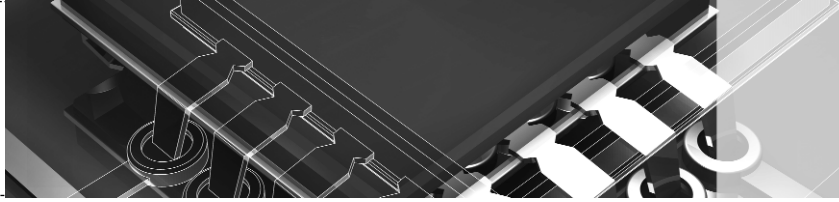


Third Edition



# CMOS Analog Circuit Design

Phillip E. Allen

*Professor Emeritus, Georgia Institute of Technology*

Douglas R. Holberg

*Consultant*

New York Oxford  
OXFORD UNIVERSITY PRESS

Page 1 of 227

[www.ebook3000.com](http://www.ebook3000.com)

REED EX1024  
IPR2024-00871

Oxford University Press, Inc., publishes works that further Oxford University's objective of excellence in research, scholarship, and education.

Oxford New York  
Auckland Cape Town Dar es Salaam Hong Kong Karachi  
Kuala Lumpur Madrid Melbourne Mexico City Nairobi  
New Delhi Shanghai Taipei Toronto

With offices in  
Argentina Austria Brazil Chile Czech Republic France Greece  
Guatemala Hungary Italy Japan Poland Portugal Singapore  
South Korea Switzerland Thailand Turkey Ukraine Vietnam

Copyright © 2012, 2002 by Oxford University Press, Inc.

For titles covered by Section 112 of the U.S. Higher Education Opportunity Act, please visit [www.oup.com/us/he](http://www.oup.com/us/he) for the latest information about pricing and alternate formats.

Published by Oxford University Press, Inc.  
198 Madison Avenue, New York, NY 10016

[www.oup.com](http://www.oup.com)

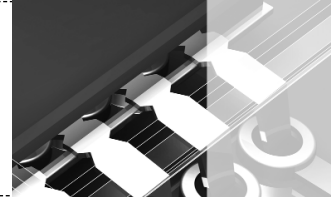
Oxford is a registered trademark of Oxford University Press

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 prior permission of Oxford University Press.

ISBN 978-0-19-976507-2

Printing number: 9 8 7 6 5 4 3 2 1  
Printed in the United States of America  
on acid-free paper

# CONTENTS



## Preface xi

## Chapter 1 Introduction and Background 1

- 1.1 Analog Integrated-Circuit Design 2
- 1.2 Notation, Symbolology, and Terminology 7
- 1.3 Analog Signal Processing 9
- 1.4 Example of Analog VLSI Mixed-Signal Circuit Design 10
- 1.5 Summary 15
- Problems 16
- References 17

## Chapter 2 CMOS Technology 18

- 2.1 Basic MOS Semiconductor Fabrication Processes 19
- 2.2 The pn Junction 33
- 2.3 The MOS Transistor 40
- 2.4 Passive Components 46
- 2.5 Other Considerations of CMOS Technology 53
- 2.6 Summary 61
- Problems 61
- References 65

## Chapter 3 CMOS Device Modeling 67

- 3.1 Simple MOS Large-Signal Model (SPICE LEVEL 1) 68
- 3.2 Other MOS Large-Signal Model Parameters 77
- 3.3 Small-Signal Model for the MOS Transistor 86

|                  |                                                   |            |
|------------------|---------------------------------------------------|------------|
| 3.4              | Computer Simulation Models                        | 90         |
| 3.5              | Subthreshold MOS Model                            | 96         |
| 3.6              | SPICE Simulation of MOS Circuits                  | 99         |
| 3.7              | Summary                                           | 108        |
|                  | <b>Problems</b>                                   | 109        |
|                  | <b>References</b>                                 | 112        |
| <b>Chapter 4</b> | <b>Analog CMOS Subcircuits</b>                    | <b>114</b> |
| 4.1              | MOS Switch                                        | 115        |
| 4.2              | MOS Diode/Active Resistor                         | 126        |
| 4.3              | Current Sinks and Sources                         | 128        |
| 4.4              | Current Mirrors                                   | 138        |
| 4.5              | Current and Voltage References                    | 147        |
| 4.6              | Temperature-Independent References                | 156        |
| 4.7              | Summary                                           | 173        |
| 4.8              | Design Problems                                   | 173        |
|                  | <b>Problems</b>                                   | 174        |
|                  | <b>References</b>                                 | 183        |
| <b>Chapter 5</b> | <b>CMOS Amplifiers</b>                            | <b>185</b> |
| 5.1              | Inverters                                         | 186        |
| 5.2              | Differential Amplifiers                           | 198        |
| 5.3              | Cascode Amplifiers                                | 218        |
| 5.4              | Current Amplifiers                                | 230        |
| 5.5              | Output Amplifiers                                 | 236        |
| 5.6              | Summary                                           | 247        |
|                  | <b>Problems</b>                                   | 248        |
|                  | <b>References</b>                                 | 260        |
| <b>Chapter 6</b> | <b>CMOS Operational Amplifiers</b>                | <b>261</b> |
| 6.1              | Design of CMOS Op Amps                            | 262        |
| 6.2              | Compensation of Op Amps                           | 271        |
| 6.3              | Design of the Two-Stage Op Amp                    | 286        |
| 6.4              | Power-Supply Rejection Ratio of Two-Stage Op Amps | 303        |
| 6.5              | Cascode Op Amps                                   | 310        |
| 6.6              | Simulation and Measurement of Op Amps             | 328        |
| 6.7              | Summary                                           | 342        |
|                  | <b>Problems</b>                                   | 342        |
|                  | <b>References</b>                                 | 351        |

## **Chapter 7 High-Performance CMOS Op Amps 353**

- 7.1 Buffered Op Amps 354
- 7.2 High-Speed/Frequency CMOS Op Amps 370
- 7.3 Differential-Output Op Amps 386
- 7.4 Micropower Op Amps 398
- 7.5 Low-Noise Op Amps 407
- 7.6 Low-Voltage Op Amps 419
- 7.7 Summary 436
- Problems 437**
- References 442**

## **Chapter 8 Comparators 444**

- 8.1 Characterization of a Comparator 444
- 8.2 Two-Stage, Open-Loop Comparators 450
- 8.3 Other Open-Loop Comparators 466
- 8.4 Improving the Performance of Open-Loop Comparators 469
- 8.5 Discrete-Time Comparators 480
- 8.6 High-Speed Comparators 488
- 8.7 Summary 493
- Problems 493**
- References 497**

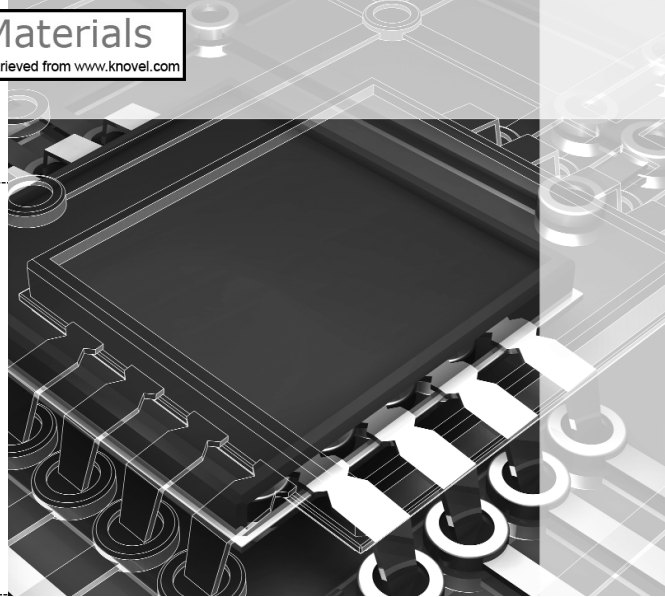
## **Chapter 9 Digital–Analog and Analog–Digital Converters 498**

- 9.1 Introduction and Characterization of Digital–Analog Converters 499
- 9.2 Parallel Digital–Analog Converters 510
- 9.3 Extending the Resolution of Parallel Digital–Analog Converters 522
- 9.4 Serial Digital–Analog Converters 534
- 9.5 Introduction and Characterization of Analog–Digital Converters 539
- 9.6 Serial Analog–Digital Converters 557
- 9.7 Medium-Speed Analog–Digital Converters 559
- 9.8 High-Speed Analog–Digital Converters 573
- 9.9 Oversampling Converters 589
- 9.10 Summary 604
- Problems 607**
- References 624**

|            |                                                                  |     |
|------------|------------------------------------------------------------------|-----|
| Appendix A | Circuit Analysis for Analog Circuit Design                       | 629 |
| Appendix B | Integrated-Circuit Layout                                        | 640 |
| Appendix C | CMOS Device Characterization                                     | 662 |
| Appendix D | Time and Frequency Domain Relationships for Second-Order Systems | 686 |
| Appendix E | Switched Capacitor Circuits                                      | 694 |
|            | Homework Problem Answers                                         | 727 |
|            | Index                                                            | 751 |

## CHAPTER 2

# CMOS Technology



Modern silicon integrated-circuit technologies consist of bipolar, BiCMOS, and CMOS, as illustrated in Fig. 2.0-1. For many years the dominant silicon integrated-circuit technology was bipolar, as evidenced by the ubiquitous monolithic operational amplifier and the TTL (transistor–transistor logic) family. In the early 1970s MOS technology was demonstrated to be viable in the area of dynamic random-access memories (DRAMs), microprocessors, and the 4000-series logic family. By the end of the 1970s, driven by the need for density, it was clear that MOS technology would be the vehicle for growth in the digital VLSI area. At this same time, several organizations were attempting analog circuit designs using MOS [1–4]. NMOS technology was the early technology of choice for the majority of both digital and analog MOS designs. The early 1980s saw the movement of the VLSI world toward silicon-gate CMOS, which has been the dominant technology for VLSI digital and mixed-signal designs ever since [5,6]. Processes that combine both CMOS and bipolar (BiCMOS) have proven themselves to be both a technological and market success where the primary market force has been improved speed for digital circuits (primarily in static random-access memories, SRAMs). BiCMOS has potential as well in analog design due to the enhanced performance that a bipolar transistor provides in the context of CMOS technology. This book focuses on the use of CMOS for analog and mixed-signal circuit design.

There are numerous references that develop the details of the physics of MOS device operation [7,8]. Therefore, this text covers only the aspects of this theory that are pertinent to the viewpoint of the circuit designer. The objective is to be able to appreciate the limits of the MOS circuit models and to understand the physical constraints on electrical performance.

This chapter covers various aspects of the CMOS process from a physical point of view. In order to understand CMOS technology, a brief review of the basic semiconductor fabrication processes is presented, followed by a description of the fabrication steps required to build the basic CMOS process. Next, the pn junction is presented and analyzed, followed by a description of how active and passive components compatible with the CMOS technology are built. Next, important limitations on the performance of CMOS technology including latch-up, temperature dependence, and noise are covered.

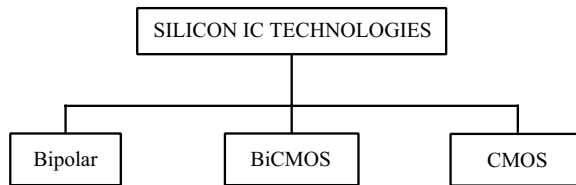


**Figure 2.0-1** Modern silicon integrated-circuit (IC) technologies.

Physical layout of an integrated circuit requires understanding of layout rules and techniques. Successful analog circuit design is tightly coupled with the physical layout. This topic is covered in Appendix B and should be studied along with Chapter 2.



## 2.1 Basic MOS Semiconductor Fabrication Processes

Semiconductor technology is based on a number of well-established process steps, which are the means for fabricating semiconductor components. In order to understand the fabrication process, it is necessary to understand these steps. The process steps described here include *oxidation*, *diffusion*, *ion implantation*, *deposition*, and *etching*. The means of defining the area of the semiconductor subject to processing is called *photolithography*. Lastly, a planarization method using *chemical-mechanical polishing* is introduced.

All processing starts with single-crystalline silicon material. There are two methods for growing such crystals [9]. Most of the material is grown by a method based on that developed by Czochralski in 1917. A second method, called the float zone technique, produces crystals of high purity and is often used for power devices. The crystals are normally grown in either a  $\langle 100 \rangle$  or  $\langle 111 \rangle$  crystal orientation. The resulting crystals are cylindrical and have a diameter of 75–300 mm and a length of 1 m. The cylindrical crystals are sliced into wafers that are approximately 0.5 mm to 0.75 mm thick for wafers of size 100 mm to 300 mm, respectively [10]. This thickness is determined primarily by the physical strength requirements. When the crystals are grown, they are doped with either an n-type or a p-type impurity to form an n- or a p-substrate. The substrate is the starting material in wafer form for the fabrication process. The doping level of most substrates is approximately  $10^{15}$  impurity atoms/cm<sup>3</sup>, which roughly corresponds to a resistivity of 3–5  $\Omega$ -cm for an n-substrate and 14–16  $\Omega$ -cm for a p-substrate [11].

An alternative to starting with a lightly doped silicon wafer is to use a heavily doped wafer that has a lightly doped epitaxial (*epi*) on top of it where subsequent devices are formed. Although epi wafers are more expensive, they can provide some benefits by reducing sensitivity to latch-up (discussed later) and reducing interference between analog and digital circuits on mixed-signal integrated circuits.

The six basic processing steps that are applied to the doped silicon wafer to fabricate semiconductor components (oxidation, diffusion, ion implantation, deposition, etching, and chemical mechanical polishing) will be described in the following paragraphs.

### Oxidation

The first basic processing step is oxide growth or oxidation [12]. Oxidation is the process by which a layer of silicon dioxide ( $\text{SiO}_2$ ) is formed on the surface of the silicon wafer. The oxide grows both into as well as on the silicon surface, as indicated in Fig. 2.1-1. Typically about 56% of the oxide thickness is above the original surface while about 44% is below the

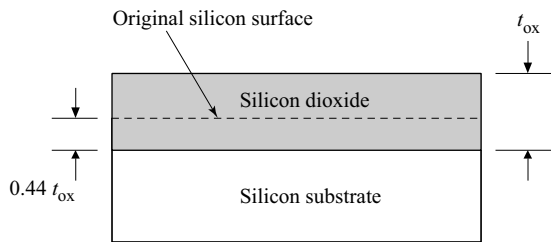


**Figure 2.1-1** Silicon dioxide growth at the surface of a silicon wafer.

original surface. The oxide thickness, designated  $t_{ox}$ , can be grown using either dry or wet techniques, with the former achieving lower defect densities. Typically oxide thickness varies from less than 50 Å for gate oxides to more than 10,000 Å for field oxides. Oxidation takes place at temperatures ranging from 700 °C to 1100 °C, with the resulting oxide thickness being proportional to the temperature at which it is grown (for a fixed amount of time).

## Diffusion

The second basic processing step is diffusion [13]. Diffusion in semiconductor material is the movement of impurity atoms at the surface of the material into the bulk of the material. Diffusion takes place at temperatures in the range of 800–1400 °C in the same way as a gas diffuses in air. The concentration profile of the impurity in the semiconductor is a function of the concentration of the impurity at the surface and the time in which the semiconductor is placed in a high-temperature environment. There are two basic types of diffusion mechanisms, which are distinguished by the concentration of the impurity at the surface of the semiconductor. One type of diffusion assumes that there is an infinite source of impurities at the surface ( $N_0 \text{ cm}^{-3}$ ) during the entire time the impurity is allowed to diffuse. The impurity profile for an infinite-source impurity as a function of diffusion time is given in Fig. 2.1-2(a). The second type of diffusion assumes that there is a finite source of impurities at the surface of the material initially. At  $t = 0$  this value is given by  $N_0$ . However, as time increases, the impurity concentration at the surface decreases as shown in Fig. 2.1-2(b). In both cases,  $N_B$  is the prediffusion impurity concentration of the semiconductor.

The infinite-source and finite-source diffusions are typical of predeposition and drive-in diffusions, respectively. The object of a predeposition diffusion is to place a large concentration of impurities near the surface of the material. There is a maximum impurity concentration that can be diffused into silicon depending on the type of impurity. This maximum concentration is due to the solid solubility limit, which is in the range of  $5 \times 10^{20}$  to  $2 \times 10^{21}$  atoms/cm<sup>3</sup>. The drive-in diffusion follows the deposition diffusion and is used to drive the impurities deeper into the semiconductor. The crossover between the prediffusion impurity level and the diffused impurities of the opposite type defines the semiconductor junction. This junction is between a p-type and an n-type material and is simply called a *pn junction*. The distance between the surface of the semiconductor and the junction is called the *junction depth*. Typical junction depths for diffusion can range from 0.1 μm for predeposition-type diffusions to greater than 10 μm for drive-in-type diffusions.

## Ion Implantation

The next basic processing step is ion implantation and is widely used in the fabrication of MOS components [14,15]. Ion implantation is the process by which ions of a particular dopant (impurity) are accelerated by an electric field to a high velocity and physically lodge

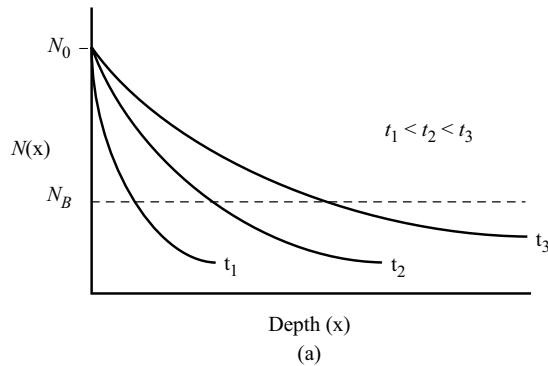
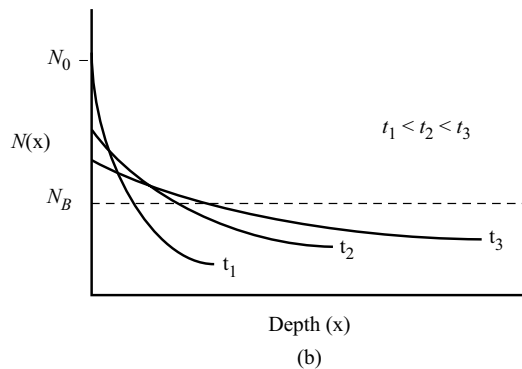


**Figure 2.1-2** Diffusion profiles as a function of time for (a) an infinite source of impurities at the surface, and (b) a finite source of impurities at the surface.



within the semiconductor material. The average depth of penetration varies from 0.1 to 0.6  $\mu\text{m}$  depending on the velocity and angle at which the ions strike the silicon wafer. The path of each ion depends on the collisions it experiences. Therefore, ions are typically implanted off-axis from the wafer so that they will experience collisions with lattice atoms, thus avoiding undesirable *channeling* of ions deep into the silicon. An alternative method to address channeling is to implant through silicon dioxide, which randomizes the implant direction before the ions enter the silicon. The ion-implantation process causes damage to the semiconductor crystal lattice, leaving many of the implanted ions electrically inactive. This damage can be repaired by an annealing process in which the temperature of the semiconductor after implantation is raised to around 800  $^{\circ}\text{C}$  to allow the ions to move to electrically active locations in the semiconductor crystal lattice.

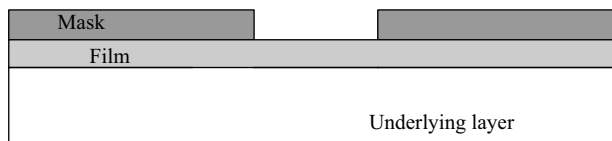
Ion implantation can be used in place of diffusion since in both cases the objective is to insert impurities into the semiconductor material. Ion implantation has several advantages over thermal diffusion. One advantage is the accurate control of doping—to within  $\pm 5\%$ . Reproducibility is very good, making it possible to adjust the thresholds of MOS devices or to create precise resistors. A second advantage is that ion implantation is a room-temperature process, although annealing at higher temperatures is required to remove the crystal damage. A third advantage is that it is possible to implant through a thin layer. Consequently, the material to be implanted does not have to be exposed to contaminants during and after the implantation process. Unlike ion implantation, diffusion requires that the surface be free of silicon dioxide or silicon nitride layers. Finally, ion implantation allows control over the profile of the implanted impurities. For example, a concentration peak can be placed below the surface of the silicon if desired.

## Deposition

The fourth basic semiconductor process is deposition. Deposition is the means by which films of various materials may be deposited on the silicon wafer. These films may be deposited using several techniques, including deposition by evaporation [16], sputtering [17], and chemical-vapor deposition (CVD) [18,19]. In evaporation deposition, a solid material is placed in a vacuum and heated until it evaporates. The evaporant molecules strike the cooler wafer and condense into a solid film on the wafer surface. Thickness of the deposited material is determined by the temperature and the amount of time evaporation is allowed to take place (a thickness of 1  $\mu\text{m}$  is typical). The sputtering technique uses positive ions to bombard the cathode, which is coated with the material to be deposited. The bombarded or target material is dislodged by direct momentum transfer and deposited on wafers, which are placed on the anode. The types of sputtering systems used for depositions in integrated circuits include dc, radio frequency (RF), or magnetron (magnetic field). Sputtering is usually done in a vacuum. Chemical vapor deposition uses a process in which a film is deposited by a chemical reaction or pyrolytic decomposition in the gas phase, which occurs in the vicinity of the silicon wafer. This deposition process is generally used to deposit polysilicon, silicon dioxide ( $\text{SiO}_2$  or simply *oxide*), or silicon nitride ( $\text{Si}_3\text{N}_4$  or simply *nitride*). While the chemical vapor deposition is usually performed at atmospheric pressure, it can also be done at low pressures where the diffusivity increases significantly. This technique is called low-pressure chemical-vapor deposition (LPCVD).

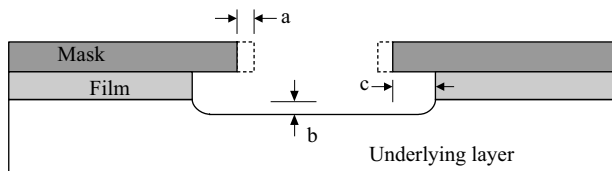
## Etching

Etching is the process of removing exposed (unprotected) material. The means by which some material is exposed and some is not will be considered later under the topic of photolithography. For the moment, we will assume that the situation illustrated in Fig. 2.1-3(a) exists. Here we see a top layer called a film and an underlying layer. A protective layer, called a mask,\* covers the film except in the area that is to be etched. The objective of etching is to remove just the section of the exposed film. To achieve this, the etching process must have two important properties: selectivity and anisotropy. Selectivity is the characteristic of the etch whereby only the desired layer is etched with no effect on either the protective layer



(a)

**Figure 2.1-3** (a) Portion of the top layer ready for etching. (b) Result of etching indicating horizontal etching and etching of underlying layer.



(b)

\*A distinction is made between a deposited masking layer referred to as a “mask” and the photographic plate used in exposing the photoresist, which is called a “photomask.”

(masking layer) or the underlying layer. Selectivity can be quantified as the ratio of the desired layer etch rate to the undesired layer etch rate as given below.

$$S_{A-B} = \frac{\text{Desired layer etch rate (A)}}{\text{Undesired layer etch rate (B)}} \quad (2.1-1)$$

Anisotropy is the property of the etch to manifest itself in one direction; that is, a perfectly anisotropic etchant will etch in one direction only. The degree of anisotropy can be quantified by the relation given below.

$$A = 1 - \frac{\text{Lateral etch rate}}{\text{Vertical etch rate}} \quad (2.1-2)$$

Reality is such that neither perfect selectivity nor perfect anisotropy can be achieved in practice, resulting in undercutting effects and partial removal of the underlying layer as illustrated in Fig. 2.1-3(b). As illustrated, the lack of selectivity with respect to the mask is given by dimension “a.” Lack of selectivity with respect to the underlying layer is given by dimension “b.” Dimension “c” shows the degree of anisotropy. There are preferential etching techniques that achieve high degrees of anisotropy and thus minimize undercutting effects, as well as maintain high selectivity. Materials that are normally etched include polysilicon, silicon dioxide, silicon nitride, and aluminum.

There are two basic types of etching techniques. *Wet etching* uses chemicals to remove the material to be etched. Hydrofluoric acid (HF) is used to etch silicon dioxide; phosphoric acid ( $\text{H}_3\text{PO}_4$ ) is used to remove silicon nitride; nitric acid, acetic acid, or hydrofluoric acid is used to remove polysilicon; potassium hydroxide is used to etch silicon; and a phosphoric acid mixture is used to remove metal. The wet-etching technique is strongly dependent on time and temperature, and care must be taken with the acids used in wet etching as they represent a potential hazard. *Dry etching* or *plasma etching* uses ionized gases that are rendered chemically active by an RF-generated plasma. This process requires significant characterization to optimize pressure, gas flow rate, gas mixture, and RF power. Dry etching is very similar to sputtering and in fact the same equipment can be used. Reactive ion etching (RIE) induces plasma etching accompanied by ionic bombardment. Dry etching is used for submicron technologies since it achieves anisotropic profiles (no undercutting).

## Chemical Mechanical Polishing

Performing the photolithographic steps on nonplanar surfaces can be challenging. Moreover, maintaining uniform thickness of metal deposited over a surface with abrupt transitions is difficult. The solution is to planarize the surface of the wafer prior to each photolithographic or deposition step. One means for achieving planarization is to use an oxide appropriately doped so that it will reflow under elevated temperature. Another, and far superior, means for planarization is to literally polish the wafer when a flat surface is needed. Modern processes use *chemical mechanical polishing (CMP)* to planarize wafers [20].

## Photolithography

Each of the basic semiconductor fabrication processes discussed thus far is applied only to selected parts of the silicon wafer with the exception of oxidation, deposition, and CMP. The selection of these parts is accomplished by a process called photolithography [12,21,22].

Photolithography refers to the complete process of transferring an image from a *photomask* or computer database to a wafer. The basic components of photolithography are the photoresist material and the photomask used to expose some areas of the photoresist to ultraviolet (UV) light while shielding the remainder.

Photoresist is an organic polymer whose characteristics can be altered when exposed to ultraviolet light. Photoresist is classified into positive and negative photoresist. *Positive photoresist* is used to create a mask where patterns exist (where the photomask is opaque to UV light). *Negative photoresist* creates a mask where patterns do not exist (where the photomask is transparent to UV light). The photolithographic process involves depositing photoresist on the wafer and conditioning it with time and temperature. Portions of the photoresist are exposed to UV light. After exposure, the exposed photoresist is hardened through a *developing* process while the unexposed areas are removed. The hardened photoresist protects selected areas from plasma or acids used in the etching process. When its protective function is complete, the photoresist is removed with solvents or plasma ashing, leaving underlying layers unharmed. This process must be repeated for each layer of the integrated circuit. Figure 2.1-4 shows, by way of example, the basic photolithographic steps in defining a polysilicon geometry using positive photoresist.

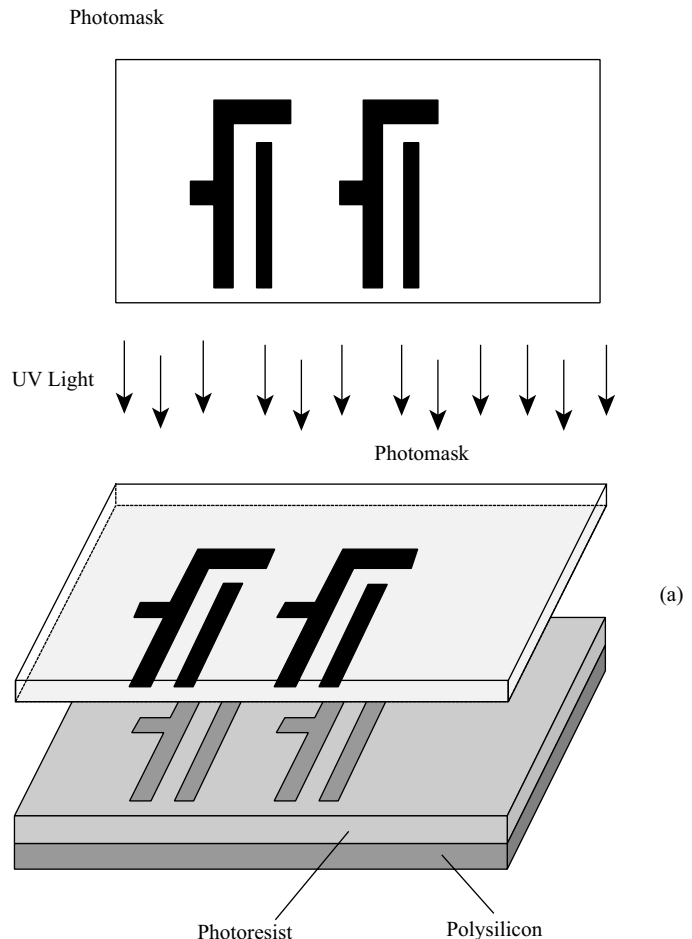


**Figure 2.1-4** Basic photolithographic steps to define a polysilicon geometry: (a) expose, (b) develop, (c) etch, (d) remove photoresist.

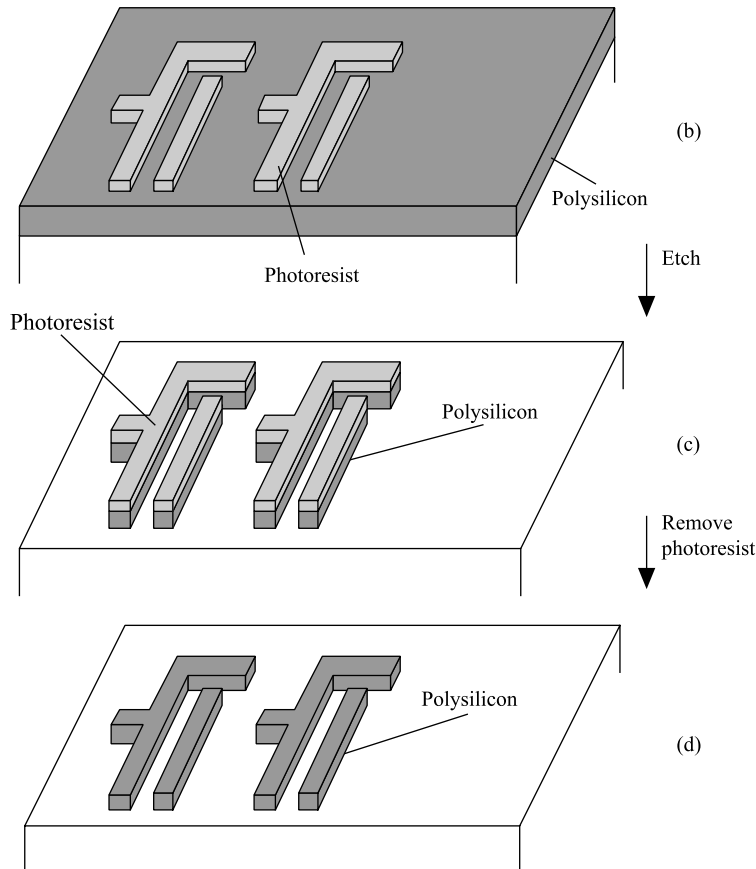


**Figure 2.1-4** (Continued)

The process of exposing selective areas of a wafer to light through a photomask is called *printing*. There are three basic types of printing systems used. They are listed below:

- Contact printing
- Proximity printing
- Projection printing

The simplest and most accurate method is *contact printing*. This method uses a glass plate a little larger than the size of the actual wafer with the image of the desired pattern on the side of the glass that comes in physical contact with the wafer. This glass plate is called a *photomask*. The system achieves high resolution, high throughput, and low cost. Unfortunately, because of the direct contact, the photomask wears out and has to be replaced after 10–25 exposures. This method also introduces impurities and defects, because of the physical contact. For these reasons, contact printing is not used in modern VLSI.

A second exposure system is called *proximity printing*. In this system, the photomask and wafer are placed very close to one another but not in intimate contact. As the gap between the photomask and the wafer increases, resolution decreases. In general, this method of patterning is not useful where minimum feature size is below 2  $\mu\text{m}$ . Therefore, proximity printing is not used in present-day VLSI.

The projection printing method separates the wafer from the photomask by a relatively large distance. Lenses or mirrors are used to focus the photomask image on the surface of the wafer. There are two approaches used for projection printing: *scanning* and *step-and-repeat*. The scanning method passes light through the scaled photomask (typically  $5\times$ ), which follows a complex optical path reflecting off multiple mirrors imaging the wafer with an arc of illumination optimized for minimum distortion. The photomask and wafer scan the illuminated arc. Minimum feature size for this method is approximately  $0.25\text{ }\mu\text{m}$ . The projection printing system most used today is step-and-repeat.

The latest advancements in photolithography, with the goal of increasing resolution, replace the air gap between the lens and the wafer surface with a liquid medium (highly purified water).

Electron beam exposure systems are often used to generate the photomasks for projection printing systems because of their high resolution (less than  $1\text{ }\mu\text{m}$ ). However, the electron beam can be used to directly pattern photoresist without using a photomask. The advantages of using the electron beam as an exposure system are accuracy and the ability to make software changes. The disadvantages are high cost and low throughput.

## Twin-Well CMOS Fabrication Steps

It is important for a circuit designer to understand some of the basic steps involved in fabricating a CMOS circuit. The fabrication steps of a generic submicron CMOS twin-well silicon-gate process will be described.

Fabrication begins with a heavily doped  $p^+$  silicon wafer with a lightly doped  $p^-$  epitaxial layer (epi layer). A thin silicon-dioxide region is grown on the surface of the  $p^-$  epi layer. Subsequent to this, the regions where n-wells are to exist are defined in a masking step by depositing a photoresist material on top of the oxide. After exposing and developing the photoresist, n-type impurities are implanted into the wafer as illustrated in Fig. 2.1-5(a). Next, photoresist is removed and p-wells are created using a p-well mask and the photolithographic process to define appropriate areas to be implanted. Photoresist is removed and a high-temperature oxidation/drive-in step is performed, causing the implanted ions to diffuse into the substrate. This is followed by oxide removal and subsequent growth of a thin pad oxide layer. (The purpose of the pad oxide is to protect the substrate from stress due to the difference in the thermal expansion of silicon and silicon nitride.) A layer of silicon nitride is deposited over the entire wafer as illustrated in Fig. 2.1-5(b). Silicon nitride is used as a *stop layer* during the CMP step. Photoresist is deposited, patterned, and developed as before, and the silicon nitride is removed from the areas where it has been patterned. The silicon nitride and photoresist remain in the areas where transistors will reside. The regions where silicon nitride remains are called *active area* (AA). In order to provide isolation between transistors, an etch is performed that cuts trenches between the remaining nitride (AA) regions as illustrated in Fig. 2.1-5(c). This is followed by a linear oxide growth and a conformal oxide deposition [Fig. 2.1-5(d)]. A CMP step removes the oxide, forming a planar surface at the top of the nitride as shown in Fig. 2.1-5(e). At this stage, the active areas are separated by *shallow trench isolation* (STI) regions.

Next, using the appropriate p-well or n-well mask and photolithographic steps, p-type and n-type field implants are performed. These will ensure that parasitic transistors are not formed in the field regions. Nitride is removed. A thin gate oxide is grown followed by a polysilicon deposition step [Fig. 2.1-5(f)]. Polysilicon is then patterned and etched, leaving only what is required to make transistor gates and interconnect lines.

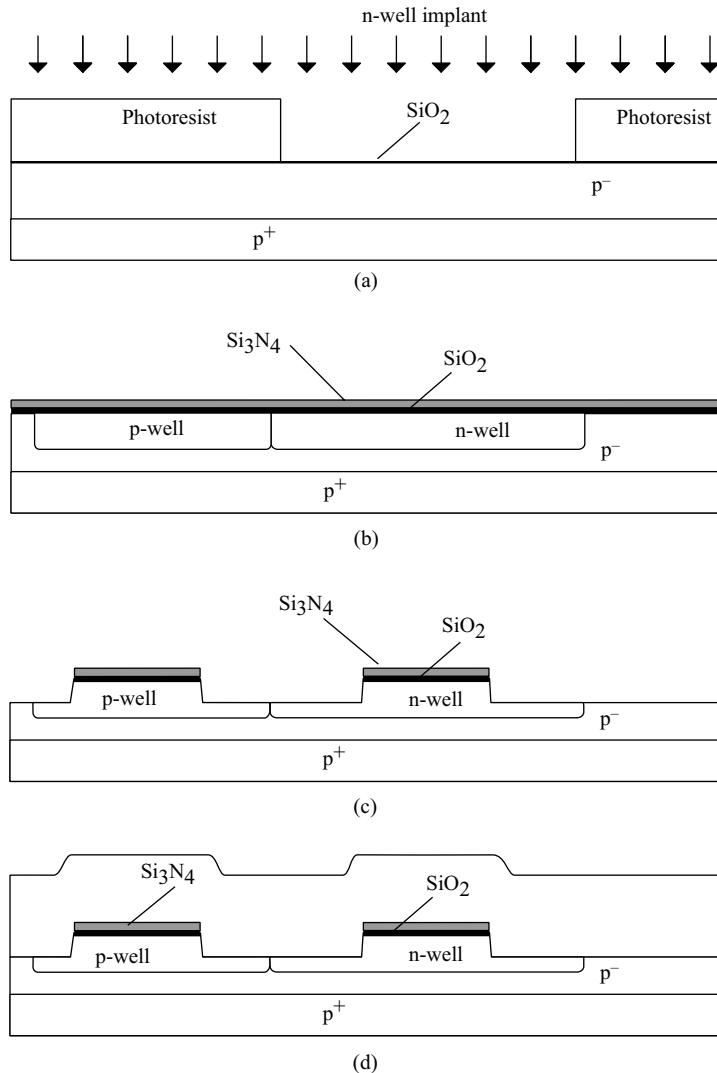


**Figure 2.1-5** The major CMOS process steps.

At this point, the drain and source areas have not been diffused into the substrate. Modern processes employ lightly doped drain/source (LDD) diffusions to minimize impact ionization. The LDD structure is created by first performing a lightly doped implant (LDD), which self-aligns with the polysilicon gate and the STI. Figure 2.1-5(g) shows the results after both p-channel and n-channel transistors have received their LDD implant of the appropriate type ( $p^-$  and  $n^-$ , respectively). A nitride spacer is formed on the sides of the polysilicon gates by depositing a thin nitride layer followed by an anisotropic etch [Fig. 2.1-5(h)]. To make  $n^+$  sources and drains, photoresist is applied and patterned everywhere n-channel transistors are required;  $n^+$  is also required where metal connections are to be made to  $n^-$  material such as the n-well. After developing, the  $n^+$  areas are implanted as illustrated in Fig. 2.1-5(i). The photoresist acts as a barrier to the implant, as does the nitride spacer. As a

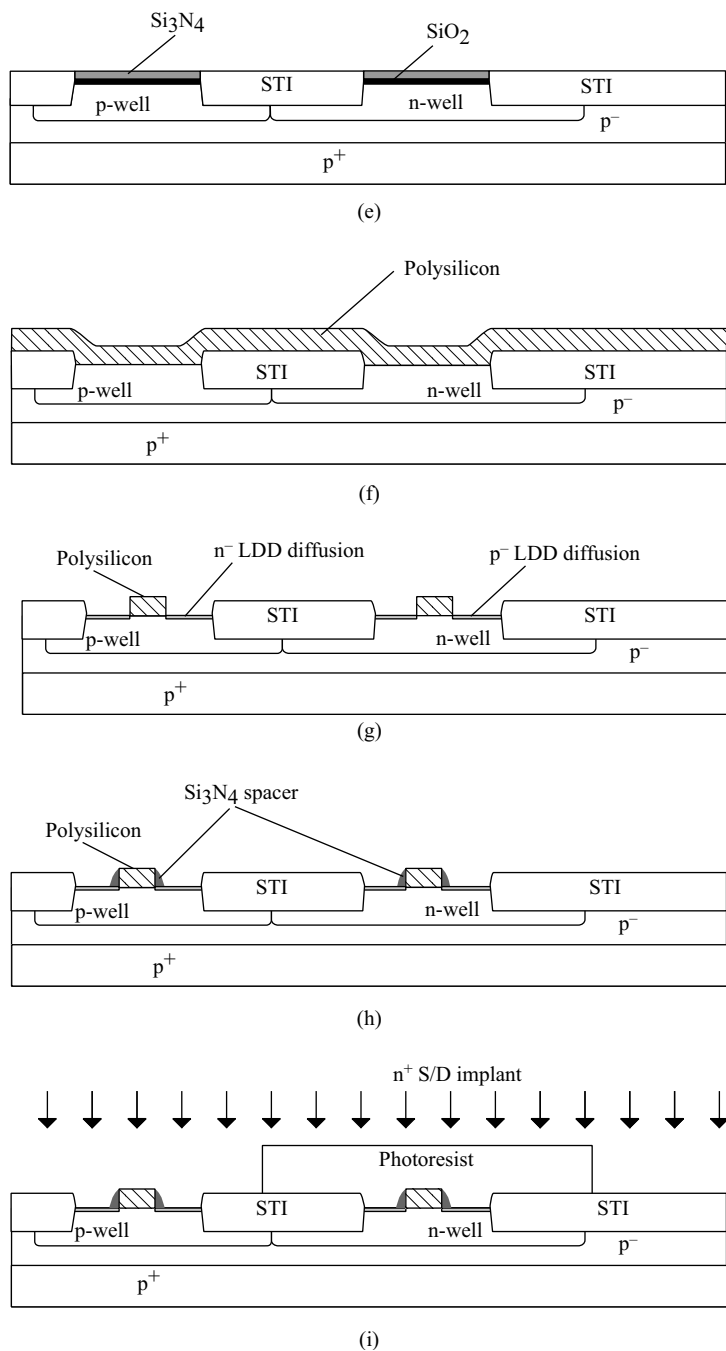
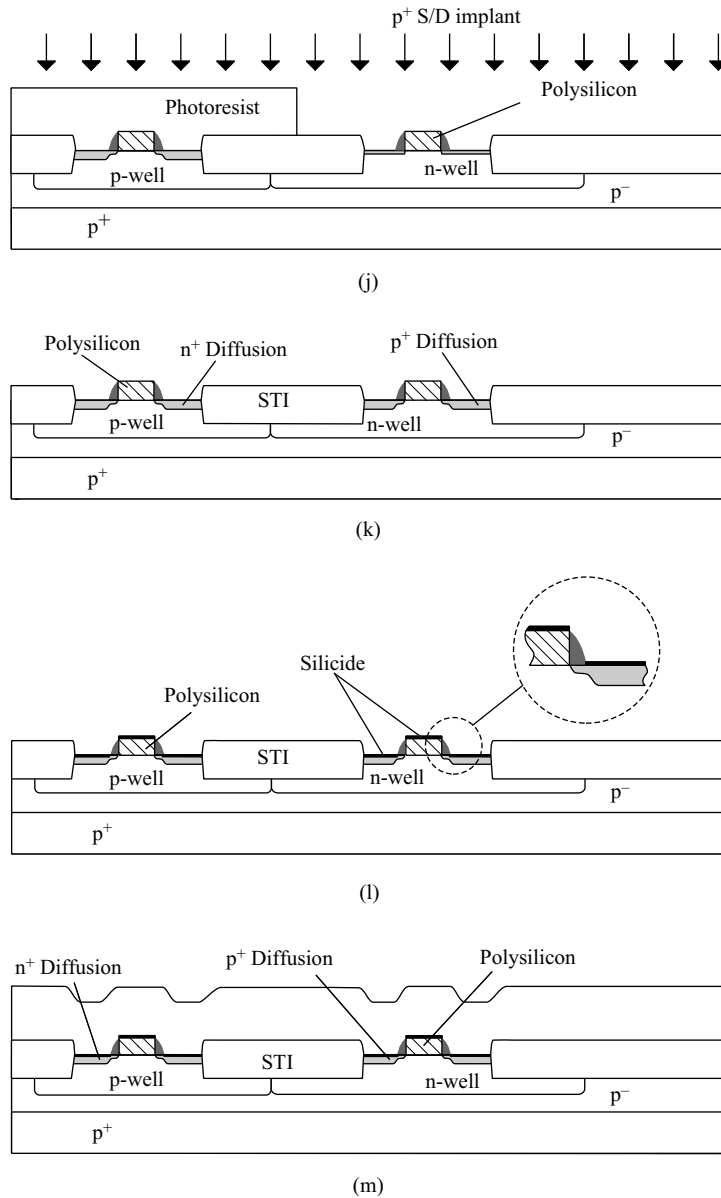


Figure 2.1-5 (Continued)

**Figure 2.1-5** (Continued)

result, the  $n^+$  regions that result are properly aligned with the spacer nitride. Similarly, the  $p$ -channel transistors receive their S/D implant after appropriate masking [Fig. 2.1-5(j)]. Annealing is performed in order to activate the implanted ions. Figure 2.1-5(k) shows the result after this step.

In order to improve the conductivity of S/D, polysilicon interconnect, and contacts, titanium (Ti) is sputtered over the entire wafer. Annealing is performed, which forms titanium silicide ( $TiSi_2$ ) at all silicon interfaces [22]. Silicide provides a much lower resistance

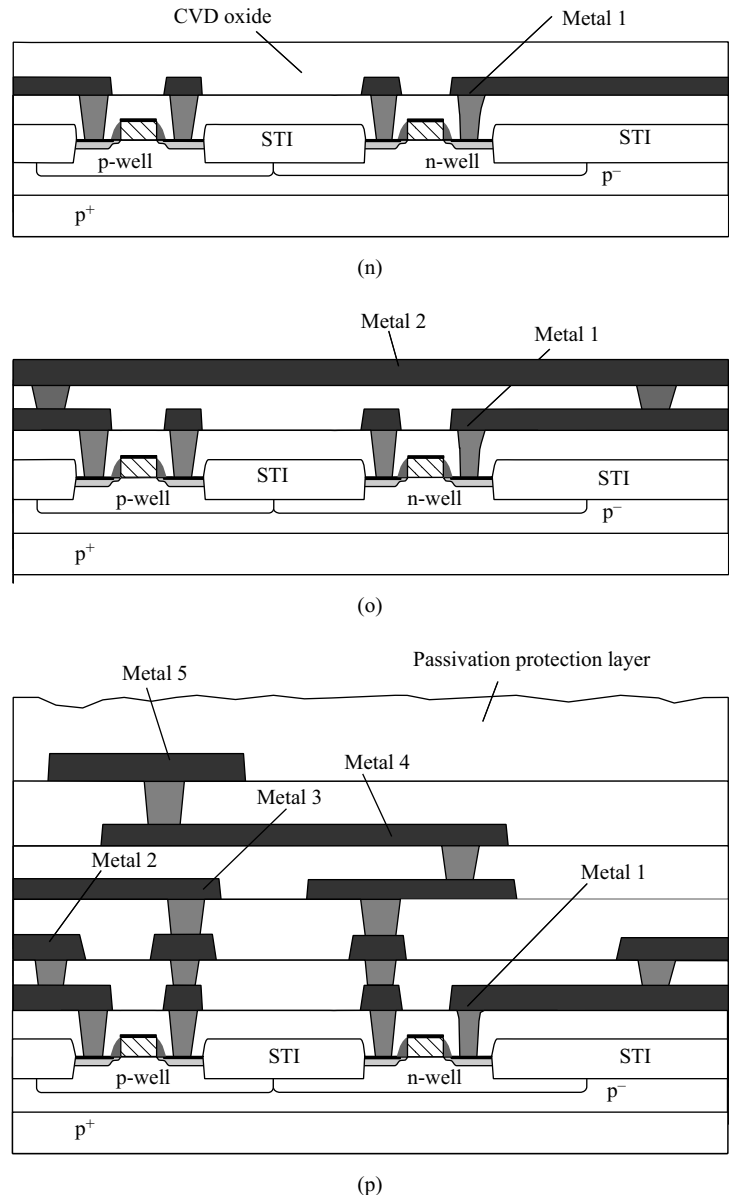


Figure 2.1-5 (Continued)

than diffusion or polysilicon. Silicide is not formed on the nitride spacer. The Ti not interfacing to silicon is removed using a chemical etch. Figure 2.1-5(l) illustrates the titanium silicide remaining on S/D and polysilicon gates. Because the titanium silicide does not react with the nitride spacer (and is later removed), the resulting silicide is self-aligned with polysilicon and diffusion. Thus, the technique for creating the silicide is called a *salicide* process (meaning a *self-aligned-silicide*). To avoid confusion, the term “silicide” will be used throughout the text. A new, thick oxide layer is deposited over the entire wafer as

illustrated in Fig. 2.1-5(m), followed by a CMP step to planarize the surface. Contacts are formed by first defining their location using the photolithographic process applied in earlier steps. Next, the oxide areas where contacts are to be made are etched down to the surface of the silicide. The remaining photoresist is removed. A thin titanium/titanium-nitride (Ti/TiN) layer is deposited followed by a thick layer of tungsten (W). The Ti/TiN provides adhesion and a diffusion barrier and the tungsten readily fills the holes formed by the contact etch. Another CMP step is performed, this time removing the tungsten so that only the *tungsten plugs* remain where contact holes were formed. Aluminum is deposited to form the Metal 1 interconnect layer. It is defined photolithographically and subsequently etched removing all unnecessary metal. To prepare for a Metal 2, another interlayer dielectric is deposited [Fig. 2.1-5(n)]. Again, CMP is performed to planarize the surface of the oxide. Intermetal connections (vias) are defined through the photolithographic patterning and etch. A thin layer of Ti/TiN is deposited followed by a thick layer of tungsten. A CMP step polishes away the tungsten above the surface of the oxide, leaving plugs. Metal 2 aluminum is deposited, patterned, and etched [Fig. 2.1-5(o)]. The steps of oxide deposition, CMP, tungsten plug formation, and aluminum metallization are repeated for as many layers of metal as desired.

In order to protect the wafer from chemical intrusion or scratching, a passivation layer of oxide or nitride is applied, covering the entire wafer. Pad regions are then defined (areas where wires will be bonded between the integrated circuit and the package containing the circuit) and the passivation layer removed only in these areas. Figure 2.1-5(p) shows a cross section of the final circuit.

To illustrate the process steps in sufficient detail, true relative dimensions are not given (i.e., the side-view drawings are not to scale). It is valuable to gain an appreciation of actual scale; thus, Fig. 2.1-6 is provided to illustrate relative dimensions.

Thus far, the basic twin-well CMOS process has been described. There are a variety of enhancements that can be applied to this process to improve circuit performance.

Prior to forming contact plugs, titanium was deposited and annealed, and titanium silicide was formed at all silicon interfaces, including polysilicon. The polysilicide forms a much lower resistance than the polysilicon alone. This is beneficial where low-resistance

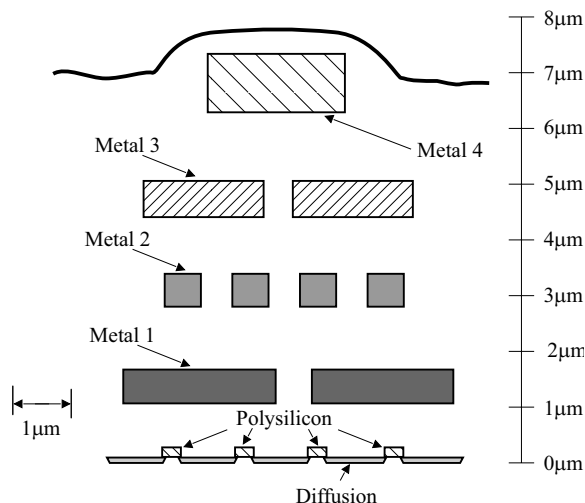


**Figure 2.1-6** Side view of a CMOS integrated circuit.

interconnect is desired. However, analog designers often want to use high-resistance components and polysilicon is an excellent choice. In order to use the sheet resistance of the polysilicon, the titanium must be blocked from being depositing onto the polysilicon. This is achieved with an additional mask—silicide block—which does precisely what its name implies. In some processes, the polysilicon is not doped. Intrinsic polysilicon has a very high resistance. To be useful as an analog component, a mask is used to selectively dope polysilicon resistors to achieve the desired resistivity.

Invariably, analog circuits require high-performance capacitors. There are at least three common ways to achieve useful capacitors: (1) poly-oxide-poly, (2) metal-oxide-metal (MOM), and (3) metal-insulator-metal (MiM). Poly-oxide-poly (or simply poly-poly) capacitors are fabricated by an additional step subsequent to the gate-poly step so that two polysilicon layers lie on top of one another with an intervening thin oxide. MOM capacitors are formed without any additional processing steps as they are simply capacitors made from the various metal layers that already exist (suitably oriented in the layout). MiM capacitors require the formation of an additional metal layer. It is usually created just prior to the final interconnect layer and sits atop a thin oxide on the second-to-last metal interconnect layer. These capacitor structures are described in Section 2.4.

When designing analog circuits integrated with complex digital circuitry, one must be careful to avoid interference due to the digital logic—digital noise. Such interference will reduce the performance of sensitive analog signal paths. There are numerous circuit design techniques that can mitigate digital noise. In addition, there are structural methods to isolate the analog circuits from the digital ones. One such method is to use *deep n-well* (DNW). Figure 2.1-7 illustrates an example of a deep n-well. The well is driven below the depth of the p-well and the standard n-well. Electrical connection to the DNW is made via n-well and  $n^+$  diffusions. Isolation and interference reduction is achieved by connecting the DNW to a quiet supply source (e.g., a quiet  $V_{DD}$ ).

As geometries shrink, the breakdown voltages for transistors also go down. Thus, core supply voltages continue to go down (e.g., 65 nm core voltage is typically around 1.0 volts). Even though core voltages are shrinking, I/O voltages are not (at the same rate); thus, interfacing between the core of an integrated circuit and the I/O presents challenges. One solution is to provide transistors with higher breakdown voltages alongside the standard devices. The primary means for accomplishing this is to offer multiple gate oxides. The smallest gate oxide is used for core transistors while larger gate oxides are used for I/O devices. Clever designers often find use of the I/O devices with the core of the integrated circuit to address special needs (e.g., low-leakage circuitry).

The process described in this section applies generally to 0.25  $\mu\text{m}$  and below. Technologies at 0.35  $\mu\text{m}$  and above typically use the LOCOS method for isolation. An example of such a process (n-well LOCOS) is illustrated in Fig. 2.1-8.

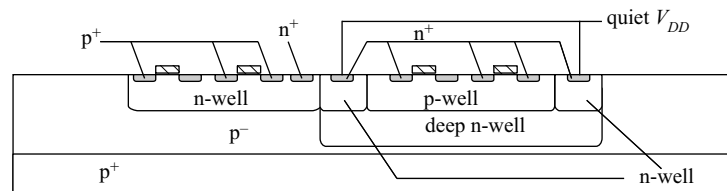


**Figure 2.1-7** Deep n-well improves isolation.

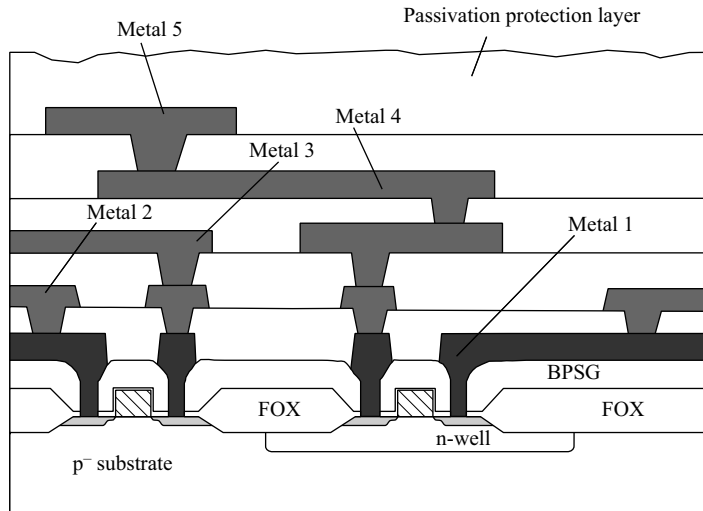


**Figure 2.1-8** Cross section of a LOCOS CMOS process.

## 2.2 The pn Junction

The pn junction plays an important role in all semiconductor devices. The objective of this section is to develop the concepts of the pn junction that will be useful to us later in our study. These include the depletion-region width, the depletion capacitance, reverse-bias or breakdown voltage, and the diode equation. Further information can be found in the references [23,24].

Figure 2.2-1(a) shows the physical model of a pn junction. In this model it is assumed that the impurity concentration changes abruptly from  $N_D$  donors in the n-type semiconductor to  $N_A$  acceptors in the p-type semiconductor. This situation is called a step junction and is illustrated in Fig. 2.2-1(b). The distance  $x$  is measured to the right from the metallurgical junction at  $x = 0$ . When two different types of semiconductor materials are formed in this manner, the free carriers in each type move across the junction by the principle of diffusion. As these free carriers cross the junction, they leave behind fixed atoms that have a charge opposite to the carrier. For example, as the electrons near the junction of the n-type material diffuse across the junction they leave fixed donor atoms of opposite charge (+) near the junction of the n-type material. This is represented in Fig. 2.2-1(c) by the rectangle with a height of  $qN_D$ . Similarly, the holes that diffuse across the junction from the p-type material to the n-type material leave behind fixed acceptor atoms that are negatively charged. The electrons and holes that diffuse across the junction quickly recombine with the free majority carriers across the junction. As positive and negative fixed charges are uncovered near the junction by the diffusion of the free carriers, an electric field develops that creates an opposing carrier movement. When the current due to the free carrier diffusion equals the current caused by the electric field, the pn junction reaches equilibrium. In equilibrium, both  $v_D$  and  $i_D$  of Fig. 2.2-1(a) are zero.

The distance over which the donor atoms have a positive charge (because they have lost their free electron) is designated as  $x_n$  in Fig. 2.2-1(c). Similarly, the distance over which the acceptor atoms have a negative charge (because they have lost their free hole) is  $x_p$ . In this

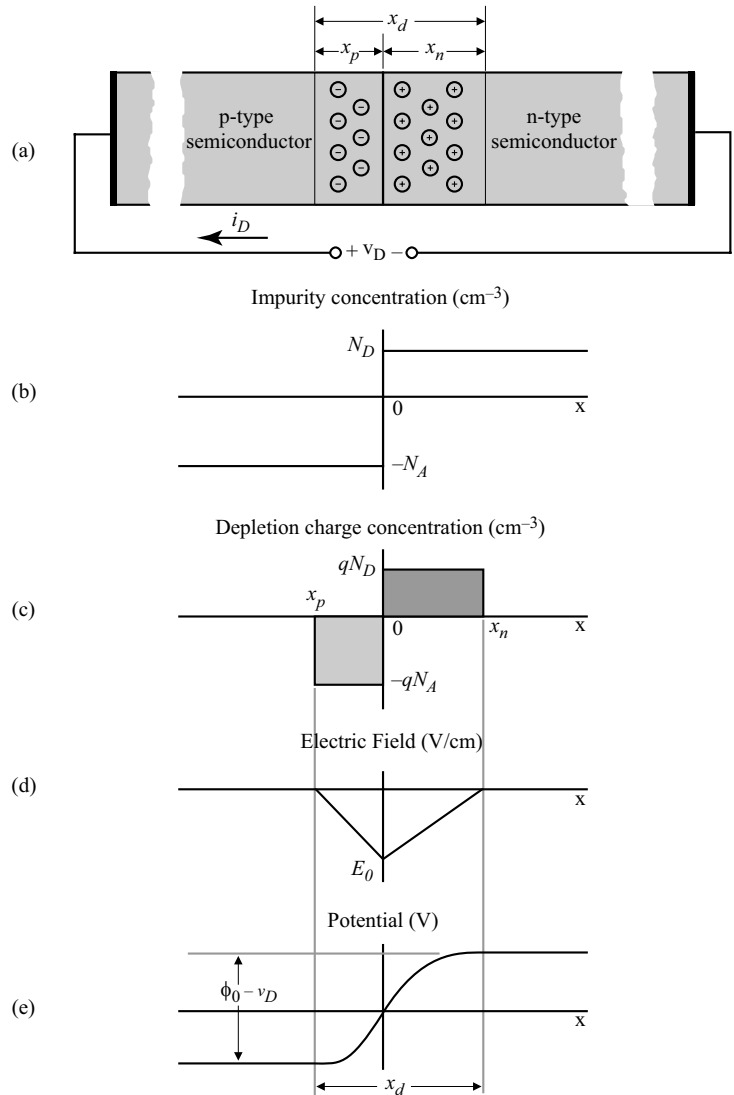


**Figure 2.2-1** Junction; (a) physical structure, (b) impurity concentration, (c) depletion charge concentration, (d) electric field, (e) electrostatic potential.

diagram,  $x_p$  is a negative number. The *depletion region* is defined as the region about the metallurgical junction that is depleted of free carriers. The depletion region is defined as

$$x_d = x_n - x_p \quad (2.2-1)$$

Note that  $x_p < 0$ .

Due to electrical neutrality, the charge on either side of the junction must be equal. Thus,

$$qN_Dx_n = -qN_Ax_p \quad (2.2-2)$$

where  $q$  is the charge of an electron ( $1.60 \times 10^{-19}$  C). The electric field distribution in the depletion region can be calculated using the point form of Gauss's law.

$$\frac{dE(x)}{dx} = \frac{qN}{\epsilon_{\text{Si}}} \quad (2.2-3)$$

By integrating either side of the junction, the maximum electric field that occurs at the junction,  $E_0$ , can be found. This is illustrated in Fig. 2.2-1(d). Therefore, the expression for  $E_0$  is

$$E_0 = \int_0^{E_0} dE = \int_{x_p}^0 \frac{-qN_A}{\epsilon_{\text{Si}}} dx = \frac{qN_A x_p}{\epsilon_{\text{Si}}} = \frac{-qN_D x_n}{\epsilon_{\text{Si}}} \quad (2.2-4)$$

where  $\epsilon_{\text{Si}}$  is the dielectric constant of silicon and is  $11.7\epsilon_0$  ( $\epsilon_0$  is  $8.85 \times 10^{-14}$  F/cm).

The voltage drop across the depletion region is shown in Fig. 2.2-1(e). The voltage is found by integrating the negative electric field, resulting in

$$\phi_0 - v_D = \frac{-E_0(x_n - x_p)}{2} \quad (2.2-5)$$

where  $v_D$  is an applied external voltage and  $\phi_0$  is called the *barrier potential* and is given as

$$\phi_0 = \frac{kT}{q} \ln \left( \frac{N_A N_D}{n_i^2} \right) = V_t \ln \left( \frac{N_A N_D}{n_i^2} \right) \quad (2.2-6)$$

Here,  $k$  is Boltzmann's constant ( $1.38 \times 10^{-23}$  J/K) and  $n_i$  is the intrinsic concentration of silicon, which is  $1.45 \times 10^{10}/\text{cm}^3$  at 300 K. At room temperature, the value of  $V_t$  is 25.9 mV. It is important to note that the notation for  $kT/q$  is  $V_t$  rather than the conventional  $V_T$ . The reason for this is to avoid confusion with  $V_T$ , which will be used to designate the threshold voltage of the MOS transistor (see Section 2.3). Although the barrier voltage exists with  $v_D = 0$ , it is not available externally at the terminals of the diode. When metal leads are attached to the ends of the diode a metal–semiconductor junction is formed. The barrier potentials of the metal–semiconductor contacts are exactly equal to  $\phi_0$  so that the open circuit voltage of the diode is zero.

Equations (2.2-2), (2.2-4), and (2.2-5) can be solved simultaneously to find the width of the depletion region in the n-type and p-type semiconductor. These widths are found as

$$x_n = \left[ \frac{2\epsilon_{\text{Si}}(\phi_0 - v_D)N_A}{qN_D(N_A + N_D)} \right]^{1/2} \quad (2.2-7)$$

and

$$x_p = - \left[ \frac{2\epsilon_{\text{Si}}(\phi_0 - v_D)N_D}{qN_A(N_A + N_D)} \right]^{1/2} \quad (2.2-8)$$

The width of the depletion region,  $x_d$ , is found from Eqs. (2.2-1), (2.2-7), and (2.2-8) and is

$$x_d = \left[ \frac{2\epsilon_{\text{Si}}(N_A + N_D)}{qN_A N_D} \right]^{1/2} (\phi_0 - v_D)^{1/2} \quad (2.2-9)$$

It can be seen from Eq. (2.2-9) that the depletion width for the pn junction of Fig. 2.2-1 is proportional to the square root of the difference between the barrier potential and the externally applied voltage. It can also be shown that  $x_d$  is approximately equal to  $x_n$  or  $x_p$  for  $N_A \gg N_D$  or  $N_D \gg N_A$ , respectively. Consequently, the depletion region will extend farther into the lightly doped semiconductor than it will into the heavily doped semiconductor.

It is also of interest to characterize the depletion charge  $Q_j$ , which is equal to the magnitude of the fixed charge on either side of the junction. The depletion charge can be expressed from the above relationships as

$$Q_j = |AqN_Ax_p| = AqN_Dx_n = A \left[ \frac{2\epsilon_{Si}qN_AN_D}{N_A + N_D} \right]^{1/2} (\phi_0 - v_D)^{1/2} \quad (2.2-10)$$

where  $A$  is the cross-sectional area of the pn junction.

The magnitude of the electric field at the junction  $E_0$  can be found from Eqs. (2.2-4) and (2.2-7) or (2.2-8). This quantity is expressed as

$$E_0 = \left[ \frac{2qN_AN_D}{\epsilon_{Si}(N_A + N_D)} \right]^{1/2} (\phi_0 - v_D)^{1/2} \quad (2.2-11)$$

Equations (2.2-9), (2.2-10), and (2.2-11) are key relationships in understanding the pn junction.

The depletion region of a pn junction forms a capacitance called the *depletion-layer capacitance*. It results from the dipole formed by uncovered fixed charges near the junction and will vary with the applied voltage. The depletion-layer capacitance  $C_j$  can be found from Eq. (2.2-10) using the following definition of capacitance:

$$C_j = \frac{dQ_j}{dv_D} = A \left[ \frac{\epsilon_{Si}qN_AN_D}{2(N_A + N_D)} \right]^{1/2} \frac{1}{(\phi_0 - v_D)^{1/2}} = \frac{C_{j0}}{[1 - (v_D/\phi_0)]^m} \quad (2.2-12)$$

$C_{j0}$  is the depletion-layer capacitance when  $v_D = 0$  and  $m$  is called a grading coefficient. The coefficient  $m$  is 1/2 for the case of Fig. 2.2-1, which is called a step junction. If the junction is fabricated using diffusion techniques described in Section 2.1, Fig. 2.2-1(b) will become more like the profile of Fig. 2.2-2. It can be shown for this case that  $m$  is 1/3. The range of values of the grading coefficient will fall between 1/3 and 1/2. Figure 2.2-3 shows a plot of the depletion layer capacitance for a pn junction. It is seen that when  $v_D$  is positive and approaches  $\phi_0$ , the depletion-layer capacitance approaches infinity. At this value of voltage, the assumptions made in deriving the above equations are no longer valid. In particular, the assumption that the depletion region is free of charged carriers is not true. Consequently, the actual curve bends over and  $C_j$  decreases as  $v_D$  approaches  $\phi_0$  [25].

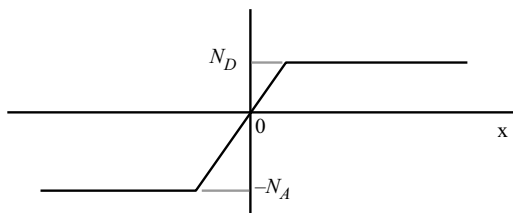


**Figure 2.2-2** Impurity concentration profile for diffused pn junction.

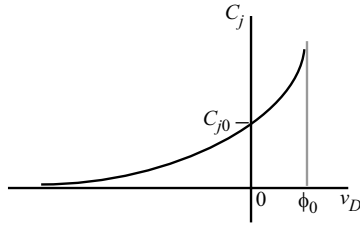


**Figure 2.2-3** Depletion capacitance as a function of externally applied junction voltage.

### Example 2.2-1

#### Characteristics of a pn Junction

Find  $x_p$ ,  $x_n$ ,  $x_d$ ,  $\phi_0$ ,  $C_{j0}$ , and  $C_j$  for an applied voltage of  $-4$  V for a pn diode with a step junction,  $N_A = 5 \times 10^{15}/\text{cm}^3$ ,  $N_D = 10^{20}/\text{cm}^3$ , and an area of  $10 \mu\text{m}$  by  $10 \mu\text{m}$ .

At room temperature, Eq. (2.2-6) gives the barrier potential as  $0.917$  V. Equations (2.2-7) and (2.2-8) give  $x_n \approx 0$  and  $x_p = 1.128 \mu\text{m}$ . Thus, the depletion width is approximately  $x_p$  or  $1.128 \mu\text{m}$ . Using these values in Eq. (2.2-12) we find that  $C_{j0}$  is  $20.3$  fF and at a voltage of  $-4$  V,  $C_j$  is  $9.18$  fF.

The voltage breakdown of a reverse-biased ( $v_D < 0$ ) pn junction is determined by the maximum electric field  $E_{\text{max}}$  that can exist across the depletion region. For silicon, this maximum electric field is approximately  $3 \times 10^5$  V/cm. If we assume that  $|v_D| > \phi_0$ , then substituting  $E_{\text{max}}$  into Eq. (2.2-11) allows us to express the maximum reverse-bias voltage or breakdown voltage ( $BV$ ) as

$$BV \cong \frac{\epsilon_{\text{Si}}(N_A + N_D)}{2qN_A N_D} E_{\text{max}}^2 \quad (2.2-13)$$

Substituting the values of Example 2.2-1 in Eq. (2.2-13) and using a value of  $3 \times 10^5$  V/cm for  $E_{\text{max}}$  gives a breakdown voltage of  $58.2$  volts. However, as the reverse-bias voltage starts to approach this value, the reverse current in the pn junction starts to increase. This increase is due to two conduction mechanisms that can take place in a reverse-biased junction between two heavily doped semiconductors. The first current mechanism is called avalanche multiplication and is caused by the high electric fields present in the pn junction; the second is called Zener breakdown. Zener breakdown is a direct disruption of valence bonds in high electric fields. However, the Zener mechanism does not require the presence of an energetic ionizing carrier. The current in most breakdown diodes will be a combination of these two current mechanisms.

If  $i_R$  is the reverse current in the pn junction and  $v_R$  is the reverse-bias voltage across the pn junction, then the actual reverse current  $i_{RA}$  can be expressed as

$$i_{RA} = M i_R = \left( \frac{1}{1 - (v_R/BV)^n} \right) i_R \quad (2.2-14)$$

$M$  is the avalanche multiplication factor and  $n$  is an exponent that adjusts the sharpness of the “knee” of the curve shown in Fig. 2.2-4. Typically,  $n$  varies between 3 and 6. If both sides of the pn junction are heavily doped, the breakdown will take place by tunneling, leading to the

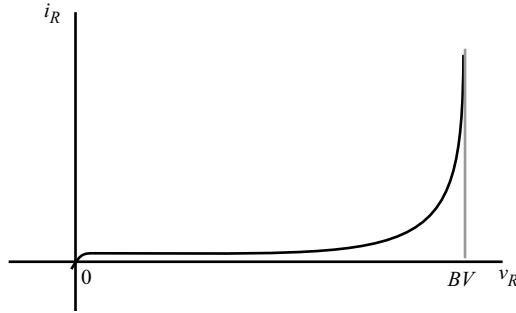


**Figure 2.2-4** Reverse-bias voltage–current characteristics of the pn junction illustrating voltage breakdown.

Zener breakdown, which generally occurs at voltages less than 6 volts. Zener diodes can be fabricated where an  $n^+$  diffusion overlaps with a  $p^+$  diffusion. Note that the Zener diode is compatible with the basic CMOS process although one terminal of the Zener must be either on the lowest power supply, ground, or the highest power supply,  $V_{DD}$ .

The diode voltage–current relationship can be derived by examining the minority-carrier concentrations in the pn junction. Figure 2.2-5 shows the minority-carrier concentration for a forward-biased pn junction. The majority-carrier concentrations are much larger and are not shown on this figure. The forward bias causes minority carriers to move across the junction where they recombine with majority carriers on the opposite side. The excess of minority-carrier concentration on each side of the junction is shown by the shaded regions. We note that this excess concentration starts at a maximum value at  $x = 0$  ( $x' = 0$ ) and decreases to the equilibrium value as  $x$  ( $x'$ ) becomes large. The value of the excess concentration at  $x = 0$ , designated as  $p_n(0)$ , or  $x' = 0$ , designated as  $n_p(0)$ , is expressed in terms of the forward-bias voltage  $v_D$  as

$$p_n(0) = p_{n0} \exp\left(\frac{v_D}{V_t}\right) \quad (2.2-15)$$

and

$$n_p(0) = n_{p0} \exp\left(\frac{v_D}{V_t}\right) \quad (2.2-16)$$

where  $p_{n0}$  and  $n_{p0}$  are the equilibrium concentrations of the minority carriers in the n-type and p-type semiconductors, respectively. We note that these values are essentially equal to the intrinsic concentration squared divided by the donor or acceptor impurity atom concentration, as shown on Fig. 2.2-5. As  $v_D$  is increased, the excess minority concentrations are increased. If  $v_D$  is zero, there is no excess minority concentration. If  $v_D$  is negative (reverse biased), the minority-carrier concentration is depleted below its equilibrium value.

The current that flows in the pn junction is proportional to the slope of the excess minority-carrier concentration at  $x = 0$  ( $x' = 0$ ). This relationship is given by the diffusion equation expressed below for holes in the n-type material.

$$J_p(x) = -qD_p \left. \frac{dp_n(x)}{dx} \right|_{x=0} \quad (2.2-17)$$

where  $D_p$  is the diffusion constant of holes in an n-type semiconductor. The excess holes in the n-type material can be defined as

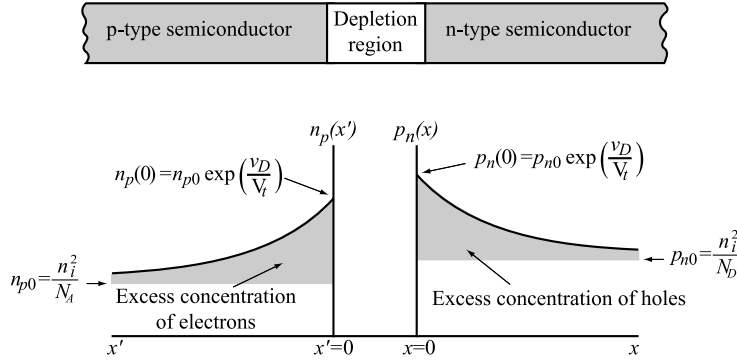


**Figure 2.2-5** Impurity concentration profile for diffused pn junction.

$$p'_n(x) = p_n(x) - p_{n0} \quad (2.2-18)$$

The decrease of excess minority carriers away from the junction is exponential and can be expressed as

$$p'_n(x) = p'_n(0) \exp\left(\frac{-x}{L_p}\right) = [p_n(0) - p_{n0}] \exp\left(\frac{-x}{L_p}\right) \quad (2.2-19)$$

where  $L_p$  is the diffusion length for holes in an n-type semiconductor. Substituting Eq. (2.2-15) into Eq. (2.2-19) gives

$$p'_n(x) = p_{n0} \left[ \exp\left(\frac{v_D}{V_t}\right) - 1 \right] \exp\left(\frac{-x}{L_p}\right) \quad (2.2-20)$$

The current density due to the excess-hole concentration in the n-type semiconductor is found by substituting Eq. (2.2-20) in Eq. (2.2-17), resulting in

$$J_p(0) = \frac{qD_p p_{n0}}{L_p} \left[ \exp\left(\frac{v_D}{V_t}\right) - 1 \right] \quad (2.2-21)$$

Similarly, for the excess electrons in the p-type semiconductor we have

$$J_n(0) = \frac{qD_n n_{p0}}{L_n} \left[ \exp\left(\frac{v_D}{V_t}\right) - 1 \right] \quad (2.2-22)$$

Assuming negligible recombination in the depletion region leads to an expression for the total current density of the pn junction, given as

$$J(0) = J_p(0) + J_n(0) = q \left[ \frac{D_p p_{n0}}{L_p} + \frac{D_n n_{p0}}{L_n} \right] \left[ \exp\left(\frac{v_D}{V_t}\right) - 1 \right] \quad (2.2-23)$$

Multiplying Eq. (2.2-23) by the pn junction area  $A$  gives the total current as

$$i_D = qA \left[ \frac{D_p p_{n0}}{L_p} + \frac{D_n n_{p0}}{L_n} \right] \left[ \exp\left(\frac{v_D}{V_t}\right) - 1 \right] = I_s \left[ \exp\left(\frac{v_D}{V_t}\right) - 1 \right] \quad (2.2-24)$$

$I_s$  is a constant called the *saturation current*. Equation (2.2-24) is the familiar voltage–current relationship that characterizes the pn junction diode.

### Example 2.2-2

#### Calculation of the Saturation Current

Calculate the saturation current of a pn junction diode with  $N_A = 5 \times 10^{15}/\text{cm}^3$ ,  $N_D = 10^{20}/\text{cm}^3$ ,  $D_n = 20 \text{ cm}^2/\text{s}$ ,  $D_p = 10 \text{ cm}^2/\text{s}$ ,  $L_n = 10 \text{ }\mu\text{m}$ ,  $L_p = 5 \text{ }\mu\text{m}$ , and  $A = 1000 \text{ }\mu\text{m}^2$ .

From Eq. (2.2-24), the saturation current is defined as

$$I_s = qA \left( \frac{D_p p_{n0}}{L_p} + \frac{D_n n_{p0}}{L_n} \right)$$

$p_{n0}$  is calculated from  $n_i^2/N_D$  to get  $2.103/\text{cm}^3$ ;  $n_{p0}$  is calculated from  $n_i^2/N_A$  to get  $4.205 \times 10^4/\text{cm}^3$ . Changing the units of area from  $\mu\text{m}^2$  to  $\text{cm}^2$  results in a saturation current magnitude of  $1.346 \times 10^{-15} \text{ A}$  or  $1.346 \text{ fA}$ .

This section has developed the depletion-region width, depletion capacitance, breakdown voltage, and voltage–current characteristics of the pn junction. These concepts will be very important in determining the characteristics and performance of MOS active and passive components.



## 2.3 The MOS Transistor

The structure of an n-channel and a p-channel MOS transistor fabricated using an n-well LOCOS process (same principles apply for a twin-well STI process) is shown in Fig. 2.3-1. The p-channel device is formed with two heavily doped  $p^+$  regions diffused into a lighter doped  $n^-$  material called the well. The two  $p^+$  regions are called drain and source, and are separated by a distance,  $L$  (referred to as the device length). At the surface between the drain and source lies a gate electrode that is separated from the silicon by a thin dielectric material (silicon dioxide). Similarly, the n-channel transistor is formed by two heavily doped  $n^+$  regions within a lightly doped  $p^-$  substrate and has a gate on the surface between the drain and source separated from the silicon by a thin dielectric material (silicon dioxide). Essentially, both types of transistors are four-terminal devices as shown in Fig. 1.2-2(c,d). The  $B$  terminal is the bulk, or substrate, which contains the drain and source diffusions. For an n-well process, the p-bulk connection is common throughout the integrated circuit and is connected to ground (the most negative supply). Multiple n-wells can be fabricated on a single circuit, and they can be connected to different potentials in various ways depending on the application.

Figure 2.3-2 shows an n-channel transistor with all four terminals connected to ground. At equilibrium, the  $p^-$  substrate and the  $n^+$  source and drain form a pn junction. Therefore, a depletion region exists between the  $n^+$  source and drain and the  $p^-$  substrate. Since the source and drain are separated by back-to-back pn junctions, the resistance between the source and drain is very high ( $>10^{12} \Omega$ ). The gate and the substrate of the MOS transistor form the parallel plates of a capacitor with the  $\text{SiO}_2$  as the dielectric. This capacitance divided by the area

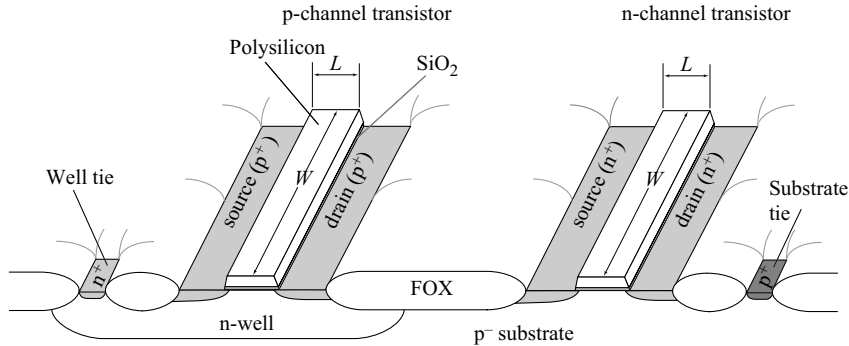


**Figure 2.3-1** Physical structure of an n-channel and a p-channel transistor in an n-well technology.

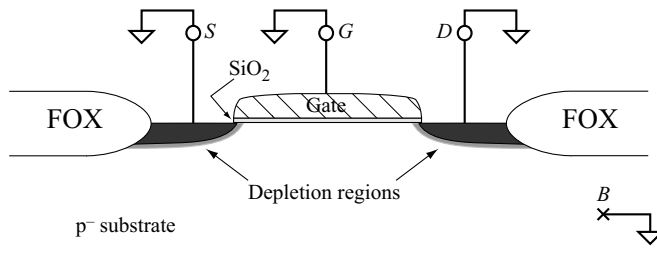


**Figure 2.3-2** Cross section of an n-channel transistor with all terminals grounded.

of the gate is designated as  $C_{ox}$ .<sup>\*</sup> When a positive potential is applied to the gate with respect to the source, a *depletion region* is formed under the gate resulting from holes being pushed away from the silicon–silicon dioxide interface. The depletion region consists of fixed ions that have a negative charge. Using one-dimensional analysis, the charge density,  $\rho$ , of the depletion region is given by

$$\rho = q(-N_A) \quad (2.3-1)$$

Applying the point form of Gauss's law, the electric field resulting from this charge is

$$E(x) = \int \frac{\rho}{\epsilon} dx = \int \frac{-qN_A}{\epsilon_{Si}} dx = \frac{-qN_A}{\epsilon_{Si}}(x) + C \quad (2.3-2)$$

The constant,  $C$ , is determined by evaluating  $E(x)$  at the edges of the depletion region ( $x = 0$  at the Si–SiO<sub>2</sub> interface;  $x = x_d$  at the boundary of the depletion region in the bulk).

$$E(0) = E_0 = \frac{-qN_A}{\epsilon_{Si}}(0) + C = C \quad (2.3-3)$$

$$E(x_d) = 0 = \frac{-qN_A}{\epsilon_{Si}}x_d + C \quad (2.3-4)$$

$$C = \frac{qN_A}{\epsilon_{Si}}x_d \quad (2.3-5)$$

<sup>\*</sup>The symbol “ $C$ ” normally has units of farads; however, in the field of MOS devices it often has units of farads per unit area (e.g., F/m<sup>2</sup>).

