Solid-State Electronic Circuits

Edward W. Maby

University of Southern California

ISBN: 978-0-615-71805-3 pdf

This book is dedicated to the memory of my parents

Ernest and Judith (Halliday) Maby

Preface

Black Box, *n.* A usually complicated electronic device that functions and is packaged as a unit and whose internal mechanism is usually hidden from or mysterious to the user. (Merriam-Webster)

"Black Box" probably originates from the 1940s as Royal Air Force slang for a type of radar that allowed pilots to see through clouds or in the dark. This text characterizes the Black Box as an *Integrated Circuit* that hopefully becomes less mysterious or "gray" through careful study of what is inside. Nevertheless,

Engineers design *with* integrated circuits—

only a relatively few design integrated circuits.

Circuit formats and especially the performance limitations of modern solid-state devices that comprise them are crucial for understanding the fragility of black-box operational rules, not just methods of integrated-circuit design. An attempt to balance electronic perspectives from both sides of the box imparts a radical similarity to classic texts on electronics.

While acclamations of superior clarity, design orientation, and computer use are welcome, the primary development from within the box is an up-to-date and undergraduate-friendly repertoire of integrated-circuit designs that occasionally feature deep-sub-micron devices. The exterior viewpoint has support from *Getting on Board* sidebars at the ends of early chapters. Practical board-level design that confronts corruptive environmental factors is a unifying theme of this endeavor.

Prerequisites

Circuit Analysis: It is presumed that students have completed a course involving Kirchhoff's equations and Ohm's law, that they are comfortable with series or parallel reductions, node-voltage analysis, and superposition, and that they can use these concepts to find Thevenin/Norton equivalents. Familiarity with analysis in the time and frequency domains is also assumed; however, the degree to which students find comfort is less important.

Computer Analysis: Most students will have been exposed to some form of computer circuit analysis such as SPICE. This text relies on computer simulations to demonstrate the validity of designs *after* hand calculations. Frequent examples develop appropriate file formats.

Device Physics: Electrical Engineering programs usually have one course devoted to physical electronics, material that is *not* prerequisite to this text. The physical basis for various devices has been incorporated to substantiate often complicated SPICE models. It can be treated as background reading, synergetic material for a *corequisite* course, or reference. In the long term, physical "devices" courses will be under pressure to have greater breadth so that students are exposed to an assortment of electronic materials, not only semiconductors, for systems such as displays and magnetic media.

Text Organization

Although not explicitly enumerated as such, *Solid-State Electronic Circuits* is organized in multiple parts that are separated by consise *Perspectives*. The latter provide periodic big-picture assessments of material to follow.

Chapter 1 presents analog and digital integrated circuits as "black boxes" that are subject to idealized interconnection rules. The principle objective is to have students "doing" electronics right away using elementary circuit analysis and common sense. As examples unfold, students are encouraged to view circuits from both sides of the box—one needs to understand lower and lower functional levels in order to appreciate the overall performance.

Perspective (A) shows a graphical method for solving non-linear circuits, and it motivates diode function as an effective "one-way street" for current. Chapter 2 establishes physical foundations that support approximations to ideal diode behavior as well as more complicated devices in later chapters. Chapter 3 exposes students to non-linear analytical techniques that require assumptions concerning a diode model and verification following a solution. The discussion emphasizes the method of "marking up a circuit diagram" to identify currents and voltages without relying on systematic node equations (unless absolutely necessary). Most students will find themselves immersed in unfamiliar patterns of thought that will prove useful throughout the text. Chapter 4 covers several practical diode applications. Much of this material (rectifier power-conversion circuits, for example) is conceptually very old, yet standard fare for undergraduates. To stimulate interest, the discussion extends to dc-to-dc converters and integrated circuits that control them. Similarly, the treatment of Zener diodes transitions from voltage regulation as the actual regulator to voltage regulation through a reference potential. Bandgap references are subsequently applied. Signal conditioning circuits (clipping and clamping) and integrated optoelectronic applications conclude the chapter as discretionary topics.

Perspective (B) identifies the transistor as a "valve" for current control. Chapters 5 and 6 subsequently describe the physical behavior of MOSFET and bipolar devices, respectively. (The chapters can be read in either order.) Both transistors have complex behavioral models. Nevertheless, the models in each chapter are confined to the support of first-order hand calculations. Complicating factors are presented later, as needed, in Chapters 9 and 10. Overviews of circuit applications are provided to promote student interest. Chapter appendices provide useful supplemental material.

