturers and sellers to distinguish their products are claimed as trademarks. Where those designations appear in this book, and Addison-Wesley was aware of a trademark claim, the designations have been printed in initial caps or all caps.

MATLAB is a registered trademark of The MathWorks, Inc.,
24 Prime Park Way, Natick, MA 01760-1500.
Phone: (508) 653-1415. Fax: (508) 653-2997
E-mail: info@mathworks.com
WWW:http://www.mathworks.com

Access the latest information about Addison-Wesley books from our Internet gopher site or our World Wide Web page:
gopher aw.com
http://www.aw.com/cseng/

Copyright: © 1996 by Addison-Wesley Publishing Company, Inc.

Reprinted with corrections March, 1996

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the written permission of the publisher.

Printed in the United States of America

8910 MA 02 01 00

15 Field Effect Introduction—The J-FET and MESFET

15.1 GENERAL INTRODUCTION

Historically, the field effect phenomenon was the basis for the first type of solid-state transistor ever proposed. Field effect transistors predate the bipolar junction transistor by approximately 20 years. As recorded in a series of patents filed in the 1920s and 1930s, J. E. Lilienfeld in the United States and O. Heil working in Germany independently conceived a transistor structure of the form shown in Fig. 15.1. The device worked on the principle that a voltage applied to the metallic plate modulated the conductance of the underlying semiconductor, which in turn modulated the current flowing between ohmic contacts A and B. This phenomenon, where the conductivity of a semiconductor is modulated by an electric field applied normal to the surface of the semiconductor, has been named the *field effect.*

The early field effect transistor proposals were of course somewhat ahead of their time. Modern-day semiconductor materials were just not available and technological immaturity, in general, retarded the development of field effect structures for many years. A practical implementation had to await the successful development of other solid-state devices, notably the bipolar junction transistor, in the late 1940s and early 1950s. The first modern-day field effect device, the junction field effect transistor (J-FET), was proposed and analyzed by W. Shockley[1] in 1952. Operational J-FETs were subsequently built by Dacey

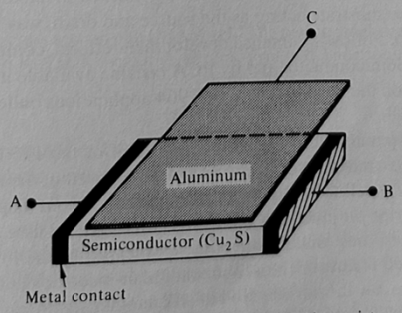


Figure 15.1 Idealization of the Lilienfeld transistor.

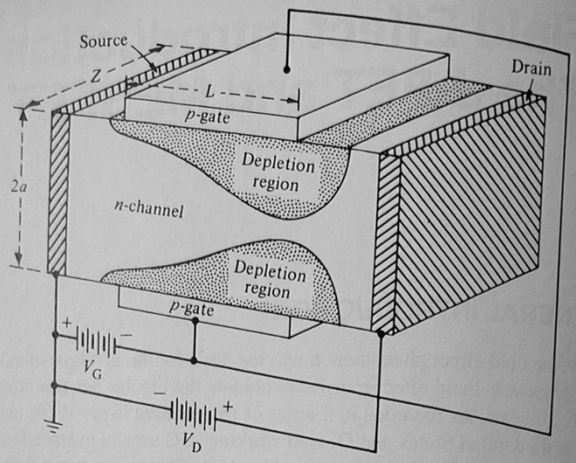


Figure 15.2 Schematic diagram of the junction field effect transistor. (From Dacey and Ross[3], © 1955 AT & T. Reprinted with permission.)

and Ross[2] in 1953. As pictured schematically in Fig. 15.2 (which was adapted from an early Dacey and Ross publication[3]), *pn* junctions replaced the metallic plate in the Lilienfeld structure, the A and B contacts became known as the source and drain, and the field-effect electrode was named the gate. In the J-FET it is the depletion regions associated with the *pn* junctions that directly modulate the semiconductor conductivity between the source and drain contacts.