This gives an expression for  $E(x)$

$$E(x) = \frac{qN_A}{\epsilon_{Si}} (x_d - x) \quad (2.3-6)$$

Applying the relationship between potential and electric field yields

$$\int d\phi = - \int E(x) dx = - \int \frac{qN_A}{\epsilon_{Si}} (x_d - x) dx \quad (2.3-7)$$

Integrating both sides of Eq. (2.3-7) with appropriate limits of integration gives

$$\int_{\phi_s}^{\phi_F} d\phi = - \int_0^{x_d} \frac{qN_A}{\epsilon_{Si}} (x_d - x) dx = - \frac{qN_A x_d^2}{2\epsilon_{Si}} = \phi_F - \phi_s \quad (2.3-8)$$

$$\frac{qN_A x_d^2}{2\epsilon_{Si}} = \phi_s - \phi_F \quad (2.3-9)$$

where  $\phi_F$  is the equilibrium electrostatic potential (Fermi potential) in the semiconductor,  $\phi_s$  is the surface potential of the semiconductor, and  $x_d$  is the thickness of the depletion region. For a p-type semiconductor,  $\phi_F$  is given as

$$\phi_F = -V_t \ln(N_A/n_i) \quad (2.3-10)$$

and for an n-type semiconductor,  $\phi_F$  is given as

$$\phi_F = V_t \ln(N_D/n_i) \quad (2.3-11)$$

Equation (2.3-9) can be solved for  $x_d$  assuming that  $|\phi_s - \phi_F| \geq 0$  to get

$$x_d = \left[ \frac{2\epsilon_{Si} |\phi_s - \phi_F|}{qN_A} \right]^{1/2} \quad (2.3-12)$$

The immobile charge due to acceptor ions that have been stripped of their mobile holes is given by

$$Q = -qN_A x_d \quad (2.3-13)$$

Substituting Eq. (2.3-12) into Eq. (2.3-13) gives

$$Q \cong -qN_A \left[ \frac{2\epsilon_{Si} |\phi_s - \phi_F|}{qN_A} \right]^{1/2} = -\sqrt{2qN_A \epsilon_{Si} |\phi_s - \phi_F|} \quad (2.3-14)$$

When the gate voltage reaches a value called the *threshold voltage*, designated as  $V_T$ , the substrate underneath the gate becomes inverted; that is, it changes from a p-type to an n-type semiconductor. Consequently, an n-type channel exists between the source and drain that allows carriers to flow. In order to achieve this inversion, the surface potential must increase from its original negative value ( $\phi_s - \phi_F$ ), to zero ( $\phi_s = 0$ ), and then to a positive value ( $\phi_s = -\phi_F$ ). The value of gate-source voltage necessary to cause this change in surface potential

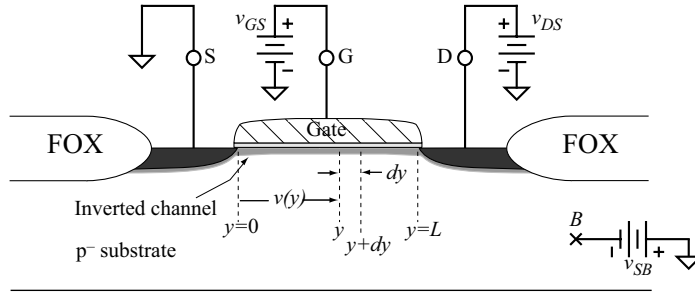


**Figure 2.3-3** Cross section of an n-channel transistor with small  $v_{DS}$  and  $v_{GS} > V_T$ .

is defined as the *threshold voltage*,  $V_T$ . This condition is known as strong inversion. The n-channel transistor in this condition is illustrated in Fig. 2.3-3. With the substrate at ground potential, the charge stored in the depletion region between the channel under the gate and the substrate is given by Eq. (2.3-14) where  $\phi_s$  has been replaced by  $-\phi_F$  to account for the fact that  $v_{GS} = V_T$ . This charge  $Q_{b0}$  is written as

$$Q_{b0} \cong -\sqrt{2qN_A\epsilon_{Si}}|-2\phi_F| \quad (2.3-15)$$

If a reverse-bias voltage  $v_{BS}$  is applied across the pn junction, Eq. (2.3-15) becomes

$$Q_b \cong \sqrt{2qN_A\epsilon_{Si}}|-2\phi_F + v_{SB}| \quad (2.3-16)$$

An expression for the threshold voltage can be developed by breaking it down into several components. First, the term  $\phi_{MS}$ \* must be included to represent the difference in the work functions between the gate material and bulk silicon in the channel region. The term  $\phi_{MS}$  is given by

$$\phi_{MS} = \phi_F(\text{substrate}) - \phi_F(\text{gate}) \quad (2.3-17)$$

where  $\phi_F(\text{metal}) = 0.6$  V. Second, a gate voltage of  $[-2\phi_F - (Q_b/C_{ox})]$  is required to change the surface potential and offset the depletion layer charge  $Q_b$ . Lastly, there is always an undesired positive charge,  $Q_{ss}$ , present in the interface between the oxide and the bulk silicon. This charge is due to impurities and imperfections at the interface and must be compensated by a gate voltage of  $-Q_{ss}/C_{ox}$ . Thus, the threshold voltage for the MOS transistor can be expressed as

$$\begin{aligned} V_T &= \phi_{MS} + \left(-2\phi_F - \frac{Q_b}{C_{ox}}\right) + \left(\frac{-Q_{ss}}{C_{ox}}\right) \\ &= \phi_{MS} - 2\phi_F - \frac{Q_{b0}}{C_{ox}} - \frac{Q_{ss}}{C_{ox}} - \frac{Q_b - Q_{b0}}{C_{ox}} \end{aligned} \quad (2.3-18)$$

\*Historically, this term has been referred to as the *metal-to-silicon* work function. We will continue the tradition even when the gate terminal is something other than metal (e.g., polysilicon).

**Table 2.3-1** Signs for the Quantities in the Threshold Voltage Equation.

| Parameter              | N-CHANNEL (p-type substrate) | P-CHANNEL (n-type substrate) |
|------------------------|------------------------------|------------------------------|
| $\phi_{MS}$            |                              |                              |
| Metal                  | —                            | —                            |
| n <sup>+</sup> Si Gate | —                            | —                            |
| p <sup>+</sup> Si Gate | +                            | +                            |
| $\phi_F$               | —                            | +                            |
| $Q_{b0}, Q_b$          | —                            | +                            |
| $Q_{ss}$               | +                            | +                            |
| $V_{SB}$               | +                            | —                            |
| $\gamma$               | +                            | —                            |

The threshold voltage can be rewritten as

$$V_T = V_{T0} + \gamma \left( \sqrt{|-2\phi_F + v_{SB}|} - \sqrt{|-2\phi_F|} \right) \quad (2.3-19)$$

where

$$V_{T0} = \phi_{MS} - 2\phi_F - \frac{Q_{b0}}{C_{ox}} - \frac{Q_{ss}}{C_{ox}} \quad (2.3-20)$$

and the body factor,  $\gamma$ , is defined as

$$\gamma = \frac{\sqrt{2q\epsilon_{Si}N_A}}{C_{ox}} \quad (2.3-21)$$

The signs of the above analysis can become very confusing. Table 2.3-1 attempts to clarify any confusion that might arise [24].

### Example 2.3-1

#### Calculation of the Threshold Voltage

Find the threshold voltage and body factor  $\gamma$  for an n-channel transistor with an n<sup>+</sup> silicon gate if  $t_{ox} = 200 \text{ \AA}$ ,  $N_A = 3 \times 10^{16} \text{ cm}^{-3}$ , gate doping,  $N_D = 4 \times 10^{19} \text{ cm}^{-3}$ , and if the positively charged ions at the oxide–silicon interface per area is  $10^{10} \text{ cm}^{-2}$ .

From Eq. (2.3-10),  $\phi_F(\text{substrate})$  is given as

$$\phi_F(\text{substrate}) = -0.0259 \ln \left( \frac{3 \times 10^{16}}{1.45 \times 10^{10}} \right) = -0.377 \text{ V}$$

The equilibrium electrostatic potential for the n<sup>+</sup> polysilicon gate is found from Eq. (2.3-11) as

$$\phi_F(\text{gate}) = 0.0259 \ln \left( \frac{4 \times 10^{19}}{1.45 \times 10^{10}} \right) = 0.563 \text{ V}$$

Equation (2.3-17) gives  $\phi_{MS}$  as

$$\phi_F(\text{substrate}) - \phi_F(\text{gate}) = -0.940 \text{ V}$$

The oxide capacitance is given as

$$C_{ox} = \epsilon_{ox}/t_{ox} = \frac{3.9 \times 8.854 \times 10^{-14}}{200 \times 10^{-8}} = 1.727 \times 10^{-7} \text{ F/cm}^2$$

The fixed charge in the depletion region,  $Q_{b0}$ , is given by Eq. (2.3-15) as

$$\begin{aligned} Q_{b0} &= -(2 \times 1.6 \times 10^{-19} \times 11.7 \times 8.854 \times 10^{-14} \times 2 \times 0.377 \times 3 \times 10^{16})^{1/2} \\ &= -8.66 \times 10^{-8} \text{ C/cm}^2 \end{aligned}$$

Dividing  $Q_{b0}$  by  $C_{ox}$  gives  $-0.501$  V. Finally,  $Q_{ss}/C_{ox}$  is given as

$$\frac{Q_{ss}}{C_{ox}} = \frac{10^{10} \times 1.60 \times 10^{-19}}{1.727 \times 10^{-7}} = 9.3 \times 10^{-3} \text{ V}$$

Substituting these values in Eq. (2.3-18) gives

$$V_{T0} = -0.940 + 0.754 + 0.501 - 9.3 \times 10^{-3} = 0.306 \text{ V}$$

The body factor is found from Eq. (2.3-21) as

$$\gamma = \frac{(2 \times 1.6 \times 10^{-19} \times 11.7 \times 8.854 \times 10^{-14} \times 3 \times 10^{16})^{1/2}}{1.727 \times 10^{-7}} = 0.577 \text{ V}^{1/2}$$

The above example shows how the value of impurity concentrations can influence the threshold voltage. In fact, the threshold voltage can be set to any value by proper choice of the variables in Eq. (2.3-18). Standard practice is to implant the appropriate ions into the substrate in the channel region to adjust the threshold voltage to the desired value. If the opposite impurities are implanted in the channel region of the substrate, the threshold for an n-channel transistor can be made negative. This type of transistor is called a *depletion transistor* and can have current flow between the drain and source for zero volts between the gate and source.

When the channel is formed between the drain and source as illustrated in Fig. 2.3-3, a drain current  $i_D$  can flow if a voltage  $v_{DS}$  exists across the channel. The dependence of this drain current on the terminal voltages of the MOS transistor can be developed by considering the characteristics of an incremental length of the channel designated as  $dy$  in Fig. 2.3-3. It is assumed that the width of the MOS transistor (into the page) is  $W$  and that  $v_{DS}$  is small. The charge per unit area in the channel,  $Q_I(y)$ , can be expressed as

$$Q_I(y) = C_{ox}[v_{GS} - v(y) - V_T] \quad (2.3-22)$$

The resistance in the channel per unit of length  $dy$  can be written as

$$dR = \frac{dy}{\mu_n Q_I(y) W} \quad (2.3-23)$$

where  $\mu_n$  is the average mobility of the electrons in the channel. The voltage drop, referenced to the source, along the channel in the  $y$  direction, is

$$dv(y) = i_D dR = \frac{i_D dy}{\mu_n Q_I(y) W} \quad (2.3-24)$$

or

$$i_D dy = W \mu_n Q_I(y) dv(y) \quad (2.3-25)$$

Integrating along the channel from  $y = 0$  to  $y = L$  gives

$$\int_0^L i_D dy = \int_0^{v_{DS}} W \mu_n Q_I(y) dv(y) = \int_0^{v_{DS}} W \mu_n C_{ox} [v_{GS} - v(y) - V_T] dv(y) \quad (2.3-26)$$

Performing the integration results in the desired expression for  $i_D$  as

$$\begin{aligned} i_D &= \frac{\mu_n C_{ox} W}{L} \left[ (v_{GS} - V_T) v(y) - \frac{v(y)^2}{2} \right]_0^{v_{DS}} \\ &= \frac{\mu_n C_{ox} W}{L} \left[ (v_{GS} - V_T) v_{DS} - \frac{v_{DS}^2}{2} \right] \end{aligned} \quad (2.3-27)$$

This equation is sometimes called the Sah equation [26] and has been used by Shichman and Hodges [27] as a model for computer simulation. Equation (2.3-27) is only valid when

$$v_{GS} \geq V_T \text{ and } v_{DS} \leq (v_{GS} - V_T) \quad (2.3-28)$$

and for values of  $L$  greater than the minimum  $L$ . The factor  $\mu_n C_{ox}$  is often called the transconductance parameter, given as

$$K' = \mu_n C_{ox} = \frac{\mu_n \epsilon_{ox}}{t_{ox}} \quad (2.3-29)$$

Equation (2.3-28) will be examined in more detail in the next chapter, concerning the modeling of MOS transistors. The operation of the p-channel transistor is essentially the same as that of the n-channel transistor, except that all voltage and current polarities are reversed.



## 2.4 Passive Components

This section examines the passive components that are compatible with fabrication steps used to build the MOS device. These passive components include the capacitor and the resistor.

### Capacitors

A good capacitor is often required when designing analog integrated circuits. They are used as compensation capacitors in amplifier designs, as bandwidth-determining components in  $g_m/C$  filters, as charge storage devices in switched-capacitor filters and digital-to-analog

converters, and in other places as well. The desired characteristics for capacitors used in these applications are given below:

- Good matching accuracy
- Low voltage coefficient
- High ratio of desired capacitance to parasitic capacitance
- High capacitance per unit area
- Low temperature dependence

Analog CMOS processes differentiate themselves from purely digital ones by providing capacitors that meet the above criteria. For such analog processes, there are basically two types of capacitors made available, capacitors utilizing polysilicon for plates and capacitors utilizing metal for plates. Modern submicron processes ( $0.18\ \mu\text{m}$  and below) use MiM (metal-insulator-metal) capacitors because they are economical to build (minimal masks, steps, and process complexity) and meet the criteria set forth above. Figure 2.4-1(a) shows an example of a MiM capacitor. It is formed by providing an additional metal layer prior to the last metal interconnect layer that lies above a thin oxide region on top of the next to last metal layer (i.e., a capacitor metal on top of M4, separated by a thin oxide region fabricated in a five-metal process). Typical performance for a MiM capacitor implemented in a  $0.18\ \mu\text{m}$  process is given in Table 2.4-1.

A second type of capacitor is illustrated in Fig. 2.4-1(b). This capacitor is constructed by putting an n-well underneath an n-channel transistor. The bottom plate (the n-well) of this

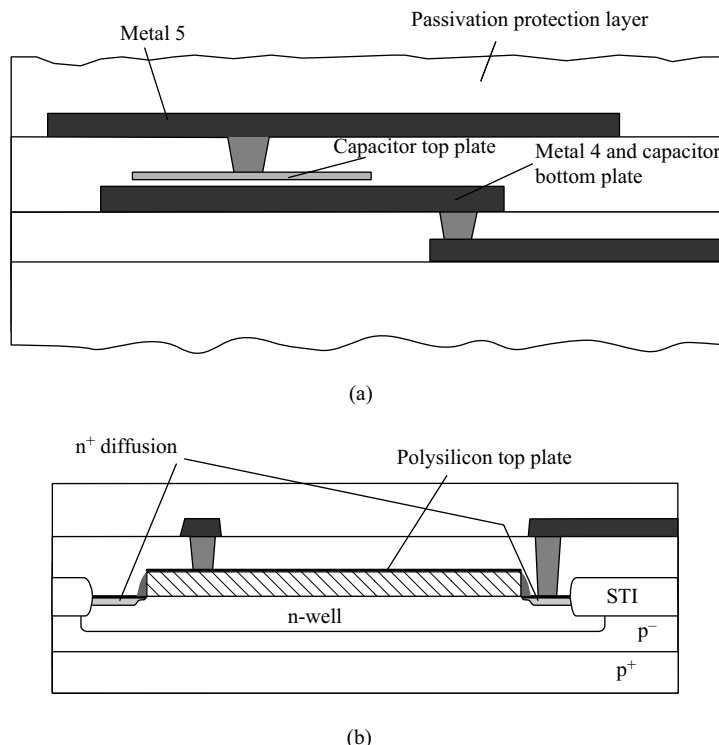


**Figure 2.4-1** (a) MiM capacitor, (b) accumulation MOS capacitor.

**Table 2.4-1 Approximate Performance Summary of Passive Components in a 0.18  $\mu\text{m}$  CMOS Process**

| Component Type                                     | Typical Value            | Typical Matching Accuracy | Temperature Coefficient       | Voltage Coefficient |
|----------------------------------------------------|--------------------------|---------------------------|-------------------------------|---------------------|
| MiM capacitor                                      | 1.0 fF/ $\mu\text{m}^2$  | 0.03%                     | 50 ppm/ $^{\circ}\text{C}$    | 50 ppm/V            |
| MOM capacitor                                      | 0.17 fF/ $\mu\text{m}^2$ | 1%                        | 50 ppm/ $^{\circ}\text{C}$    | 50 ppm/V            |
| P <sup>+</sup> Diffused resistor<br>(nonsilicide)  | 80–150 $\Omega/\square$  | 0.4%                      | 1500 ppm/ $^{\circ}\text{C}$  | 200 ppm/V           |
| N <sup>+</sup> Diffused resistor<br>(non-silicide) | 50–80 $\Omega/\square$   | 0.4%                      | 1500 ppm/ $^{\circ}\text{C}$  | 200 ppm/V           |
| N <sup>+</sup> Poly resistor<br>(non-silicide)     | 300 $\Omega/\square$     | 2%                        | –2000 ppm/ $^{\circ}\text{C}$ | 100 ppm/V           |
| P <sup>+</sup> Poly resistor<br>(non-silicide)     | 300 $\Omega/\square$     | 0.5%                      | –500 ppm/ $^{\circ}\text{C}$  | 100 ppm/V           |
| P <sup>–</sup> Poly resistor<br>(non-silicide)     | 1000 $\Omega/\square$    | 0.5%                      | –1000 ppm/ $^{\circ}\text{C}$ | 100 ppm/V           |
| n-well resistor                                    | 1–2 k $\Omega/\square$   |                           | 8000 ppm/ $^{\circ}\text{C}$  | 10k ppm/V           |

capacitor has a much higher resistivity (as compared to a MiM bottom plate). Because of this fact, it is not used in circuits where a low voltage coefficient is important. It is, however, often used when one terminal of the capacitor is connected to ground. It offers a very high capacitance per unit area, it can be matched well, and it is available in all CMOS processes because no unique steps or masks are required. A conventional transistor can be used as a capacitor if properly biased in the inversion region and has similar performance to the capacitor illustrated in Fig. 2.4-1(b).

Quite often, the processing performance required by the digital component of a mixed-signal integrated circuit necessitates the use of a process targeted for digital applications. Such processes do not provide tailored capacitors for analog applications. Therefore, when a capacitor is needed, it must be derived from two or more of the existing interconnect layers. Capacitors constructed of only native metal layers are called *MOM* (metal-oxide-metal) capacitors (or *finger capacitors* if constructed using interdigitated fingers). Figure 2.4-2 illustrates symbolically several schemes for making MOM capacitors using only metal layers available in a purely digital process. Figure 2.4-2(a) is a finger capacitor constructed vertically, taking advantage of both vertical and sidewall capacitance. This structure is very efficient because the vertical space between metal layers is greater than the allowed space between them (see Fig. 2.1-6). An example of a same-layer parallel-plate MOM capacitor is illustrated in Fig. 2.4-2(b). Compared to MiM capacitors, these capacitors typically suffer from lower per-unit-area capacitance and lower ratio of desired capacitance to parasitic capacitance; however, they generally have a higher  $Q$ , which makes them the capacitor of choice for RF circuit applications. Matching accuracy of capacitors implemented like those in Fig 2.4-2 is on the order of 1–2% and voltage coefficient is low. A typical value for a finger capacitor implemented in four metal layers is about 0.2 fF/ $\mu\text{m}^2$  (depends greatly on layout).

The voltage coefficient of MiM and MOM capacitors is generally less than 50 ppm/V and their temperature coefficient less than 50 ppm/ $^{\circ}\text{C}$ . When considering the ratio of two capacitors

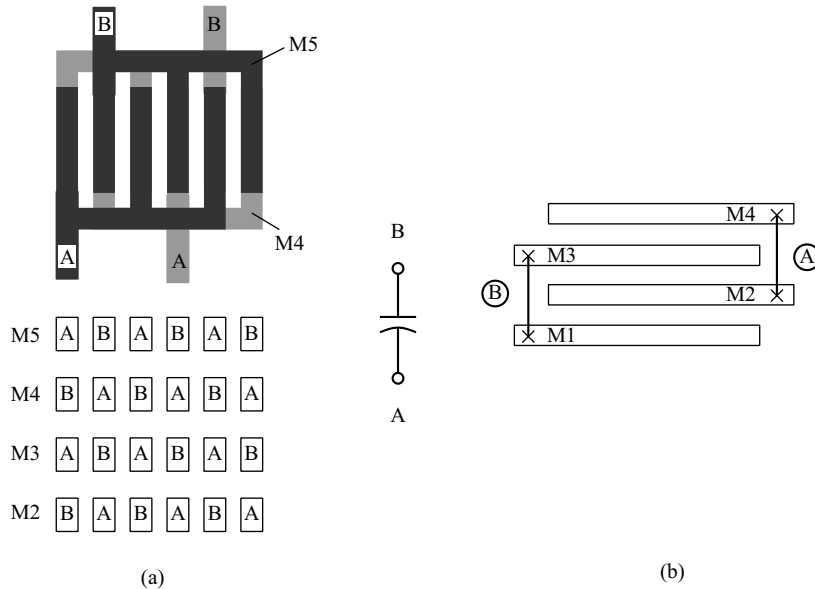


**Figure 2.4-2** Various ways to implement capacitors using available interconnect layers illustrated symbolically: (a) interleaved finger capacitor, (b) horizontal parallel plate capacitor.

on the same substrate, note that the variations on the absolute value of the capacitor due to temperature tend to cancel. Therefore, temperature variations have little effect on the matching accuracy of capacitors. When capacitors are switched to different voltages, as in the case of sampled-data circuits, their voltage coefficients can have a negative impact on performance if not kept to a minimum.

The parasitic capacitance associated with the capacitors of Figs. 2.4-1 and 2.4-2 can give rise to a significant source of error in analog sampled-data circuits. Figure 2.4-3 illustrates a general model for a capacitor with its top- and bottom-plate parasitics. These parasitic capacitances depend on the capacitor size, layout, and technology and are unavoidable.

The capacitor plate with the smallest parasitic associated with it is referred to as the *top plate*. It is not necessarily physically the top plate although quite often it is. The *bottom plate* is that plate having the larger parasitic capacitance associated with it. Schematically, the top plate is represented by the flat (straight) line in the capacitor symbol while the curved plate represents the bottom plate. For the MiM capacitor illustrated in Fig. 2.4-1(a) the parasitic capacitor associated with the top plate of the capacitor itself is due to interconnect lines leading to the capacitor and the bottom-plate parasitic capacitance is due to the capacitance

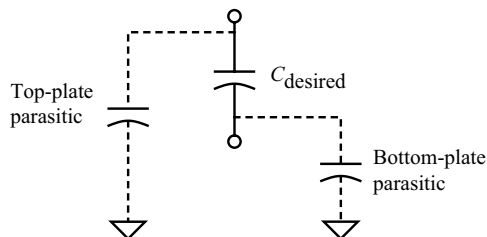


**Figure 2.4-3** A model for the integrated capacitors showing top- and bottom-plate parasitics.

between the bottom plate and the substrate or any intervening metal lines below the bottom plate. Parasitic bottom-plate capacitance of the MiM capacitor is minimized by constructing it in such a way so that there are no metal or polysilicon interconnects running below the bottom-plate metal. MOM capacitors available in a digital process and shown in Fig. 2.4-2 have parasitics that are not so easily generalized.

## Resistors

The other passive component compatible with MOS technology is the resistor. Even though we shall use circuits consisting of primarily MOS active devices and capacitors, some applications, such as digital-to-analog conversion, use resistors. Resistors compatible with the MOS technology of this section include diffused, polysilicon, and n-well resistors. Though not as common, metal can be used as a resistor as well.

To understand the dimensions that are important in accessing the performance of a resistor, it is necessary to review the relationship for the resistance of a conductive bar.

For a conductive bar of material as shown in Fig. 2.4-4, the resistance,  $R$ , is given as

$$R = \frac{\rho L}{A} \quad (\Omega) \quad (2.4-1)$$

where  $\rho$  is resistivity in  $\Omega\text{-cm}$  and  $A$  is a plane perpendicular to the direction of current flow. In terms of the dimensions given in Fig. 2.4-4, Eq. (2.4-1) can be rewritten as

$$R = \frac{\rho L}{WT} \quad (\Omega) \quad (2.4-2)$$

Since the nominal values for  $\rho$  and  $T$  are generally fixed for a given process and material type, they are grouped together to form a new term  $\rho_s$  called sheet resistivity. This is clarified by the following expression:

$$R = \left( \frac{\rho}{T} \right) \frac{L}{W} = \rho_s \frac{L}{W} \quad (\Omega) \quad (2.4-3)$$

It is conventional to give  $\rho_s$  the units of  $\Omega/\square$  (read *ohms per square*). From the layout point of view, a resistor has the value determined by the *number of squares* of resistance multiplied by  $\rho_s$ .

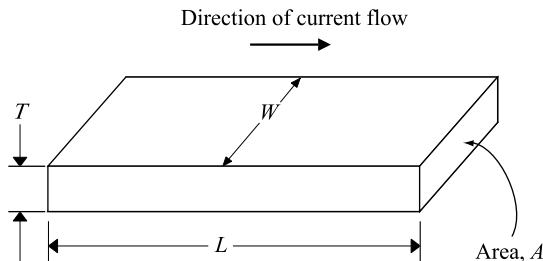


**Figure 2.4-4** Current flow in a conductive bar.

### Example 2.4-1

#### Example 2.4-1 Resistance Calculation

Given a polysilicon resistor like that drawn in Fig. 2.4-4 with  $W = 0.8 \mu\text{m}$  and  $L = 20 \mu\text{m}$ , calculate  $\rho_s$  (in  $\Omega/\square$ ), the number of squares of resistance, and the resistance value. Assume that  $\rho$  for polysilicon is  $9 \times 10^{-4} \Omega\text{-cm}$  and polysilicon is  $3000 \text{ \AA}$  thick. Ignore any contact resistance.

First calculate  $\rho_s$ .

$$\rho_s = \frac{\rho}{T} = \frac{9 \times 10^{-4} \Omega\text{-cm}}{3000 \times 10^{-8} \text{ cm}} = 30 \Omega/\square$$

The number of squares of resistance,  $N$ , is

$$N = \frac{L}{W} = \frac{20 \mu\text{m}}{0.8 \mu\text{m}} = 25$$

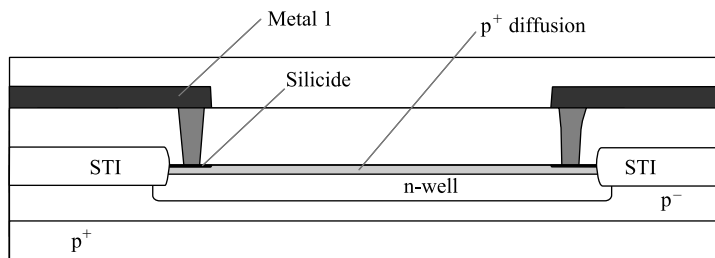
giving the total resistance as

$$R = \rho_s \times N = 30 \times 25 = 750 \Omega$$

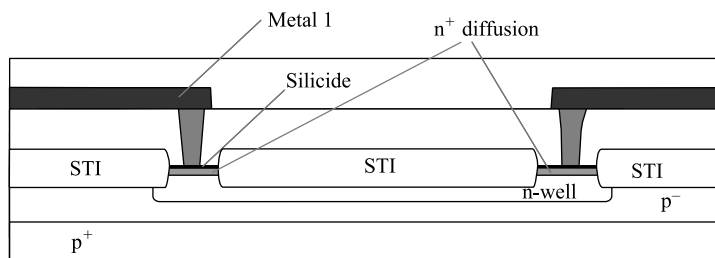
Figure 2.4-5(a) illustrates a  $p^+$  diffusion resistor. It is formed using source/drain diffusion in an n-well. The sheet resistance of such resistors in a nonsilicide process is usually in the range of  $50\text{--}150 \Omega/\square$ . For a silicide process, these resistors are in the range of  $5\text{--}15 \Omega/\square$ . The fact that the source/drain diffusion is needed as a good conductor (i.e., as interconnect) in integrated circuits conflicts with its use as a resistor. Clearly the goal of a silicide process is to achieve “interconnect-like” performance from source/drain diffusion. In these processes, a *silicide block* can be used to mask the titanium film from forming a silicide, thus allowing for a high-resistance source/drain diffusion where desired. The diffused resistor is found to have a voltage coefficient of resistance in the  $100\text{--}500 \text{ ppm/V}$  range. The parasitic capacitance to ground is also voltage dependent in this type of resistor. Figure 2.4-5(b) shows an example of an n-well resistor. n-Well resistors have much higher resistivity (on the order of  $1000 \Omega/\square$ ) and a high-voltage coefficient. In cases where accuracy is not required, such as pull-up resistors, or protection resistors, this structure is very useful.

Polysilicon makes an excellent choice for use as a resistor because it is least affected by voltage coefficient and parasitic capacitance. Polysilicon also offers very good matching performance. A polysilicon resistor is shown in Fig. 2.4-6(a) where a silicide block was not implemented. Resistance for a polysilicide resistor is on the order of  $5\text{--}15 \Omega/\square$ . Nonsilicided polysilicon resistors have a sheet resistance in the range of  $30\text{--}200 \Omega/\square$  depending on doping levels. As noted earlier, an additional mask is generally required to allow for selective doping of the polysilicon resistor to achieve a useful resistivity.

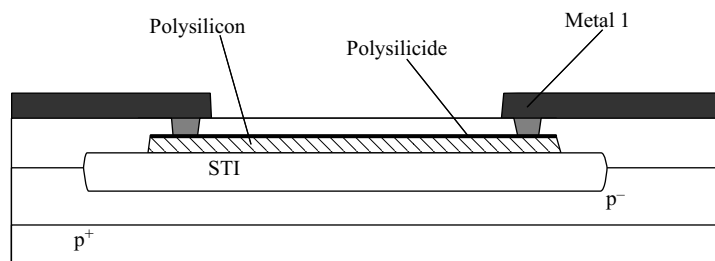
Other types of resistors are possible if the process is altered. The three categories above represent those most commonly applied with standard MOS technology. Table 2.4-1 summarizes the characteristics of the passive components presented.



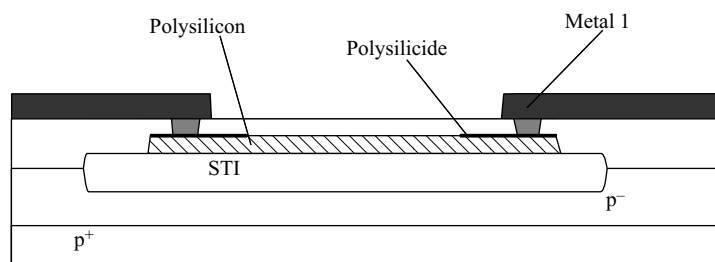
(a)



(b)

**Figure 2.4-5** Resistors: (a) diffusion, (b) n-well.

(a)



(b)

**Figure 2.4-6** Polysilicon resistors: (a) without silicide block, (b) with silicide block.



## 2.5 Other Considerations of CMOS Technology

In the previous two sections, the active and passive components of the basic CMOS process have been introduced. In this section we wish to consider some other components that are also available in the basic CMOS process but that are not used as extensively. We will further consider some of the limitations of CMOS technology, including latch-up, temperature, and noise. This information will become useful later, when the performance of CMOS circuits is characterized.

So far we have seen that it is possible to make resistors, capacitors, and pn diodes that are compatible with the basic twin-well CMOS fabrication process illustrated in Fig. 2.3-1. It is also possible to implement a *bipolar junction transistor* (BJT) that is compatible with this process, even though the collector terminal is constrained to ground (for an n-well process). Figure 2.5-1 shows how the BJT is implemented for an n-well process. The emitter is the source/drain diffusion, the base is the n-well (with a base width of  $w_B$ ), and the  $p^-$  substrate is the collector. Because the pn junction between the n-well and the  $p^-$  substrate must be reverse biased, the collector must always be connected to the most negative power-supply voltage (typically ground). The BJT will still find many useful applications even though the collector is constrained to ground. The BJT illustrated in Fig. 2.5-1 is often referred to as a *substrate BJT*. The substrate BJT functions like the BJT fabricated in a process designed for BJTs. The only difference is that the collector terminal is constrained and the base width is not well controlled, resulting in a wide variation of (and typically not very high) current gains.

Figure 2.5-2 shows the minority-carrier concentrations in the BJT. Normally, the base-emitter (*BE*) pn junction is forward biased and the collector-base (*CB*) pn junction is reverse biased. The forward-biased *BE* junction causes free holes to be injected into the base region. If the base width  $w_B$  is small, most of these holes reach the *CB* junction and are swept into the collector by the reverse-bias voltage. If the minority-carrier concentrations are much less than the majority-carrier concentrations, then the collector current can be found by solving for the current in the base region. In terms of current densities, the collector current density is

$$J_C = -J_p|_{\text{base}} = -qD_p \frac{dp_n(x)}{dx} = qD_p \frac{p_n(0)}{w_B} \quad (2.5-1)$$

From Eq. (2.2-16) we can write

$$p_n(0) = p_{n0} \exp\left(\frac{v_{EB}}{V_t}\right) \quad (2.5-2)$$

Combining Eqs. (2.5-1) and (2.5-2) and multiplying by the area of the *BE* junction  $A$  gives the collector current as

$$i_C = AJ_C = \frac{qAD_p p_{n0}}{w_B} \exp\left(\frac{v_{EB}}{V_t}\right) = I_s \exp\left(\frac{v_{EB}}{V_t}\right) \quad (2.5-3)$$

where  $I_s$  is defined as

$$I_s = \frac{qAD_p p_{n0}}{w_B} \quad (2.5-4)$$

As the holes travel through the base, a small fraction will recombine with electrons, which are the majority carriers in the base. As this occurs, an equal number of electrons must enter

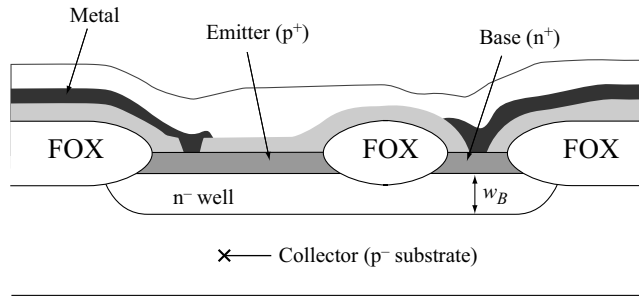


**Figure 2.5-1** Substrate BJT available from a bulk CMOS process.

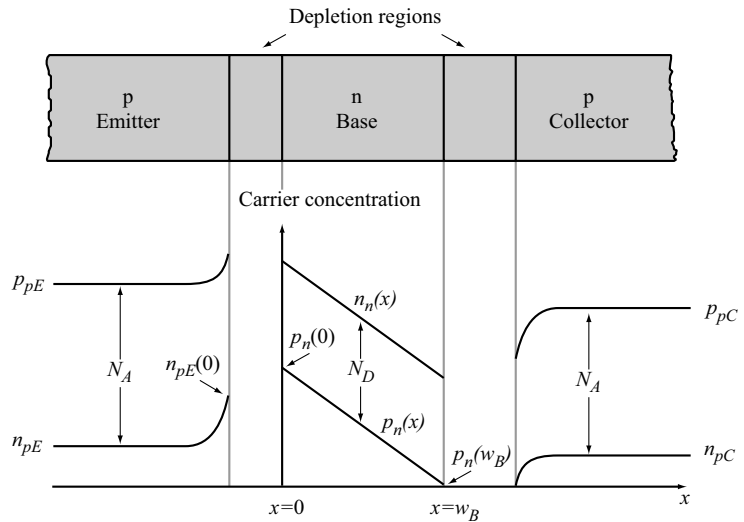


**Figure 2.5-2** Minority-carrier concentrations for a bipolar junction transistor.

the base from the external base circuit in order to maintain electrical neutrality in the base region. Also, there will be injection of the electrons from the base to the emitter due to the forward-biased  $BE$  junction. This injection is much smaller than the hole injection from the emitter because the emitter is more heavily doped than the base. The injection of electrons into the emitter and the recombination of electrons with holes in the base both constitute the external base current  $i_B$  that flows out of the base. The ratio of collector current to base current,  $i_C/i_B$ , is defined as  $\beta_F$  or the common-emitter current gain. Thus, the base current is expressed as

$$i_B = \frac{i_C}{\beta_F} = \frac{I_s}{\beta_F} \exp\left(\frac{v_{EB}}{V_t}\right) \quad (2.5-5)$$

The emitter current can be found from the base current and the collector current because the sum of all three currents must equal zero. Although  $\beta_F$  has been assumed constant it varies with  $i_C$ , having a maximum for moderate currents and falling off from this value for large or small currents.

In addition to the substrate BJT, it is also possible to have a lateral BJT. Figure 2.3-1 can be used to show how the lateral BJT can be implemented. The emitter could be the  $n^+$  source of the n-channel device, the base the  $p^-$  substrate, and the collector the n-well. Although the



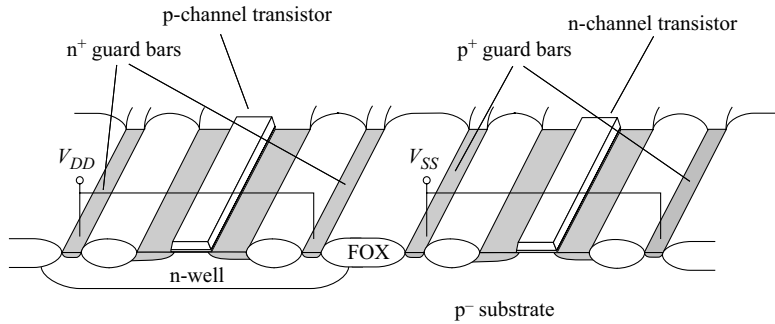


**Figure 2.5-4** Preventing latch-up using guard bars in an n-well technology.

clearly illustrated. The resistor  $R_N^-$  is the n-well resistance from the base of the vertical *PNP* (Q2) to  $V_{DD}$ . The resistor  $R_P^-$  is the substrate resistance from the base of the lateral *NPN* (Q2) to ground.

Regeneration occurs when three conditions are satisfied. The first condition is that the loop gain must exceed unity. This condition is stated as

$$\beta_{NPN}\beta_{PNP} \geq 1 \quad (2.5-6)$$

where  $\beta_{NPN}$  and  $\beta_{PNP}$  are the common-emitter, current-gain ratios of Q2 and Q1, respectively. The second condition occurs when both of the base-emitter junctions become forward biased. The third condition is enabled when the circuits connected to the emitter are capable of sinking and sourcing a current greater than the holding current of the PNPN device.

To prevent latch-up, several standard techniques should be applied by the designer. One is to keep the source/drain of the n-channel device as far away from the n-well as possible. This reduces the value of  $\beta_{NPN}$  and helps to prevent latch-up. Unfortunately, this is very costly in terms of area. A second approach is to reduce the values of  $R_N^-$  and  $R_P^-$ . Smaller resistor values are helpful because more current must flow through them in order to forward bias the base-emitter regions of Q1 and Q2. These resistances can be reduced by surrounding the p-channel devices with an  $n^+$  guard ring connected to  $V_{DD}$  and by surrounding n-channel transistors with  $p^+$  guard rings tied to ground as shown in Fig. 2.5-4.

Latch-up can also be prevented by keeping the potential of the source/drain of the p-channel device [A in Fig. 2.5-3(b)] from being higher than  $V_{DD}$  or the potential of the source/drain of the n-channel device [B in Fig. 2.5-3(b)] from going below ground. By careful design and layout, latch-up can be avoided in most cases. In the design of various circuits, particularly those that have high currents, one must use care to avoid circuit conditions that will initiate latch-up.

Another type of EOS event that must be considered is an *electrostatic discharge* (ESD). An ESD event occurs when a very high voltage (several thousand volts) appears across the pins of an integrated circuit. This can occur when a circuit is handled by a person who has built up a large static charge as a result of, for example, walking across a carpeted floor. When the person touches the circuit and a discharge path exists through the circuit, circuit components can be damaged. Thin gate oxides of input transistors are particularly vulnerable. Protection of the gates of these transistors, which are externally accessible, is the goal of an ESD protection methodology/circuit. The simplest structure to prevent accidental destruction

of the gate oxide is a resistor and two reverse-biased pn junction diodes forming an input protection circuit. One of the diodes is connected with the cathode to the highest circuit potential ( $V_{DD}$ ) and the anode side to the gate to be protected. The other diode is connected with the cathode to the gate to be protected and the anode to the lowest circuit potential (ground). This is illustrated in Fig. 2.5-5. For an n-well process, the first diode is usually made by a  $p^+$  diffusion into the n-well. The second diode is made by an  $n^+$  diffusion into the substrate. The resistor is connected between the external contact and the junction between the diodes and the gate to be protected. If a large voltage is applied to the input, one of the diodes will break down depending on the polarity of the voltage. If the resistor is large enough, it will limit the breakdown current so that the diode is not destroyed. This circuit should be used whenever the gates of a transistor (or transistors) are taken to external circuits.

The topic of ESD protection is considerably broader than the previous paragraph might imply. More complex protection circuits as well as consideration for total integrated circuit topology and interconnect must generally be taken into account to achieve optimal protection. When a designer arrives at the pins of an integrated circuit, he should consult the design guidelines for the process being used (generally provided by the fabrication vendor/company). Quite often, I/O cells are carefully engineered and are available to the designer as standard cells. In this way, the complexities of ESD protection need not be understood (only appreciated) by the designer, in general.

The temperature dependence of MOS components is an important performance characteristic in analog circuit design. The temperature behavior of passive components is usually expressed in terms of a *fractional temperature coefficient*,  $TC_F$ , defined as

$$TC_F = \frac{1}{X} \cdot \frac{dX}{dT} \quad (2.5-7)$$

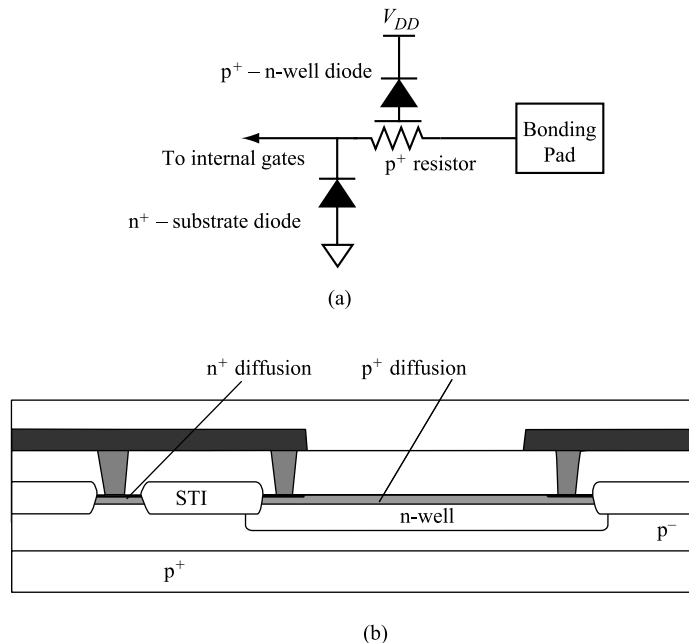


**Figure 2.5-5** Electrostatic discharge protection circuitry. (a) Electrical equivalent circuit. (b) Implementation in CMOS technology.

where  $X$  can be the resistance or capacitance of the passive component. Generally, the fractional temperature coefficient is multiplied by  $10^6$  and expressed in units of parts per million per °C, or ppm/°C. The fractional temperature coefficient of various CMOS passive components has been given in Table 2.4-1.

The temperature dependence of the MOS device can be found from the expression for drain current given in Eq. (2.3-27). The primary temperature-dependent parameters are the mobility  $\mu$  and the threshold voltage  $V_T$ . The temperature dependence of the carrier mobility  $\mu$  is given as [28]

$$\mu = K_\mu T^{-1.5} \quad (2.5-8)$$

The temperature dependence of the threshold voltage can be approximated by the following expression [29]:

$$V_T(T) = V_T(T_0) - \alpha(T - T_0) \quad (2.5-9)$$

where  $\alpha$  is approximately 2.3 mV/°C. This expression is valid over the range of 200–400 K, with  $\alpha$  depending on the substrate doping level and the dosages of the implants used during fabrication. These expressions for the temperature dependence of mobility and threshold voltage will be used later to determine the temperature performance of MOS circuits and are valid only for limited ranges of temperature variation about room temperature. Other modifications are necessary for extreme temperature ranges.

The temperature dependence of the pn junction is also important in this study. For example, the pn-junction diode can be used to create a reference voltage whose temperature stability will depend on the temperature characteristics of the pn-junction diode. We shall consider the reverse-biased pn-junction diode first. Equation (2.2-24) shows that when  $v_D < 0$ , the diode current is given as

$$-i_D \cong I_s = qA \left[ \frac{D_p p_{n0}}{L_p} + \frac{D_n n_{p0}}{L_n} \right] \cong \frac{qAD}{L} \frac{n_i^2}{N} = KT^3 \exp\left(\frac{-V_{G0}}{V_t}\right) \quad (2.5-10)$$

where it has been assumed that one of the terms in the brackets is dominant and that  $L$  and  $N$  correspond to the diffusion length and impurity concentration of the dominant term. Also,  $T$  is the absolute temperature in kelvin and  $V_{G0}$  is the bandgap voltage of silicon at 300 K (1.205 V). Differentiating Eq. (2.5-10) with respect to  $T$  results in

$$\frac{dI_s}{dT} = \frac{3KT^3}{T} \exp\left(\frac{-V_{G0}}{V_t}\right) + \frac{qKT^3 V_{G0}}{KT^2} \exp\left(\frac{-V_{G0}}{V_t}\right) = \frac{3I_s}{T} + \frac{I_s V_{G0}}{T V_t} \quad (2.5-11)$$

The  $TC_F$  for the reverse diode current can be expressed as

$$\frac{1}{I_s} \frac{dI_s}{dT} = \frac{3}{T} + \frac{1}{T} \frac{V_{G0}}{V_t} \quad (2.5-12)$$

The reverse diode current is seen to double approximately every 5 °C increase as illustrated in the following example.

**Example**  
2.5-1**Calculation of the Reverse Diode Current Temperature Dependence and  $TC_F$** 

Assume that the temperature is 300 K (room temperature) and calculate the reverse diode current change and the  $TC_F$  for a 5 K increase.

The  $TC_F$  can be calculated from Eq. (2.5-12) as

$$TC_F = 0.01 + 0.155 = 0.165$$

Since the  $TC_F$  is change per degree, the reverse current will increase by a factor of 1.165 for every kelvin (or °C) change in temperature. Multiplying by 1.165 five times gives an increase of approximately 2. This implies that the reverse saturation current will approximately double for every 5 °C temperature increase. Experimentally, the reverse current doubles for every 8 °C increase in temperature because the reverse current is in part leakage current.

The forward-biased pn-junction diode current is given by

$$i_D \cong I_s \exp\left(\frac{v_D}{V_t}\right) \quad (2.5-13)$$

Differentiating this expression with respect to temperature and assuming that the diode voltage is a constant ( $v_D = v_D$ ) gives

$$\frac{di_D}{dT} = \frac{i_D}{I_s} \cdot \frac{dI_s}{dT} - \frac{1}{T} \cdot \frac{V_D}{V_t} i_D \quad (2.5-14)$$

The fractional temperature coefficient for  $i_D$  results from Eq. (2.5-14) as

$$\frac{1}{i_D} \cdot \frac{di_D}{dT} = \frac{1}{I_s} \cdot \frac{dI_s}{dT} - \frac{V_D}{TV_t} = \frac{3}{T} + \left(\frac{V_{G0} - V_D}{TV_t}\right) \quad (2.5-15)$$

If  $v_D$  is assumed to be 0.6 volts, then the fractional temperature coefficient is equal to  $0.01 + (0.155 - 0.077) = 0.0879$ . It can be seen that the forward diode current will double for approximately a 10 °C increase in temperature.

The above analysis for the forward-biased pn-junction diode assumed that the diode voltage  $v_D$  was held constant. If the forward current is held constant ( $i_D = I_D$ ), then the fractional temperature coefficient of the forward diode voltage can be found. From Eq. (2.5-13) we can solve for  $v_D$  to get

$$v_D = V_t \ln\left(\frac{I_D}{I_s}\right) \quad (2.5-16)$$

Differentiating Eq. (2.5-16) with respect to temperature gives

$$\frac{dv_D}{dT} = \frac{v_D}{T} - V_t \left(\frac{1}{I_s} \cdot \frac{dI_s}{dT}\right) = \frac{v_D}{T} - \frac{3V_t}{T} - \frac{V_{G0}}{T} = -\left(\frac{V_{G0} - v_D}{T}\right) - \frac{3V_t}{T} \quad (2.5-17)$$

Assuming that  $v_D = v_D = 0.6$  V, the temperature dependence of the forward diode voltage at room temperature is approximately  $-2.3$  mV/°C.

Another limitation of CMOS components is noise. Noise is a phenomenon caused by small fluctuations of the analog signal within the components themselves. Noise results from the fact that electrical charge is not continuous but the result of quantized behavior and is associated with the fundamental processes in a semiconductor component. In essence, noise acts like a random variable and is often treated as one. Our objective is to introduce the basic concepts concerning noise in CMOS components. More detail can be found in several excellent references [23,30].

Several sources of noise are important in CMOS components. *Shot noise* is associated with the dc current flow across a pn junction. It typically has the form of

$$i_n^2 = 2qI_D \Delta f \quad (\text{amperes}^2) \quad (2.5-18)$$

where  $i_n^2$  is the mean-square value of the noise current,  $q$  is the charge of an electron,  $I_D$  is the average dc current of the pn junction, and  $\Delta f$  is the bandwidth in Hertz. Noise-current spectral density can be found by dividing  $i_n^2$  by  $\Delta f$ . The noise-current spectral density is denoted as  $i_n^2/\Delta f$ .

Another source of noise, called *thermal noise*, is due to random thermal motion of the electron and is independent of the dc current flowing in the component. It generally has the form of

$$e_n^2 = 4kTR \Delta f \quad (2.5-19)$$

where  $k$  is Boltzmann's constant and  $R$  is the resistor or equivalent resistor in which the thermal noise is occurring.

An important source of noise for MOS components is the *flicker noise* or the *1/f noise*. This noise is associated with carrier traps in semiconductors, which capture and release carriers in a random manner. The time constants associated with this process give rise to a noise signal with energy concentrated at low frequency. The typical form of the *1/f* noise is given as

$$i_n^2 = K_f \left( \frac{I}{f^b} \right)^a \Delta f \quad (2.5-20)$$

where  $K_f$  is a constant,  $a$  is a constant (0.5–2), and  $b$  is a constant ( $\cong 1$ ). The current-noise spectral density for typical *1/f* noise is shown in Fig. 2.5-6. Other sources of noise exist, such as burst noise and avalanche noise, but are not important in CMOS components and are not discussed here.

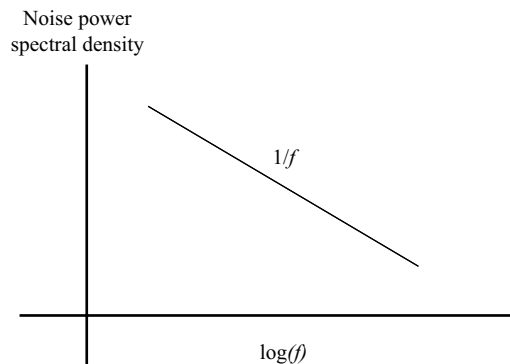


**Figure 2.5-6** *1/f* noise spectrum.



## 2.6 Summary

This chapter has introduced CMOS technology from the viewpoint of its use to implement analog circuits. The basic semiconductor fabrication processes were described in order to understand the fundamental elements of this technology. The basic fabrication steps include diffusion, implantation, deposition, etching, and oxide growth. These steps are implemented by the use of photolithographic methods, which limit the processing steps to certain physical areas of the silicon wafer. The basic processing steps needed to implement a typical twin-well, STI, silicon-gate CMOS process were described next.

The pn junction was reviewed following the introduction to CMOS technology because it plays an important role in all semiconductor devices. This review examined a step pn junction and developed the physical dimensions, the depletion capacitance, and the voltage–current characteristics of the pn junction. Next, the MOS transistor was introduced and characterized with respect to its behavior. It was shown how the channel between the source and drain is formed and the influence of the gate voltage on this channel was discussed. The MOS transistor is physically a very simple component.

A discussion of the passive components available in CMOS technology followed. These components include only resistors and capacitors. The various ways to implement these along with their merits were discussed.

The next section presented further considerations of CMOS technology. These included the substrate and lateral BJTs compatible with the CMOS process; electrical overstress and means for protecting against it; the temperature dependence of CMOS components; and the noise sources in these components.

Support material for this chapter covers the topic of device layout and is presented in Appendix B.

## Problems (refer to Appendix B for problems on layout)

- |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>2.1-1. List the five basic MOS fabrication processing steps and give the purpose or function of each step.</p> <p>2.1-2. What is the difference between positive and negative photoresist and how is photoresist used?</p> <p>2.1-3. Consider a mask that is opaque everywhere except for a transparent circle in the center. Metal is deposited on a substrate followed by an application of negative photoresist which is patterned with the mask described. After exposure, developing, and subsequent etching, what will remain?</p> <p>2.1-4. Illustrate the impact on source and drain diffusions of a <math>7^\circ</math> angle off-perpendicular ion implant. Assume that the thickness of polysilicon is <math>8000 \text{ \AA}</math> and that outdiffusion from point of ion impact is <math>0.07 \text{ }\mu\text{m}</math>.</p> | <p>2.1-5. Given the process flow described in Section 2.1, would it be possible to create a structure with polysilicon on top of a heavily doped diffusion separated by a thin oxide? Why or why not? Explain.</p> <p>2.2-1. Repeat Example 2.2-1 if the applied voltage is <math>-2 \text{ V}</math>.</p> <p>2.2-2. Develop Eq. (2.2-9) using Eqs. (2.2-1), (2.2-7), and (2.2-8) of the same section.</p> <p>2.2-3. Redevelop Eqs. (2.2-7) and (2.2-8) if the impurity concentration of a pn junction is given by Fig. 2.2-2 rather than the step junction of Fig. 2.2-1(b).</p> <p>2.2-4. Plot the normalized reverse current, <math>i_{RA}/i_R</math>, versus the reverse voltage <math>v_R</math> of a silicon pn diode that has <math>BV = 12 \text{ V}</math> and <math>n = 6</math>.</p> <p>2.2-5. What is the breakdown voltage of a pn junction with <math>N_A = N_D = 10^{16}/\text{cm}^3</math>?</p> |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

- 2.2-6. What change in  $v_D$  of a silicon pn diode will cause an increase of 10 (an order of magnitude) in the forward diode current?
- 2.3-1. Explain in your own words why the magnitude of the threshold voltage in Eq. (2.3-19) increases as the magnitude of the source-bulk voltage increases. (The source-bulk pn diode remains reverse biased.)
- 2.3-2. If  $V_{SB} = 2$  V, find the value of  $V_T$  for the n-channel transistor of Example 2.3-1.
- 2.3-3. Rederive Eq. (2.3-27) given that  $V_T$  is not constant in Eq. (2.3-22) but rather varies linearly with  $v(y)$  according to the following equation:

$$V_T = V_{T0} + \alpha v(y)$$

- 2.3-4. If the mobility of an electron is  $500 \text{ cm}^2/(\text{V} \cdot \text{s})$  and the mobility of a hole is  $200 \text{ cm}^2/(\text{V} \cdot \text{s})$ , compare the performance of an n-channel with a p-channel transistor. In particular, consider the value of the transconductance parameter and speed of the MOS transistor.
- 2.3-5. Using Example 2.3-1 as a starting point, calculate the difference in threshold voltage between two devices whose gate oxide is different by 5% (i.e.,  $t_{ox} = 210 \text{ \AA}$ ).
- 2.3-6. Repeat Example 2.3-1 using  $N_A = 7 \times 10^{16} \text{ cm}^{-3}$ , gate doping,  $N_D = 1 \times 10^{19} \text{ cm}^{-3}$ .
- 2.4-1. If a capacitor is needed to shunt high-frequency signals from an internal signal node to ground, which of the two capacitors illustrated in Fig. 2.4-1 would be best suited and why? Assume that both capacitors have the same capacitance per unit area.
- 2.4-2. A simple first-order filter shown in Fig. P2.4-2 is to be built with a polysilicon resistor and an MOS capacitor. The polysilicon resistor has a sheet resistance of  $50 \text{ } \Omega/\text{sq.} \pm 30\%$  and is  $5 \text{ } \mu\text{m}$  wide. The MOS capacitor is  $2 \text{ fF}/\mu\text{m}^2 \pm 10\%$ . The  $-3 \text{ dB}$  frequency of the low-pass filter is  $1 \text{ MHz}$ . (a) Choose the size of the resistor (the number of squares,  $N$ ) to minimize the total area of the filter including both the resistor and the capacitor. Find the area of the resistor and the capacitor in  $\mu\text{m}^2$  and their values. (b)

Using the worst-case tolerance of the resistor and capacitor, find the maximum and minimum  $-3 \text{ dB}$  frequencies.

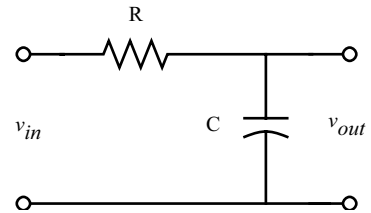


**Figure P2.4-2**

- 2.4-3. List two sources of error that can make the actual capacitor, fabricated using a CMOS process, differ from its designed value.
- 2.4-4. The capacitor shown in Fig. 2.4-1(a) has, in addition to the desired capacitance between the top and bottom plates, a capacitance between Metal 4 and Metal 5. Illustrate (sketch) a way to make such a capacitor, minimizing this additional capacitance due to M5–M4. For this problem, draw a square capacitor and use relative dimensions.
- 2.4-5. Consider the circuit in Fig. P2.4-5. Resistor  $R_1$  is an n-well resistor with a nominal value of  $10 \text{ k}\Omega$  when the voltage at both terminals is  $2.5 \text{ V}$ . The input voltage,  $v_{in}$ , is  $2.5 \text{ V}$ . Under these conditions, the value of  $R_1$  is given as

$$R_1 = R_{nom} \left[ 1 + K \left( \frac{v_{in} + v_{out}}{2} \right) \right]$$

where  $R_{nom}$  is  $10 \text{ K}$  and the coefficient  $K$  is the voltage coefficient of an n-well resistor and has a value of  $20 \text{ K ppm/V}$ . Resistor  $R_2$  is an ideal resistor with a value of  $10 \text{ k}\Omega$ . Calculate the value for  $v_{out}$ .

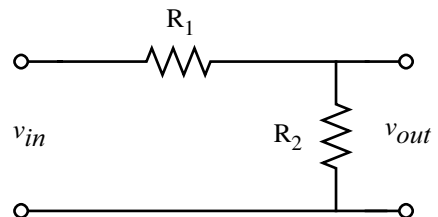
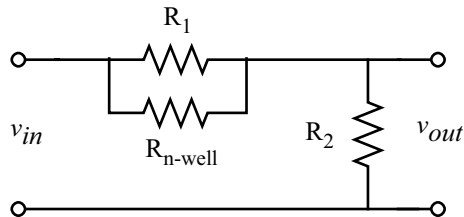


**Figure P2.4-5**

- 2.4-6. Repeat Problem 2.4-5 using a  $n^+$  diffused resistor for  $R_1$ . Assume that a  $n^+$  resistor's voltage coefficient is  $200 \text{ ppm/V}$ .

2.4-7. Consider Problem 2.4-6 again but assume that the n-well in which  $R_1$  lies is not connected to a 5 volt supply, but rather is connected as shown in Fig. P2.4-7. What effect does the n-well have on the voltage dependency of the diffused resistor?



2.4-8. Given the polysilicon resistor in Fig. P2.4-8 with a resistivity of  $\rho = 4 \times 10^{-3} \Omega\text{-cm}$ , calculate the resistance of the structure. Consider only the resistance defined by the silicide block.

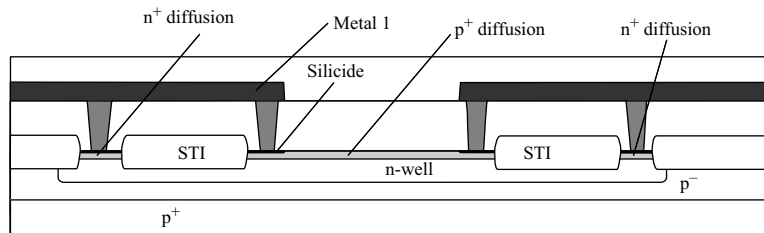


Figure P2.4-7

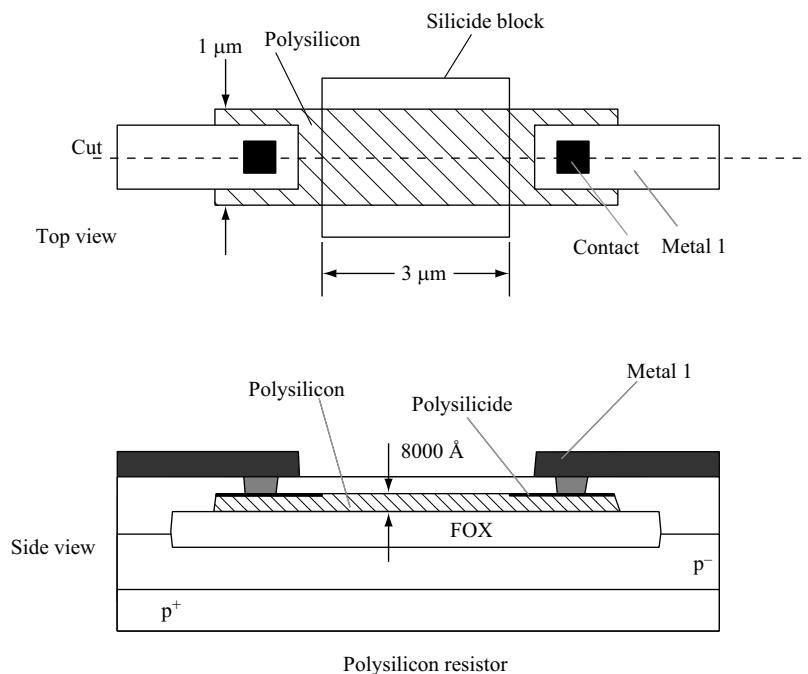


Figure P2.4-8

- 2.5-1. Assume  $v_D = 0.7$  V and find the fractional temperature coefficient of  $I_s$  and  $v_D$ .
- 2.5-2. Plot the noise voltage as a function of the frequency if the thermal noise is  $100 \text{ nV}/\sqrt{\text{Hz}}$  and the junction of the  $1/f$  and thermal noise (the  $1/f$  noise corner) is  $10,000$  Hz.

Refer to Appendix B for the following problems.

- A-1. Given that you wish to match two transistors having a  $W/L$  of  $40 \mu\text{m}/0.3 \mu\text{m}$  each, sketch the layout of these two transistors to achieve the best possible matching.
- A-2. Assume that the edge variation of the top plate of a capacitor is  $0.05 \mu\text{m}$  and that capacitor top plates are to be laid out as squares. It is desired to match two equal capacitors to an accuracy of  $0.1\%$ . Assume that there is no variation in oxide thickness.

- A-3. Show that a circular geometry minimizes the perimeter-to-area ratio for a given area requirement. In your proof, compare against a rectangle and square.
- A-4. Show analytically how the Yiannoulos-path technique illustrated in Fig. A-5 maintains a constant area-to-perimeter ratio with noninteger ratios.
- A-5. Design an optimal layout of a matched pair of transistors whose  $W/L$  is  $16 \mu\text{m}/0.5 \mu\text{m}$ . The matching should be a photolithographic invariant as well as a common centroid.
- A-6. Figure PA-6 illustrates various ways to implement the layout of a resistor divider. Choose the layout that BEST achieves the goal of a  $2:1$  ratio. Explain why the other choices are not optimal.

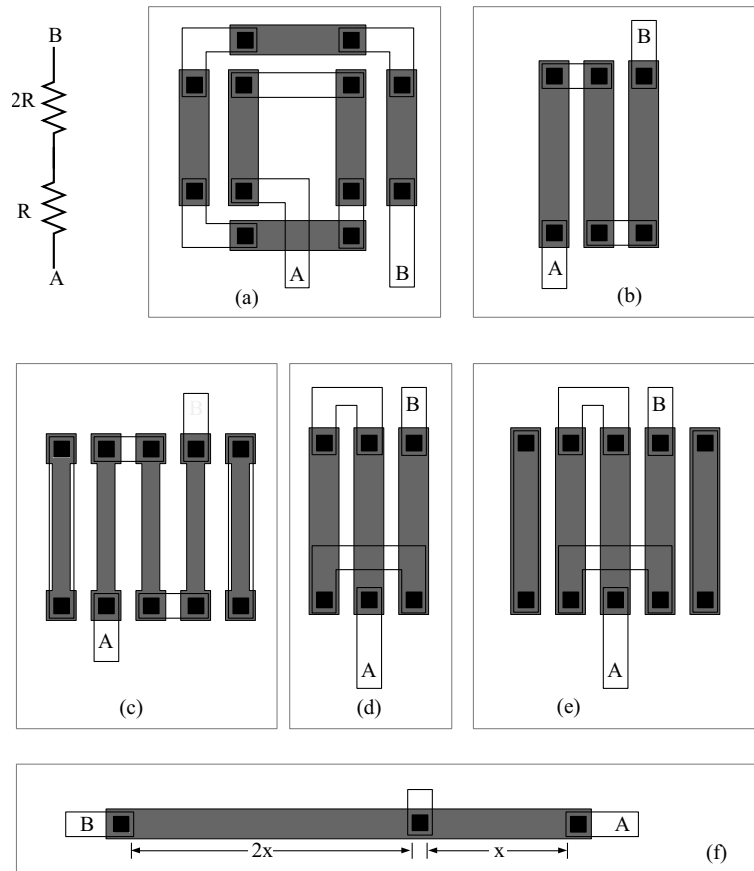


Figure PA-6

## References

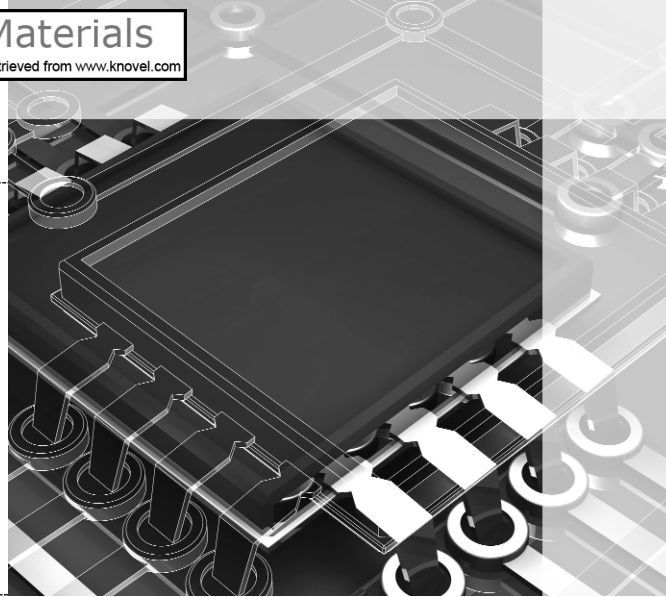
---

1. Y. P. Tsividis and P. R. Gray, "A Segmented  $\mu 255$  Law PCM Voice Encoder Utilizing NMOS Technology," *IEEE J. Solid-State Circuits*, Vol. SC-11, pp. 740–747, Dec. 1976.
2. B. Fotouhi and D. A. Hodges, "High-Resolution A/D Conversion in MOS/LSI," *IEEE J. Solid-State Circuits*, Vol. SC-14, pp. 920–926, Dec. 1979.
3. J. T. Caves, C. H. Chan, S. D. Rosenbaum, L. P. Sellers, and J. B. Terry, "A PCM Voice Codec with On-Chip Filters," *IEEE J. Solid-State Circuits*, Vol. SC-14, pp. 65–73, Feb. 1979.
4. Y. P. Tsividis and P. R. Gray, "An Integrated NMOS Operational Amplifier with Internal Compensation," *IEEE J. Solid-State Circuits*, Vol. SC-11, pp. 748–754, Dec. 1976.
5. B. K. Ahuja, P. R. Gray, W. M. Baxter, and G. T. Uehara, "A Programmable CMOS Dual Channel Interface Processor for Telecommunications Applications," *IEEE J. Solid-State Circuits*, Vol. SC-19, pp. 892–899, Dec. 1984.
6. H. Shirasu, M. Shibukawa, E. Amada, Y. Hasegawa, F. Fujii, K. Yasunari, and Y. Toba, "A CMOS SLIC with an Automatic Balancing Hybrid," *IEEE J. Solid-State Circuits*, Vol. SC-18, pp. 678–684, Dec. 1983.
7. A. S. Grove, *Physics and Technology of Semiconductor Devices*. New York: John Wiley & Sons, 1967.
8. R. S. Muller and T. I. Kamins, *Device Electronics for Integrated Circuits*. New York: John Wiley & Sons, 1977.
9. S. Wolf and R. N. Tauber, *Silicon Processing for the VLSI Era*, Vol. 1: *Process Technology*. Sunset Beach, CA: Lattice Press, 1986, pp. 5–11.
10. S. Wolf and R. N. Tauber, *Silicon Processing for the VLSI Era*, Vol. 1: *Process Technology*. Sunset Beach, CA: Lattice Press, 1986, p. 27.
11. J. C. Irvin, "Resistivity of Bulk Silicon and Diffused Layers in Silicon," *Bell Syst. Technical J.*, Vol. 41, pp. 387–410, Mar. 1962.
12. D. G. Ong, *Modern MOS Technology—Processes, Devices, & Design*. New York: McGraw-Hill, 1984, Chap. 8.
13. D. J. Hamilton and W. G. Howard, *Basic Integrated Circuit Engineering*. New York: McGraw-Hill, 1975, Chap. 2.
14. D. H. Lee and J. W. Mayer, "Ion Implanted Semiconductor Devices," *Proceedings of IEEE*, pp. 1241–1255, Sept. 1974.
15. J. F. Gibbons, "Ion Implantation in Semiconductors," *Proceedings of IEEE*, Part I, Vol. 56, pp. 295–319, March 1968; Part II, Vol. 60, pp. 1062–1096, Sept. 1972.
16. S. Wolf and R. N. Tauber, *Silicon Processing for the VLSI Era*, Vol. 1: *Process Technology*. Sunset Beach, CA: Lattice Press, 1986, pp. 374–381.
17. S. Wolf and R. N. Tauber, *Silicon Processing for the VLSI Era*, Vol. 1: *Process Technology*. Sunset Beach, CA: Lattice Press, 1986, pp. 335–374.
18. J. L. Vossen and W. Kern (eds.), *Thin Film Processes*, Part III-2. New York: Academic Press, 1978.
19. P. E. Gise and R. Blanchard, *Semiconductor and Integrated Circuit Fabrication Technique*. Reston, VA: Reston Publishers, 1979, Chaps. 5, 6, 10, and 12.
20. S. Wolf and R. N. Tauber, *Silicon Processing for the VLSI Era*, Vol. 2: *Process Integration*. Sunset Beach, CA: Lattice Press, 1990, pp. 238–239.
21. D. J. Elliot, *Integrated Circuit Fabrication Technology*. New York: McGraw-Hill, 1982.
22. S. Wolf and R. N. Tauber, *Silicon Processing for the VLSI Era*, Vol. 1: *Process Technology*. Sunset Beach, CA: Lattice Press, 1986, pp. 384–406.
23. P. R. Gray and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*. 2nd ed. New York: John Wiley & Sons, 1984, Chap. 1.
24. D. A. Hodges and H. G. Jackson, *Analysis and Design of Analog Integrated Circuits*. New York: McGraw-Hill, 1983.
25. B. R. Chawla and H. K. Gummel, "Transition Region Capacitance of Diffused pn Junctions," *IEEE Transactions on Electron Devices*, Vol. ED-18, pp. 178–195, Mar. 1971.
26. C. T. Sah, "Characteristics of the Metal-Oxide-Semiconductor Transistor," *IEEE Transactions on Electron Devices*, Vol. ED-11, pp. 324–345, July 1964.
27. H. Shichman and D. Hodges, "Modelling and Simulation of Insulated-Gate Field-Effect Transistor Switching Circuits," *IEEE J. Solid-State Circuits*, Vol. SC-13, pp. 285–289, Sept. 1968.

28. D. B. Estreich and R. W. Dutton, "Modeling Latch-Up in CMOS Integrated Circuits and Systems," *IEEE Transactions on CAD*, Vol. CAD-1, pp. 157–162, Oct. 1982.
29. S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed. New York: John Wiley & Sons, 1981, p. 28.
30. R. A. Blauschild, P. A. Tucci, R. S. Muller, and R. G. Meyer, "A New Temperature-Stable Voltage Reference," *IEEE J. Solid-State Circuits*, Vol. SC-19, Dec. 1978, pp. 767–774.
31. C. D. Motchenbacher and F. G. Fitchen, *Low-Noise Electronic Design*. New York: John Wiley & Sons, 1973.

## CHAPTER 3

# CMOS Device Modeling



**B**efore one can design a circuit to be integrated in CMOS technology, one must first have a model describing the behavior of all the components available for use in the design. A model can take the form of mathematical equations, circuit representations, or tables. Most of the modeling used in this text will focus on the active and passive devices discussed in the previous chapter as opposed to higher-level modeling such as macromodeling or behavioral modeling.

It should be stressed at the outset that a model is just that and no more—it is not the real thing! In an ideal world, we would have a model that accurately describes the behavior of a device under all possible conditions. Realistically, we are happy to have a model that predicts simulated performance to within a few percent of measured performance. There is no clear agreement as to which model comes closest to meeting this “ideal” model [1]. This lack of agreement is illustrated by the fact that, at this writing, HSPICE [2] offers the user 43 different MOS transistor models from which to choose!

This text will concentrate on only three of these models. The simplest model, which is appropriate for hand calculations, was described in Section 2.3 and will be further developed here to include capacitance, noise, and ohmic resistance. In SPICE terminology, this simple model is called the LEVEL 1 model. Next, a small-signal model is derived from the LEVEL 1 large-signal model and is presented in Section 3.3.

A far more complex model, the SPICE LEVEL 3 model, is presented in Section 3.4. This model includes many effects that are more evident in modern short-channel technologies as well as subthreshold conduction. It is adequate for device geometries down to about 0.8  $\mu\text{m}$ . Finally, the BSIM3v3 [3] model is presented. This model is the closest to becoming a standard for computer simulation.

### Notation

SPICE was originally implemented in FORTRAN where all input was required to be uppercase ASCII characters. Lowercase, Greek, and super-subscripting were not allowed. Modern SPICE implementations generally accept (but do not distinguish between) uppercase and lowercase but the tradition of using uppercase ASCII still lives on. This is particularly evident in the device model parameters. Since Greek characters are not available, these were simply spelled out, for example,  $\gamma$  entered as GAMMA. Superscripts and subscripts were simply not used.

It is inconvenient to adopt the SPICE naming convention throughout the book because equations would appear unruly and would not be familiar to what is commonly seen in the literature. On the other hand, it is necessary to provide the correct notation where application to SPICE is intended. To address this dilemma, we have decided to use SPICE uppercase (nonitalic) notation for all model parameters except those applied to the simple model (SPICE LEVEL 1).



### 3.1 Simple MOS Large-Signal Model (SPICE LEVEL 1)

All large-signal models will be developed for the n-channel MOS device with the positive polarities of voltages and currents shown in Fig. 3.1-1(a). The same models can be used for the p-channel MOS device if all voltages and currents are multiplied by  $-1$  and the absolute value of the p-channel threshold is used. This is equivalent to using the voltages and currents defined by Fig. 3.1-1(b), which are all positive quantities. As mentioned in Chapter 1, lower-case variables with capital subscripts will be used for the variables of large-signal models and lowercase variables with lowercase subscripts will be used for the variables of small-signal models. When the voltage or current is a model parameter, such as threshold voltage, it will be designated by an uppercase variable and an uppercase subscript.

When the length and width of the MOS device is greater than about  $10\text{ }\mu\text{m}$ , the substrate doping is low, and when a simple model is desired, the model suggested by Sah [4] and used in SPICE by Shichman and Hodges [5] is very appropriate. This model was developed in Eq. (2.3-27) and is given below.

$$i_D = \frac{\mu_0 C_{ox} W}{L} \left[ (v_{GS} - V_T) - \left( \frac{v_{DS}}{2} \right) \right] v_{DS} \quad (3.1-1)$$

The terminal voltages and currents have been defined in the previous chapter. The various parameters of Eq. (3.1-1) are defined as

$\mu_0$  = surface mobility of the channel for the n-channel  
or p-channel device ( $\text{cm}^2/\text{V}\cdot\text{s}$ )

$C_{ox} = \frac{\epsilon_{ox}}{t_{ox}}$  = capacitance per unit area of the gate oxide ( $\text{F}/\text{cm}^2$ )

$W$  = effective channel width

$L$  = effective channel length

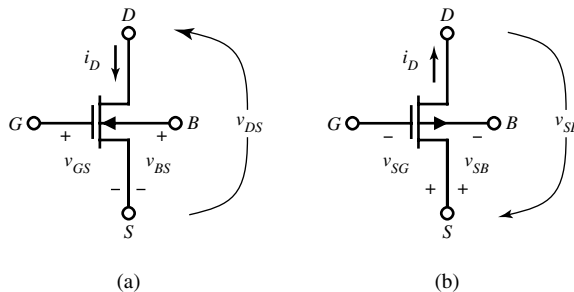


**Figure 3.1-1** Positive sign convention for (a) n-channel and (b) p-channel MOS transistor.

The threshold voltage  $V_T$  is given by Eq. (2.3-19) for an n-channel transistor:

$$V_T = V_{T0} + \gamma \left( \sqrt{2|\phi_F| + v_{SB}} - \sqrt{2|\phi_F|} \right) \quad (3.1-2)$$

$$V_{T0} = V_T(v_{SB} = 0) = V_{FB} + 2|\phi_F| + \frac{\sqrt{2q\epsilon_{Si}N_{SUB}2|\phi_F|}}{C_{ox}} \quad (3.1-3)$$

$$\gamma = \text{bulk threshold parameter (V}^{1/2}\text{)} = \frac{\sqrt{2\epsilon_{Si}qN_{SUB}}}{C_{ox}} \quad (3.1-4)$$

$$\phi_F = \text{strong inversion surface potential (V)} = \frac{kT}{q} \ln \left( \frac{N_{SUB}}{n_i} \right) \quad (3.1-5)$$

$$V_{FB} = \text{flatband voltage (V)} = \phi_{MS} - \frac{Q_{ss}}{C_{ox}} \quad (3.1-6)$$

$$\phi_{MS} = \phi_F(\text{substrate}) - \phi_F(\text{gate}) \quad [\text{see Eq. (2.3-17)}] \quad (3.1-7)$$

$$\phi_F(\text{substrate}) = -\frac{kT}{q} \ln \left( \frac{N_{SUB}}{n_i} \right) [\text{n-channel with p-substrate}] \quad (3.1-8)$$

$$\phi_F(\text{gate}) = -\frac{kT}{q} \ln \left( \frac{N_{GATE}}{n_i} \right) [\text{n-channel with n}^+\text{ polysilicon gate}] \quad (3.1-9)$$

$$Q_{ss} = \text{oxide-charge} = qN_{ss} \quad (3.1-10)$$

$k$  = Boltzmann's constant

$T$  = temperature (K)

$n_i$  = intrinsic carrier concentration

Table 3.1-1 gives some of the pertinent constants for silicon.

A unique aspect of the MOS device is its dependence on the voltage from the source to bulk as shown by Eq. (3.1-2). This dependence means that the MOS device must be treated as a four-terminal element. It will be shown later how this behavior can influence both the large- and small-signal performance of MOS circuits.

**Table 3.1-1 Constants for Silicon**

| Constant Symbol | Constant Description                    | Value                   | Units            |
|-----------------|-----------------------------------------|-------------------------|------------------|
| $V_{G0}$        | Silicon bandgap (27 °C)                 | 1.205                   | V                |
| $k$             | Boltzmann's constant                    | $1.381 \times 10^{-23}$ | J/K              |
| $n_i$           | Intrinsic carrier concentration (27 °C) | $1.45 \times 10^{10}$   | cm <sup>-3</sup> |
| $\epsilon_0$    | Permittivity of free space              | $8.854 \times 10^{-14}$ | F/cm             |
| $\epsilon_{Si}$ | Permittivity of silicon                 | $11.7 \epsilon_0$       | F/cm             |
| $\epsilon_{ox}$ | Permittivity of SiO <sub>2</sub>        | $3.9 \epsilon_0$        | F/cm             |

**Table 3.1-2 Model Parameters for a Typical CMOS Bulk Process Suitable for Hand Calculations Using the Simple Model with Values Based on a 0.8  $\mu\text{m}$  Silicon-Gate Bulk CMOS n-Well Process**

| Parameter<br>Symbol | Parameter Description                         | Typical Parameter Value                                          |                                                                  | Units                    |
|---------------------|-----------------------------------------------|------------------------------------------------------------------|------------------------------------------------------------------|--------------------------|
|                     |                                               | n-Channel                                                        | p-Channel                                                        |                          |
| $V_{T0}$            | Threshold voltage ( $V_{BS} = 0$ )            | $0.7 \pm 0.15$                                                   | $-0.7 \pm 0.15$                                                  | V                        |
| $K'$                | Transconductance parameter<br>(in saturation) | $110.0 \pm 10\%$                                                 | $50.0 \pm 10\%$                                                  | $\mu\text{A}/\text{V}^2$ |
| $\gamma$            | Bulk threshold parameter                      | 0.4                                                              | 0.57                                                             | $\text{V}^{1/2}$         |
| $\lambda$           | Channel length modulation<br>parameter        | $0.04$ ( $L = 1 \mu\text{m}$ )<br>$0.01$ ( $L = 2 \mu\text{m}$ ) | $0.05$ ( $L = 1 \mu\text{m}$ )<br>$0.01$ ( $L = 2 \mu\text{m}$ ) | $\text{V}^{-1}$          |
| $2 \phi_F $         | Surface potential at strong<br>inversion      | 0.7                                                              | 0.8                                                              | V                        |

In the realm of circuit design, it is more desirable to express the model equations in terms of electrical rather than physical parameters. For this reason, the drain current is often expressed as

$$i_D = \beta \left[ (v_{GS} - V_T) - \frac{v_{DS}}{2} \right] v_{DS} \quad (3.1-11)$$

or

$$i_D = K' \frac{W}{L} \left[ (v_{GS} - V_T) - \frac{v_{DS}}{2} \right] v_{DS} \quad (3.1-12)$$

where the transconductance parameter  $\beta$  is given in terms of physical parameters as

$$\beta = K' \frac{W}{L} \cong \mu_0 C_{ox} \frac{W}{L} \quad (\text{A}/\text{V}^2) \quad (3.1-13)$$

When devices are characterized in the nonsaturation region with low gate and drain voltages, the value for  $K'$  is approximately equal to  $\mu_0 C_{ox}$  in the simple model. This is not the case when devices are characterized with larger voltages introducing effects such as mobility degradation. For these latter cases,  $K'$  is usually smaller. Typical values for the model parameters of Eq. (3.1-12) are given in Table 3.1-2.

There are various regions of operation of the MOS transistor based on the model of Eq. (3.1-1). These regions of operation depend on the value of  $v_{GS} - V_T$ . If  $v_{GS} - V_T$  is zero or negative, then the MOS device is in the cutoff\* region and Eq. (3.1-1) becomes

$$i_D = 0, \quad v_{GS} - V_T \leq 0 \quad (3.1-14)$$

In this region, the channel acts like an open circuit.

---

\*We will learn later that MOS transistors can operate in the subthreshold region where the gate-source voltage is less than the threshold voltage.

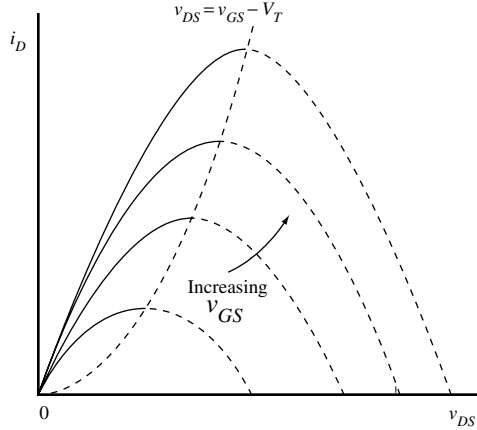


**Figure 3.1-2** Graphical illustration of the modified Sah equation.

A plot of Eq. (3.1-1) with  $\lambda = 0$  as a function of  $v_{DS}$  is shown in Fig. 3.1-2 for various values of  $v_{GS} - V_T$ . At the maximum of these curves the MOS transistor is said to saturate. The value of  $v_{DS}$  at which this occurs is called the saturation voltage and is given as

$$v_{DS}(\text{sat}) = v_{GS} - V_T \quad (3.1-15)$$

Thus,  $v_{DS}(\text{sat})$  defines the boundary between the remaining two regions of operation. If  $v_{DS}$  is less than  $v_{DS}(\text{sat})$ , then the MOS transistor is in the nonsaturated region and Eq. (3.1-1) becomes

$$i_D = K' \frac{W}{L} \left[ (v_{GS} - V_T) - \frac{v_{DS}}{2} \right] v_{DS}, \quad 0 < v_{DS} \leq (v_{GS} - V_T) \quad (3.1-16)$$

In Fig. 3.1-2, the nonsaturated region lies between the vertical axis ( $v_{DS} = 0$ ) and the  $v_{DS} = v_{GS} - V_T$  curve.