Perspective (C) describes the amplifier mode of transistor operation in which linearly-dependent small-signal perturbations are superimposed over a non-linear operating point. Favorable device attributes are enumerated. Chapter 7 examines biasing processes that promote a stable operating point and performance metrics for three basic transistor amplifier configurations. Chapter 8 considers limiting amplifier behavior at high and low frequencies and design methods that extend or localize (tune) high-frequency operation. Apart from biasing details, which are unique to the type of transistor used, both chapters pursue general analytical procedures that determine amplifier characteristics in terms of a few ubiquitous small-signal device parameters and various resistances looking into or away from the transistor terminals. This approach leads to a small but powerful set of analytical expressions that are easily applied with minimal algebraic haze.

An "Interlude" exhibits integrated-circuit fabrication processes and layout design rules for a typical CMOS technology. In turn, Chapter 9 probes the interior of a CMOS operational amplifier and factors leading to its non-ideal performance. Chapter 10 addresses the influence of process and layout at the digital integrated-circuit exterior, strategies for the implementation of simple logical functions at the IC interior, and static/dynamic inverter design for deep sub-micron processes. Together, Chapters 9 and 10 serve as introductions to more advanced courses in mixed-signal VLSI design.

Perspective (D) motivates four fundamental forms of analog feedback. Chapter 11 examines each form separately. Nevertheless, general analytical tools avoid the otherwise tedious application of two-port system parameters. Chapter 12 concerns the identification of dynamic feedback effects that have the potential to disrupt the circuits of Chapter 11. Compensation methods are used to mitigate these effects for successful board-level op-amp design. Deliberate promotion of instability is the basis for oscillator presentation.

The remaining chapters address topics of importance for mixed-signal electronic systems. Chapter 13 (Conditionings and Corruptions) explores analog filters, digital processes for data communications over serial links, noise, distortion, and transmission-line effects. Chapter 14 (Conversions) examines digital-to-analog and analog-to-digital conversions. The objective is to guide judicious assessments of numerous design alternatives.

Reading Aids

Concept Summaries at the end of each chapter outline issues that students are expected to understand. Whereas Chapters 2, 5, and 6 contain more than average physical discussion, Concept Summaries are provided at the end of each section. Occasional difficult or supporting material that is not crucial for a first reading is marked with a heavy blue line in the left margin.

Course Organization

It should come as no surprise that much of this text reflects an ever-evolving *Analog Electronics* course that the author has taught over 20+ years at USC and Rensselaer. The following Table suggests an ordering of topics with sectional support for a one-semester "analog" course.

Week	**Topic**	**Section**
1	Op-amp overview, non-ideal op-amps	1.1, 1.4
	Semiconductor principles (overview/review)	Ch. 2
2	Diode circuit analysis	Ch. 3
3	Power supplies	4.1
	Regulation, voltage references	4.2
4	Elementary MOSFETs, circuit overview (switch emphasis)	Ch. 5
5	Elementary BJTs, circuit overview ("valve" emphasis)	Ch. 6
6	Amplification strategies, biasing	7.1
7	Small-signal transistor models and analysis	7.2
	One-transistor "midfrequency" amplifiers	7.3
8	High-frequency transistor models	8.1
	High-frequency response	8.2
9	Circuits for improved or localized bandwidth	8.3
	Low-frequency response	8.4
10	CMOS fabrication processes, layout overview	Interlude
	Differential Amplifier	9.1

11	Analog Integrated-Circuit Complications	9.2
	Front-end op-amp design	9.3
12	Intermediate and output op-amp stages	9.4
	Elementary CMOS Inverter, CMOS inputs/outputs	10.1, 10.2
13	Series-shunt feedback	11.1
	Shunt-shunt feedback	11.2
14	Dynamic feedback effects, tests for stability	12.1
	Compensation methods	12.2, 12.3
15	Noise	13.4
	Distortion	13.5

Treatment of semiconductor principles, MOSFET and BJT overviews, and CMOS processes and layout should provide or reinforce broad concepts without becoming entangled in analytical detail. This form of presentation is not recommended for topics that are computationally intensive.

The following Table suggests an ordering of topics with sectional support for a one-semester "digital" course for computer engineering students.