A key component of present-day microelectronics, the metal-oxide-semiconductor field effect transistor (MOSFET), achieved practical status during the early 1960s. The now-familiar planar version of the structure pictured in Fig. 15.3, a structure with thermally grown SiO_2 functioning as the gate insulator, a surface-inversion channel, and islands doped opposite to the substrate acting as the source and drain, was first reported by Kahng and Atalla[4] in 1960. Following intense development efforts, commercial MOSFETs produced by Fairchild Semiconductor and by RCA became available in late 1964. Figure 15.3 originally appeared on the cover page of the 1964 applications bulletin describing the early Fairchild MOSFET[5].

A major development occurring in the latter half of the 1960s was the invention by Dennard[6] of the one-transistor dynamic memory cell used in the random access memory (DRAM). The DRAM cell (see Fig. 15.4) is an integrated combination of a charge storage element (a capacitor or *pn* junction) and a MOSFET utilized as a switch. A quote from a recent review article[8] best underscores the significance of the invention and its impact. "The 1-T DRAM cell is now the most abundant man-made object on this planet earth."

The early 1970s saw the introduction of still another significant field effect structure—the charge-coupled device (CCD). Physically, a CCD might be viewed as a MOSFET with

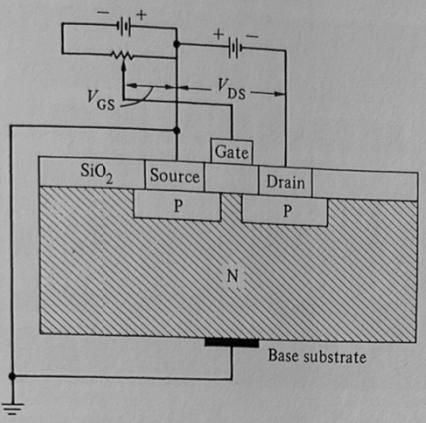


Figure 15.3 Cross section of a MOSFET with correct biasing polarity. (From MacDougall[5].)

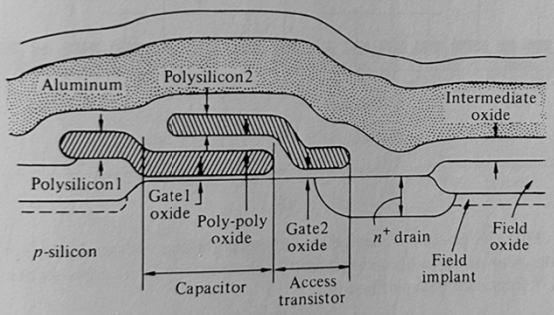
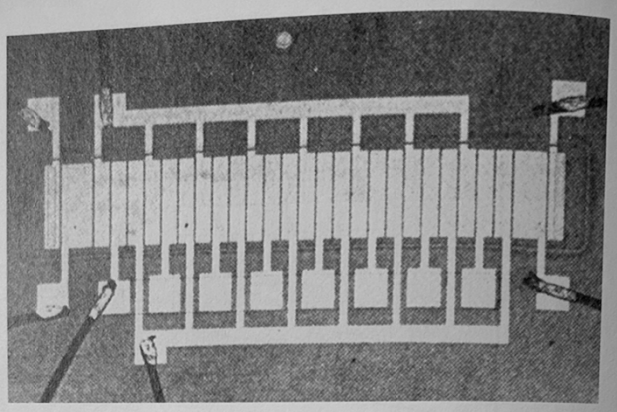