The third region occurs when  $v_{DS}$  is greater than  $v_{DS}(\text{sat})$  or  $v_{GS} - V_T$ . At this point the current  $i_D$  becomes independent of  $v_{DS}$ . Therefore,  $v_{DS}$  in Eq. (3.1-1) is replaced by  $v_{DS}(\text{sat})$  of Eq. (3.1-15) to get

$$i_D = K' \frac{W}{2L} (v_{GS} - V_T)^2, \quad 0 < (v_{GS} - V_T) \leq v_{DS} \quad (3.1-17)$$

Equation (3.1-17) indicates that drain current remains constant once  $v_{DS}$  is greater than  $v_{GS} - V_T$ . In reality, this is not true. As drain voltage increases, the channel length is reduced, resulting in increased current. This phenomenon is called *channel length modulation* and is accounted for in the saturation model with the addition of the factor  $(1 + \lambda v_{DS})$ , where  $v_{DS}$  is the actual drain-source voltage and not  $v_{DS}(\text{sat})$ . The saturation region model modified to include channel length modulation is given in Eq. (3.1-18):

$$i_D = K' \frac{W}{2L} (v_{GS} - V_T)^2 (1 + \lambda v_{DS}), \quad 0 < (v_{GS} - V_T) \leq v_{DS} \quad (3.1-18)$$

Up to this point in time, it has been convention to refer to the various regions of the MOS transistor as cutoff when  $(v_{GS} - V_T) \leq 0$ ; as active, ohmic, triode, or nonsaturation when  $0 \leq v_{DS} \leq (v_{GS} - V_T)$ ; and as saturation when  $(v_{GS} - V_T) \leq v_{DS}$ . While we will continue to use this convention throughout this text, the reader needs to be aware that a different convention

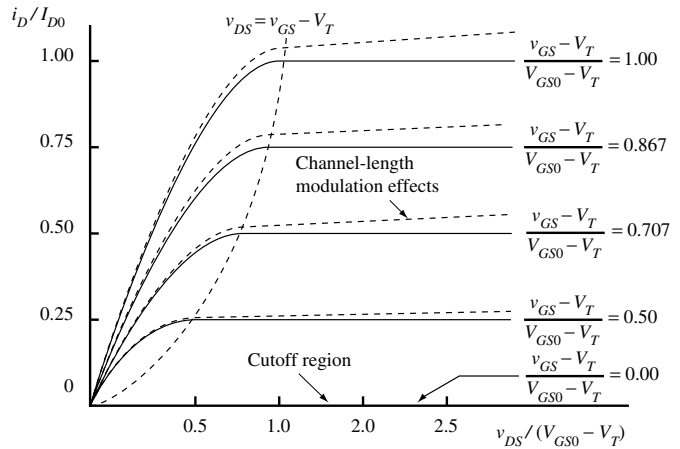


**Figure 3.1-3** Output characteristics of the MOS device.

was introduced recently in the fourth edition of Gray, Hurst, Lewis, and Meyer [6] that defines cutoff when  $(v_{GS} - V_T) \leq 0$ , triode or ohmic when  $0 \leq v_{DS} \leq (v_{GS} - V_T)$ , and active when  $(v_{GS} - V_T) \leq v_{DS}$ . The important difference is that “active” has replaced “saturation” when describing the region when  $(v_{GS} - V_T) \leq v_{DS}$ . This new convention makes the names of the regions of operation for an MOS transistor similar to the bipolar transistor.

The output characteristics of the MOS transistor can be developed from Eqs. (3.1-14), (3.1-16), and (3.1-18). Figure 3.1-3 shows these characteristics plotted on a normalized basis. These curves have been normalized to the upper curve, where  $V_{GS0}$  is defined as the value of  $v_{GS}$  that causes a drain current of  $I_{D0}$  in the saturation region. The entire characteristic is developed by extending the solid curves of Fig. 3.1-2 horizontally to the right from the maximum points. The solid curves of Fig. 3.1-3 correspond to  $\lambda = 0$ . If  $\lambda \neq 0$ , then the curves are the dashed lines.

Another important characteristic of the MOS transistor can be obtained by plotting  $i_D$  versus  $v_{GS}$  using Eq. (3.1-18). Figure 3.1-4 shows this result. This characteristic of the MOS

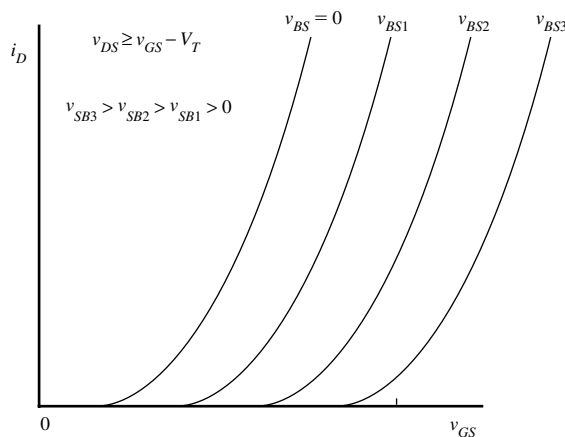


**Figure 3.1-4** Transconductance characteristic of the MOS transistor as a function of the source–bulk voltage,  $v_{SB}$ .

transistor is called the transconductance characteristic. We note that the transconductance characteristic in the saturation region can be obtained from Fig. 3.1-3 by drawing a vertical line to the right of the parabolic dashed line and plotting values of  $i_D$  versus  $v_{GS}$ . Figure 3.1-4 is also useful for illustrating the effect of the source–bulk voltage,  $v_{SB}$ . As the value of  $v_{SB}$  increases, the value of  $V_T$  increases for the enhancement, n-channel devices (for a p-channel device,  $|V_T|$  increases as  $v_{BS}$  increases).  $V_T$  also increases positively for the n-channel depletion device, but since  $V_T$  is negative, the value of  $V_T$  approaches zero from the negative side. If  $v_{SB}$  is large enough,  $V_T$  will actually become positive and the depletion device becomes an enhancement device.

Since the MOS transistor is a bidirectional device, determining which physical node is the drain and which the source may seem arbitrary. This is not really the case. For an n-channel transistor, the source is always at the lower potential of the two nodes. For the p-channel transistor, the source is always at the higher potential. It is obvious that the drain and source designations are not constrained to a given node of a transistor but can switch back and forth depending on the terminal voltages applied to the transistor.

A circuit version of the large-signal model of the MOS transistor consists of a current source connected between the drain and source terminals, that depends on the drain, source, gate, and bulk terminal voltages defined by the simple model described in this section. This simple model has five electrical and process parameters that completely define it. These parameters are  $K'$ ,  $V_T$ ,  $\gamma$ ,  $\lambda$ , and  $2\phi_F$ . The subscript  $n$  or  $p$  will be used when the parameter refers to an n-channel or p-channel device, respectively. They constitute the LEVEL 1 model parameters of SPICE [7]. Typical values for these model parameters are given in Table 3.1-2.

The function of the large-signal model is to solve for the drain current given the terminal voltages of the MOS device. An example will help to illustrate this as well as show how the model is applied to the p-channel device.

### Example 3.1-1

#### Application of the Simple MOS Large-Signal Model

Assume that the transistors in Fig. 3.1-1 have a  $W/L$  ratio of  $5\ \mu\text{m}/1\ \mu\text{m}$  and that the large-signal model parameters are those given in Table 3.1-2. If the drain, gate, source, and bulk voltages of the n-channel transistor are 3 V, 2 V, 0 V, and 0 V, respectively, find the drain current. Repeat for the p-channel transistor if the drain, gate, source, and bulk voltages are  $-3\ \text{V}$ ,  $-2\ \text{V}$ , 0 V, and 0 V, respectively.

#### SOLUTION

We must first determine in which region the transistor is operating. Equation (3.1-15) gives  $v_{DS}(\text{sat})$  as  $2\ \text{V} - 0.7\ \text{V} = 1.3\ \text{V}$ . Since  $v_{DS}$  is 3 V, the n-channel transistor is in the saturation region. Using Eq. (3.1-18) and the values from Table 3.1-2, we have

$$\begin{aligned} i_D &= \frac{K'_N W}{2L} (v_{GS} - V_{TN})^2 (1 + \lambda_N v_{DS}) \\ &= \frac{110 \times 10^{-6} (5\ \mu\text{m})}{2(1\ \mu\text{m})} (2 - 0.7)^2 (1 + 0.04 \times 3) = 520\ \mu\text{A} \end{aligned}$$

Evaluation of Eq. (3.1-15) for the p-channel transistor is given as

$$v_{SD}(\text{sat}) = v_{SG} - |V_{TP}| = 2\ \text{V} - 0.7\ \text{V} = 1.3\ \text{V}$$

Since  $v_{SD}$  is 3 V, the p-channel transistor is also in the saturation region, and Eq. (3.1-17) is applicable. The drain current of Fig. 3.1-1(b) can be found using the values from Table 3.1-2 as

$$\begin{aligned} i_D &= \frac{K'_p W}{2L} (v_{SG} - |V_{TP}|)^2 (1 + \lambda_p v_{SD}) \\ &= \frac{50 \times 10^{-6} (5 \mu\text{m})}{2(1 \mu\text{m})} (2 - 0.7)^2 (1 + 0.05 \times 3) = 243 \mu\text{A} \end{aligned}$$

It is often useful to describe  $v_{GS}$  in terms of  $i_D$  in saturation as shown below:

$$v_{GS} = V_T + \sqrt{2i_D/\beta} \quad (3.1-19)$$

This expression illustrates that there are two components to  $v_{GS}$ —an amount to invert the channel plus an additional amount to support the desired drain current. This second component is often referred to in the literature as  $V_{ON}$  or overdrive. Thus,  $V_{ON}$  can be defined as

$$V_{ON} = \sqrt{2i_D/\beta} \quad (3.1-20)$$

The term  $V_{ON}$  should be recognized as the term for saturation voltage  $V_{DS}$  (sat). They can be used interchangeably.

As the channel length of the technology decreases, the carriers in the channel will experience velocity saturation and the velocity is no longer proportional to the electric field [8]. An approximation for the carrier velocity is given as

$$v_d \approx \frac{\mu E}{1 + \frac{E}{E_c}} \quad (3.1-21)$$

where  $E_c$  is the critical electrical field at which velocity saturation of the carrier occurs. Eq. (2.3-25) can be expressed as

$$i_D = WQ_I(y)\mu_n \frac{dv(y)}{dy} = WQ_I(y)\mu_n E(y) = WQ_I(y)\mu_n v_d \quad (3.1-22)$$

If we substitute Eq. (3.1-21) into Eq. (3.1-22) and replace the electric field  $E$  by  $dv/dy$ , we obtain

$$i_D \left( 1 + \frac{1}{E_c} \frac{dv}{dy} \right) dy = WQ_I(y)\mu_n dv(y) \quad (3.1-23)$$

which is the form of Eq. (2.3-25) that includes velocity saturation for electrons. Integrating the distance and voltage along the channel from 0 to  $L$  and 0 to  $v_{DS}$  as done in Section 2.4 gives

$$i_D = \frac{\mu_n C_{ox}}{2 \left( 1 + \frac{1}{E_c} \frac{v_{DS}}{L} \right)} \frac{W}{L} [2(v_{GS} - V_T)v_{DS} - v_{DS}^2] \quad (3.1-24)$$

which is equivalent to Eq. (2.3-27) but incorporates the velocity saturation effect. It is convenient to define a constant  $\theta$  as

$$\theta = \frac{1}{E_c L} (V^{-1}) \quad (3.1-25)$$

Using this definition, Eq. (3.1-24) can be rewritten as

$$i_D = \frac{\mu_n C_{ox}}{2(1 + \theta v_{DS})} \frac{W}{L} [2(v_{GS} - V_T)v_{DS} - v_{DS}^2] \quad (3.1-26)$$

This expression is also equivalent to Eq. (3.1-1) but includes the influence of velocity saturation.

The next step, which follows the previous development in this section, is to solve for  $v_{DS}(\text{sat})$  as was done in Eq. (3.1-15). Unfortunately, this is not very straightforward because of  $v_{DS}$  in the denominator of Eq. (3.1-26). An approximation for  $v_{DS}(\text{sat})$  is given as [6]

$$\begin{aligned} v'_{DS}(\text{sat}) &\approx (V_{GS} - V_T) \left( 1 - \frac{\theta(V_{GS} - V_T)}{2} + \dots \right) \\ &= v_{DS}(\text{sat}) \left( 1 - \frac{\theta(V_{GS} - V_T)}{2} + \dots \right) \end{aligned} \quad (3.1-27)$$

The large-signal model for the saturation region that includes velocity saturation can be developed by assuming that  $0.5\theta(v_{GS} - V_T) < 1$  so that Eq. (3.1-27) reduces to the definition of  $v_{DS}(\text{sat})$ . With this assumption, Eq. (3.1-26) becomes

$$i_D = \frac{\mu_n C_{ox}}{2[1 + \theta(v_{GS} - V_T)]} \frac{W}{L} [(v_{GS} - V_T)^2], v_{DS} \geq (V_{GS} - V_T) \left( 1 - \frac{\theta(V_{GS} - V_T)}{2} + \dots \right) \quad (3.1-28)$$

which extends the large-signal model of Eq. (3.1-17) to include the effect of velocity saturation. Figure 3.1-5 shows the influence of the  $\theta$  parameter on the transconductance characteristics of an n-channel MOSFET with  $K' = 110 \mu\text{A}/\text{V}^2$  and  $W/L = 1$ . We see that the effect of velocity saturation (when  $\theta \neq 0$ ) causes the slope of the transconductance curve to decrease. The transconductance characteristic moves from square law to linear as  $\theta$  increases from zero.

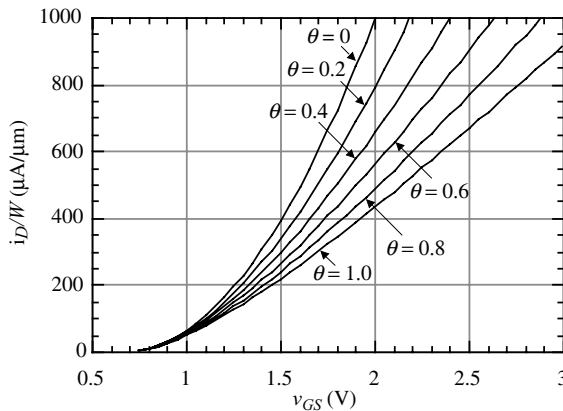


**Figure 3.1-5** Influence of velocity saturation on the transconductance characteristics.

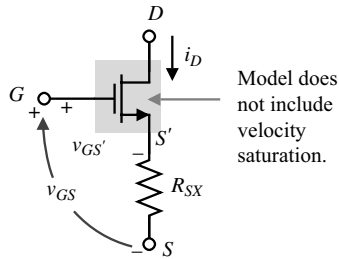


**Figure 3.1-6** n-Channel MOSFET with a degeneration resistor  $R_{SX}$  to model velocity saturation.

A simple way of modeling velocity saturation is found by considering degeneration of the MOSFET [6]. Degeneration is the insertion of a resistor in series with the source as illustrated in Fig. 3.1-6. In terms of this figure, we can write

$$i_D = \frac{K'W}{2L}(v_{GS}' - V_T)^2 \quad (3.1-29)$$

where

$$v_{GS}' = v_{GS} - i_D R_{SX} \quad (3.1-30)$$

Solving for  $i_D$  by substituting Eq. (3.1-30) into Eq. (3.1-29) results in

$$i_D = \frac{K'}{2 \left[ 1 + K' \frac{W}{L} R_{SX} (v_{GS} - V_T) \right]} \frac{W}{L} (v_{GS} - V_T)^2 \quad (3.1-31)$$

Comparing Eq. (3.1-31) with Eq. (3.1-28) gives the result that

$$R_{SX} = \frac{\theta L}{K'W} = \frac{1}{E_c K'W} \quad (3.1-32)$$

Therefore, given the critical field,  $E_c$ , one can calculate a value of  $R_{SX}$  that can be inserted in series with the source of the MOSFET to model velocity saturation. Therefore, given  $E_c$ ,  $K'$ , and  $W$ , we can calculate a resistor to be placed in series with the source of the MOS transistor to model velocity saturation using the MOS transistor model that does not include velocity saturation.

### Example 3.1-2

#### Application of the Simple MOS Large-Signal Model for Velocity Saturation

Assume that the transistor in Fig. 3.1-6 has a  $W/L$  ratio of  $1 \mu\text{m}/0.18 \mu\text{m}$  and that the large-signal model parameters are those given in Table 3.1-2. Assume also that this transistor experiences velocity saturation where  $E_c$  is  $1.5 \times 10^6 \text{ V/m}$ . Find the value of  $R_{SX}$ .

#### SOLUTION

The value of  $R_{SX}$  can be calculated as

$$R_{SX} = \frac{1}{E_c K'W} = \frac{1}{1.5 \times 10^6 \cdot 110 \times 10^{-6} \cdot 1 \times 10^{-6}} = 6.06 \text{ k}\Omega$$



## 3.2 Other MOS Large-Signal Model Parameters

The large-signal model also includes several other characteristics such as the source/drain bulk junctions, source/drain ohmic resistances, various capacitors, noise, and temperature dependence. The complete version of the large-signal model is given in Fig. 3.2-1.

The diodes of Fig. 3.2-1 represent the pn junctions between the source and substrate and the drain and substrate. For proper transistor operation, these diodes must always be reverse biased. Their purpose in the dc model is primarily to model leakage currents. These currents are expressed as

$$i_{BD} = I_s \left[ \exp\left(\frac{qV_{BD}}{kT}\right) - 1 \right] \quad (3.2-1)$$

and

$$i_{BS} = I_s \left[ \exp\left(\frac{qV_{BS}}{kT}\right) - 1 \right] \quad (3.2-2)$$

where  $I_s$  is the reverse saturation current of a pn junction,  $q$  is the charge of an electron,  $k$  is Boltzmann's constant, and  $T$  is temperature in kelvin units.

The resistors  $r_D$  and  $r_S$  represent the ohmic resistance of the drain and source, respectively. Typically, these resistors may be 50–100  $\Omega$ \* and can often be ignored at low drain currents.

The capacitors of Fig. 3.2-1 can be separated into three types. The first type includes capacitors  $C_{BD}$  and  $C_{BS}$ , which are associated with the backbiased depletion region between

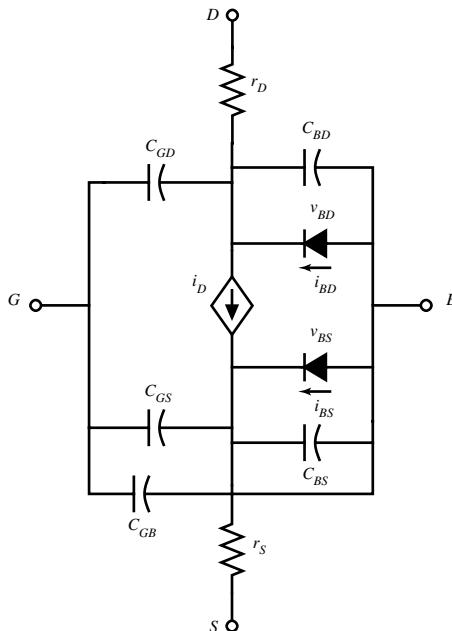


**Figure 3.2-1** Complete large-signal model for the MOS transistor.

\*For silicide process, these resistances will be much less—on the order of 5–10  $\Omega$ .

the drain and substrate and the source and substrate. The second type includes capacitors  $C_{GD}$ ,  $C_{GS}$ , and  $C_{GB}$ , which are all common to the gate and are dependent on the operating condition of the transistor. The third type includes parasitic capacitors, which are independent of the operating conditions.

The depletion capacitors are a function of the voltage across the pn junction. The expression of this junction-depletion capacitance is divided into two regions to account for the high injection effects. The first is given as

$$C_{BX} = (CJ)(AX) \left[ 1 - \frac{v_{BX}}{PB} \right]^{-MJ}, \quad v_{BX} \leq (FC)(PB) \quad (3.2-3)$$

where

$X = D$  for  $C_{BD}$  or  $X = S$  for  $C_{BS}$

$AX$  = area of the source ( $X = S$ ) or drain ( $X = D$ )

$CJ$  = zero-bias ( $v_{BX} = 0$ ) junction capacitance (per unit area)

$$CJ \cong \sqrt{\frac{q\epsilon_{Si}N_{SUB}}{2PB}}$$

$PB$  = bulk junction potential [same as  $\phi_0$  given in Eq. (2.2-6)]

$FC$  = forward-bias nonideal junction-capacitance coefficient ( $\cong 0.5$ )

$MJ$  = bulk junction grading coefficient ( $\frac{1}{2}$  for step junctions and  $\frac{1}{3}$  for graded junctions)

The second region is given as

$$C_{BX} = \frac{(CJ)(AX)}{(1 - FC)^{1+MJ}} \left[ 1 - (1 + MJ)FC + MJ \frac{v_{BX}}{PB} \right], \quad v_{BX} > (FC)(PB) \quad (3.2-4)$$

Figure 3.2-2 illustrates how the junction-depletion capacitances of Eqs. (3.2-3) and (3.2-4) are combined to model the large-signal capacitances  $C_{BD}$  and  $C_{BS}$ . It is seen that Eq. (3.2-4) prevents  $C_{BX}$  from approaching infinity as  $v_{BX}$  approaches  $PB$ .

A closer examination of the depletion capacitors in Fig. 3.2-3 shows that this capacitor is like a tub. It has a bottom with an area equal to the area of the drain or source. However, there are the sides that are also part of the depletion region. This area is called the sidewall.  $AX$  in Eqs. (3.2-3) and (3.2-4) should include both the bottom and sidewall assuming the zero-bias capacitances of the two regions are similar. To more closely model the depletion capacitance, it is separated into the bottom and sidewall components, given as follows:

$$C_{BX} = \frac{(CJ)(AX)}{\left[ 1 - \left( \frac{v_{BX}}{PB} \right) \right]^{MJ}} + \frac{(CJSW)(PX)}{\left[ 1 - \left( \frac{v_{BX}}{PB} \right) \right]^{MJSW}}, \quad v_{BX} \leq (FC)(PB) \quad (3.2-5)$$

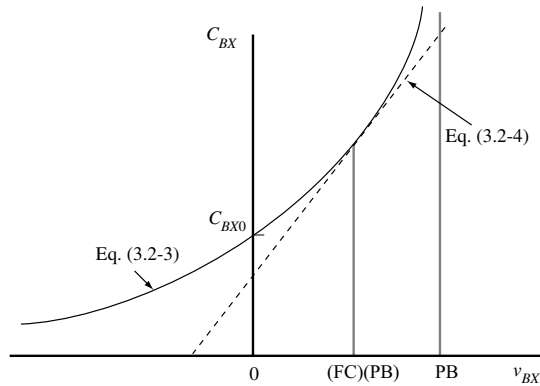


**Figure 3.2-2** Example of the method of modeling the voltage dependence of the bulk junction capacitances.

and

$$C_{BX} = \frac{(CJ)(AX)}{(1 - FC)^{1+MJ}} \left[ 1 - (1 + MJ)FC + MJ \frac{v_{BX}}{PB} \right] + \frac{(CJSW)(PX)}{(1 - FC)^{1+MJSW}} \left[ 1 - (1 + MJSW)FC + \frac{v_{BX}}{PB} (MJSW) \right],$$

$$v_{BX} \geq (FC)(PB) \quad (3.2-6)$$

where

AX = area of the source (X = S) or drain (X = D)

PX = perimeter of the source (X = S) or drain (X = D)

CJSW = zero-bias, bulk–source/drain sidewall capacitance

MJSW = bulk–source/drain sidewall grading coefficient

Table 3.2-1 gives the values for CJ, CJSW, MJ, and MJSW for an MOS device that has an oxide thickness of 140 Å resulting in a  $C_{ox} = 24.7 \times 10^{-4} \text{ F/m}^2$ . It can be seen that the

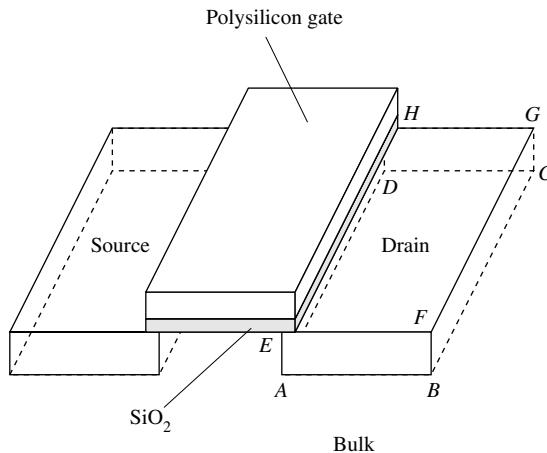


**Figure 3.2-3** Illustration showing the bottom (ABCD) and sidewall (ABFE + BCGF + DCGH + ADHE) components of the bulk junction capacitors.

**Table 3.2-1** Capacitance Values and Coefficients for the MOS Model

| Type | p-Channel             | n-Channel             | Units            |
|------|-----------------------|-----------------------|------------------|
| CGSO | $220 \times 10^{-12}$ | $220 \times 10^{-12}$ | F/m              |
| CGDO | $220 \times 10^{-12}$ | $220 \times 10^{-12}$ | F/m              |
| CGBO | $700 \times 10^{-12}$ | $700 \times 10^{-12}$ | F/m              |
| CJ   | $560 \times 10^{-6}$  | $770 \times 10^{-6}$  | F/m <sup>2</sup> |
| CJSW | $350 \times 10^{-12}$ | $380 \times 10^{-12}$ | F/m              |
| MJ   | 0.5                   | 0.5                   |                  |
| MJSW | 0.35                  | 0.38                  |                  |

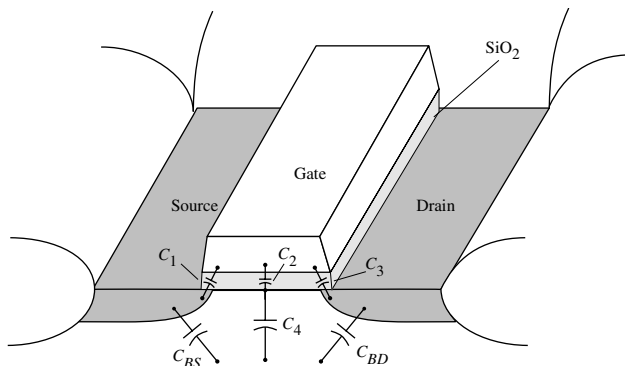
Based on an oxide thickness of 140 Å or  $C_{ox} = 24.7 \times 10^{-4}$  F/m<sup>2</sup>.

depletion capacitors cannot be accurately modeled until the geometry of the device is known, for example, the area and perimeter of the source and drain. However, values can be assumed for the purpose of design. For example, one could consider a typical source or drain to be 1.8 μm by 5 μm. Thus, a value for  $C_{BX}$  of 12.1 F and 9.8 F results for n-channel and p-channel devices, respectively, for  $V_{BX} = 0$ .

The large-signal, charge-storage capacitors of the MOS device consist of the gate-to-source ( $C_{GS}$ ), gate-to-drain ( $C_{GD}$ ), and gate-to-bulk ( $C_{GB}$ ) capacitances. Figure 3.2-4 shows a cross section of the various capacitances that constitute the charge-storage capacitors of the MOS device.  $C_{BS}$  and  $C_{BD}$  are the bulk-to-source and bulk-to-drain capacitors discussed above. The following discussion represents a heuristic development of a model for the large-signal charge-storage capacitors.

$C_1$  and  $C_3$  are overlap capacitances and are due to an overlap of two conducting surfaces separated by a dielectric. The overlapping capacitors are shown in more detail in Fig. 3.2-5. The amount of overlap is designated as LD. This overlap is due to the lateral diffusion of the source and drain underneath the polysilicon gate. For example, a 0.8 μm CMOS process might have a lateral diffusion component, LD, of approximately 16 nm. The overlap capacitances can be approximated as

$$C_1 = C_3 \cong (LD)(W_{\text{eff}})C_{ox} = (CGXO)W_{\text{eff}} \quad (3.2-7)$$

**Figure 3.2-4** Large-signal, charge-storage capacitors of the MOS device.

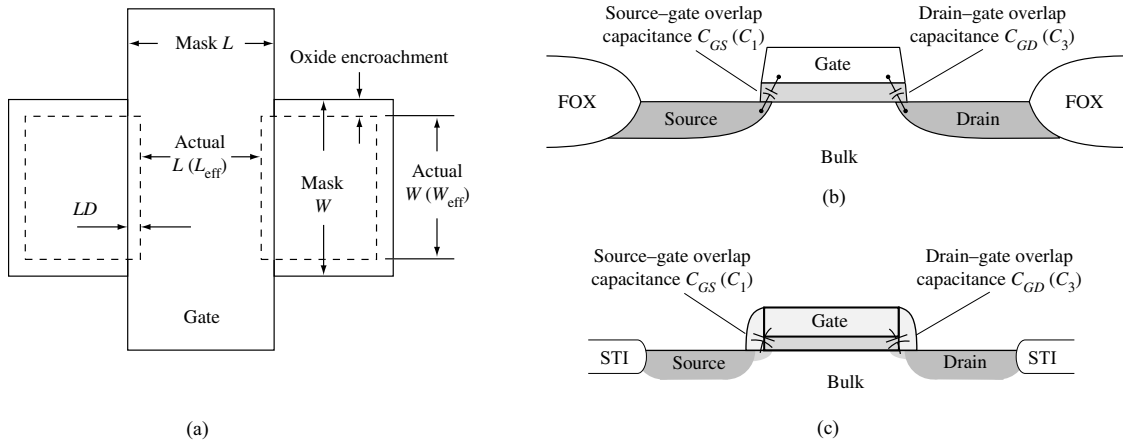


**Figure 3.2-5** Overlap capacitances of an MOS transistor. (a) Top view showing the overlap between the source or drain and the gate for LOCOS technology. (b) Side view for LOCOS technology. (c) Side view for STI technology.

where  $W_{eff}$  is the effective channel width and  $CGXO$  ( $X = S$  or  $D$ ) is the overlap capacitance in F/m for the gate–source or gate–drain overlap. The difference between the mask  $W$  and actual  $W$  is due to the encroachment of the field oxide under the silicon nitride. Table 3.2-1 gives a value for  $CGSO$  and  $CGDO$  based on a device with an oxide thickness of  $140\text{ \AA}$ . A third overlap capacitance that can be significant is the overlap between the gate and the bulk. Figure 3.2-6 shows this overlap capacitor ( $C_5$ ) in more detail. This is the capacitance that occurs between the gate and bulk at the edges of the channel and is a function of the effective length of the channel,  $L_{eff}$ . Table 3.2-1 gives a typical value for  $CGBO$  for a device based on an oxide thickness of  $140\text{ \AA}$ .

If the device illustrated in Fig. 3.2-4 was in the saturated state, the channel would extend almost to the drain and would extend completely to the drain if the MOS device was in the nonsaturated state.  $C_2$  is the gate-to-channel capacitance and is given as

$$C_2 = W_{eff}(L - 2LD)C_{ox} = W_{eff}(L_{eff})C_{ox} \quad (3.2-8)$$

The term  $L_{eff}$  is the effective channel length resulting from the mask-defined length being reduced by the amount of lateral diffusion (note that up until now, the symbols  $L$  and  $W$  were

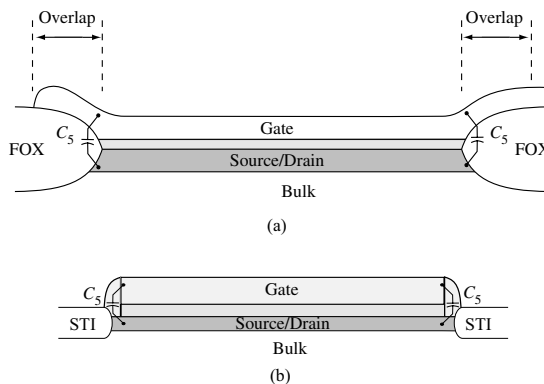


**Figure 3.2-6** Gate–bulk overlap capacitances. (a) LOCOS technology. (b) STI technology.

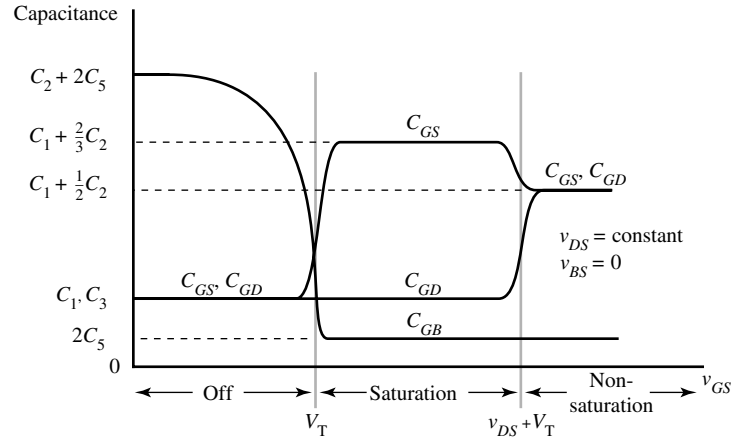


**Figure 3.2-7** Voltage dependence of  $C_{GS}$ ,  $C_{GD}$ , and  $C_{GB}$  as a function of  $V_{GS}$  with  $V_{DS}$  constant and  $V_{BS} = 0$ .

used to refer to “effective” dimensions whereas now these have been changed for added clarification).  $C_4$  is the channel-to-bulk capacitance, which is a depletion capacitance that will vary with voltage like  $C_{BS}$  or  $C_{BD}$ .

It is of interest to examine  $C_{GB}$ ,  $C_{GS}$ , and  $C_{GD}$  as  $v_{DS}$  is held constant and  $v_{GS}$  is increased from zero. To understand the results, one can imagine following a vertical line on Fig. 3.1-3 at, say,  $v_{DS} = 0.5(V_{GS0} - V_T)$ , as  $v_{GS}$  increases from zero. The MOS device will first be off until  $v_{GS}$  reaches  $V_T$ . Next, it will be in the saturated region until  $v_{GS}$  becomes equal to  $v_{DS}(\text{sat}) + V_T$ . Finally, the MOS device will be in the nonsaturated region. The approximate variation of  $C_{GB}$ ,  $C_{GS}$ , and  $C_{GD}$  under these conditions is shown in Fig. 3.2-7. In cutoff, there is no channel and  $C_{GB}$  is approximately equal to  $C_2 + 2C_5$ . As  $v_{GS}$  approaches  $V_T$  from the off region, a thin depletion layer is formed, creating a large value of  $C_4$ . Since  $C_4$  is in series with  $C_2$ , little effect is observed. As  $v_{GS}$  increases, this depletion region widens, causing  $C_4$  to decrease and reducing  $C_{GB}$ . When  $v_{GS} = V_T$ , an inversion layer is formed that prevents further decreases of  $C_4$  (and thus  $C_{GB}$ ).

$C_1$ ,  $C_2$ , and  $C_3$  constitute  $C_{GS}$  and  $C_{GD}$ . The problem is how to allocate  $C_2$  to  $C_{GS}$  and  $C_{GD}$ . The approach used is to assume in saturation that approximately two-thirds of  $C_2$  belongs to  $C_{GS}$  and none to  $C_{GD}$ . This is, of course, an approximation. However, it has been found to give reasonably good results. Figure 3.2-7 shows how  $C_{GS}$  and  $C_{GD}$  change values in going from the off to the saturation region. Finally, when  $v_{GS}$  is greater than  $v_{DS} + V_T$ , the MOS device enters the nonsaturated region. In this case, the channel extends from the drain to the source and  $C_2$  is simply divided evenly between  $C_{GD}$  and  $C_{GS}$  as shown in Fig. 3.2-7.

As a consequence of the above considerations, we shall use the following formulas for the charge-storage capacitances of the MOS device in the indicated regions.

**Off**

$$C_{GB} = C_2 + 2C_5 = C_{ox}(W_{\text{eff}})(L_{\text{eff}}) + C_{GBO}(L_{\text{eff}}) \quad (3.2-9a)$$

$$C_{GS} = C_1 \cong C_{ox}(\text{LD})(W_{\text{eff}}) = C_{GSO}(W_{\text{eff}}) \quad (3.2-9b)$$

$$C_{GD} = C_3 \cong C_{ox}(\text{LD})(W_{\text{eff}}) = C_{GDO}(W_{\text{eff}}) \quad (3.2-9c)$$

**Saturation**

$$C_{GB} = 2C_5 = \text{CGBO}(L_{\text{eff}}) \quad (3.2-10a)$$

$$\begin{aligned} C_{GS} &= C_1 + \frac{2}{3}C_2 = C_{\text{ox}}(\text{LD} + 0.67L_{\text{eff}})(W_{\text{eff}}) \\ &= \text{CGSO}(W_{\text{eff}}) + 0.67C_{\text{ox}}(W_{\text{eff}})(L_{\text{eff}}) \end{aligned} \quad (3.2-10b)$$

$$C_{GD} = C_3 \cong C_{\text{ox}}(\text{LD})(W_{\text{eff}}) = \text{CGDO}(W_{\text{eff}}) \quad (3.2-10c)$$

**Nonsaturated**

$$C_{GB} = 2C_5 = \text{CGBO}(L_{\text{eff}}) \quad (3.2-11a)$$

$$\begin{aligned} C_{GS} &= C_1 + 0.5C_2 = C_{\text{ox}}(\text{LD} + 0.5L_{\text{eff}})(W_{\text{eff}}) \\ &= (\text{CGSO} + 0.5C_{\text{ox}}L_{\text{eff}})W_{\text{eff}} \end{aligned} \quad (3.2-11b)$$

$$\begin{aligned} C_{GD} &= C_3 + 0.5C_2 = C_{\text{ox}}(\text{LD} + 0.5L_{\text{eff}})(W_{\text{eff}}) \\ &= (\text{CGDO} + 0.5C_{\text{ox}}L_{\text{eff}})W_{\text{eff}} \end{aligned} \quad (3.2-11c)$$

Equations that provide a smooth transition between the three regions can be found in the literature [9].

Other capacitor parasitics associated with transistors are due to interconnect to the transistor, for example, polysilicon over field (substrate). This type of capacitance typically constitutes the major portion of  $C_{GB}$  in the nonsaturated and saturated regions and thus is very important and should be considered in the design of CMOS circuits.

Another important aspect of modeling the CMOS device is noise. The existence of noise is due to the fact that electrical charge is not continuous but is carried in discrete amounts equal to the charge of an electron. In electronic circuits, noise manifests itself by representing a lower limit below which electrical signals cannot be amplified without significant deterioration in the quality of the signal. Noise can be modeled by a current source connected in parallel with  $i_D$  of Fig. 3.2-1. This current source represents two sources of noise, called thermal noise and flicker noise [10,11]. These sources of noise were discussed in Section 2.5. The mean-square current-noise source is defined as

$$i_n^2 = \left[ \frac{8kTg_m(1 + \eta)}{3} + \frac{(\text{KF})I_D}{f C_{\text{ox}} L^2} \right] \Delta f \quad (A^2) \quad (3.2-12)$$

where

$\Delta f$  = a small bandwidth (typically 1 Hz) at a frequency  $f$

$\eta = g_{\text{mbs}}/g_m$  [see Eq. (3.3-8)]

$k$  = Boltzmann's constant

$T$  = temperature (K)

$g_m$  = small-signal transconductance from gate to channel [see Eq. (3.3-6)]

KF = flicker noise coefficient (F-A)

$f$  = frequency (Hz)

KF has a typical value of  $10^{-28}$  F-A. Both sources of noise are process dependent and the values are usually different for enhancement and depletion mode field effect transistors (FETs).

The mean-square current noise can be reflected to the gate of the MOS device by dividing Eq. (3.2-12) by  $g_m^2$ , if the source is on ac ground, giving

$$e_n^2 = \frac{i_n^2}{g_m^2} = \left[ \frac{8kT(1 + \eta)}{3 g_m} + \frac{KF}{2f C_{ox} WLK'} \right] \Delta f \quad (V^2) \quad (3.2-13)$$

The equivalent input-mean-square voltage-noise form of Eq. (3.2-13) will be useful for analyzing the noise performance of CMOS circuits in later chapters.

The experimental noise characteristics of n-channel and p-channel devices are shown in Figs. 3.2-8(a) and 3.2-8(b). These devices were fabricated using a submicron, silicon-gate, n-well, CMOS process. The data in Figs. 3.2-8(a) and 3.2-8(b) are typical for MOS devices and show that the  $1/f$  noise is the dominant source of noise for frequencies below 100 kHz (at the given bias conditions).\* Consequently, in many practical cases, the equivalent input-mean-square voltage noise of Eq. (3.2-13) is simplified to

$$e_{eq}^2 = \left[ \frac{KF}{2f C_{ox} WLK'} \right] \Delta f \quad (V^2) \quad (3.2-14)$$

or in terms of the input-voltage-noise spectral density we can rewrite Eq. (3.2-14) as

$$e_{eq}^2 = \frac{e_{eq}^2}{\Delta f} = \frac{KF}{2f C_{ox} WLK'} = \frac{B}{f WL} \quad (V^2/Hz) \quad (3.2-15)$$

where  $B$  is a constant for an n-channel or a p-channel device of a given process.\*\* The right-hand expression of Eq. (3.2-15) will be important in optimizing the design with respect to noise performance.

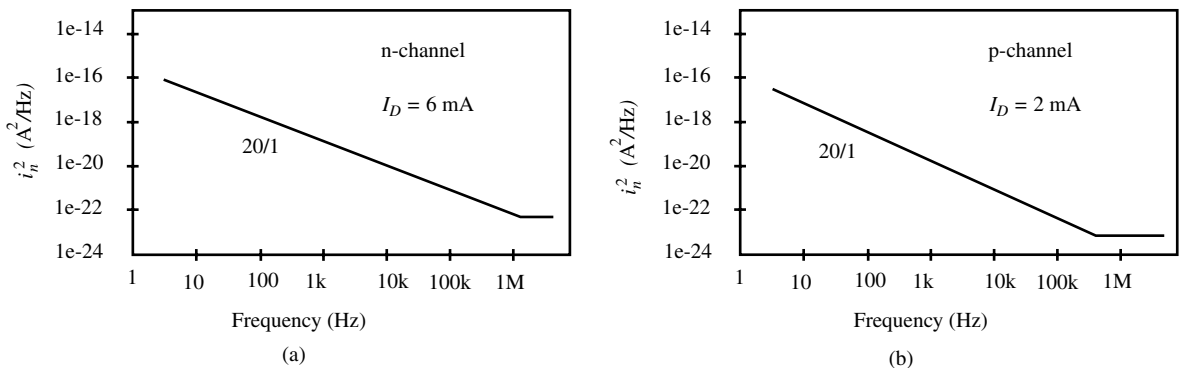


**Figure 3.2-8** Drain-current noise for (a) an n-channel and (b) a p-channel MOSFET measured on a silicon-gate submicron process.

\*If the bias current is reduced, the thermal noise floor increases, thus moving the  $1/f$  noise corner to a lower frequency. Therefore, the  $1/f$  noise corner is a function of the thermal noise floor.

\*\*Since the same symbol is used for voltage (current) noise and voltage (current) spectral density, the units are generally used to distinguish the difference if it is not clear in the text.

The temperature behavior of the MOS was given in Eqs. (2.5-8) and (2.5-9). Equation (2.5-8) can be expressed as

$$K'(T) = K'(T_0) (T/T_0)^{-1.5} \quad (3.2-16)$$

Equation (2.5-9) is a series approximation of the threshold voltage temperature dependence and can be written as

$$V_T(T) = V_T(T_0) + \alpha(T - T_0) + \dots \quad (3.2-17)$$

where a plus sign is used instead of the minus sign in Eq. (2.5-9). This means that  $\alpha$  will be negative for NMOS and positive for PMOS. Substituting Eqs. (3.2-16) and (3.2-17) into the simple, large-signal model of the MOSFET in Eq. (3.1-17) gives the temperature dependence of the transistor:

$$I_D(T) = \frac{\mu_o C_{ox} W}{2L} \left( \frac{T}{T_0} \right)^{-1.5} [V_{GS} - V_{T0} - \alpha(T - T_0)]^2 \quad (3.2-18)$$

where  $V_{T0}$  is the value of the threshold voltage at the reference temperature,  $T_0$ .

While Eq. (3.2-18) can be used to determine the influence of temperature on the transistor current, there is an interesting characteristic called the *zero temperature coefficient* point that we want to examine. The differentiation of Eq. (3.2-18) is given as

$$\begin{aligned} \frac{dI_D}{dT} = & \frac{-1.5\mu_o C_{ox} W}{2LT_0} \left( \frac{T}{T_0} \right)^{-2.5} [V_{GS} - V_{T0} - \alpha(T - T_0)]^2 \\ & - \alpha \frac{\mu_o C_{ox} W}{L} \left( \frac{T}{T_0} \right)^{-1.5} [V_{GS} - V_{T0} - \alpha(T - T_0)] \end{aligned} \quad (3.2-19)$$

If Eq. (3.2-19) is set to zero, the value of  $V_{GS}$  that gives a zero temperature coefficient is given as

$$V_{GS}(\text{ZTC}) = V_{T0} - \alpha T_0 - \frac{\alpha T}{3} \quad (3.2-20)$$

The zero temperature characteristics of an NMOS transistor are illustrated in Fig. 3.2-9. Note that in general, the ZTC point is well defined only for temperatures less than 150 °C. For higher temperatures the ZTC drifts from the low-temperature value. Also, note that for values of gate–source voltage above the ZTC value, the drain current *decreases* as the temperature *increases*. This reduces the thermal runaway characteristics of the MOSFET compared to the BJT. The following example illustrates the application of the ZTC concept.

### Example 3.2-1

#### Example 3.2-1 Application of the ZTC Point to an NMOS Transistor

Assume that  $K' = 110 \mu\text{A/V}^2$ ,  $V_{T0} = 0.7 \text{ V}$ , and  $\alpha = -2.3 \text{ mV/}^\circ\text{C}$  for an NMOS transistor. Find the ZTC point if the reference temperature is 300 K and the new temperature is 400 K. Find the drain current under these conditions if  $W/L = 10$ .

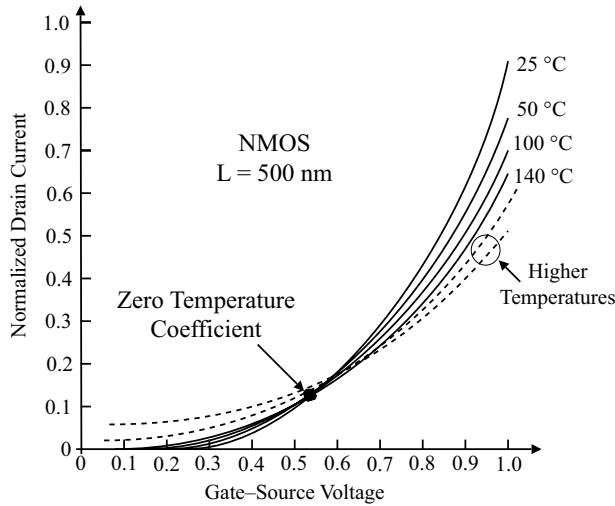


**Figure 3.2-9** Zero temperature coefficient characteristics of a modern CMOS technology ( $L_{\min} = 50$  nm).

#### SOLUTION

Evaluating Eq. (3.2-20) gives

$$V_{GS}(ZTC) = 0.7 \text{ V} + 0.0023(\text{V}/^\circ\text{C})(300 \text{ K}) + \frac{0.0023(\text{V}/^\circ\text{C})(400 \text{ K})}{3}$$

$$= 1.697 \text{ V}$$

At this value of VGS the drain current is

$$I_D(400^\circ\text{K}) = \frac{110 \mu\text{A}/\text{V}^2(10)}{2} \left( \frac{400}{300} \right)^{-1.5} [1.697 - 0.7 - 0.0023(100)]^2$$

$$= 176 \mu\text{A}$$



### 3.3 Small-Signal Model for the MOS Transistor

Up to this point, we have been considering the large-signal model of the MOS transistor shown in Fig. 3.2-1. However, after the large-signal model has been used to find the dc conditions, the small-signal model becomes important. The small-signal model is a linear model that helps to simplify calculations. It is valid only over voltage or current regions where the large-signal voltage and currents can adequately be represented by a straight line.

Figure 3.3-1 shows a linearized small-signal model for the MOS transistor. The parameters of the small-signal model will be designated by lowercase subscripts. The various parameters of this small-signal model are all related to the large-signal model parameters and dc variables. The normal relationship between these two models assumes that the small-signal parameters are defined in terms of the ratio of small perturbations of the large-signal variables or as the partial differentiation of one large-signal variable with respect to another.

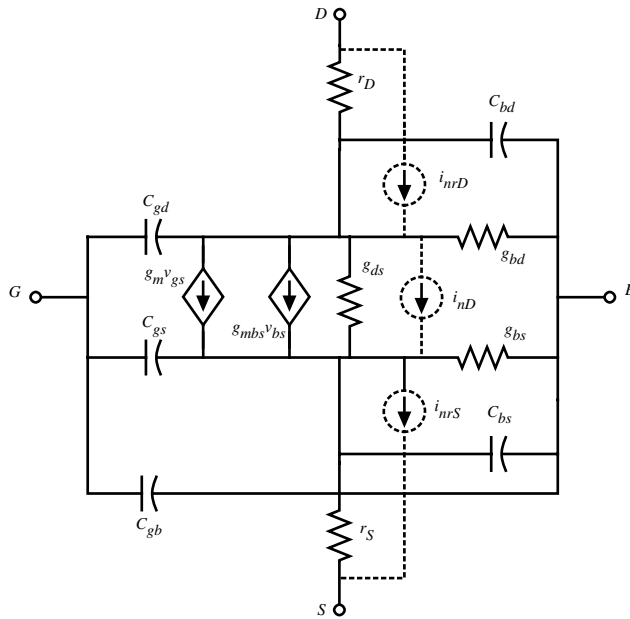


**Figure 3.3-1** Small-signal model of the MOS transistor.

The conductances  $g_{bd}$  and  $g_{bs}$  are the equivalent conductances of the bulk-to-drain and bulk-to-source junctions. Since these junctions are normally reverse biased, the conductances are very small. They are defined as

$$g_{bd} = \frac{\partial i_{BD}}{\partial v_{BD}} \text{ (evaluated at the quiescent point) } \cong 0 \quad (3.3-1)$$

and

$$g_{bs} = \frac{\partial i_{BS}}{\partial v_{BS}} \text{ (evaluated at the quiescent point) } \cong 0 \quad (3.3-2)$$

The channel transconductances  $g_m$  and  $g_{mbs}$  and the channel conductance  $g_{ds}$  are defined as

$$g_m = \frac{\partial i_D}{\partial v_{GS}} \text{ (evaluated at the quiescent point) } \quad (3.3-3)$$

$$g_{mbs} = \frac{\partial i_D}{\partial v_{BS}} \text{ (evaluated at the quiescent point) } \quad (3.3-4)$$

and

$$g_{ds} = \frac{\partial i_D}{\partial v_{DS}} \text{ (evaluated at the quiescent point) } \quad (3.3-5)$$

The values of these small-signal parameters depend on which region the quiescent point occurs in. For example, in the saturated region  $g_m$  can be found from Eq. (3.1-18) as

$$g_m = \sqrt{(2K'W/L)|I_D|(1 + \lambda V_{DS})} \cong \sqrt{(2K'W/L)|I_D|} \quad (3.3-6)$$

which emphasizes the dependence of the small-signal parameters on the large-signal operating conditions. The small-signal channel transconductance due to  $v_{SB}$  is found by rewriting Eq. (3.3-4) as

$$g_{mbs} = \frac{-\partial i_D}{\partial v_{SB}} = -\left(\frac{\partial i_D}{\partial V_T}\right)\left(\frac{\partial V_T}{\partial v_{SB}}\right) \quad (3.3-7)$$

Using Eq. (3.1-2) and noting that  $\partial i_D / \partial V_T = -\partial i_D / \partial v_{GS}$ , we get\*

$$g_{mbs} = g_m \frac{\gamma}{2(|\phi_F| + |V_{SB}|)^{1/2}} = \eta g_m \quad (3.3-8)$$

This transconductance will become important in our small-signal analysis of the MOS transistor when the ac value of the source–bulk potential  $v_{sb}$  is not zero.

The small-signal channel conductance,  $g_{ds}$  ( $g_0$ ), is given as

$$g_{ds} = g_0 = \frac{I_D \lambda}{1 + \lambda V_{DS}} \cong I_D \lambda \quad (3.3-9)$$

The channel conductance will be dependent on  $L$  through  $\lambda$ , which is inversely proportional to  $L$ . We have assumed the MOS transistor is in saturation for the results given by Eqs. (3.3-6), (3.3-8), and (3.3-9).

The important dependence of the small-signal parameters on the large-signal model parameters and dc voltages and currents is illustrated in Table 3.3-1. In this table we see that the three small-signal model parameters of  $g_m$ ,  $g_{mbs}$ , and  $g_{ds}$  have several alternate forms. An example of the typical values of the small-signal model parameters follows.

### Example 3.3-1

#### Typical Values of Small-Signal Model Parameters

Find the values of  $g_m$ ,  $g_{mbs}$ , and  $g_{ds}$  using the large-signal model parameters in Table 3.1-2 for both an n-channel and a p-channel device if the dc value of the magnitude of the drain current is 50  $\mu\text{A}$  and the magnitude of the dc value of the source–bulk voltage is 2 V. Assume that the  $W/L$  ratio is 1  $\mu\text{m}/1 \mu\text{m}$ .

#### SOLUTION

Using the values of Table 3.1-2 and Eqs. (3.3-6), (3.3-8), and (3.3-9) gives  $g_m = 105 \mu\text{A/V}$ ,  $g_{mbs} = 12.8 \mu\text{A/V}$ , and  $g_{ds} = 2.0 \mu\text{A/V}$  for the n-channel device and  $g_m = 70.7 \mu\text{A/V}$ ,  $g_{mbs} = 12.0 \mu\text{A/V}$ , and  $g_{ds} = 2.5 \mu\text{A/V}$  for the p-channel device.

\*Note that absolute signs are used for  $V_{SB}$  in order to prevent  $g_{mbs}$  from becoming infinite. However, in a few rare cases the source–bulk junction is forward biased and in this case the absolute signs must be removed and  $V_{SB}$  becomes negative (for an n-channel transistor).

**Table 3.3-1** Dependence of the Small-Signal Model Parameters on the dc Values of Voltage and Current in the Saturation Region

| Small-Signal Model Parameters | dc Current                 | dc Current and Voltage                                          | dc Voltage                                                                |
|-------------------------------|----------------------------|-----------------------------------------------------------------|---------------------------------------------------------------------------|
| $g_m$                         | $\cong (2K'I_D W/L)^{1/2}$ | —                                                               | $\cong \frac{K'W}{L}(V_{GS} - V_T)$                                       |
| $g_{mbs}$                     | —                          | $\frac{\gamma(2I_D\beta)^{1/2}}{2(2 \phi_F  +  V_{SB} )^{1/2}}$ | $\frac{\gamma[\beta(V_{GS} - V_T)]^{1/2}}{2(2 \phi_F  +  V_{SB} )^{1/2}}$ |
| $g_{ds}$                      | $\cong \lambda I_D$        | —                                                               | —                                                                         |

Although MOS devices are not often used in the nonsaturation region in analog circuit design, the relationships of the small-signal model parameters in the nonsaturation region are given as

$$g_m = \frac{\partial i_D}{\partial v_{GS}} \cong \beta V_{DS} \quad (3.3-10)$$

$$g_{mbs} = \frac{\partial i_D}{\partial v_{BS}} = \frac{\beta \gamma V_{DS}}{2(2|\phi_F| + |V_{SB}|)^{1/2}} \quad (3.3-11)$$

and

$$g_{ds} \cong \beta(V_{GS} - V_T - V_{DS}) \quad (3.3-12)$$

Table 3.3-2 summarizes the dependence of the small-signal model parameters on the large-signal model parameters and dc voltages and currents for the nonsaturated region. The typical values of the small-signal model parameters for the nonsaturated region are illustrated in the following example.

### Example 3.3-2

#### Typical Values of the Small-Signal Model Parameters in the Nonsaturated Region

Find the values of the small-signal model parameters in the nonsaturation region for an n-channel and a p-channel transistor if  $V_{GS} = 5$  V,  $V_{DS} = 1$  V, and  $|V_{BS}| = 2$  V. Assume that the  $W/L$  ratio for both transistors is  $1 \mu\text{m}/1 \mu\text{m}$ . Also assume that the value for  $K'$  in the nonsaturation region is the same as that for the saturation (generally a poor assumption).

#### SOLUTION

First, it is necessary to calculate the threshold voltage of each transistor using Eq. (3.1-2). The results are a  $V_T$  of 1.02 V for the n-channel and  $-1.14$  V for the p-channel. This gives a dc current of  $383 \mu\text{A}$  and  $168 \mu\text{A}$ , respectively. Using Eqs. (3.3-10), (3.3-11), and (3.3-12), we get  $g_m = 110 \mu\text{A/V}$ ,  $g_{mbs} = 13.4 \mu\text{A/V}$ , and  $r_{ds} = 3.05 \text{ k}\Omega$  for the n-channel transistor and  $g_m = 50 \mu\text{A/V}$ ,  $g_{mbs} = 8.52 \mu\text{A/V}$ , and  $r_{ds} = 6.99 \text{ k}\Omega$  for the p-channel transistor.

**Table 3.3-2** Dependence of the Small-Signal Model Parameters on the dc Values of Voltage and Current in the Nonsaturation Region

| Small-Signal Model Parameters | dc Voltage and/or Current Dependence                        |
|-------------------------------|-------------------------------------------------------------|
| $g_m$                         | $\cong \beta V_{DS}$                                        |
| $g_{mbs}$                     | $\frac{\beta \gamma V_{DS}}{2(2 \phi_F  +  V_{SB} )^{1/2}}$ |
| $g_{ds}$                      | $\cong \beta (V_{GS} - V_T - V_{DS})$                       |

The values of  $r_d$  and  $r_s$  are assumed to be the same as  $r_D$  and  $r_S$  of Fig. 3.2-1. Likewise, for small-signal conditions  $C_{gs}$ ,  $C_{gd}$ ,  $C_{gb}$ ,  $C_{bd}$ , and  $C_{bs}$  are evaluated for  $C_{gs}$ ,  $C_{gd}$ , and  $C_{gb}$  by knowing the region of operation (cutoff, saturation or nonsaturation) and for  $C_{bd}$  and  $C_{bs}$  by knowing the value of  $V_{BD}$  and  $V_{BS}$ . With this information,  $C_{gs}$ ,  $C_{gd}$ ,  $C_{gb}$ ,  $C_{bd}$ , and  $C_{bs}$  can be found from  $C_{GS}$ ,  $C_{GD}$ ,  $C_{GB}$ ,  $C_{BD}$ , and  $C_{BS}$ , respectively.

If the noise of the MOS transistor is to be modeled, then three additional current sources are added to Fig. 3.3-1 as indicated by the dashed lines. The values of the mean-square noise-current sources are given as

$$i_{nrD}^2 = \left( \frac{4kT}{r_D} \right) \Delta f \quad (A^2) \quad (3.3-13)$$

$$i_{nrS}^2 = \left( \frac{4kT}{r_S} \right) \Delta f \quad (A^2) \quad (3.3-14)$$

and

$$i_{nD}^2 = \left[ \frac{8kT g_m (1 + \eta)}{3} + \frac{(KF)I_D}{f C_{ox} L^2} \right] \Delta f \quad (A^2) \quad (3.3-15)$$

The various parameters for these equations have previously been defined. With the noise modeling capability, the small-signal model of Fig. 3.3-1 is a very general model.

It will be important to be familiar with the small-signal model for the saturation region developed in this section. This model, along with the circuit simplification techniques given in Appendix A, will be the key element in analyzing the circuits in the following chapters.



### 3.4 Computer Simulation Models

The large-signal model of the MOS device previously discussed is simple to use for hand calculations but neglects many important second-order effects. While a simple model for hand calculation and design intuition is critical, a more accurate model is required for computer simulation. There are many model choices available for the designer when choosing a device model to use for computer simulation. At one time, HSPICE\* supported 43 different MOSFET

\*HSPICE is now owned by Avant! Inc. and has been renamed Star-Hspice.

models [2] (many of which were company proprietary) while SmartSpice publishes support for 14 [12]. Which model is the right one to use? In the fabless semiconductor environment, the user must use the model provided by the wafer foundry. In companies where the foundry is captive (i.e., the company owns its own wafer fabrication facility) a modeling group provides the model to circuit designers. It is seldom that a designer chooses a model and performs parameter extraction to get the terms for the model chosen.

The SPICE LEVEL 3 dc model will be covered in some detail because it is a relatively straightforward extension of the LEVEL 2 model. The BSIM3v3 model will be introduced but the detailed equations will not be presented because of the volume of equations required to describe it—there are other good texts that deal with the subject of modeling exclusively [13,14], and there is little additional design intuition derived from covering the details.

Models developed for computer simulation have improved over the years but no model has yet been developed that, with a single set of parameters, covers device operation for all possible geometries. Therefore, many SPICE simulators offer a feature called “model binning.” Parameters are derived for transistors of different geometry ( $W$ ’s and  $L$ ’s) and the simulator determines which set of parameters to use based on the particular  $W$  and  $L$  called out in the device instantiation line in the circuit description. The circuit designer need only be aware of this since the binning is done by the model provider.

### SPICE LEVEL 3 Model

The large-signal model of the MOS device previously discussed is simple to use for hand calculations but neglects many important second-order effects. Most of these second-order effects are due to narrow or short channel dimensions (less than about 3  $\mu\text{m}$ ). In this section, we will consider a more complex model that is suitable for computer-based analysis (circuit simulation, i.e., SPICE simulation). In particular, the SPICE LEVEL 3 model will be covered (see Table 3.4-1). This model is typically good for MOS technologies down to about 0.8  $\mu\text{m}$ . We will also consider the effects of temperature on the parameters of the MOS large-signal model.

We first consider second-order effects due to small geometries (Fig. 3.4-1). When  $v_{GS}$  is greater than  $V_T$ , the drain current for a small device can be given as [2] follows:

#### Drain Current

$$i_{DS} = \text{BETA} \left[ v_{GS} - V_T - \left( \frac{1 + f_b}{2} \right) v_{DE} \right] v_{DE} \quad (3.4-1)$$

$$\text{BETA} = \text{KP} \frac{W_{\text{eff}}}{L_{\text{eff}}} = \mu_{\text{eff}} \text{COX} \frac{W_{\text{eff}}}{L_{\text{eff}}} \quad (3.4-2)$$

$$L_{\text{eff}} = L - 2(\text{LD}) \quad (3.4-3)$$

$$W_{\text{eff}} = W - 2(\text{WD}) \quad (3.4-4)$$

$$v_{DE} = \min(v_{DS}, v_{DS}(\text{sat})) \quad (3.4-5)$$

$$f_b = f_n + \frac{\text{GAMMA} \cdot f_s}{4(\text{PHI} + v_{SB})^{1/2}} \quad (3.4-6)$$

**Table 3.4-1 Typical Model Parameters Suitable for SPICE Simulations Using LEVEL-3 Model (Extended Model)\***

| Parameter Symbol | Parameter Description                                | Typical Parameter Value |                       |                                     |
|------------------|------------------------------------------------------|-------------------------|-----------------------|-------------------------------------|
|                  |                                                      | n-Channel               | p-Channel             | Units                               |
| VTO              | Threshold                                            | $0.7 \pm 0.15$          | $-0.7 \pm 0.15$       | V                                   |
| UO               | Mobility                                             | 660                     | 210                   | $\text{cm}^2/\text{V}\cdot\text{s}$ |
| DELTA            | Narrow-width threshold adjustment factor             | 2.4                     | 1.25                  | —                                   |
| ETA              | Static-feedback threshold adjustment factor          | 0.1                     | 0.1                   | —                                   |
| KAPPA            | Saturation field factor in channel length modulation | 0.15                    | 2.5                   | 1/V                                 |
| THETA            | Mobility degradation factor                          | 0.1                     | 0.1                   | 1/V                                 |
| NSUB             | Substrate doping                                     | $3 \times 10^{16}$      | $6 \times 10^{16}$    | $\text{cm}^{-3}$                    |
| TOX              | Oxide thickness                                      | 140                     | 140                   | Å                                   |
| XJ               | Metallurgical junction depth                         | 0.2                     | 0.2                   | $\mu\text{m}$                       |
| WD               | Delta width                                          |                         |                       | $\mu\text{m}$                       |
| LD               | Lateral diffusion                                    | 0.016                   | 0.015                 | $\mu\text{m}$                       |
| NFS              | Parameter for weak inversion modeling                | $7 \times 10^{11}$      | $6 \times 10^{11}$    | $\text{cm}^{-2}$                    |
| CGSO             |                                                      | $220 \times 10^{-12}$   | $220 \times 10^{-12}$ | F/m                                 |
| CGDO             |                                                      | $220 \times 10^{-12}$   | $220 \times 10^{-12}$ | F/m                                 |
| CGBO             |                                                      | $700 \times 10^{-12}$   | $700 \times 10^{-12}$ | F/m                                 |
| CJ               |                                                      | $770 \times 10^{-6}$    | $560 \times 10^{-6}$  | $\text{F}/\text{m}^2$               |
| CJSW             |                                                      | $380 \times 10^{-12}$   | $350 \times 10^{-12}$ | F/m                                 |
| MJ               |                                                      | 0.5                     | 0.5                   |                                     |
| MJSW             |                                                      | 0.38                    | 0.35                  |                                     |

\*These values are based on a 0.8  $\mu\text{m}$  silicon-gate bulk CMOS n-well process and include capacitance parameters from Table 3.2-1.

Note that PHI is the SPICE model term for the quantity  $2\phi_F$ . Also be aware that PHI is always positive in SPICE regardless of the transistor type (p- or n-channel). In this text, the term PHI will always be positive while the term  $2\phi_F$  will have a polarity determined by the transistor type as shown in Table 2.3-1.

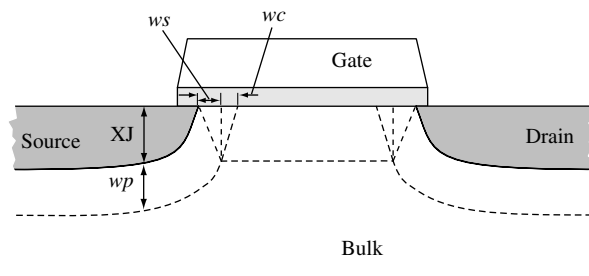


**Figure 3.4-1** Illustration of the short-channel effects in the MOS transistor.

$$f_n = \frac{\text{DELTA}}{W_{\text{eff}}} \frac{\pi \epsilon_{\text{Si}}}{2 \cdot C_{\text{ox}}} \quad (3.4-7)$$

$$f_s = 1 - \frac{\text{XJ}}{L_{\text{eff}}} \left\{ \frac{\text{LD} + wc}{\text{XJ}} \left[ 1 - \left( \frac{wp}{\text{XJ} + wp} \right)^2 \right]^{1/2} - \frac{\text{LD}}{\text{XJ}} \right\} \quad (3.4-8)$$

$$wp = xd(\text{PHI} + v_{SB})^{1/2} \quad (3.4-9)$$

$$xd = \left( \frac{2 \cdot \epsilon_{\text{Si}}}{q \cdot \text{NSUB}} \right)^{1/2} \quad (3.4-10)$$

$$wc = \text{XJ} \left[ k_1 + k_2 \left( \frac{wp}{\text{XJ}} \right) - k_3 \left( \frac{wp}{\text{XJ}} \right)^2 \right] \quad (3.4-11)$$

$$k_1 = 0.0631353, \quad k_2 = 0.08013292, \quad k_3 = 0.01110777$$

### Threshold Voltage

$$V_T = V_{bi} - \left( \frac{\text{ETA} - 8.14 \times 10^{-22}}{C_{\text{ox}} L_{\text{eff}}^3} \right) v_{DS} + \text{GAMMA} \cdot f_s(\text{PHI} + v_{SB})^{1/2} + f_n(\text{PHI} + v_{SB}) \quad (3.4-12)$$

$$v_{bi} = v_{fb} + \text{PHI} \quad (3.4-13)$$

or

$$v_{bi} = \text{VTO} - \text{GAMMA} \cdot \sqrt{\text{PHI}} \quad (3.4-14)$$

### Saturation Voltage

$$v_{\text{sat}} = \frac{v_{gs} - V_T}{1 + f_b} \quad (3.4-15)$$

$$v_{DS}(\text{sat}) = v_{\text{sat}} + v_C - \left( v_{\text{sat}}^2 + v_C^2 \right)^{1/2} \quad (3.4-16)$$

$$v_C = \frac{\text{VMAX} \cdot L_{\text{eff}}}{\mu_s} \quad (3.4-17)$$

If VMAX is not given, then  $v_{DS}(\text{sat}) = v_{\text{sat}}$ .

### Effective Mobility

$$\mu_s = \frac{U_0}{1 + \text{THETA} (v_{Gs} - V_T)}, \quad \text{when VMAX} = 0 \quad (3.4-18)$$

$$\mu_{\text{eff}} = \frac{\mu_s}{1 + \frac{v_{DE}}{v_C}}, \quad \text{when VMAX} > 0; \quad \text{otherwise } \mu_{\text{eff}} = \mu_s \quad (3.4-19)$$

**Channel Length Modulation**

$$\Delta L = xd [KAPPA (v_{DS} - v_{DS}(\text{sat}))]^{1/2}, \quad \text{when } VMAX = 0 \quad (3.4-20)$$

$$\Delta L = -\frac{ep \cdot xd^2}{2} + \left[ \left( \frac{ep \cdot xd^2}{2} \right)^2 + KAPPA \cdot xd^2 (v_{DS} - v_{DS}(\text{sat})) \right]^{1/2}, \quad (3.4-21)$$

when  $VMAX > 0$

where

$$ep = \frac{v_C (v_C + v_{DS}(\text{sat}))}{L_{\text{eff}} v_{DS}(\text{sat})} \quad (3.4-22)$$

$$i_{DS} = \frac{i_{DS}}{1 - \Delta L} \quad (3.4-23)$$

The temperature-dependent variables in the models developed so far include the Fermi potential, PHI, EG, bulk junction potential of the source–bulk and drain–bulk junctions, PB, the reverse currents of the pn junctions,  $I_S$ , and the dependence of mobility on temperature. The temperature dependence of most of these variables is found in the equations given previously or from well-known expressions. The dependence of mobility on temperature is given as

$$U0(T) = U0(T_0) \left( \frac{T}{T_0} \right)^{\text{BEX}} \quad (3.4-24)$$

where BEX is the temperature exponent for mobility and is typically  $-1.5$ .

$$v_{\text{therm}}(T) = \frac{kT}{q} \quad (3.4-25)$$

$$EG(T) = 1.16 - 7.02 \cdot 10^{-4} \left[ \frac{T^2}{T + 1108.0} \right] \quad (3.4-26)$$

$$PHI(T) = PHI(T_0) \cdot \left( \frac{T}{T_0} \right) - v_{\text{therm}}(T) \left[ 3 \ln \left( \frac{T}{T_0} \right) + \frac{EG(T_0)}{v_{\text{therm}}(T_0)} - \frac{EG(T)}{v_{\text{therm}}(T)} \right] \quad (3.4-27)$$

$$v_{bi}(T) = v_{bi}(T_0) + \frac{PHI(T) - PHI(T_0)}{2} + \frac{EG(T_0) - EG(T)}{2} \quad (3.4-28)$$

$$VT0(T) = v_{bi}(T) + GAMMA \left[ \sqrt{PHI(T)} \right] \quad (3.4-29)$$

$$PHI(T) = 2v_{\text{therm}} \ln \left( \frac{NSUB}{n_i(T)} \right) \quad (3.4-30)$$

$$n_i(T) = 1.45 \cdot 10^{16} \left( \frac{T}{T_0} \right)^{3/2} \exp \left[ EG \cdot \left( \frac{T}{T_0} - 1 \right) \left( \frac{1}{2 \cdot v_{\text{therm}}(T_0)} \right) \right] \quad (3.4-31)$$

For drain and source junction diodes, the following relationships apply:

$$PB(T) = PB \cdot \left(\frac{T}{T_0}\right) - v_{\text{therm}}(T) \left[ 3 \ln\left(\frac{T}{T_0}\right) + \frac{EG(T_0)}{v_{\text{therm}}(T_0)} - \frac{EG(T)}{v_{\text{therm}}(T)} \right] \quad (3.4-32)$$

and

$$I_S(T) = \frac{I_S(T_0)}{N} \cdot \exp \left[ \frac{EG(T_0)}{v_{\text{therm}}(T_0)} - \frac{EG(T)}{v_{\text{therm}}(T)} + 3 \ln\left(\frac{T}{T_0}\right) \right] \quad (3.4-33)$$

where  $N$  is the diode emission coefficient. The nominal temperature,  $T_0$ , is 300 K.

An alternate form of the temperature dependence of the MOS model can be found elsewhere [15].

### BSIM 3v3 Model

MOS transistor models introduced thus far in this chapter have been used successfully when applied to 0.8  $\mu\text{m}$  technologies and above. As geometries shrink below 0.8  $\mu\text{m}$ , better models are required. Researchers in the Electrical Engineering and Computer Sciences Department at the University of California at Berkeley have been leaders in the development of SPICE and the models used in it. In 1984 they introduced the BSIM1 model [16] to address the need for a better submicron MOS transistor model. The BSIM1 model approached the modeling problem as a multiparameter curve-fitting exercise. The model contained 60 parameters covering the dc performance of the MOS transistor. There was some relationship to device physics, but in large part, it was a nonphysical model. Later, in 1991, UC Berkeley released the BSIM2 model that improved performance related to the modeling of output resistance changes due to hot-electron effects, source/drain parasitic resistance, and inversion-layer capacitance. This model contained 99 dc parameters, making it more unwieldy than the 60-parameter (dc parameters) BSIM1 model. In 1994, UC Berkeley introduced the BSIM3 model (version 2), which, unlike the earlier BSIM models, returned to a more device-physics-based modeling approach. The model is simpler to use and has only 40 dc parameters. Moreover, the BSIM3 model provides good performance when applied to analog as well as digital circuit simulation. In its third version, BSIM3v3 [3], it has become the industry standard MOS transistor model.

The BSIM3 model addresses the following important effects seen in deep-submicron MOSFET operation:

- Threshold voltage reduction
- Mobility degradation due to a vertical field
- Velocity saturation effects
- Drain-induced barrier lowering (DIBL)
- Channel length modulation
- Subthreshold (weak inversion) conduction
- Parasitic resistance in the source and drain
- Hot-electron effects on output resistance

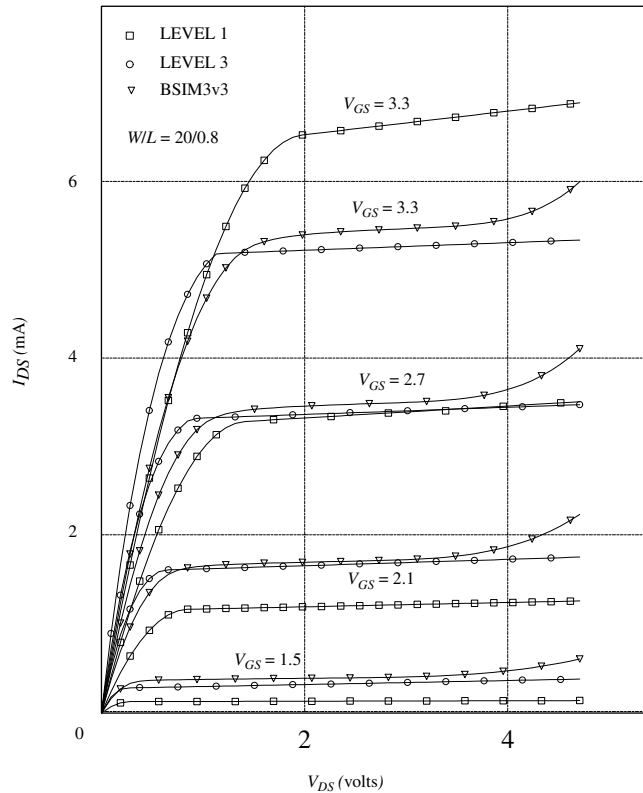


**Figure 3.4-2** Simulation of MOSFET transconductance characteristic using LEVEL = 1, LEVEL = 3 and the BSIM3v3 models.

The plot shown in Fig. 3.4-2 shows a comparison of a 20/0.8 device using the LEVEL 1, LEVEL 3, and BSIM3v3 models. The model parameters were adjusted to provide similar characteristics (given the limitations of each model). Assuming that the BSIM3v3 model closely approximates actual transistor performance, this figure indicates that the LEVEL 1 model is grossly in error, while the LEVEL 3 model shows a significant difference in modeling the transition from the nonsaturation to linear region.



### 3.5 Subthreshold MOS Model

The models discussed in previous sections predict that no current will flow in a device when the gate-source voltage is at or below the threshold voltage. In reality, this is not the case. As  $v_{GS}$  approaches  $V_T$ , the  $i_D - v_{GS}$  characteristics change from square-law to exponential. Whereas the region where  $v_{GS}$  is above the threshold is called the *strong inversion* region, the region below (actually, the transition between the two regions is not well defined as will be explained later) is called the *subthreshold*, or *weak inversion* region. This is illustrated in Fig. 3.5-1 where the transconductance characteristic of a MOSFET in saturation is shown with the square root of current plotted as a function of the gate-source voltage. When the gate-source voltage reaches the value designated as  $V_{ON}$  (this relates to the SPICE model

formulation), the current changes from square-law to an exponential-law behavior. It is the objective of this section to present two models suitable for the subthreshold region. The first is the SPICE LEVEL 3 [2] model for computer simulation while the second is useful for hand calculations.

In the SPICE LEVEL 3 model, the transition point from the region of strong inversion to the weak inversion characteristic of the MOS device is designated as  $V_{ON}$  and is greater than  $V_T$ .  $V_{ON}$  is given by

$$V_{ON} = V_T + fast \quad (3.5-1)$$

where

$$fast = \frac{kT}{q} \left[ 1 + \frac{q \cdot NFS}{COX} + \frac{GAMMA \cdot f_s (PHI + v_{SB})^{1/2} + f_n (PHI + v_{SB})}{2(PHI + v_{SB})} \right] \quad (3.5-2)$$

$NFS$  is a parameter used in the evaluation of  $V_{ON}$  and can be extracted from measurements. The drain current in the weak inversion region,  $v_{GS} < V_{ON}$ , is given as

$$i_{DS} = i_{DS}(V_{ON}, v_{DE}, v_{SB}) \exp \left( \frac{v_{GS} - V_{ON}}{fast} \right) \quad (3.5-3)$$

where  $i_{DS}$  is given as [from Eq. (3.4.1), with  $v_{GS}$  replaced with  $V_{ON}$ ]

$$i_{DS} = BETA \left[ V_{ON} - V_T - \left( \frac{1 + f_b}{2} \right) v_{DE} \right] \cdot v_{DE} \quad (3.5-4)$$

For hand calculations, a simple model describing weak inversion operation is given as

$$i_D \cong \frac{W}{L} I_{D0} \exp \left( \frac{v_{GS}}{n(kT/q)} \right) \quad (3.5-5)$$

where the term  $n$  is the subthreshold slope factor, and  $I_{D0}$  is a process-dependent parameter that is dependent also on  $v_{SB}$  and  $V_T$ . These two terms are best extracted from experimental

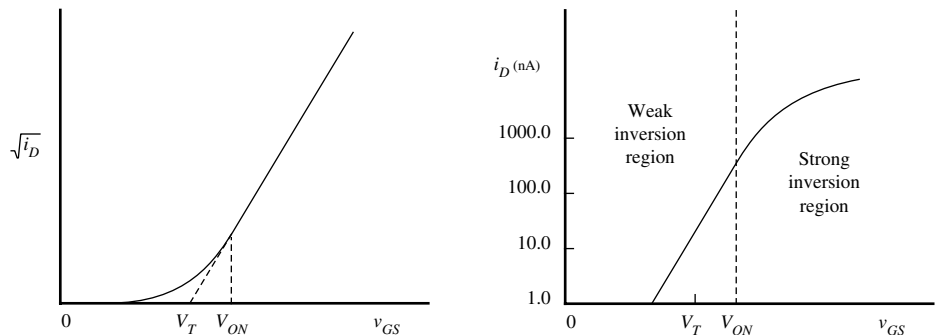


**Figure 3.5-1** Weak inversion characteristics of the MOS transistor as modeled by Eq. (3.5-4).

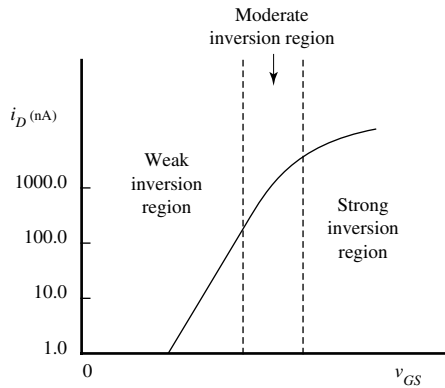


**Figure 3.5-2** The three regions of operation of an MOS transistor.

data. Typically  $n$  is greater than 1 and less than 3 ( $1 < n < 3$ ). The point at which a transistor enters the weak inversion region can be approximated as

$$v_{gs} < V_T + n \frac{kT}{q} \quad (3.5-6)$$

Unfortunately, the model equations given here do not properly model the transistor as it makes the transition from strong to weak inversion. In reality, there is a transition region of operation between strong and weak inversion called the “moderate inversion” region [17]. This is illustrated in Fig. 3.5-2. A complete treatment of the operation of the transistor through this region is given in the literature [17,18].

It is important to consider the temperature behavior of the MOS device operating in the subthreshold region. As is the case for strong inversion, the temperature coefficient of the threshold voltage is negative in the subthreshold region. The variation of current due to temperature of a device operating in weak inversion is dominated by the negative temperature coefficient of the threshold voltage. Therefore, for a given gate–source voltage, subthreshold current increases as the temperature increases. This is illustrated in Fig. 3.5-3 [19].

Operation of the MOS device in the subthreshold region is very important when low-power circuits are desired. A whole class of CMOS circuits have been developed based on the weak inversion operation characterized by the above model [20–23]. We will consider some of these circuits in later chapters.

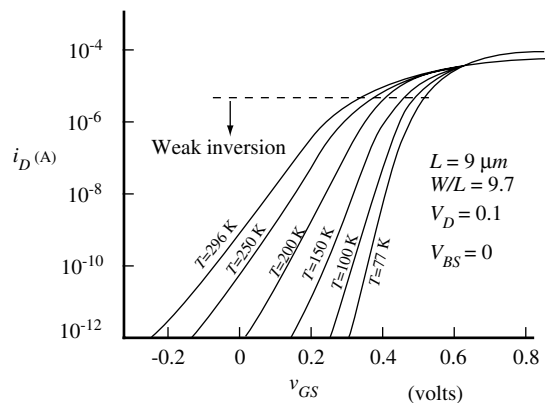


**Figure 3.5-3** Transfer characteristics of a long-channel device as a function of temperature. (Copyright © 1977).



## 3.6 SPICE Simulation of MOS Circuits

The objective of this section is to show how to use SPICE to verify the performance of an MOS circuit. It is assumed that the reader already has experience using SPICE to simulate circuits containing resistors, capacitors, sources, and so on. This section will extend the reader's knowledge to include the application of MOS transistors into SPICE simulations. The models used in this section are the LEVEL 1 and LEVEL 3 models.

In order to simulate MOS circuits in SPICE, two components of the SPICE simulation file are needed. They are *instance* declarations and model descriptions. Instance declarations are simply descriptions of MOS devices appearing in the circuit along with characteristics unique to each instance. A simple example that shows the minimum required terms for a transistor instance follows:

---

```
M1 3 6 7 0 NCH W=100U L=1U
```

---

Here, the first letter in the instance declaration, *M*, tells SPICE that the instance is an MOS transistor (just like *R* tells SPICE that an instance is a resistor). The *1* makes this instance unique (different from *M2*, *M99*, etc.). The four numbers following *M1* specify the nets (or nodes) to which the drain, gate, source, and substrate (bulk) are connected. These nets have a specific order as indicated below:

---

```
M<number> <DRAIN> <GATE> <SOURCE> <BULK> . . .
```

---

Following the net numbers is the model name governing the character of the particular instance. In the example given above, the model name is *NCH*. There must be a model description somewhere in the simulation file that describes the model *NCH*. The transistor width and length are specified for the instance by the  $W = 100U$  and  $L = 1U$  expressions. The default units for width and length are meters so the *U* following the number 100 is a multiplier of  $10^{-6}$ . (Recall that the following multipliers can be used in SPICE: *M*, *U*, *N*, *P*, *F*, for  $10^{-3}$ ,  $10^{-6}$ ,  $10^{-9}$ ,  $10^{-12}$ ,  $10^{-15}$ , respectively.)

Additional information can be specified for each instance. Some of these are:

- Drain area and periphery (*AD* and *PD*)
- Source area and periphery (*AS* and *PS*)
- Drain and source resistance in squares (*NRD* and *NRS*)
- Multiplier designating how many devices are in parallel (*M*)
- Initial conditions (for initial transient analysis)

Drain and source area and periphery terms are used in calculating depletion capacitance and diode currents (remember, the drain and source are pn diodes to the bulk or well). The numbers of squares of resistance in the drain and source (*NRD* and *NRS*) are used to calculate the drain and source resistances for the transistor. The multiplier designator is very important and thus deserves extended discussion here.