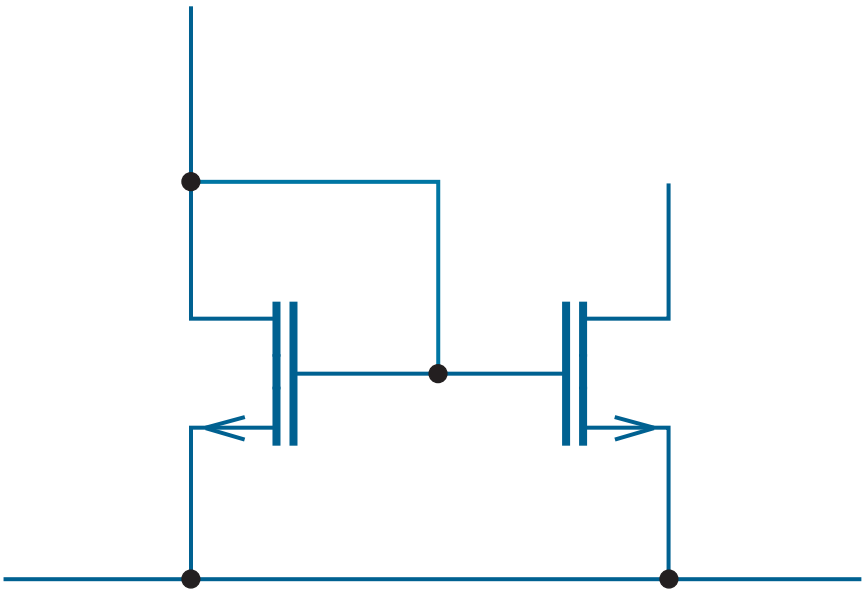
Week	**Topic**	**Section**
1	Digital and mixed-signal integrated circuits	1.2, 1.3
	Semiconductor principles (overview)	Ch. 2
2	Diode circuit analysis	Ch. 3
3	Power supplies	4.1
	Regulation, voltage references	4.2
4	Elementary MOSFETs, circuit overview (switch emphasis)	Ch. 5
5	Elementary BJTs, circuit overview (include brief consideration of saturation logic)	Ch. 6
6	CMOS fabrication processes, design rules	Interlude
	Elementary CMOS inverter	10.1
7	Layout exercise	Interlude
	CMOS inputs/outputs	10.1

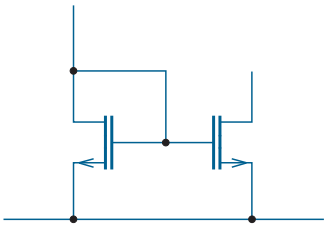
8	Combinational CMOS logic, memory	10.2
9	Static CMOS inverter design	10.3
	MOSFET capacitance	6.4
10	Dynamic CMOS inverter design	10.4
11	Op-amp overview	1.1
	Analog filters	13.1, 13.2
12	UARTs, RS-232	13.3
	Fully differential amplifiers, RS-485	13.3
13	Transmission-line effects	13.6
14	A/D and D/A conversions, errors, sampling	14.1, 14.2
	Nyquist-rate analog-to-digital converters	14.3
15	Oversampling analog-to-digital converters	14.4
	Digital-to-analog converters	14.6

A subset of these students will continue with a course in VLSI design.

Contents

Solid-State Electronic Circuits





Chapter 1

Black-Box Electronics

Just look at a circuit board in your personal computer and you will only begin to realize what amazing things can be done with *Integrated Circuits*—those little "chips" with the wire leads that look like feet. Your PC is not a trivial system. Nevertheless, it is remarkable that today's electronic hacker needs only some "black-box" connection rules, some $v = iR$ -variety circuit concepts, and some common sense to put together practical circuits that once would have required years of design experience. The integration of complex systems into functionally simple systems will certainly continue to be a highlight of electronic engineering throughout the 21st century.

This chapter examines a few powerful black-box circuit elements that (hopefully) set the stage for an appreciation of elementary electronic devices and concepts—not to mention more interesting end-of-chapter problems—throughout the remaining text.

Objectives

After studying this chapter and working through the problems, you should be able to perform each of the following:

- Design simple analog circuits using the ideal operational amplifier (Section 1.1).
- Design a comparator circuit that reacts to a preset threshold voltage (Section 1.1).
- Design a Schmitt trigger with specific hysteresis (Section 1.1).
- Design inverter, NAND, and NOR logical functions with ideal switches, and describe the related hierarchy of digital circuits (Section 1.2).
- Design simple pulse and asynchronous circuits using the 555 timer (Section 1.3).
- Simulate ideal op-amp circuits using SPICE (Section 1.4).

1.1 Analog: OpAmps and Comparators

An **analog** signal is a time-dependent current or voltage that extends over a *continuous* range of equally likely values. These signals often originate from some form of sensing device, such as a microphone (to detect sound) or a thermocouple (to detect temperature), and their ranges can be very small. Thus, we frequently require electronic circuits to process analog information into forms that are more easily accommodated.