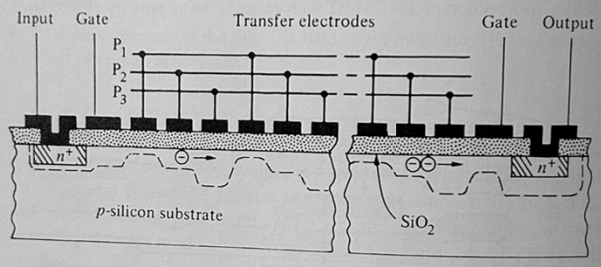
Figure 15.4 Schematic cross section of a DRAM cell. The planar, single-capacitor cell of the type shown was employed in 16 K–256 K memories. (From Wyns and Anderson[7], © 1989 IEEE.)

a segmented gate. As pictured in Fig. 15.5, the proper application of biases to the CCD gates induces the systematic movement or transfer of stored charge along the surface channel and into the device output. The image sensing element in the camcorder, a combination home TV camera and video recorder, is a CCD containing a two-dimensional array of gates. Stored charge is produced in proportion to the optical image incident on the two-dimensional gate array. The electrical analog of the optical image is subsequently transferred to the output and converted into an electrical signal.

The invention, development, and evolution of field effect devices continues to the pres-



(a)



(b)

Figure 15.5 First charge-coupled device comprising eight three-phase elements and input-output gates and diodes, shown (a) in plan view and (b) schematically in its cross-sectional view. (From Tompsett et al.[9] Reprinted with permission.)

ent. Notably, in progressing through the 1980s, complementary MOS or CMOS became the integrated circuit technology of choice. Figure 15.6 shows a cut-away view of a fundamental building block in CMOS circuitry, the CMOS inverter. The term "complementary" denotes the fact that, unlike other MOS circuit technologies, both *n*-channel (electron current carrying) MOSFETs and *p*-channel (hole current carrying) MOSFETs are fabricated on the same chip. Although CMOS came to the forefront in the 1980s because of its lower power dissipation and other circuit-related advantages, the change was evolutionary—CMOS is not a new circuit technology. It was initially conceived by Wanlass[10] in the early 1960s and was even mass produced as part of LED watch circuitry in 1972.

Whereas the use of CMOS exemplifies continuing evolutionary changes, there is also continuing field effect device invention and development. The modulation doped field effect transistor (MODFET) pictured in Fig. 15.7 is an excellent example. Although the elec-

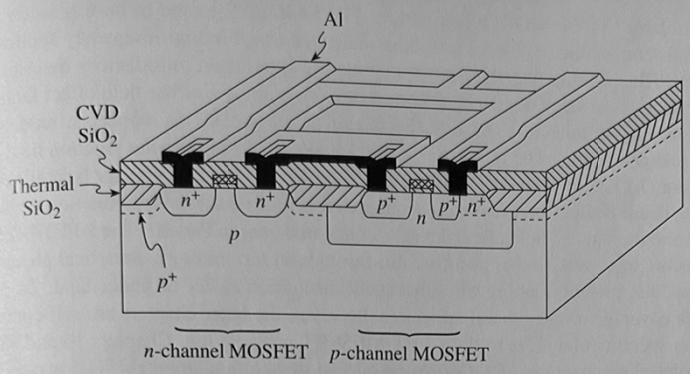


Figure 15.6 Cut-away schematic view of a CMOS inverter. (From Fonstad[11], © 1994 by McGraw-Hill, Inc. Reprinted by permission of the publisher.)

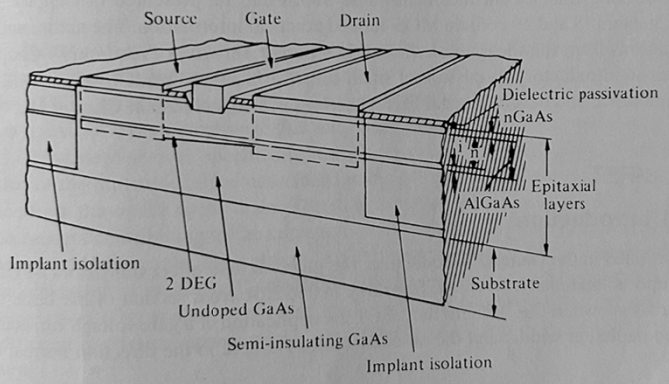


Figure 15.7 Perspective view of the basic MODFET structure. (From Drummond et al.[12], © 1986 IEEE.)