In Appendix B, layout matching techniques were developed. One of the fundamental principles described was the “unit-matching” principle. This principle prescribes that when one device needs to be *M* times larger than another device, then the larger device should be

made from  $M$  units of the smaller device. In the layout, the larger device would be drawn using  $M$  copies of the smaller device—all of them in parallel (i.e., all of the gates tied together, all of the drains tied together, and all of the sources tied together). In SPICE, one must account for the multiple components tied in parallel. One way to do this would be to instantiate the larger device by instantiating  $M$  of the smaller devices. A more convenient way to handle this is to use the multiplier parameter when the larger device is instantiated. Figure 3.6-1 illustrates two methods for implementing a 2X device (unit device implied). In Fig. 3.6-1(a) the correct way to instantiate the device in SPICE is

---

```
M1 3 2 1 0 NCH W=20U L=1U
```

---

whereas in Fig. 3.6-1(b) the correct SPICE instantiation is

---

```
M1 3 2 1 0 NCH W=10U L=1U M=2
```

---

Clearly, from the point of view of matching (again, it is implied that an attempt is made to achieve a 2:1 ratio), case (b) is the better choice and thus the instantiation with the multiplier is required. For the sake of completeness, it should be noted that the following pair of instantiations are equivalent to the use of the multiplier:

---

```
M1A 3 2 1 0 NCH W=10U L=1U
M1B 3 2 1 0 NCH W=10U L=1U
```

---

Some SPICE simulators offer additional terms further describing an instance of an MOS transistor.

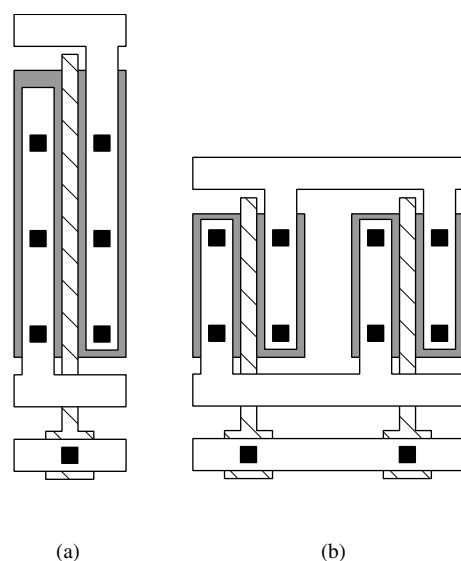


**Figure 3.6-1** (a) M1 3 2 1 0 NCH  
W = 20U L = 1U. (b) M1 3 2 1 0  
NCH W = 10U L = 1U M = 2.

A SPICE simulation file for an MOS circuit is incomplete without a description of the model to be used to characterize the MOS transistors used in the circuit. A model is described by placing a line in the simulation file using the following format:

---

```
.MODEL <MODEL NAME> <MODEL TYPE> <MODEL PARAMETERS>
```

---

The model line must always begin with `.MODEL` and be followed by a model name such as `NCH` in our example. Following the model name is the model type. The appropriate choices for model type in MOS circuits is either `PMOS` or `NMOS`. The final group of entries is model parameters. If no entries are provided, SPICE uses a default set of model parameters. Except for the crudest of simulations, you will always want to avoid the default parameters. Most of the time you should expect to get a model from the foundry where the wafers will be fabricated, or from the modeling group within your company. For times where it is desired to check hand calculations that were performed using the simple model (LEVEL 1 model) it is useful to know the details of entering model information. An example model description line follows.

---

```
.MODEL NCH NMOS LEVEL=1 VT0=1 KP=50U GAMMA=0.5
+LAMBDA=0.01
```

---

In this example, the model name is `NCH` and the model type is `NMOS`. The model parameters dictate that the LEVEL 1 model is used with `VT0`, `KP`, `GAMMA`, and `LAMBDA` specified. Note that the `+` is SPICE syntax for a continuation line.

The information on the model line is much more extensive and will be covered in this and the following paragraphs. The model line is preceded by a period to flag the program that this line is not a component. The model line identifies the model `LEVEL` (e.g., `LEVEL=1`) and provides the electrical and process parameters. If the user does not input the various parameters, default values are used. These default values are indicated in the user's guide for the version of SPICE being used (e.g., SmartSpice). The LEVEL 1 model parameters were covered in Section 3.1 and are the zero-bias threshold voltage, `VT0` ( $V_{T0}$ ), in volts extrapolated to  $i_D = 0$  for large devices; the intrinsic transconductance parameter, `KP` ( $K'$ ), in amperes/volt<sup>2</sup>; the bulk threshold parameter, `GAMMA` ( $\gamma$ ), in volt<sup>1/2</sup>; the surface potential at strong inversion, `PHI` ( $2\phi_F$ ), in volts; and the channel length modulation parameter, `LAMBDA` ( $\lambda$ ), in volt<sup>-1</sup>. Values for these parameters can be found in Table 3.1-2.

Sometimes, one would rather let SPICE calculate the above parameters from the appropriate process parameters. This can be done by entering the surface state density in cm<sup>-2</sup> (NSS); the oxide thickness in meters (TOX); the surface mobility, `U0` ( $\mu_0$ ), in cm<sup>2</sup>/V-s; and the substrate doping in cm<sup>-3</sup> (NSUB). The equations used to calculate the electrical parameters are

$$VT0 = \phi_{MS} - \frac{q(NSS)}{(\epsilon_{ox}/TOX)} + \frac{(2q \cdot \epsilon_{Si} \cdot NSUB \cdot PHI)^{1/2}}{(\epsilon_{ox}/TOX)} + PHI \quad (3.6-1)$$

$$KP = U0 \frac{\epsilon_{ox}}{TOX} \quad (3.6-2)$$

$$GAMMA = \frac{(2q \cdot \epsilon_{Si} \cdot NSUB)^{1/2}}{(\epsilon_{ox}/TOX)} \quad (3.6-3)$$

and

$$\text{PHI} = |2\phi_F| = \frac{2kT}{q} \ln \left( \frac{\text{NSUB}}{n_i} \right) \quad (3.6-4)$$

LAMBDA is not calculated from the process parameters for the LEVEL 1 model. The constants for silicon, given in Table 3.1-1, are contained within the SPICE program and do not have to be entered.

The next model parameters considered are those that were considered in Section 3.2. The first parameters considered were associated with the bulk–drain and bulk–source pn junctions. These parameters include the reverse current of the drain–bulk or source–bulk junctions in A (IS) or the reverse-current density of the drain–bulk or source–bulk junctions in A/m<sup>2</sup> (JS). JS requires the specification of AS and AD on the model line. If IS is specified, it overrides JS. The default value of IS is usually 10<sup>−14</sup> A. The next parameters considered in Section 3.2 were the drain ohmic resistance in ohms (RD), the source ohmic resistance in ohms (RS), and the sheet resistance of the source and drain in ohms/square (RSH). RSH is overridden if RD or RS is entered. To use RSH, the values of NRD and NRS must be entered on the model line.

The drain–bulk and source–bulk depletion capacitors can be specified by the zero-bias bulk junction bottom capacitance in farads per m<sup>2</sup> of junction area (CJ). CJ requires NSUB and assumes a step junction using a formula similar to Eq. (2.2-12). Alternately, the drain–bulk and source–bulk depletion capacitances can be specified using Eqs. (3.2-5) and (3.2-6). The necessary parameters include the zero-bias bulk–drain junction capacitance (CBD) in farads, the zero-bias bulk–source junction capacitance (CBS) in farads, the bulk junction potential (PB) in volts, the coefficient for forward-bias depletion capacitance (FC), the zero-bias bulk junction sidewall capacitance (CJSW) in farads per meter of junction perimeter, and the bulk junction sidewall capacitance grading coefficient (MJSW). If CBD or CBS is specified, then CJ is overridden. The values of AS, AD, PS, and PD must be given on the device line to use the above parameters. Typical values of these parameters are given in Table 3.2-1.

The next parameters discussed in Section 3.2 were the gate overlap capacitances. These capacitors are specified by the gate–source overlap capacitance (CGSO) in farads/meter, the gate–drain overlap capacitance (CGDO) in farads/meter, and the gate–bulk overlap capacitance (CGBO) in farads/meter. Typical values of these overlap capacitances can be found in Table 3.2-1. Finally, the noise parameters include the flicker noise coefficient (KF) and the flicker noise exponent (AF). Typical values of these parameters are 10<sup>−28</sup> and 1, respectively.

Additional parameters not discussed in Section 3.4 include the type of gate material (TPG), the thin oxide capacitance model flag, and the coefficient of channel charge allocated to the drain (XQC). The choices for TPG are +1 if the gate material is opposite to the substrate, −1 if the gate material is the same as the substrate, and 0 if the gate material is aluminum. A charge-controlled model is used in the SPICE simulator if the value of the parameter XQC has a value smaller than or equal to 0.5. This model attempts to keep the sum of charge associated with each node equal to zero. If XQC is larger than 0.5, charge conservation is not guaranteed.

In order to illustrate its use and to provide examples for the novice user to follow, several examples will be given showing how to use SPICE to perform various simulations.

**Example  
3.6-1****Use of SPICE to Simulate MOS Output Characteristics**

Use SPICE to obtain the output characteristics of the n-channel transistor shown in Fig. 3.6-2 using the LEVEL 1 model and the parameter values of Table 3.1-2. The output curves are to be plotted for drain–source voltages from 0 to 5 V and for gate–source voltages of 1, 2, 3, 4, and 5 V. Assume that the bulk voltage is zero.

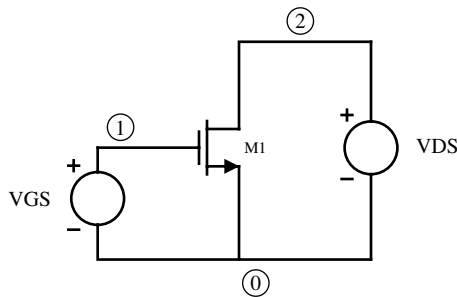
**Figure 3.6-2** Circuit for Example 3.6-1.**SOLUTION**

Table 3.6-1 shows the input file for SPICE to solve this problem. The first line is a title for the simulation file and must be present. The lines not preceded by “.” define the interconnection of the circuit. The second line describes how the transistor is connected, defines the model to be used, and gives the  $W$  and  $L$  values. Note that because the units are meters, the suffix  $U$  is used to convert to  $\mu\text{m}$ . The third and fourth lines describe the independent voltages.  $V_{DS}$  and  $V_{GS}$  are used to bias the MOSFET. The fifth line is the model description for  $M1$ . The remaining lines instruct SPICE to perform a dc sweep and print desired results. `.DC` asks for a dc sweep. In this particular case, a nested dc sweep is specified in order to avoid seven consecutive analyses. The `.DC. . .` line will set  $V_{GS}$  to a value of 1 V and then sweep  $V_{DS}$  from 0 to 5 V in increments of 0.2 V. Next, it will increment  $V_{GS}$  to 2 V and repeat the  $V_{DS}$  sweep. This is continued until five  $V_{DS}$  sweeps have been made with the desired values of  $V_{GS}$ . The `.PRINT. . .` line directs the program to print the values of the dc sweeps. The last line of every SPICE input file must be `.END`. Figure 3.6-3 shows the output plot of this analysis.

**Table 3.6-1** SPICE Input File for Example 3.6-1

```
Ex. 3.6-1 Use of SPICE to Simulate MOS Output
M1 2 1 0 0 MOS1 W=5U L=1.0U
VDS 2 0 5
VGS 1 0 1
.MODEL MOS1 NMOS VTO=0.7 KP=110U GAMMA=0.4 LAMBDA=0.04 PHI=0.7
.DC VDS 0 5 0.2 VGS 1 5 1
.PRINT DC V(2) I(VDS)
.END
```

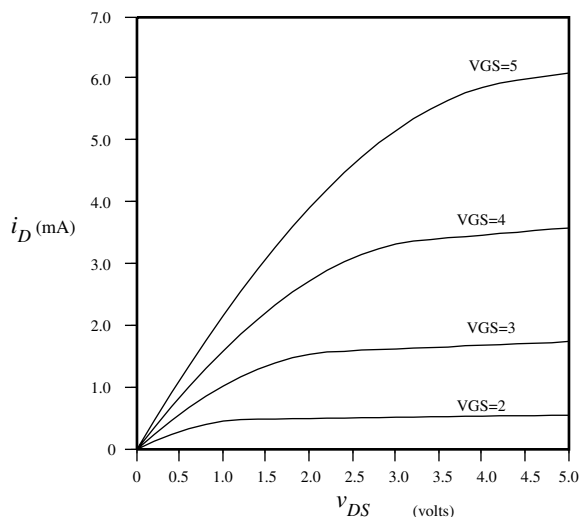


**Figure 3.6-3** Output from Example 3.6-1.

### Example 3.6-2

#### DC Analysis of Fig. 3.6-4

Use the SPICE simulator to obtain a plot of the value of  $v_{OUT}$  as a function of  $v_{IN}$  of Fig. 3.6-4. Identify the dc value of  $v_{IN}$  that gives  $v_{OUT} = 2.5$  V.

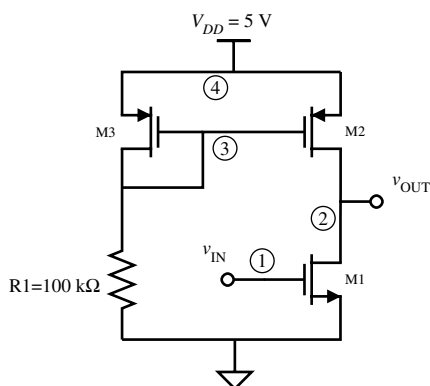


**Figure 3.6-4** A simple MOS amplifier for Example 3.6-2.

#### SOLUTION

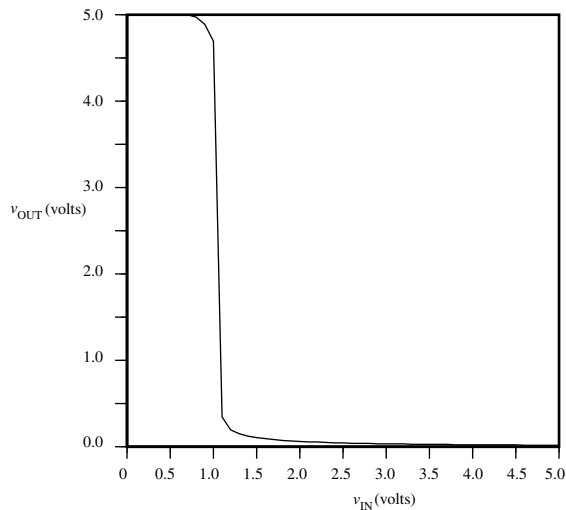
The input file for SPICE is shown in Table 3.6-2. It follows the same format as the previous example except that two types of transistors are used. These models are designated by `MOSN` and `MOSP`. A dc sweep is requested starting from  $v_{IN} = 0$  V and going to +5 V. Figure 3.6-5 shows the resulting output of the dc sweep.  $v_{OUT} = 2.5$  V when  $v_{IN} \approx 1.0$  V.

**Table 3.6-2** SPICE Input File for Example 3.6-2

```

Ex. 3.6-2 DC Analysis of Fig. 3.6-4
M1 2 1 0 0 MOSN W=5U L=1U
M2 2 3 4 4 MOSP W=5U L=1U
M3 3 3 4 4 MOSP W=5U L=1U
R1 3 0 100K
VDD 4 0 DC 5.0
VIN 1 0 DC 5.0
.MODEL MOSN NMOS VTO=0.7 KP=110U GAMMA=0.4 LAMBDA=0.04 PHI=0.7
.MODEL MOSP PMOS VTO=-0.7 KP=50U GAMMA=0.57 LAMBDA=0.05 PHI=0.8
.DC VIN 0 5 0.1
.PRINT DC V(2)
.END

```

**Figure 3.6-5** Output of Example 3.6-2.**Example  
3.6-3****AC Analysis of Fig. 3.6-4**

Use SPICE to obtain a small-signal frequency response of  $V_{out}(\omega)/V_{in}(\omega)$  when the amplifier is biased in the transition region. Assume that a 5 pF capacitor is attached to the output of Fig. 3.6-4 and find the magnitude and phase response over the frequency range of 100 Hz to 100 MHz.

**SOLUTION**

The SPICE input file for this example is shown in Table 3.6-3. It is important to note that  $V_{IN}$  has been defined as both an ac and a dc voltage source with a dc value of 1.07 V. If the dc voltage were not included, SPICE would find the dc solution for  $V_{IN} = 0$  V, which is not in the transition region. Therefore, the small-signal solution would not be evaluated in the transition region.

**Table 3.6-3 SPICE Input File for Example 3.6-3**

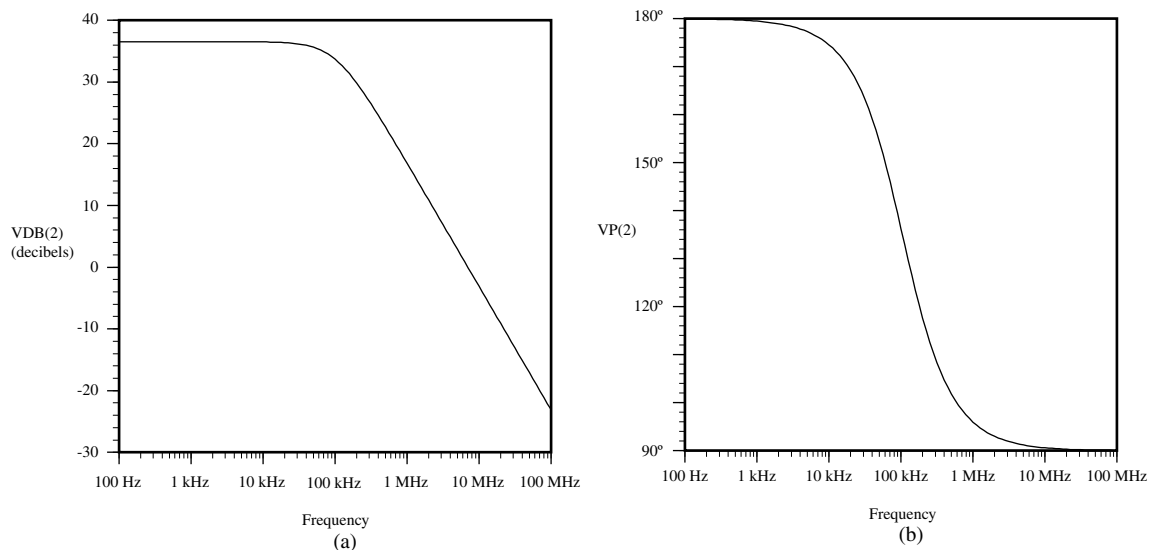
```

Ex. 3.6-3 AC Analysis of Fig. 3.6-4
M1 2 1 0 0 MOSN W=5U L=1U
M2 2 3 4 4 MOSP W=5U L=1U
M3 3 3 4 4 MOSP W=5U L=1U
CL 2 0 5P
R1 3 0 100K
VDD 4 0 DC 5.0
VIN 1 0 DC 1.07 AC 1.0
.MODEL MOSN NMOS VTO = 0.7 KP = 110U GAMMA = 0.4 LAMBDA = 0.04
+ PHI = 0.7
.MODEL MOSP PMOS VTO = -0.7 KP = 50U GAMMA = 0.57 LAMBDA = 0.05
+ PHI = 0.8
.AC DEC 20 100 100MEG
.OP
.PRINT AC VM(2) VDB(2) VP(2)
.END

```

Once the dc solution has been evaluated, the amplitude of the signal applied as the ac input has no influence on the simulation. Thus, it is convenient to use ac inputs of unity in order to treat the output as a gain quantity. Here, we have assumed an ac input of 1.0 V peak.

The simulation desired is defined by the `.AC DEC 20 100 100MEG` line. This line directs SPICE to make an ac analysis over a log frequency with 20 points per decade from 100 Hz to 100 MHz. The `.OP` option has been added to print out the dc voltages of all circuit nodes in order to verify that the ac solution is in the desired region. The program will calculate the linear magnitude, dB magnitude, and phase of the output voltage. Figures 3.6-6(a) and 3.6-6(b) show the magnitude (dB) and the phase of this simulation.

**Figure 3.6-6** (a) Magnitude response and (b) phase response of Example 3.6-3.

**Example  
3.6-4****Transient Analysis of Fig. 3.6-4**

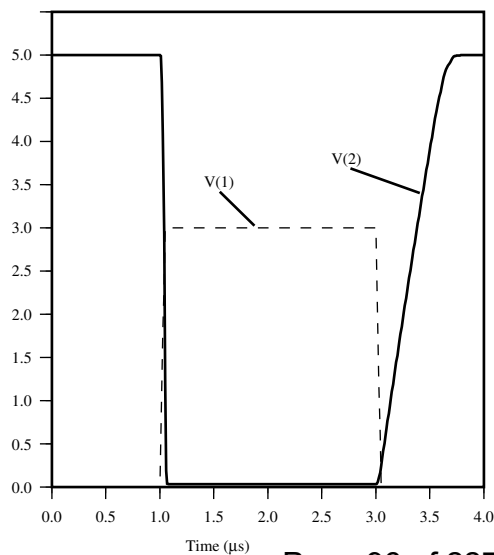
The last simulation to be made with Fig. 3.6-4 is the transient response to an input pulse. This simulation will include the 5 pF output capacitor of the previous example and will be made from time 0 to 4  $\mu\text{s}$ .

**SOLUTION**

Table 3.6-4 shows the SPICE input file. The input pulse is described using the piecewise linear capability (PWL) of SPICE. The output desired is defined by `.TRAN 0.01U 4U`, which asks for a transient analysis from 0 to 4  $\mu\text{s}$  at points spaced every 0.01  $\mu\text{s}$ . The output will consist of both  $v_{\text{IN}}(t)$  and  $v_{\text{OUT}}(t)$  and is shown in Fig. 3.6-7. The use of an asterisk at the beginning of a line causes that line to be ignored.

**Table 3.6-4 SPICE Output for Example 3.6-4**

```
Ex. 3.6-4 Transient Analysis of Fig. 3.6-4
M1 2 1 0 0 MOSN W=5U L=1U
M2 2 3 4 4 MOSP W=5U L=1U
M3 3 3 4 4 MOSP W=5U L=1U
CL 2 0 5P
R1 3 0 100K
VDD 4 0 DC 5.0
VIN 1 0 PWL(0 0V 1U 0V 1.05U 3V 3U 3V 3.05U 0V 6U 0V)
*VIN 1 0 DC 21.07 AC 1.0
.MODEL MOSN NMOS VTO = 0.7 KP = 110U GAMMA = 0.4 LAMBDA = 0.04
+ PHI = 0.7
.MODEL MOSP PMOS VTO = -0.7 KP = 50U GAMMA = 0.57 LAMBDA =
+ 0.05 PHI = 0.8
.TRAN 0.01U 4U
.PRINT TRAN V(2) V(1)
.END
```

**Figure 3.6-7** Transient response of Example 3.6-4.

The above examples will serve to introduce the reader to the basic ideas and concepts of using the SPICE program. In addition to what the reader has distilled from these examples, a useful set of guidelines is offered, which has resulted from extensive experience in using SPICE:

1. Never use a simulator unless you know the range of answers beforehand.
2. Never simulate more of the circuit than is necessary.
3. Always use the simplest model that will do the job.
4. Always start a dc solution from the point at which the majority of the devices are on.
5. Use a simulator in the same manner as you would make the measurement on the bench.
6. Never change more than one parameter at a time when using the simulator for design.
7. Learn the basic operating principles of the simulator so that you can enhance its capability. Know how to use its options.
8. Watch out for syntax problems like O and 0.
9. Use the correct multipliers for quantities.
10. Use common sense.

Most problems with simulators can be traced back to a violation of one or more of these guidelines.

There are many SPICE simulators in use today. The discussion here focused on the more general versions of SPICE and should apply in most cases. However, there is nothing fundamental about the syntax or use of a circuit simulator, so it is prudent to carefully study the manual of the SPICE simulator you are using.



### 3.7 Summary

This chapter has tried to give the reader the background necessary to be able to simulate CMOS circuits. The approach used has been based on the SPICE simulation program. This program normally has three levels of MOS models that are available to the user. The function of models is to solve for the dc operating conditions and then use this information to develop a linear small-signal model. Section 3.1 described the LEVEL 1 model used by SPICE to solve for the dc operating point. This model also uses the additional model parameters presented in Section 3.2. These parameters include bulk resistance, capacitance, and noise. A small-signal model that was developed from the large-signal model was described in Section 3.3. These three sections represent the basic modeling concepts for MOS transistors.

Models for computer simulation were presented. The SPICE LEVEL 3 model, which is effective for device lengths of 0.8  $\mu\text{m}$  and greater, was covered. The BSIM3v3 model, which is effective for deep-submicron devices, was introduced. Large-signal models suitable for weak inversion were also described. Further details of these models and other models are found in the references for this chapter. A brief background of simulation methods was presented in Section 3.6. Simulation of MOS circuits using SPICE was discussed. After studying this chapter, the reader should be able to use the model information presented along with a SPICE simulator to analyze MOS circuits. This ability will be very important in the remainder of this text. It will be used to verify intuitive design approaches and to perform analyses beyond the scope of the techniques presented. One of the important aspects of modeling is to determine the model parameters that best fit the MOS process being used. Appendix C will be devoted to this subject.

## Problems

- 3.1-1. Sketch to scale the transconductance characteristics of an enhancement n-channel device if  $V_T = 0.7$  V and  $I_D = 500$   $\mu$ A when  $V_{GS} = 5$  V in saturation. Choose values of  $V_{GS} = 1, 2, 3, 4$ , and 5 V. Assume that the channel modulation parameter is zero.
- 3.1-2. Sketch to scale the output characteristics of an enhancement p-channel device if  $V_T = -0.7$  V and  $I_D = -500$   $\mu$ A when  $V_{GS} = -5$  V in saturation. Choose values of  $V_{GS} = -1, -2, -3, -4$ , and  $-6$  V. Assume that the channel modulation parameter is zero.
- 3.1-3. In Table 3.1-2, why is  $\gamma_P$  greater than  $\gamma_N$  for an n-well, CMOS technology?
- 3.1-4. A large-signal model for the MOSFET that features symmetry for the drain and source is given as

$$i_D = K' \frac{W}{L} \{ [(v_{GS} - V_{TS})^2 u(v_{GS} - V_{TS})] - [(v_{GD} - v_{TD})^2 u(v_{GD} - v_{TD})] \}$$

where  $u(x)$  is 1 if  $x$  is greater than or equal to zero and 0 if  $x$  is less than zero (step function), and  $V_{TX}$  is the threshold voltage evaluated from the gate to  $X$ , where  $X$  is either  $S$  (source) or  $D$  (drain). Sketch this model in the form of  $i_D$  versus  $v_{DS}$  for a constant value of  $v_{GS}$  ( $v_{GS} > V_{TS}$ ) and identify the saturated and nonsaturated regions. Be sure to extend this sketch for both positive and negative values of  $v_{DS}$ . Repeat the sketch of  $i_D$  versus  $v_{DS}$  for a constant value of  $v_{GD}$  ( $v_{GD} > V_{TD}$ ). Assume that both  $V_{TS}$  and  $V_{TD}$  are positive.

- 3.1-5. Equations (3.1-12) and (3.1-18) describe the MOS model in the nonsaturation and saturation region, respectively. These equations do not agree at the point of transition between saturation and nonsaturation regions. For hand calculations, this is not an issue, but for computer analysis, it is. How would you change Eq. (3.1-18) so that it would agree with Eq. (3.1-12) at  $v_{DS} = v_{DS}(\text{sat})$ ?

- 3.2-1. Using the values of Tables 3.1-1 and 3.2-1, calculate the values of  $C_{GB}$ ,  $C_{GS}$ , and  $C_{GD}$  for an MOS device that has a  $W$  of 5  $\mu$ m and an  $L$  of 1  $\mu$ m for all three regions of operation.
- 3.2-2. Find  $C_{BX}$  at  $|V_{BX}| = 0$  V and 0.75 V (with the junction always reverse biased) of Fig. P3.2-2. The values of Table 3.2-1 apply to the MOS device, where  $FC = 0.5$  and  $PB = 1$  V. Assume the device is n-channel and repeat for a p-channel device.

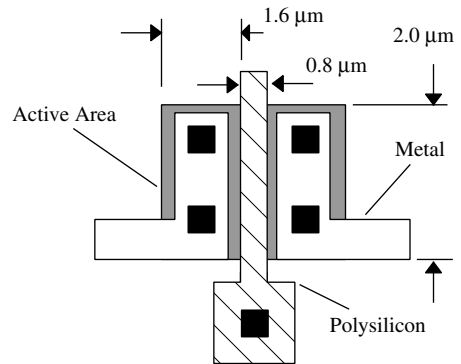


Figure P3.2-2

- 3.2-3. Calculate the values of  $C_{GB}$ ,  $C_{GS}$ , and  $C_{GD}$  for an n-channel device with a length of 1  $\mu$ m and a width of 5  $\mu$ m. Assume  $V_D = 2$  V,  $V_G = 2.4$  V, and  $V_S = 0.5$  V and let  $V_B = 0$  V. Use model parameters from Tables 3.1-1, 3.1-2, and 3.2-1.
- 3.2-4. A layout of an NMOS transistor is shown in Fig. P3.2-4. (a) Find the values of  $RD$  and  $RS$  in the schematic shown if the sheet resistance of the  $n^+$  is 35  $\Omega/\text{sq.}$  and the resistance of a single contact is 1  $\Omega$ . (b) Find the values of  $C_{BD}$  and  $C_{BS}$  assuming the transistor is cutoff and the drain and source are at ground potential if  $CJ$  and  $CJSW$  for an NMOS transistor are  $770 \times 10^{-6}$  F/m<sup>2</sup> and  $380 \times 10^{-12}$  F/m. Assume the capacitors are lumped and appear on the source/drain side of the bulk resistors in part (a). (c) What is the  $W$  and  $L$  of this transistor? (d) If the overlap capacitor/unit length is  $220 \times 10^{-12}$  F/m, what is  $C_{GD}$ ?

- 3.2-5 A CMOS inverter is shown in Fig. P3.2-5 along with the top view of the circuit layout assuming a p-well CMOS technology. If this inverter is driving an identical inverter with the same layout, find the magnitude of the pole at the output of the first inverter ( $v_x$ ) and the input of the second inverter which

is equal to the reciprocal product of the sum of all capacitances connected to this node and the output resistance which is assumed to be  $1\text{ M}\Omega$ . Express this pole magnitude in Hz. Use the information in Table 3.2-1 to calculate the capacitances.

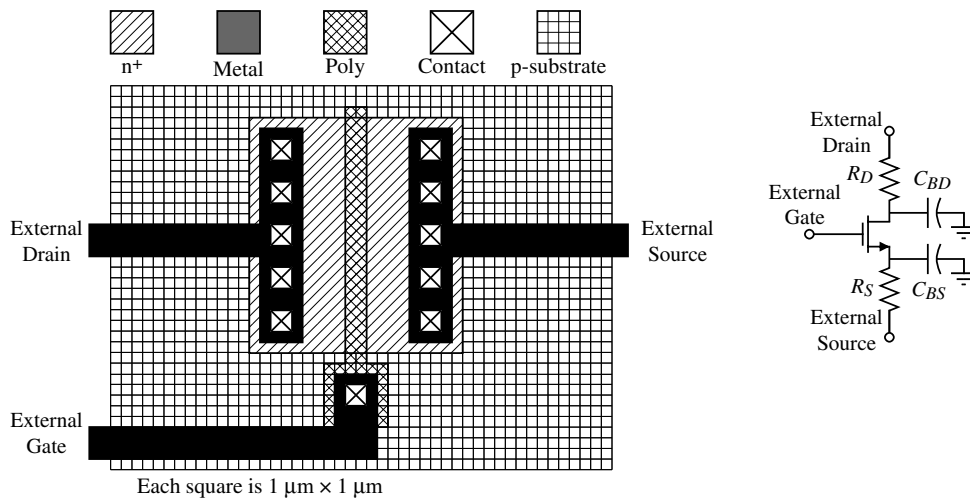


Figure P3.2-4

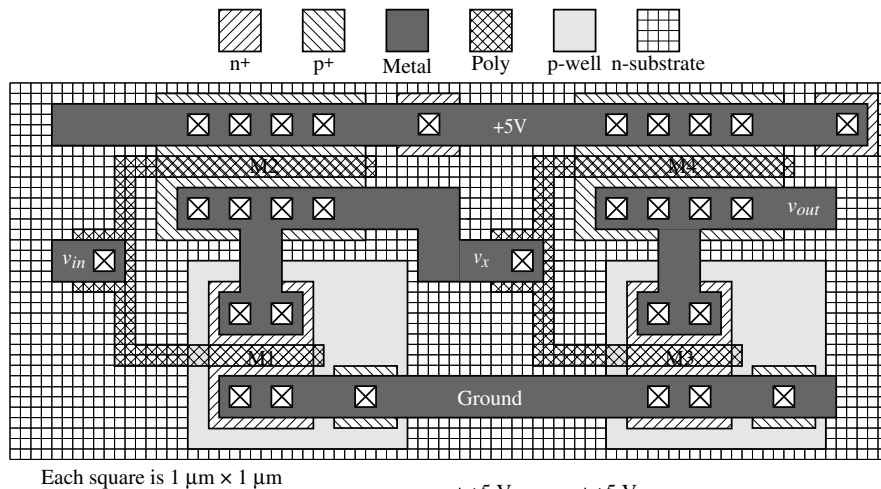


Figure P3.2-5

- 3.3-1. Calculate the transfer function  $v_{out}(s)/v_{in}(s)$  for the circuit shown in Fig. P3.3-1. The  $W/L$  of M1 is  $2\ \mu\text{m}/0.8\ \mu\text{m}$  and the  $W/L$  of M2 is  $4\ \mu\text{m}/4\ \mu\text{m}$ . Note that this is a small-signal analysis and the input voltage has a dc value of 2 V.

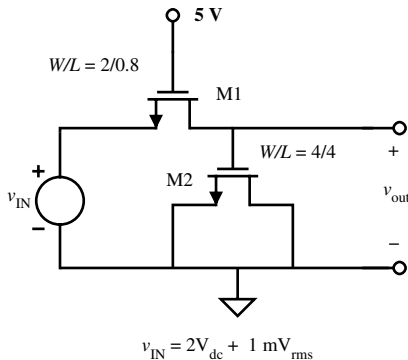


Figure P3.3-1

- 3.3-2. Design a low-pass filter patterned after the circuit in Fig. P3.3-1 that achieves a  $-3\ \text{dB}$  frequency of  $100\ \text{kHz}$ .
- 3.3-3. Repeat Examples 3.3-1 and 3.3-2 if the  $W/L$  ratio is  $100\ \mu\text{m}/10\ \mu\text{m}$ . Assume  $\lambda = 0.01\ \text{V}^{-1}$ .
- 3.3-4. Find the complete small-signal model for an n-channel transistor with the drain at 4 V, gate at 4 V, source at 2 V, and bulk at 0 V. Assume the model parameters from Tables 3.1-1, 3.1-2, and 3.2-1, and  $W/L = 10\ \mu\text{m}/1\ \mu\text{m}$ .
- 3.3-5. Consider the circuit in Fig P3.3-5. It is a parallel connection of  $n$  MOSFET transistors. Each transistor has the same length,  $L$ , but each transistor can have a different width,  $W$ . Derive an expression for  $W$  and  $L$  for a single transistor that replaces, and is equivalent to, the multiple parallel transistors.

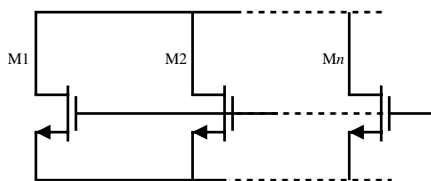


Figure P3.3-5

- 3.3-6. Consider the circuit in Fig P3.3-6. It is a series connection of  $n$  MOSFET transistors. Each transistor has the same width,  $W$ , but each transistor can have a different length,  $L$ . Derive an expression for  $W$  and  $L$  for a single transistor that replaces, and is equivalent to, the multiple series transistors. When using the simple model, you must ignore body effect.

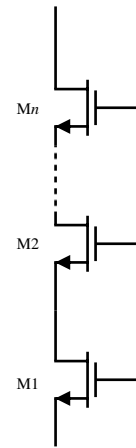


Figure P3.3-6

- 3.5-1. Calculate the value for  $V_{ON}$  for an NMOS transistor in weak inversion assuming that  $f_s$  and  $f_n$  can be approximated to be unity (1.0).
- 3.5-2. Develop an expression for the small-signal transconductance of an MOS device operating in weak inversion using the large-signal expression of Eq. (3.5-5).
- 3.5-3. Another way to approximate the transition from strong inversion to weak inversion is to find the current at which the weak inversion transconductance and the strong inversion transconductance are equal. Using this method and the approximation for drain current in weak inversion [Eq. (3.5-5)], derive an expression for drain current at the transition between strong and weak inversion.
- 3.6-1. Consider the circuit illustrated in Fig. P3.6-1. (a) Write a SPICE netlist that describes this circuit. (b) Repeat part (a) with M2 being  $2\ \mu\text{m}/1\ \mu\text{m}$  and M3 and M2 are ratio matched, 1:2.

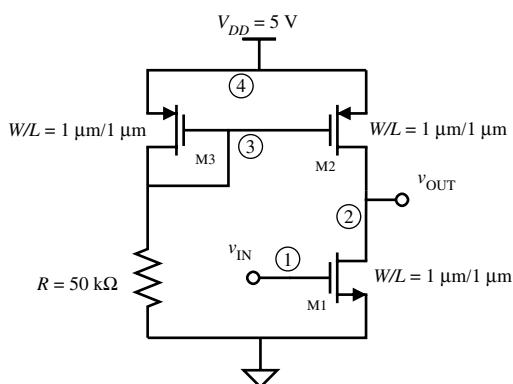


Figure P3.6-1

- 3.6-2. Use SPICE to perform the following analyses on the circuit shown in Fig. P3.6-1: (a) Plot  $v_{OUT}$  versus  $v_{IN}$  for the nominal parameter set shown. (b) Separately, vary  $K'$  and  $V_T$  by  $+10\%$  and repeat part (a)—four simulations.

| Parameter | n-Channel | p-Channel | Units                    |
|-----------|-----------|-----------|--------------------------|
| $V_T$     | 0.7       | -0.7      | V                        |
| $K'$      | 110       | 50        | $\mu\text{A}/\text{V}^2$ |
| $\lambda$ | 0.04      | 0.05      | $\text{V}^{-1}$          |

- 3.6-3. Use SPICE to plot  $i_2$  as a function of  $v_2$  when  $i_1$  has values of 10, 20, 30, 40, 50, 60,

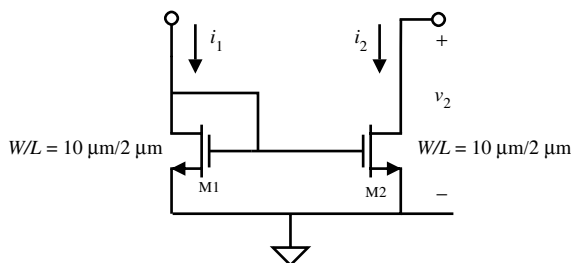


Figure P3.6-3

- and 70  $\mu\text{A}$  for Fig. P3.6-3. The maximum value of  $v_2$  is 5 V. Use the model parameters of  $V_T = 0.7$  V and  $K' = 110 \mu\text{A}/\text{V}^2$  and  $\lambda = 0.01 \text{ V}^{-1}$ . Repeat with  $\lambda = 0.04 \text{ V}^{-1}$ .
- 3.6-4. Use SPICE to plot  $i_D$  as a function of  $v_{DS}$  for values of  $v_{GS} = 1, 2, 3, 4,$  and 5 V for an n-channel transistor with  $V_T = 1$  V,  $K' = 110 \mu\text{A}/\text{V}^2$ , and  $\lambda = 0.04 \text{ V}^{-1}$ . Show how SPICE can be used to generate and plot these curves simultaneously as illustrated by Fig. 3.1-3.
- 3.6-5. Repeat Example 3.6-1 if the transistor of Fig. 3.6-2 is a PMOS having the model parameters given in Table 3.1-2.
- 3.6-6. Repeat Examples 3.6-2 through 3.6-4 for the circuit of Fig. 3.6-4 if  $R1 = 200 \text{ k}\Omega$ .

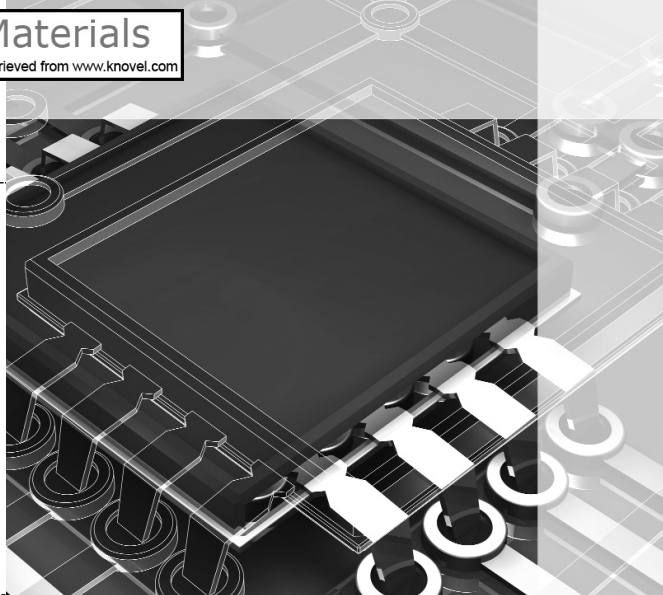
## References

1. Y. Tsividis, "Problems with Modeling of Analog MOS LSI," *IEDM*, pp. 274–277, 1982.
2. *Star-Hspice User's Manual*. Fremont, CA: Avant! 2000.
3. Yuhua Cheng and Chenming Hu, *MOSFET Modeling & BSIM3 User's Guide*. Norwell, MA: Kluwer Academic Publishers, 1999.
4. C. T. Sah, "Characteristics of the Metal-Oxide-Semiconductor Transistor," *IEEE Transactions on Electron Devices*, Vol. ED-11, No. 7, pp. 324–345, July 1964.
5. H. Shichman and D. Hodges, "Modelling and Simulation of Insulated-Gate Field-Effect Transistor Switching Circuits," *IEEE J. Solid-State Circuits*, Vol. SC-3, No. 3, pp. 285–289, Sept. 1968.
6. P. R. Gray, P. J. Hurst, S. H. Lewis, and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 4th ed. New York: John Wiley & Sons, 2001, pp. 62–63.
7. A. Vladimerescu, A. R. Newton, and D. O. Pederson, *SPICE Version 2G.0 User's Guide*. University of California, Berkeley, Sept. 1980.
8. S. M. Sze and Kwok K. Ng, *Physics of Semiconductor Devices*, 3rd ed. Hoboken, NJ: John Wiley & Sons, Inc., 2007, pp. 36–38.
9. D. R. Alexander, R. J. Antinone, and G. W. Brown, *SPICE Modelling Handbook*, Report BDM/A-77-071-TR, BDM Corporation, 2600 Yale Blvd., Albuquerque, NM 87106.

10. P. R. Gray and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 2nd ed. New York: John Wiley & Sons, 1984, p. 646.
11. P. E. Allen and E. Sanchez-Sinencio, *Switched Capacitor Circuits*. New York: Van Nostrand Reinhold, 1984, p. 589.
12. *SmartSpice Modeling Manual*, Vols. 1 and 2. Santa Clara, CA: Silvaco International, Sept. 1999.
13. Daniel P. Foty, *MOSFET Modeling with SPICE: Principles and Practice*. Scarborough, ON: Prentice Hall Canada, 1997.
14. G. Massobrio and P. Antognetti, *Semiconductor Device Modeling with SPICE*, 2nd ed. New York: McGraw-Hill, 1993.
15. F. H. Gaensslen and R. C. Jaeger, "Temperature Dependent Threshold Behavior of Depletion Mode MOSFET's," *Solid-State Electronics*, Vol. 22, No. 4, pp. 423–430, 1979.
16. J. R. Pierret, *A MOS Parameter Extraction Program for the BSIM Model*, Electronics Research Laboratory, University of California, Berkeley, CA, 94720. Memorandum No. UCB/ERL M84/99, Nov. 21, 1984.
17. Y. Tsividis, "Moderate Inversion In MOS Devices," *Solid State Electronics*, Vol. 25, No. 11, pp. 1099–1104, 1982.
18. P. Antognetti, D. D. Caviglia, and E. Profumo, "CAD Model for Threshold and Subthreshold Conduction in MOSFET's," *IEEE J. Solid-State Circuits*, Vol. SC-17, No. 2, pp. 454–458, June 1982.
19. S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed. New York: John Wiley and Sons, 1981.
20. E. Vittoz and J. Fellrath, "CMOS Analog Integrated Circuits Based on Weak Inversion Operation," *IEEE J. Solid-State Circuits*, Vol. SC-12, No. 3, pp. 231–244, June 1977.
21. M. G. DeGrauwe, J. Rigenants, E. Vittoz, and H. J. DeMan, "Adaptive Biasing CMOS Amplifiers," *IEEE J. Solid-State Circuits*, Vol. SC-17, No. 3, pp. 522–528, June 1982.
22. W. Steinhagen and W. L. Engl, "Design of Integrated Analog CMOS Circuits—A Multichannel Telemetry Transmitter," *IEEE J. Solid-State Circuits*, Vol. SC-13, No. 6, pp. 799–805, Dec. 1978.
23. Y. Tsividis and R. Ulmer, "A CMOS Voltage Reference," *IEEE J. Solid-State Circuits*, Vol. SC-13, No. 6, pp. 774–778, Dec. 1978.

## CHAPTER 4

# Analog CMOS Subcircuits



From the viewpoint of Table 1.1-2, the previous two chapters have provided the background for understanding the technology and modeling of CMOS devices and components compatible with the CMOS process. The next step toward our objective—methodically developing the subject of CMOS analog circuit design—is to develop subcircuits. These simple circuits consist of one or more transistors and generally perform only one function. A subcircuit is typically combined with other simple circuits to generate a more complex circuit function. Consequently, the circuits of this and the next chapter can be considered as building blocks.

The *operational amplifier*, or *op amp*, to be covered in Chapters 6 and 7, is a good example of how simple circuits are combined to perform a complex function. Figure 4.0-1 presents a hierarchy showing how an operational amplifier—a complex circuit—might be related to various simple circuits. Working our way backward, we note that one of the stages of an op amp is the differential amplifier. The differential amplifier consists of simple circuits that might include a current sink, a current-mirror load, and a source-coupled pair. Another stage of the op amp is a second gain stage, which might consist of an inverter and a current-sink load. If the op amp is to be able to drive a low-impedance load, an output stage is necessary. The output stage might consist of a source follower and a current-sink load. It is also necessary to provide a stabilized bias for each of the previous stages. The biasing stage could consist of a current sink and current mirrors to distribute the bias currents to the other stages.

The subject of basic CMOS analog circuits has been divided into two chapters to avoid one lengthy chapter and yet provide sufficient detail. Chapter 4 covers the simpler subcircuits, including the MOS switch, active loads, current sinks/sources, current mirrors and current amplifiers, and voltage and current references. Chapter 5 will examine more complex circuits like CMOS amplifiers. That chapter represents a natural extension of the material presented in Chapter 4. Taken together, these two chapters are fundamental for the analog CMOS designer's understanding and capability, as most designs will start at this level and progress upward to synthesize the more complex circuits and systems of Table 1.1-2.

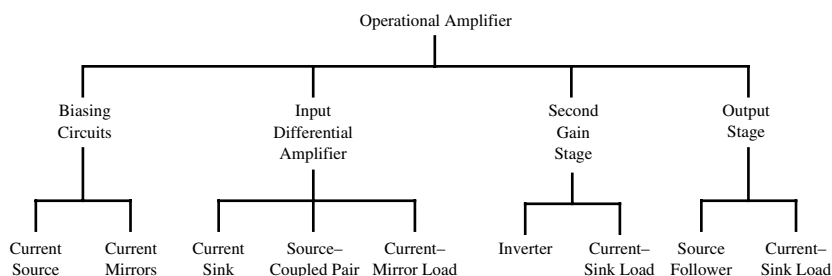


**Figure 4.0-1** Illustration of the hierarchy of analog circuits for an operational amplifier.

## 4.1 MOS Switch

The switch finds many applications in integrated-circuit design. In analog circuits, the switch is used primarily to connect capacitors in various configurations. This application is called *switched capacitor circuits*. The switch is also useful for multiplexing, modulation, and a number of other applications. The switch is used as a transmission gate in digital circuits and adds a dimension of flexibility not found in standard logic circuits. The objective of this section is to study the characteristics of switches that are compatible with CMOS integrated circuits.

We begin with the characteristics of a voltage-controlled switch. Figure 4.1-1 shows a model for such a device. The voltage  $v_C$  controls the state of the switch—ON or OFF. The voltage-controlled switch is a three-terminal network with terminals  $A$  and  $B^*$  comprising the switch and terminal  $C$  providing the means of applying the control voltage  $v_C$ . The most important characteristics of a switch are its ON resistance,  $r_{ON}$ , and its OFF resistance,  $r_{OFF}$ . Ideally,  $r_{ON}$  is zero and  $r_{OFF}$  is infinite. Reality is such that  $r_{ON}$  is never zero and  $r_{OFF}$  is never infinite. Moreover, these values are never constant with respect to terminal conditions. In general, switches can have some form of voltage offset, which is modeled by  $V_{OS}$  of Fig. 4.1-1.  $V_{OS}$  represents the small voltage that may exist between terminals  $A$  and  $B$  when the switch

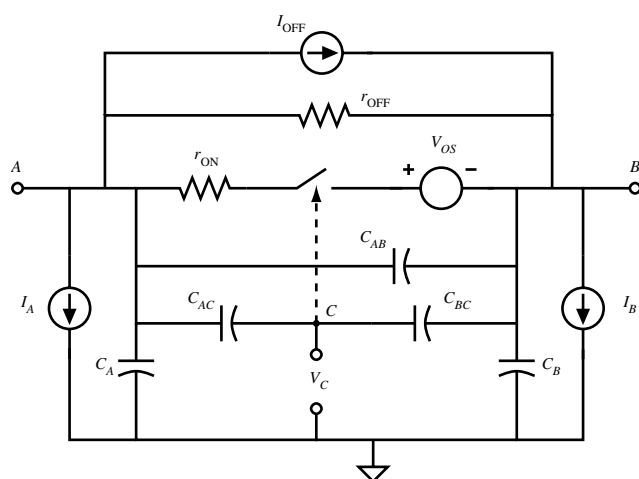


**Figure 4.1-1** Model for a non-ideal switch.

\*The reader is cautioned not to confuse  $B$  used in this section with  $B$  used for the bulk terminal (see Fig. 3.1-1).

is in the ON state and the current is zero.  $I_{\text{OFF}}$  represents the leakage current that may flow in the OFF state of the switch. Currents  $I_A$  and  $I_B$  represent leakage currents from the switch terminals to ground (or some other supply potential). The polarities of the offset sources and leakage currents are not known and have arbitrarily been assigned the directions indicated in Fig. 4.1-1. The parasitic capacitors are an important consideration in the application of analog sampled-data circuits. Capacitors  $C_A$  and  $C_B$  are the parasitic capacitors between the switch terminals  $A$  and  $B$  and ground. Capacitor  $C_{AB}$  is the parasitic capacitor between the switch terminals  $A$  and  $B$ . Capacitors  $C_{AC}$  and  $C_{BC}$  are parasitic capacitors that may exist between the voltage-control terminal  $C$  and the switch terminals  $A$  and  $B$ . Capacitors  $C_{AC}$  and  $C_{BC}$  contribute to the effect called *charge feedthrough*—where a portion of the control voltage appears at the switch terminals  $A$  and  $B$ .

One advantage of MOS technology is that it provides a good switch. Figure 4.1-2 shows an MOS transistor that is to be used as a switch. Its performance can be determined by comparing Fig. 4.1-1 with the large-signal model for the MOS transistor. We see that either terminal,  $A$  or  $B$ , can be the drain or the source of the MOS transistor depending on the terminal voltages (e.g., for an n-channel transistor, if terminal  $A$  is at a higher potential than  $B$ , then terminal  $A$  is the drain and terminal  $B$  is the source). The ON resistance consists of the series combination of  $r_D$ ,  $r_S$ , and whatever channel resistance exists. Typically, by design, the contribution from  $r_D$  and  $r_S$  is small such that the primary consideration is the channel resistance. An expression for the channel resistance can be found as follows. In the ON state of the switch, the voltage across the switch should be small and  $v_{GS}$  should be large. Therefore, the MOS device is assumed to be in the nonsaturation region. Equation (3.1-1), repeated below, is used to model this state:

$$i_D = \frac{K'W}{L} \left[ (v_{GS} - V_T)v_{DS} - \frac{v_{DS}^2}{2} \right] \quad (4.1-1)$$

where  $v_{DS}$  is less than  $v_{GS} - V_T$  but greater than zero. ( $v_{GS}$  becomes  $v_{GD}$  if  $v_{DS}$  is negative.) The small-signal channel resistance is given as

$$r_{\text{ON}} = \frac{1}{\partial i_D / \partial v_{DS}} \bigg|_Q = \frac{L}{K'W(V_{GS} - V_T - V_{DS})} \quad (4.1-2)$$

where  $Q$  in Eq. (4.1-2) designates the quiescent point of the transistor. Figure 4.1-3 illustrates the drain current of an n-channel transistor as a function of the voltage across the drain and source terminals, plotted for equal increasing steps of  $V_{GS}$  for  $W/L = 5/1$ . This figure illustrates some very important principles about MOS transistor operation. Note that the curves are not symmetrical about  $V_1 = 0$ . This is because the transistor terminals (drain and source) switch roles as  $V_1$  crosses 0 V. For example, when  $V_1$  is positive, node  $B$  is the drain and node  $A$  is the source and  $V_{BS}$  is fixed at  $-2.5$  V and  $V_{GS}$  is fixed as well (for a given  $V_G$ ). When  $V_1$  is negative, node  $B$  is the source and node  $A$  is the drain and as  $V_1$  continues to decrease,  $V_{BS}$  decreases and  $V_{GS}$  increases, resulting in an increase in current.

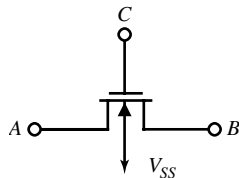


**Figure 4.1-2** An n-channel transistor used as a switch.

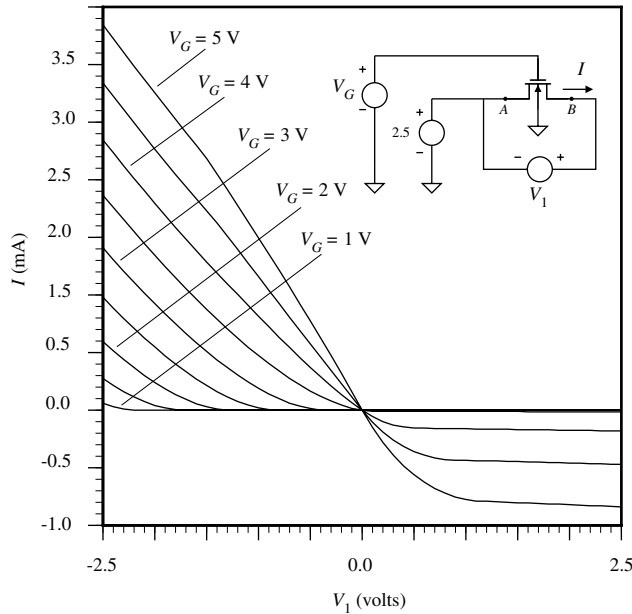


**Figure 4.1-3** I–V characteristic of an n-channel transistor operating as a switch.

A plot of  $r_{ON}$  as a function of  $V_{GS}$  is shown in Fig. 4.1-4 for  $V_{DS} = 0.1$  V and for  $W/L = 1, 2, 5,$  and  $10$ . It is seen that a lower value of  $r_{ON}$  is achieved for larger values of  $W/L$ . When  $V_{GS}$  approaches  $V_T$  ( $V_T = 0.7$  V in this case),  $r_{ON}$  approaches infinity because the switch is turning off.

When  $V_{GS}$  is less than or equal to  $V_T$ , the switch is OFF and  $r_{OFF}$  is ideally infinite. Of course, it is never infinite, but because it is so large, the performance in the OFF state is dominated by the drain–bulk and source–bulk leakage current as well as subthreshold leakage from drain to source. The leakage from drain and source to bulk is primarily due to the pn junction leakage current and is modeled in Fig. 4.1-1 as  $I_A$  and  $I_B$ . Typically, this leakage current is on the order of  $1 \text{ fA}/\mu\text{m}^2$  at room temperature and doubles for every  $8^\circ\text{C}$  increase (see Example 2.5-1).

The offset voltage modeled in Fig. 4.1-1 does not exist in MOS switches and thus is not a consideration in MOS switch performance. The capacitors  $C_A$ ,  $C_B$ ,  $C_{AC}$ , and  $C_{BC}$  of Fig. 4.1-1

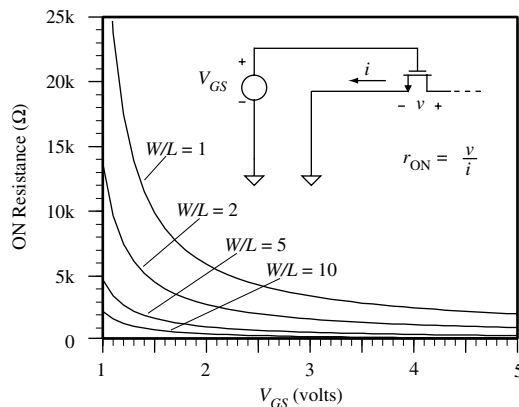


**Figure 4.1-4** Illustration of ON resistance for an n-channel transistor.

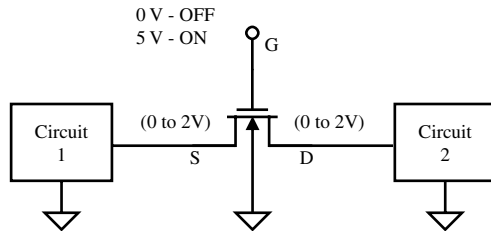


**Figure 4.1-5** Application of an n-channel transistor as a switch with typical terminal voltages indicated.

correspond directly to the capacitors  $C_{BS}$ ,  $C_{BD}$ ,  $C_{GS}$ , and  $C_{GD}$  of the MOS transistor (see Fig. 3.2-1).  $C_{AB}$  is small for the MOS transistor and is usually negligible.

One important aspect of the switch is the range of voltages on the switch terminals compared to the control voltage. For the n-channel MOS transistor we see that the gate voltage must be considerably larger than either the drain or source voltage in order to ensure that the MOS transistor is ON. (For the p-channel transistor, the gate voltage must be considerably less than either the drain or source voltage.) Typically, the bulk is taken to the most negative potential for the n-channel switch (positive for the p-channel switch). This requirement can be illustrated as follows for the n-channel switch. Suppose that the ON voltage of the gate is the positive power supply  $V_{DD}$ . With the bulk to ground this should keep the n-channel switch ON until the signal on the switch terminals (which should be approximately identical at the source and drain) approaches  $V_{DD} - V_T$ . As the signal approaches  $V_{DD} - V_T$  the switch begins to turn OFF. Typical voltages used for an n-channel switch are shown in Fig. 4.1-5, where the switch is connected between the two networks shown.

Consider the use of a switch to charge a capacitor as shown in Fig. 4.1-6. An n-channel transistor is used as a switch and  $V_\phi$  is the control voltage (clock) applied to the gate. The ON resistance of the switch is important during the charge transfer phase of this circuit. For example, when  $V_\phi$  goes high ( $V_\phi > v_{in} + V_T$ ), M1 connects  $C$  to the voltage source  $v_{in}$ . The equivalent circuit at this time is shown in Fig. 4.1-7. It can be seen that  $C$  will charge to  $v_{in}$  with the time constant of  $r_{ON}C$ . For successful operation  $r_{ON}C \ll T$ , where  $T$  is the time  $V_\phi$  is high. Clearly,  $r_{ON}$  varies greatly with  $v_{GS}$ , as illustrated in Fig. 4.1-4. The worst-case value for  $r_{ON}$  (the highest value), during the charging of  $C$ , is when  $v_{DS} = 0$  and  $v_{GS} = V_\phi - v_{in}$ . This value should be used when sizing the transistor to achieve the desired charging time.

Consider a case where the time  $V_\phi$  is high is  $T = 0.1 \mu s$  and  $C = 0.2 pF$ ; then  $r_{ON}$  must be less than  $100 k\Omega$  if sufficient charge transfer occurs in five time constants. For a 5 V clock swing and  $v_{in}$  of 2.5 V, the MOS device of Fig. 4.1-4 with  $W = L$  gives  $r_{ON}$  of approximately  $6.4 k\Omega$ , which is sufficiently small to transfer the charge in the desired time. It is desirable

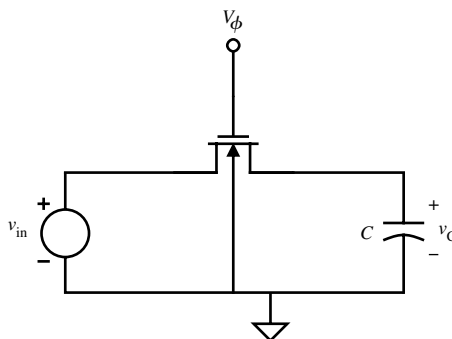


**Figure 4.1-6** An application of an MOS switch.

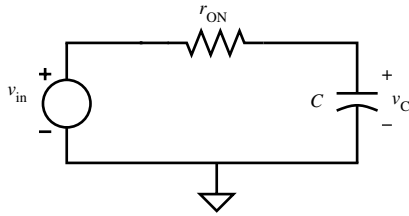


**Figure 4.1-7** Model for the ON state of the switch in Fig. 4.1-6.

to keep the switch size as small as possible (minimize  $W \times L$ ) to minimize charge feedthrough from the gate.

The OFF state of the switch has little influence on the performance of the circuit in Fig. 4.1-6 except for the leakage current. Figure 4.1-8 shows a sample-and-hold circuit where the leakage current can create serious problems. If  $C_H$  is not large enough, then in the hold mode where the MOS switch is OFF the leakage current can charge or discharge  $C_H$  a significant amount.

Switches are often used to switch capacitors from one configuration to another. In this application, the capacitance of the switch becomes important. There are two types of capacitors that are associated with the MOSFET switch: capacitors with one terminal connected to the gate and the drain- and source-to-bulk depletion capacitors. The gate capacitors can cause charge injection onto the capacitors that they are switching while the depletion capacitors influence the value of the switched capacitors as parasitics. Since the impact of the depletion capacitors can be eliminated by circuit techniques [1] we will deal only with the influence of the gate capacitors.

Figure 4.1-9 is a useful model for studying the charge injection of the MOSFET switch. Figure 4.1-9(b) illustrates modeling a transistor with the channel symbolized as a resistor,  $R_{channel}$ , and gate-channel coupling capacitance, denoted  $C_{channel}$ . The values of  $C_{channel}$  and  $R_{channel}$  depend on the terminal conditions of the device. The gate-channel coupling is distributed across the channel as is the channel resistance,  $R_{channel}$ . In addition to the channel capacitance, there is the overlap capacitance,  $C_{GSO}$  and  $C_{GDO}$ .

It is convenient to approximate the total channel capacitance by splitting it into two capacitors of equal size placed at the gate-source and gate-drain terminals as illustrated in Fig. 4.1-9(c). These capacitors represent the charge that is stored in the channel of the MOSFET switch. In addition to these capacitors, there are the overlap capacitors,  $C_{GSO}$  and  $C_{GDO}$ . We will keep these two capacitors distinct because they cause different forms of charge transfer during the operation of the switch. For example, when the switch is ON, there is a charge in the channel, which must be removed before the switch can turn OFF. The flow of this charge from the channel to the switch terminals is called *channel charge injection*. In addition to the channel charge injection, the clock at the gate of the switch is connected to the switch terminals by the overlap capacitors. If there is any change of voltage across these

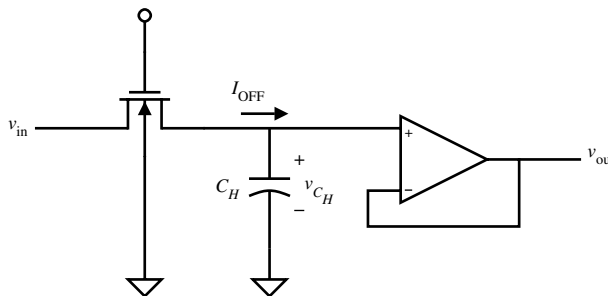


**Figure 4.1-8** Example of the influence of  $I_{OFF}$  in a sample-and-hold circuit.

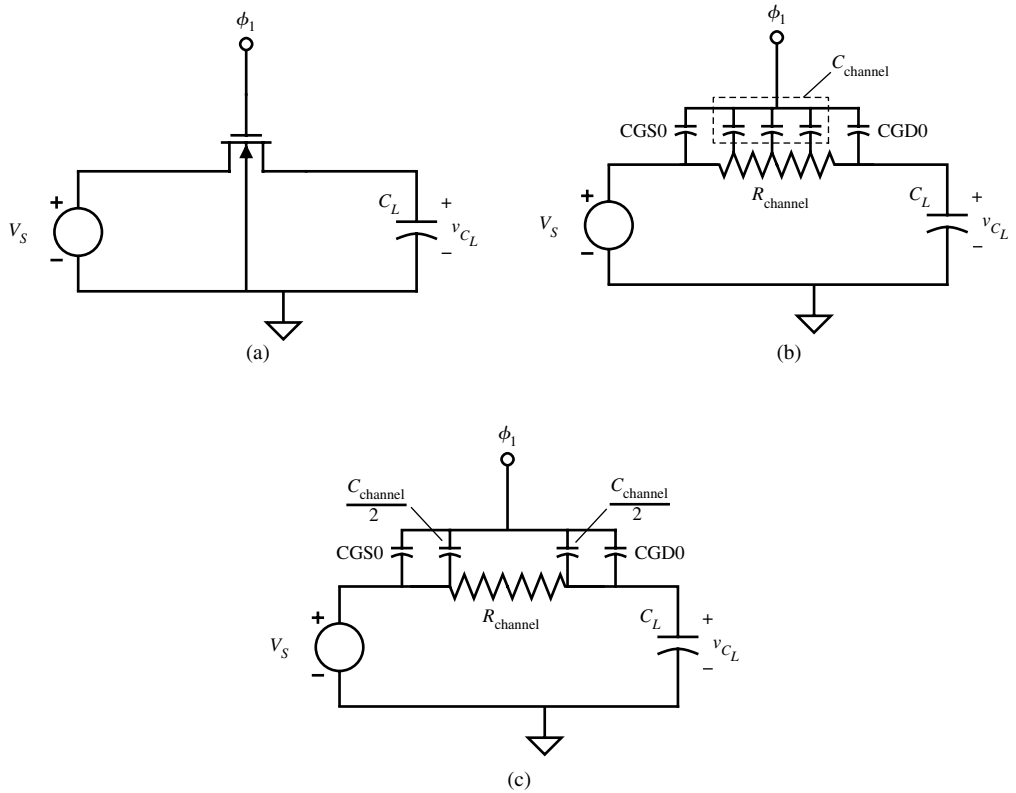


**Figure 4.1-9** (a) Simple switch circuit useful for studying charge injection. (b) Distributed model for the transistor switch. (c) Lumped model of Fig. 4.1-9(a).

capacitors, there will be a corresponding charge flow, which is called *clock feedthrough*. We see there can be clock feedthrough whenever there is a change in the gate voltage with respect to the switch terminals regardless of whether the switch is ON or OFF.

Let us first consider the channel charge injection using Fig. 4.1-9(a). When the switch is ON, a charge will be stored in the channel, which is equal to

$$Q_{channel} = -WLC_{ox}(V_H - v_{in} - V_T) \quad (4.1-3)$$

where  $V_H$  is the value of the clock waveform applied to the gate when the switch is ON, for example,  $V_H \approx V_{DD}$  and  $v_{in} = V_S$ . When the switch turns OFF, this charge is injected into the source and drain terminals. If we assume the charge splits evenly, then half goes to the input voltage source,  $v_{in}$ , and the other half to  $C_L$ . The half that goes through  $C_L$  creates a change in voltage across  $C_L$  given as

$$\Delta V = \frac{Q_{ch}}{2C_L} = \frac{-WLC_{ox}(V_H - v_{in} - V_T)}{2C_L} \quad (4.1-4)$$

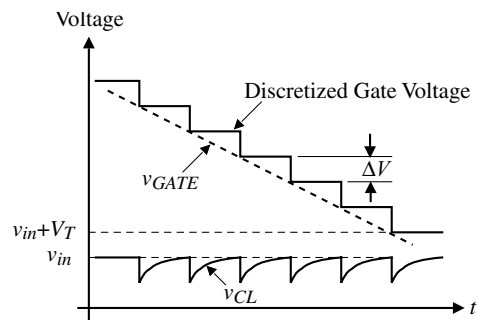
The charge injection does not influence  $v_{in}$  because it is a voltage source. We see from Eq. (4.1-4) that the channel charge injection is a function of the input voltage and will vary as the input voltage changes. We also know that the channel charge will have a time constant that is equal to the resistance of the channel times the channel capacitance.

Clock feedthrough (sometimes called *charge injection* and *charge feedthrough*) allows charge to be transferred from the gate signal (which is generally a clock) to the drain and source nodes—an undesirable but unavoidable effect. Clock feedthrough involves a complex process whose resulting effects depend on a number of factors such as the layout of the transistor, its dimensions, impedance levels at the source and drain nodes, and gate waveform. It is hopeless to attempt to describe all of these effects precisely analytically—we have computers to do that! Nevertheless, it is useful to develop a qualitative understanding of this important effect.

For the circuit in Fig. 4.1-9, it is of interest to examine the charge flow in the switch during a high-to-low transition at the gate of  $\phi_1$ . Moreover, it is convenient to consider two cases regarding the gate transition—a fast transition time and a slow transition time. Consider the slow-transition case first (what is meant by slow and fast will be covered shortly). As the gate is falling, some charge is being injected into the channel. However, initially, the transistor remains on so that whatever charge is injected flows in the input voltage source,  $V_S$ . None of this charge will appear on the load capacitor,  $C_L$ . As the gate voltage falls, at some point, the transistor turns off (when the gate voltage reaches  $V_S + V_T$ ). After the transistor turns off, there is no other path for the charge injection other than through  $C_L$ .

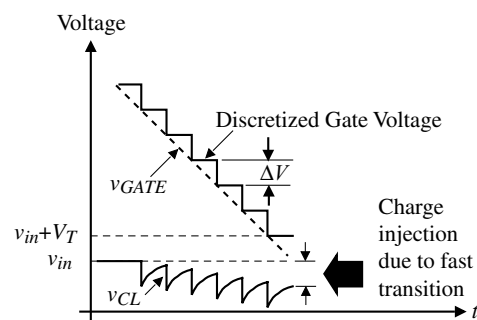
For the fast case, the time constant associated with the channel resistance and the channel capacitance limits the amount of charge that can flow to the source voltage so that some of the channel charge that is injected while the transistor is on contributes to the total charge on  $C_L$ .

To develop some intuition about the fast and slow cases, it is useful to model the gate voltage as a piecewise constant waveform (a quantized waveform) and consider the charge flow at each transition as illustrated in Figs. 4.1-10(a) and 4.1-10(b). In these figures, the range of voltage across  $C_L$  illustrated represents the period while the transistor is on. In both



(a)

**Figure 4.1-10** Illustration of (a) slow ramp and (b) fast ramp using a quantized voltage ramp to illustrate the effects due to the time constant of the channel resistance and capacitance.



(b)

cases, the quantized voltage step,  $\Delta V$ , is the same, but the time between steps is different. The voltage across  $C_L$  is observed to be an exponential whose time constant is due to the channel resistance and channel capacitance and does not change from the fast case to the slow case.

Analytical expressions have been derived that describe the approximate operation of a transistor in the slow and fast regimes and include both channel charge injection and clock feedthrough [2]. Consider the gate voltage traversing from  $V_H$  to  $V_L$  (e.g., 5.0 V to 0.0 V, respectively) described in the time domain as

$$v_G = V_H - Ut \quad (4.1-5)$$

where  $U$  is the magnitude of the slope of  $v_G(t)$ . When operating in the slow regime defined by the relationship

$$\frac{\beta V_{HT}^2}{2C_L} \gg U \quad (4.1-6)$$

where  $V_{HT}$  is defined as

$$V_{HT} = V_H - V_S - V_T \quad (4.1-7)$$

the error (the difference between the desired voltage  $V_S$  and the actual voltage  $V_{CL}$ ) due to charge injection can be described as

$$V_{\text{error}} = \left( \frac{W \cdot \text{CGDO} + \frac{C_{\text{channel}}}{2}}{C_L} \right) \sqrt{\frac{\pi U C_L}{2\beta}} + \frac{W \cdot \text{CGDO}}{C_L} (V_S + V_T - V_L) \quad (4.1-8)$$

In the fast switching regime defined by the relationship

$$\frac{\beta V_{HT}^2}{2C_L} \ll U \quad (4.1-9)$$

the error voltage is given as

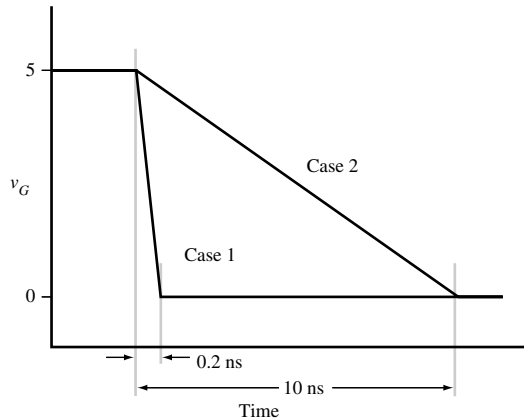
$$V_{\text{error}} = \left( \frac{W \cdot \text{CGDO} + \frac{C_{\text{channel}}}{2}}{C_L} \right) \left( V_{HT} - \frac{\beta V_{HT}^3}{6UC_L} \right) + \frac{W \cdot \text{CGDO}}{C_L} (V_S + V_T - V_L) \quad (4.1-10)$$

The following example illustrates the application of the charge feedthrough model given by Eqs. (4.1-5) through (4.1-10).

#### Example 4.1-1

#### Calculation of Charge Feedthrough Error

Calculate the effect of charge feedthrough on the circuit shown in Fig. 4.1-9, where  $V_S = 1.0$  V,  $C_L = 200$  fF,  $W/L = 0.8 \mu\text{m}/0.8 \mu\text{m}$ , and  $V_G$  is given for two cases illustrated below. Use model parameters from Tables 3.1-2 and 3.2-1. Neglect  $\Delta L$  and  $\Delta W$  effects.



### SOLUTION

*Case 1:* The first step is to determine the value of  $U$  in the expression

$$v_G = V_H - Ut$$

For a transition from 5 V to 0 V in 0.2 ns,  $U = 25 \times 10^9$  V/s.

In order to determine operating regime, the following relationship must be tested:

$$\frac{\beta V_{HT}^2}{2C_L} \gg U \text{ for slow} \quad \text{or} \quad \frac{\beta V_{HT}^2}{2C_L} \ll U \text{ for fast}$$

Observing that there is a backbias on the transistor switch affecting  $V_T$ ,  $V_{HT}$  is

$$V_{HT} = V_H - V_S - V_T = 5 - 1 - 0.887 = 3.113$$

giving

$$\frac{\beta V_{HT}^2}{2C_L} = \frac{110 \times 10^{-6} \times 3.113^2}{2 \times 200 \times 10^{-15}} = 2.66 \times 10^9 \ll 25 \times 10^9 \quad \text{thus fast regime}$$

Applying Eq. (4.1-10) for the fast regime yields

$$V_{\text{error}} = \left( \frac{176 \times 10^{-18} + \frac{1.58 \times 10^{-15}}{2}}{200 \times 10^{-15}} \right) \left( 3.113 - \frac{3.32 \times 10^{-3}}{30 \times 10^{-3}} \right) + \frac{176 \times 10^{-18}}{200 \times 10^{-15}} (5 + 0.887 - 0)$$

$$V_{\text{error}} = 19.7 \text{ mV}$$

*Case 2:* The first step is to determine the value of  $U$  in the expression

$$v_G = V_H - Ut$$

For a transition from 5 V to 0 V in 10 ns,  $U = 5 \times 10^8$ , thus indicating the slow regime according to the following test:

$$2.66 \times 10^9 \gg 5 \times 10^8$$

$$V_{\text{error}} = \left( \frac{176 \times 10^{-18} + \frac{1.58 \times 10^{-15}}{2}}{200 \times 10^{-15}} \right) \left( \frac{314 \times 10^{-6}}{220 \times 10^{-6}} \right)^{1/2} + \frac{176 \times 10^{-18}}{200 \times 10^{-15}} (5 + 0.887 - 0)$$

$$V_{\text{error}} = 10.95 \text{ mV}$$

This example illustrates the application of the charge feedthrough model. The reader should be cautioned not to expect Eqs. (4.1-5) through (4.1-10) to give precise answers regarding the amount of charge feedthrough one should expect in an actual circuit. The model should be used as a guide in understanding the effects of various circuit elements and terminal conditions in order to minimize unwanted behavior by design.

It is possible to partially cancel some of the feedthrough effects using the technique illustrated in Fig. 4.1-11. Here a dummy MOS transistor MD (with source and drain both attached to the signal line and the gate attached to the inverse clock) is used to apply an opposing clock feedthrough due to M1. The area of MD can be designed to provide minimum clock feedthrough. Typically,  $W_D/L_D \approx 0.5 W_1/L_1$ . Unfortunately, this method never completely removes the feedthrough and in some cases may worsen it. Also, it is necessary to generate an inverted clock, which is applied to the dummy switch. Clock feedthrough can be reduced by using the largest capacitors possible, using minimum-geometry switches, and keeping the clock swings as small as possible. Typically, these solutions will create problems in other areas, requiring some compromises.

The dynamic range limitations associated with single-channel MOS switches can be avoided with the CMOS switch shown in Fig. 4.1-12. Using CMOS technology, a switch is usually constructed by connecting p-channel and n-channel enhancement transistors in parallel as illustrated. For this configuration, when  $\phi$  is low, both transistors are off, creating an effective open circuit. When  $\phi$  is high, both transistors are on, giving a low-impedance state. The bulk potentials of the p-channel and the n-channel devices are taken to the highest and lowest potentials, respectively. The primary advantage of the CMOS switch over the single-channel MOS switch is that the dynamic analog-signal range in the ON state is greatly increased.

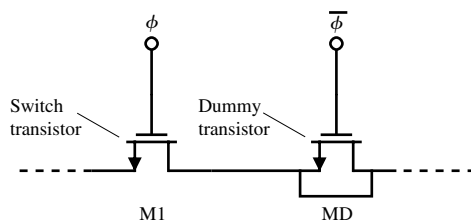


**Figure 4.1-11** The use of a dummy transistor to cancel clock feedthrough.

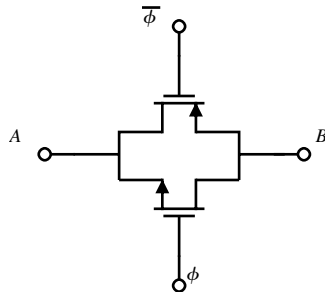
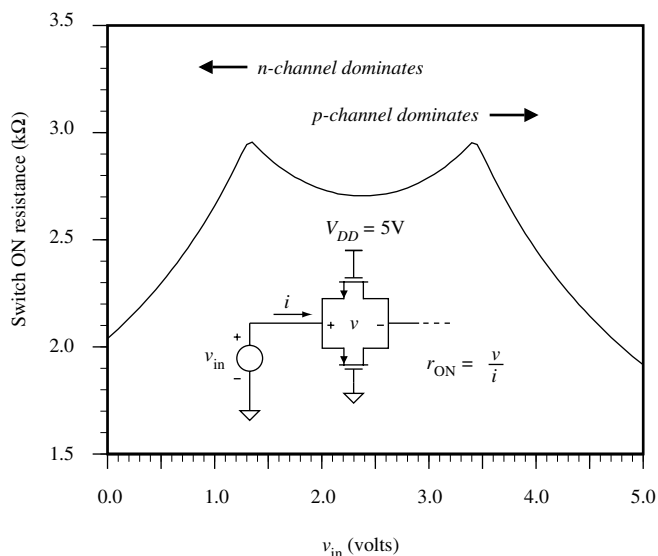


Figure 4.1-12 A CMOS switch.

The increased dynamic range of the analog signal is evident in Fig. 4.1-13, where the on resistance of a CMOS switch is plotted as a function of the input voltage. In this figure, the p-channel and n-channel devices are sized in such a way that they have equivalent resistance with identical terminal conditions. The double-peak behavior is due to the n-channel device dominating when  $v_{in}$  is low and the p-channel dominating when  $v_{in}$  is high (near  $V_{DD}$ ). At the midrange (near  $V_{DD}/2$ ), the parallel combination of the two devices results in a minimum. The dip at midrange is due to mobility degradation effects and is not evident when analyzed using the LEVEL 1 model.

In many switch applications, the voltage available to turn on the switch is not large enough. This is particularly true as the power supply voltages decrease or if the voltage to be switched is midway between the power supply rails. In these cases, it is necessary to generate a gate overdrive voltage. *Charge pumps* are used to accomplish the gate overdrive voltage. Figure 4.1-14 shows one method of generating a gate overdrive in a CMOS technology. The operation of this circuit is straightforward. To more easily understand the basic principle of the circuit, first ignore M4, M5, M7, and M8. Consider the other transistors (M1, M2, M3, and M6) as simple switches. During the phase when  $\phi_A$  is high and  $\phi_B$  is low,  $C_{PUMP}$  is charged with the full supply voltage. When  $\phi_A$  goes low and  $\phi_B$  goes high, M1 turns on along with M6, dumping a portion of the charge in  $C_{PUMP}$  into  $C_{HOLD}$ , causing  $v_{DBL}$  to go positive.

Figure 4.1-13  $r_{ON}$  of Fig 4.1-12 as a function of the voltage  $v_{in}$ .

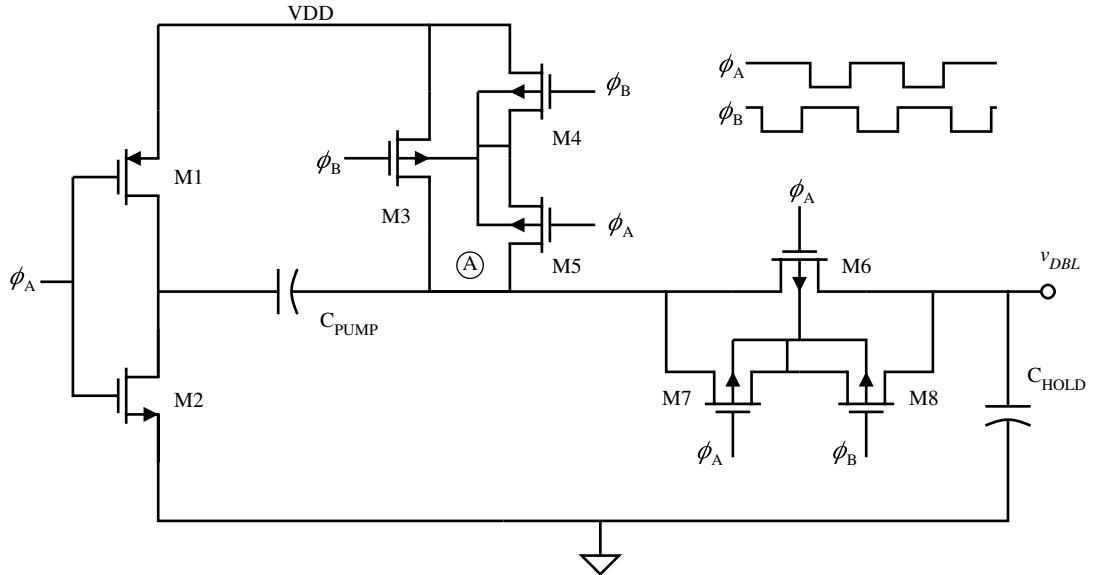


**Figure 4.1-14** Voltage doubler used to provide gate overdrive.

If both capacitors are equal and the charge is transferred ideally, then after the first pump,  $v_{DBL}$  will be equal to  $V_{DD} + 0.5(V_{DD} - V_{SS})$ . After many cycles,  $v_{DBL}$  will be at  $2V_{DD}$  with respect to  $V_{SS}$  (assuming there is no current drain on the  $v_{DBL}$  node). The switches M4, M5, M7, and M8 are for the purpose of keeping the bulk sources of M3 and M6 reverse biased during the operation of the charge pump.

In this section we have seen that MOS transistors make one of the best switch realizations available in integrated-circuit form. They require small area, dissipate very little power, and provide reasonable values of  $r_{ON}$  and  $r_{OFF}$  for most applications. The inclusion of a good realization of a switch into the designer's basic building blocks will produce some interesting and useful circuits and systems that will be studied in the following chapters.



## 4.2 MOS Diode/Active Resistor

When the gate and drain of an MOS transistor are tied together as illustrated in Figs. 4.2-1(a) and 4.2-1(b), the  $I$ - $V$  characteristics are qualitatively similar to a pn-junction diode, thus the name *MOS diode*. The MOS diode is used as a component of a current mirror (Section 4.4) and for level translation (voltage drop).