Op-Amps

Perhaps the most useful "black box" for analog signal processing is the **operational amplifier** or "**op-amp**" shown in Fig. 1.1. This component features a **non-inverting input** (v^+), an **inverting input** (v^-), and a single output (v_{out}) that are related through the expression

$$v_{out} = A_{vd}\,(v^+ - v^-)\,, \tag{1.1}$$

where A_{vd} is a *very large* **differential voltage gain**. Meanwhile, the input currents i^+ and i^- are exceptionally small—so small, in fact, that under practical conditions,

$$i^+ \approx 0\,, \tag{1.2a}$$
$$i^- \approx 0\,. \tag{1.2b}$$

Finally, there are two power-supply connections V^+ and V^- that establish a range for the output voltage:

$$V^- \le v_{out} \le V^+ \tag{1.3}$$

—while providing output current, as needed. The power supplies play an important but unexciting role, so the connections are often omitted from an op-amp circuit diagram to avoid clutter. Yet be warned: neglect of the power supplies can lead to a misapplication of Kirchhoff's Current Law. Supply voltages are typically equal in magnitude, say ±5 V, but there are some op-amps that allow $V^- = 0$ (a ground connection).

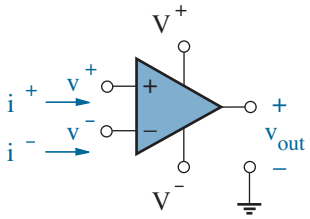


Figure 1.1: Operational amplifier with terminal voltages and currents.

To get a feel for the op-amp in action, consider the circuit of Fig. 1.2a and the equivalent circuit of Fig. 1.2b. In the latter, we model the op-amp as a dependent voltage source that is controlled by voltage $v_d = v^+ - v^-$. This voltage is observed across an internal open circuit to account for the zero-current input conditions of Eq. 1.2.

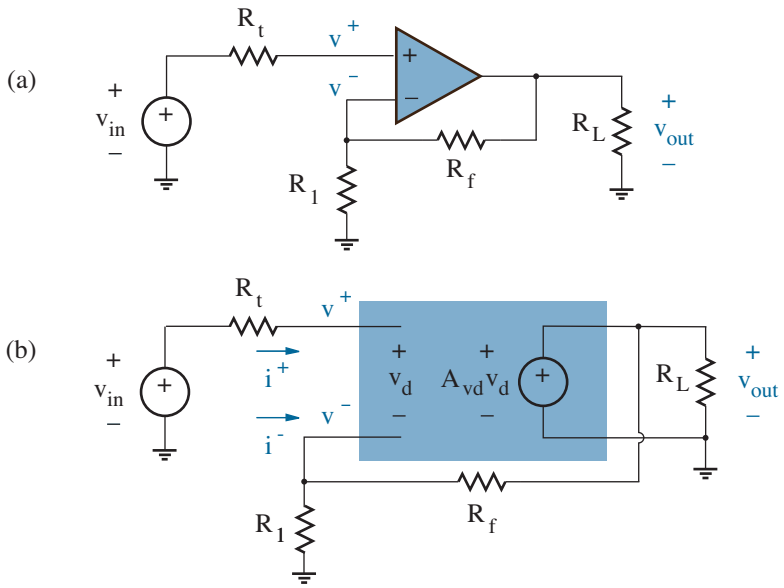


Figure 1.2: Trial op-amp circuit and its equivalent.

The circuit analysis is straightforward. With current $i^+ = 0$, $v^+ = v_{in}$. And upon application of a voltage divider relationship for v^-, we have

$$v_d = v_{in} - v_{out}\left(\frac{R_1}{R_1 + R_f}\right) \tag{1.4}$$

at the op-amp inputs. So with $v_{out} = A_{vd}\, v_d$, we find

$$v_{out} = \frac{A_{vd}\, v_{in}}{1 + A_{vd}\left(\dfrac{R_1}{R_1 + R_f}\right)} \,. \tag{1.5}$$

Then subject to $A_{vd}\, R_1/(R_1 + R_f) \gg 1$,

$$v_{out} \approx v_{in}\left(1 + \frac{R_f}{R_1}\right) \,. \tag{1.6}$$

Equation 1.6 is a *linear* relation that describes a **non-inverting amplifier** with a **voltage gain** v_{out}/v_{in} equal to one plus a *ratio* of two resistor values. It is valid provided v_{out} lies within the V^+ and V^- power-supply bounds. Any attempt to exceed these bounds leads to output signal **distortion**.

Despite the straightforward claim, it is much easier to work out the ideal performance of the preceding op-amp circuit (and most op-amp circuits) by observing two simple "black-box" rules:

Rule 1: Subject to **negative feedback**, in which the op-amp output is connected to the *inverting* input terminal through a set of circuit elements, *the non-inverting and inverting terminal voltages are ideally equal.*

This is apparent from Eq. 1.1. If v_{out} lies within the range of the op-amp power supplies and $A_{vd} \to \infty$, $v^+ - v^- = 0$.