tron transport properties of GaAs are superior to silicon, GaAs technology lacks a quality homomorphic (Ga or As containing) insulator comparable to SiO_2. Moreover, the deposition of other insulators on GaAs has consistently failed to produce devices with acceptable electrical characteristics. The emergence of sophisticated deposition techniques such as molecular beam epitaxy (MBE), however, has made it possible to grow layers of different lattice-matched semiconductors on top of each other. The alloy semiconductor AlGaAs appearing in Fig. 15.7 has a larger band gap than GaAs and essentially assumes the role

of "insulator" in the MODFET structure. The MODFET, touted to be the fastest three-terminal semiconductor device, is finding increased usage in high-frequency applications.

Part III, spanning five chapters, is intended to provide an introductory description of field effect devices and device operation. The core members of the field effect family are examined in some detail to expose the reader to relevant terms, concepts, models, and analytical procedures. The remainder of this chapter is devoted to the junction field effect transistor (J-FET) and the closely related MESFET (metal-semiconductor field effect transistor). These devices are viewed as conceptual bridges between the junction devices in Part II and the "pure" field effect devices considered later in Part III. The J-FET/MESFET discussion, moreover, serves as a first run-through on terminology, analytical procedures, and the like, thereby making the subsequent discussion easier to understand. To permit flexible coverage of the general topic area, however, the latter chapters are self-contained, with no specific references to the J-FET/MESFET presentation. Chapters 16 and 17 provide a baseline coverage of MOS structural and device fundamentals. The two-terminal MOS-capacitor or MOS-C is the structural heart of all MOS devices. In Chapter 16 we define the ideal MOS-C structure, present a qualitative and quantitative description of the structure in the static state, and examine the capacitance-voltage characteristics derived from the device. The operational description and analysis of the basic transistor configuration, the long-channel enhancement-mode MOSFET, are presented in Chapter 17. Finally, Chapters 18 and 19 contain MOS-related practical information. The nature and effect of deviations from the ideal are described in Chapter 18, while "read-only" Chapter 19 provides an introductory discussion of small-dimension effects in MOSFETs. A brief survey of modified and developmental FET structures is also included in Chapter 19.

15.2 J-FET

15.2.1 Introduction

As mentioned in the General Introduction, the junction field effect transistor (J-FET) was first proposed and analyzed by W. Shockley in 1952. A cross section of the basic device structure is shown in Fig. 15.8. In the J-FET the application of a gate voltage varies the *pn*-junction depletion widths and the associated electric field in the direction normal to the

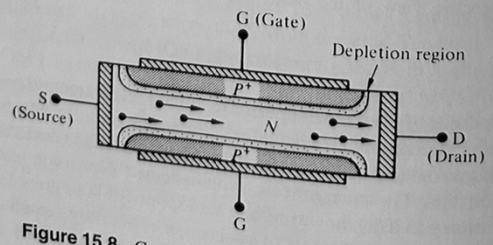


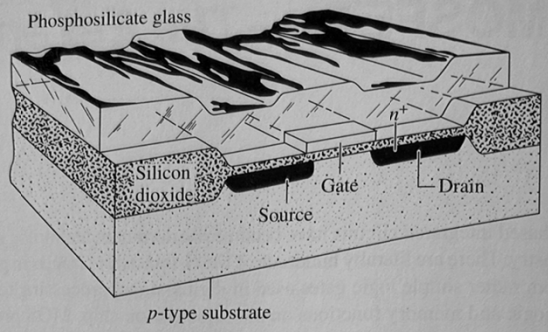
Figure 15.8 Cross section of the basic J-FET structure.