The  $I$ - $V$  characteristics of the MOS diode are illustrated in Fig. 4.2-1(c) and described by the large-signal equation for drain current in saturation (the connection of the gate to the drain guarantees operation in the saturation region) shown below.

$$I = I_D = \left( \frac{K'W}{2L} \right) [(V_{GS} - V_T)^2] = \frac{\beta}{2} (V_{GS} - V_T)^2 \quad (4.2-1)$$

or

$$V = V_{GS} = V_{DS} = V_T + \sqrt{2I_D/\beta} \quad (4.2-2)$$

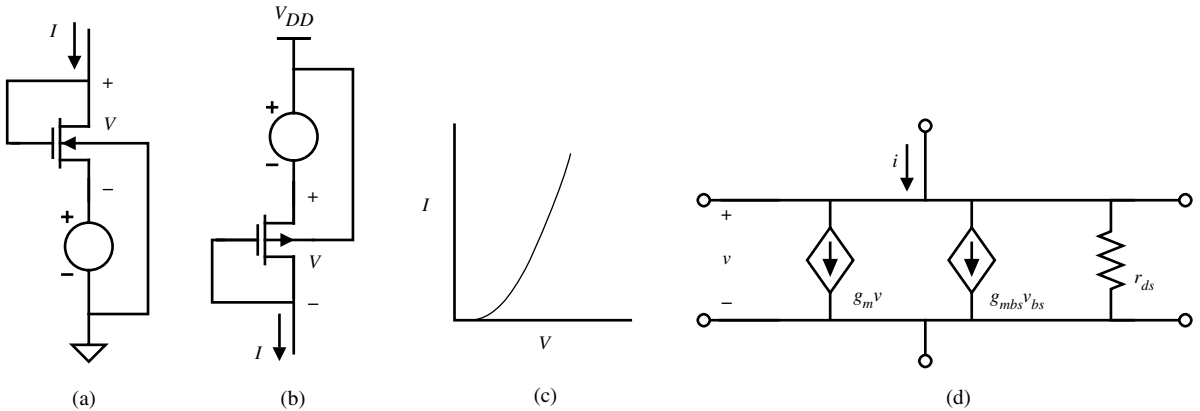


**Figure 4.2-1** Active resistor. (a) n-Channel. (b) p-Channel. (c)  $I$ - $V$  characteristics for n-channel case. (d) Small-signal model.

If  $V$  or  $I$  is given, then the remaining variable can be designed using either Eq. (4.2-1) or Eq. (4.2-2) and solving for the value of  $\beta$ .

Connecting the gate to the drain means that  $v_{DS}$  controls  $i_D$  and therefore the channel transconductance becomes a channel conductance. The small-signal model of an MOS diode (excluding capacitors) is shown in Fig. 4.2-1(d). It is easily seen that the small-signal resistance of an MOS diode is

$$r_{out} = \frac{1}{g_m + g_{mbs} + g_{ds}} \cong \frac{1}{g_m} \quad (4.2-3)$$

where  $g_m$  is greater than  $g_{mbs}$  or  $g_{ds}$ .

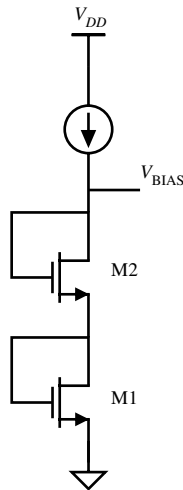
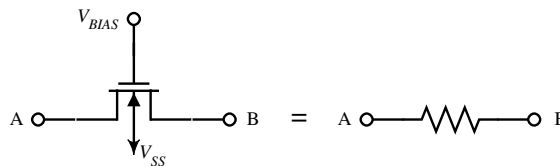
An illustration of the application of the MOS diode is shown in Fig. 4.2-2, where a bias voltage is generated with respect to ground (the value of such a circuit will become obvious later). Noting that  $V_{DS} = V_{GS}$  for both devices,

$$V_{DS} = \sqrt{2I/\beta} + V_T = V_{ON} + V_T \quad (4.2-4)$$

$$V_{BIAS} = V_{DS1} + V_{DS2} = 2V_{ON} + 2V_T \quad (4.2-5)$$

The MOS switch described in Section 4.1 and illustrated in Fig. 4.1-2 can be viewed as a resistor, albeit rather nonlinear, as illustrated in Fig. 4.1-4. The nonlinearity can be mitigated where the drain and source voltages vary over a small range so that the transistor ON resistance can be approximated as small-signal resistance. Figure 4.2-3 illustrates this point showing a configuration where the transistor's drain and source form the two ends of a "floating" resistor. For the small-signal premise to be valid,  $v_{DS}$  is assumed small. The  $I$ - $V$  characteristics of the floating resistor are given by Fig. 4.1-3. Consequently, the range of resistance values is large but nonlinear. When the transistor is operated in the nonsaturation region, the resistance can be calculated from Eq. (4.1-2), repeated below, where  $v_{DS}$  is assumed small.

$$r_{ds} = \frac{L}{K'W(V_{GS} - V_T)} \quad (4.2-6)$$

**Figure 4.2-2** Voltage division using active resistors.**Figure 4.2-3** Floating active resistor using a single MOS transistor.**Example  
4.2-1****Calculation of the Resistance of an Active Resistor**

The floating active resistor of Fig. 4.2-3 is to be used to design a  $1\text{ k}\Omega$  resistance. The dc value of  $V_{A,B} = 2\text{ V}$ . Use the device parameters in Table 3.1-2 and assume the active resistor is an n-channel transistor with the gate voltage at  $5\text{ V}$ . Assume that  $V_{DS} = 0.0$ . Calculate the required  $W/L$  to achieve  $1\text{ k}\Omega$  resistance. The bulk terminal is  $0.0\text{ V}$ .

**SOLUTION**

Before applying Eq. (4.2-6), it is necessary to calculate the new threshold voltage,  $V_T$ , due to  $V_{BS}$  not being zero ( $|V_{BS}| = 2\text{ V}$ ). From Eq. (3.1-2) the new  $V_T$  is found to be  $1.022\text{ V}$ . Equating Eq. (4.2-6) to  $1000\text{ }\Omega$  gives a  $W/L$  of  $4.597 \approx 4.6$ .

**4.3 Current Sinks and Sources**

A current sink and current source are two terminal components whose current at any instant of time is independent of the voltage across their terminals. The current of a current sink or source flows from the positive node, through the sink or source, to the negative node. A current sink typically has the negative node at  $V_{SS}$  and the current source has the positive node at  $V_{DD}$ . Figure 4.3-1(a) shows the MOS implementation of a current sink. The gate is taken

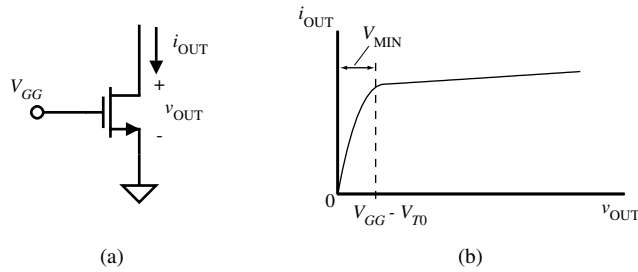


**Figure 4.3-1** (a) Current sink.  
(b) Current–voltage characteristics of (a).

to whatever voltage is necessary to create the desired value of current. The voltage divider of Fig. 4.2-2 can be used to provide this voltage. We note that in the nonsaturation region the MOS device is not a good current source. In fact, the voltage across the current sink must be larger than  $V_{\text{MIN}}$  in order for the current sink to perform properly. For Fig. 4.3-1(a) this means that

$$v_{\text{OUT}} \geq V_{\text{GG}} - V_{T0} \quad (4.3-1)$$

If the gate–source voltage is held constant, then the large-signal characteristics of the MOS transistor are given by the output characteristics of Fig. 3.1-3. An example is shown in Fig. 4.3-1(b). If the source and bulk are both connected to ground, then the small-signal output resistance is given by [see Eq. (3.3-9)]

$$r_{\text{out}} = \frac{1 + \lambda V_{DS}}{\lambda I_D} \cong \frac{1}{\lambda I_D} \quad (4.3-2)$$

If the source and bulk are not connected to the same potential, the characteristics will not change as long as  $V_{BS}$  is a constant.

Figure 4.3-2(a) shows an implementation of a current source using a p-channel transistor. Again, the gate is taken to a constant potential as is the source. With the definition of  $v_{\text{OUT}}$  and  $i_{\text{OUT}}$  of the source as shown in Fig. 4.3-2(a), the large-signal  $I$ – $V$  characteristic is shown in Fig. 4.3-2(b). The small-signal output resistance of the current source is given by Eq. (4.3-2). The source–drain voltage must be larger than  $V_{\text{MIN}}$  for this current source to work properly. This current source works only for values of  $v_{\text{OUT}}$  given by

$$v_{\text{OUT}} \leq V_{\text{GG}} + |V_{T0}| \quad (4.3-3)$$

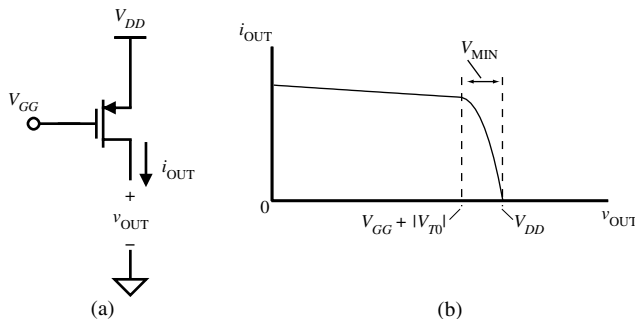


**Figure 4.3-2** (a) Current source.  
(b) Current–voltage characteristics of (a).

The advantage of the current sink and source of Figs. 4.3-1(a) and 4.3-2(a) is their simplicity. However, there are two areas in which their performance may need to be improved for certain applications. One improvement is to increase the small-signal output resistance—resulting in a more constant current over the range of  $v_{OUT}$  values. The second is to reduce the value of  $V_{MIN}$ , thus allowing a larger range of  $v_{OUT}$  over which the current sink/source works properly. We shall illustrate methods to improve both areas of performance. First, the small-signal output resistance can be increased using the principle illustrated in Fig. 4.3-3(a). This principle uses the common-gate configuration to multiply the source resistance  $r$  by the approximate voltage gain of the common-gate configuration with an infinite load resistance. The exact small-signal output resistance  $r_{out}$  can be calculated from the small-signal model of Fig. 4.3-3(b) as

$$r_{out} = \frac{v_{out}}{i_{out}} = r + r_{ds2} + [(g_{m2} + g_{mbs2})r_{ds2}]r \cong (g_{m2}r_{ds2})r \quad (4.3-4)$$

where  $g_{m2}r_{ds2} \gg 1$  and  $g_{m2} > g_{mbs2}$ .

The above principle is implemented in Fig. 4.3-4(a), where the output resistance ( $r_{ds1}$ ) of the current sink of Fig. 4.3-1(a) should be increased by the common-gate voltage gain of M2. To verify the principle, the small-signal output resistance of the cascode current sink of Fig. 4.3-4(a) will be calculated using the model of Fig. 4.3-4(b). Since  $v_{gs2} = -v_1$  and  $v_{gs1} = 0$ , summing the currents at the output node gives

$$i_{out} + g_{m2}v_1 + g_{mbs2}v_1 = g_{ds2}(v_{out} - v_1) \quad (4.3-5)$$

Since  $v_1 = i_{out}r_{ds1}$ , we can solve for  $r_{out}$  as

$$\begin{aligned} r_{out} = \frac{v_{out}}{i_{out}} &= r_{ds2}(1 + g_{m2}r_{ds1} + g_{mbs2}r_{ds1} + g_{ds2}r_{ds1}) \\ &= r_{ds1} + r_{ds2} + g_{m2}r_{ds1}r_{ds2}(1 + \eta_2) \end{aligned} \quad (4.3-6)$$

Typically,  $g_{m2}r_{ds2}$  is greater than unity so that Eq. (4.3-6) simplifies to

$$r_{out} \cong (g_{m2}r_{ds2})r_{ds1} \quad (4.3-7)$$

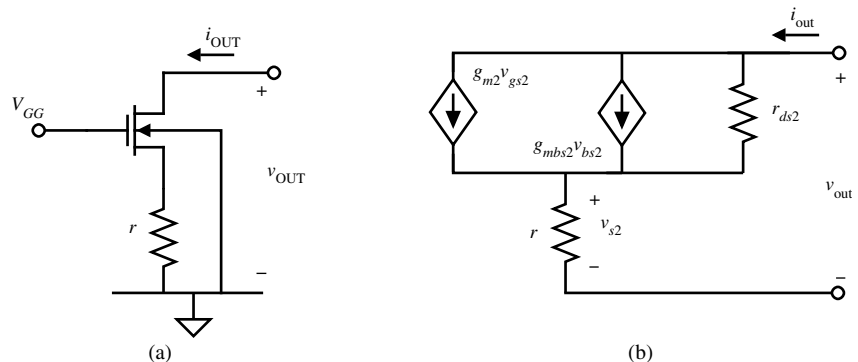


**Figure 4.3-3** (a) Technique for increasing the output resistance of a resistor  $r$ . (b) Small-signal model for the circuit in (a).

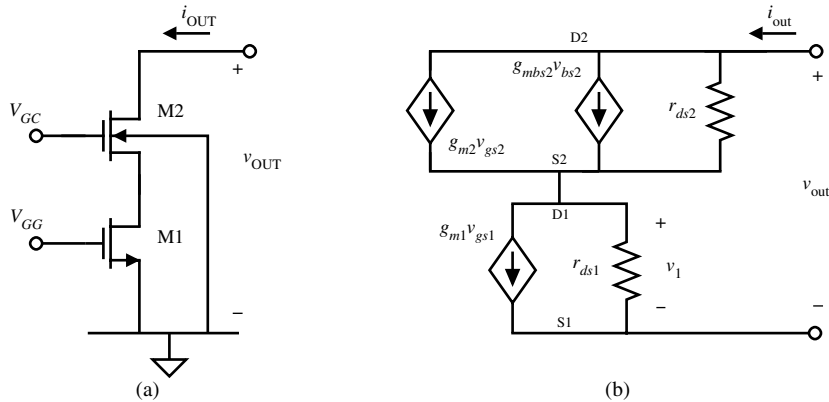


**Figure 4.3-4** (a) Circuit for increasing  $r_{out}$  of a current sink. (b) Small-signal model for the circuit in (a).

We see that the small-signal output resistance of the current sink of Fig. 4.3-4(a) is increased by the factor of  $g_{m2}r_{ds2}$ .

#### Example 4.3-1

#### Calculation of Output Resistance for a Current Sink

Use the model parameters of Table 3.1-2 to calculate (a) the small-signal output resistance for the simple current sink of Fig. 4.3-1(a) if  $I_{OUT} = 100 \mu\text{A}$ ; and (b) the small-signal output resistance if the simple current sink of (a) is inserted into the cascode current-sink configuration of Fig. 4.3-4(a). Assume that  $W_1/L_1 = W_2/L_2 = 1$ .

#### SOLUTION

- (a) Using  $\lambda = 0.04$  and  $I_{OUT} = 100 \mu\text{A}$  gives a small-signal output resistance of  $250 \text{ k}\Omega$ .  
 (b) The body-effect term,  $g_{mbs2}$ , can be ignored with little error in the result. Equation (3.3-6) gives  $g_{m1} = g_{m2} = 148 \mu\text{A/V}$ . Substituting these values into Eq. (4.3-7) gives the small-signal output resistance of the cascode current sink as  $9.25 \text{ M}\Omega$ .

The other performance limitation of the simple current sink/source was the fact that the constant output current could not be obtained for all values of  $v_{OUT}$ . This was illustrated in Figs. 4.3-1(b) and 4.3-2(b). While this problem may not be serious in the simple current sink/source, it becomes more severe in the cascode current-sink/source configuration that was used to increase the small-signal output resistance. It therefore becomes necessary to examine methods of reducing the value of  $V_{MIN}$  [3]. Obviously,  $V_{MIN}$  can be reduced by increasing the value of  $W/L$  and adjusting the gate-source voltage to get the same output current. However, another method that works well for the cascode current-sink/source configuration will be presented.

We must introduce an important principle used in biasing MOS devices before showing the method of reducing  $V_{MIN}$  of the cascode current sink/source. This principle can best be illustrated by considering two MOS devices, M1 and M2. Assume that the applied dc gate-source voltage  $V_{GS}$  can be divided into two parts, given as

$$V_{GS} = V_{ON} + V_T \quad (4.3-8)$$

where  $V_{ON}$  is that part of  $V_{GS}$  that is in excess of the threshold voltage,  $V_T$ . This definition allows us to express the minimum value of  $v_{DS}$  for which the device will remain in saturation as

$$v_{DS}(\text{sat}) = V_{GS} - V_T = V_{ON} \quad (4.3-9)$$

Thus,  $V_{ON}$  can be thought of as the minimum drain–source voltage for which the device remains saturated. In saturation, the drain current can be written as

$$i_D = \frac{K'W}{2L} (V_{ON})^2 \quad (4.3-10)$$

The principle to be illustrated is based on Eq. (4.3-10). If the currents of two MOS devices are equal (because they are in series), then the following relationship holds:

$$\frac{K'_1 W_1}{L_1} (V_{ON1})^2 = \frac{K'_2 W_2}{L_2} (V_{ON2})^2 \quad (4.3-11)$$

If both MOS transistors are of the same type, then Eq. (4.3-11) reduces to

$$\frac{W_1}{L_1} (V_{ON1})^2 = \frac{W_2}{L_2} (V_{ON2})^2 \quad (4.3-12)$$

or

$$\frac{\left(\frac{W_1}{L_1}\right)}{\left(\frac{W_2}{L_2}\right)} = \frac{(V_{ON2})^2}{(V_{ON1})^2} \quad (4.3-13)$$

The principle above can also be used to define a relationship between the current and  $W/L$  ratios. If the gate–source voltages of two similar MOS devices are equal (because they are physically connected), then  $V_{ON1}$  is equal to  $V_{ON2}$ . From Eq. (4.3-10) we can write

$$i_{D1} \left(\frac{W_2}{L_2}\right) = i_{D2} \left(\frac{W_1}{L_1}\right) \quad (4.3-14)$$

Equation (4.3-13) is useful even though the gate–source terminals of M1 and M2 may not be physically connected because voltages can be identical without being physically connected, as will be seen in later material. Equations (4.3-13) and (4.3-14) represent a very important principle that will be used not only in the material immediately following but throughout this text to determine biasing relationships.

Consider the cascode current sink of Fig. 4.3-5(a). Our objective is to use the above principle to reduce the value of  $V_{\text{MIN}}$ . If we ignore the bulk effects on M2 and M4 and assume that M1, M2, M3, and M4 are all matched with identical  $W/L$  ratios, then the gate–source voltage of each transistor can be expressed as  $V_T + V_{\text{ON}}$  as shown in Fig. 4.3-5(a). At the gate of M2 we see that the voltage with respect to the lower power supply is  $2V_T + 2V_{\text{ON}}$ . In order to maintain current-sink/source operation, it will be assumed that M1 and M2 must have at least a voltage of  $V_{\text{ON}}$  as given in Eq. (4.3-9). In order to find  $V_{\text{MIN}}$  of Fig. 4.3-5(a) we can rewrite Eq. (3.1-15) as

$$v_D \geq v_G - V_T \quad (4.3-15)$$

Since  $V_{G2} = 2V_T + 2V_{\text{ON}}$ , substituting this value into Eq. (4.3-15) gives

$$V_{D2}(\text{min}) = V_{\text{MIN}} = V_T + 2V_{\text{ON}} \quad (4.3-16)$$

The current–voltage characteristics of Fig. 4.3-5(a) are illustrated in Fig. 4.3-5(b), where the value of  $V_{\text{MIN}}$  of Eq. (4.3-16) is shown.

$V_{\text{MIN}}$  of Eq. (4.3-16) is dropped across both M1 and M2. The drop across M2 is  $V_{\text{ON}}$  while the drop across M1 is  $V_T + V_{\text{ON}}$ . From the results of Eq. (4.3-9), this implies that  $V_{\text{MIN}}$  of Fig. 4.3-5 could be reduced by  $V_T$  and still keep both M1 and M2 in saturation. Figure 4.3-6(a) shows how this can be accomplished [4]. The  $W/L$  ratio of M4 is made one-quarter of the identical  $W/L$  ratios of M1 through M3. This causes the gate–source voltage across M4 to be  $V_T + 2V_{\text{ON}}$  rather than  $V_T + V_{\text{ON}}$ . Consequently, the voltage at the gate of M2 is now  $V_T + 2V_{\text{ON}}$ . Substituting this value into Eq. (4.3-15) gives

$$V_{D2}(\text{min}) = V_{\text{MIN}} = 2V_{\text{ON}} \quad (4.3-17)$$

The resulting current–voltage relationship is shown in Fig. 4.3-6(b). It can be seen that a voltage of  $2V_{\text{ON}}$  is across both M1 and M2, giving the lowest value of  $V_{\text{MIN}}$  and still keeping both M1 and M2 in saturation. Using this approach and increasing the  $W/L$  ratios will result in minimum values of  $V_{\text{MIN}}$ .

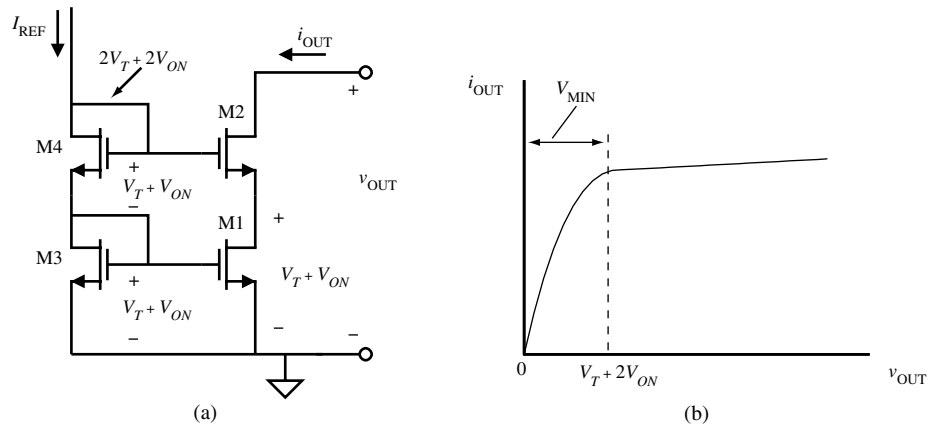


**Figure 4.3-5** (a) Standard cascode current sink. (b) Output characteristics of circuit in (a).

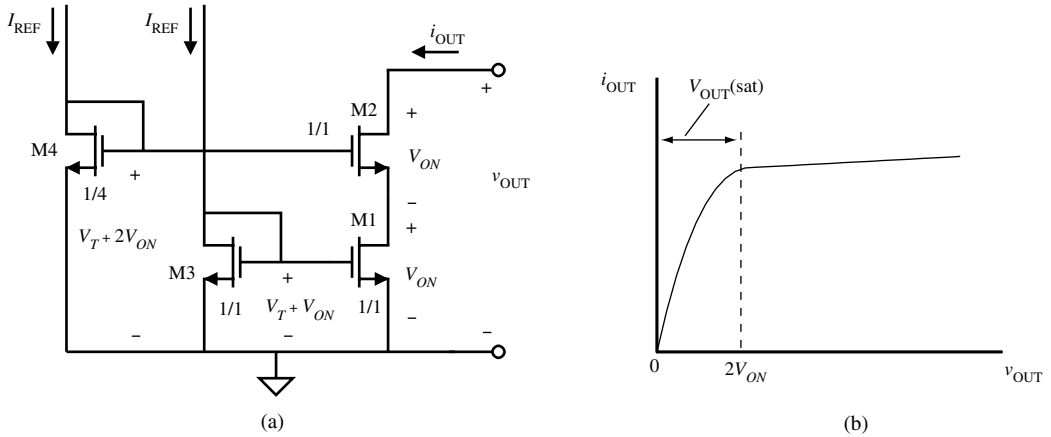


**Figure 4.3-6** (a) High-swing cascode. (b) Output characteristics of circuit in (a).

### Example 4.3-2

#### Designing the Cascode Current Sink for a Given $V_{MIN}$

Use the cascode current-sink configuration of Fig. 4.3-6(a) to design a current sink of  $100\ \mu\text{A}$  and a  $V_{MIN}$  of 1 V. Assume the device parameters of Table 3.1-2.

#### SOLUTION

With  $V_{MIN}$  of 1 V, choose  $V_{ON} = 0.5$  V. Using the saturation model, the  $W/L$  ratio of M1 through M3 can be found from

$$\frac{W}{L} = \frac{2 i_{OUT}}{K' V_{ON}^2} = \frac{2 \times 100 \times 10^{-6}}{110 \times 10^{-6} \times 0.25} = 7.27$$

The  $W/L$  ratio of M4 will be one-quarter this value or 1.82.

A problem exists with the circuit in Fig. 4.3-6. The  $V_{DS}$  of M1 and the  $V_{DS}$  of M3 are not equal. Therefore, the current  $i_{OUT}$  will not be an accurate replica of  $I_{REF}$  due to channel length modulation as well as drain-induced threshold shift. If precise mirroring of the current  $I_{REF}$  to  $I_{OUT}$  is desired, a slight modification of the circuit of Fig. 4.3-6 will minimize this problem. Figure 4.3-7 illustrates this fix. An additional transistor, M5, is added in series with M3 so as to force the drain voltages of M3 and M1 to be equal, thus eliminating any errors due to channel length modulation and drain-induced threshold shift.

The above technique will be useful in maximizing the voltage-signal swings of cascode configurations to be studied later. This section has presented implementations of the current sink/source and has shown how to boost the output resistance of an MOS device. A very important principle that will be used in biasing was based on relationships between the excess gate-source voltage  $V_{ON}$ , the drain current, and the  $W/L$  ratios of MOS devices. This principle was applied to reduce the voltage  $V_{MIN}$  of the cascode current source.

When power dissipation must be kept at a minimum, the circuit in Fig. 4.3-7 can be modified to eliminate one of the  $I_{REF}$  currents. Figure 4.3-8 illustrates a self-biased cascode current source that requires only one reference current [5]. While this circuit eliminates a bias

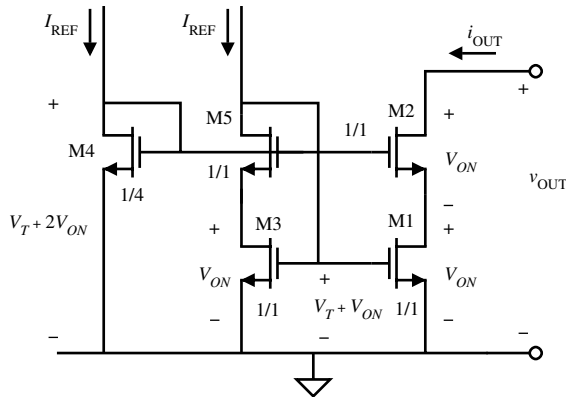


**Figure 4.3-7** Improved high-swing cascode.

voltage, it has two disadvantages. The first is that the resistor  $R$  can become quite large if the current is small and the second is that another node has been introduced into the circuit, creating a parasitic pole.

### Example 4.3-3

#### Designing the Self-Biased High-Swing Cascode Current Sink for a Given $V_{\text{MIN}}$

Use the cascode current-sink configuration of Fig. 4.3-8 to design a current sink of  $250\ \mu\text{A}$  and a  $V_{\text{MIN}}$  of  $0.5\ \text{V}$ . Assume the device parameters of Table 3.1-2.

#### SOLUTION

With  $V_{\text{MIN}}$  of  $0.5\ \text{V}$ , choose  $V_{\text{ON}} = 0.25\ \text{V}$ . Using the saturation model, the  $W/L$  ratio of M1 and M3 can be found from

$$\frac{W}{L} = \frac{2 i_{\text{OUT}}}{K' V_{\text{ON}}^2} = \frac{500 \times 10^{-6}}{110 \times 10^{-6} \times 0.0625} = 72.73$$

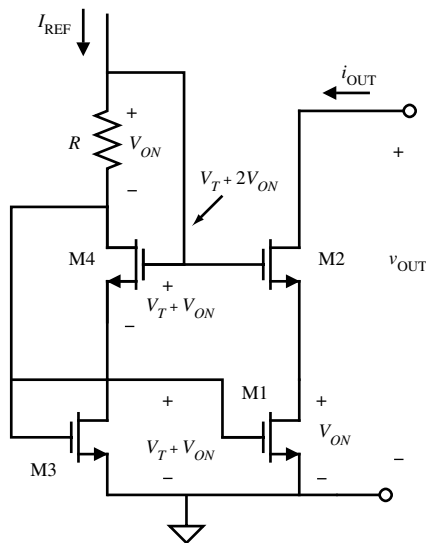


**Figure 4.3-8** Self-biased high-swing cascode current source.

The back-gate bias on M2 and M4 is  $-0.25$  V. Therefore, the threshold voltage for M2 and M4 is calculated to be

$$V_{TH} = 0.7 + 0.4 \left( \sqrt{0.25 + 0.7} - \sqrt{0.7} \right) = 0.755$$

Taking into account the increased value of the threshold voltage, the gate voltage of M4 and M2 is

$$V_{G4} = 0.755 + 0.25 + 0.25 = 1.255$$

The gate voltage of M1 and M3 is

$$V_{G1} = 0.70 + 0.25 = 0.95$$

Both terminals of the resistor are now defined so that the required resistance value is easily calculated to be

$$R = \frac{V_{G4} - V_{G1}}{250 \times 10^{-6}} = \frac{1.255 - 0.95}{250 \times 10^{-6}} = 1220 \, \Omega$$

It was seen that if a resistance is inserted between the source and ground of a current sink, the output resistance is increased. This was demonstrated by small-signal analysis in Eqs. (4.3-4) and (4.3-6). However, from a design viewpoint we should understand the principle that enables this result so that it can be used in future designs to either increase or decrease the resistance at a port. The principle at hand is negative feedback applied to a single port. It is characterized by a concept called Blackman's formula [6]. Blackman's formula is based on Fig. 4.3-9 where the port of interest is designated as  $X$ . In Fig. 4.3-9(a), a controlled or dependent source is shown. The controlling variable can be a voltage,  $v_c$ , or current,  $i_c$ . The source that is controlled can be a voltage source,  $kv_c$ , or a current source,  $ki_c$ , where the controlling parameter is  $k$ . In Fig. 4.3-9(b), the dependent source has been converted into an independent voltage source designated as  $kv_c'$  or an independent current source designated as  $ki_c'$ . When this is done, the feedback loop is opened and a quantity called *return ratio* can be calculated. The return ratio is simply the negative ratio of the controlling variable divided by the primed controlling variable. There are two types of return ratio, the return ratio with port  $X$  open-circuited ( $i_x = 0$ ) and the return ratio with port  $X$  short-circuited ( $v_x = 0$ ). It can be shown that the input resistance at port  $X$  of Fig. 4.3-9(a) is [7]

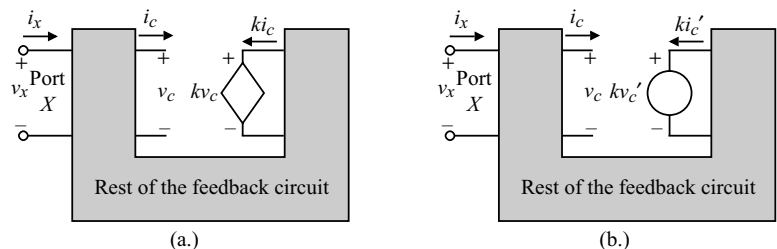


**Figure 4.3-9** (a) Application of Blackman's formula to find the resistance at port  $X$ . (b) Circuit used to find the return ratio of Blackman's formula.

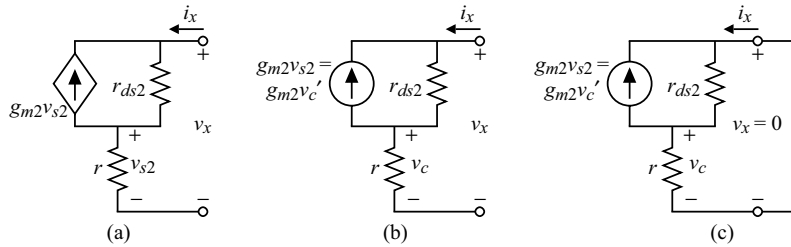


**Figure 4.3-10** (a) Figure 4.3-3(b) with  $v_{g2} = 0$  and  $v_{bs2} = 0$  V. (b) Conversion of Fig. 4.3-10(a) from Fig 4.3-9(a) to Fig. 4.3-9(b), which can also be used for the open-circuit port if  $i_x = 0$ . (c) Short-circuit form of Fig. 4.3-10(b).

$$R_x = R_x(k=0) \left[ \frac{1 + RR(\text{port shorted})}{1 + RR(\text{port opened})} \right] \quad (4.3-18)$$

where  $RR$  is the return ratio ( $-v_c/v_c'$  or  $-i_c/i_c'$ ). Let us apply this formula to Fig. 4.3-3(a) to derive Eq. (4.3-4) in the following example.

#### Example 4.3-4

#### Application of Blackman's Formula to Fig. 4.3-3(a)

Apply Blackman's formula given in Eq. (4.3-18) to the current sink of Fig. 4.3-3(a). To simplify the example, we will assume that  $v_{bs} = 0$  V.

#### SOLUTION

Figure 4.3-10 shows the process of converting Fig. 4.3-3(b) into the form of Fig. 4.3-9(b). The output resistance with  $g_{m2} = 0$  is

$$R_{out}(g_{m2} = 0) = r_{ds2} + r$$

The return ratio with the port shorted ( $v_x = 0$ ) is found as

$$\text{Return ratio } (v_x = 0) = -\frac{v_c}{v_c'} = g_{m2} \frac{r_{ds2} r}{r_{ds2} + r}$$

The return ratio with the port open ( $i_x = 0$ ) is found as

$$\text{Return ratio } (i_x = 0) = -\frac{v_c}{v_c'} = 0$$

because if  $i_x = 0$ , then all of the current  $g_{m2}v_c'$  flows through  $r_{ds2}$  and none through  $r$ . Substituting these values in Eq. (4.3-18) gives the output resistance as

$$\begin{aligned} R_{out} &= R_{out}(g_{m2} = 0) \left[ \frac{1 + RR(\text{port shorted})}{1 + RR(\text{port opened})} \right] = (r_{ds2} + r) \left( g_{m2} \frac{r_{ds2} r}{r_{ds2} + r} \right) \\ &= r_{ds2} + g_{m2} r_{ds2} r \end{aligned}$$

which is identical to Eq. (4.3-4) if  $g_{mbs2} = 0$ .

The key principle we learn from the influence of feedback through Blackman's formula is that if the resistance is to be increased, we want to find a feedback circuit where the return ratio goes to zero when the port is opened. On the other hand, if we want the resistance at a port to decrease, we want to find a feedback circuit where the return ratio goes to zero when the port is shorted. In terms of feedback concepts, the first case is called *series* feedback and the second case is called *shunt* feedback. Generally, if the controlling variable is voltage,  $v_c$ , the feedback is series, and if the controlling variable is current,  $i_c$ , the feedback is shunt.



## 4.4 Current Mirrors

Current mirrors are simply an extension of the current sink/source of the previous section. In fact, it is unlikely that one would ever build a current sink/source that was not biased as a current mirror. The current mirror uses the principle that if the gate-source potentials of two identical MOS transistors are equal, the channel currents should be equal. Figure 4.4-1 shows the implementation of a simple n-channel current mirror. The current  $i_I$  is assumed to be defined by a current source or some other means and  $i_O$  is the output or "mirrored" current. M1 is in saturation because  $v_{DS1} = v_{GS1}$ . Assuming that  $v_{DS2} \geq v_{GS2} - V_{T2}$  allows us to use the equations in the saturation region of the MOS transistor. In the most general case, the ratio of  $i_O$  to  $i_I$  is

$$\frac{i_O}{i_I} = \left( \frac{L_1 W_2}{W_1 L_2} \right) \left( \frac{V_{GS} - V_{T2}}{V_{GS} - V_{T1}} \right)^2 \left[ \frac{1 + \lambda v_{DS2}}{1 + \lambda v_{DS1}} \left( \frac{K'_2}{K'_1} \right) \right] \quad (4.4-1)$$

Normally, the components of a current mirror are processed on the same integrated circuit and thus all of the physical parameters such as  $V_T$  and  $K'$  are identical for both devices. As a result, Eq. (4.4-1) simplifies to

$$\frac{i_O}{i_I} = \left( \frac{L_1 W_2}{W_1 L_2} \right) \left( \frac{1 + \lambda v_{DS2}}{1 + \lambda v_{DS1}} \right) \quad (4.4-2)$$

If  $v_{DS2} = v_{DS1}$  (not always a good assumption), then the ratio of  $i_O/i_I$  becomes

$$\frac{i_O}{i_I} = \left( \frac{L_1 W_2}{W_1 L_2} \right) \quad (4.4-3)$$

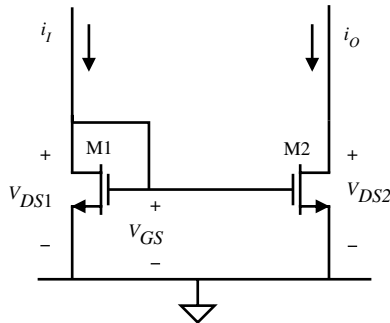


**Figure 4.4-1** n-Channel current mirror.

Consequently,  $i_O/i_I$  is a function of the aspect ratios that are under the control of the designer.

There are three effects that cause the current mirror to be different from the ideal situation of Eq. (4.4-3). These effects are (1) channel length modulation, (2) threshold offset between the two transistors, and (3) imperfect geometrical matching. Each of these effects will be analyzed separately.

Consider the channel length modulation effect. Assuming all other aspects of the transistor are ideal and the aspect ratios of the two transistors are both unity, then Eq. (4.4-2) simplifies to

$$\frac{i_O}{i_I} = \frac{1 + \lambda v_{DS2}}{1 + \lambda v_{DS1}} \quad (4.4-4)$$

with the assumption that  $\lambda$  is the same for both transistors. This equation shows that differences in drain–source voltages of the two transistors can cause a deviation from the ideal unity current gain or current mirroring. Figure 4.4-2 shows a plot of current ratio error versus  $v_{DS2} - v_{DS1}$  for different values of  $\lambda$  with both transistors in the saturation region. Two important facts should be recognized from this plot. The first is that significant ratio error can exist when the mirror transistors do not have the same drain–source voltage and second, for a given difference in drain–source voltages, the ratio of the mirror current to the reference current improves as  $\lambda$  becomes smaller (output resistance becomes larger). Thus, a good current mirror or current amplifier should have identical drain–source voltages and a high output resistance.

The second nonideal effect is that of offset between the threshold voltage of the two transistors. For clean silicon-gate CMOS processes, the threshold offset is typically less than 10 mV for transistors that are identical and in close proximity to one another.

Consider two transistors in a mirror configuration where both have the same drain–source voltage and all other aspects of the transistors are identical except  $V_T$ . In this case, Eq. (4.4-1) simplifies to

$$\frac{i_O}{i_I} = \left( \frac{v_{GS} - V_{T2}}{v_{GS} - V_{T1}} \right)^2 \quad (4.4-5)$$

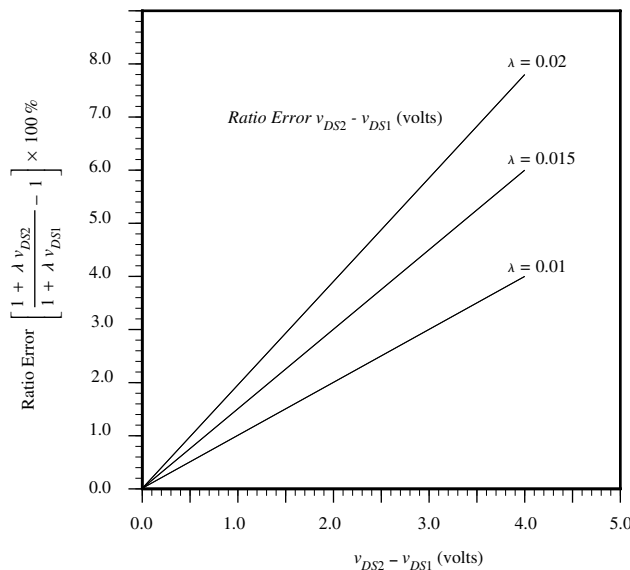


**Figure 4.4-2** Plot of ratio error (in %) versus drain voltage difference for the current mirror of Fig. 4.4-1. For this plot,  $v_{DS1} = 2.0$  V.

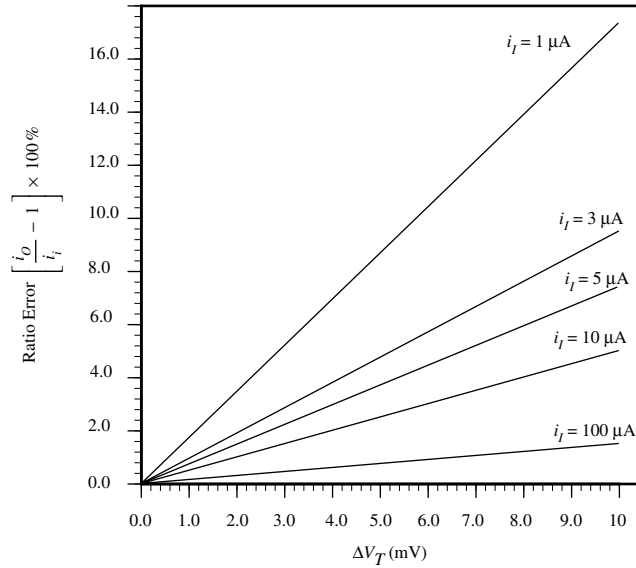


**Figure 4.4-3** Plot of ratio error (in %) versus offset voltage for the current mirror of Fig. 4.4-1. For this plot,  $v_{T1} = 0.7$  V and  $K'W/L = 110 \mu\text{A}/\text{V}^2$ .

Figure 4.4-3 shows a plot of the ratio error versus  $\Delta V_T$ , where  $\Delta V_T = V_{T1} - V_{T2}$ . It is obvious from this graph that better current-mirror performance is obtained at higher currents, because  $v_{GS}$  is higher for higher currents and thus  $\Delta V_T$  becomes a smaller percentage of  $v_{GS}$ .

It is also possible that the transconductance gain  $K'$  of the current mirror is also mismatched (due to oxide gradients). A quantitative analysis approach to variations in both  $K'$  and  $V_T$  is now given. Let us assume that the  $W/L$  ratios of the two mirror devices are exactly equal but that  $K'$  and  $V_T$  may be mismatched. Equation (4.4-5) can be rewritten as

$$\frac{i_O}{i_I} = \frac{K'_2 (v_{GS} - V_{T2})^2}{K'_1 (v_{GS} - V_{T1})^2} \quad (4.4-6)$$

where  $v_{GS1} = v_{GS2} = v_{GS}$ . Defining  $\Delta K' = K'_2 - K'_1$  and  $K' = 0.5(K'_2 + K'_1)$  and  $\Delta V_T = V_{T2} - V_{T1}$  and  $V_T = 0.5(V_{T2} + V_{T1})$  gives

$$K'_1 = K' - 0.5\Delta K' \quad (4.4-7)$$

$$K'_2 = K' + 0.5\Delta K' \quad (4.4-8)$$

$$V_{T1} = V_T - 0.5\Delta V_T \quad (4.4-9)$$

$$V_{T2} = V_T + 0.5\Delta V_T \quad (4.4-10)$$

Substituting Eqs. (4.4-7) through (4.4-10) into Eq. (4.4-6) gives

$$\frac{i_O}{i_I} = \frac{(K' + 0.5\Delta K')(v_{GS} - V_T - 0.5\Delta V_T)^2}{(K' - 0.5\Delta K')(v_{GS} - V_T + 0.5\Delta V_T)^2} \quad (4.4-11)$$

Factoring out  $K'$  and  $(v_{GS} - V_T)$  gives

$$\frac{i_O}{i_I} = \frac{\left(1 + \frac{\Delta K'}{2K}\right) \left(1 - \frac{\Delta V_T}{2(v_{GS} - V_T)}\right)^2}{\left(1 - \frac{\Delta K'}{2K}\right) \left(1 + \frac{\Delta V_T}{2(v_{GS} - V_T)}\right)^2} \quad (4.4-12)$$

Assuming that the quantities in Eq. (4.4-12) following the “1” are small, Eq. (4.4-12) can be approximated as

$$\frac{i_O}{i_I} \cong \left(1 + \frac{\Delta K'}{2K'}\right) \left(1 + \frac{\Delta K'}{2K'}\right) \left(1 - \frac{\Delta V_T}{2(v_{GS} - V_T)}\right)^2 \left(1 - \frac{\Delta V_T}{2(v_{GS} - V_T)}\right)^2 \quad (4.4-13)$$

Retaining only first-order products gives

$$\frac{i_O}{i_I} \cong 1 + \frac{\Delta K'}{K'} - \frac{2\Delta V_T}{v_{GS} - V_T} \quad (4.4-14)$$

If the percentage change of  $K'$  and  $V_T$  are known, Eq. (4.4-14) can be used on a worst-case basis to predict the error in the current-mirror gain. For example, assume that  $\Delta K'/K' = \pm 5\%$  and  $\Delta V_T/(v_{GS} - V_T) = \pm 10\%$ . Then the current-mirror gain would be given as  $i_O/i_I \cong 1 \pm 0.05 \pm (-0.20)$  or  $1 \pm (-0.15)$  amounting to a 15% error in gain assuming the tolerances of  $K'$  and  $V_T$  are correlated.

The third nonideal effect of current mirrors is the error in the aspect ratio of the two devices. We saw in Chapter 3 that there are differences in the drawn values of  $W$  and  $L$ . These are due to mask, photolithographic, etch, and outdiffusion variations. These variations can be different even for two transistors placed side by side. One way to avoid the effects of these variations is to make the dimensions of the transistors much larger than the typical variation one might see. For transistors of identical size with  $W$  and  $L$  greater than  $10 \mu\text{m}$ , the errors due to geometrical mismatch will generally be insignificant compared to offset-voltage and  $v_{DS}$ -induced errors.

In some applications, the current mirror is used to multiply current and function as a current amplifier. In this case, the aspect ratio of the multiplier transistor (M2) is much greater than the aspect ratio of the reference transistor (M1). To obtain the best performance, the geometrical aspects must be considered. An example will illustrate this concept.

#### Example 4.4-1

#### Aspect Ratio Errors in Current Amplifiers

Figure 4.4-4 shows the layout of a one-to-four current amplifier. Assume that the lengths are identical ( $L_1 = L_2$ ) and find the ratio error if  $W_1 = 5 \pm 0.1 \mu\text{m}$  and  $W_2 = 20 \pm 0.1 \mu\text{m}$ .

#### SOLUTION

The actual widths of the two transistors are

$$W_1 = 5 \pm 0.1 \mu\text{m}$$

and

$$W_2 = 20 \pm 0.1 \mu\text{m}$$

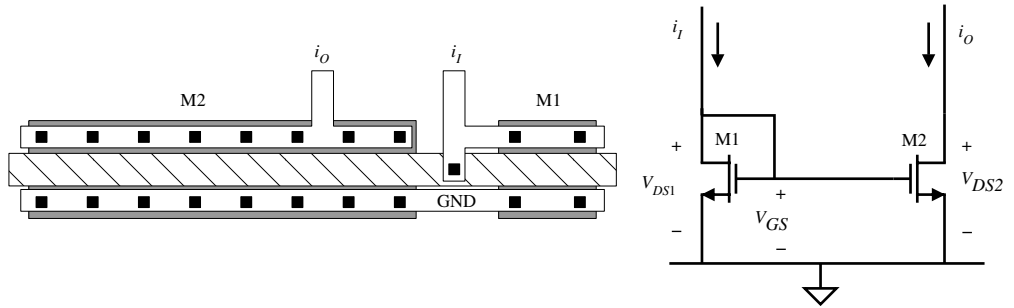


**Figure 4.4-4** Layout of current mirror without  $\Delta W$  correction.

We note that the tolerance is not multiplied by the nominal gain factor of 4. The ratio of  $W_2$  to  $W_1$  and consequently the gain of the current amplifier is

$$\begin{aligned} \frac{i_O}{i_I} &= \frac{W_2}{W_1} = \frac{20 + 0.1}{5 \pm 0.1} = 4 \left( \frac{1 \pm (0.1/20)}{1 \pm (0.1/5)} \right) \approx 4 \left( 1 \pm \frac{0.1}{20} \right) \left( 1 - \frac{\pm 0.1}{5} \right) \\ &\approx 4 \left( 1 \pm \frac{0.1}{20} - \frac{\pm 0.4}{20} \right) = 4 - (\pm 0.03) \end{aligned}$$

where we have assumed that the variations would both have the same sign. It is seen that this ratio error is 0.75% of the desired current ratio or gain.

The error noted above would be valid if every other aspect of the transistor were matched perfectly. A solution to this problem can be achieved by using proper layout techniques as illustrated in Appendix B. The correct one-to-four ratio should be implemented using four duplicates of the transistor M1. In this way, the tolerance on  $W_2$  is multiplied by the nominal current gain. Let us reconsider the above example using this approach.

#### Example 4.4-2

#### Reduction of the Aspect Ratio Error in Current Amplifiers

Use the layout technique illustrated in Fig. 4.4-5 and calculate the ratio error of a current amplifier having the specifications of the previous example.

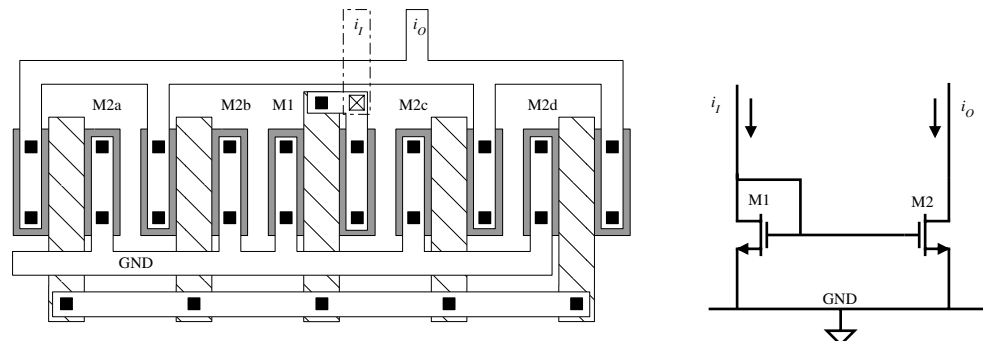


**Figure 4.4-5** Layout of current mirror with  $\Delta W$  correction as well as common-centroid layout techniques.

**SOLUTION**

The actual widths of M1 and M2 are

$$W_1 = 5 \pm 0.05 \mu\text{m}$$

and

$$W_2 = 4(5 \pm 0.05) \mu\text{m}$$

The ratio of  $W_2$  to  $W_1$  and consequently the current gain is seen to be

$$\frac{i_O}{i_I} = \frac{4(5 \pm 0.05)}{5 \pm 0.05} = 4$$

In the above examples we made the assumption that  $\Delta W$  should be the same for all transistors. Unfortunately, this is not true, but the  $\Delta W$  matching errors will be small compared to the other error contributions. If the widths of two transistors are equal but the lengths differ, the scaling approach discussed above for the width is also applicable to the length. Usually one does not try to scale the length because the tolerances are greater than the width tolerances due to diffusion (outdiffusion) under the polysilicon gate.

We have seen that the small-signal output resistance is a good measure of the perfection of the current mirror or amplifier. The output resistance of the simple n-channel mirror of Fig. 4.4-1 is given as

$$r_{\text{out}} = \frac{1}{g_{ds}} \cong \frac{1}{\lambda I_D} \quad (4.4-15)$$

Higher-performance current mirrors will attempt to increase the value of  $r_{\text{out}}$ . Equation (4.4-15) will be the point of comparison.

Up to this point we have discussed aspects of and improvements on the current mirror or current amplifier shown in Fig. 4.4-1, but there are ways of improving current-mirror performance using the same principles employed in Section 4.3. The current mirror shown in Fig. 4.4-6 applies the cascode technique, which reduces ratio errors due to differences in output and input voltages.

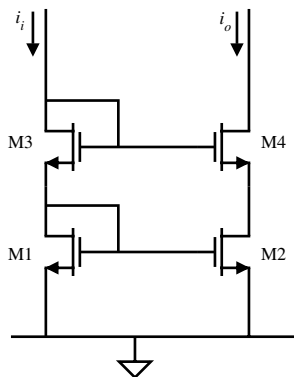


**Figure 4.4-6** Standard cascode current sink.

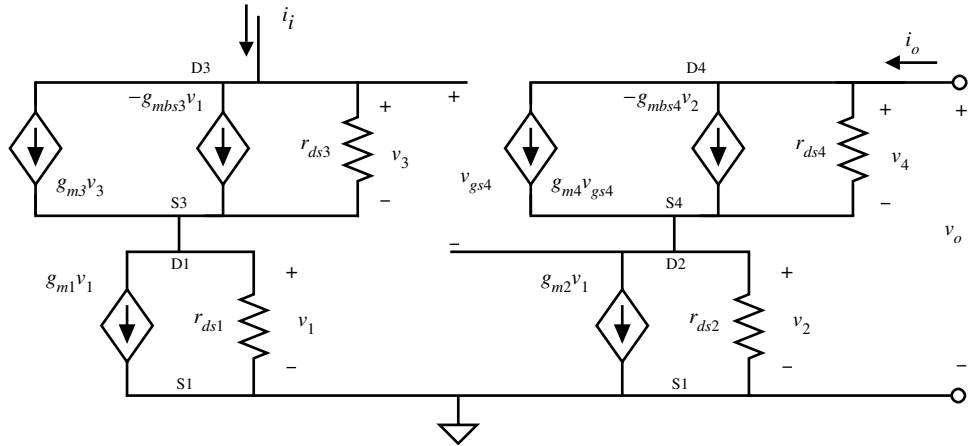


**Figure 4.4-7** Small-signal model for the circuit of Fig. 4.4-6.

Figure 4.4-7 shows an equivalent small-signal model of Fig. 4.4-6. To find the small-signal output resistance, set  $i_i = 0$ . This causes the small-signal voltages  $v_1$  and  $v_3$  to be zero. Therefore, Fig. 4.4-7 is exactly equivalent to the circuit of Example 4.3-1. Using the correct subscripts for Fig. 4.4-7, we can use the results of Eq. (4.3-6) to write

$$r_{\text{out}} = r_{ds2} + r_{ds4} + g_{m4}r_{ds2}r_{ds4}(1 + \eta_4) \quad (4.4-16)$$

We have already seen from Example 4.3-1 that the small-signal output resistance of this configuration is much larger than for the simple mirror of Eq. (4.4-15).

Another current mirror is shown in Fig. 4.4-8. This circuit is an n-channel implementation of the well-known Wilson current mirror [8]. The output resistance of the Wilson current mirror is increased through the use of negative current feedback. If  $i_o$  increases, then the current through M2 also increases. However, the mirroring action of M1 and M2 causes the current in M1 to increase. If  $i_i$  is constant and if we assume there is some resistance from the gate of M3 (drain of M1) to ground, then the gate voltage of M3 is decreased if the current  $i_o$  increases. The loop gain is essentially the product of  $g_{m1}$  and the small-signal resistance seen from the drain of M1 to ground.

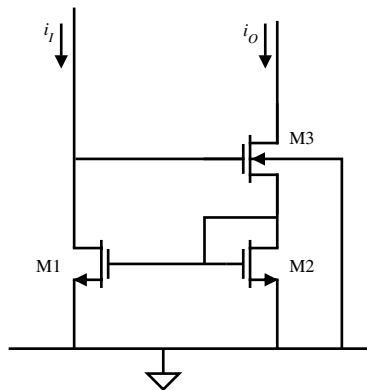


**Figure 4.4-8** Wilson current mirror.

It can be shown that the small-signal output resistance of the Wilson current source of Fig. 4.4-8 is

$$r_{\text{out}} = r_{ds3} + r_{ds2} \left( \frac{1 + r_{ds3}g_{m3}(1 + \eta_3) + g_{m1}r_{ds1}g_{m3}r_{ds3}}{1 + g_{m2}r_{ds2}} \right) \quad (4.4-17)$$

The output resistance of Fig. 4.4-8 is seen to be comparable with that of Fig. 4.4-6.

Unfortunately, the behavior described above for the current mirrors or amplifier requires a nonzero voltage at the input and output before it is achieved. Consider the cascode current mirror of Fig. 4.4-6 from a large-signal viewpoint. This voltage at the input, designated as  $V_I(\text{min})$ , can be shown to depend on the value of  $i_I$  as follows. Since  $v_{DG} = 0$  for both M1 and M3, these devices are always in saturation. Therefore, we may express  $V_I(\text{min})$  as

$$V_I(\text{min}) = \left( \frac{2i_I}{K'} \right)^{1/2} \left[ \left( \frac{L_1}{W_1} \right)^{1/2} + \left( \frac{L_3}{W_3} \right)^{1/2} \right] + (V_{T1} + V_{T3}) \quad (4.4-18)$$

It is seen that for a given  $i_I$  the only way to decrease  $V_I(\text{min})$  is to increase the  $W/L$  ratios of both M1 and M3. One must also remember that  $V_{T3}$  will be larger due to the back-gate bias on M3. The techniques used to reduce  $V_{\text{MIN}}$  at the output of the cascode current sink/source in Section 4.3 are not applicable to the Wilson current source.

For the cascoded current sink of Fig. 4.4-6 we are also interested in the voltage,  $V_{\text{MIN}}$ , where M4 makes the transition from the nonsaturated region to the saturated region. This voltage can be found from the relationship

$$v_{DS4} \geq (v_{GS4} - V_{T4}) \quad (4.4-19)$$

or

$$v_{D4} \geq v_{G4} - V_{T4} \quad (4.4-20)$$

which is when M4 is on the threshold between the two regions. Equation (4.4-20) can be used to obtain the value of  $V_{\text{MIN}}$  as

$$V_{\text{MIN}} = V_I - V_{T4} = \left( \frac{2I_I}{K'} \right)^{1/2} \left[ \left( \frac{L_1}{W_1} \right)^{1/2} + \left( \frac{L_3}{W_3} \right)^{1/2} \right] + (V_{T1} + V_{T3} - V_{T4}) \quad (4.4-21)$$

For voltages above  $V_{\text{MIN}}$ , the transistor M4 is in saturation and the output resistance should be that calculated in Eq. (4.4-16). Since the value of voltage across M2 is greater than necessary for saturation, the technique used to decrease  $V_{\text{MIN}}$  in Section 4.3 can be used to decrease  $V_{\text{MIN}}$ .

Similar relationships can be developed for the Wilson current mirror or amplifier. If M3 is saturated, then  $V_I(\text{min})$  is expressed as

$$V_I(\text{min}) = \left( \frac{2I_O}{K'} \right)^{1/2} \left[ \left( \frac{L_2}{W_2} \right)^{1/2} + \left( \frac{L_3}{W_3} \right)^{1/2} \right] + (V_{T2} + V_{T3}) \quad (4.4-22)$$

For M3 to be saturated,  $v_{\text{OUT}}$  must be greater than  $V_{\text{OUT}}(\text{sat})$  given as

$$V_{\text{OUT}}(\text{sat}) = V_I - V_{T3} = \left( \frac{2I_O}{K'} \right)^{1/2} \left[ \left( \frac{L_2}{W_2} \right)^{1/2} + \left( \frac{L_3}{W_3} \right)^{1/2} \right] + V_{T2} \quad (4.4-23)$$

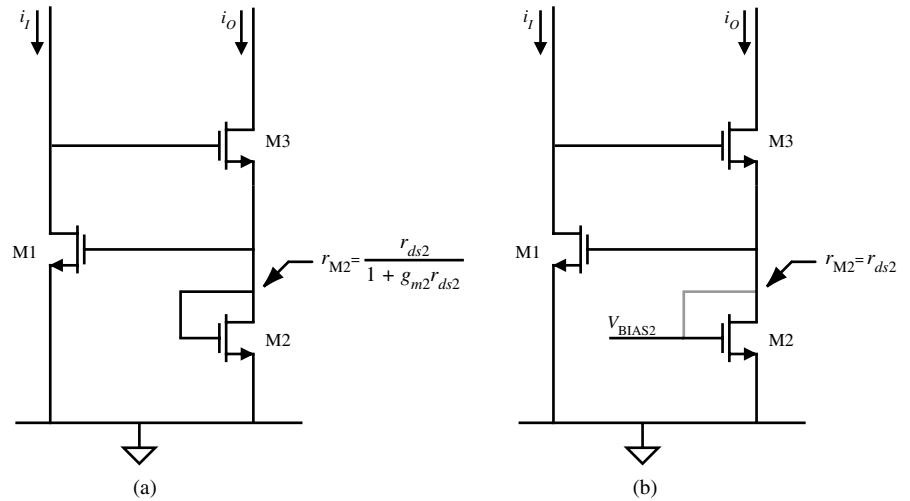


**Figure 4.4-9** (a) Wilson current mirror redrawn. (b) Wilson current mirror modified to increase  $r_{out}$  at M2.

It is seen that both of these circuits require at least  $2V_T$  across the input before they behave as described above. Larger  $W/L$  ratios will decrease  $V_I(\min)$  and  $V_{OUT}(\text{sat})$ .

An improvement on the Wilson current mirror can be developed by viewing from a different perspective. Consider the Wilson current mirror redrawn in Fig. 4.4-9. Note that the resistance looking into the diode connection of M2 is

$$r_{M2} = \frac{r_{ds2}}{1 + g_{m2}r_{ds2}} \quad (4.4-24)$$

If the gate of M2 is tied to a bias voltage so that  $r_{M2}$  becomes

$$r_{M2} = \frac{r_{ds2}}{1 + g_{m2}r_{ds2}} \Rightarrow r_{M2} = r_{ds2} \quad (4.4-25)$$

then the expression for  $r_{out}$  is given as

$$r_{out} = r_{ds3} + r_{ds2} \left( \frac{1 + r_{ds3}g_{m3}(1 + \eta_3) + g_{m1}r_{ds1}g_{m3}r_{ds3}}{1} \right) \quad (4.4-26)$$

$$r_{out} \cong r_{ds2}g_{m1}r_{ds1}g_{m3}r_{ds3} \quad (4.4-27)$$

This new current mirror illustrated fully in Fig. 4.4-10 is called a *regulated cascode* [9] and it achieves an output resistance on the order of  $g_m^2 r_{ds}^3$ .

Each of the current mirrors discussed above can be implemented using p-channel devices. The circuits perform in an identical manner and exhibit the same small-signal output resistance. The use of n-channel and p-channel current mirrors will be useful in dc biasing of CMOS circuits.

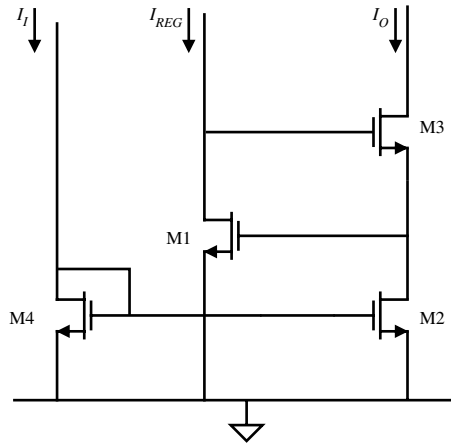


Figure 4.4-10 Regulated cascode current mirror.

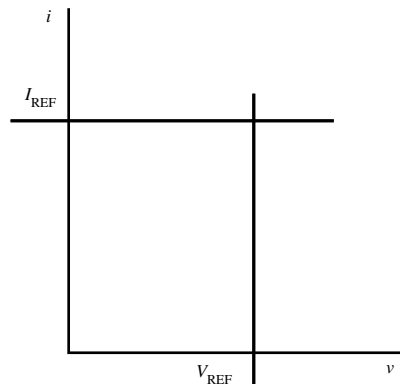


## 4.5 Current and Voltage References

An ideal current or voltage reference is independent of power supply and temperature. Many applications in analog circuits require such a building block, which provides a stable current or voltage. The large-signal current and voltage characteristics of an ideal current and voltage reference are shown in Fig. 4.5-1. These characteristics are identical to those of the ideal current and voltage source. The term *reference* is used when the current or voltage values have more precision and stability than ordinarily found in a source. A reference is typically dependent on the load connected to it. It will always be possible to use a buffer amplifier to isolate the reference from the load and maintain the high performance of the reference. In the discussion that follows, it will be assumed that a high-performance voltage reference can be used to implement a high-performance current reference and vice versa.

A very crude voltage reference can be made from a voltage divider between the power supplies. Passive or active components can be used as the divider elements. Figure 4.5-2 shows an example of each. Unfortunately, the value of  $V_{REF}$  is directly proportional to the power supply. Let us quantify this relationship by introducing the concept of *sensitivity*  $S$ . The sensitivity of  $V_{REF}$  of Fig. 4.5-2(a) to  $V_{DD}$  can be expressed as

$$S_{V_{DD}}^{V_{REF}} = \frac{(\partial V_{REF}/V_{REF})}{(\partial V_{DD}/V_{DD})} = \frac{V_{DD}}{V_{REF}} \left( \frac{\partial V_{REF}}{\partial V_{DD}} \right) \quad (4.5-1)$$

Figure 4.5-1  $I$ - $V$  characteristics of ideal current and voltage references.

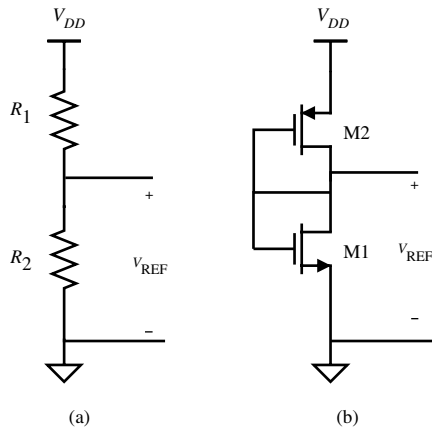


**Figure 4.5-2** Voltage references using voltage division. (a) Resistor implementation. (b) Active device implementation.

Equation (4.5-1) can be interpreted as follows: if the sensitivity is 1, then a 10% change in  $V_{DD}$  will result in a 10% change in  $V_{REF}$  (which is undesirable for a voltage reference). It may also be shown that the sensitivity of  $V_{REF}$  of Fig. 4.5-2(b) with respect to  $V_{DD}$  is unity (see Problem 4.5-1).

A simple way of obtaining a better voltage reference is to use an active device as shown in Fig. 4.5-3. In Fig. 4.5-3(a), the substrate BJT has been connected to the power supply through a resistance  $R$ . The voltage across the pn junction is given as

$$V_{REF} = V_{EB} = \frac{kT}{q} \ln \left( \frac{I}{I_s} \right) \quad (4.5-2)$$

where  $I_s$  is the junction-saturation current defined in Eq. (2.5-4). If  $V_{DD}$  is much greater than  $V_{EB}$ , then the current  $I$  is given as

$$I = \frac{V_{DD} - V_{EB}}{R} \cong \frac{V_{DD}}{R} \quad (4.5-3)$$

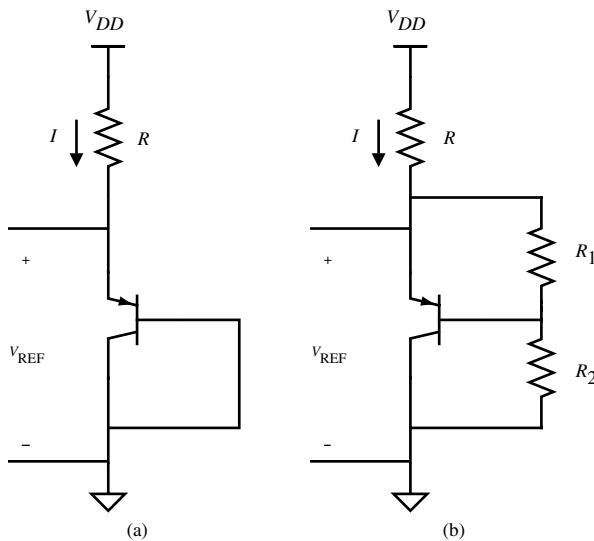


**Figure 4.5-3** (a) pn Junction voltage reference. (b) Increasing  $V_{REF}$  of (a).

Thus, the reference voltage of this circuit is given as

$$V_{\text{REF}} \cong \frac{kT}{q} \ln \left( \frac{V_{DD}}{RI_s} \right) \quad (4.5-4)$$

The sensitivity of  $V_{\text{REF}}$  of Fig. 4.5-3(a) to  $V_{DD}$  is shown to be

$$S_{V_{DD}}^{V_{\text{REF}}} = \frac{1}{\ln[V_{DD}/(RI_s)]} = \frac{1}{\ln(I/I_s)} \quad (4.5-5)$$

Interestingly enough, since  $I$  is normally greater than  $I_s$ , the sensitivity of  $V_{\text{REF}}$  of Fig. 4.5-3(a) is less than unity. For example, if  $I = 1$  mA and  $I_s = 10^{-15}$  A, then Eq. (4.5-5) becomes 0.0362. Thus, a 10% change in  $V_{DD}$  creates only a 0.362% change in  $V_{\text{REF}}$ . Figure 4.5-3(b) shows a method of increasing the value of  $V_{\text{REF}}$  in Fig. 4.5-3(a). The reference voltage of Fig. 4.5-3(b) can be written as

$$V_{\text{REF}} \cong V_{EB} \left( \frac{R_1 + R_2}{R_1} \right) \quad (4.5-6)$$

In order to find the value of  $V_{EB}$ , it is necessary to assume that the transistor  $\beta_F$  (common-emitter current gain) is large and/or the resistance  $R_1 + R_2$  is large. The larger  $V_{\text{REF}}$  becomes in Fig. 4.5-3(b), the more the current  $I$  becomes a function of  $V_{\text{REF}}$  and eventually an iterative solution is necessary.

The BJT of Fig. 4.5-3(a) may be replaced with an MOS enhancement device to achieve a voltage that is less dependent on  $V_{DD}$  than Fig. 4.5-2(a) as shown in Fig. 4.5-4(a).  $V_{\text{REF}}$  can be found from Eq. (4.2-2), which gives  $V_{GS}$  as

$$V_{GS} = V_T + \sqrt{\frac{2I}{\beta}} \quad (4.5-7)$$

Ignoring channel length modulation,  $V_{\text{REF}}$  is

$$V_{\text{REF}} = V_T - \frac{1}{\beta R} + \sqrt{\frac{2(V_{DD} - V_T)}{\beta R} + \frac{1}{\beta^2 R^2}} \quad (4.5-8)$$

If  $V_{DD} = 5$  V,  $W/L = 2$ , and  $R$  is 100 k $\Omega$ , the values of Table 3.1-2 give a reference voltage of 1.281 V. The sensitivity of Fig. 4.5-4(a) can be found as

$$S_{V_{DD}}^{V_{\text{REF}}} = \left( \frac{1}{\sqrt{1 + 2\beta(V_{DD} - V_T)R}} \right) \left( \frac{V_{DD}}{V_{\text{REF}}} \right) \quad (4.5-9)$$

Using the previous values gives a sensitivity of  $V_{\text{REF}}$  to  $V_{DD}$  of 0.283. This sensitivity is not as good as the BJT because the logarithmic function is much less sensitive to its argument than the square root. The value of  $V_{\text{REF}}$  of Fig. 4.5-4(a) can be increased using the technique

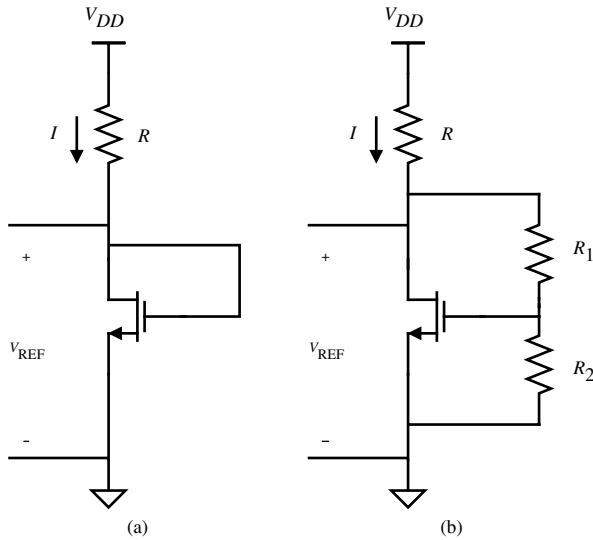


**Figure 4.5-4** (a) MOS equivalent of the pn junction voltage reference. (b) Increasing  $V_{REF}$  of (a).

employed for the BJT reference of Fig. 4.5-3(b), with the result shown in Fig. 4.5-4(b), where the reference voltage is given as

$$V_{REF} = V_{GS} \left( 1 + \frac{R_1}{R_2} \right) \quad (4.5-10)$$

In the types of voltage references illustrated in Figs. 4.5-3 and 4.5-4, the designer can use geometry to adjust the value of  $V_{REF}$ . In the BJT reference the geometric-dependent parameter is  $I_s$  and for the MOS reference it is  $W/L$ . The small-signal output resistance of these references is a measure of how dependent the reference will be on the load (see Problem 4.5-5).

We have noted in Figs. 4.5-3(a) and 4.5-4(a) that the sensitivity of the voltage across an active device is less than unity. If the voltage across the active device is used to create a current and this current is somehow used to provide the original current through the device, then a current or voltage will be obtained that is for all practical purposes independent of  $V_{DD}$ . This technique is called a  $V_T$  referenced source. This technique is also called a *bootstrap reference*. Figure 4.5-5(a) shows an example of this technique using all MOS devices. M3 and M4 cause the currents  $I_1$  and  $I_2$  to be equal.  $I_1$  flows through M1 creating a voltage  $V_{GS1}$ .  $I_2$  flows through  $R$  creating a voltage  $I_2 R$ . Because these two voltages are connected together, an equilibrium point is established. Figure 4.5-5(b) illustrates how the equilibrium point is achieved. On this curve,  $I_1$  and  $I_2$  are plotted as a function of  $V$ . The intersection of these curves defines the equilibrium point indicated by  $Q$ . The equation describing this equilibrium point is given as

$$I_2 R = V_{T1} + \left( \frac{2I_1 L_1}{K'_N W_1} \right)^{1/2} \quad (4.5-11)$$

This equation can be solved for  $I_1 = I_2 = I_Q$ , giving (ignoring  $\lambda$ )

$$I_Q = I_2 = \frac{V_{T1}}{R} + \frac{1}{\beta_1 R^2} + \frac{1}{R} \sqrt{\frac{2V_{T1}}{\beta_1 R} + \frac{1}{\beta_1 R^2}} \quad (4.5-12)$$

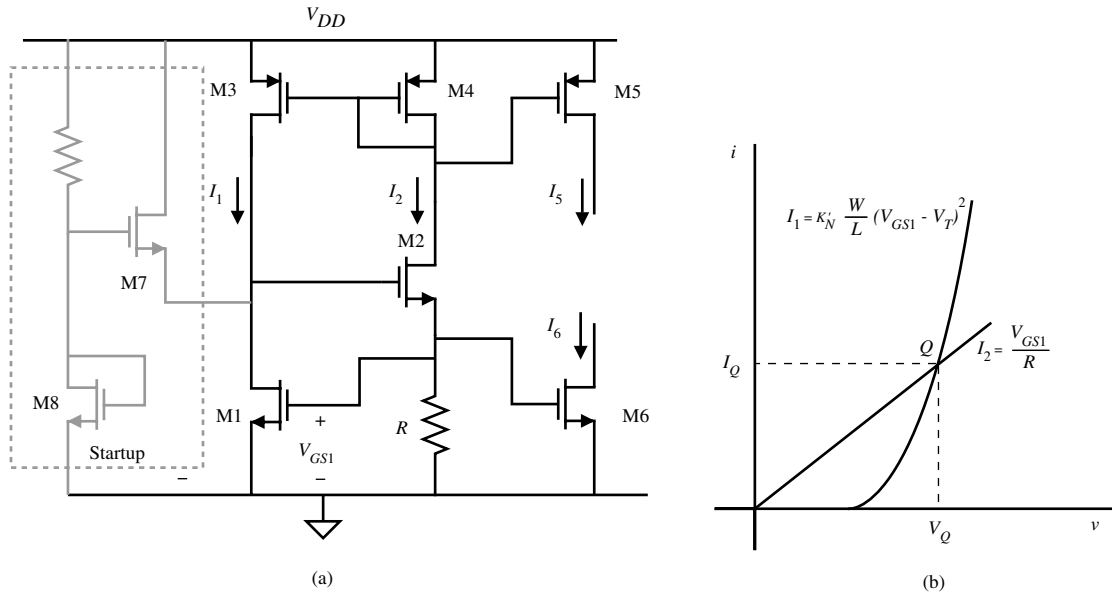


**Figure 4.5-5** (a) Threshold-referenced circuit. (b)  $I$ - $V$  characteristics of (a), illustrating how the bias point is established.

To first order, neither  $I_1$  nor  $I_2$  changes as a function of  $V_{DD}$ ; thus, the sensitivity of  $I_Q$  to  $V_{DD}$  is essentially zero. A voltage reference can be achieved by mirroring  $I_2$  ( $= I_Q$ ) through M5 or M6 and using a resistor. A closer examination shows that channel modulation effects on the M3–M4 current mirrors causes a weak dependence on  $V_{DD}$ .

Unfortunately, there are two possible equilibrium points on Fig. 4.5-5(b). One is at  $Q$  and the other is at the origin. In order to prevent the circuit from choosing the wrong equilibrium point, a startup circuit is necessary. The circuit within the dotted box in Fig. 4.5-5(a) functions as a startup circuit. If the circuit is at the undesired equilibrium point, then  $I_1$  and  $I_2$  are zero. However, M7 will provide a current in M1 that will cause the circuit to move to the equilibrium point at  $Q$ . As the circuit approaches the point  $Q$ , the source voltage of M7 increases, causing the current through M7 to decrease. At  $Q$  the current through M1 is essentially the current through M3.

An alternate version of Fig. 4.5-5(a) that uses  $V_{BE}$  to reference the voltage or current is shown in Fig. 4.5-6. It can be shown that the equilibrium point is defined by the relationship

$$I_2 R = V_{BE1} = V_T \ln \left( \frac{I_1}{I_s} \right) \quad (4.5-13)$$

This reference circuit also has two equilibrium points and a startup circuit similar to Fig. 4.5-5(a) is necessary. The reference circuits in Figs. 4.5-5(a) and 4.5-6 represent a very good method of implementing power-supply-independent references. Either circuit can be operated in the weak-threshold inversion in order to develop a low-power, low-supply-voltage reference.

Unfortunately, supply-independent references are not necessarily temperature independent because the pn junction and gate–source voltage drops are temperature dependent, as noted in Section 2.5. The concept of *fractional temperature coefficient* ( $TC_F$ ), defined in

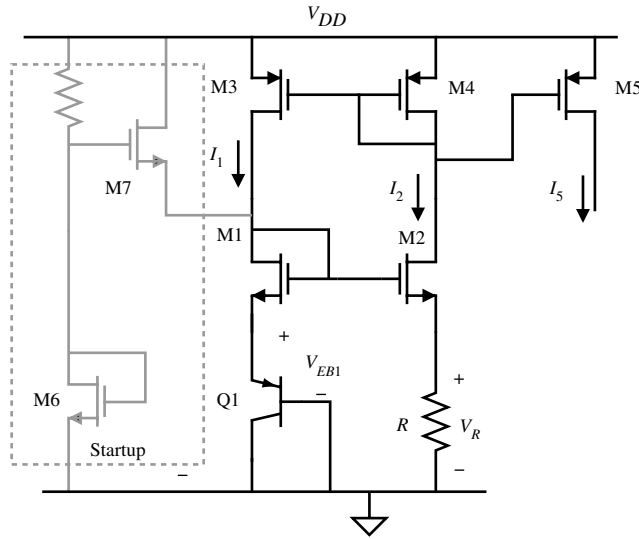


**Figure 4.5-6** Base-emitter voltage-referenced circuit.

Eq. (2.5-7), will be used to characterize the temperature dependence of voltage and current references. We see that  $TC_F$  is related to the sensitivity as defined in Eq. (4.5-1)

$$TC_F = \frac{1}{T} \left( S_T^X \right) \quad (4.5-14)$$

where  $X = V_{REF}$  or  $I_{REF}$ . Let us now consider the temperature characteristics of the simple pn junction of Fig. 4.5-3(a). If we assume that  $V_{DD}$  is much greater than  $V_{REF}$ , then Eq. (4.5-4) describes the reference voltage. Although  $V_{DD}$  is independent of temperature,  $R$  is not and must be considered. The fractional temperature coefficient of this voltage reference can be expressed using the results of Eq. (2.5-17) as

$$TC_F = \frac{1}{V_{REF}} \frac{dV_{REF}}{dT} \cong \frac{V_{REF} - V_{G0}}{V_{REF}T} - \frac{3k}{V_{REF}q} - \frac{kT}{V_{REF}q} \left( \frac{dR}{RdT} \right) \quad (4.5-15)$$

if  $v_D = V_{REF}$ . Assuming a  $V_{REF}$  of 0.6 V at room temperature, the  $TC_F$  of the simple pn voltage reference is approximately  $-2500$  ppm/°C.

Figure 4.5-4(a) is the MOS equivalent of the simple pn junction voltage reference. The temperature dependence of  $V_{REF}$  of this circuit can be written as

$$\frac{dV_{REF}}{dT} = \frac{-\alpha + \sqrt{\frac{V_{DD} - V_{REF}}{2\beta R}} \left( \frac{1.5}{T} - \frac{1}{R} \frac{dR}{dT} \right)}{1 + \frac{1}{\sqrt{2\beta R (V_{DD} - V_{REF})}}} \quad (4.5-16)$$

**Example**  
4.5-1**Calculation of Threshold Voltage Reference Circuit**

Calculate the temperature coefficient of the circuit in Fig. 4.5-4(a), where  $W/L = 2$ ,  $V_{DD} = 5$  V, and  $R = 100$  k $\Omega$ , using the parameters of Table 3.1-2. Resistor  $R$  is polysilicon and has a temperature coefficient of 1500 ppm/ $^{\circ}\text{C}$ .

**SOLUTION**

Using Eq. (4.5-8),

$$V_{\text{REF}} = V_T - \frac{1}{\beta_R} + \sqrt{\frac{2(V_{DD} - V_T)}{\beta R} + \frac{1}{\beta^2 R^2}}$$

$$\beta R = 220 \times 10^{-6} \times 10^5 = 22$$

$$V_{\text{REF}} = 0.7 - \frac{1}{22} + \sqrt{\frac{2(5 - 0.7)}{22} + \left(\frac{1}{22}\right)^2}$$

$$V_{\text{REF}} = 1.281$$

$$\frac{1}{R} \frac{dR}{dT} = 1500 \text{ ppm}/^{\circ}\text{C}$$

$$\frac{dV_{\text{REF}}}{dT} = \frac{-\alpha + \sqrt{\frac{V_{DD} - V_{\text{REF}}}{2\beta R}} \left( \frac{1.5}{T} - \frac{1}{R} \frac{dR}{dT} \right)}{1 + \frac{1}{\sqrt{2\beta R (V_{DD} - V_{\text{REF}})}}}$$

$$\frac{dV_{\text{REF}}}{dT} = \frac{-2.3 \times 10^{-3} + \sqrt{\frac{5 - 1.281}{2(22)}} \left( \frac{1.5}{300} - 1500 \times 10^{-6} \right)}{1 + \frac{1}{\sqrt{2(22)(5 - 1.281)}}}$$

$$\frac{dV_{\text{REF}}}{dT} = -1.189 \times 10^{-3} \text{ V}/^{\circ}\text{C}$$

The fractional temperature coefficient is given by

$$TC_F = \frac{1}{V_{\text{REF}}} \frac{dV_{\text{REF}}}{dT}$$

giving, for this example,

$$TC_F = -1.189 \times 10^{-3} \left( \frac{1}{1.281} \right) ^{\circ}\text{C}^{-1} = -928 \text{ ppm}/^{\circ}\text{C}$$

Unfortunately, the  $TC_F$  of this example is not realistic because the values of  $\alpha$  and the  $TC_F$  of the resistor do not have the implied accuracy.

The bootstrap reference circuit of Fig. 4.5-5(a) has its current  $I_2$  given by Eq. (4.5-12). If the product of  $R$  and  $\beta$  is large, the  $TC_F$  of the bootstrap reference circuit can be approximated as

$$TC_F = \frac{1}{V_T} \frac{dV_T}{dT} - \frac{1}{R} \frac{dR}{dT} = \frac{-\alpha}{V_T} - \frac{1}{R} \frac{dR}{dT} \quad (4.5-17)$$

### Example 4.5-2

#### Calculation of Bootstrap Reference Circuit

Calculate the temperature coefficient of the circuit in Fig. 4.5-5(a), where  $(W/L)_1 = 20$ ,  $V_{DD} = 5$  V, and  $R = 100$  k $\Omega$  using the parameters of Table 3.1-2. Resistor  $R$  is polysilicon and has a temperature coefficient of 1500 ppm/ $^{\circ}\text{C}$ .  $\alpha = 2.3 \times 10^{-3}$  V/ $^{\circ}\text{C}$ .