Rule 2: *The non-inverting and inverting terminal currents are ideally zero.*

This is apparent from Eq. 1.2.

Back to the circuit of Fig. 1.2a. Following Rule 2, we neglect the op-amp input currents to find $v^+ = v_{in}$ and $v^- = v_{out}\, R_1/(R_1 + R_f)$ at the non-inverting and inverting terminals, respectively. Rule 1 requires $v^+ = v^-$. Thus,

$$v_{in} = v_{out} \left(\frac{R_1}{R_1 + R_f} \right) , \tag{1.7}$$

which rearranges to

$$v_{out} = v_{in} \left(1 + \frac{R_f}{R_1} \right) . \tag{1.8}$$

Much easier, indeed.

It is important to note that the circuit of Fig. 1.2 satisfies the negative-feedback restriction of Rule 1: the op-amp output is connected back to the inverting input terminal through R_f. We require this condition to ensure **stability** against undesired perturbations. For example, suppose there is spurious electrical "noise" that induces a small *increase* in the voltage v_{out}. Equation 1.4 shows that increased v_{out} produces a *decrease* in v_d. In turn, the dependent source decreases v_{out} to cancel the original perturbing factor. If we had used **positive feedback** with the non-inverting and inverting terminal connections interchanged, then in theory we would find

$$v_{out} = -A_{vd}\, v_{in} \left[1 - A_{vd} \left(\frac{R_1}{R_1 + R_f} \right) \right]^{-1} \approx v_{in} \left(1 + \frac{R_f}{R_1} \right) . \tag{1.9}$$

But the stability is lost. An increase in v_{out} leads to an increase in v_d, which, in turn, leads to an *increase* in v_{out}. The result is a runaway condition.

When viewed in terms of black-box rules, the non-inverting amplifier is perhaps most remarkable for what we do *not* observe:

- The voltage gain is independent of R_t, the Thevenin resistance that is linked to the input signal source. This reflects a **high-impedance** op-amp input that is consistent with $i^+ \approx 0$.
- The voltage gain is independent of R_L, the output "load" resistance. This reflects a **low-impedance** op-amp output. The circuit of Fig. 1.2 features zero Thevenin resistance when looking back from the load. The result generally applies when $A_{vd} \to \infty$ (see Problem 1.57).
- The behavior of the input is independent of conditions at the output. In this sense, the op-amp provides **buffer** action—a signal propagates from input to output, but not in the reverse direction.

The non-inverting amplifier has minimum unity voltage gain if $R_f = 0$ (as for a short circuit), and resistor R_1 is not needed to assert the minimum. These conditions apply to the special **unity-gain buffer** shown in Fig. 1.3. While the voltage gain is hardly exciting, the buffer action is often useful. For example, suppose $R_t = 99.9\ \text{k}\Omega$ and $R_L = 100\ \Omega$. If the Thevenin signal source is reconnected directly to the load, there is significant attenuation: $v_{out} = v_{in}\, R_L/(R_t + R_L) = v_{in}/1000$. The buffer allows $v_{out} = v_{in}$.

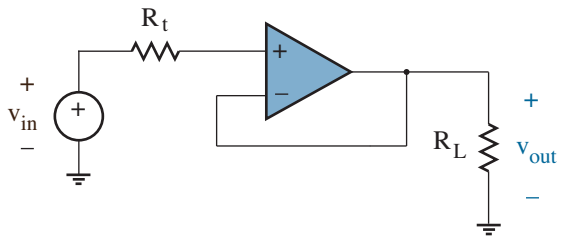


Figure 1.3: Unity-gain buffer.

The non-inverting amplifier exhibits a voltage gain that is dimensionless. Nevertheless, a common specification is in dB (decibels), specifically

$$\text{Voltage Gain}\,|_{\text{dB}} = 20\,\log_{10}\,|v_{out}/v_{in}|\,. \tag{1.10}$$

Thus, for example, a voltage gain of 10 is equivalent to 20 dB. In contrast, the amplifier **power gain** is specified in decibels through the relation

$$\text{Power Gain}\,|_{\text{dB}} = 10\,\log_{10}\,|P_{out}/P_{in}|\,, \tag{1.11}$$

where P_{out} and P_{in} are the output and input signal powers, respectively. The black-box non-inverting amplifier has *infinite* power gain.

Example 1.1

Determine v_{out} for the circuit of Fig. 1.4.