17 MOSFETs—*The Essentials*

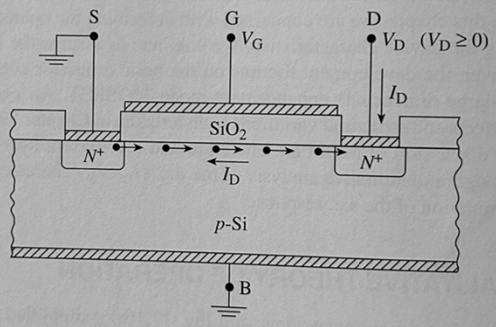
MOSFET-based integrated circuits have become the dominant technology in the semiconductor industry. There are literally hundreds of MOS-transistor circuits in production today, ranging from rather simple logic gates used in digital-signal processing to custom designs with both logic and memory functions on the same silicon chip. MOS products are found in a mind-boggling number of electronic systems, including the now commonplace personal computer. Initially the MOS-transistor was identified by several competing acronyms, namely, metal–oxide–semiconductor transistor (MOST), insulated gate field effect transistor (IGFET), and metal–oxide–semiconductor field effect transistor (MOSFET). (PIGFET and MISFET were even suggested with a smile at one time or another.) With the passage of time, however, the transistor structure has commonly come to be known as the MOSFET. In this chapter we are concerned with describing the operation of the MOSFET and modeling the device characteristics. We continue to assume the MOS structure to be ideal. Moreover, the development focuses on the basic transistor configuration, the long-channel (or large-dimension) enhancement-mode MOSFET. An examination of small-dimension effects and structural variations is undertaken in Chapter 19. We begin here with a qualitative discussion of MOSFET operation and d.c. current flow inside the structure, progress through a quantitative analysis of the d.c. (I_D–V_D) characteristics, and conclude with an examination of the a.c. response.

17.1 QUALITATIVE THEORY OF OPERATION

Figure 17.1(a) shows a cut-away view, and Fig. 17.1(b) a simplified cross-sectional view, of the basic MOS-transistor configuration. Physically, the MOSFET is essentially nothing more than an MOS-capacitor with two *pn* junctions placed immediately adjacent to the region of the semiconductor controlled by the MOS-gate. The Si substrate can be either *p*-type (as pictured) or *n*-type; p^+ junction islands are of course required in *n*-bulk devices. Also shown in Fig. 17.1(b) are the standard terminal and d.c. voltage designations. The drain current (I_D), which flows in response to the applied terminal voltages, is the primary d.c. observable. Consistent with the naming of the device leads, the current flow is always such that carriers (electrons in the present case) enter the structure through the source (S), leave through the drain (D), and are subject to the control or gating action of the gate (G). The voltage applied to the gate relative to ground is V_G, while the drain voltage relative to ground is V_D. Unless stated otherwise, we will assume that the source and back are grounded. Please note that under normal operational conditions the drain bias is always



(a)



(b)

Figure 17.1 The basic MOSFET structure. (a) Idealized cut-away view of a p-bulk (n-channel) MOSFET. (b) Simplified cross-sectional view of the p-bulk (n-channel) structure showing the terminal designations, carrier and current flow directions, and standard biasing conditions. [(a) From Beadle, Tsai, and Plummer[15]. Reprinted with permission of AT&T.]

such as to reverse bias the drain *pn* junction ($V_D \geq 0$ for the Fig. 17.1(b) device). Finally, reflecting the normal flow of electrons within the structure, the drain current for the *p*-bulk device is taken to be positive when flowing from the external circuit into the drain terminal.