#### SOLUTION

Using Eq. (4.5-12),

$$I_Q = I_2 = \frac{V_{T1}}{R} + \frac{1}{\beta_1 R^2} + \frac{1}{R} \sqrt{\frac{2V_{T1}}{\beta_1 R} + \frac{1}{\beta_1^2 R^2}}$$

$$\beta_1 R = 220 \times 10^{-5} \times 10^5 = 220$$

$$\beta_1 R^2 = 220 \times 10^{-5} \times 10^{10} = 22 \times 10^6$$

$$I_Q = \frac{0.7}{10^5} - \frac{1}{22 \times 10^6} + \frac{1}{10^5} \sqrt{\frac{2 \times 0.7}{220} + \left(\frac{1}{220}\right)^2}$$

$$I_Q = 7.75 \mu\text{A}$$

$$\frac{1}{R} \frac{dR}{dT} = 1500 \text{ ppm}/^{\circ}\text{C}$$

$$TC_F = \frac{-2.3 \times 10^{-3}}{0.7} - 1500 \times 10^{-6} ^{\circ}\text{C}^{-1} = -4.79 \times 10^{-3} ^{\circ}\text{C}^{-1} = -4790 \text{ ppm}/^{\circ}\text{C}$$

The temperature behavior of the base-emitter-referenced circuit of Fig. 4.5-6 is similar to that of the threshold-referenced circuit of Fig. 4.5-5(a). Equation (4.5-13) showed that  $I_2$  is equal to  $V_{BE1}$  divided by  $R$ . Thus, Eq. (4.5-17) above expresses the  $TC_F$  of this reference if  $V_T$  is replaced by  $V_{BE}$  as follows:

$$TC_F = \frac{1}{V_{BE}} \frac{dV_{BE}}{dT} - \frac{1}{R} \frac{dR}{dT} \quad (4.5-18)$$

Assuming  $V_{BE}$  of 0.6 V gives a  $TC_F$  of  $-2333$  ppm/ $^{\circ}\text{C}$ .



**Example**  
4.5-2**Design of the Simple Reference of Fig. 4.5-7**

Use Fig. 4.5-7 to design (a) a 20  $\mu\text{A}$  current sink and (b) a 1 V bias voltage. Assume the MOSFET parameters of Table 3.1-2 and the  $W/L$  of  $M1 = 5$ .

**SOLUTION**

From Eq. (4.5-21) we can design the resistor,  $R$ , as

$$R = \frac{1}{\sqrt{2}\beta_1 I_1} = \frac{1}{\sqrt{2} \times 110(\mu\text{A}/\text{V}^2) \times 5 \times 20 \mu\text{A}} = 6.74 \text{ k}\Omega$$

To design the 1 V bias voltage we use Eq. (4.5-22) to redesign the resistor  $R$  as

$$R = \frac{1}{\beta_1(V_{\text{Bias}} - V_{TN})} = \frac{1}{110(\mu\text{A}/\text{V}^2) \times 5 \times (1 \text{ V} - 0.7 \text{ V})} = 6.06 \text{ k}\Omega$$

The voltage and current references presented in this section have the objective of providing a stable value of current with respect to changes in power supply and temperature. It was seen that while power-supply independence could thus be obtained, satisfactory temperature performance could not. To illustrate this conclusion, consider the accuracy requirement for a voltage reference for 8-bit accuracy. Maintaining this accuracy over a 100  $^{\circ}\text{C}$  change requires the  $TC_F$  of the reference to be  $1/(256 \times 100^{\circ}\text{C})$  or 39 ppm/ $^{\circ}\text{C}$ .

**4.6 Temperature-Independent References**

In the last section, we showed voltage and current references that were reasonably independent of power supply but were dependent on temperature. In this section, we will show how to design voltage and current references that are reasonably independent of *both* power supply and temperature. These references are typically called *bandgap references* [10–14] although they have very little to do with the bandgap voltage. These references typically have a temperature coefficient (TC) of 10–50 ppm/ $^{\circ}\text{C}$  over a temperature range of 0–70  $^{\circ}\text{C}$  without correction. With correction [15–20], they are capable of achieving 1 ppm/ $^{\circ}\text{C}$  or less over 0–70  $^{\circ}\text{C}$ .

**Voltage References with Moderate Temperature Stability**

The principle of temperature-independent references is very simple. It begins with identifying a voltage that increases with temperature and a voltage that decreases with temperature. Figure 4.6-1(a) shows two voltages, one that increases proportionately with temperature and one that decreases proportionately with temperature. The increasing voltage is called *proportional to absolute temperature* or *PTAT* and the decreasing voltage is called *complementary to absolute temperature* or *CTAT*. Next, the voltage with the smallest magnitude of slope (in this case  $V_{PTAT}$ ) is multiplied by a temperature independent constant,  $K$ , so that the magnitudes of the slopes are equal. Finally, if  $K \cdot V_{PTAT}(T)$  is added to  $V_{CTAT}(T)$ , as illustrated in Fig. 4.6-1(b), the resulting voltage should be independent of temperature.

To implement the principle above, we must develop  $V_{PTAT}$  and  $V_{CTAT}$  voltages. We will begin by showing how to generate a PTAT voltage. Figure 4.6-2 shows two current sources

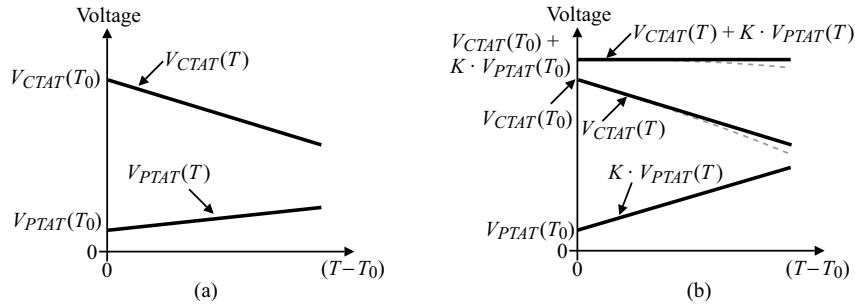


**Figure 4.6-1** (a) Illustration of voltages that increase and decrease with increasing temperature. (b) Multiplication of  $V_{PTAT}(T)$  by  $K$  to make the slopes of  $V_{PTAT}(T)$  and  $V_{CTAT}(T)$  equal but opposite.

each in series with a diode. The diode could also be a bipolar junction transistor with the collector connected to the base. The voltage,  $\Delta V_D$ , will be PTAT and can be expressed as

$$\begin{aligned} V_{PTAT} = \Delta V_D &= V_{D1} - V_{D2} = V_t \ln\left(\frac{I_1}{I_{s1}}\right) - V_t \ln\left(\frac{I_2}{I_{s2}}\right) \\ &= V_t \ln\left(\frac{I_1 I_{s2}}{I_2 I_{s1}}\right) = V_t \ln\left(\frac{I_{s2}}{I_{s1}}\right) = V_t \ln\left(\frac{A_2}{A_1}\right) = \frac{kT}{q} \ln\left(\frac{A_2}{A_1}\right) \end{aligned} \quad (4.6-1)$$

if  $I_1 = I_2$ . Equation (4.6-1) shows that  $\Delta V_D$  is proportional to absolute temperature with a proportionality constant or slope equal to Boltzmann's constant divided by the charge of an electron ( $k/q = 1.381 \times 10^{-23} \text{ J}^\circ\text{K}/1.6 \times 10^{-19} \text{ Coul} = 0.086 \text{ mV}^\circ\text{C}$ ).

In developing voltage-independent references, it is useful to show how to generate PTAT and CTAT currents. A pseudo-PTAT current can be created by placing the PTAT voltage  $\Delta V_D$  across a resistor. Unfortunately, all resistors have some sort of temperature dependence so that the current is not truly PTAT, thus the name *pseudo*-PTAT. This will not be a problem because in most cases, the pseudo-PTAT current runs through a second resistor, creating a new PTAT voltage. If the two resistors have the same temperature dependence, then the new PTAT voltage is truly PTAT. Pseudo-PTAT currents will be designated as PTAT'.

Figures 4.6-3(a) and 4.6-3(b) show two ways of creating a pseudo-PTAT currents. Figure 4.6-3(a) is similar to Fig. 4.5-6 except for the diode  $D_2$ , which has been added in series with the resistor  $R$ . The transistors M1 and M2 ensure that the voltage across diode  $D_1$  is the same as the voltage across the series combination of the resistor  $R$  and the diode  $D_2$ . The pseudo-PTAT current is

$$I_{PTAT}' = \frac{V_{D1} + V_{GS1} - V_{GS2} - V_{D2}}{R} = \frac{kT}{Rq} \ln\left(\frac{A_2}{A_1}\right) \quad (4.6-2)$$

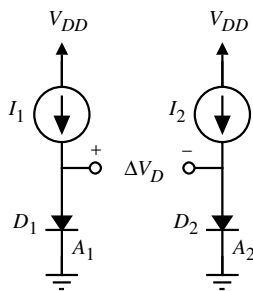


**Figure 4.6-2** Implementation of a PTAT voltage.

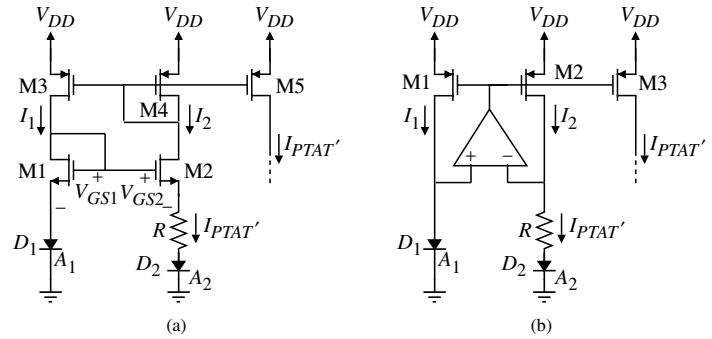


**Figure 4.6-3** Creation of pseudo-PTAT currents using (a) MOSFETs only and (b) MOSFETs and an op amp.

if  $V_{GS1} = V_{GS2}$ . Figure 4.6-3(b) shows a method of generating the pseudo-PTAT current using an op amp to make the voltage across diode  $D_1$  the same as the voltage across the series combination of the resistor  $R$  and the diode  $D_2$ . Note that the op amp has both positive and negative feedback, which is also the case for Fig. 4.6-3(a). For these circuits to work properly, the negative feedback loop gain should be at least twice as large as the positive feedback loop gain.

A true PTAT current,  $I_{PTAT}$ , can be generated from a pseudo-PTAT current using the zero temperature coefficient property described in Section 3.2. Figure 4.6-4 shows the application of this property to obtain a true PTAT current. The pseudo-PTAT current can be provided by either Fig. 4.6-3(a) or 4.6-3(b). If the resistor in Fig. 4.6-3(a) or 4.6-3(b) is designated  $R_1$ , then the gate–source voltage of the MOSFET can be written as

$$V_{GS6}(ZTC) = I_{PTAT}' R_2 = \frac{R_2}{R_1} \frac{kT}{q} \ln\left(\frac{A_2}{A_1}\right) \quad (4.6-3)$$

If the voltage in Eq. (4.6-3) is designed to be the ZTC point of the MOSFET, then because the gate–source voltage is PTAT, the drain current should also be PTAT. This of course assumes that the temperature change is not large enough to make the ZTC point shift. The influence of the resistors,  $R_1$  and  $R_2$ , can be found by differentiating their ratio with respect to temperature. The result is

$$\frac{d(R_2/R_1)}{dT} = \frac{1}{R_1} \frac{dR_2}{dT} - \frac{R_2}{R_1^2} \frac{dR_1}{dT} = \frac{R_2}{R_1} \left[ \frac{1}{R_2} \frac{dR_2}{dT} - \frac{1}{R_1} \frac{dR_1}{dT} \right] \quad (4.6-4)$$

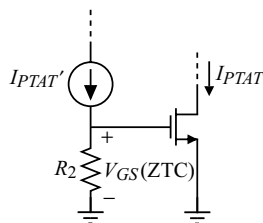


**Figure 4.6-4** Generation of a true PTAT current from a pseudo-PTAT current.

If the temperature coefficients of  $R_1$  and  $R_2$  are the same, then the resistor ratio does not influence the temperature characteristics. At this point, we have shown how to develop voltages and currents that are PTAT.

The next step is to show how to create a voltage that is complementary to absolute temperature (CTAT). This becomes more challenging because a voltage with a true complementary dependence on temperature does not exist. The best candidate for finding a voltage that decreases linearly with absolute temperature is the  $pn$  junction. The current density in a diode is given as

$$J_D = \left[ \frac{qD_n n_{po}}{L_n} + \frac{qD_p p_{no}}{L_p} \right] \exp\left(\frac{v_D}{V_t}\right) \quad (4.6-5)$$

where

$J_D$  = diode current density (A/m<sup>2</sup>)

$n_{po}$  = equilibrium concentration of electrons in the p-side

$p_{no}$  = equilibrium concentration of holes in the n-side

$D_n$  = average diffusion constant for electrons

$D_p$  = average diffusion constant for holes

$L_n$  = diffusion length of electrons in the p-side

$L_p$  = diffusion length of holes in the n-side

The equilibrium concentration of the electrons and holes can be expressed as

$$n_{po} = \frac{n_i^2}{N_A} \quad (4.6-6)$$

and

$$p_{no} = \frac{n_i^2}{N_D} \quad (4.6-7)$$

where

$$n_i^2 = DT^3 \exp(-V_{G0}/V_t) \quad (4.6-8)$$

The term  $D$  is a temperature-independent constant and  $V_{G0}$  is the *bandgap voltage* at room temperature (1.205 volts). Combining Eqs. (4.6-5) through (4.6-8) results in the following equation for diode current density:

$$J_D = \left[ \frac{qD_n}{L_n N_A} + \frac{qD_p}{L_p N_D} \right] n_i^2 \exp\left(\frac{V_D}{V_t}\right) \quad (4.6-9)$$

or

$$J_D = AT^\gamma \exp\left(\frac{V_D - V_{G0}}{V_t}\right) \quad (4.6-10)$$

In Eq. (4.6-10), the temperature-independent constants of Eq. (4.6-9) are combined into a single constant  $A$ . The term  $T^3$  in Eq. (4.6-8) has been replaced with  $T^\gamma$  due to the influence of the temperature dependence of  $D_n$  and  $D_p$ .

Solving for  $V_D$  from Eq. (4.6-10) gives

$$V_D = \frac{kT}{q} \ln\left(\frac{J_D}{AT^\gamma}\right) + V_{G0} \quad (4.6-11)$$

Now consider  $J_D$  at a temperature  $T_0$ .

$$J_{D0} = AT_0^\gamma \exp\left[\frac{q}{kT_0}(V_{D0} - V_{G0})\right] \quad (4.6-12)$$

The ratio of  $J_D$  to  $J_{D0}$  is

$$\frac{J_D}{J_{D0}} = \left(\frac{T}{T_0}\right)^\gamma \exp\left[\frac{q}{k}\left(\frac{V_D - V_{G0}}{T} - \frac{V_{D0} - V_{G0}}{T_0}\right)\right] \quad (4.6-13)$$

Equation (4.6-13) can be rearranged to get  $V_D$  which is a CTAT voltage as

$$V_{CTAT} = V_D = V_{G0}\left(1 - \frac{T}{T_0}\right) + V_{D0}\left(\frac{T}{T_0}\right) + \frac{\gamma kT}{q} \ln\left(\frac{T_0}{T}\right) + \frac{kT}{q} \ln\left(\frac{J_D}{J_{D0}}\right) \quad (4.6-14)$$

A close examination of Eq. (4.6-14) shows that the term  $\frac{\gamma kT}{q} \ln\left(\frac{T_0}{T}\right)$  is not linear with temperature. This term will cause  $V_{CTAT}(T)$  in Fig. 4.6-1(b) to follow the shaded, dashed line instead of the straight bold line. This problem is called the *bandgap curvature* problem. We will look at ways to overcome this problem later.

The last step is to generate currents that are CTAT. We will use the CTAT voltage of the pn junction to accomplish this objective. Figures 4.6-5(a) and 4.6-5(b) show two ways of accomplishing this using negative feedback [15]. The current flowing in the resistor  $R$  is given as

$$I_{CTAT}' = \frac{V_{BE}}{R} = \frac{V_D}{R} \quad (4.6-15)$$

The negative feedback loop is indicated on both circuits. This loop causes the current flowing in  $R$  to be equal to the current in Eq. (4.6-15). The resistor,  $R$ , will be dependent on temperature, which causes this current to be pseudo-CTAT as indicated by the primed notation. It is not necessary to find a true CTAT current because in all applications, the pseudo-CTAT current flows through a second resistor to create a true CTAT voltage assuming the two resistors have the same temperature dependence.

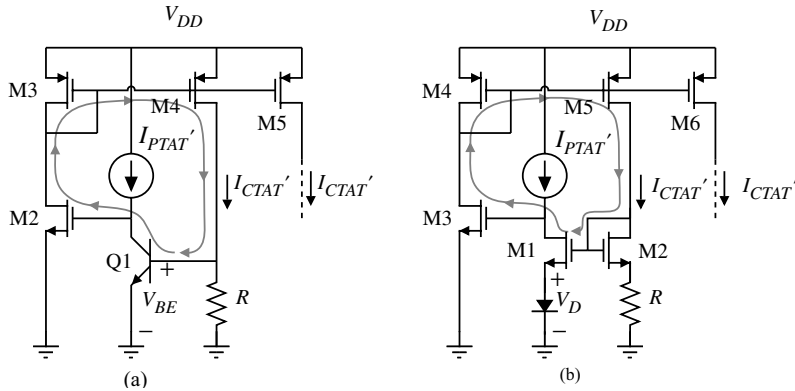


**Figure 4.6-5** Generation of a pseudo-CTAT current using (a) a bipolar transistor and (b) a diode.

A temperature-independent voltage reference can be generated using any of the techniques shown in Fig. 4.6-6. The temperature-independent voltage reference has two basic structures. They are the series realization in Fig. 4.6-6(a) and the parallel realization in Fig. 4.6-6(b). There are realizations that use a combination of the series and parallel realizations [16]. For the series realization, the temperature-independent reference voltage can be written as

$$V_{\text{REF}} = I_{PTAT}' R_2 + V_D = \left( \frac{R_2}{R_1} \right) V_{PTAT} + V_{CTAT} \quad (4.6-16)$$

The resistor,  $R_1$ , comes from Eq. (4.6-3) where  $R = R_1$  or from the realizations for  $I_{PTAT}'$  like those in Figure 4.6-3.

For the parallel realization, the temperature-independent reference voltage can be written as

$$V_{\text{REF}} = (I_{PTAT}' + I_{CTAT}') R_3 = \left( \frac{R_3}{R_1} \right) V_{PTAT} + \left( \frac{R_3}{R_2} \right) V_{CTAT} \quad (4.6-17)$$

The resistor,  $R_2$ , comes from Eq. (4.6-15) where  $R = R_2$  or from the realizations for  $I_{CTAT}'$  like those in Figure 4.6-5.

In order to achieve temperature independence, Eqs. (4.6-16) and (4.6-17) must be differentiated with respect to temperature and set equal to zero. The resistor ratios and other parameters can be used to achieve temperature independence. Let us first examine the series realization.

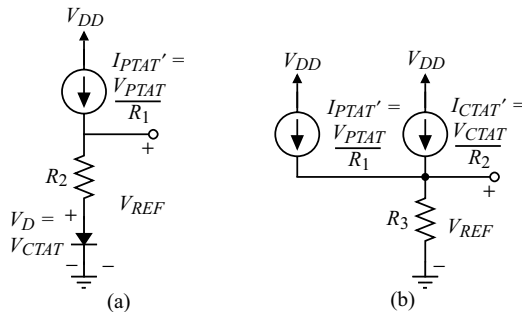


**Figure 4.6-6** Temperature-independent voltage references: (a) series form and (b) parallel form.

Differentiating Eq. (4.6-16) with respect to absolute temperature gives

$$\frac{dV_{\text{REF}}}{dT} = \left(\frac{R_2}{R_1}\right) \frac{dV_{PTAT}}{dT} + \frac{dV_{CTAT}}{dT} \quad (4.6-18)$$

assuming the  $R_1$  and  $R_2$  have identical dependence on temperature. Differentiating Eqs. (4.6-1) and (4.6-14) gives

$$\left. \frac{dV_{PTAT}}{dT} \right|_{T=T_0} = \frac{k}{q} \ln \left( \frac{J_{D2}}{J_{D1}} \right) = \frac{k}{q} \ln \left( \frac{A_2}{A_1} \right) = \frac{V_{t0}}{T_0} \ln \left( \frac{A_2}{A_1} \right) = \frac{V_{PTAT}}{T_0} \quad (4.6-19)$$

and

$$\left. \frac{dV_{CTAT}}{dT} \right|_{T=T_0} = \frac{V_D - V_{GO}}{T_0} + (\alpha - \gamma) \left( \frac{k}{q} \right) = \frac{V_{CTAT} - V_{GO}}{T_0} + (\alpha - \gamma) \left( \frac{V_{t0}}{T_0} \right) \quad (4.6-20)$$

where for Eq. (4.6-20) we have assumed that the current density,  $J_D$ , in Eq. (4.6-14) has a temperature dependence of  $T^\alpha$ . Substituting Eqs. (4.6-19) and (4.6-20) into Eq. (4.6-18) and equating to zero gives the desired result as

$$\text{Temperature-independent constant} = K = \frac{R_2}{R_1} = \frac{V_{GO} - V_{CTAT} + (\gamma - \alpha)V_{t0}}{V_{PTAT}} \quad (4.6-21)$$

Substituting Eq. (4.6-21) into Eq. (4.6-16) gives the value of  $V_{\text{REF}}$  as

$$V_{\text{REF}} = V_{GO} - V_{CTAT} + (\gamma - \alpha)V_{t0} + V_{CTAT} = V_{GO} + (\gamma - \alpha)V_{t0} \quad (4.6-22)$$

The second term of Eq. (4.6-22) is small compared to  $V_{GO}$  so that the reference voltage is slightly larger than the bandgap voltage,  $V_{GO}$ . Historically, this type of reference is called the *bandgap voltage reference* [10]. However as stated earlier, the operation of the bandgap voltage reference has nothing to do with the bandgap voltage.

#### Example 4.6-1

#### Design of $R_2/R_1$ for a Series Temperature-Independent Voltage Reference

Assuming that  $A_2/A_1 = 10$  and that  $V_{CTAT} = 0.6$  V, find the value of the temperature-independent constant,  $R_2/R_1$ , and the value of the reference voltage at room temperature where  $V_{t0} = 0.026$  V.

#### SOLUTION

Equation (4.6-21) gives

$$\frac{R_2}{R_1} = \frac{1.205 - 0.6 + 2.2(0.026)}{0.026(2.3026)} = 11.04$$

Equation (4.6-22) gives the value of  $V_{\text{REF}}$  as

$$V_{\text{REF}} = 1.205 + 2.2(0.026) = 1.262 \text{ V}$$

The parallel form of the temperature-independent voltage reference of Eq. (4.6-17) is independent of temperature under exactly the same conditions as the series form given in Eq. (4.6-21). In the case of the parallel form, the value of  $V_{REF}$  is

$$V_{REF} = \frac{R_3}{R_2} [(V_{GO} + (\gamma - \alpha)V_{t0}] \quad (4.6-23)$$

We see that the ratio of  $R_3$  to  $R_2$  can be used to achieve any desired value of temperature-independent voltage.

### Example 4.6-2

#### Design of $R_2/R_1$ and $R_3/R_2$ for a Parallel Temperature-Independent Voltage Reference

Assuming that  $A_2/A_1 = 10$  and that  $V_{CTAT} = 0.6$  V, find the value of  $R_1$ ,  $R_2$ , and  $R_3$  that gives a temperature-independent value of the reference voltage at room temperature of 0.5 V where  $V_{t0} = 0.026$  V if  $R_1 = 1$  k $\Omega$ .

#### SOLUTION

Equation (4.6-21) gives that  $R_2/R_1 = 11.04$ . Therefore,  $R_2 = 11.04$  k $\Omega$ . From Eq. (4.6-23) and Example 4.6-1, we find that

$$\frac{R_3}{R_2} = \frac{V_{REF}}{V_{GO} + (\gamma - \alpha)V_{t0}} = \frac{0.5 \text{ V}}{1.262 \text{ V}} = 0.3962$$

Therefore,  $R_3 = 0.3962(11.04 \text{ k}\Omega) = 4.376 \text{ k}\Omega$ .

One of the first realizations of the series form of the temperature-independent reference is shown in Fig. 4.6-7 [11]. This realization is compatible with a normal n-well CMOS technology using substrate bipolar transistors. In fact, the bipolar transistors could be replaced by diodes. If we assume for the present that  $V_{OS}$  is zero, then the voltage  $V_{R1}$  can be written as

$$V_{R1} = V_{EB2} - V_{EB1} = V_t \ln \left( \frac{J_2}{J_{S2}} \right) - V_t \ln \left( \frac{J_1}{J_{S1}} \right) = V_t \ln \left( \frac{I_2 A_{E1}}{I_1 A_{E2}} \right) \quad (4.6-24)$$

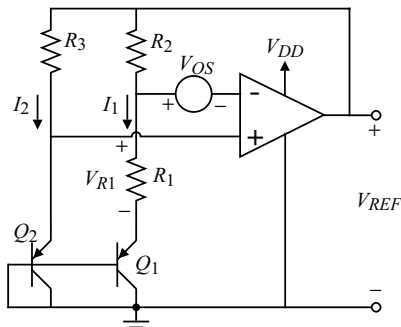


**Figure 4.6-7** A realization of the series temperature-independent voltage reference that uses an op amp.

However, the op amp also forces the relationship

$$I_1 R_2 = I_2 R_3 \quad (4.6-25)$$

Therefore, the reference voltage of Fig. 4.6-7 can be written as

$$V_{\text{REF}} = V_{EB2} + I_2 R_3 = V_{EB2} + V_{R1} \left( \frac{R_2}{R_1} \right) \quad (4.6-26)$$

Substituting Eq. (4.6-25) into Eq. (4.6-24) and the result into Eq. (4.6-26) gives

$$V_{\text{REF}} = V_{EB2} + \left( \frac{R_2}{R_1} \right) V_t \ln \left( \frac{R_2 A_{E1}}{R_3 A_{E2}} \right) \quad (4.6-27)$$

Differentiating Eq. (4.6-27) with respect to absolute temperature and setting the result to zero gives

$$\left( \frac{R_2}{R_1} \right) \ln \left( \frac{R_2 A_{E1}}{R_3 A_{E2}} \right) = \frac{V_{GO} - V_{CTAT} + (\gamma - \alpha) V_{t0}}{V_{PTAT}} \quad (4.6-28)$$

Thus, the temperature-independent constant,  $K$ , is defined in terms of resistor and emitter-base area ratios. It can be shown that if the input-offset voltage is not zero, Eq. (4.6-27) becomes

$$V_{\text{REF}} = V_{EB2} - \left( 1 + \frac{R_2}{R_1} \right) V_{OS} + \frac{R_2}{R_1} V_t \ln \left[ \frac{R_2 A_{E1}}{R_3 A_{E2}} \left( 1 - \frac{V_{OS}}{I_1 R_2} \right) \right] \quad (4.6-29)$$

It is clear that the input-offset voltage of the op amp should be small and independent of temperature in order not to deteriorate the performance of  $V_{\text{REF}}$ .

The dependence of  $V_{\text{REF}}$  on power supply can now be investigated. In Eq. (4.6-29), the only possible parameters that may depend on power supply are  $V_{EB2}$ ,  $V_{OS}$ , and  $I_1$ . Since  $V_{EB2}$  and  $I_1$  are derived from  $V_{\text{REF}}$ , the only way in which  $V_{\text{REF}}$  can depend on the power supply is through a finite power-supply rejection ratio of the op amp (manifesting itself as a variation in  $V_{OS}$ ). If the PSRR of the op amp is large, then Fig. 4.6-7 is for all practical purposes a power-supply-independent as well as a temperature-independent voltage reference.

### Example 4.6-3

#### The Design of a Bandgap-Voltage Reference

Assume that  $A_{E1} = 10 A_{E2}$ ,  $V_{EB2} = 0.7 \text{ V}$ ,  $R_2 = R_3$ , and  $V_t = 0.026 \text{ V}$  at room temperature for Fig. 4.6-7. Find  $R_2/R_1$  to give a zero temperature coefficient at room temperature. If  $V_{OS} = 10 \text{ mV}$ , find the change in  $V_{\text{REF}}$ . Note that  $I_1 R_2 = V_{\text{REF}} - V_{EB2} - V_{OS}$ .

#### SOLUTION

Evaluating Eq. (4.6-28) gives

$$\left( \frac{R_2}{R_1} \right) \ln \left( \frac{R_2 A_{E1}}{R_3 A_{E2}} \right) = \frac{V_{GO} - V_{CTAT} + (\gamma - \alpha) V_{t0}}{V_{PTAT}} = \frac{1.205 - 0.7 + (2.2)(0.026)}{0.026} = 21.62$$

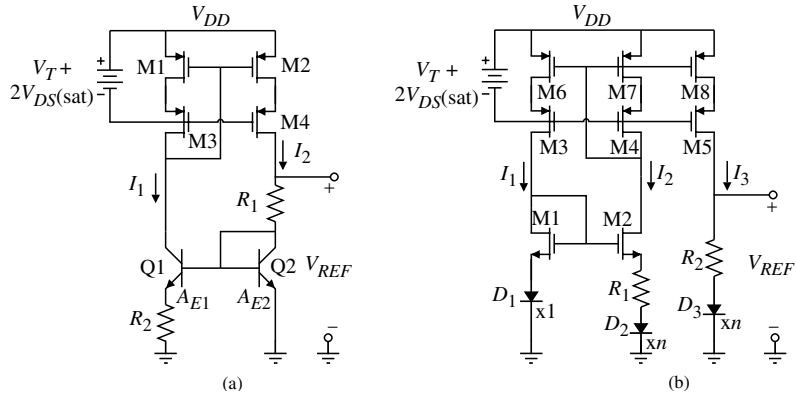


**Figure 4.6-8** Series temperature-independent voltage references using (a) MOSFETs and BJTs and (b) only MOSFETs and diodes.

Therefore,  $R_2/R_1 = 9.39$ . In order to use Eq. (4.6-29), we must know the approximate value, of  $V_{REF}$  and iterate if necessary because  $I_1$  is a function of  $V_{REF}$ . Assuming  $V_{REF}$  to be 1.262, we obtain from Eq. (4.6-29) a new value,  $V_{REF} = 1.153$  V. The second iteration makes little difference on the result because  $V_{REF}$  is in the argument of the logarithm.

While an op amp can be used to implement the temperature-independent voltage reference, it has several disadvantages, including the influence of the input-offset voltage (which is a function of temperature), the input noise of the op amp, and the additional complexity of the op amp. For the remainder of this section, we will examine temperature-independent voltage references that do not use op amps.

Figure 4.6-8 shows two examples of the series form of the temperature-independent voltage references that do not use op amps. Figure 4.6-8(a) uses a cascode current mirror (M1–M4) to more precisely force  $I_1$  to be equal to  $I_2$ . We see that  $I_1$  is pseudo-PTAT and is given as

$$\begin{aligned} I_1 = I_{PTAT}' &= \frac{V_{BE2} - V_{BE1}}{R_2} = \frac{V_t}{R_2} \left[ \ln\left(\frac{I_2}{I_{s2}}\right) - \ln\left(\frac{I_1}{I_{s1}}\right) \right] \\ &= \frac{V_t}{R_2} \ln\left(\frac{I_{s1}}{I_{s2}}\right) = \frac{V_t}{R_2} \ln\left(\frac{A_{E1}}{A_{E2}}\right) \end{aligned} \quad (4.6-30)$$

Since  $I_1 = I_2$ , we can write  $V_{REF}$  as

$$\begin{aligned} V_{REF} &= V_{BE2} + I_1 R_1 = V_{BE1} + \left[ \frac{R_1}{R_2} \ln\left(\frac{A_{E1}}{A_{E2}}\right) \right] V_t \\ &= V_{CTAT} + \left[ \frac{R_1}{R_2} \ln\left(\frac{A_{E1}}{A_{E2}}\right) \right] V_{PTAT} \end{aligned} \quad (4.6-31)$$

We see that Eq. (4.6-31) is similar in form to the previous equations for a series temperature-independent voltage reference. To implement the reference, simply set the temperature-independent multiplier of  $V_{PTAT}$  equal to Eq. (4.6-21).

Figure 4.6-8(b) gives another realization of the series temperature-independent voltage reference. M1 and M2 are used to enforce the following relationship:

$$V_{D1} = I_2 R_1 + V_{D2} \quad (4.6-32)$$

or

$$I_3 = I_2 = I_{PTAT}' = \frac{V_t}{R_1} \ln(n) \quad (4.6-33)$$

which is a pseudo-PTAT current. The reference voltage is written as

$$\begin{aligned} V_{REF} &= V_{D3} + I_3(R_2) = V_{D3} + \frac{R_2}{R_1} V_t \ln(n) \\ &= V_{CTAT} + \frac{R_2}{R_1} \ln(n) V_{PTAT} \end{aligned} \quad (4.6-34)$$

The temperature-independent constant,  $(R_2/R_1) \ln(n)$ , is equated to Eq. (4.6-21) to complete the design. Since  $n$  is an integer greater than 1,  $R_2/R_1$  and  $n$  can be used to satisfy Eq. (4.6-21). It is interesting to note that all of the series forms of the temperature-independent voltage reference will have a voltage close to 1.262 V at room temperature.

The parallel form of the temperature-independent voltage reference requires a pseudo-PTAT and CTAT current. Combining Fig. 4.6-3(a) and Fig. 4.6-5(a) gives one possible realization shown in Fig. 4.6-9. If the gate of a PMOS transistor is connected to one of the darkened horizontal busses and the source to  $V_{DD}$ , the drain current will be either  $I_{PTAT}'$  or  $I_{CTAT}'$ . The bipolar transistor in the  $I_{CTAT}'$  generator can be replaced by the M1, M2, and diode of Fig. 4.6-5(b) to achieve a MOSFET–diode parallel temperature-independent voltage reference. Example 4.6-2 gave an example of how to design the resistors of Fig. 4.6-9.

There are many different realizations of the *bandgap voltage* reference [17]. A typical family of reference-voltage variations as a function of  $T$  for various values of  $T_0$  is shown in Fig. 4.6-10. The influence of the nonlinear term of the CTAT voltage can clearly be seen in this figure. True temperature independence is achieved only over a small range of temperatures. Most of the temperature-independent voltage references that do not use a curvature correction technique have a temperature dependence of 10–50 ppm/°C over 0–70°C. This undesirable characteristic is called the *bandgap curvature* problem.

## Voltage References with Excellent Temperature Stability

To achieve temperature stability that exceeds the above voltage references, it is necessary to solve the bandgap curvature problem. There have been many ideas forwarded on how to

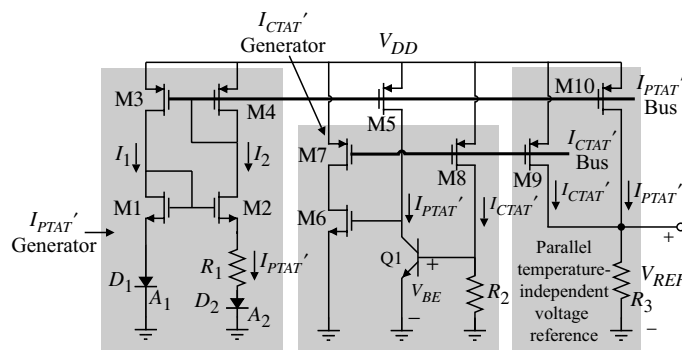


**Figure 4.6-9** A parallel form of the temperature-independent voltage reference.

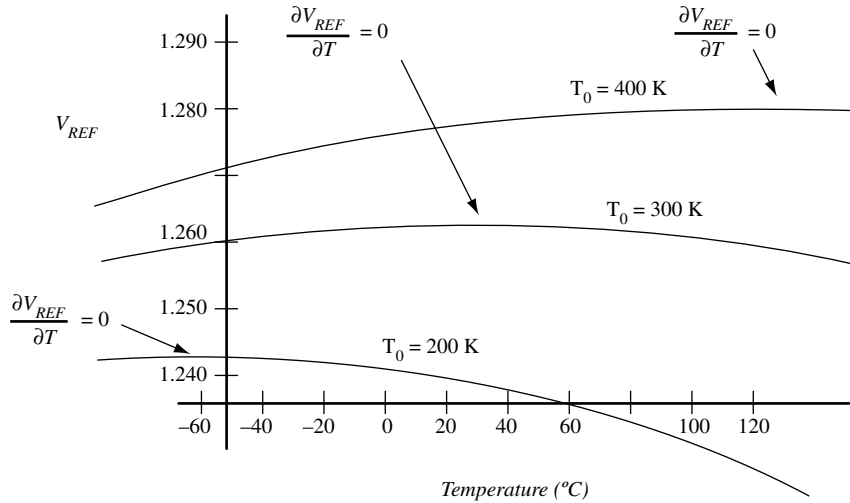


**Figure 4.6-10** Variation of bandgap reference output with temperature (© John Wiley and Sons, Inc.).

solve this problem [15–20]. We will examine one of these solutions that allows a temperature dependence of less than 1 ppm/°C over a 0–100°C range. This method was first developed for the series form of the temperature-independent voltage reference [20]. The idea is to cancel out the nonlinear term in the CTAT voltage of Eq. (4.6-14). Figure 4.6-11(a) shows a slightly modified form of the series temperature-independent voltage reference. Figure 4.6-11(b) shows one that is capable of reducing the bandgap curvature problem.  $V_{REF}$  can be written as

$$V_{REF} = V_{PTAT} + 3V_{CTAT} - 2V_{constant} \quad (4.6-35)$$

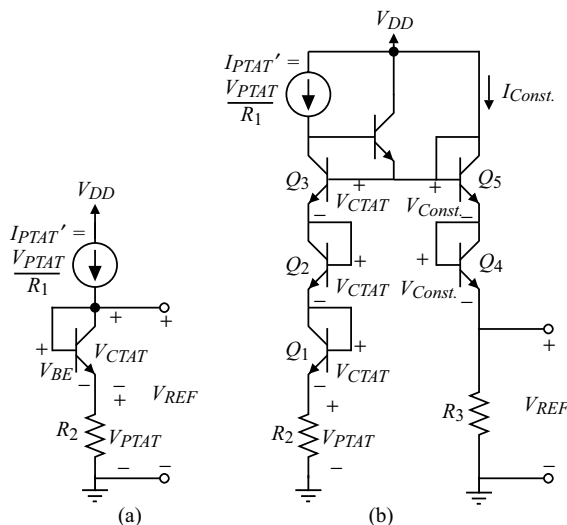


**Figure 4.6-11** (a) Series temperature-independent voltage reference. (b) Method of correcting for the bandgap curvature problem.

As before, if we assume that the current density,  $J_D$ , in Eq. (4.6-14) has a temperature dependence of  $T^\alpha$ , then we can rewrite Eq. (4.6-14) as

$$V_{CTAT} = V_{BE}(T) = V_{GO} - \frac{T}{T_0} [V_{GO} - V_{CTAT}(T_0)] - (\gamma - \alpha) V_t \ln\left(\frac{T}{T_0}\right) \quad (4.6-36)$$

If the current through a pn junction is pseudo-PTAT, then  $\alpha = 1$  and we can express the voltage across the pn junction as

$$V_{CTAT} = V_{GO} - \frac{T}{T_0} [V_{GO} - V_{CTAT}(T_0)] - (\gamma - 1) V_t \ln\left(\frac{T}{T_0}\right) \quad (4.6-37)$$

If we put a temperature-independent current ( $I_{Const.}$ ) through a pn junction, then  $\alpha = 0$  and we can express the voltage across the pn junction as

$$V_{Const} = V_{GO} - \frac{T}{T_0} [V_{GO} - V_{CTAT}(T_0)] - \gamma V_t \ln\left(\frac{T}{T_0}\right) \quad (4.6-38)$$

Substituting Eqs. (4.6-37) and (4.6-38) into Eq. (4.6-35) gives

$$V_{REF} = V_{PTAT} + V_{GO} - \frac{T}{T_0} [V_{GO} - V_{CTAT}(T_0)] - (\gamma - 3) V_t \ln\left(\frac{T}{T_0}\right) \quad (4.6-39)$$

If  $\gamma \approx 3$ , then Eq. (4.6-39) becomes

$$V_{REF} \approx V_{PTAT} + V_{GO} - \frac{T}{T_0} [V_{GO} - V_{CTAT}(T_0)] \quad (4.6-40)$$

To complete this example, the PTAT voltage will have to be multiplied by the constant  $K$  given in Eq. (4.6-21). This is accomplished by the resistors  $R_1$  and  $R_2$  and any component area ratios that are part of the pseudo-PTAT current generation. The unique aspect of Eq. (4.6-40) is that the nonlinear CTAT term,  $\ln(T/T_0)$ , has been cancelled, eliminating or at least minimizing the bandgap curvature problem.

The stacking of pn junctions is not attractive as power supplies decrease. The previous idea can be implemented more precisely and at a smaller value of power supply in the parallel form of the temperature-independent voltage reference. The implementation of Fig. 4.6-11(b) in a parallel form is shown as blocks in Fig. 4.6-12. The various forms of the  $I_{PTAT}'$  generator and  $I_{CTAT}'$  generator have already been described in the previous discussion. The  $I_{Const.}'$  generator is developed using the principle of Fig. 4.6-4 and is shown in Fig. 4.6-13. We solve for  $V_{GS}(ZTC)$  as

$$V_{GS}(ZTC) = (I_{PTAT}' + I_{CTAT}')R_3 = \left(\frac{R_3}{R_1}\right)V_{PTAT} + \left(\frac{R_3}{R_2}\right)V_{CTAT} \quad (4.6-41)$$

The value of the constant current depends on the ZTC characteristics of the NMOS transistor and the  $W/L$  that is selected and is designated as  $I_{Const.}$

If the pseudo-PTAT current in the pseudo-CTAT generators of Fig. 4.6-5 is replaced with the constant current generated above, then the current generated will be CTAT but  $\alpha$  of Eq. (4.6-36) becomes zero. This current is designated as  $I_{\alpha=0}'$ . Finally, the three currents,  $I_{PTAT}'$ ,  $I_{CTAT}'$ , and  $I_{\alpha=0}'$ , are combined in the low TC voltage reference of Fig. 4.6-14 to achieve a

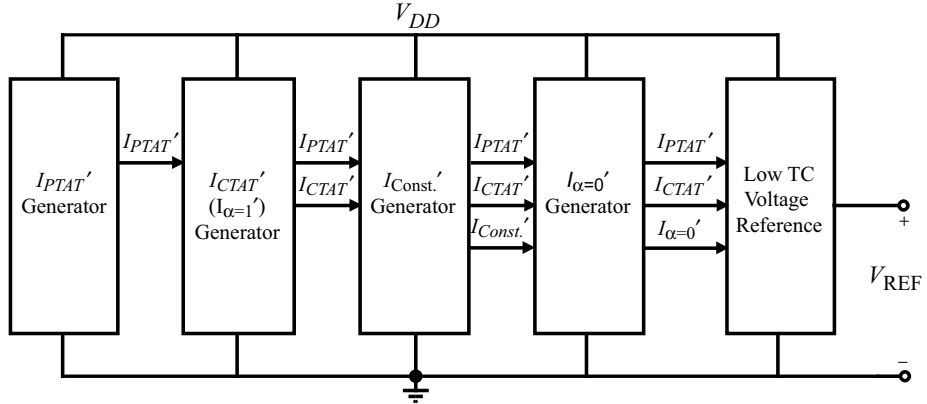


**Figure 4.6-12** Block diagram of the parallel form of the curvature corrected temperature stable reference.

highly temperature-stable voltage reference. To simplify the situation, we will assume the  $W/L$  ratios of all NMOS transistors are equal and the  $W/L$  ratios of all PMOS transistors are equal to twice the  $W/L$  ratio of the NMOS transistors. The output reference voltage can be written as

$$\begin{aligned} V_{REF} &= R_5(I_{PTAT}' + I_{CTAT}' - I_{\alpha=0}') \\ &= \left(\frac{R_5}{R_1}\right)V_{PTAT} + \left(\frac{R_5}{R_2}\right)V_{CTAT}(\alpha = 1) - \left(\frac{R_5}{R_4}\right)V_{CTAT}(\alpha = 0) \end{aligned} \quad (4.6-42)$$

Substituting Eqs. (4.6-37) and (4.6-38) into Eq. (4.6-42) gives

$$\begin{aligned} V_{REF} &= \frac{R_5}{R_1} V_{PTAT} + \frac{R_5}{R_2} \left[ V_{GO} - \frac{T}{T_0} [V_{GO} - V_{CTAT}(T_0)] - (\gamma - 1)V_t \ln\left(\frac{T}{T_0}\right) \right] \\ &\quad - \frac{R_5}{R_4} \left[ V_{GO} - \frac{T}{T_0} [V_{GO} - V_{CTAT}(T_0)] - \gamma V_t \ln\left(\frac{T}{T_0}\right) \right] \end{aligned} \quad (4.6-43)$$

To cancel the nonlinear CTAT term, we want the following relationship to be true:

$$\frac{R_5}{R_2}(\gamma - 1) = \frac{R_5}{R_4}\gamma \quad (4.6-44)$$

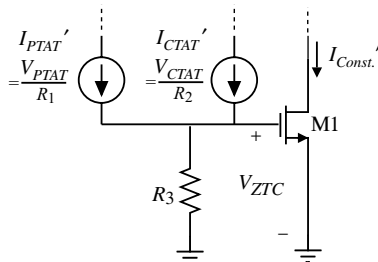


**Figure 4.6-13** Implementation of the constant-current generator.

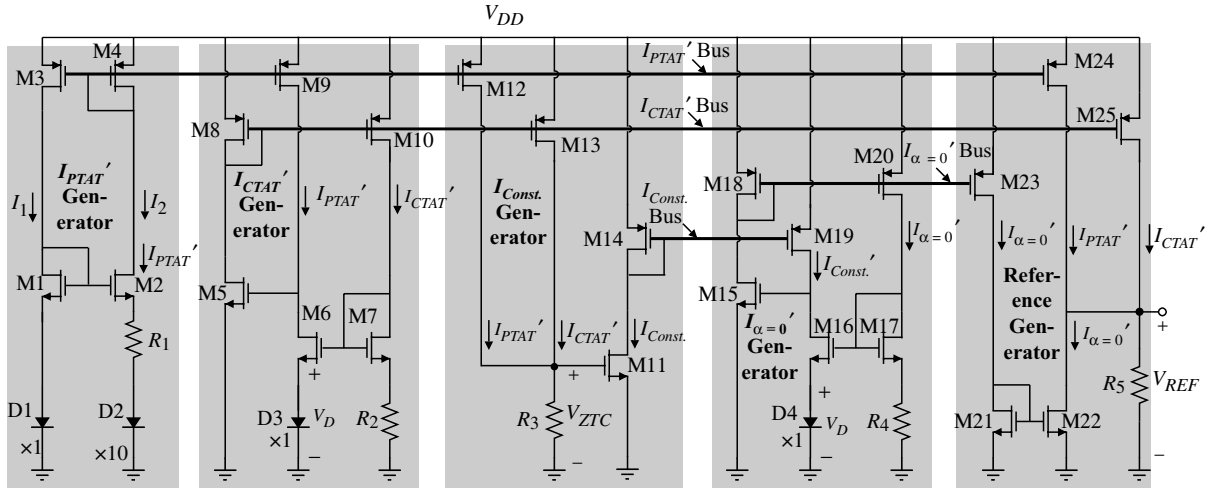


**Figure 4.6-14** Transistor-level implementation of the low TC voltage reference in Fig. 4.6-13.

This relationship gives

$$\frac{R_2}{R_4} = \frac{(\gamma - 1)}{\gamma} \quad (4.6-45)$$

Fortunately,  $\gamma$  is always greater than one. With the relationship of Eq. (4.6-45) valid, Eq. (4.6-43) becomes

$$V_{REF} = \frac{R_5}{R_1} V_{PTAT} + \left( \frac{R_5}{R_4(\gamma - 1)} \right) \left[ V_{GO} - \frac{T}{T_0} [V_{GO} - V_{CTAT}(T_0)] \right] \quad (4.6-46)$$

Rewriting Eq. (4.6-46) as follows gives the form necessary to find the relationship that will ensure the positive and negative temperature coefficients are cancelled:

$$V_{REF} = \frac{R_5}{R_4(\gamma - 1)} \left[ \frac{R_4(\gamma - 1)}{R_1} V_{PTAT} + V_{GO} - \frac{T}{T_0} [V_{GO} - V_{CTAT}(T_0)] \right] \quad (4.6-47)$$

Differentiating Eq. (4.6-47) with respect to temperature and setting  $T = T_0$  gives

$$\frac{\partial V_{REF}}{\partial T} \bigg|_{T=T_0} = \frac{R_5}{R_4(\gamma - 1)} \left[ \frac{R_4(\gamma - 1)}{R_1} \frac{V_{PTAT}}{T_0} - \frac{1}{T_0} [V_{GO} - V_{CTAT}(T_0)] \right] \quad (4.6-48)$$

Setting Eq. (4.6-48) equal to zero gives

$$\frac{R_4}{R_1} = \frac{V_{GO} - V_{CTAT}(T_0)}{(\gamma - 1)V_{PTAT}} \quad (4.6-49)$$

Substituting Eq. (4.6-49) into Eq. (4.6-47) for  $T = T_0$  gives

$$\begin{aligned} V_{\text{REF}} &= \frac{R_5}{R_4(\gamma - 1)} [V_{GO} - V_{CTAT}(T_0) + V_{GO} - V_{GO} + V_{CTAT}(T_0)] \\ &= \frac{R_5}{R_4(\gamma - 1)} V_{GO} \end{aligned} \quad (4.6-50)$$

Equation (4.6-50) will be used to design the value of  $R_5/R_4$  as

$$\frac{R_5}{R_1} = \frac{V_{\text{REF}}(\gamma - 1)}{V_{GO}} \quad (4.6-51)$$

Thus, the design equations for Fig. 4.6-14 are Eqs. (4.6-45), (4.6-49), and (4.6-51). The value of  $R_3$  is designed based on the value of  $V_{GS}(\text{ZTC})$  from Eq. (4.6-41). The values of these resistors could be modified by adjusting the  $W/L$  ratios of M23, M24, and M25.

### Example 4.6-4

#### Design of a Zero Temperature Coefficient Voltage Reference

Assume that  $V_{CTAT} = 0.7 \text{ V}$ ,  $R_4 = 10 \text{ k}\Omega$ ,  $\gamma = 3.2$ , and  $V_t = 0.026 \text{ V}$  at room temperature for Fig. 4.6-14. Find  $R_1$ ,  $R_2$ , and  $R_5$  to give a zero temperature coefficient at room temperature and a reference voltage of  $1.0 \text{ V}$ .

#### SOLUTION

From Eq. (4.6-45), we get  $R_2 = 0.6875 R_4 = 6.88 \text{ k}\Omega$ . From Eq. (4.6-49) we get

$$\frac{R_4}{R_1} = \frac{V_{GO} - V_{CTAT}(T_0)}{(\gamma - 1)V_{PTAT}} = \frac{1.205 \text{ V} - 0.7 \text{ V}}{2.2(0.026 \text{ V})2.306} = 3.83$$

Thus,  $R_1 = R_4/3.83 = 2.61 \text{ k}\Omega$ . From Eq. (4.6-51) we get

$$\frac{R_5}{R_1} = \frac{V_{\text{REF}}(\gamma - 1)}{V_{GO}} = \frac{1 \text{ V}(2.2)}{1.205 \text{ V}} = 1.83$$

Thus,  $R_5 = 1.83 R_1 = 1.83(2.61 \text{ k}\Omega) = 4.77 \text{ k}\Omega$ .

It is important to keep any influences on the temperature-independent voltage reference less than the inherent temperature variation. If, for example, dc drifts or noise causes a variation of the voltage outside of the inherent temperature variation, then this disturbance becomes larger than the temperature variation. If a very good temperature-independent voltage reference has a temperature dependence of  $1 \text{ ppm}/^\circ\text{C}$  over a  $100^\circ\text{C}$  range, then the maximum temperature deviation of the reference is  $100 \text{ ppm}$  or  $0.01\%$ . If the nominal reference voltage is  $1 \text{ V}$ , then the deviation is  $100 \mu\text{V}$ . If the noise on the output of the voltage reference, or the dc offset or drift, or the PSRR of the op amp (if one is used) is greater than  $100 \mu\text{V}$ , then the possible temperature independence is not achieved. As the deviation of the voltage reference with respect to temperature becomes small, the designer must be aware of these other influences.

Suppose that a temperature-independent current is desired. There are at least two ways of accomplishing this. The first is to generate a temperature-independent voltage that is equal to the ZTC gate-source voltage of a MOSFET. We have illustrated this approach in our previous considerations. The only disadvantage of this approach is that if the temperature variation becomes too large, the ZTC point will move. A second approach is to place the temperature-independent reference voltage across a resistor, thus generating a  $V_{\text{REF}}/R$  current. The obvious problem with this is the lack of a temperature-independent resistor. If we divide Eq. (4.6-16) by a resistor,  $R$ , to get a reference current, the result is

$$I_{\text{REF}} = \frac{V_{\text{REF}}}{R} = \frac{1}{R} \left[ \left( \frac{R_2}{R_1} \right) V_{PTAT} + V_{CTAT} \right] \quad (4.6-52)$$

Differentiating with respect to absolute temperature gives

$$\begin{aligned} \frac{dI_{\text{REF}}}{dT} &= \frac{1}{R} \left[ \left( \frac{R_2}{R_1} \right) \frac{dV_{PTAT}}{dT} + \frac{dV_{CTAT}}{dT} \right] - \frac{1}{R^2} \frac{dR}{dT} \left[ \left( \frac{R_2}{R_1} \right) V_{PTAT} + V_{CTAT} \right] \\ &= \frac{1}{R} \left[ \left( \frac{R_2}{R_1} \right) \frac{dV_{PTAT}}{dT} + \frac{dV_{CTAT}}{dT} \right] - \frac{1}{R} \frac{dR}{dT} I_{\text{REF}} = 0 \end{aligned} \quad (4.6-53)$$

To achieve a temperature-independent reference current, the value of the temperature-independent constant,  $R_2/R_1$ , must be

$$\frac{R_2}{R_1} = \frac{I_{\text{REF}} T_0 \frac{dR}{dT} - V_{CTAT} + (\gamma - \alpha) V_{T0}}{V_{PTAT}} \quad (4.6-54)$$

Assuming the values of Example 4.6-1 and that  $dR/dT = 100 \, \Omega/^\circ\text{C}$  and  $I_{\text{REF}} = 100 \, \mu\text{A}$ , we get a value of  $R_2/R_1$  equal to

$$\frac{R_2}{R_1} = \frac{(100 \, \mu\text{A})(300 \, \text{K})(100 \, \Omega/\text{K}) - 0.6 \, \text{V} + 2.2(0.026)}{2.306(0.026)} = 40.98 \quad (4.6-55)$$

Depending on the value of  $dR/dT$  and  $I_{\text{REF}}$ , the value of  $R_2/R_1$  may not be positive and this method will not work unless the current or the temperature behavior of the resistor is different.

Although other techniques have been used to develop power-supply and temperature-independent references, the bandgap circuit has proven the best to date. In this section we have used the bandgap concept to develop precision references. As the requirement for higher precision increases, the designer will find it necessary to begin including second-order and sometimes third-order effects that might normally be neglected. These higher-order effects require the designer to be familiar with the physics and operation of the MOS devices and passive components.



## 4.7 Summary

This chapter has introduced CMOS subcircuits, including the switch, active resistors, current sinks/sources, current mirrors or amplifiers, and voltage and current references. The general principles of each circuit were covered, as was their large-signal and small-signal performance. Remember that the circuits presented in this chapter are rarely used by themselves; rather, they are joined with other such circuits to implement a desired analog function.

The approach used in each case was to present a general understanding of the circuit and how it works. This presentation was followed by analysis of large-signal performance, typically a voltage-transfer function or a voltage–current characteristic. Limitations such as signal swing or nonlinearity were identified and characterized. This was followed by the analysis of small-signal performance. The important parameters of small-signal performance include ac resistance, voltage gain, and bandwidth.

The subject matter presented in this chapter will be continued and extended in the next chapter. A good understanding of the circuits in this and the next chapter will provide a firm foundation for the later chapters and subject material.



## 4.8 Design Problems

It is common to hear an experienced designer say, “I am not sure what works, but I know several things that do not work.” This understanding comes from the experience acquired through various design experiences. As an analog designer, it is very important to be able to identify which topology gives the best compromise under certain design constraints. Starting with the problems of this chapter, design problems are introduced to get the reader acquainted with the various design trade-offs of analog circuit design and to increase your understanding of the practical problem of design. These problems will be clearly identified in each chapter that has them.

Below are some guidelines that will help you approach these design problems.

1. Select a topology that you believe will best achieve the design goals. Do not forget, you can always make things more complex: what is important is to keep things as simple as possible while meeting the specifications.
2. Calculate the bias currents, transistor aspect ratios, and bias voltages that will optimize the circuit toward achieving the specifications (e.g., slew rate, GBW, signal swing, etc.). In this step you should do the small-signal analysis and/or derivations by hand for the circuit. Show all your calculations and briefly explain your reasoning.
3. Simulate the circuit in SPICE or any SPICE-like simulator (e.g., Spectre, SpectreS, etc.). You may need to do some tweaking in this step. If the simulation results are significantly different from your hand calculations (say, more than  $\pm 30\%$ ), then make sure that your reasoning and hand calculations are correct.
4. Provide any relevant simulation results (e.g., frequency response, step response, etc.).
5. Provide a table summarizing your simulation results (e.g., slew rate, GBW, dc gain, signal swing, etc.).
6. Provide your final schematic along with the transistor aspect ratios and bias currents. Doing so may result in feedback and suggestions on your design, which may improve your understanding of the circuit.

All of the design problems are self-grading. By this we mean that you will be given a formula for the score of the problem. Various specifications of the problem have been given a particular weighting. It is your responsibility to make the best choices and trade-offs to maximize your score. Some guidelines for helping you with the problem and its scoring are as follows:

1. These problems are normally designed so that to get a full score requires an inordinate amount of design time. One of the things you are to learn is the trade-off between the time you can invest in the problem and the score you want to achieve.
2. If you submit your design to an instructor, you must provide the simulation information in the form of at least the input file or schematic input information and the output of the simulator using whatever postprocessing necessary to clearly present the results.
3. Remember, there is no “right” answer, and your responsibility is to find the best answer under the given circumstances including how much time you are willing to spend on the problem.

# Problems

- 4.1-1. Using SPICE, generate a set of parametric  $I$ - $V$  curves similar to Fig. 4.1-3 for a transistor with a  $W/L = 10/1$ . Use model parameters from Table 3.1-2.
- 4.1-2. The circuit shown in Fig. P4.1-2 illustrates a single-channel MOS resistor with a  $W/L$  of  $2\text{ }\mu\text{m}/1\text{ }\mu\text{m}$ . Using Table 3.1-2 model parameters, calculate the small-signal ON resistance of the MOS transistor at various values for  $V_S$  and fill in the table below.

| $V_S$ (volts) | $R$ (ohms) |
|---------------|------------|
| 1.0           |            |
| 3.0           |            |
| 5.0           |            |

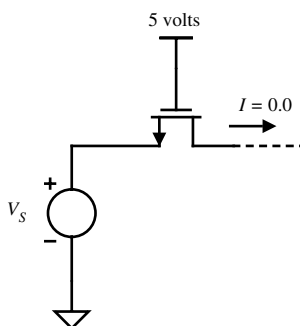


Figure P4.1-2

- 4.1-3. The circuit shown in Fig. P4.1-3 illustrates a single-channel MOS resistor with a  $W/L$  of  $4\text{ }\mu\text{m}/1\text{ }\mu\text{m}$ . Using Table 3.1-2 model parameters, calculate the small-signal ON resistance of the MOS transistor at various values for  $V_S$  and fill in the table below. Note that the most positive supply voltage is 5 V.

| $V_S$ (volts) | $R$ (ohms) |
|---------------|------------|
| 1.0           |            |
| 3.0           |            |
| 5.0           |            |

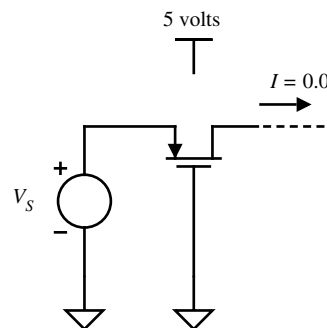


Figure P4.1-3

- 4.1-4. The circuit shown in Fig. P4.1-4 illustrates a complementary MOS resistor with an n-channel  $W/L$  of  $2\text{ }\mu\text{m}/1\text{ }\mu\text{m}$  and a p-channel

$W/L$  of  $4\ \mu\text{m}/1\ \mu\text{m}$ . Using Table 3.1-2 model parameters, calculate the small-signal ON resistance of the complementary MOS resistor at various values for  $V_S$  and fill in the table below. Note that the most positive supply voltage is 5 V.

| $V_S$ (volts) | $R$ (ohms) |
|---------------|------------|
| 1.0           |            |
| 3.0           |            |
| 5.0           |            |

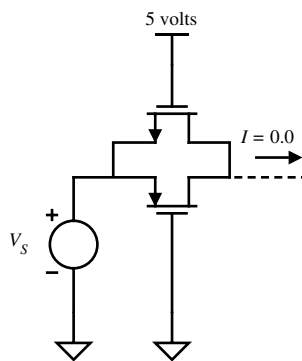


Figure P4.1-4

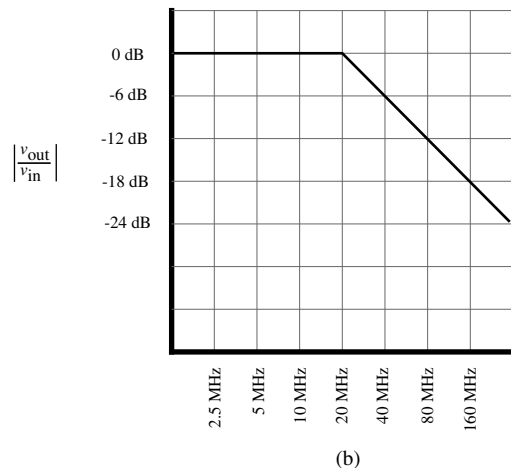
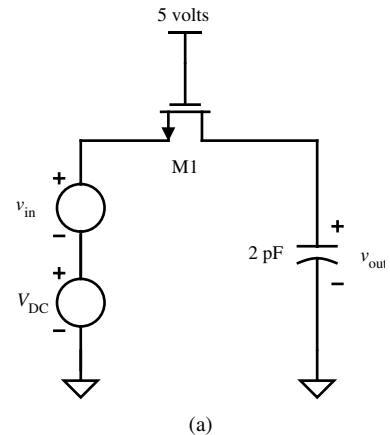


Figure P4.1-5

- 4.1-5. For the circuit in Fig. P4.1-5(a) assume that there are NO capacitance parasitics associated with M1. The voltage source  $v_{in}$  is a small-signal value, whereas voltage source  $V_{DC}$  has a dc value of 3 V. Design M1 to achieve the asymptotic frequency response shown in Fig. P4.1-5(b).
- 4.1-6. Using the result of Problem 4.1-5, calculate the frequency response resulting from changing the gate voltage of M1 to 4.5 V. Draw a Bode diagram of the resulting frequency response.
- 4.1-7. Consider the circuit shown in Fig. P4.1-7. Assume that the *slow regime* of charge injection is valid for this circuit. Initially, the charge on  $C_1$  is zero. Calculate  $v_{out}$  at time  $t_1$  after  $\phi_1$  pulse occurs. Assume that CGSO and CGDO are both 5 fF.  $C_1 = 30$  fF. You cannot ignore body effect.  $L = 1.0\ \mu\text{m}$  and  $W = 5.0\ \mu\text{m}$ . Use the model parameters from Tables 3.1-2 and 3.2-1 as required.

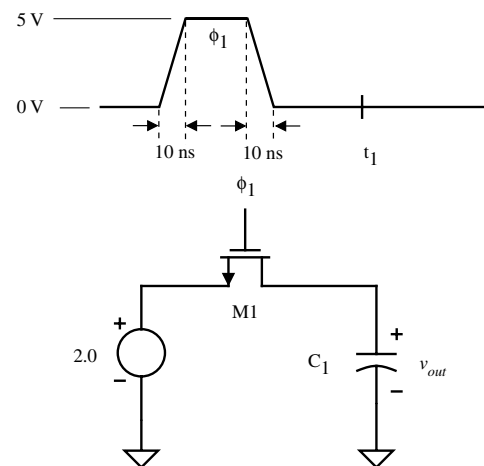


Figure P4.1-7

- 4.1-8. In Problem 4.1-7, how long must  $\phi_1$  remain high for  $C_1$  to charge up to 99% of the desired final value (2.0 V)?
- 4.1-9. In Problem 4.1-7, the charge feedthrough could be reduced by reducing the size of M1 or increasing the size of  $C_1$ . What impact do both of these changes have on the width of the  $\phi_1$  pulse?
- 4.1-10. Considering charge feedthrough due to slow regime only, will reducing the magnitude of the  $\phi_1$  pulse impact the resulting charge feedthrough? What impact does reducing the magnitude of the  $\phi_1$  pulse have on the accuracy of the voltage transfer to the output?
- 4.1-11. Repeat Example 4.1-1 with the following conditions. Calculate the effect of charge feedthrough on the circuit shown in Fig. 4.1-9, where  $V_S = 1.5$  V,  $C_L = 150$  fF,  $W/L = 1.6 \mu\text{m}/0.8 \mu\text{m}$ . The fall time for Case 1 and Case 2 is 0.1 ns and 8 ns, respectively.
- 4.1-12. Figure P4.1-12 illustrates a circuit that contains a charge-cancellation scheme. Design the size of M2 to minimize the effects of charge feedthrough. Assume a slow rise and fall time of the clock.

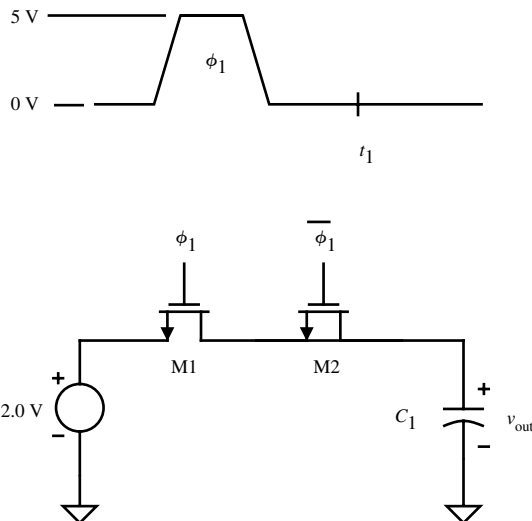


Figure P4.1-12

- 4.1-13. Calculate the effect of charge injection on the circuit shown in Fig. P4.1-13. Assume that  $U = 5$  V/50 ns =  $10^8$  V/s and  $v_{in} = 2.5$  V and ignore the bulk effect. Use the model parameters from Tables 3.1-2 and 3.2-1.

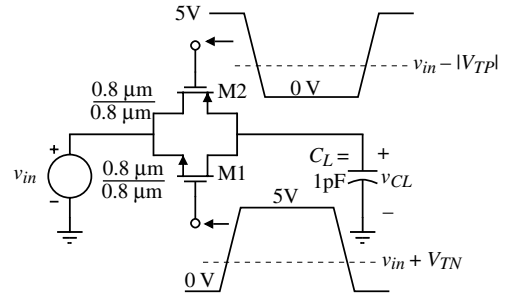


Figure P4.1-13

- 4.1-14. The four MOS diodes shown in Fig. P4.1-14 form a switch connecting  $v_{in}$  to  $1 \text{ k}\Omega$  load resistor when  $I_{control} = 200 \mu\text{A}$ . Find the small-signal value of the on-resistance of this switch configuration when  $I_{control} = 200 \mu\text{A}$ . What is the switch loss in dB when the switch is on? If this switch (the four MOS diodes and controlling current) was inserted in place of the MOSFET switch in Fig. 4.1-6, compare the performance of the two switch realizations.
- 4.2-1 Find an expression for the resistance  $R_{AB}$  for the parallel combination of identical NMOS transistors shown in Fig. P4.2-1. Assume that the transistors are in the active or triode region of operation and  $v_{DS}$  is small but not zero. Rederive the expression for resistance of Fig. 4.2-3 given in Eq. (4.2-6) assuming the same large-signal model parameters and  $W/L$  used for Fig. P4.2-1 when  $v_{DS}$  is small but not zero. Which of the two realizations, Fig. 4.2-3 or Fig. P4.2-1, is more linear?
- 4.3-1 Figure P4.3-1 illustrates a source-degenerated current source. Using Table 3.1-2 model parameters, calculate the output resistance at the given current bias. Ignore the body effect.

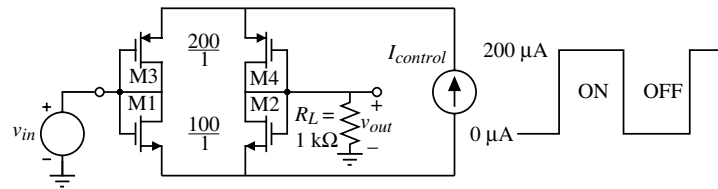


Figure P4.1-14

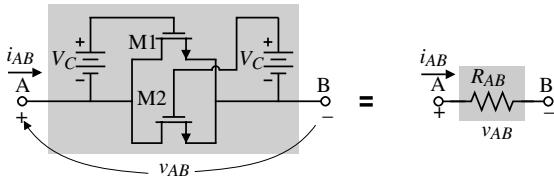


Figure P4.2-1

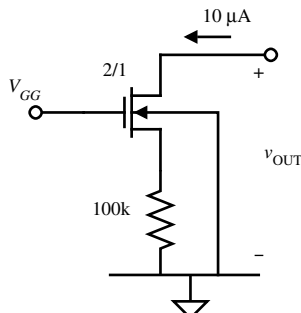


Figure P4.3-1

- 4.3-2 Using the cascode circuit shown in Fig. P4.3-2, design the  $W/L$  of M1 to achieve the same output resistance as the circuit in Fig. P4.3-1. Ignore the body effect.

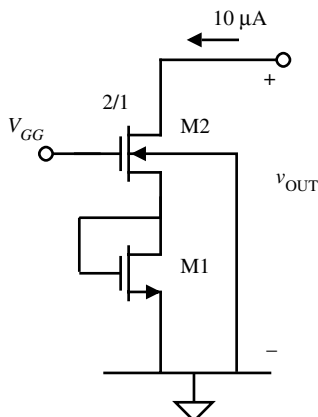


Figure P4.3-2

- 4.3-3. Calculate the minimum output voltage required to keep the device in saturation in Problem 4.3-1. Compare this result with that of Problem 4.3-2. Which circuit is a better choice in most cases?

- 4.3-4. Calculate the output resistance and the minimum output voltage, while maintaining all devices in saturation, for the circuit shown in Fig. P4.3-4. Assume that  $i_{OUT}$  is actually  $10 \mu A$ . Simulate this circuit using SPICE LEVEL 3 model (Table 3.4-1) and determine the actual output current,  $i_{OUT}$  and small-signal output resistance,  $r_{OUT}$ . Use Table 3.1-2 for device model information.

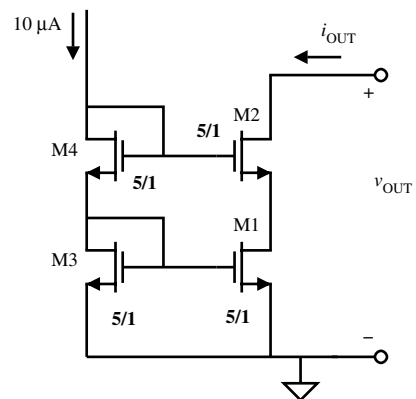


Figure P4.3-4

- 4.3-5. Calculate the output resistance and the minimum output voltage, while maintaining all devices in saturation, for the circuit shown in Fig. P4.3-5. Assume that  $i_{OUT}$  is actually  $10 \mu A$ . Simulate this circuit using SPICE LEVEL 3 model (Table 3.4-1) and determine the actual output current,  $i_{OUT}$ . Use Table 3.1-2 for device model information.

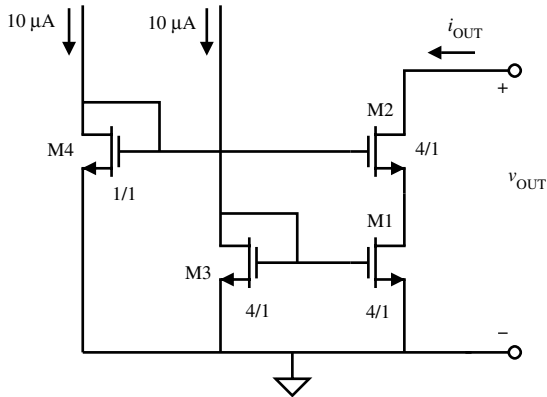


Figure P4.3-5

- 4.3-6. Design M3 and M4 of Fig. P4.3-6 so that the output characteristics are identical to the circuit shown in Fig. P4.3-5. It is desired that  $i_{OUT}$  is ideally  $10\ \mu\text{A}$ .

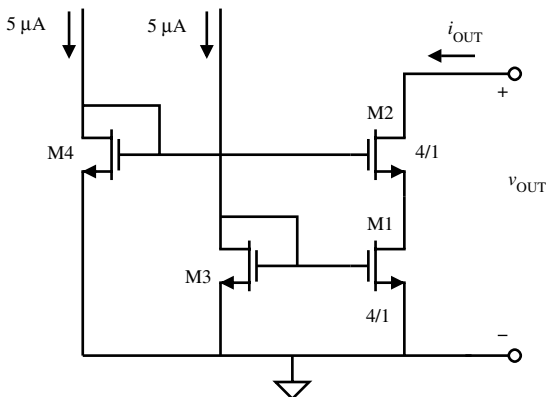


Figure P4.3-6

- 4.3-7. For the circuit shown in Fig. P4.3-7, determine  $i_{OUT}$  by simulating it using SPICE LEVEL 3 model (Table 3.4-1). Use Table 3.1-2 for device model information. Compare the results with the SPICE results from Problem 4.3-5.

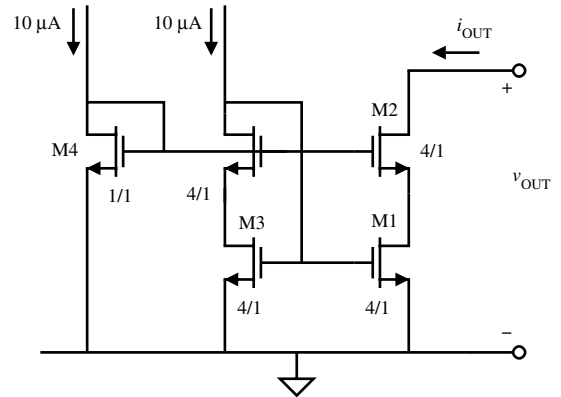


Figure P4.3-7

- 4.3-8. Use Blackman's formula to calculate the small-signal output resistance of the stacked MOSFET configuration of Fig. P4.3-8 having identical drain-source drops for both transistors. Express your answer in terms of all the pertinent small-signal parameters and then simplify your answer if  $g_m > g_{ds} > (1/R)$ . Assume the MOSFETs are identical and ignore body effects.

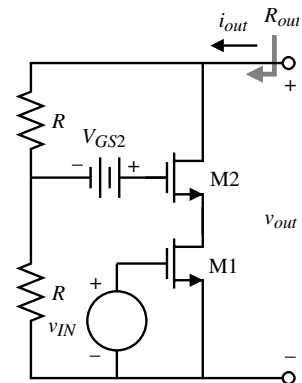


Figure P4.3-8

- 4.3-9. In Fig. P4.3-9, transistors M1 through M4 are identical and have a  $W/L$  ratio of  $10\ \mu\text{m}/1\ \mu\text{m}$ . What value of  $V_{Bias}$  gives the smallest value of  $V_{min}$  and what is this value of  $V_{min}$  assuming all transistors are in saturation? What is the small-signal output resistance in ohms?

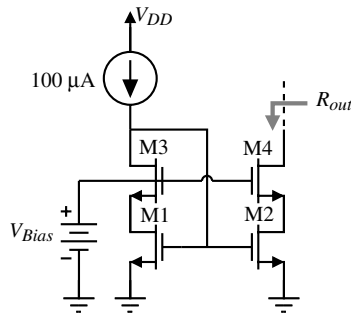


Figure P4.3-9

#### 4.3-10 (Design Problem—this problem is self-grading.)

You are to design a CMOS current sink using only a 5 volt power supply and MOSFETs that will attempt to minimize  $V_{MIN}$  and  $\Delta I_o$  as illustrated in Fig. P4.3-10.  $\Delta I_o$  is arithmetically centered about  $I_o$ , which is 100  $\mu\text{A}$ . Ignore bulk effects and keep  $W/L$ s between 1 and 100.

$V_{MIN}$  and  $\Delta I_o$  are determined by simulating your design once using the HIGH model parameters and once using the LOW model parameters. These two simulations are to be plotted as  $i_{OUT}$  versus  $v_{OUT}$  plot as shown in Fig. P4.3-10.  $V_{MIN}$  is the distance horizontally from the origin to where the line AB intersects the right-most IV curve.  $\Delta I_o$  is the biggest vertical difference between the two simulations from  $V_{MIN}$  to 5 V.

The PMOS and NMOS model parameters are as follows:

| Model      | $K'$ [ $\mu\text{A}/\text{V}^2$ ] | $V_{T0}$ [ $\text{V}^2$ ] | $\lambda$ [ $\text{V}^{-1}$ ] |
|------------|-----------------------------------|---------------------------|-------------------------------|
| NMOS(HIGH) | 26.4                              | 0.675                     | 0.011                         |
| NMOS(LOW)  | 21.6                              | 0.825                     | 0.009                         |
| PMOS(HIGH) | 8.8                               | -0.675                    | 0.022                         |
| PMOS(LOW)  | 7.2                               | -0.875                    | 0.018                         |

Your score for this problem will be determined by the following formula:

$$\text{SCORE} = \min\left(\frac{5}{V_{MIN}}, 25\right) + \min\left(\frac{500 \mu\text{A}}{\Delta I_o}, 25\right)$$

The values entered in the above SCORE formula must come from computer simulation. You are responsible for providing the supporting information necessary to verify how you determined the value of your SCORE along with a labeled schematic of your circuit.

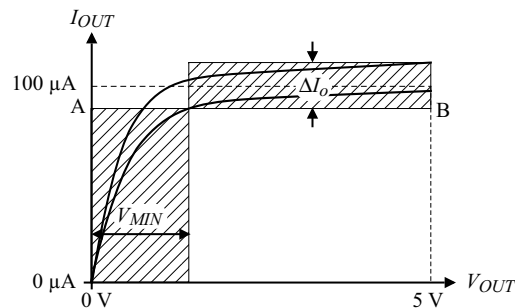


Figure P4.3-10

#### 4.4-1.

Consider the simple current mirror illustrated in Fig. P4.4-1. Over the process, the absolute variations of physical parameters are as follows:

|                  |                    |
|------------------|--------------------|
| Width variation  | $\pm 5\%$          |
| Length variation | $\pm 5\%$          |
| $K'$ variation   | $\pm 5\%$          |
| $V_T$ variation  | $\pm 5 \text{ mV}$ |

Assuming that the drain voltages are identical, what is the minimum and maximum output current measured over the process variations given above?

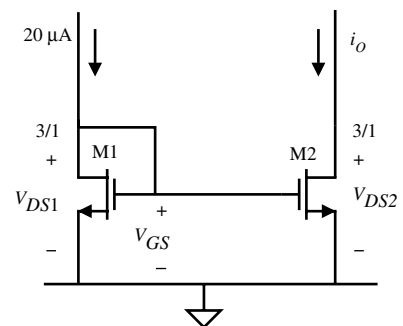


Figure P4.4-1

#### 4.4-2.

Find  $R_{in}$ ,  $i_{out}/i_{in}$ ,  $R_{out}$ ,  $V_{MIN}(\text{in})$ , and  $V_{MIN}(\text{out})$  for the current mirror shown in Figure P4.4-2.  $R$  has the value of 12.9 k $\Omega$ .

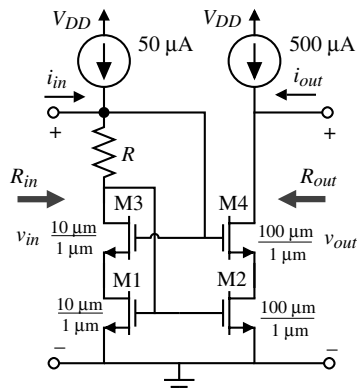


**Figure P4.4-2**

- 4.4-3.** The capacitances of M1 and M2 in Fig. P4.4-3 are  $C_{gs1} = C_{gs2} = 200$  fF,  $C_{gd1} = C_{gd2} = 50$  fF, and  $C_{bd1} = C_{bd2} = 100$  fF. Find the low-frequency current gain,  $i_{out}/i_{in}$ , the input resistance seen by  $i_{in}$ , the output resistance looking into the drain of M2, and the  $-3$ dB frequency in Hz.

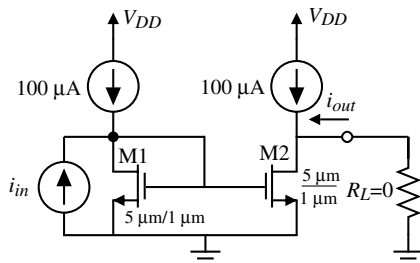


**Figure P4.4-3**

- 4.4-4.** An NMOS 1:1 current mirror layout is shown in Fig. P4.4-4. Assume both transistors are in saturation and that  $V_{DS1} = V_{DS2}$ . (a) If  $I_{in} = 100 \mu\text{A}$ , the value of  $I_{out}$  should be  $100 \mu\text{A}$ . Due to the layout, find the actual value of  $I_{out}$ . Assume that  $K' = 110 \mu\text{A/V}^2$ ,  $V_T = 0.7 \text{ V}$ , the various sheet resistances are

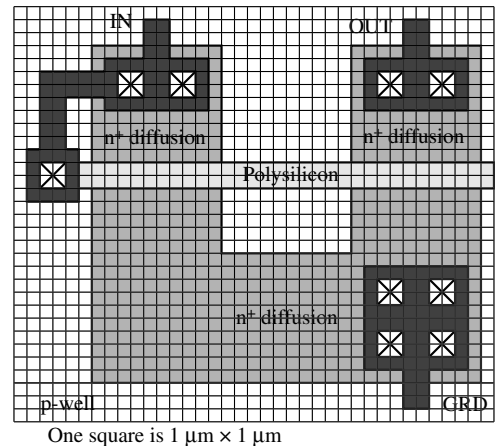


**Figure P4.4-4**

- 4.4-5** Four different layouts for an NMOS 1:2 current mirror are shown in Fig. P4.4-5. (a) Show how to connect the  $n^+$  regions and the poly regions to form the current in mirror in each layout. Label the IN, OUT, and GRD nodes. (Just draw a line from the region to wherever to indicate the connection.) (b) Which of the four layouts has the most accurate current gain? Why? (c) Which of the four layouts has the least accurate current gain from physical parasitic considerations? Why?
- 4.5-1.** Show that the sensitivity of the reference circuit shown in Fig. 4.5-2(b) is unity.
- 4.5-2.** Fig. P4.5-2 illustrates a reference circuit that provides an interesting reference voltage output. Derive an expression for  $V_{\text{REF}}$ .