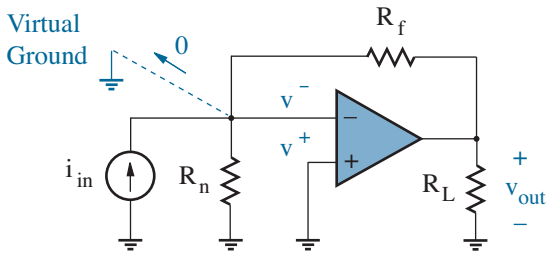


Figure 1.4: Circuit for Example 1.1.

Solution

The non-inverting op-amp terminal is connected to ground, so $v^+ = 0$. In turn, we apply Rule 1 to set $v^- = 0$, thereby forming a **virtual ground** with the same potential as actual ground at the inverting op-amp terminal. Next, we write Kirchhoff's Current Law at the same terminal:

$$i_{in} + \frac{0-0}{R_n} + \frac{v_{out}-0}{R_f} = 0 \tag{1.12}$$

—while taking care to apply Rule 2 by neglecting op-amp input current. Thus,

$$v_{out} = -\,i_{in}\,R_f\,. \tag{1.13}$$

This is the characteristic of a **current-to-voltage converter**.

The input-output characteristic of Eq. 1.13 features a **transresistance** v_{out}/i_{in} equal to the negative of feedback resistance R_f. Here, the "trans" descriptor applies because the voltage and current in the Ohm's-law relation appear across the circuit from one another. Note that the transresistance is independent of of R_n, the Norton resistance tied to the input signal source. This reflects a low-impedance input that is an artifact of the virtual ground. The transresistance is independent of R_L (low-impedance op-amp output), and there is effective buffer action.

The ability to create a virtual ground has enormous potential for design. Given a single feedback path through R_f and a virtual ground as in Fig. 1.4, the op-amp output voltage is

$$v_{out} = -R_f \sum i_{vg}\,, \tag{1.14}$$

where $\sum i_{vg}$ is the sum of currents into the virtual ground from *other* paths.

Exercise 1.1 Determine v_{out} for the circuit of Fig. 1.5.

Ans: $v_{out} = -R_f \left(\frac{v_{in}}{R_1} \right)$

This is the characteristic of an **inverting amplifier**.

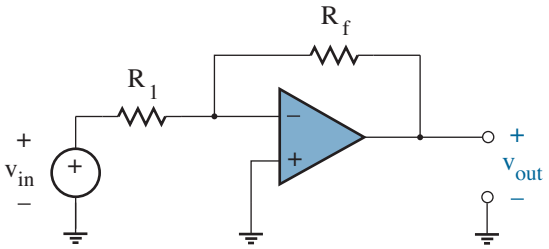


Figure 1.5: Circuit for Exercise 1.1.

Exercise 1.2 Determine v_{out} for the circuit of Fig. 1.6.

Ans: $v_{out} = -R_f \left(\frac{v_1}{R_1} + \frac{v_2}{R_2} \right)$

This is the characteristic of a **summing amplifier**. (The input voltages are transformed to currents, which are added and converted back to a voltage.)

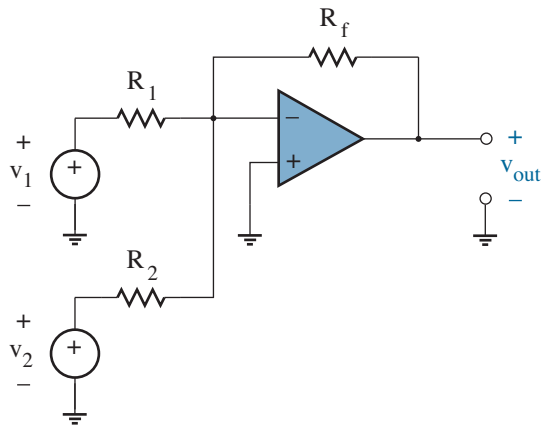


Figure 1.6: Circuit for Exercise 1.2.

Example 1.2

Determine v_{out} for the circuit of Fig. 1.7.

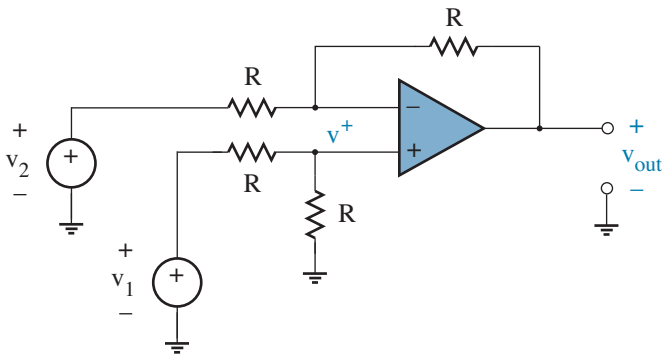


Figure 1.7: Circuit for Example 1.2.