To determine how the drain current is expected to vary as a function of the applied terminal voltages, let us first conceptually set $V_D = 0$ and examine the situation inside the structure as a function of the imposed gate voltage. When V_G is accumulation or depletion biased ($V_G \leq V_T$, where V_T is the depletion–inversion transition-point voltage), the gated region between the source and drain islands contains either an excess or deficit of holes and very few electrons. Thus, looking along the surface between the n^+ islands under the cited conditions, one effectively sees an open circuit. When V_G is inversion biased ($V_G > V_T$), however, an inversion layer containing mobile electrons is formed adjacent to the Si surface. Now looking along the surface between the n^+ islands one sees, as pictured in Fig. 17.2(a), an induced "*n*-type" region (the inversion layer) or conducting *channel* connecting the source and drain islands. Naturally, the greater the inversion bias, the greater the pileup of electrons at the Si surface and the greater the conductance of the inversion layer. An inverting gate bias, therefore, creates or induces a source-to-drain channel and determines the maximum conductance of the channel.

Turning next to the action of the drain bias, suppose an inversion bias $V_G > V_T$ is applied to the gate and the drain voltage is increased in small steps starting from $V_D = 0$. At $V_D = 0$ the situation inside the device is as pictured in Fig. 17.2(a), thermal equilibrium obviously prevails, and the drain current is identically zero. With V_D stepped to small positive voltages, the surface channel merely acts like a simple resistor and a drain current proportional to V_D begins to flow into the drain terminal. The portion of the I_D–V_D relationship corresponding to small V_D biases is shown as the line from the origin to point A in Fig. 17.3. Any $V_D > 0$ bias, it should be interjected, simultaneously reverse biases the drain *pn* junction, and the resulting reverse bias junction current flowing into the Si substrate does contribute to I_D. In well-made devices, however, the junction current is totally negligible compared to the channel current, provided V_D is less than the junction breakdown voltage.

Once V_D is increased above a few tenths of a volt, the device enters a new phase of operation. Specifically, the voltage drop from the drain to the source associated with the flow of current in the channel starts to negate the inverting effect of the gate. As pictured in Fig. 17.2(b), the depletion region widens in going down the channel from the source to the drain and the number of inversion layer carriers correspondingly decreases. The reduced number of carriers decreases the channel conductance, which in turn is reflected as a decrease in the slope of the observed I_D–V_D characteristic. Continuing to increase the drain voltage causes a progressive reduction in the channel carrier concentration and the systematic slope-over in the I_D–V_D characteristic noted in Fig. 17.3. The greatest decrease in channel carriers occurs near the drain, and eventually the inversion layer completely vanishes ($n_{|\text{surface}}$ drops below N_A) in the near vicinity of the drain (see Fig. 17.2c). The onset of surface depletion at the drain end of the channel, the special situation where the channel carrier concentration at the Si–SiO_2 interface immediately adjacent to the drain becomes equal to the bulk doping concentration, is referred to as *pinch-off*. When the

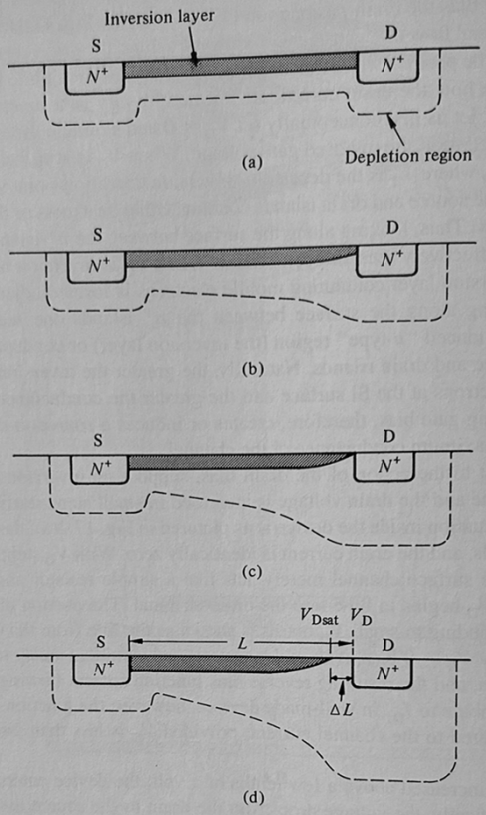


Figure 17.2 Visualization of various phases of $V_G > V_T$ MOSFET operation. (a) $V_D = 0$; (b) channel (inversion layer) narrowing under moderate V_D biasing; (c) pinch-off; and (d) postpinch-off ($V_D > V_{Dsat}$) operation. (Note that the inversion layer widths, depletion widths, etc. are not drawn to scale.)