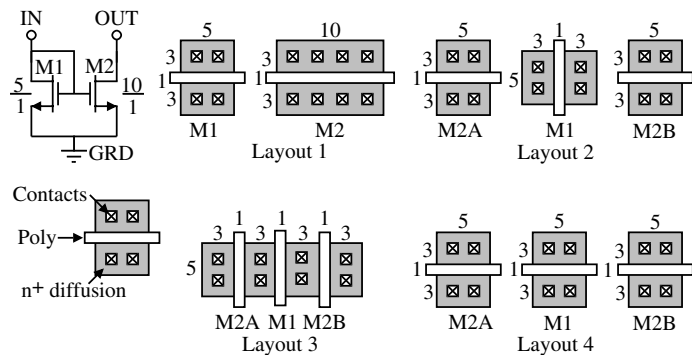


Figure P4.4-5

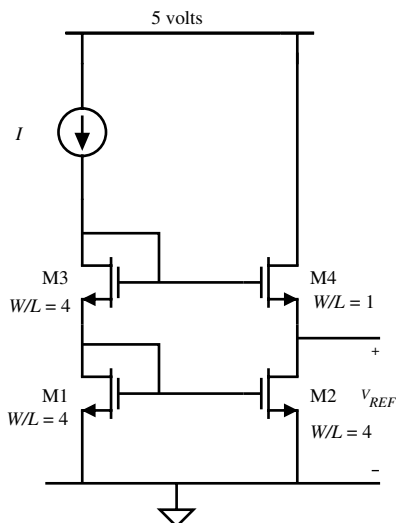


Figure P4.5-2

- 4.5-3. Figure P4.5-3 illustrates a current reference. The  $W/L$  of M1 and M2 is 100/1. The resistor is made from n-well and its nominal value is 500 k $\Omega$  at 25  $^{\circ}\text{C}$ . Using Table 3.1-2 and an n-well resistor with a sheet resistivity of 1 k $\Omega/\square \pm 40\%$  and temperature coefficient of 8000 ppm/ $^{\circ}\text{C}$ , calculate the total variation of output current seen over

process, temperature of 0–70  $^{\circ}\text{C}$ , and supply voltage variation of  $\pm 10\%$ . Assume that the temperature coefficient of the threshold voltage is  $-2.3 \text{ mV}/^{\circ}\text{C}$ .

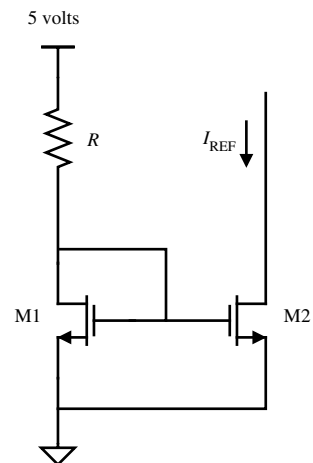


Figure P4.5-3

- 4.5-4. Figure P4.5-4 illustrates a current reference circuit. Assume that M3 and M4 are identical in size. The sizes of M1 and M2 are different. Derive an expression for the output current  $I_{\text{out}}$ .



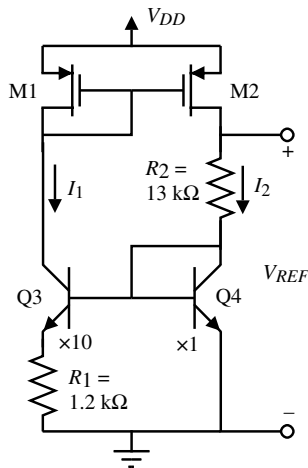


Figure P4.6-3

- 4.6-4. Assuming all transistors in Fig. 4.6-8(b) are matched,  $n = 10$ , and  $R_1 = 10 \text{ k}\Omega$ , find the value of  $V_{REF}$  at room temperature and  $R_2$  if  $V_{DD}$  is large enough that all transistors are operating in the saturation region.

4.6-5. If the bipolar transistors in Fig. 4.6-8(a) are true vertical bipolar transistors (can be isolated from the substrate), compare the current flowing in the substrate between the two series, temperature-independent references in Fig. 4.6-8.

4.6-6. Repeat Example 4.6-2 if  $V_{REF} = 1 \text{ V}$ .

4.6-7. Repeat Example 4.6-4 and assume that all resistors (except  $R_3$ ) are equal. Assume all  $W/L$  ratios of all transistors are 10 except for M23, M24, and M25 (diodes have the areas indicated in Fig. 4.6-14). Design the  $W/L$  ratios of M23, M24, and M25 and the value of the equal resistors.

4.6-8. Extend Example 4.6-3 to the design of a temperature-independent current based on the circuit shown in Fig. P4.6-8. The temperature coefficient of the resistor  $R_4$  is  $+1500 \text{ ppm}/^\circ\text{C}$ .

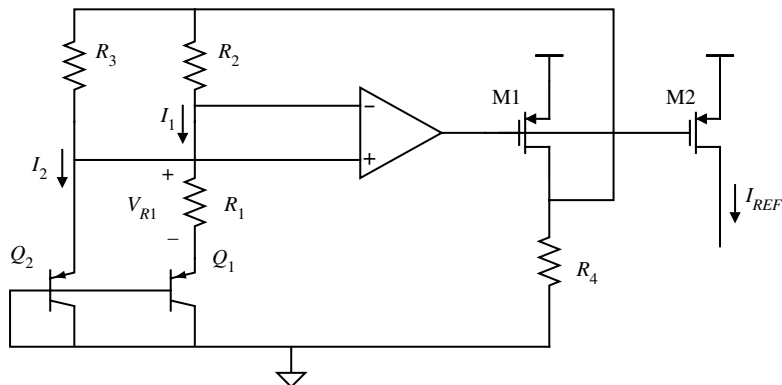


Figure P4.6-8 A temperature-independent reference current.

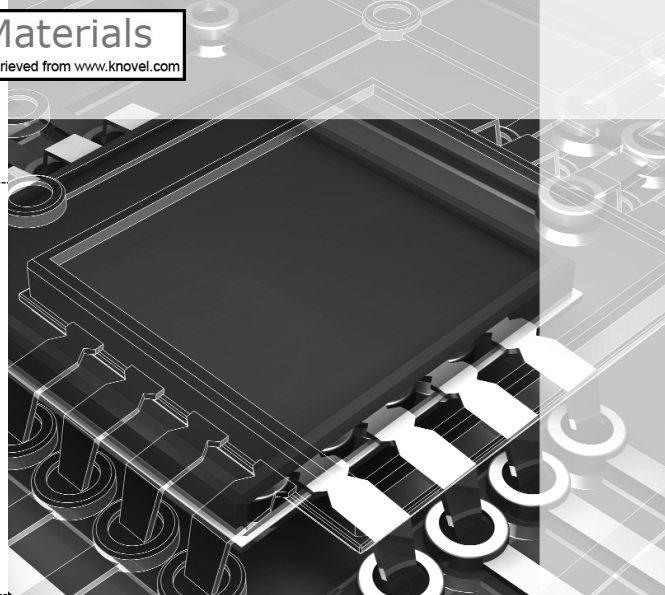
## References

1. K. Martin, "Improved Circuits for the Realization of Switched-Capacitor Filters," *IEEE Trans. Circuits and Systems*, Vol. CAS-27, No. 4, pp. 237–244, Apr. 1980.
2. B. J. Sheu and Chenming Hu, "Switch-Induced Error Voltage on a Switched Capacitor," *IEEE J. Solid-State Circuits*, Vol. SC-19, No. 4, pp. 519–525, Aug. 1984.
3. T. C. Choi, R. T. Kaneshiro, R. W. Brodersen, P. R. Gray, W. B. Jett, and M. Wilcox, "High-Frequency CMOS Switched-Capacitor Filters for Communications Applications," *IEEE J. Solid-State Circuits*, Vol. SC-18, No. 6, pp. 652–664, Dec. 1983.

4. E. J. Swanson, "Compound Current Mirror," U.S. Patent #4,477,782.
5. Todd L. Brooks and Mathew A. Rybicki, "Self-Biased Cascode Current Mirror Having High Voltage Swing and Low Power Consumption," U.S. Patent #5,359,296.
6. R. B. Blackman, "Effect of Feedback on Impedance," *Bell Sys. Tech. J.*, Vol. 23, pp. 269–277, Oct. 1943.
7. P. R. Gray, P. J. Hurst, S. H. Lewis, and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 4th ed. New York: John Wiley & Sons, 2001, pp. 607–612.
8. G. R. Wilson, "A Monolithic Junction FET-npn Operational Amplifier," *IEEE J. Solid-State Circuits*, Vol. SC-3, No. 5, pp. 341–348, Dec. 1968.
9. E. Sackinger and W. Guggenbuhl, "A Versatile Building Block: The CMOS Differential Difference Amplifier," *IEEE J. Solid-State Circuits*, Vol. SC-22, No. 2, pp. 287–294, Apr. 1987.
10. R. J. Widlar, "New Developments in IC Voltage Regulators," *IEEE J. Solid-State Circuits*, Vol. SC-6, No. 1, pp. 2–7, Feb. 1971.
11. K. E. Kujik, "A Precision Reference Voltage Source," *IEEE J. Solid-State Circuits*, Vol. SC-8, No. 3, pp. 222–226, June 1973.
12. B. S. Song and P. R. Gray, "A Precision Curvature-Corrected CMOS Bandgap Reference," *IEEE J. Solid-State Circuits*, Vol. SC-18, No. 6, pp. 634–643, Dec. 1983.
13. Y. P. Tsividis and R. W. Ulmer, "A CMOS Voltage Reference," *IEEE J. Solid-State Circuits*, Vol. SC-13, No. 6, pp. 774–778, Dec. 1982.
14. G. Tzanateas, C. A. T. Salama, and Y. P. Tsividis, "A CMOS Bandgap Voltage Reference," *IEEE J. Solid-State Circuits*, Vol. SC-14, No. 3, pp. 655–657, June 1979.
15. I. M. Gunawan, G. C. M. Meijer, J. Fonderie, and J. H. Huijsing, "A Curvature-Corrected Low-Voltage Bandgap Reference," *IEEE J. Solid-State Circuits*, Vol. SC-28, No. 6, pp. 667–670, June 1993.
16. G. A. Rincon-Mora and P. E. Allen, "A 1.1-V Current-Mode and Piecewise-Linear Curvature-Corrected Bandgap Reference," *IEEE J. Solid-State Circuits*, Vol. 33, No. 10, pp. 1551–1554, Oct. 1998.
17. B. Gilbert, "Monolithic Voltage and Current References: Themes and Variations." In: *Analog Circuit Design*, J. H. Huijsing, R. J. van de Plassche, and W. M. C. Sansen, eds. Boston, MA: Kluwer Academic Publishers, 1996, pp. 2269–2352.
18. B. S. Song and P. R. Gray, "A Precision Curvature-Corrected CMOS Bandgap Reference," *IEEE J. Solid-State Circuits*, Vol. SC-18, No. 6, pp. 634–643, Dec. 1983.
19. I. Lee et al., "Exponential Curvature-Compensated BiCMOS Bandgap References," *IEEE J. Solid-State Circuits*, Vol. SC-29, No. 11, pp. 1396–1403, Nov. 1994.
20. G. M. Meijer et al., "A New Curvature-Corrected Bandgap Reference," *IEEE J. Solid-State Circuits*, Vol. SC-17, No. 11, pp. 1139–1143, Dec. 1982.

## CHAPTER 8

# Comparators



In the previous two chapters, we have focused on the op amp and its design. Before considering how the op amp is used to accomplish various types of analog signal processing and analog–digital conversion, we will examine the comparator. The comparator is a circuit that compares an analog signal with another analog signal or reference and outputs a binary signal based on the comparison. What is meant here by an analog signal is one that can have any of a continuum of amplitude values at a given point in time (see Section 1.1). In the strictest sense a binary signal can have only one of two given values at any point in time, but this concept of a binary signal is too ideal for real-world situations, where there is a transition region between the two binary states. It is important for the comparator to pass quickly through the transition region.

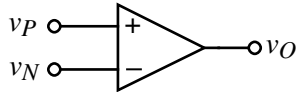
The comparator is widely used in the process of converting analog signals to digital signals. In the analog-to-digital conversion process, it is necessary to first sample the input. This sampled signal is then applied to a combination of comparators to determine the digital equivalent of the analog signal. In its simplest form, the comparator can be considered as a 1-bit analog–digital converter. The application of the comparator in analog-to-digital converters will be presented in Chapter 9.

The presentation on comparators will first examine the requirements and characterization of comparators. It will be seen that comparators can be divided into open-loop and regenerative comparators. The open-loop comparators are basically op amps without compensation. Regenerative comparators use positive feedback, similar to sense amplifiers or flip-flops, to accomplish the comparison of the magnitude between two signals. A third type of comparator emerges that is a combination of the open-loop and regenerative comparators. This combination results in comparators that are extremely fast.



### 8.1 Characterization of a Comparator

Figure 8.1-1 shows the circuit symbol for the comparator that will be used throughout this book. This symbol is identical to that for an operational amplifier, because a comparator has many of the same characteristics as a high-gain amplifier. A positive voltage applied at the  $v_P$

**Figure 8.1-1** Circuit symbol for a comparator.

input will cause the comparator output to go positive, whereas a positive voltage applied at the  $v_N$  input will cause the comparator output to go negative. The upper and lower voltage limits of the comparator output are defined as  $V_{OH}$  and  $V_{OL}$ , respectively.

## Static Characteristics

A comparator was defined above as a circuit that has a binary output whose value is based on a comparison of two analog inputs. This is illustrated in Fig. 8.1-2. As shown in this figure, the output of the comparator is high ( $V_{OH}$ ) when the difference between the noninverting and inverting inputs is positive, and low ( $V_{OL}$ ) when this difference is negative. Even though this type of behavior is impossible in a real-world situation, it can be modeled with ideal circuit elements with mathematical descriptions. One such circuit model is shown in Fig. 8.1-3. It comprises a voltage-controlled voltage source (VCVS) whose characteristics are described by the mathematical formulation given on the figure.

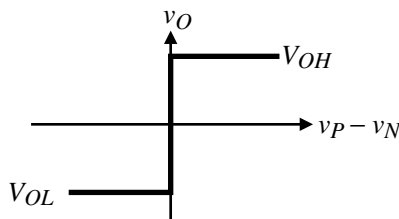
The ideal aspect of this model is the way in which the output makes a transition between  $V_{OL}$  and  $V_{OH}$ . The output changes states for an input change of  $\Delta V$ , where  $\Delta V$  approaches zero. This implies a gain of infinity, as shown below.

$$\text{Gain} = A_v = \lim_{\Delta V \rightarrow 0} \frac{V_{OH} - V_{OL}}{\Delta V} \quad (8.1-1)$$

Figure 8.1-4 shows the dc transfer curve of a first-order model that is an approximation to a realizable comparator circuit. The difference between this model and the previous one is the gain, which can be expressed as

$$A_v = \frac{V_{OH} - V_{OL}}{V_{IH} - V_{IL}} \quad (8.1-2)$$

where  $V_{IH}$  and  $V_{IL}$  represent the input-voltage difference  $v_P - v_N$  needed to just saturate the output at its upper and lower limit, respectively. This input change is called the *resolution* of the comparator. Gain is a very important characteristic describing comparator operation, for it defines the minimum amount of input change (resolution) necessary to make the output swing between the two binary states. These two output states are usually defined by the input requirements of the

**Figure 8.1-2** Ideal transfer curve of a comparator.

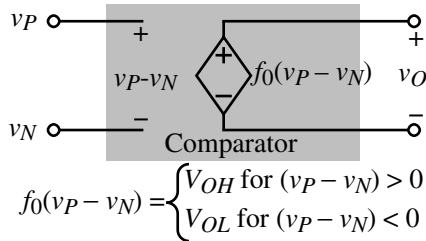


Figure 8.1-3 Model for an ideal comparator.

digital circuitry driven by the comparator output. The voltages  $V_{OH}$  and  $V_{OL}$  must be adequate to meet the  $V_{IH}$  and  $V_{IL}$  requirements of the following digital stage. For CMOS technology, these values are usually 70% and 30%, respectively, of the rail-to-rail supply voltage.

The transfer curve of Fig. 8.1-4 is modeled by the circuit of Fig. 8.1-5. This model looks similar to the model of Fig. 8.1-3, the only difference being the functions  $f_1$  and  $f_0$ .

The second nonideal effect seen in comparator circuits is input-offset voltage,  $V_{OS}$ . In Fig. 8.1-2 the output changes as the input difference crosses zero. If the output did not change until the input difference reached a value  $+V_{OS}$ , then this difference would be defined as the offset voltage. This would not be a problem if the offset could be predicted, but it varies randomly from circuit to circuit [1] for a given design. Figure 8.1-6 illustrates offset in the transfer curve for a comparator, with the circuit model including an offset generator shown in Fig. 8.1-7. The  $\pm$  sign of the offset voltage accounts for the fact that  $V_{OS}$  is unknown in polarity.

In addition to the above characteristics, the comparator can have a differential-input resistance and capacitance and an output resistance. In addition, there will also be an input common-mode resistance,  $R_{icm}$ . All these aspects can be modeled in the same manner as was done for the op amp in Section 6.1. Because the input to the comparator is usually differential, the input common-mode range is also important. The ICMR for a comparator would be that range of input common-mode voltage over which the comparator functions normally. This input common-mode range is generally the range where all transistors of the comparator remain in saturation. Even though the comparator is not designed to operate in the transition region between the two binary output states, noise is still important to the comparator. The noise of a comparator is modeled as if the comparator were biased in the transition region of the voltage-transfer characteristics. The noise will lead to an uncertainty in the transition region as shown in Fig. 8.1-8. The uncertainty in the transition region will lead to jitter or phase noise in the circuits where the comparator is employed.

## Dynamic Characteristics

The dynamic characteristics of the comparator include both small-signal and large-signal behavior. We do not know, at this point, how long it takes for the comparator to respond to

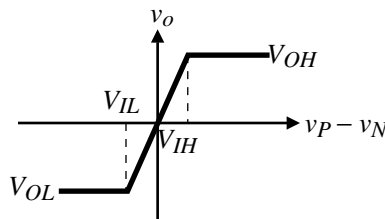


Figure 8.1-4 Transfer curve of a comparator with finite gain.

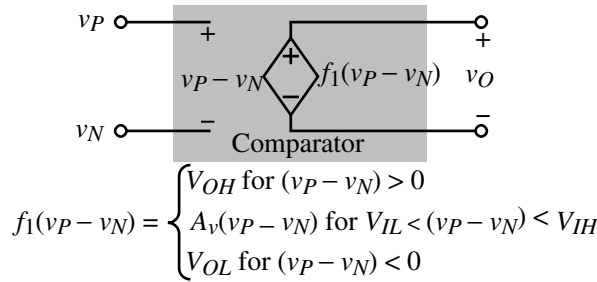


**Figure 8.1-5** Model for a comparator with finite gain.

the given differential input. The characteristic delay between input excitation and output transition is the time response of the comparator. Figure 8.1-9 illustrates the response of a comparator to an input as a function of time. Note that there is a delay between the input excitation and the output response. This time difference is called the *propagation delay time* of the comparator. It is a very important parameter since it is often the speed limitation in the conversion rate of an A/D converter. The propagation delay time in comparators generally varies as a function of the amplitude of the input. A larger input will result in a smaller delay time. There is an upper limit at which a further increase in the input voltage will no longer affect the delay. This mode of operation is called *slewing* or *slew rate*.

The small-signal dynamics are characterized by the frequency response of the comparator. A simple model of this behavior assumes that the differential voltage gain,  $A_v$ , is given as

$$A_v(s) = \frac{A_v(0)}{\frac{s}{\omega_c} + 1} = \frac{A_v(0)}{s\tau_c + 1} \quad (8.1-3)$$

where  $A_v(0)$  is the dc gain of the comparator and  $\omega_c = 1/\tau_c$  is the  $-3$  dB frequency of the single (dominant) pole approximation to the comparator frequency response. Normally, the  $A_v(0)$  and  $\omega_c$  of the comparator are smaller and larger, respectively, than for an op amp.

Let us assume that the minimum change of voltage at the input of the comparator is equal to the resolution of the comparator. We will define this minimum input voltage to the comparator as

$$V_{in(min)} = \frac{V_{OH} - V_{OL}}{A_v(0)} \quad (8.1-4)$$

For a step input voltage, the output of the comparator modeled by Eq. (8.1-3) rises (or falls) with a first-order exponential time response from  $V_{OL}$  to  $V_{OH}$  (or  $V_{OH}$  to  $V_{OL}$ ) as shown in

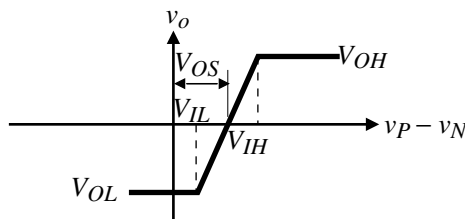


**Figure 8.1-6** Transfer curve of a comparator including input-offset voltage.

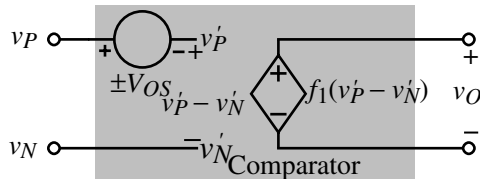


**Figure 8.1-7** Model for a comparator including input-offset voltage.

Fig. 8.1-10. If  $V_{in}$  is larger than  $V_{in(min)}$ , the output rise or fall time is faster. When  $V_{in(min)}$  is applied to the comparator, we can write the following equation:

$$\frac{V_{OH} - V_{OL}}{2} = A_v(0) [1 - e^{-t_p/\tau_c}] V_{in(min)} = A_v(0) [1 - e^{-t_p/\tau_c}] \left( \frac{V_{OH} - V_{OL}}{A_v(0)} \right) \quad (8.1-5)$$

Therefore, the propagation delay time for an input step of  $v_{in(min)}$  can be expressed as

$$t_p(max) = \tau_c \ln(2) = 0.693\tau_c \quad (8.1-6)$$

This propagation delay time will be valid for either positive-going or negative-going comparator outputs. Note in Fig. 8.1-10 that if the applied input is  $k$  times greater than  $V_{in(min)}$ , then the propagation delay time is given as

$$t_p = \tau_c \ln\left(\frac{2k}{2k - 1}\right) \quad (8.1-7)$$

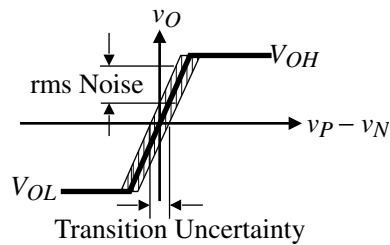


**Figure 8.1-8** Influence of noise on a comparator.

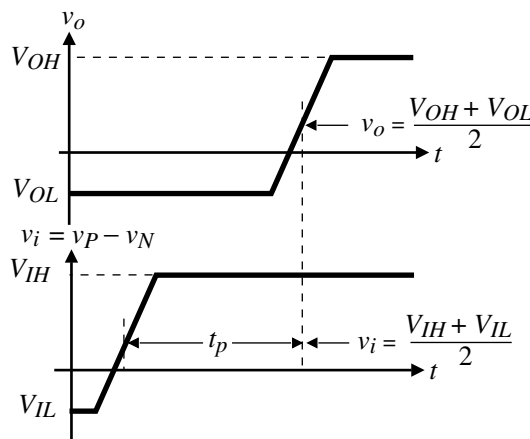


**Figure 8.1-9** Propagation delay time of a noninverting comparator.

where

$$k = \frac{V_{in}}{V_{in(min)}} \quad (8.1-8)$$

Obviously, the more overdrive applied to the input of this comparator, the smaller the propagation delay time.

As the overdrive increases to the comparator eventually the comparator enters a large-signal mode of operation. Under large-signal operation, a slew-rate limit will occur due to limited current to charge or discharge capacitors. If the propagation delay time is determined by the slew-rate of the comparator, then this time can be written as

$$t_p = \Delta T = \frac{\Delta V}{SR} = \frac{V_{OH} - V_{OL}}{2 \cdot SR} \quad (8.1-9)$$

In the case where the propagation time is determined by the slew rate, the most important factor to decrease the propagation time is increasing the sinking or sourcing capability of the comparator.

### Example 8.1-1

#### Propagation Delay Time of a Comparator

Find the propagation delay time of an open-loop comparator that has a dominant pole at  $10^3$  rad/s, a dc gain of  $10^4$ , a slew rate of  $1 \text{ V}/\mu\text{s}$ , and a binary output voltage swing of  $1 \text{ V}$ . Assume the applied input voltage is  $10 \text{ mV}$ .

#### SOLUTION

The input resolution for this comparator is  $1 \text{ V}/10^4$  or  $0.1 \text{ mV}$ . Therefore, the  $10 \text{ mV}$  input is 100 times larger than  $v_{in(min)}$ , giving a  $k$  of 100. From Eq. (8.1-7) we get

$$t_p = \frac{1}{10^3} \ln\left(\frac{2 \cdot 100}{2 \cdot 100 - 1}\right) = 10^{-3} \ln\left(\frac{200}{199}\right) = 5.01 \mu\text{s}$$

From Eq. (8.1-9) we get

$$t_p = \frac{1}{2 \cdot 1 \times 10^6} = 0.5 \mu\text{s}$$

Therefore, the propagation delay time for this case is the larger or  $5.01 \mu\text{s}$ .

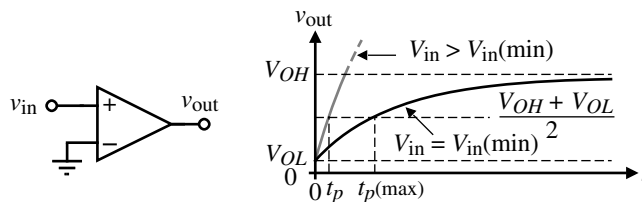


**Figure 8.1-10** Small-signal transient response for a comparator.



The comparator of Fig. 8.2-1 has two poles of interest. One is the output pole of the first stage,  $p_1$ , and the second is the output pole of the second stage,  $p_2$ . These poles are expressed as

$$p_1 = \frac{-(g_{ds2} + g_{ds4})}{C_I} \quad (8.2-4a)$$

and

$$p_2 = \frac{-(g_{ds6} + g_{ds7})}{C_{II}} \quad (8.2-4b)$$

where  $C_I$  is the sum of the capacitances connected to the output of the first stage and  $C_{II}$  is the sum of the capacitances connected to the output of the second stage.  $C_{II}$  will generally be dominated by  $C_L$ . The frequency response of the two-stage comparator can be expressed using the above results as

$$A_v(s) = \frac{A_v(0)}{\left(\frac{s}{p_1} - 1\right)\left(\frac{s}{p_2} - 1\right)} \quad (8.2-5)$$

The following example illustrates the practical values for the two-stage, open-loop comparator.

### Example 8.2-1

#### Performance of a Two-Stage Comparator

Evaluate  $V_{OH}$ ,  $V_{OL}$ ,  $A_v(0)$ ,  $V_{in}(\min)$ ,  $p_1$ , and  $p_2$  for the two-stage comparator shown in Fig. 8.2-1. Assume that this comparator is the circuit of Example 6.3-1 with no compensation capacitor,  $C_c$ , and the minimum value of  $V_{G6} = 0$  V. Also, assume that  $C_I = 0.2$  pF and  $C_{II} = 5$  pF.

#### SOLUTION

Using Eq. (8.2-1), we find that

$$V_{OH} = 2.5 - (2.5 - 0 - 0.7) \left[ 1 - \sqrt{1 - \frac{2 \cdot 95 \times 10^{-6}}{50 \times 10^{-6} \cdot 14(2.5 - 0 - 0.7)^2}} \right] = 2.42 \text{ V}$$

The value of  $V_{OL}$  is found from Eq. (8.2-2) and is  $-2.5$  V. Equation (8.2-3) was evaluated in Example 6.3-1 as  $A_v(0) = 7696$ . Thus, from Eq. (8.1-2) we find the input resolution as

$$V_{in}(\min) = \frac{V_{OH} - V_{OL}}{A_v(0)} = \frac{4.92 \text{ V}}{7696} = 0.640 \text{ mV}$$

Next, we find the poles of the comparator,  $p_1$  and  $p_2$ . From Example 6.3-1 we find that

$$p_1 = -\frac{g_{ds2} + g_{ds4}}{C_I} = -\frac{15 \times 10^{-6}(0.04 + 0.05)}{0.2 \times 10^{-12}} = -6.75 \times 10^6 \text{ (1.074 MHz)}$$

and

$$p_2 = -\frac{g_{ds6} + g_{ds7}}{C_{II}} = -\frac{95 \times 10^{-6}(0.04 + 0.05)}{5 \times 10^{-12}} = -1.71 \times 10^6 \text{ (0.670 MHz)}$$

The response of a two-stage, open-loop comparator having two poles to a step input of  $V_{in}$  is given as

$$v_{out}(t) = A_v(0)V_{in} \left[ 1 + \frac{p_2 e^{-tp_1}}{p_1 - p_2} - \frac{p_1 e^{-tp_2}}{p_1 - p_2} \right] \quad (8.2-6)$$

where  $p_1 \neq p_2$ . Equation (8.2-6) is valid for the output unless the rate of rise or fall of the output exceeds the slew rate of the output of the comparator. The slew rate of the output is similar to a Class A inverter and the negative slew rate is

$$SR^- = \frac{I_7}{C_{II}} \quad (8.2-7)$$

The positive slew rate is determined by the current sourced by M6. This slew rate is expressed as

$$SR^+ = \frac{I_6 - I_7}{C_{II}} = \frac{\beta_6(V_{DD} - V_{G6}(\min) - |V_{TP}|)^2 - I_7}{C_{II}} \quad (8.2-8)$$

If the rate of rise or fall of Eq. (8.2-6) exceeds the positive or negative slew rate, then the output response is simply a ramp whose slope is given by Eq. (8.2-7) or (8.2-8).

Assuming that slew does not occur, the step response of the two-pole comparator of Eq. (8.2-6) can be plotted by normalizing the amplitude and time. The result is

$$v'_{out}(t_n) = \frac{v_{out}(t_n)}{A_v(0)V_{in}} = 1 - \frac{m}{m-1} e^{-t_n} + \frac{1}{m-1} e^{-mt_n} \quad (8.2-9)$$

where

$$m = \frac{p_2}{p_1} \neq 1 \quad (8.2-10)$$

and

$$t_n = -tp_1 \quad (8.2-11)$$

If  $m = 1$ , then Eq. (8.2-9) becomes

$$v'_{\text{out}}(t_n) = 1 - e^{tp_1} + tp_1 e^{-tp_1} = 1 - e^{-t_n} - t_n e^{-t_n} \quad (8.2-12)$$

Equations (8.2-9) and (8.2-12) are shown in Fig. 8.2-2 for values of  $m$  from 0.25 to 4.

If the input step is larger than  $v_{\text{in}}(\text{min})$ , then the amplitudes of the curves in Fig. 8.2-2 become limited at  $V_{OH}$ . We note with interest that the slope at  $t = 0$  is zero. This can be seen by differentiating Eq. (8.2-9) and letting  $t = 0$ . The steepest slope of Eq. (8.2-9) occurs at

$$t_n(\text{max}) = \frac{\ln(m)}{m - 1} \quad (8.2-13)$$

which is found by differentiating Eq. (8.2-9) twice and setting the result equal to zero. The slope at  $t_n(\text{max})$  can be written as

$$\frac{dv'_{\text{out}}(t_n(\text{max}))}{dt_n} = \frac{m}{m - 1} \left[ \exp\left(\frac{-\ln(m)}{m - 1}\right) - \exp\left(-m \frac{\ln(m)}{m - 1}\right) \right] \quad (8.2-14)$$

If the slope of the linear response exceeds the slew rate, then the step response becomes slew limited. If the slew rate is close to the value of Eq. (8.2-14) it is not clear how to model the step response. One could assume a slew limited response until the slope of the linear response becomes less than the slew rate but this point is not easy to find. If the comparator is overdriven,  $V_{\text{in}} > V_{\text{in}}(\text{min})$ , then if the slew rate is smaller than Eq. (8.2-14), the slew rate can be used to predict the step response.

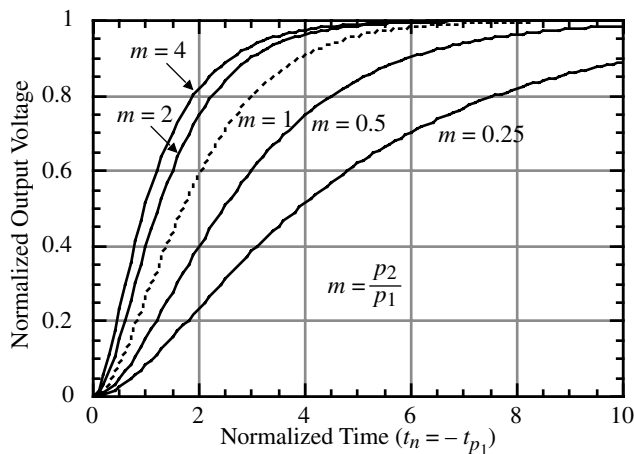


**Figure 8.2-2** Linear step response of a comparator with two real axis poles,  $p_1$  and  $p_2$ .

**Example  
8.2-2**
**Step Response of Example 8.2-1**

Find the maximum slope of Example 8.2-1 and the time at which it occurs if the magnitude of the input step is  $V_{in(min)}$ . If the dc bias current in M7 of Fig. 8.2-1 is  $100\text{ }\mu\text{A}$ , at what value of load capacitance,  $C_L$ , would the transient response become slew limited? If the magnitude of the input step is  $100V_{in(min)}$  and  $V_{OH} - V_{OL} = 1\text{V}$ , what would be the new value of  $C_L$  at which slewing would occur?

**SOLUTION**

The poles of the comparator were given in Example 8.2-1 as  $p_1 = -6.75 \times 10^6\text{ rad/s}$  and  $p_2 = -1.71 \times 10^6\text{ rad/s}$ . This gives a value of  $m = 0.253$ . From Eq. (8.2-13), the maximum slope occurs at  $t_n(max) = 1.84\text{ s}$ . Dividing by  $|p_1|$  gives  $t(max) = 0.273\text{ }\mu\text{s}$ . The slope of the transient response at this time is found from Eq. (8.2-14) as

$$\frac{dv'_{out}(t_n(max))}{dt_n} = -0.338[\exp(-1.84) - \exp(-0.253 \cdot 1.84)] = 0.159\text{ V/s}$$

Multiplying the above by  $|p_1|$  gives

$$\frac{dv'_{out}(t(max))}{dt} = 1.072\text{ V}/\mu\text{s}$$

Therefore, if the slew rate of the comparator is less than  $1.072\text{ V}/\mu\text{s}$ , the transient response will experience slewing. If the load capacitance,  $C_L$ , becomes larger than  $(100\text{ }\mu\text{A})/(1.072\text{ V}/\mu\text{s})$  or  $93.3\text{ pF}$ , the comparator will experience slewing.

If the comparator is overdriven by a factor of  $100V_{in(min)}$ , then we must unnormalize the output slope as follows:

$$\frac{dv'_{out}(t(max))}{dt} = \frac{v_{in}}{V_{in(min)}} \frac{dv'_{out}(t(max))}{dt} = 100 \cdot 1.072\text{ V}/\mu\text{s} = 107.2\text{ V/s}$$

Therefore, the comparator will now slew with a load capacitance of  $0.933\text{ pF}$ . For large overdrives, the comparator will generally experience slewing.

It is of interest to predict the propagation delay time for the two-pole comparator when slewing does not occur. To do this, we set Eq. (8.2-9) equal to  $0.5(V_{OH} + V_{OL})$  and solve for the propagation delay time,  $t_p$ . Unfortunately, this equation is not easily solved. An alternate approach is to replace the exponential terms in Eq. (8.2-9) with their power series representations. This results in

$$v_{out}(t_n) \approx A_v(0)V_{in} \left[ 1 - \frac{m}{m-1} \left( 1 - t_n + \frac{t_n^2}{2} + \dots \right) + \frac{1}{m-1} \left( 1 - mt_n + \frac{m^2 t_n^2}{2} + \dots \right) \right] \quad (8.2-15)$$

Equation (8.2-15) can be simplified to

$$v_{\text{out}}(t_n) \approx \frac{m t_n^2 A_v(0) V_{\text{in}}}{2} \quad (8.2-16)$$

Setting  $v_{\text{out}}(t_n)$  equal to  $0.5(V_{OH} + V_{OL})$  and solving for  $t_n$  gives the normalized propagation delay time,  $t_{pn}$ , as

$$t_{pn} \approx \sqrt{\frac{V_{OH} + V_{OL}}{mA_v(0)V_{\text{in}}}} = \sqrt{\frac{V_{\text{in}}(\text{min})}{mV_{\text{in}}}} = \frac{1}{\sqrt{mk}} \quad (8.2-17)$$

where  $k$  was defined in Eq. (8.1-8). This result approximates the response of Fig. 8.2-2 as a parabola. For values of  $t_n$  that are less than 1, this is a reasonable approximation. If one considers the influence of overdriving the input, then Eq. (8.2-17) is even a better approximation. The influence of overdriving is to only use the initial part of the response (i.e., it is like  $V_{OH}$  is being lowered and brought closer to zero). The following example explores the application of Eq. (8.2-17) to predict the propagation time delay for a two-pole comparator.

### Example 8.2-3

#### Propagation Delay Time of a Two-Pole Comparator That Is Not Slewing

Find the propagation delay time of the comparator of Example 8.2-1 if  $V_{\text{in}} = 10 \text{ mV}$ ,  $100 \text{ mV}$ , and  $1 \text{ V}$ .

#### SOLUTION

From Example 8.2-1 we know that  $V_{\text{in}}(\text{min}) = 0.642 \text{ mV}$  and  $m = 0.253$ . For  $V_{\text{in}} = 10 \text{ mV}$ ,  $k = 15.576$ , which gives  $t_{pn} = 0.504$ . This corresponds well with Fig. 8.2-2, where the normalized propagation time delay is the time at which the amplitude is  $1/(2k)$  or  $0.032$ . Dividing by  $|p_1|$  gives the propagation time delay as  $72.7 \text{ ns}$ . Similarly, for  $V_{\text{in}} = 100 \text{ mV}$  and  $1 \text{ V}$  we get propagation delays time of  $23.6 \text{ ns}$  and  $7.5 \text{ ns}$ , respectively.

### Initial Operating States for the Two-Stage, Open-Loop Comparator

In order to analyze the propagation time delay of a slewing, two-stage, open-loop comparator, it is necessary to first be able to find the initial operating states of the output voltages of the first and second stages. Consider the two-stage, open-loop comparator of Fig. 8.2-3. The capacitances at the outputs of the first and second stages are  $C_I$  and  $C_{II}$ , respectively.

Our approach will be to choose one of the input gates equal to a dc voltage and find the output voltage of the first and second stages when the other input gate is above and below the dc voltage on the first gate. Actually, we need to consider two cases for each of the previous possibilities. These cases are when the currents in M1 and M2 are different but neither is zero and when one of the input transistors has a current of  $I_{SS}$  and the other current is zero.

We begin by first assuming that  $v_{G2}$  is equal to a dc voltage,  $V_{G2}$ , and that  $v_{G1} > V_{G2}$  with  $i_1 < I_{SS}$  and  $i_2 > 0$ . In this case, as long as M4 remains in saturation,  $i_4 = i_3 = i_1$ , which is greater than  $i_2$ . Consequently,  $v_{o1}$  increases because of the difference current flowing into  $C_I$ .

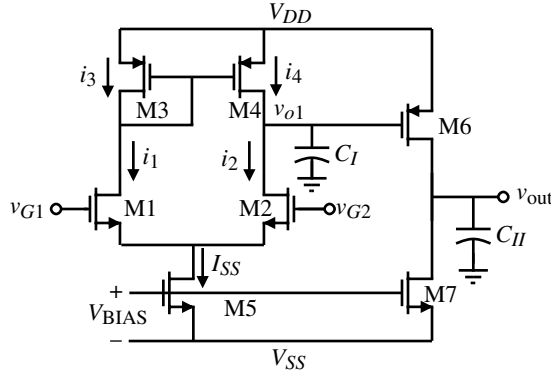


**Figure 8.2-3** Two-stage, open-loop comparator used to find initial states.

As  $v_{o1}$  continues to increase, M4 will become active and  $i_4 < i_3$ . When the voltage across the source–drain of M4 decreases to the point where  $i_4 = i_2$ , the output voltage of the first stage,  $v_{o1}$ , stabilizes. This value of voltage is

$$V_{DD} - V_{SD4}(\text{sat}) < v_{o1} < V_{DD}, \quad v_{G1} > V_{G2}, \quad i_1 < I_{SS} \quad \text{and} \quad i_2 > 0 \quad (8.2-18)$$

Under the conditions of Eq. (8.2-18), the value of  $v_{SG6} < |V_{TP}|$  and M6 will be off and the output voltage will be

$$v_{out} = V_{SS}, \quad v_{G1} > V_{G2}, \quad i_1 < I_{SS} \quad \text{and} \quad i_2 > 0 \quad (8.2-19)$$

If  $v_{G1} \gg V_{G2}$ , then  $i_1 = I_{SS}$  and  $i_2 = 0$  and  $v_{o1}$  will be at  $V_{DD}$  and  $v_{out}$  is still at  $V_{SS}$ .

Next, assume that  $v_{G2}$  is still equal to  $V_{G2}$ , but now  $v_{G1} < V_{G2}$  with  $i_1 > 0$  and  $i_2 < I_{SS}$ . In this case,  $i_4 = i_3 = i_1$  is less than  $i_2$  and  $v_{o1}$  decreases. When  $v_{o1} \leq V_{G2} - V_{TN}$ , then M2 becomes active. As  $v_{o1}$  continues to decrease,  $v_{DS2} < V_{DS2}(\text{sat})$ , the current through M2 decreases until  $i_1 = i_2 = I_{SS}/2$  at which point  $v_{o1}$  stabilizes. At this point,

$$V_{G2} - V_{GS2} < v_{o1} < V_{G2} - V_{GS2} + V_{DS2}(\text{sat}) \quad (8.2-20)$$

or

$$V_{S2} < v_{o1} < V_{S2} + V_{DS2}(\text{sat}), \quad v_{G1} < V_{G2}, \quad i_1 > 0 \quad \text{and} \quad i_2 < I_{SS} \quad (8.2-21)$$

Under the conditions of Eq. (8.2-21), the output voltage,  $v_{out}$ , will be near  $V_{DD}$  and can be found following the approach used in Example 5.1-2. If  $v_{G1} \ll V_{G2}$ , the previous results are still valid until the source voltage of M1 or M2 causes M5 to leave the saturated region. When that happens,  $I_{SS}$  decreases and  $v_{o1}$  can approach  $V_{SS}$  and  $v_{out}$  will be determined as was illustrated in Example 5.1-2.

If the gate of M1 is equal to a dc voltage of  $V_{G1}$ , we can now examine the initial states of the outputs of the first and second stages by repeating the above process. First assume that  $v_{G1} = V_{G1}$  and  $v_{G2} > V_{G1}$  with  $i_2 < I_{SS}$  and  $i_1 > 0$ . As a result,  $i_1 < i_2$  gives  $i_4 < i_2$  as long as M4 is saturated. Therefore,  $v_{o1}$  decreases because of the difference current flowing out of  $C_I$ . As  $v_{o1}$  decreases, M2 will become active and  $i_2$  will decrease to the point where  $i_1 = i_2 = I_{SS}/2$ . Therefore,  $v_{o1}$  stabilizes and has a value given as

$$V_{G1} - V_{GS2}(I_{SS}/2) < v_{o1} < V_{G1} - V_{GS2}(I_{SS}/2) + V_{DS2}(\text{sat}) \quad (8.2-22)$$

or

$$V_{S2}(I_{SS}/2) < v_{o1} < V_{S2}(I_{SS}/2) + V_{DS2}(\text{sat}), \quad v_{G2} > V_{G1}, \quad i_1 > 0 \quad \text{and} \quad i_2 < I_{SS} \quad (8.2-23)$$

Under the conditions of Eq. (8.2-23), the output voltage,  $v_{\text{out}}$ , will be near  $V_{DD}$  and can be found following the approach used in Example 5.1-2. If  $v_{G2} \gg V_{G1}$ , the previous results are still valid until the source voltage of M1 or M2 causes M5 to leave the saturated region. When that happens,  $I_{SS}$  decreases and  $v_{o1}$  can approach  $V_{SS}$  and  $v_{\text{out}}$  will be determined as was illustrated in Example 5.1-2.

Next, assume that  $v_{G1}$  is still equal to  $V_{G1}$ , but now  $v_{G2} < V_{G1}$  with  $i_1 < I_{SS}$  and  $i_2 > 0$ . As a result,  $i_1 > i_2$ , which gives  $i_4 > i_2$ , causing  $v_{o1}$  to increase. As long as M4 is saturated,  $i_4 > i_2$ . When M4 enters the active region,  $i_4$  will decrease until  $i_4 = i_2$  at which point  $v_{o1}$  stabilizes and is given as

$$V_{DD} - V_{SD4}(\text{sat}) < v_{o1} < V_{DD}, \quad v_{G2} < V_{G1}, \quad i_1 < I_{SS} \quad \text{and} \quad i_2 > 0 \quad (8.2-24)$$

Under the conditions of Eq. (8.2-24), the value of  $v_{SG6} < |V_{TP}|$  and M6 will be off and the output voltage will be

$$v_{\text{out}} = V_{SS}, \quad v_{G2} < V_{G1}, \quad i_1 < I_{SS} \quad \text{and} \quad i_2 > 0 \quad (8.2-25)$$

If  $v_{G2} \ll V_{G1}$ , then  $i_1 = I_{SS}$  and  $i_2 = 0$  and  $v_{o1}$  will be at  $V_{DD}$  and  $v_{\text{out}}$  is still at  $V_{SS}$ . The above results are summarized in Table 8.2-1.

### Propagation Delay Time of a Slewing, Two-Stage, Open-Loop Comparator

In most applications, the two-stage, open-loop comparator is overdriven to the point where the propagation delay time is determined by the slewing performance of the comparator. In this case, the propagation delay time is found by evaluating the equation

$$i_i = C_i \frac{dv_i}{dt_i} = C_i \frac{\Delta v_i}{\Delta t_i} \quad (8.2-26)$$

**Table 8.2-1 Initial Operating States of the Two-Stage, Open-Loop Comparator of Fig. 8.2-3**

| Conditions                                         | Initial State of $v_{o1}$                                            | Initial State of $v_{\text{out}}$ |
|----------------------------------------------------|----------------------------------------------------------------------|-----------------------------------|
| $v_{G1} > V_{G2}$ , $i_1 < I_{SS}$ and $i_2 > 0$   | $V_{DD} - V_{SD4}(\text{sat}) < v_{o1} < V_{DD}$                     | $V_{SS}$                          |
| $v_{G1} \gg V_{G2}$ , $i_1 = I_{SS}$ and $i_2 = 0$ | $V_{DD}$                                                             | $V_{SS}$                          |
| $v_{G1} < V_{G2}$ , $i_1 > 0$ and $i_2 < I_{SS}$   | $V_{S2} < v_{o1} < V_{S2} + V_{DS2}(\text{sat})$                     | Eq. (5.1-19) for PMOS             |
| $v_{G1} \ll V_{G2}$ , $i_1 = 0$ and $i_2 = I_{SS}$ | $V_{DD}$                                                             | Eq. (5.1-19) for PMOS             |
| $v_{G2} > V_{G1}$ , $i_1 > 0$ and $i_2 < I_{SS}$   | $V_{S2}(I_{SS}/2) < v_{o1} < V_{S2}(I_{SS}/2) + V_{DS2}(\text{sat})$ | Eq. (5.1-19) for PMOS             |
| $v_{G2} \gg V_{G1}$ , $i_1 = 0$ and $i_2 = I_{SS}$ | $V_{SS}$                                                             | Eq. (5.1-19) for PMOS             |
| $v_{G2} < V_{G1}$ , $i_1 < I_{SS}$ and $i_2 > 0$   | $V_{DD} - V_{SD4}(\text{sat}) < v_{o1} < V_{DD}$                     | $V_{SS}$                          |
| $v_{G2} \ll V_{G1}$ , $i_1 = I_{SS}$ and $i_2 = 0$ | $V_{DD}$                                                             | $V_{SS}$                          |

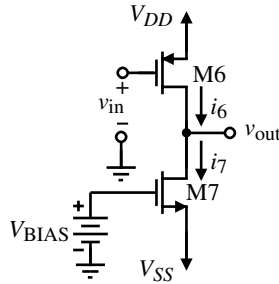


**Figure 8.2-4** Second stage of the two-stage, open-loop comparator.

where  $C_i$  is the capacitance to ground at the output of the  $i$ th stage. The propagation delay time of the  $i$ th stage is found by solving for  $\Delta t_i$  from Eq. (8.2-26) to get

$$t_i = \Delta t_i = C_i \frac{\Delta V_i}{I_i} \quad (8.2-27)$$

The propagation delay time is found by summing the delays of each stage.

The term  $\Delta V_i$  in Eq. (8.2-27) is generally equal to half of the output swing of the  $i$ th stage. In some cases, the value of  $\Delta V_i$  is determined by the threshold or trip point of the next stage. The trip point of an amplifier is the value of input that causes the output to be midway between its limits. In the two-stage, open-loop comparator, the second stage is generally a Class A inverting amplifier similar to that shown in Fig. 8.2-4. The trip point can be calculated by assuming both transistors are in saturation and equating the currents. The only unknown is the input voltage. Solving for this input voltage gives

$$v_{in} = V_{TRP} = V_{DD} - |V_{TP}| - \sqrt{\frac{K_N(W_7/L_7)}{K_P(W_6/L_6)}} (V_{BIAS} - V_{SS} - V_{TN}) \quad (8.2-28)$$

If we refer back to Fig. 5.1-5, we see that the trip point is really a range and not a specific value. However, if the slope of the inverting amplifier in the region where both transistors are saturated is steep enough, then the trip point can be considered to be a point.

#### Example 8.2-4

#### Calculation of the Trip Point of an Inverting Amplifier

Use the model parameters of Table 3.1-2 and calculate the trip point of the inverter in Fig. 5.1-5.

#### SOLUTION

It is probably easier to redevelop Eq. (8.2-28) for the inverter of Fig. 5.1-5 than to use Eq. (8.2-28). Equating the currents of M1 and M2 of Fig. 5.1-5 gives

$$V_{TRP} = V_{TN} + \sqrt{\frac{\beta_2}{\beta_1}} (V_{DD} - V_{BIAS} - |V_{TP}|) = 0.7 + \sqrt{\frac{50(2)}{110(2)}} (2.5 - 0.7) = 1.913 \text{ V}$$

We note that this value corresponds very closely with the value of input voltage at which the output of Fig. 5.1-5 is midway between power supplies.

The propagation delay time of a two-stage, open-loop comparator that is slewing can be calculated using the above ideas. The delay of the first stage is added to the delay of the second stage to get the total propagation delay time. Let us illustrate the procedure with an example.

### Example 8.2-5

### Calculation of the Propagation Delay Time of a Two-Stage, Open-Loop Comparator

For the two-stage comparator shown in Fig. 8.2-5, assume that  $C_I = 0.2$  pF and  $C_{II} = 5$  pF. Also, assume that  $v_{G1} = 0$  V and that  $v_{G2}$  has the waveform shown in Fig. 8.2-6. If the input voltage is large enough to cause slew to dominate, find the propagation delay time of the rising and falling output of the comparator and give the propagation delay time of the comparator.

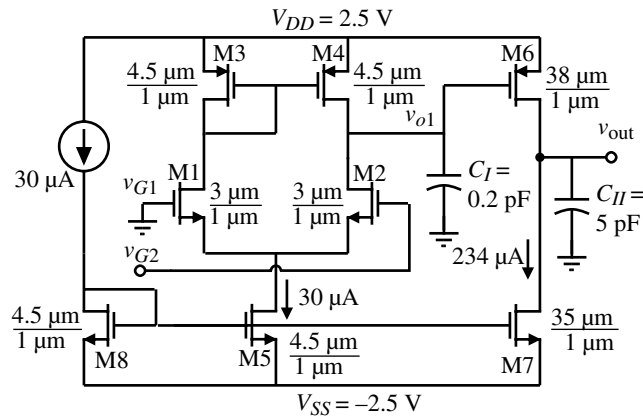


**Figure 8.2-5** Two-stage comparator for Example 8.2-5.

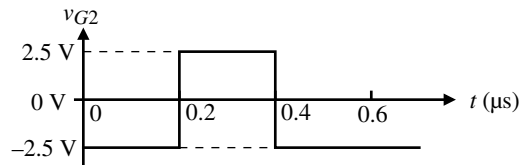


**Figure 8.2-6** Input to the comparator of Example 8.2-5.

### SOLUTION

The total delay will be given as the sum of the first- and second-stage delays,  $t_1$  and  $t_2$ , respectively. First, consider the change of  $v_{G2}$  from  $-2.5$  to  $2.5$  V at  $0.2$   $\mu$ s. From Table 8.2-1, the last row, the initial states of  $v_{o1}$  and  $v_{out}$  are  $+2.5$  and  $-2.5$  V, respectively. To find the falling delay of the first stage,  $t_{f1}$ , requires  $C_I$ ,  $\Delta V_{o1}$ , and  $I_5$ .  $C_I = 0.2$  pF,  $I_5 = 30$   $\mu$ A, and  $\Delta V_1$  can be calculated by finding the trip point of the output stage. Although we could use Eq. (8.2-8), it is easier to set the current of M6 when saturated equal to  $234$   $\mu$ A to get the trip point. Doing this we get

$$\frac{\beta_6}{2} (V_{SG6} - |V_{TP}|)^2 = 234 \mu\text{A} \rightarrow V_{SG6} = 0.7 + \sqrt{\frac{234 \cdot 2}{50 \cdot 38}} = 0.496 \text{ V}$$

Subtracting this voltage from 2.5 V gives the trip point of the second stage as

$$V_{TRP2} = 2.5 - 1.196 = 1.304 \text{ V}$$

Therefore,  $\Delta V_1 = 2.5 \text{ V} - 1.304 \text{ V} = V_{SG6} = 1.196 \text{ V}$ . Thus, the falling propagation time delay of the first stage is

$$t_{f01} = (0.2 \text{ pF}) \left( \frac{1.196 \text{ V}}{30 \text{ } \mu\text{A}} \right) = 8 \text{ ns}$$

The rising propagation time delay of the second stage requires knowledge of  $C_{II}$ ,  $\Delta V_{out}$ , and  $I_6$ .  $C_{II}$  is given as 5 pF,  $\Delta V_{out} = 2.5 \text{ V}$  (assuming the trip point of the circuit connected to the output of the comparator is 0 V), and  $I_6$  can be found as follows. When the gate of M6 is at 1.304 V, the current is 234  $\mu\text{A}$ . However, the output of the first stage will continue to fall, so what value should be used for the gate in order to calculate  $I_6$ ? The lowest value of  $V_{G6}$  is given as

$$V_{G6} = V_{G1} - V_{GS2}(I_{SS}/2) + V_{DS2} \approx -V_{GS2}(I_{SS}/2) = -0.7 - \sqrt{\frac{2 \cdot 15}{110 \cdot 3}} = -1.00 \text{ V}$$

Let us take the approximate value of  $V_{G6}$  as midway between 1.304 and  $-1.00 \text{ V}$ , which is 0.152 V. Therefore,  $V_{SG6} = 2.348 \text{ V}$  and the value of  $I_6$  is

$$I_6 = \frac{\beta_6}{2} (V_{SG6} - |V_{TP}|)^2 = \frac{38 \cdot 50}{2} (2.348 - 0.7)^2 = 2580 \text{ } \mu\text{A}$$

which is a reminder that the active transistor can generally sink or source more current than the fixed transistor in the Class A inverting stage. The rising propagation time delay for the output can be expressed as

$$t_{r,out} = (5 \text{ pF}) \left( \frac{2.5 \text{ V}}{2580 \text{ } \mu\text{A} - 234 \text{ } \mu\text{A}} \right) = 5.3 \text{ ns}$$

Thus, the total propagation time delay of the rising output of the comparator is approximately 13.3 ns and most of this delay is attributable to the first stage.

Next, consider the change of  $v_{G2}$  from 2.5 to  $-2.5 \text{ V}$ , which occurs at 0.4  $\mu\text{s}$ . We shall assume that  $v_{G2}$  has been at 2.5 V long enough for the conditions of Table 8.2-1 to be valid. Therefore,  $v_{o1} \approx V_{SS} = -2.5 \text{ V}$  and  $v_{out} \approx V_{DD}$ . Rather than use Eq. (5.1-19), we have assumed that  $v_{out}$  is approximately  $V_{DD}$ . The propagation delay times for the first and second stages are calculated as

$$t_{r01} = (0.2 \text{ pF}) \left( \frac{1.304 \text{ V} - (-1.000)}{30 \text{ } \mu\text{A}} \right) = 15.4 \text{ ns}$$

and

$$t_{f,\text{out}} = (5 \text{ pF}) \left( \frac{2.5 \text{ V}}{234 \text{ } \mu\text{A}} \right) = 53.42 \text{ ns}$$

The total propagation delay time of the falling output is 68.82 ns. Taking the average of the rising and falling propagation delays time gives a propagation delay time for this two-stage, open-loop comparator of about 41 ns. These values compare favorably with the simulation of the comparator shown in Fig. 8.2-7.

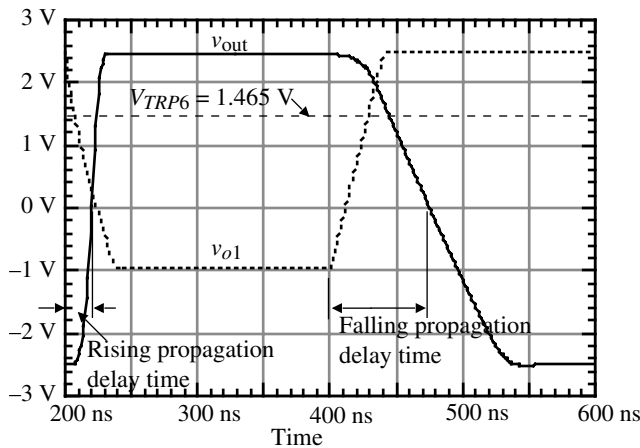


**Figure 8.2-7** Response of Fig. 8.2-3 to the input shown in Fig. 8.2-6.

The previous example can be reworked for the case where  $v_{G2}$  is held constant and  $v_{G1}$  varies (see Problem 8.2-7). If the comparator is not slewing, then the propagation time delay will be determined by the linear step response of the comparator. The propagation time delay analysis becomes more complex if the comparator slews only for part of the step response. This is a case where simulation would provide the desired details.

## Design of a Two-Stage, Open-Loop Comparator

The design of a two-stage, open-loop comparator is similar in many respects to the two-stage op amp. The primary difference is that the comparator is not compensated. The typical input specifications include the propagation time delay, the output voltage swing, the resolution, and input common-mode range. There are two approaches to the design of the comparator depending on whether it is slewing or not. If the comparator is not slewing, then the pole locations are extremely important. On the other hand, if the comparator slews, then the ability to charge and discharge capacitors becomes more important.

For the nonslewing comparator, the propagation time is reduced by a large overdrive and the pole placement. The larger the magnitude of the poles, the shorter the propagation delay time. Since the poles are due to the capacitance and resistance to ground at the output of each stage, it is important to keep the value of capacitance and resistance small. The design

procedure illustrated in Table 8.2-2 can be used to accomplish a minimum propagation delay time for the two-stage, open-loop comparator. This procedure attempts to design the pole locations to correspond to a desired propagation delay time. The procedure is simplified by assuming that  $m = 1$  or the poles are equal.

**Table 8.2-2 Design of the Two-Stage, Open-Loop Comparator of Fig. 8.2-3 for a Linear Response**

| Specifications: $t_p$ , $C_{II}$ , $V_{in}(\min)$ , $V_{OH}$ , $V_{OL}$ , $V_{icm}^+$ , $V_{icm}^-$ , and overdrive<br>Constraints: Technology, $V_{DD}$ , and $V_{SS}$ |                                                                                                                                                                                                           |                                                                                                                                          |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------|
| Step                                                                                                                                                                    | Design Relationships                                                                                                                                                                                      | Comments                                                                                                                                 |
| 1                                                                                                                                                                       | $ p_I  =  p_{II}  = \frac{1}{t_p \sqrt{mk}}$ , and $I_7 = I_6 = \frac{ p_{II}  C_{II}}{\lambda_N + \lambda_P}$                                                                                            | Choose $m = 1$                                                                                                                           |
| 2                                                                                                                                                                       | $\frac{W_6}{L_6} = \frac{2 \cdot I_6}{K'_P(V_{SD6}(\text{sat}))^2}$ and $\frac{W_7}{L_7} = \frac{2 \cdot I_7}{K'_N(V_{DS7}(\text{sat}))^2}$                                                               | $V_{SD6}(\text{sat}) = V_{DD} - V_{OH}$<br>$V_{DS7}(\text{sat}) = V_{OL} - V_{SS}$                                                       |
| 3                                                                                                                                                                       | Guess $C_I$ as 0.1–0.5 pF $\therefore I_5 = I_7 \frac{2C_I}{C_{II}}$                                                                                                                                      | A result of choosing $m = 1$ ; will check $C_I$ later                                                                                    |
| 4                                                                                                                                                                       | $\frac{W_3}{L_3} = \frac{W_4}{L_4} = \frac{I_5}{K'_P(V_{SG3} -  V_{TP} )^2}$                                                                                                                              | $V_{SG3} = V_{DD} - V_{icm}^+ + V_{TN}$                                                                                                  |
| 5                                                                                                                                                                       | $g_{m1} = \frac{A_v(0)(g_{ds2} + g_{ds4})(g_{ds6} + g_{ds7})}{g_{m6}}$                                                                                                                                    | $g_{m6} = \sqrt{\frac{2K'_P W_6 I_6}{L_6}}$                                                                                              |
|                                                                                                                                                                         | $\frac{W_1}{L_1} = \frac{W_2}{L_2} = \frac{g_{m1}^2}{K'_N I_5}$                                                                                                                                           | $A_v(0) = \frac{V_{OH} - V_{OL}}{V_{in}(\min)}$                                                                                          |
| 6                                                                                                                                                                       | Find $C_I$ and check assumption<br>$C_I = C_{gd2} + C_{gd4} + C_{gs6} + C_{bd2} + C_{bd4}$<br>If $C_I$ is greater than the guess in step 3, then increase the value of $C_I$ and repeat steps 4 through 6 | $AD_2 = W_2(L_1 + L_2 + L_3)$<br>$PD_2 = 2(W_2 + L_1 + L_2 + L_3)$<br>$AD_4 = W_4(L_1 + L_2 + L_3)$<br>$PD_4 = 2(W_4 + L_1 + L_2 + L_3)$ |
| 7                                                                                                                                                                       | $V_{DSS}(\text{sat}) = V_{icm}^- - V_{GS1} - V_{SS}$<br>$\frac{W_5}{L_5} = \frac{2 \cdot I_5}{K'_N(V_{DSS}(\text{sat}))^2}$                                                                               | If $V_{DSS}(\text{sat})$ is less than 100 mV, increase $W_1/L_1$                                                                         |

### Example 8.2-6

### Design of the Two-Stage, Open-Loop Comparator of Fig. 8.2-3 for a Linear Response

Assume the specifications of Fig. 8.2-3 are given below:

$$\begin{array}{lll}
 t_p = 50 \text{ ns} & V_{OH} = 2 \text{ V} & V_{OL} = -2 \text{ V} \\
 V_{DD} = 2.5 \text{ V} & V_{SS} = -2.5 \text{ V} & C_{II} = 5 \text{ pF} \\
 V_{in}(\min) = 1 \text{ mV} & V_{icm}^+ = 2 \text{ V} & V_{icm}^- = -1.25 \text{ V}
 \end{array}$$

Also assume that the overdrive will be a factor of 10. Design a two-stage, open-loop comparator using Fig. 8.2-3 to achieve the above specifications and assume that all channel lengths are to be 1  $\mu\text{m}$ .

## SOLUTION

Following the procedure outlined in Table 8.2-2, we choose  $m = 1$  to get

$$|p_I| = |p_{II}| = \frac{10^9}{50\sqrt{10}} = 6.32 \times 10^6 \text{ rad/s}$$

This gives

$$I_6 = I_7 = \frac{6.32 \times 10^6 \cdot 5 \times 10^{-12}}{0.04 + 0.05} = 351 \text{ } \mu\text{A} \rightarrow I_6 = I_7 = 400 \text{ } \mu\text{A}$$

Therefore,

$$\frac{W_6}{L_6} = \frac{2 \cdot 400}{(0.5)^2 \cdot 50} = 64 \quad \text{and} \quad \frac{W_7}{L_7} = \frac{2 \cdot 400}{(0.5)^2 \cdot 110} = 29$$

Next, we guess  $C_I = 0.2 \text{ pF}$ . This gives  $I_5 = 32 \text{ } \mu\text{A}$  and we will increase it to  $40 \text{ } \mu\text{A}$  for a margin of safety. Step 4 gives  $V_{SG3}$  as  $1.2 \text{ V}$ , which results in

$$\frac{W_3}{L_3} = \frac{W_4}{L_4} = \frac{40}{50(1.2 - 0.7)^2} = 3.2 \rightarrow \frac{W_3}{L_3} = \frac{W_4}{L_4} = 4$$

The desired gain is found to be 4000, which gives an input transconductance of

$$g_{m1} = \frac{4000 \cdot 0.09 \cdot 20}{44.44} = 162 \text{ } \mu\text{S}$$

This gives the  $W/L$  ratios of M1 and M2 as

$$\frac{W_1}{L_1} = \frac{W_2}{L_2} = \frac{(162)^2}{110 \cdot 40} = 5.96 \rightarrow \frac{W_1}{L_1} = \frac{W_2}{L_2} = 6$$

To check the guess for  $C_I$  we need to calculate it, which is done as

$$\begin{aligned} C_I &= C_{gd2} + C_{gd4} + C_{gs6} + C_{bd2} + C_{bd4} = 0.9 \text{ fF} + 1.3 \text{ fF} + 119.5 \text{ fF} + 20.4 \text{ fF} + 36.8 \text{ fF} \\ &= 178.9 \text{ fF} \end{aligned}$$

which is less than what was guessed so we will make no changes.

Finally, the  $W/L$  value of M5 is found by finding  $V_{GS1}$  as  $0.946 \text{ V}$ , which gives  $V_{DS5}(\text{sat}) = 0.304 \text{ V}$ . This gives

$$\frac{W_5}{L_5} = \frac{2 \cdot 40}{(0.304)^2 \cdot 110} = 7.87 \approx 8$$

Obviously, M5 and M7 cannot be connected gate–gate and source–source. The values of  $I_5$  and  $I_7$  must be derived separately as illustrated in Fig. 8.2-8. The  $W$  values are summarized below, assuming that all channel lengths are  $1 \text{ } \mu\text{m}$ .

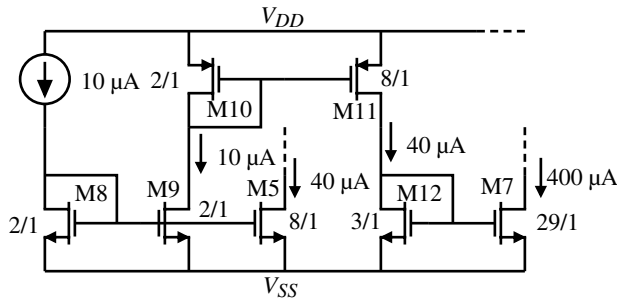


**Figure 8.2-8** Biasing scheme for the comparator of Example 8.2-6.

$$W_1 = W_2 = 6 \mu\text{m} \quad W_3 = W_4 = 4 \mu\text{m} \quad W_5 = 8 \mu\text{m}$$

$$W_6 = 64 \mu\text{m} \quad W_7 = 29 \mu\text{m}$$

The design of the slewing, two-stage, open-loop comparator has a different focus than the design suggested in Table 8.2-2 for the linear response. In this case we assume that the large-signal conditions dominate the design. A typical design procedure for the slewing case is shown in Table 8.2-3.

**Table 8.2-3** Design of the Two-Stage, Open-Loop Comparator of Fig. 8.2-3 for a Slewing Response

| Specifications: $t_p$ , $C_{II}$ , $V_{in}(\min)$ , $V_{OH}$ , $V_{OL}$ , $V_{icm}^+$ , $V_{icm}^-$<br>Constraints: Technology, $V_{DD}$ , and $V_{SS}$ |                                                                                                                                                                                                      |                                                                                                                                          |
|---------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------|
| Step                                                                                                                                                    | Design Relationships                                                                                                                                                                                 | Comments                                                                                                                                 |
| 1                                                                                                                                                       | $I_7 = I_6 = C_{II} \frac{dv_{out}}{dt} = \frac{C_{II}(V_{OH} - V_{OL})}{t_p}$                                                                                                                       | Assume the trip point of the output is $(V_{OH} - V_{OL})/2$                                                                             |
| 2                                                                                                                                                       | $\frac{W_6}{L_6} = \frac{2 \cdot I_6}{K'_P(V_{SD6}(\text{sat}))^2}$ and $\frac{W_7}{L_7} = \frac{2 \cdot I_7}{K'_N(V_{DS7}(\text{sat}))^2}$                                                          | $V_{SD6}(\text{sat}) = V_{DD} - V_{OH}$<br>$V_{DS7}(\text{sat}) = V_{OL} - V_{SS}$                                                       |
| 3                                                                                                                                                       | Guess a value of $C_I$ and check later                                                                                                                                                               | Typically $0.1 \text{ pF} < C_I < 0.5 \text{ pF}$                                                                                        |
| 4                                                                                                                                                       | $I_5 = C_I \frac{dv_{o1}}{dt} \approx \frac{C_I(V_{OH} - V_{OL})}{t_p}$                                                                                                                              | Assume that $v_{o1}$ swings between $V_{OH}$ and $V_{OL}$                                                                                |
| 5                                                                                                                                                       | $\frac{W_3}{L_3} = \frac{W_4}{L_4} = \frac{I_5}{K'_P(V_{SG3} -  V_{TP} )^2}$                                                                                                                         | $V_{SG3} = V_{DD} - V_{icm}^+ + V_{TN}$                                                                                                  |
| 6                                                                                                                                                       | $g_{m1} = \frac{A_v(0)(g_{ds2} + g_{ds4})(g_{ds6} + g_{ds7})}{g_{m6}}$<br>$\frac{W_1}{L_1} = \frac{W_2}{L_2} = \frac{g_{m1}^2}{K'_N I_5}$                                                            | $g_{m6} = \sqrt{\frac{2K'_P W_6 I_6}{L_6}}$<br>$A_v(0) = \frac{V_{OH} - V_{OL}}{V_{in}(\min)}$                                           |
| 7                                                                                                                                                       | Find $C_I$ and check assumption<br>$C_I = C_{gd2} + C_{gd4} + C_{gs6} + C_{bd2} + C_{bd4}$<br>If $C_I$ is greater than the guess in step 3, increase the value of $C_I$ and repeat steps 4 through 6 | $AD_2 = W_2(L_1 + L_2 + L_3)$<br>$PD_2 = 2(W_2 + L_1 + L_2 + L_3)$<br>$AD_4 = W_4(L_1 + L_2 + L_3)$<br>$PD_4 = 2(W_4 + L_1 + L_2 + L_3)$ |
| 8                                                                                                                                                       | $V_{DSS}(\text{sat}) = V_{icm}^- - V_{GS1} - V_{SS}$ , $\frac{W_5}{L_5} = \frac{2 \cdot I_5}{K'_N(V_{DSS}(\text{sat}))^2}$                                                                           | If $V_{DSS}(\text{sat})$ is less than 100 mV, increase $W_1/L_1$                                                                         |

**Example  
8.2-7****Design of the Two-Stage, Open-Loop Comparator  
of Fig. 8.2-3 for a Slewing Response**

Assume the specifications of Fig. 8.2-3 are given below.

$$\begin{array}{lll} t_p = 50 \text{ ns} & V_{OH} = 2 \text{ V} & V_{OL} = -2 \text{ V} \\ V_{DD} = 2.5 \text{ V} & V_{SS} = -2.5 \text{ V} & C_{II} = 5 \text{ pF} \\ V_{in}(\text{min}) = 1 \text{ mV} & V_{icm}^+ = 2 \text{ V} & V_{icm}^- = -1.25 \text{ V} \end{array}$$

Design a two-stage, open-loop comparator using the circuit of Fig. 8.2-3 to the above specifications and assume all channel lengths are to be  $1 \mu\text{m}$ .

**SOLUTION**

Following the procedure outlined in Table 8.2-3, we calculate  $I_6$  and  $I_7$  as

$$I_6 = I_7 = \frac{5 \times 10^{-12} \cdot 4}{50 \times 10^{-9}} = 400 \mu\text{A}$$

Therefore,

$$\frac{W_6}{L_6} = \frac{2 \cdot 400}{(0.5)^2 \cdot 50} = 64 \quad \text{and} \quad \frac{W_7}{L_7} = \frac{2 \cdot 400}{(0.5)^2 \cdot 110} = 29$$

Next, we guess  $C_I = 0.2 \text{ pF}$ . This gives

$$I_5 = \frac{(0.2 \text{ pF})(4 \text{ V})}{50 \text{ ns}} = 16 \mu\text{A} \rightarrow I_5 = 20 \mu\text{A}$$

Step 5 gives  $V_{SG3}$  as  $1.2 \text{ V}$ , which results in

$$\frac{W_3}{L_3} = \frac{W_4}{L_4} = \frac{20}{50(1.2 - 0.7)^2} = 1.6 \rightarrow \frac{W_3}{L_3} = \frac{W_4}{L_4} = 2$$

The desired gain is found to be 4000, which gives an input transconductance of

$$g_{m1} = \frac{4000 \cdot 0.09 \cdot 10}{44.44} = 81 \mu\text{S}$$

This gives the  $W/L$  ratios of M1 and M2 as

$$\frac{W_3}{L_3} = \frac{W_4}{L_4} = \frac{(81)^2}{110 \cdot 40} = 1.49 \rightarrow \frac{W_1}{L_1} = \frac{W_2}{L_2} = 2$$

To check the guess for  $C_I$  we need to calculate it, which is done as

$$C_I = C_{gd2} + C_{gd4} + C_{gs6} + C_{bd2} + C_{bd4} = 0.9 \text{ fF} + 0.4 \text{ fF} + 119.5 \text{ fF} + 20.4 \text{ fF} + 15.3 \text{ fF} \\ = 156.5 \text{ fF}$$

which is less than what was guessed.

Finally, the  $W/L$  value of M5 is found by finding  $V_{GS1}$  as 1.00 V, which gives  $V_{DSS}(\text{sat}) = 0.25$  V. This gives

$$\frac{W_5}{L_5} = \frac{2 \cdot 20}{(0.25)^2 \cdot 110} = 5.8 \approx 6$$

As in the previous example, M5 and M7 cannot be connected gate–gate and source–source and a scheme like Fig. 8.2-8 must be used. The  $W$  values are summarized below assuming that all channel lengths are 1  $\mu\text{m}$ .

$$\begin{array}{lll} W_1 = W_2 = 2 \mu\text{m} & W_3 = W_4 = 4 \mu\text{m} & W_5 = 6 \mu\text{m} \\ W_6 = 64 \mu\text{m} & W_7 = 29 \mu\text{m} & \end{array}$$

The above examples should help to illustrate the design of a two-stage, open-loop comparator. It is interesting to note there may be little difference in a comparator designed for linear response or for slewing response. This comes about from the fact that what causes the poles to be large is also what causes the slew rate to be fast. For example, large poles require low resistance and small capacitance from a node to ac ground. Low resistances come from larger bias currents. Hence, large bias currents and smaller capacitors lead to improved slew rate. Generally, a fast comparator will also be one that dissipates more power. In the next section, we will examine other forms of the open-loop comparator that offer more trade-offs between speed and power.



### 8.3 Other Open-Loop Comparators

There are many other types of comparators besides the two-stage comparator of the previous section. In fact, most of the op amps of the previous two chapters could be used as comparators. In this section we will examine comparators with push–pull outputs, the folded-cascode comparator, and comparators capable of driving very large capacitive loads.

#### Push–Pull Output Comparators

We noted in the two-stage comparator of the previous section that the propagation delay time was due to both the transition of the first-stage output and the second-stage output. If we replace the current-mirror load in the first stage with MOS diodes (gate–drain-connected MOSFETs), then the signal at the output of the first stage will be reduced in magnitude. This type of comparator is called a *clamped* comparator and is shown in Fig. 8.3-1.

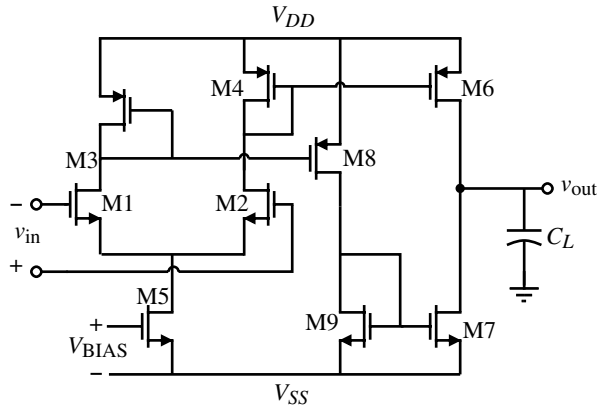


**Figure 8.3-1** Clamped push-pull output comparator.

There are several interesting features of Fig. 8.3-1. First, the gain has been reduced because the current-mirror load of the first stage has been replaced by MOS diodes. Second, we note that the output is push-pull. The maximum sinking and sourcing current at the output would be  $I_5$  times any current gain in M3–M8 (M4–M6). The gain equivalent to a two-stage comparator can be achieved by cascoding the output as shown in Fig. 8.3-2. We note that this circuit is simply the op amp of Fig. 6.5-6 using cascode circuits that are not self-biased. The large output resistance leads to a single-pole response. This pole is at lower frequencies than the poles of the two-stage, open-loop comparator and therefore the linear response will be slower with the same amount of overdrive. However, because the comparator is push-pull, it will be able to sink and source large currents into the output capacitance,  $C_{II}$ . The design procedure used in Example 6.5-2 can be applied to this circuit as well.

The dynamics of Fig. 8.3-2 depend on the output resistance,  $R_{II}$ . From Example 6.5-2, the output resistance is about  $11\text{ M}\Omega$ . If  $C_{II}$  is  $5\text{ pF}$ , the dominant pole is  $-18.12\text{ krad/s}$ . For an overdrive factor of 10, Eq. (8.1-7) gives a propagation delay time of  $2.83\text{ }\mu\text{s}$ . Compared to the two-stage, open-loop comparator of Example 8.2-6 this is much slower. However, if the comparator is slewing, then the clamped comparator with a cascoded output stage can compete with the two-stage, open-loop comparator.

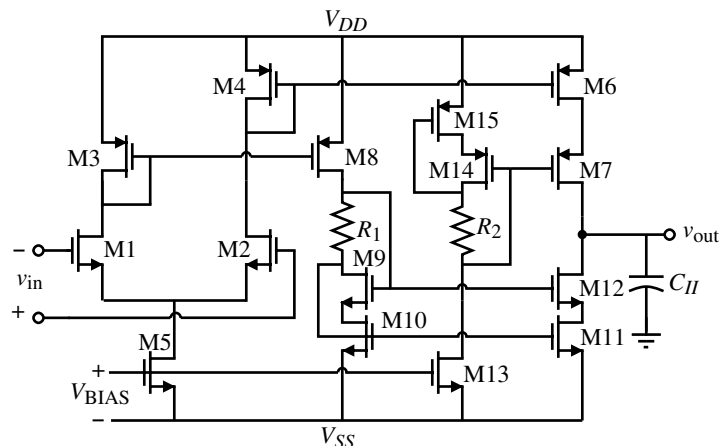


**Figure 8.3-2** Clamped comparator of Fig. 8.3-1 using a cascode output stage.

The folded-cascode op amp of Chapter 6 also makes a satisfactory comparator. Its performance is similar to the comparator of Fig. 8.3-2. The primary difference would be a better input common-mode voltage range because the MOS diodes are not used to load the first stage. From a linear speed viewpoint, the comparators with cascoded output stage are slow. In general, these types of comparators should not be used if the response is to be linear. They do have satisfactory performance as a comparator if the response is slewing.

### Comparators That Can Drive Large Capacitive Loads

If a comparator has a large capacitive load, the chances are that it is slew rate limited. Under these conditions, we will show several methods capable of driving large values of  $C_{II}$ . The first method is simply to add several push-pull inverters in cascade with the output of a two-stage, open-loop comparator as shown in Fig. 8.3-3.

The inverters, M8–M9 and M10–M11, allow  $C_{II}$  to be large without sacrificing the speed. The principle is well understood in high-speed digital buffers. If a large capacitance was connected to the drains of M6 and M7, the slew rate would be poor because the current sinking and sourcing is not large. The M8–M9 inverter allows the current driving capability to be increased without sacrificing the slew rate. The  $W/L$  value of M8 and M9 must be large enough to increase the current sinking and sourcing capability without loading M6 and M7. Similarly, the inverter M10–M11 allows the current sinking and sourcing to be further increased without loading M8–M9. It can be shown that, if the  $W/L$  increases by a factor of 2.72, the minimum propagation delay time is achieved. However, this is a very broad optimum so that larger factors like 10 are used to reduce the required number of stages.

The overdrive techniques introduced for low-power op amps in Section 7.4 can be used for comparators to increase the output current sinking and sourcing capability of the comparator. These methods included boosting the tail current (Fig. 7.4-4) and using current mirrors whose output transistor switches from the active to the saturation region when output current is required (Fig. 7.4-6). Another circuit that has the ability to sink and source large amounts of current is called the self-biased differential amplifier [2] and is shown in Fig. 8.3-4. This amplifier consists of two differential amplifiers each serving as the load for the other as illustrated in Fig. 8.3-4(a). The tail currents of the differential amplifiers become adaptive by connecting the gates of M5 and M6 to the drains of M1 and M3 as shown in

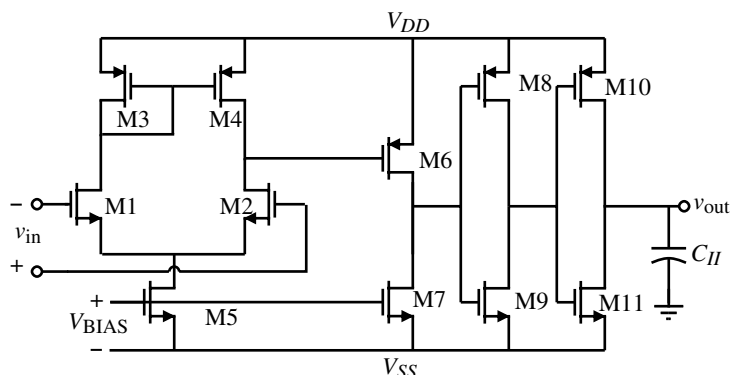


**Figure 8.3-3** Increasing the capacitive drive of a two-stage, open-loop comparator.

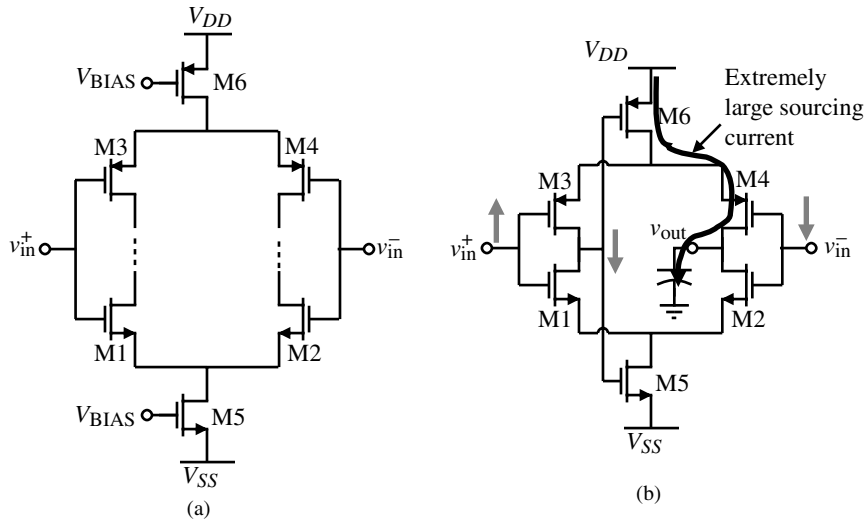


**Figure 8.3-4** (a) Two differential amplifiers loading each other. (b) Evolution of (a) into the self-biased differential amplifier.

Fig. 8.3-4(b). When the positive input voltage,  $v_{in}^+$ , is increased, the drains of M1 and M3 fall and turn on M6 to a large current, which is sourced to the output capacitance connected to the drains of M2 and M4 through M4. During this condition, the current in M5 is zero. When  $v_{in}^+$  is decreased, M5 turns on and a large current is sunk through the output capacitance via M2. Thus, this circuit has the ability to source and sink large currents without a large quiescent current. A disadvantage of the circuit is that the delay time from  $v_{in}^+$  to the output is slower than from  $v_{in}^-$  to the output.



## 8.4 Improving the Performance of Open-Loop Comparators

There are two areas in which the performance of an open-loop, high-gain comparator can be improved with little extra effort. These areas are the input-offset voltage and a single transition of the comparator in a noisy environment. The first problem can be solved by *autozeroing* and the second can be solved by the introduction of hysteresis using a bistable circuit. These two techniques will be examined in the following.

### Autozeroing Techniques

Input-offset voltage can be a particularly difficult problem in comparator design. In precision applications, such as high-resolution A/D converters, large input-offset voltages cannot be tolerated. While systematic offset can nearly be eliminated with proper design (though still affected by process variations), random offsets still remain and are unpredictable. Fortunately, there are techniques in MOS technology to remove a large portion of the input offset using offset-cancellation techniques. These techniques are available in MOS because of the nearly infinite input resistance of MOS transistors. This characteristic allows long-term storage of voltages on the transistor's gate. As a result, offset voltages can be measured, stored on capacitors, and summed with the input so as to cancel the offset.