Solution

It is sometimes helpful to break a problem into familiar pieces.

Provided v_{out} is within the bounds of the two op-amp supply voltages, the circuit is linear and subject to the rules of superposition. Black-box Rule 2 disallows any current into the non-inverting v^+ terminal. Thus,

$$v^+ = v_1 \left(\frac{R}{R+R} \right) = \frac{1}{2} v_1 \, . \tag{1.15}$$

If we let $v_2 = 0$, the circuit looks like a non-inverting amplifier with

$$v_{out} = \frac{1}{2} v_1 \left(1 + \frac{R}{R} \right) = v_1 \, . \tag{1.16}$$

On the other hand, if we let $v_1 = 0$, v^+ is also zero, and the circuit looks like an inverting amplifier with

$$v_{out} = -R \left(\frac{v_2}{R} \right) = - v_2 \, . \tag{1.17}$$

So by superposition,

$$v_{out} = v_1 - v_2 \, . \tag{1.18}$$

This is the characteristic of a **differential amplifier**.

Example 1.3

Determine i for the circuit of Fig. 1.8.

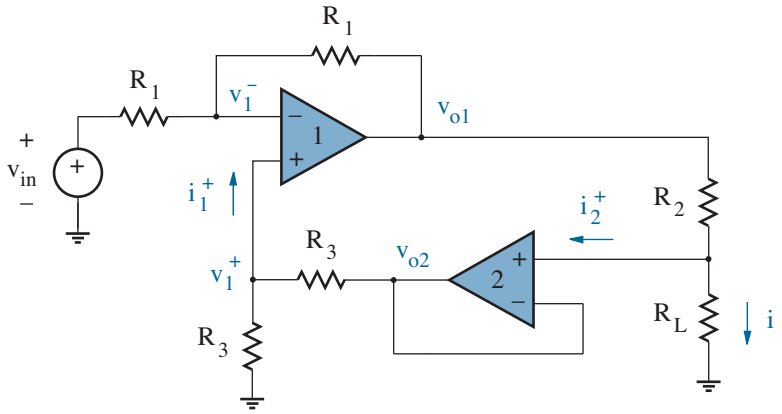


Figure 1.8: Circuit for Example 1.3.

Solution

Our first objective is to find the node voltage at the non-inverting input to op-amp 1 so that we can apply black-box rules to obtain node voltage v_{o1}. With this in mind, we assign iR_L as the node voltage at the non-inverting input to op-amp 2, which is configured to function as a unity-gain buffer. Thus, $v_{o2} = iR_L$. And with ${i_1}^+ = 0$, a voltage divider relation reveals

$$ {v_1}^+ = iR_L \left(\frac{R_3}{R_3 + R_3} \right) = \frac{iR_L}{2} \, . \tag{1.19} $$

So black-box Rule 1 forces ${v_1}^- = iR_L/2$, and Kirchhoff's Current Law has

$$ \frac{v_{in} - iR_L/2}{R_1} + \frac{v_{o1} - iR_L/2}{R_1} = 0 \tag{1.20} $$

at the inverting node for op-amp 1. In turn, $v_{o1} = iR_L - v_{in}$.

Since ${i_2}^+ = 0$, the current through R_L is also the current through R_2. Then in consideration of the node voltages at each end of R_2,

$$ i = \frac{(iR_L - v_{in}) - iR_L}{R_2} = \frac{-v_{in}}{R_2} \, . \tag{1.21} $$

This is the characteristic of a **voltage-to-current converter**. The output current i is independent of the load resistance R_L (as for a current source).

Comparators

In the absence of any feedback, Rule 1 no longer applies, and the op-amp inputs are free to assume independent voltages (while maintaining Rule 2). The op-amp output wants to obey Eq. 1.1. Nevertheless, in the limit as $A_{vd} \to \infty$, the output tends to assume **saturation** with

$$v_{out} = V^+ \qquad \text{for } v^+ > v^- \,, \tag{1.22a}$$

$$v_{out} = V^- \qquad \text{for } v^+ < v^- \,. \tag{1.22b}$$

This is the action of an ideal **comparator**, since v_{out} only depends upon the relative values of v^+ and v^-. The large difference between V^+ and V^- allows easy distinction between the two input conditions.

Exercise 1.3 As a robot arm moves with angle θ to the horizontal, potentiometer R_2 varies linearly between 3 kΩ ($\theta = 0$) and 4.8 kΩ ($\theta = 90°$). The circuit of Fig. 1.9 is intended to produce a "HIGH" v_{out} —the signal closes the robot hand—when $\theta > 60°$. Complete the design by choosing R_1.