channel pinches off inside the device, the point B is reached on the Fig. 17.3 characteristic; that is, the slope of the I_D–V_D characteristic becomes approximately zero.

For drain voltages in excess of the pinch-off voltage, V_{Dsat}, the pinched-off portion of the channel widens from just a point into a depleted channel section ΔL in extent (see Fig. 17.2d). Being a region with few carriers, and hence low conductance, the pinched-off ΔL section absorbs most of the voltage drop in excess of V_{Dsat}. Given a long-channel

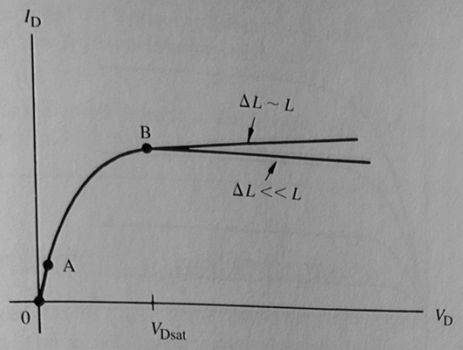


Figure 17.3 General variation of I_D with V_D for a given $V_G > V_T$.

device where $\Delta L \ll L$, the source to pinch-off region of the MOSFET will be essentially identical in shape and will have the same endpoint voltages for all $V_D \geq V_{Dsat}$. When the shape of a conductive region and the potential applied across the region do not change, the current through the region must also remain invariant. Thus, I_D remains approximately constant for drain voltages in excess of V_{Dsat} provided $\Delta L \ll L$. If ΔL is comparable to L, the same voltage drop (V_{Dsat}) will appear across a shorter channel ($L - \Delta L$) and, as noted in Fig. 17.3, the post-pinch-off I_D in such devices will increase somewhat with increasing $V_D > V_{Dsat}$.

Thus far we have examined the response of the MOSFET to the separate manipulation of the gate and drain biases. To establish a complete set of I_D–V_D characteristics, it is necessary to combine the results derived from the separate considerations. Clearly, for $V_G \leq V_T$, the gate bias does not create a surface channel and $I_D \simeq 0$ for all drain biases below the junction breakdown voltage. For all $V_G > V_T$ a characteristic of the form shown in Fig. 17.3 will be observed. Since the conductance of the channel increases with increasing V_G, it follows that the initial slope of the I_D–V_D characteristic will likewise increase with increasing V_G. Moreover, the greater the number of inversion layer carriers present when $V_D = 0$, the larger the drain voltage required to achieve pinch-off. Thus V_{Dsat} must increase with increasing V_G. From the foregoing arguments one concludes that the variation of I_D with V_D and V_G must be of the form displayed in Fig. 17.4.

The I_D–V_D characteristics just established confirm the "transistor" nature of the MOSFET structure. Transistor action results, of course, when I_D flowing in an output circuit is modulated by an input voltage applied to the gate. Relative to terminology, we should note that the portion of the characteristics where $V_D > V_{Dsat}$ for a given V_G is referred to as the *saturation* region of operation; the portion of the characteristics where $V_D < V_{Dsat}$ is called the *linear* (or sometimes *triode*) region of operation. Also, the MOSFET is known as an *n-channel* device when the channel carriers are electrons; when the channel carriers are holes the MOSFET is designated a *p-channel* device.