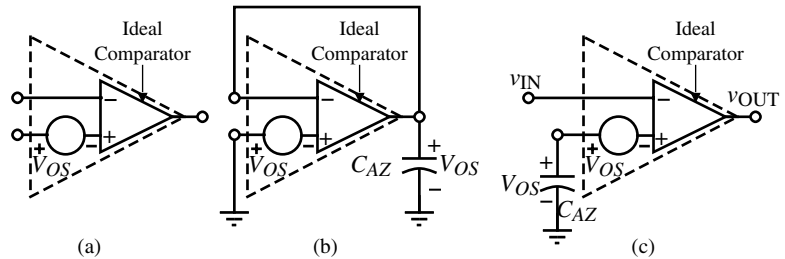


**Figure 8.4-1** (a) Simple model of a comparator including offset. (b) Comparator in unity-gain configuration storing the offset on autozero capacitor  $C_{AZ}$  during the first half of the autozero cycle. (c) Comparator in open-loop configuration with offset cancellation achieved at the noninverting input during the second half of the autozero cycle.

Figure 8.4-1 shows symbolically an offset-cancellation algorithm. A model of a comparator with an input-offset voltage is shown in Fig. 8.4-1(a). A known polarity is given to the offset voltage for convenience. Neither the value nor the polarity can be predicted in reality. Figure 8.4-1(b) shows the comparator connected in the unity-gain configuration so that the input offset is available at the output. In order for this circuit to work properly, it is necessary that the comparator be stable in the unity-gain configuration. This implies that only self-compensated high-gain amplifiers would be suitable for autozeroing. One could use the two-stage, open-loop comparator but a compensation circuit should be switched into the circuit during autozeroing. In the final operation of the autozero algorithm  $C_{AZ}$  is placed at the input of the comparator in series with  $V_{OS}$ . The voltage across  $C_{AZ}$  adds to  $V_{OS}$ , resulting in zero volts at the noninverting input of the comparator. Since there is no dc path to discharge the autozero capacitor, the voltage across it remains indefinitely (in the ideal case). In reality, there are leakage paths in shunt with  $C_{AZ}$  that can discharge it over a period of time. The solution to this problem is to repeat the autozero cycle periodically.

A practical implementation of a differential-input, autozeroed comparator is shown in Fig. 8.4-2(a). The comparator is modeled with an offset-voltage source as before. Figure 8.4-2(b) shows the state of the circuit during the first phase of the cycle when  $\phi_1$  is high. The offset is stored across  $C_{AZ}$ . Figure 8.4-2(c) shows the circuit in the second phase of the autozero cycle when  $\phi_2$  is high. The offset is canceled by the addition of  $V_{OS}$  on  $C_{AZ}$ . It is during this portion of the cycle that the circuit functions as a comparator.

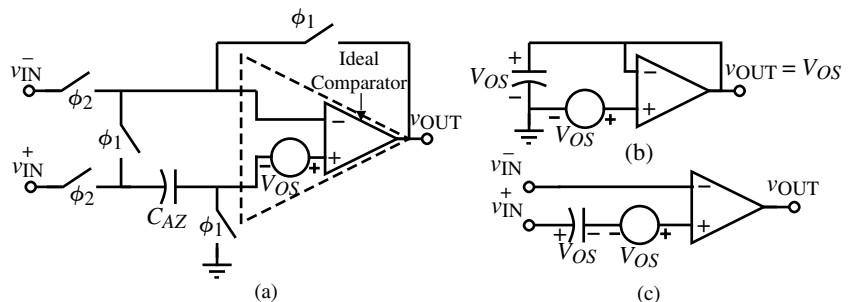
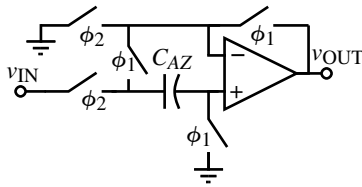


**Figure 8.4-2** (a) Differential circuit implementation of an autozeroed comparator. (b) Comparator during  $\phi_1$  phase. (c) Comparator during  $\phi_2$  phase.

**Figure 8.4-3** Noninverting autozeroed comparator.

There are other ways to implement an autozeroed comparator other than the one just described. A slight variation of Fig. 8.4-2(a) is shown in Fig. 8.4-3. This is a noninverting version of the previous autozeroed comparator. Yet another version is shown in Fig. 8.4-4. This one is simpler in operation since the noninverting input is always connected to ground.

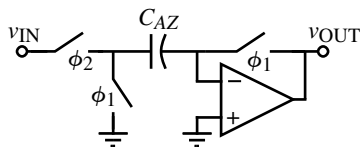
The switch implementations of the autozeroed comparator can be single-channel MOSFETs or complementary MOSFETs. It is important to use nonoverlapping clocks to drive the switches so that any given switch turns off before another turns on. Autozeroing techniques can be very effective in removing a large amount of a comparator's input offset, but the offset cancellation is not perfect. Charge injection resulting from clock feedthrough (Section 4.1) can by itself introduce an offset. This too can be canceled, but it is usually the cause of the lower limit of the offset voltage being greater than zero.

### Comparator Using Hysteresis

Often a comparator is placed in a very noisy environment in which it must detect signal transitions at the threshold point. If the comparator is fast enough (depending on the frequency of the most prevalent noise) and the amplitude of the noise is great enough, the output will also be noisy. In this situation, a modification on the transfer characteristic of the comparator is desired. Specifically, hysteresis is needed in the comparator.

Hysteresis is the quality of the comparator in which the input threshold changes as a function of the input (or output) level. In particular, when the input passes the threshold, the output changes and the input threshold is subsequently reduced so that the input must return beyond the previous threshold before the comparator's output changes state again. This can be illustrated much more clearly with the diagram shown in Fig. 8.4-5. Note that as the input starts negative and goes positive, the output does not change until it reaches the positive trip point,  $V_{TRP}^+$ . Once the output goes high, the effective trip point is changed. When the input returns in the negative direction, the output does not switch until it reaches the negative trip point,  $V_{TRP}^-$ .

The advantage of hysteresis in a noisy environment can clearly be seen from the illustration given in Fig. 8.4-6. In this figure, a noisy signal is shown as the input to a comparator without hysteresis. The intent is to have the comparator output follow the low-frequency signal. Because of noise variations near the threshold points, the comparator output is too noisy. The response of the comparator can be improved by adding hysteresis equal to or greater than the amount of the largest expected noise amplitude. The response of such a comparator is shown in Fig. 8.4-6(b).

**Figure 8.4-4** Inverting autozeroed comparator.

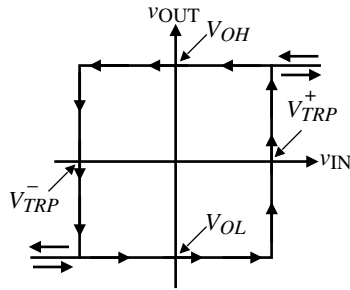


**Figure 8.4-5** Comparator transfer curve with hysteresis.

The voltage-transfer function shown in Fig. 8.4-5 is called a *bistable* characteristic. A bistable circuit can be clockwise or counterclockwise. Figure 8.4-5 is an example of a counterclockwise bistable characteristic. Sometimes, the counterclockwise bistable circuit is called noninverting and the clockwise bistable is called inverting. The bistable characteristic is defined by its width and height and whether it is clockwise or counterclockwise. The width is given by the difference between  $V_{TRP}^+$  and  $V_{TRP}^-$ . The height is generally the difference between  $V_{OH}$  and  $V_{OL}$ . In addition, the bistable characteristic can be shifted horizontally to the left or right by addition of a dc offset voltage.

There are many ways to introduce hysteresis in a comparator. All of them use some form of positive feedback. The methods can be categorized into external methods and internal methods. External hysteresis uses external positive feedback to implement hysteresis and can be accomplished after the comparator is built. Internal hysteresis is implemented into the comparator and does not require external feedback. We shall examine these two methods in the following.

Figure 8.4-7 shows a noninverting bistable circuit using external positive feedback to accomplish the hysteresis. This bistable characteristic is counterclockwise. We assume that the maximum and minimum output voltages of the comparator are  $V_{OH}$  and  $V_{OL}$ , respectively. The trip points are found as follows. Assume that  $v_{IN}$  is much less than the voltage at the positive input to the comparator. In this case, the output voltage will be at  $V_{OL}$ . As  $v_{IN}$  is increased, the upper trip point,  $V_{TRP}^+$ , is found by setting the voltage due to  $v_{IN}$  and  $V_{OL}$  at the positive input to the comparator equal to zero. This results in

$$0 = \left( \frac{R_1}{R_1 + R_2} \right) V_{OL} + \left( \frac{R_2}{R_1 + R_2} \right) V_{TRP}^+ \quad (8.4-1)$$

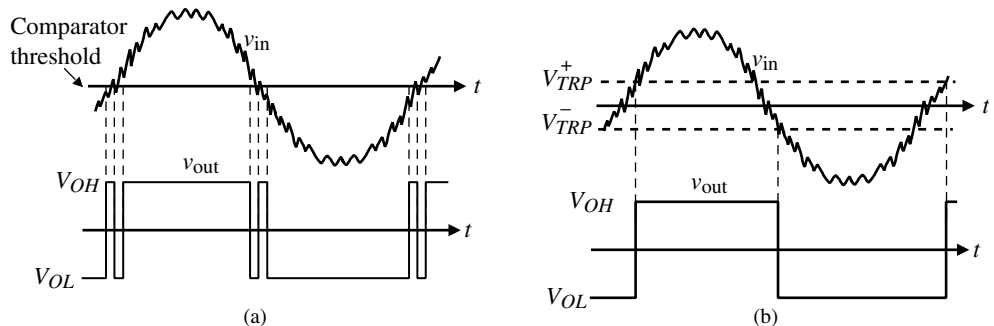


**Figure 8.4-6** (a) Comparator response to a noisy input. (b) Comparator response to a noisy input when hysteresis is added.

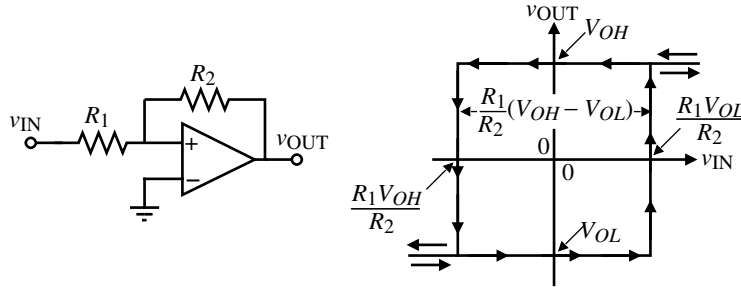


**Figure 8.4-7** Noninverting bistable circuit using external positive feedback.

Solving for the upper trip point gives

$$V_{TRP}^+ = -\frac{R_1 V_{OL}}{R_2} \quad (8.4-2)$$

Generally,  $V_{OL}$  is negative so that the upper trip point is a positive voltage.

The lower trip point,  $V_{TRP}^-$ , can be found by assuming that  $v_{IN}$  is much greater than the voltage at the positive input to the comparator so that the output state, which is  $V_{OH}$ , is known. As  $v_{IN}$  is decreased,  $V_{TRP}^-$  occurs when the voltage at the positive input to the comparator is equal to zero. Therefore,

$$0 = \left(\frac{R_1}{R_1 + R_2}\right)V_{OH} + \left(\frac{R_2}{R_1 + R_2}\right)V_{TRP}^- \quad (8.4-3)$$

which gives

$$V_{TRP}^- = -\frac{R_1 V_{OH}}{R_2} \quad (8.4-4)$$

The width of the bistable characteristic is given as

$$\Delta V_{in} = V_{TRP}^+ - V_{TRP}^- = \left(\frac{R_1}{R_2}\right)(V_{OH} - V_{OL}) \quad (8.4-5)$$

The counterclockwise bistable circuit resulting from this analysis is shown in Fig. 8.4-7.

Figure 8.4-8 shows the clockwise bistable circuit using external positive feedback. Assuming that the input is much less than the voltage at the positive input of the comparator allows us to define the output state at  $V_{OH}$ . The upper trip point can be found by setting the input equal to the voltage at the positive input of the comparator. The result is

$$v_{IN} = V_{TRP}^+ = \left(\frac{R_1}{R_1 + R_2}\right)V_{OH} \quad (8.4-6)$$

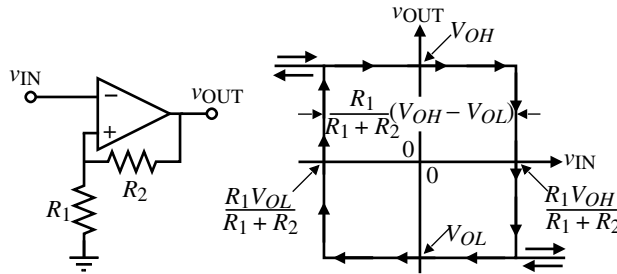


**Figure 8.4-8** Inverting bistable circuit using external positive feedback.

Next, assuming that the input is much greater than the voltage at the positive input of the comparator defines the output voltage as  $V_{OL}$ . The lower trip point can be found by setting the input equal to the voltage at the positive input of the comparator. Therefore,

$$v_{IN} = V_{TRP}^- = \left( \frac{R_1}{R_1 + R_2} \right) V_{OL} \quad (8.4-7)$$

The width of the bistable characteristic is given as

$$\Delta V_{in} = V_{TPR}^+ - V_{TRP}^- = \left( \frac{R_1}{R_1 + R_2} \right) (V_{OH} - V_{OL}) \quad (8.4-8)$$

The clockwise bistable circuit resulting from this analysis is shown in Fig. 8.4-8.

The center point of the bistable characteristics of Figs. 8.4-7 and 8.4-8 can be shifted horizontally by inserting a battery,  $V_{REF}$ , as shown in Fig. 8.4-9 for the counterclockwise bistable circuit. If we solve for the trip points of the counterclockwise bistable circuit in Fig. 8.4-9 we get

$$V_{REF} = \left( \frac{R_1}{R_1 + R_2} \right) V_{OL} + \left( \frac{R_2}{R_1 + R_2} \right) V_{TRP}^+ \quad (8.4-9)$$

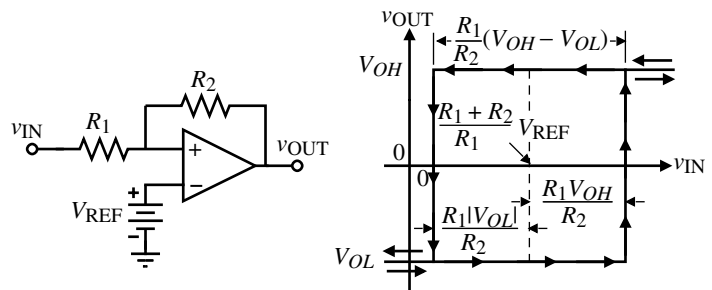


**Figure 8.4-9** Horizontally shifted noninverting bistable circuit using external positive feedback.

or

$$V_{TRP}^+ = \left( \frac{R_1 + R_2}{R_2} \right) V_{REF} - \frac{R_1}{R_2} V_{OL} \quad (8.4-10)$$

and

$$V_{REF} = \left( \frac{R_1}{R_1 + R_2} \right) V_{OH} + \left( \frac{R_2}{R_1 + R_2} \right) V_{TRP}^- \quad (8.4-11)$$

or

$$V_{TRP}^- = \left( \frac{R_1 + R_2}{R_2} \right) V_{REF} - \frac{R_1}{R_2} V_{OH} \quad (8.4-12)$$

The width of the bistable characteristic has not changed but is now centered at  $(R_1 + R_2)/R_2$  times  $V_{REF}$ .

Figure 8.4-10 shows how the bistable characteristic for the inverting or clockwise bistable circuit can be shifted horizontally by inserting a voltage,  $V_{REF}$ , in series with  $R_1$ . Setting the input voltage equal to the voltage at the positive input terminal to the comparator gives the upper trip point as

$$v_{IN} = V_{TRP}^+ = \left( \frac{R_1}{R_1 + R_2} \right) V_{OH} + \left( \frac{R_2}{R_1 + R_2} \right) V_{REF} \quad (8.4-13)$$

The lower trip point can be found by setting the input equal to the voltage at the positive input of the comparator. Therefore,

$$v_{IN} = V_{TRP}^- = \left( \frac{R_1}{R_1 + R_2} \right) V_{OL} + \left( \frac{R_2}{R_1 + R_2} \right) V_{REF} \quad (8.4-14)$$

The width of the bistable characteristic has not changed but is now centered at  $R_2/(R_1 + R_2)$  times  $V_{REF}$ .

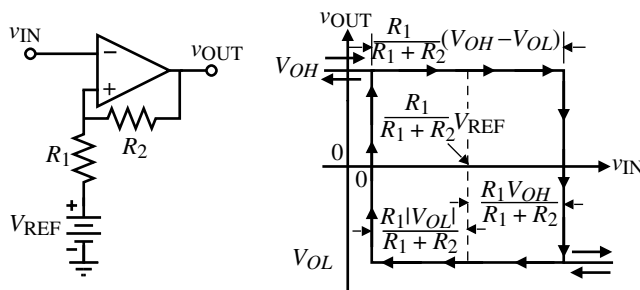


**Figure 8.4-10** Horizontally shifted inverting bistable circuit using external positive feedback.

**Example  
8.4-1****Example 8.4-1 Design of an Inverting Comparator with Hysteresis**

Use the circuit of Fig. 8.4-10 to design a high-gain, open-loop comparator having an upper trip point of 1 V and a lower trip point of 0 V if  $V_{OH} = 2$  V and  $V_{OL} = -2$  V.

**SOLUTION**

Putting the values of this example into Eqs. (8.4-13) and (8.4-14) gives

$$1 = \left( \frac{R_1}{R_1 + R_2} \right) 2 + \left( \frac{R_2}{R_1 + R_2} \right) V_{REF}$$

and

$$0 = \left( \frac{R_1}{R_1 + R_2} \right) (-2) + \left( \frac{R_2}{R_1 + R_2} \right) V_{REF}$$

Solving these two equations gives  $3R_1 = R_2$  and  $V_{REF} = 0.667$  V.

The previous circuits are examples of using external positive feedback to implement hysteresis in a high-gain, open-loop comparator. The hysteresis can also be accomplished by using internal positive feedback. Figure 8.4-11 shows the differential-input stage of a comparator such as Fig. 8.3-1 or 8.3-2 [1]. In this circuit there are two paths of feedback. The first is current-series [3,4] feedback through the common-source node of transistors M1 and M2. This feedback path is negative. The second path is the voltage-shunt feedback through the gate-drain connections of transistors M6 and M7. This path of feedback is positive. If the positive-feedback factor is less than the negative-feedback factor, then the overall feedback will be negative and no hysteresis will result. If the positive-feedback factor becomes greater, the overall feedback will be positive, which will give rise to hysteresis in the voltage-transfer curve. As long as the ratio  $\beta_6/\beta_3$  is less than one, there is no hysteresis in the transfer function. When this ratio is greater than one, hysteresis will result. The following analysis will develop the equations for the trip points when there is hysteresis.

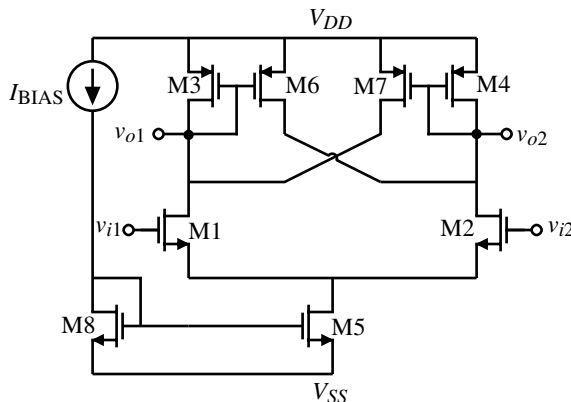


**Figure 8.4-11** Implementation of hysteresis using internal positive feedback in the input stage of a high-gain, open-loop comparator.

Assume that plus and minus supplies are used and that the gate of M1 is tied to ground. With the input of M2 much less than zero, M1 is on and M2 off, thus turning on M3 and M6 and turning off M4 and M7. All of  $i_5$  flows through M1 and M3, so  $v_{o2}$  is high. The resulting circuit is shown in Fig. 8.4-12(a). Note that M2 is shown even though it is off. At this point, M6 is attempting to source the following amount of current:

$$i_6 = \frac{(W/L)_6}{(W/L)_3} i_5 \quad (8.4-15)$$

As  $v_{in}$  increases toward the threshold point (which is not yet known), some of the tail current  $i_5$  begins to flow through M2. This continues until the point where the current through M2 equals the current in M6. Just beyond this point the comparator switches state. To approximately calculate one of the trip points, the circuit must be analyzed right at the point where  $i_2$  equals  $i_6$ . Mathematically this is

$$i_6 = \frac{(W/L)_6}{(W/L)_3} i_3 \quad (8.4-16)$$

$$i_2 = i_6 \quad (8.4-17)$$

$$i_5 = i_2 + i_1 \quad (i_1 = i_3) \quad (8.4-18)$$

Therefore,

$$i_3 = \frac{i_5}{1 + [(W/L)_6/(W/L)_3]} = i_1 \quad (8.4-19)$$

$$i_2 = i_5 - i_1 \quad (8.4-20)$$

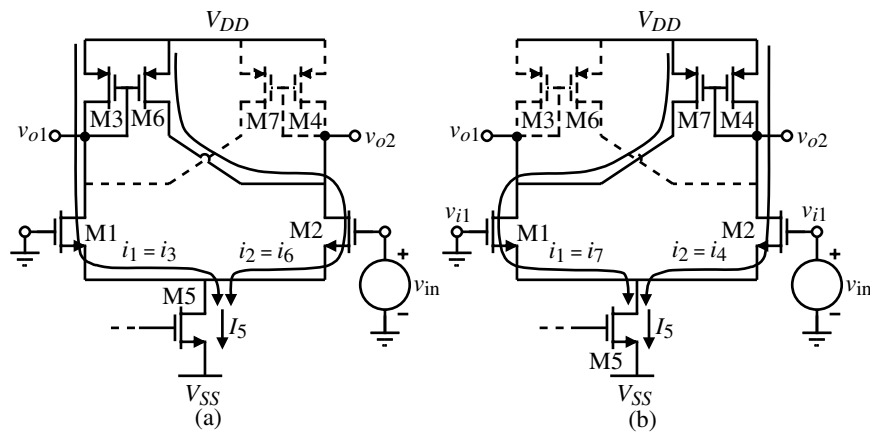


**Figure 8.4-12** (a) Comparator of Fig. 8.4-11 where  $v_{in}$  is very negative and increasing toward  $V_{TRP}^+$ . (b) Comparator of Fig. 8.4-11 where  $v_{in}$  is very positive and decreasing toward  $V_{TRP}^-$ .

Knowing the currents in both M1 and M2, it is easy to calculate their respective  $v_{GS}$  voltages. Since the gate of M1 is at ground, the difference in their gate–source voltages will yield the positive trip point as given below:

$$v_{GS1} = \left( \frac{2i_1}{\beta_1} \right)^{1/2} + V_{T1} \quad (8.4-21)$$

$$v_{GS2} = \left( \frac{2i_2}{\beta_2} \right)^{1/2} + V_{T2} \quad (8.4-22)$$

$$V_{TRP}^+ = v_{GS2} - v_{GS1} \quad (8.4-23)$$

Once the threshold is reached, the comparator changes state so that the majority of the tail current now flows through M2 and M4. As a result, M7 is also turned on, thus turning off M3, M6, and M1. As in the previous case, as the input decreases the circuit reaches a point at which the current in M1 increases until it equals the current in M7. The input voltage at this point is the negative trip point  $V_{TRP}^-$ . The equivalent circuit in this state is shown in Fig. 8.4-12(b). To calculate the trip point, the following equations apply:

$$i_7 = \frac{(W/L)_7}{(W/L)_4} i_4 \quad (8.4-24)$$

$$i_1 = i_7 \quad (8.4-25)$$

$$i_5 = i_2 + i_1 \quad (8.4-26)$$

Therefore,

$$i_4 = \frac{i_5}{1 + [(W/L)_7/(W/L)_4]} = i_2 \quad (8.4-27)$$

$$i_1 = i_5 - i_2 \quad (8.4-28)$$

Using Eqs. (8.4-21) and (8.4-22) to calculate  $v_{GS}$ , the trip point is

$$V_{TRP}^- = v_{GS2} - v_{GS1} \quad (8.4-29)$$

These equations do not take into account the effect of channel length modulation. Their use is illustrated in the following example.

#### Example 8.4-2

#### Calculation of Trip Voltages for a Comparator with Hysteresis

Consider the circuit shown in Fig. 8.4-11. Using the transistor device parameters given in Table 3.1-2 calculate the positive and negative threshold points if the device lengths are all 1  $\mu\text{m}$  and the widths are given as  $W_1 = W_2 = W_{10} = W_{11} = 10 \mu\text{m}$  and  $W_3 = W_4 = 2 \mu\text{m}$ .

The gate of M1 is tied to ground and the input is applied to the gate of M2. The current  $i_5 = 20 \mu\text{A}$ . Simulate the results using simulation.

### SOLUTION

To calculate the positive trip point, assume that the input has been negative and is heading positive.

$$i_6 = \frac{(W/L)_6}{(W/L)_3} i_3 = (5/1)(i_3)$$

$$i_3 = \frac{i_5}{1 + [(W/L)_6/(W/L)_3]} = i_1 = \frac{20 \mu\text{A}}{1 + 5} = 3.33 \mu\text{A}$$

$$i_2 = i_5 - i_1 = 20 - 3.33 = 16.67 \mu\text{A}$$

$$v_{GS1} = \left( \frac{2i_1}{\beta_1} \right)^{1/2} + V_{T1} = \left( \frac{2 \cdot 3.33}{(5)110} \right)^{1/2} + 0.7 = 0.81 \text{ V}$$

$$v_{GS2} = \left( \frac{2i_2}{\beta_2} \right)^{1/2} + V_{T2} = \left( \frac{2 \cdot 16.67}{(5)110} \right)^{1/2} + 0.7 = 0.946 \text{ V}$$

$$V_{TRP}^+ \cong v_{GS2} - v_{GS1} = 0.946 - 0.810 = 0.136 \text{ V}$$

Determining the negative trip point, similar analysis yields

$$i_4 = 3.33 \mu\text{A}$$

$$i_1 = 16.67 \mu\text{A}$$

$$v_{GS2} = 0.81 \text{ V}$$

$$v_{GS1} = 0.946 \text{ V}$$

$$V_{TRP}^- \cong v_{GS2} - v_{GS1} = 0.81 - 0.946 = -0.136 \text{ V}$$

PSPICE simulation results of this circuit are shown in Fig. 8.4-13.

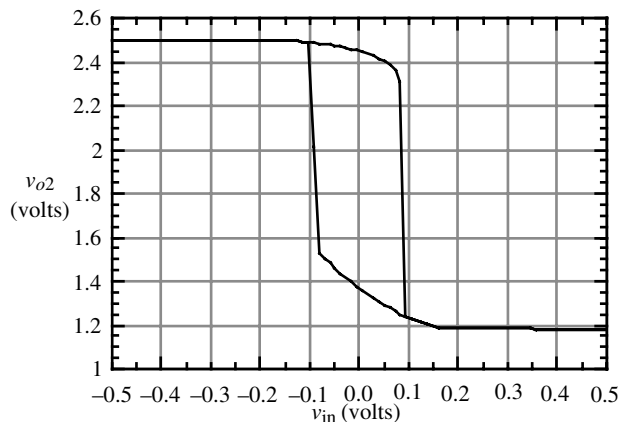


**Figure 8.4-13** Simulation of the comparator of Example 8.4-2.

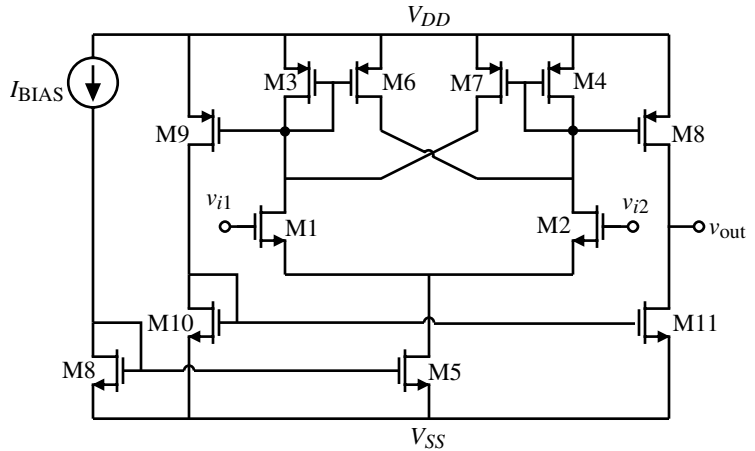


**Figure 8.4-14** Complete comparator with internal hysteresis including an output stage.

The differential stage described thus far is generally not useful alone and thus requires an output stage to achieve reasonable voltage swings and output resistance. There are a number of ways to implement an output stage for this type of input stage. One of these is given in Fig. 8.4-14. Differential-to-single-ended conversion is accomplished at the output and thus provides a Class AB type of driving capability.



## 8.5 Discrete-Time Comparators

In many applications the comparator only functions over a portion of a time period. Such circuits are driven by a clock and will have a portion of time or phase when the comparator is functioning as a comparator and a phase when the comparator is not being used. In this circumstance, other forms of comparators can be used that are efficient and have a small propagation delay time. We will examine two such comparators in this section. They are the switched capacitor comparator and the regenerative comparator.

### Switched Capacitor Comparators

The switched capacitor comparator uses a combination of switched capacitors and open-loop comparators. The advantages of the switched capacitor comparator are that differential signals can be compared using single-ended circuits and the switched capacitor comparator naturally lends itself to autozeroing the dc offset voltage of the open-loop comparator. Figure 8.5-1(a) shows a typical switched capacitor comparator. The voltages applied to the circuit are normally sampled and held so that capital variables are used.

When the  $\phi_1$  switches are closed in Fig. 8.5-1, the capacitor  $C$  autozeros the offset voltage of the comparator,  $V_{OS}$ . The capacitor  $C_p$  represents the parasitic capacitance from the input of the comparator to ground. We recall that the comparator must be stable in unity-gain

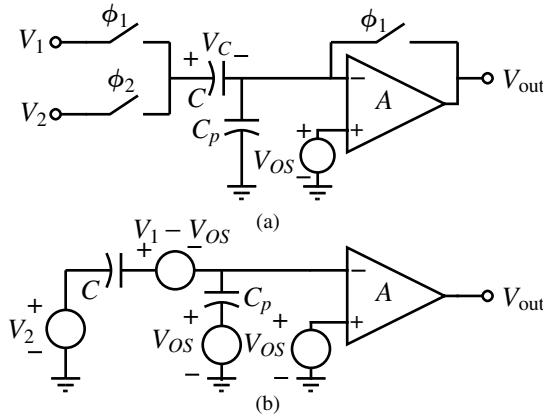


**Figure 8.5-1** (a) A switched capacitor comparator. (b) Equivalent circuit of (a) when the  $\phi_2$  switches are closed.

operation for this circuit to work properly. It can be shown that the voltage across  $C_1$  and  $C_p$  at the end of the  $\phi_1$  phase period is

$$V_C(\phi_1) = V_1 - V_{Os} \quad (8.5-1)$$

and

$$V_{C_p}(\phi_1) = V_{Os} \quad (8.5-2)$$

Now when the  $\phi_2$  switch is closed, the equivalent circuit at the beginning of the  $\phi_2$  phase period is shown in Fig. 8.5-1(b). In this circuit, the voltage across each capacitor has been removed and represented by a step voltage source. This allows us to use the principle of superposition to easily solve for the output as illustrated below.

$$\begin{aligned} V_{out}(\phi_2) &= -A \left[ \frac{V_2 C}{C + C_p} - \frac{(V_1 - V_{Os})C}{C + C_p} + \frac{V_{Os} C_p}{C + C_p} \right] + A V_{Os} \\ &= -A \left[ (V_2 - V_1) \frac{C}{C + C_p} - V_{Os} \left( \frac{C}{C + C_p} + \frac{C_p}{C + C_p} \right) \right] + A V_{Os} = -A (V_2 - V_1) \frac{C}{C + C_p} \end{aligned} \quad (8.5-3)$$

If  $C_p$  is smaller than  $C$ , then Eq. (8.5-3) can be simplified to

$$V_{out}(\phi_2) \approx A(V_1 - V_2) \quad (8.5-4)$$

Therefore, the difference between the voltages  $V_1$  and  $V_2$  is amplified by the gain of the comparator.

The gain of the comparator used for the switched capacitor comparator must be large enough to satisfy the resolution requirements. In many cases, the resolution is large (i.e., 100 mV), so that a very simple single-stage amplifier is sufficient for the comparator. The speed of the comparator depends on how long it takes the circuit to settle to its steady state after the switches have been closed for a given period. During the  $\phi_1$  phase, the response of the circuit in Fig. 8.5-1(a) is very fast. The time constants of the circuit are associated with the product of the switch on-resistances and the capacitor  $C$  and the dynamics of the comparator

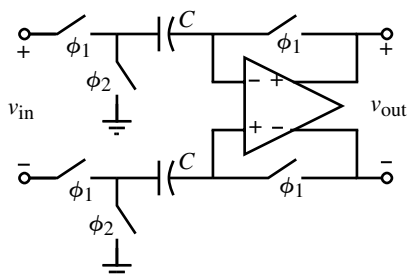


**Figure 8.5-2** A differential-in, differential-out, switched capacitor comparator.

in unity-gain configuration. Both of these time constants can be small. During  $\phi_2$ , the open-loop response of the comparator will determine the speed. This subject has been covered in detail in Section 8.1 for a single-pole approximation and in Section 8.2 for a multiple-pole approximation of the comparator.

A differential switched capacitor comparator is shown in Fig. 8.5-2. The input is sampled on the two identical capacitors,  $C$ , during the  $\phi_1$  phase period. The dc offsets of the differential-in, differential-out comparator are also autozeroed during this period. During the  $\phi_2$  phase period, the voltages sampled across the capacitors are applied to the comparator input. As the  $\phi_1$  switches open at the end of the  $\phi_1$  phase period, the charge injection will be reduced because the signal is a differential voltage and the charge injection is a common-mode voltage.

## Regenerative Comparators

Regenerative comparators use positive feedback to accomplish the comparison of two signals. The regenerative comparator is also called a *latch* or a *bistable* [5]. The simplest form of a latch is shown in Fig. 8.5-3 and consists of two cross-coupled MOSFETs. Figure 8.5-3(a) uses NMOS transistors while Fig. 8.5-3(b) is a PMOS latch. The current sources/sinks are used to identify the dc currents in the transistors. Normally, the latch has two modes of operation. The first mode disables the positive feedback and applies the input signal to the terminals designated as  $v_{o1}$  and  $v_{o2}$ . The initial voltages applied during this mode will be designated as  $v'_{o1}$  and  $v'_{o2}$ . The second mode enables the latch and depending on the relative values of  $v'_{o1}$  and  $v'_{o2}$ , one of the outputs will go high and the other will go low. A two-phase clock is used to determine the modes of operation.

It is important to characterize the time it takes for the latch to go from its initial state to the final state during the enabled mode of operation. Figure 8.5-4(a) is a redrawing of

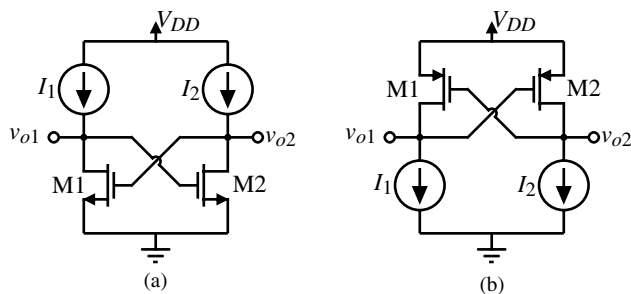


**Figure 8.5-3** (a) NMOS latch.  
(b) PMOS latch.

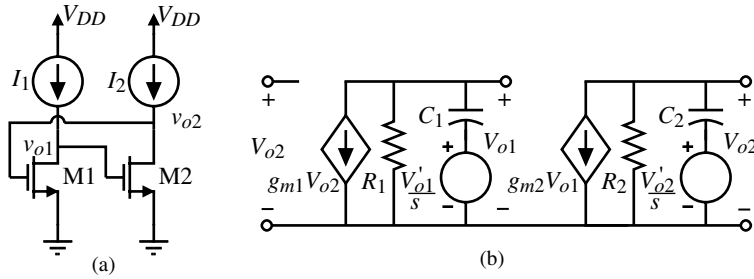


**Figure 8.5-4** (a) Redrawing of Fig. 8.5-3(a). (b) Equivalent model of (a).

the latch in Fig. 8.5-3(a). Let us assume that the initial values of  $v_{o1}$  and  $v_{o2}$  have been established and find the time it takes for the latch to operate. We will use the model of Fig. 8.5-4(b) to analyze the latch of Fig. 8.5-4(a). The voltage sources in series with the capacitances represent the initial values of  $v_{o1}$  and  $v_{o2}$  and are step functions.

We may write the nodal equations of Fig. 8.5-4(b) as

$$g_{m1}V_{o2} + G_1V_{o1} + sC_1\left(V_{o1} - \frac{V'_{o1}}{s}\right) = g_{m1}V_{o2} + G_1V_{o1} + sC_1V_{o1} - C_1V'_{o1} = 0 \quad (8.5-5)$$

and

$$g_{m2}V_{o1} + G_2V_{o2} + sC_2\left(V_{o2} - \frac{V'_{o2}}{s}\right) = g_{m2}V_{o1} + G_2V_{o2} + sC_2V_{o2} - C_2V'_{o2} = 0 \quad (8.5-6)$$

where  $G_1$  and  $G_2$  are the conductances seen from the drains of M1 and M2 to ground and  $C_1$  and  $C_2$  are the capacitances seen from the drains of M1 and M2 to ground. Solving Eqs. (8.5-5) and (8.5-6) for  $V_{o1}$  and  $V_{o2}$  gives

$$V_{o1} = \frac{R_1C_1}{sR_1C_1 + 1}V'_{o1} - \frac{g_{m1}R_1}{sR_1C_1 + 1}V_{o2} = \frac{\tau_1}{s\tau_1 + 1}V'_{o1} - \frac{g_{m1}R_1}{s\tau_1 + 1}V_{o2} \quad (8.5-7)$$

and

$$V_{o2} = \frac{R_2C_2}{sR_2C_2 + 1}V'_{o2} - \frac{g_{m2}R_2}{sR_2C_2 + 1}V_{o1} = \frac{\tau_2}{s\tau_2 + 1}V'_{o2} - \frac{g_{m2}R_2}{s\tau_2 + 1}V_{o1} \quad (8.5-8)$$

where  $\tau_i$  is the time constant  $R_iC_i$ . Assume that both transistors are identical so that  $g_{m1} = g_{m2} = g_m$ ,  $R_1 = R_2 = R$ , and  $C = C_1 = C_2$ , which gives  $\tau_1 = \tau_2 = \tau$ . Let us define the difference between  $V_{o2}$  and  $V_{o1}$  as  $\Delta V_o$  and the difference between  $V'_{o2}$  and  $V'_{o1}$  as  $\Delta V_i$ . Therefore,

$$\Delta V_o = V_{o2} - V_{o1} = \frac{\tau}{s\tau + 1}\Delta V_i + \frac{g_mR}{s\tau + 1}\Delta V_o \quad (8.5-9)$$

Solving for  $\Delta V_o$  gives

$$\Delta V_o = \frac{\tau \Delta V_i}{s\tau + (1 - g_m R)} = \frac{\frac{\tau \Delta V_i}{1 - g_m R}}{\frac{s\tau}{1 - g_m R} + 1} = \frac{\tau' \Delta V_i}{s\tau' + 1} \quad (8.5-10)$$

where

$$\tau' = \frac{\tau}{1 - g_m R} \quad (8.5-11)$$

Taking the inverse Laplace transform of Eq. (8.5-10) gives

$$\Delta v_o(t) = \Delta V_i' e^{-t/\tau'} = \Delta V_i e^{-t(1 - g_m R)/\tau} \approx e^{g_m R t/\tau} \Delta V_i \quad (8.5-12)$$

if  $g_m R$  is greater than one. The time constant of the latch as expressed in Eq. (8.5-12) is given as

$$\tau_L \approx \frac{\tau}{g_m R} = \frac{C}{g_m} \quad (8.5-13)$$

If  $C$  is mostly the gate-source capacitance, then the latch time constant can be expressed as

$$\tau_L = \frac{0.67 W L C_{ox}}{\sqrt{2K'(W/L)I}} = 0.67 C_{ox} \sqrt{\frac{W L^3}{2K'I}} \quad (8.5-14)$$

Equation (8.5-14) shows the strong dependence of the latch time constant on the channel length. Therefore, the latch time response can be expressed as

$$\Delta V_{out}(t) = e^{t/\tau_L} \Delta V_i \quad (8.5-15)$$

Equation (8.5-15) gives the difference between the latch output voltages,  $\Delta V_{out}$ , at a time  $t$  after enabling the latch. The voltage  $\Delta V_i$  is the difference between the latch output voltages,  $V_{o2}$  and  $V_{o1}$ , before the latch is enabled.

Figure 8.5-5 shows the time-domain response of the latch for various values of  $\Delta V_i$ . This figure has been normalized by dividing Eq. (8.5-15) by  $V_{OH} - V_{OL}$  to get

$$\frac{\Delta V_{out}(t)}{V_{OH} - V_{OL}} = e^{t/\tau_L} \frac{\Delta V_i}{V_{OH} - V_{OL}} \quad (8.5-16)$$

It is important to remember that  $\Delta V_i$  will always be less than  $V_{OH} - V_{OL}$ . The propagation time delay of a latch can be found by setting Eq. (8.5-16) equal to 0.5. The result is

$$t_p = \tau_L \ln\left(\frac{V_{OH} - V_{OL}}{2 \Delta V_i}\right) \quad (8.5-17)$$

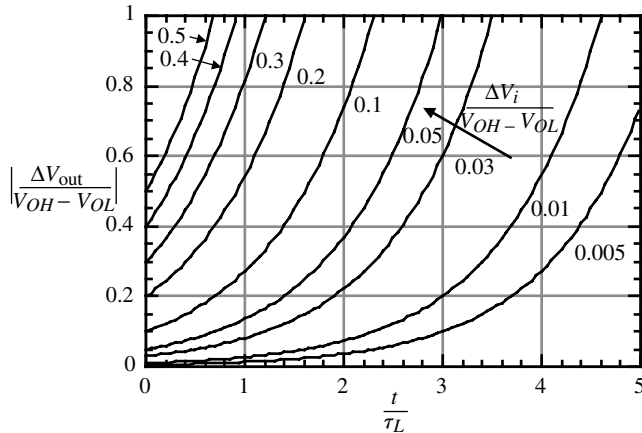


**Figure 8.5-5** Normalized time-domain response of a latch.

Since  $\Delta V_i$  is always less than  $0.5 (V_{OH} - V_{OL})$ , the argument of the logarithm is always greater than unity.

Figure 8.5-5 illustrates the time-domain response characteristics of a latch. There are several important observations to be made. The first is that the time required for  $\Delta V_{out}$  to reach  $V_{OH} - V_{OL}$  is decreased by applying a larger input to the latch,  $\Delta V_i$ , before enabling the latch. The second and more obvious is that the smaller the latch time constant, the faster the response. If the input to the latch before enabling the latch,  $\Delta V_i$ , is small, the latch takes a long time for the output voltage,  $\Delta V_{out}$ , to reach  $V_{OH} - V_{OL}$ . Therefore, it is desirable to apply a reasonably large value of  $\Delta V_i$  in order to take advantage of the rapidly increasing slope of the positive exponential characteristics of the latch.

### Example 8.5-1

#### Time-Domain Characteristics of a Latch

Find the time it takes from the time the latch is enabled until the output voltage,  $\Delta V_{out}$ , equals  $V_{OH} - V_{OL}$  if the  $W/L$  of the latch NMOS transistors is  $10 \mu\text{m}/1 \mu\text{m}$  and the latch dc current is  $10 \mu\text{A}$  when  $\Delta V_i = 0.01(V_{OH} - V_{OL})$  and  $\Delta V_i = 0.1(V_{OH} - V_{OL})$ . Find the propagation time delay for the latch for each of these conditions.

#### SOLUTION

The transconductance of the latch transistors is

$$g_m = \sqrt{2 \cdot 110 \cdot 10 \cdot 10} = 148 \mu\text{S}$$

The output conductance gives a latch gain of  $370 \text{ V/V}$ . Since  $g_m R$  is greater than one, we can use Eq. (8.5-14). Therefore, the latch time constant is found as

$$\tau_L = 0.67 C_{ox} \sqrt{\frac{WL^3}{2K'}} = 0.67(24 \times 10^{-4}) \sqrt{\frac{(10 \cdot 1) \times 10^{-24}}{2 \cdot 110 \times 10^{-6} \cdot 10 \times 10^{-6}}} = 0.112 \text{ ns}$$

Using Eq. (8.5-17) or Fig. 8.5-5 we find for  $\Delta V_i = 0.01(V_{OH} - V_{OL})$  that  $t = 4.6\tau_L = 0.515$  ns and for  $\Delta V_i = 0.1(V_{OH} - V_{OL})$  that  $t = 2.3\tau_L = 0.258$  ns.

The propagation time delay can be found using Eq. (8.5-17) and is 0.438 ns and 0.180 ns for  $\Delta V_i = 0.01(V_{OH} - V_{OL})$  and  $\Delta V_i = 0.1(V_{OH} - V_{OL})$ , respectively.

A practical latch comparator is shown in Fig. 8.5-6 [6]. M7 and M8 are the latch transistors and are PMOS in this case. M9 and M10 are used to reset the latch by setting the source–drain voltages of M7 and M8 to zero. The input to the latch is applied to the gates of M1A and M1B. The transistors M1A, M1B, M2A, and M2B are operating in the triode region. The values of the inputs will cause the resistance seen by the sources of M3 and M4 to ground to vary. When the latch is enabled, the drains of M3 and M4 are connected to the latch outputs. M3 and M4 form a parallel positive-feedback path for the latch. For example, the signal at the gate of M7 can go through M7 or can go through M3 (M5 is a closed switch). The gain of the M3 and M4 feedback paths depends on the value of the resistor,  $R_1$  or  $R_2$ , respectively. If the resistor is small the gain is large and that side of the latch will go high.

When the latch/reset goes high, the latch goes into its regenerative mode. The drain currents of M5 and M6 are steered to obtain a final state determined by the mismatch between the  $R_1$  and  $R_2$  resistances. These resistances are given as

$$\frac{1}{R_1} = K'_N \left[ \frac{W_{1A}}{L} (v_{in}^+ - V_T) + \frac{W_{2A}}{L} (V_{REF}^- - V_T) \right] \quad (8.5-18)$$

and

$$\frac{1}{R_2} = K'_N \left[ \frac{W_{1B}}{L} (v_{in}^- - V_T) + \frac{W_{2B}}{L} (V_{REF}^+ - V_T) \right] \quad (8.5-19)$$

The value of the input voltage that causes  $R_1$  and  $R_2$  to be equal is given as

$$v_{in}(\text{threshold}) = \left( \frac{W_2}{W_1} \right) V_{REF} \quad (8.5-20)$$

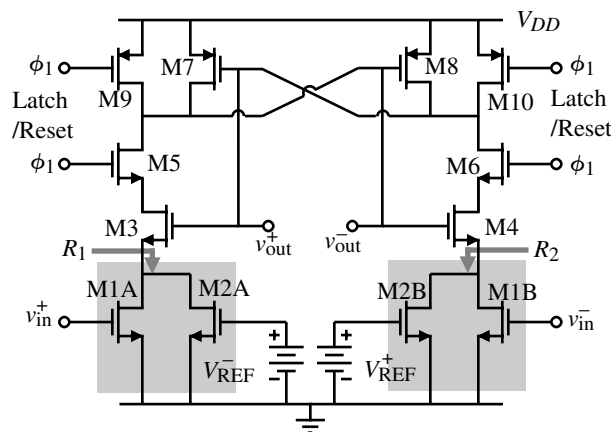


**Figure 8.5-6** Comparator using a latch with a built-in threshold.



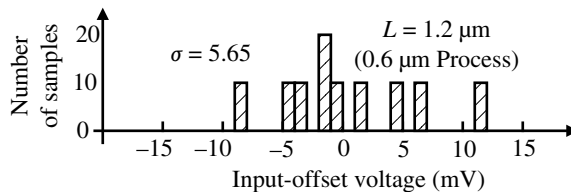


**Figure 8.5-9** Input-offset voltage distribution histogram of the dynamic latch of Fig. 8.5-8.

input-offset voltage is distributed over a range of  $\pm 10$  mV. This comparator can only discriminate the difference between two signals to a 5 bit accuracy if the reference voltage was 1 V. This range can be reduced by preceding the latch by a preamplifier, which will be discussed in the next section.



## 8.6 High-Speed Comparators

A high-speed comparator should have a propagation delay time as small as possible. In order to achieve this goal, one must understand the requirements for a fast comparator. This is best understood by separating the comparator into a number of cascaded stages. Figure 8.6-1 illustrates how this can be accomplished. A number of stages with a gain of  $A_0$  and a single pole at  $1/\tau$  are shown. If the input change is slightly larger than  $V_{in(min)}$ , then the function of the stages is to amplify the input with as little delay per stage as possible. We note that the signal swings in the initial stages will be small. As the signal swing begins to approach the desired range, the amplifiers will be limited by their slew rate. Thus, for initial stages, the important parameter is to have a high bandwidth so that there is little delay in amplifying the signal and passing it on to the next stage. However, at the end of the cascade of amplifiers, it is more important to have a high slew rate capability so that the voltage across the interstage capacitors and the load capacitor rises or falls quick enough. Therefore, the stages at the beginning should be designed differently than the stages at the end of the amplifier chain.

The basic principle behind the high-speed comparator is to use a preamplifier to build up the input change to a sufficiently large value and then apply it to the latch. This combines the best aspects of circuits with a negative exponential response (the preamplifier) with circuits with a positive exponential response (the latch). This is illustrated in Fig. 8.6-2 where the time-domain response for a preamplifier and a latch are illustrated. In this figure, the gain of the preamplifier times the input voltage is not sufficient to reach the desired output level. Rather, during time  $t_1$  the preamplifier amplifies the input voltage to a value of  $V_X$ . The voltage  $V_X$  is applied to the latch input, which then goes to the desired output voltage in time  $t_2$ . Thus, the total response time is  $t_1 + t_2$ . If the comparator consisted only of the preamplifier,

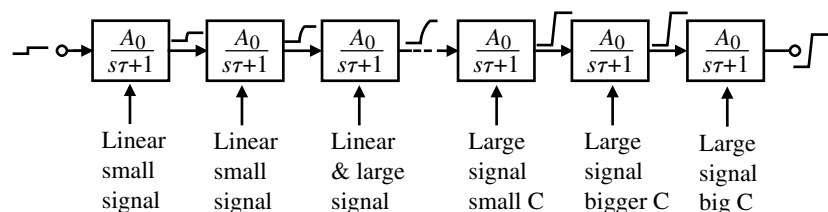


**Figure 8.6-1** Conceptual illustration of a cascaded comparator.

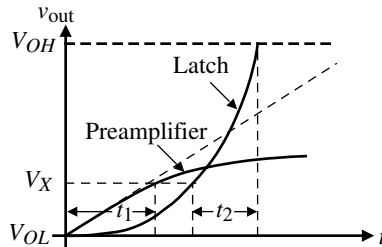


Figure 8.6-2 Preamplifier and latch step response.

the gain would have to be larger and the delay to make the transition from  $V_{OL}$  to  $V_{OH}$  would be longer than  $t_1 + t_2$ . On the other hand, the latch would require more time than  $t_1 + t_2$  if the input was small. We saw from Fig. 8.5-5 that the larger the input to the latch the shorter the time for the output to reach its maximum value.

The design of the preamplifier must be done in such a manner that the desired latch input voltage,  $V_X$ , is achieved in minimum time. Since the preamplifier is working in the linear region, this means that the bandwidth must be as large as possible. We know that the gain bandwidth of an amplifier is normally constant. Therefore, a single amplifier has a limited capability. If a number of low-gain, wide-bandwidth amplifiers are cascaded, the delay time,  $t_1$ , can be minimized. In fact, it has been shown that the optimum number of identical low-gain amplifiers is six each with a gain of 2.72. However, this optimum is very broad and three amplifiers each with a gain of 6 gives equally good results with less area [9].

A high-speed comparator using three cascaded low-gain amplifiers as the preamplifier and a latch at the output is shown in Fig. 8.6-3. When the FB and Reset switches are closed, the capacitors designated as  $C_v$  are autozeroed for each amplifier. Unfortunately, each amplifier must be autozeroed by itself. More switches would be required to autozero all three amplifiers at the same time. The input is applied through the capacitors  $C_1$  and  $C_2$ . For high-speed applications, the clocks to the comparator of Fig. 8.6-3 can be as high as 100 MHz, resulting in a comparator that can run at 100 Msps. As the sample time is increased, the power consumption of the comparator also increases.

The low-gain preamplifiers must compromise between a high bandwidth and sufficient gain. A simple preamplifier is shown in Fig. 8.6-4. The connection with the latch (for the last preamplifier) is shown. The gain of this preamplifier is

$$A_v = g_{m1}/g_{m3} = g_{m2}/g_{m4} \quad (8.6-1)$$

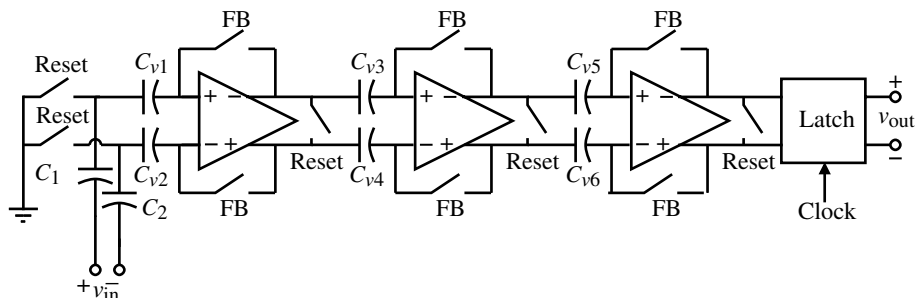


Figure 8.6-3 Fully differential, three-stage comparator and latch.

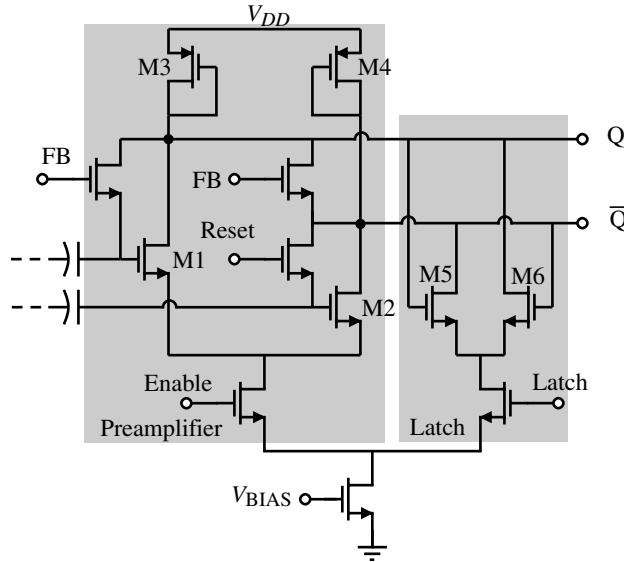


**Figure 8.6-4** Example of a pre-amplifier and latch.

The dominant pole is given by

$$|p_{\text{dominant}}| = \frac{g_{m3}}{C} = \frac{g_{m4}}{C} \quad (8.6-2)$$

where  $C$  is the capacitance seen from the output nodes to ground. If the bias current in the transistors is  $25 \mu\text{A}$  and the  $W/L$  values of M1 and M2 are 100 and M3 and M4 are 1, then the gain is  $3.85 \text{ V/V}$  and if  $C = 0.5 \text{ pF}$ , the pole is  $10^8 \text{ rad/s}$  or  $15.9 \text{ MHz}$ . We see that it is important to keep the pole as large as possible. Sometimes it is necessary to buffer the output of the amplifier from the capacitive load of the next stage. Unfortunately, the buffer will have a gain loss  $2\text{--}3 \text{ dB}$ , which is undesirable.

There are several problems with the preamplifier of Fig. 8.6-4. One is that the gain is very small even for large differences of  $W/L$  values. Another is that there is no isolation between the latch outputs and the inputs to the preamplifier. Rapid changes in the output of the latch can propagate through the drain-gate capacitances of M1 and M2 and appear at the input of the latch. Figure 8.6-5 shows a preamplifier that solves these two problems. Transistors M5 and M6 are used to increase the current in M1 and M2 so that the gain is enhanced by the square root of the difference of currents in M1 and M2 to the currents in M3 and M4. This can be illustrated as follows. The gain of the preamplifier was given in Eq. (8.6-1) and can be expressed further as

$$A_v = -\frac{g_{m1}}{g_{m3}} = -\sqrt{\frac{K'_N(W_1/L_1)I_1}{K'_P(W_3/L_3)I_3}} = -\sqrt{\frac{K'_N(W_1/L_1)}{K'_P(W_3/L_3)}} \sqrt{1 + \frac{I_5}{I_3}} \quad (8.6-3)$$

If  $I_5$  is greater than  $I_3$ , the gain can be enhanced by the square root of 1 plus the ratio of  $I_5$  to  $I_3$ . If  $I_5 = 24I_3$ , the gain is boosted by a factor of 5. Transistors M7 and M8 isolate the input from rapid changes in the latch output.

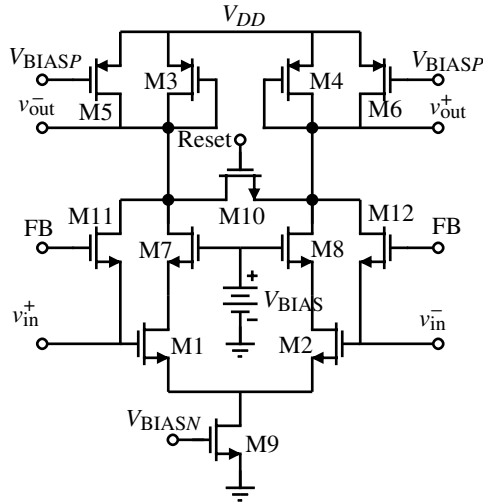


Figure 8.6-5 An improved preamplifier.

The use of a preamplifier before the latch also has the advantage of reducing the input-offset voltage of the latch by the gain of the preamplifier. The input-offset voltage of the comparator will now become that of the preamplifier, which can be autozeroed, resulting in small values of input-offset voltage.

The preamplifier can be replaced by a charge-transfer circuit resulting in simplification of the preamplifier [8]. A simple charge-transfer circuit is shown in Fig. 8.6-6(a), where the capacitor  $C_T$  is larger than  $C_O$ . This charge-transfer amplifier has three phases of operation. The first phase of operation is a reset phase, where switch S1 is closed discharging  $C_T$ . The second phase is the precharge phase shown in Fig. 8.6-6(b), where switch S2 is closed. In this phase, the capacitor  $C_T$  is charged to  $v_{in} - V_T$  and the output capacitor,  $C_O$ , is charged to the preset voltage,  $V_{PR}$ . The third phase is the amplification phase and is shown in Fig. 8.6-6(c), where both switches are open. In this phase, the input has been assumed to change by an amount  $\Delta V$ . This causes a current to flow in  $C_T$ , which also flows in  $C_O$ . Since the charge change must be equal for both  $C_T$  and  $C_O$ , the change across  $C_O$  is  $-(C_T/C_O)\Delta V$ . If  $C_T$  is greater than  $C_O$ , the input change,  $\Delta V$ , is amplified at the output by the factor of  $C_T/C_O$ .

The charge-transfer amplifier has several problems that must be solved. The first is that only positive values of voltage will be amplified and the second is that large offset voltages result as a function of the subthreshold current. A circuit that solves these problems is shown in Fig. 8.6-7 and uses both NMOS and PMOS transistors. Switches S3 are used to prevent a

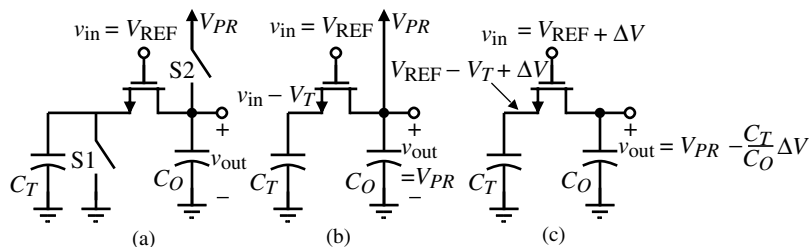


Figure 8.6-6 (a) Charge-transfer amplifier. (b) Precharge phase. (c) Amplification phase.

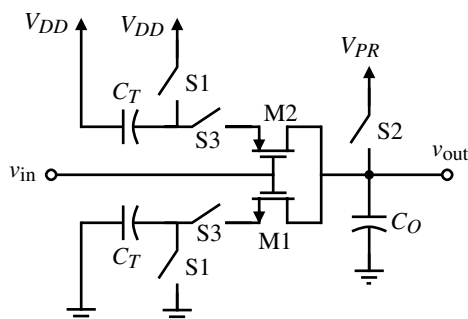


**Figure 8.6-7** A CMOS charge-transfer preamplifier.

current path during the reset phase. This circuit was used in cascade with a dynamic latch to function as a comparator up to 20 Ms/s with 8 bit linearity and a power dissipation of less than  $5\ \mu\text{W}$ . One of the limitations of the charge-transfer comparator is charge feedthrough. Dummy switches have been used to cancel some of the effects of charge feedthrough that occur when the various switches are opened.

When a comparator must drive a significant amount of output capacitance in very short times, the latch is generally not sufficient. In this case it is advisable to follow the latch by circuits that can quickly generate large amounts of current. A high-speed comparator following these principles is shown in Fig. 8.6-8 [10]. The first stage is a low-gain, high-bandwidth preamplifier that drives a latch. The latch outputs are used to drive a self-biased differential amplifier. The output of the self-biased differential amplifier drives a push-pull output driver.

The comparator of Fig. 8.6-8 was designed to have a propagation time delay of 10 ns with a 5 pF load capacitor and a 10 mV overdrive. It is interesting to note that the comparator is synchronous (clocks are not used). The gain of the comparator in midswing is greater than 2000 V/V and the quiescent current is 100  $\mu\text{A}$ . More design details can be found in Baker et al. [10].

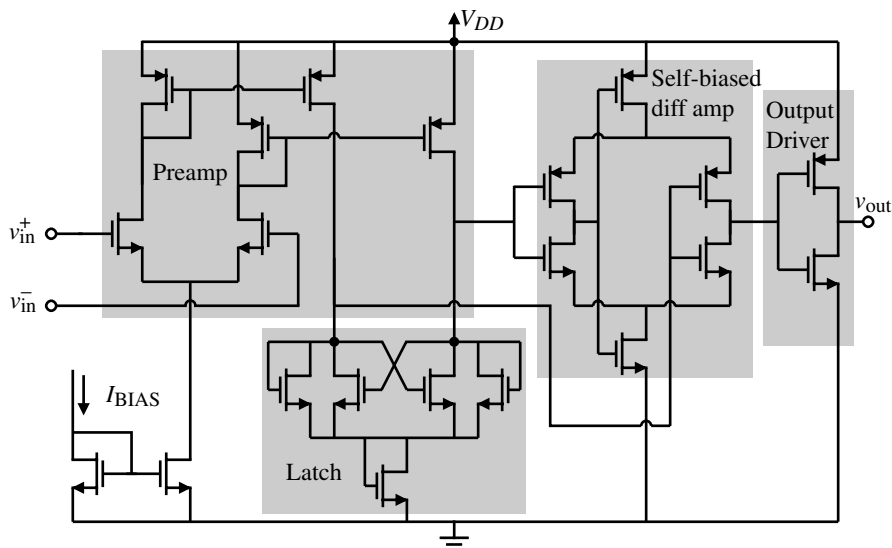


**Figure 8.6-8** A high-speed comparator using a cascade of a preamplifier, latch, self-biased differential amplifier, and output driver.



## 8.7 Summary

This chapter has introduced the comparator function along with the CMOS implementation of that function. The development of the CMOS comparator relied on previous work covering basic CMOS circuits in Chapters 4 and 5. The op amp of Chapter 6 makes an excellent open-loop, high-gain comparator realization. The characterization of the two-stage, open-loop comparator illustrated the performance capabilities. It was seen that two modes of response can exist in a linear comparator. One is small signal and the other is large signal. The difference is whether or not slewing occurs. When a comparator does not slew (small signal) the bandwidth is key to reducing the propagation delay time. However, when the comparator slews, then current sinking and sourcing capability is crucial to fast operation.

The performance of the open-loop comparator can be improved by the use of hysteresis to remove the influence of a noisy input signal. Autozeroing can be used in comparators to reduce the input-offset voltage. It is necessary that the comparator be stable in the unity-gain mode during autozeroing, which can be difficult to accomplish in some types of comparators. It was seen that self-biased comparators were always stable in the autozeroing operation. Although we did not examine the noise of the comparator, it would be analyzed in the same manner as was done for op amps. We assume that the noise, which is a linear phenomenon, is important during the switching of the comparator, which is in the linear region.

In many comparator applications, the signal is discrete time rather than continuous. In this case it is possible to use regenerative circuits as comparators. Unfortunately, the transient response of regenerative circuits is characterized by a positive exponential argument. This means that if the input signal is small a long time will be required for the signal to reach the region of the exponential response where the slope is steep. A solution for this problem is to cascade the latch with a preamplifier. The function of the preamplifier is to quickly build up the input to the latch so that the slow rate of rise of the exponential response can be avoided. This has resulted in a comparator capable of operating up to 20 Ms/s and greater.

The comparator will play an important role in the remainder of the text, particularly in the area of analog–digital converters. It will be one of the major factors determining how accurately and how quickly analog signal processing can be accomplished.

## Problems

- |                                                                                                                                                                                            |                                                                                                                                                                                                                                                                                               |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>8.1-1. Give the equivalent figures for Figs. 8.1-2, 8.1-4, 8.1-6, and 8.1-9 for an inverting comparator.</p>                                                                            | <p>8.1-3. Repeat Example 8.1-1 if the pole of the comparator is <math>-10^5</math> rad/s rather than <math>-10^3</math> rad/s.</p>                                                                                                                                                            |
| <p>8.1-2. Draw the first-order time response of an inverting comparator with a <math>20\ \mu\text{s}</math> propagation delay time. The input is described by the following equations:</p> | <p>8.1-4. What value of <math>V_{\text{in}}</math> in Example 8.1-1 will give a slewing response?</p>                                                                                                                                                                                         |
| $v_{\text{in}} = 0 \quad \text{for } t < 5\ \mu\text{s}$                                                                                                                                   | <p>8.2-1. Repeat Example 8.2-1 for the two-stage comparator of Fig. 8.2-5.</p>                                                                                                                                                                                                                |
| $v_{\text{in}} = 5(t - 5\ \mu\text{s}) \quad \text{for } 5\ \mu\text{s} < t < 7\ \mu\text{s}$                                                                                              | <p>8.2-2. If the poles of a two-stage comparator are both equal to <math>-10^7</math> rad/s, find the maximum slope and the time it occurs if the magnitude of the input step is <math>10V_{\text{in(min)}}</math>. What must be the <math>SR</math> of this comparator to avoid slewing?</p> |
| $v_{\text{in}} = 10 \quad \text{for } t > 7\ \mu\text{s}$                                                                                                                                  |                                                                                                                                                                                                                                                                                               |

- 8.2-3. Repeat Example 8.2-3 if  $p_1 = -5 \times 10^6$  rad/s and  $p_2 = -10 \times 10^6$  rad/s.
- 8.2-4. For Fig. 8.2-5, find all of the possible initial states listed in Table 8.2-1 of the first-stage output voltage and the comparator output voltage.
- 8.2-5. Calculate the trip voltage for the comparator shown in Fig. 8.2-4. Use the parameters given in Table 3.1-2. Also,  $(W/L)_2 = 100$  and  $(W/L)_1 = 10$ .  $V_{BIAS} = 1$  V,  $V_{SS} = 0$  V, and  $V_{DD} = 4$  V. Compute the worst-case variations of the trip voltage assuming a  $\pm 10\%$  variation on  $V_T$ ,  $K'$ ,  $V_{DD}$ , and  $V_{BIAS}$ .
- 8.2-6. Sketch the output response of the circuit in Problem 8.2-5, given a step input that goes from 4 to 1 V. Assume a 10 pF capacitive load. Also assume the input has been at 4 V for a very long time. What is the delay time from the step input to when the output changes logical (CMOS) states?
- 8.2-7. Repeat Example 8.2-5 with  $v_{G2}$  constant and the waveform of Fig. 8.2-6 applied to  $v_{G1}$ .
- 8.2-8. Repeat Example 8.2-5 using the two-stage op amp designed in Example 6.3-1 if the compensation capacitor is removed.
- 8.2-9. Repeat Example 8.2-6 if the propagation time is  $t_p = 25$  ns.
- 8.2-10. An open-loop comparator has a gain of  $10^4$ , a dominant pole of  $10^5$  rad/s, a slew rate of 5 V/ $\mu$ s, and an output swing of 1 V. (a) If  $V_{in} = 1$  mV, find the propagation delay time of this comparator (the time for the output to go halfway from one state to the other). (b) Repeat part (a) if  $V_{in} = 10$  mV. (c) Repeat part (a) if  $V_{in} = 100$  mV.
- 8.2-11. The comparator shown in Fig. P8.2-11 has an input applied as shown. Assuming the pulse width is wide enough, calculate the propagation delay time for this comparator. Assume that the trip point of the output is at 0 V.
- 8.2-12. Design a comparator given the following requirements:  $P_{diss} < 2$  mW,  $V_{DD} = 3$  V,  $V_{SS} = 0$  V,  $C_{load} = 3$  pF,  $t_{prop} < 1$   $\mu$ s,  $ICMR = 1.5$ – $2.5$  V,  $A_{v0} > 2200$ , and output voltage swing within 1.5 V of either

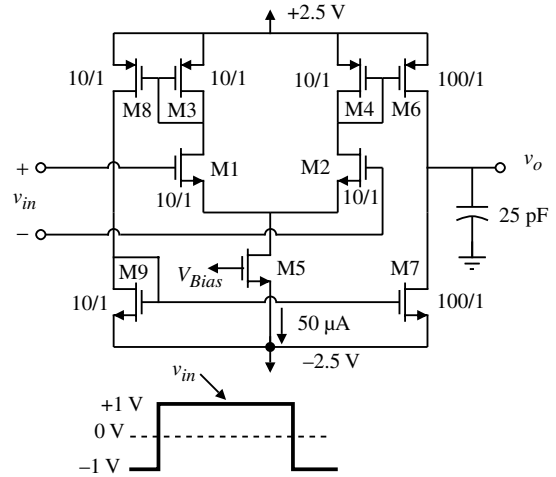


Figure P8.2-11

- rail. Use Tables 3.1-2 and 3.3-1. Use 1  $\mu$ m channel lengths for all transistors.
- 8.3-1. Assume that the dc current in M5 of Fig. 8.3-1 is 100  $\mu$ A. If  $W_6/L_6 = 5(W_4/L_4)$  and  $W_8/L_8 = 5(W_3/L_3)$ , what is the slew rate limited propagation delay time of this comparator if  $C_L = 10$  pF and  $V_{DD} = -V_{SS} = 2$  V?
- 8.3-2. If the folded-cascode op amp of Example 6.5-3 is used as a comparator, find the dominant pole if  $C_L = 5$  pF. If the input step is 10 mV, determine whether the response is linear or slewing and find the propagation delay time.
- 8.3-3. Find the open-loop gain of Fig. 8.3-3 if the two-stage op amp is the same as Example 6.3-1 without the compensation and  $W_{10}/L_{10} = 10(W_8/L_8) = 100(W_6/L_6)$ ,  $W_9/L_9 = (K'_P/K'_N)(W_8/L_8)$ ,  $W_{11}/L_{11} = (K'_P/K'_N)(W_{10}/L_{10})$  and the quiescent current in M8 and M9 is 100  $\mu$ A and in M10 and M11 is 500  $\mu$ A. What is the propagation time delay if  $C_{II} = 100$  pF and the step input is large enough to cause slewing?
- 8.3-4. Figure P8.3-4 shows a circuit called a clamped comparator. Use the parameters of Table 3.1-2 and calculate the gain of this comparator. What is the positive and negative slew rate of this comparator if the load capacitance is 5 pF?
- 8.4-1. If the comparator used in Fig. 8.4-1 has a dominant pole at  $10^4$  rad/s and a gain of  $10^3$ ,

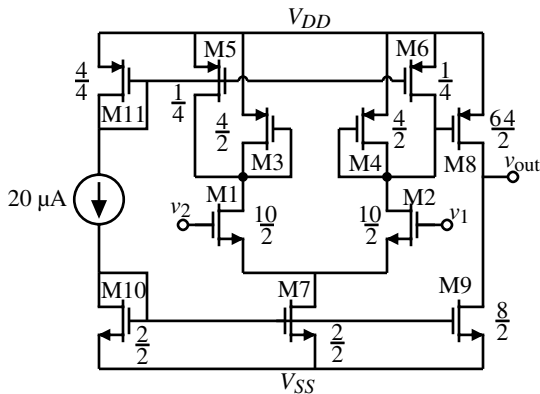


Figure P8.3-4

how long does it take  $C_{AZ}$  to charge to 99% of its final value,  $V_{OS}$ ? What is the final value that the capacitor,  $C_{AZ}$ , will charge to if left in the configuration of Fig. 8.4-1(b) for a long time?

- 8.4-2. Use the circuit of Fig. 8.4-9 and design a hysteresis characteristic that has  $V_{TRP}^- = 0$  V and  $V_{TRP}^+ = 1$  V if  $V_{OH} = 2$  V and  $V_{OL} = 0$  V. Let  $R_1 = 100$  k $\Omega$ .
- 8.4-3. Repeat Problem 8.4-2 for Fig. 8.4-10.
- 8.4-4. Assume that all transistors in Fig. 8.4-11 are operating in the saturation mode. What is the gain of the positive-feedback loop, M6–M7, using the  $W/L$  values and currents of Example 8.4-2?
- 8.4-5. Repeat Example 8.4-1 to design  $V_{TRP}^+ = -V_{TRP}^- = 0.5$  V.
- 8.4-6. Repeat Example 8.4-2 if  $i_5 = 50$   $\mu$ A. Confirm using a simulator.
- 8.4-7. Review the development of the switching points for the comparator of Fig. 8.4-11. Discuss why the simulation data in Fig. 8.4-13 does not match closely with the calculations of Example 8.4-2.
- 8.5-1. List the advantages and disadvantages of the switched capacitor comparator of Fig. 8.5-1 over an open-loop comparator having the same gain and frequency response.
- 8.5-2. If the current and  $W/L$  values of the two latches in Fig. 8.5-3 are identical, which latch will be faster? Why?

- 8.5-3. Repeat Example 8.5-1 if  $\Delta V_{out} = (0.5)(V_{OH} - V_{OL})$ .
- 8.5-4. Repeat Example 8.5-1 if the dc latch current is 50  $\mu$ A.
- 8.5-5. Redevelop the expression for  $\Delta V_{out}/\Delta V_i$  for the circuit of Fig. P8.5-5, where  $\Delta V_{out} = v_{o2} - v_{o1}$  and  $\Delta V_i = v_{i1} - v_{i2}$ .

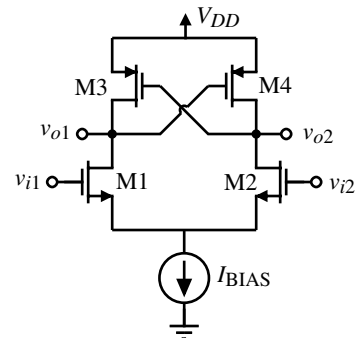


Figure P8.5-5

- 8.5-6. Compare the dynamic latch of Fig. 8.5-8 with the NMOS and PMOS latches of Fig. 8.5-3. What are the advantages and disadvantages of the two latches?
- 8.5-7. Use the worst-case values of the transistor parameters in Table 3.1-2 and calculate the worst-case voltage offset for the NMOS latch of Fig. 8.5-3(a).
- 8.5-8. Consider the latch shown in Fig. P8.5-8. Assume the capacitors connected to the drains of M1 and M2 ( $C_1$  and  $C_2$ ) are initially discharged. Express  $\Delta V_{out} = v_{o2} - v_{o1}$  as

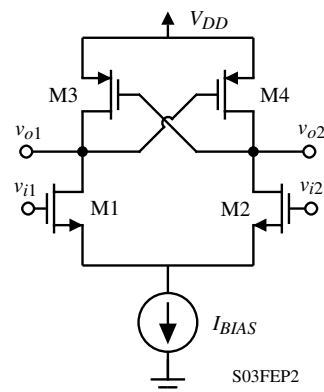


Figure P8.5-8

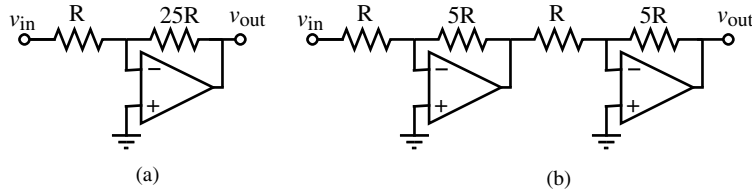


Figure P8.6-1

a function of the applied input,  $\Delta V_{in} = v_{i1} - v_{i2}$ , in the time domain assuming  $\Delta V_{in}$  is a step input. If  $g_{m1} = g_{m2} = 1 \text{ mS}$ ,  $g_{m3} = g_{m4} = 100 \mu\text{S}$ , and  $C_1 = C_2 = 1 \text{ pF}$ , what is the propagation delay time ( $\Delta V_{out} = 0.5(V_{OH} - V_{OL})$ ) for a step input of  $\Delta V_{in} = 0.01(V_{OH} - V_{OL})$ ?

- 8.6-1. Assume an op amp has a low-frequency gain of 1000 V/V and a dominant pole at  $-10^4 \text{ rad/s}$ . Compare the  $-3 \text{ dB}$  bandwidths of the configurations in Fig. P8.6-1 using this op amp.
- 8.6-2. What is the gain and  $-3 \text{ dB}$  bandwidth (in Hz) of Fig. P8.6-2 if  $C_L = 1 \text{ pF}$ ? Ignore reverse-bias voltage effects on the pn junctions and assume the bulk-source and bulk-drain areas are given by  $W \times 5 \mu\text{m}$ .

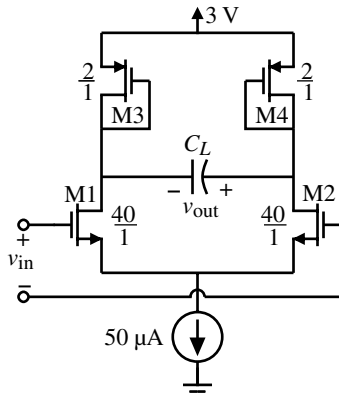


Figure P8.6-2

- 8.6-3. Repeat Problem 8.6-2 for Fig. P8.6-3. The  $W/L$  ratios for M1 and M2 are  $10 \mu\text{m}/1 \mu\text{m}$  and for the remaining PMOS transistors the  $W/L$  ratios are all  $2 \mu\text{m}/1 \mu\text{m}$ .

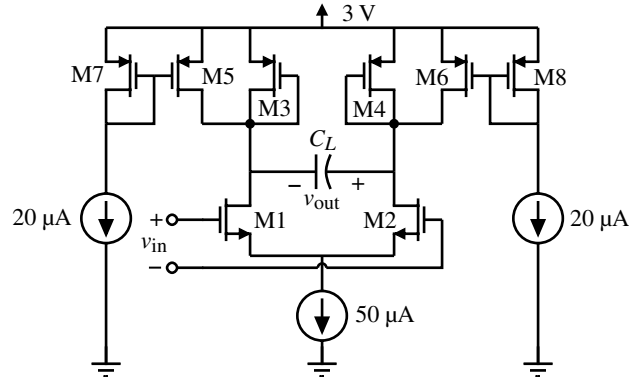


Figure P8.6-3

- 8.6-4. Assume that a comparator consists of an amplifier cascaded with a latch. Assume the amplifier has a gain of 5 V/V and a  $-3 \text{ dB}$  bandwidth of  $1/\tau_L$ , where  $\tau_L$  is the latch time constant. Find the normalized propagation delay time for the overall configuration if the applied input voltage is  $0.05(V_{OH} - V_{OL})$  and the voltage applied to the latch is (a)  $\Delta V_i = 0.05(V_{OH} - V_{OL})$ , (b)  $\Delta V_i = 0.1(V_{OH} - V_{OL})$ , (c)  $\Delta V_i = 0.15(V_{OH} - V_{OL})$ , and (d)  $\Delta V_i = 0.2(V_{OH} - V_{OL})$ . From your results, what value of  $\Delta V_i$  would give minimum propagation delay time?
- 8.6-5. Assume that a comparator consists of two identical amplifiers cascaded with a latch. Assume the amplifier has the characteristics given in Problem 8.6-4. What would be the normalized propagation delay time if the applied input voltage is  $0.05(V_{OH} - V_{OL})$  and the voltage applied to the latch is  $\Delta V_i = 0.1(V_{OH} - V_{OL})$ ?
- 8.6-6. Repeat Problem 8.6-5 if there are three identical amplifiers cascaded with a latch. What

would be the normalized propagation delay time if the applied input voltage is  $0.05(V_{OH} - V_{OL})$  and the voltage applied to the latch is  $\Delta V_i = 0.2(V_{OH} - V_{OL})$ ?

- 8.6-7. A comparator consists of an amplifier cascaded with a latch as shown in Fig. P8.6-7. The amplifier has voltage gain of 10 V/V and  $f_{-3\text{ dB}} = 100\text{ MHz}$  and the latch has a time constant of 10 ns. The maximum and minimum voltage swings of the amplifier and

latch are  $V_{OH}$  and  $V_{OL}$ . When should the latch be enabled after the application of a step input to the amplifier of  $0.05(V_{OH} - V_{OL})$  to get minimum overall propagation time delay? What is the value of the minimum propagation time delay? It may be useful to recall that the propagating time delay of the latch is given as  $t_p = \tau_L \ln\left(\frac{V_{OH} - V_{OL}}{2v_{il}}\right)$ ,

where  $v_{il}$  is the latch input ( $\Delta V_i$  of the text).

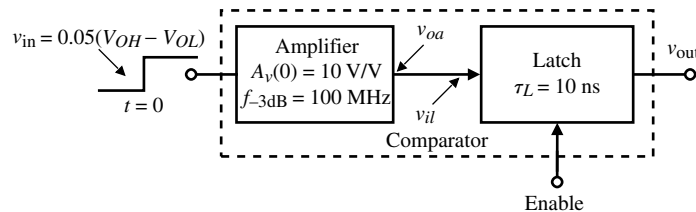


Figure P8.6-7

## References

1. D. J. Allstot, "A Precision Variable-Supply CMOS Comparator," *IEEE J. Solid-State Circuits*, Vol. SC-17, No. 6, pp. 1080–1087, Dec. 1982.
2. M. Bazes, "Two Novel Full Complementary Self-Biased CMOS Differential Amplifiers," *IEEE J. Solid-State Circuits*, Vol. 26, No. 2, pp. 165–168, Feb. 1991.
3. J. Millman and C. C. Halkias, *Integrated Electronics: Analog and Digital Circuits and Systems*. New York: McGraw-Hill, 1972.
4. A. S. Sedra and K. C. Smith, *Microelectronic Circuits*, 4th ed. New York: Oxford University Press, 1998.
5. J. Millman and H. Taub, *Pulse, Digital, and Switching Waveforms*. New York: McGraw-Hill, 1965.
6. T. B. Cho and P. R. Gray, "A 10b, 20 Msamples/s, 35 mW Pipeline A/D Converter," *IEEE J. Solid-State Circuits*, Vol. 30, No. 3, pp. 166–172, Mar. 1995.
7. A. L. Coban and P. E. Allen, "A 1.5 V, 1 mW Audio  $\Delta\Sigma$  Modulator with 98 dB Dynamic Range," *Proc. Int. Solid-State Circuit Conf.*, pp. 50–51, Feb. 1999.
8. K. Kotani, T. Shibata, and T. Ohmi, "CMOS Charge-Transfer Preamplifier for Offset-Fluctuation Cancellation in Low-Power A/D Converters," *IEEE J. Solid-State Circuits*, Vol. 33, No. 5, pp. 762–769, May 1998.
9. J. Doernberg, P. R. Gray, and D. A. Hodges, "A 10-bit 5-Msample/s CMOS Two-Step Flash ADC," *IEEE J. Solid-State Circuits*, Vol. 24, No. 2, pp. 241–249, Apr. 1989.
10. R. J. Baker, H. W. Li, and D. E. Boyce, *CMOS Circuit Design, Layout, and Simulation*. Piscataway, NJ: IEEE Press, 1998, Chap. 26.