Ans: $R_1 = 4.2$ kΩ

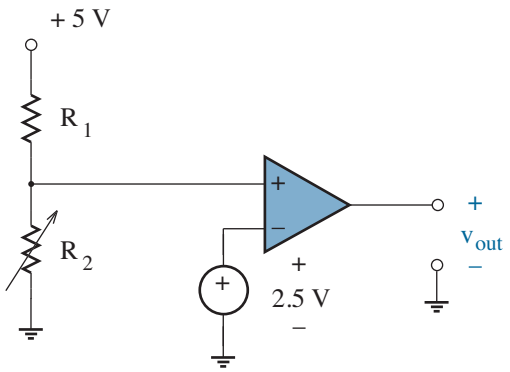


Figure 1.9: Circuit for Exercise 1.3.

Much later, we will find that real op-amps are typically designed with internal circuitry that suppresses unwanted dynamic behavior (oscillations), which sometime accompany negative feedback. There being no free lunch, the suppression is achieved at the expense of relatively sluggish response. "Comparator" integrated circuits are not so designed, and thus they tend to promote rapid interlevel transitions—an example is the popular LM311. A choice of "LOW" output level (V^- or ground) is often available.

A practical complication to Exercise 1.3 muddles the *intended* design: the robot arm moves with slowly increasing θ, the negative comparator input "sees" a linearly related voltage, and the comparator reacts with a single LOW-to-HIGH transition when this signal corresponds to $\theta > 60°$. In a less perfect world, the signal at the negative comparator input linearly tracks the robot arm, but with an added component of electrical "noise." Unable to reason out the true intentions of a naive designer, the comparator is satisfied to make heroic efforts to conform to the inequality relationships of Eq. 1.22 under all input conditions. Thus, the noisy input variation of Fig. 1.10a yields the wild series of output gyrations shown in Fig. 1.10b.

Exercise 1.3 implied a separate circuit designed to close the robot hand. No doubt that circuit can become quite confused.

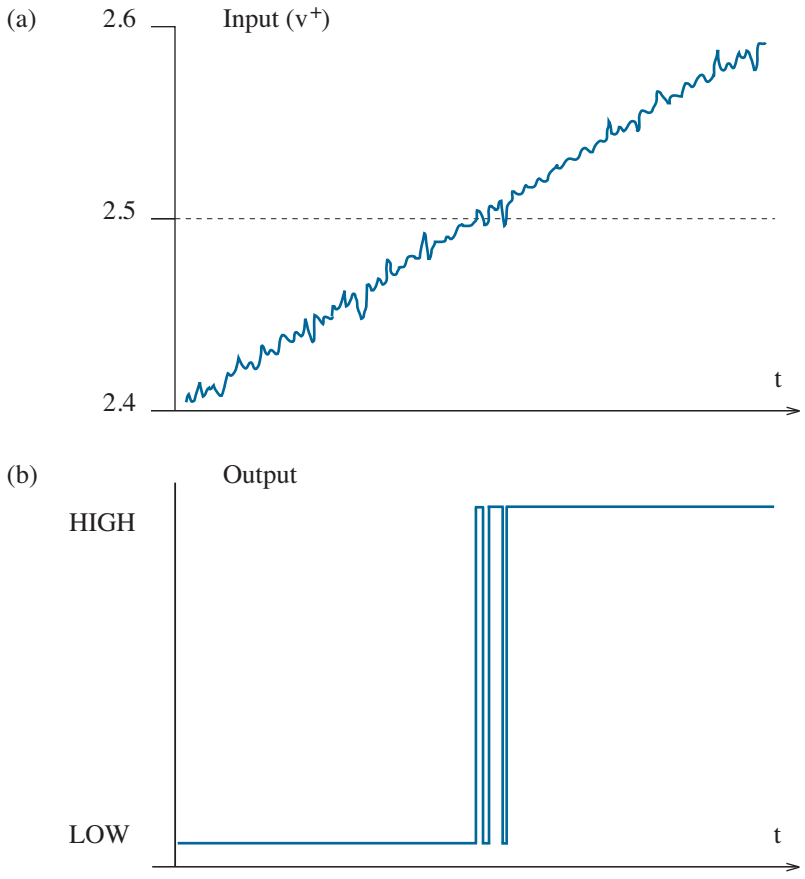


Figure 1.10: Comparator output subject to superimposed input noise.

Positive Feedback Redeemed: The Schmitt Trigger

Despite our earlier disparaging comments, *positive* feedback can be used to avoid the unintended effect of noise in the comparator circuit of Fig. 1.11. Comparator outputs are either HIGH or LOW, so stability at intermediate voltage levels is no longer an issue.

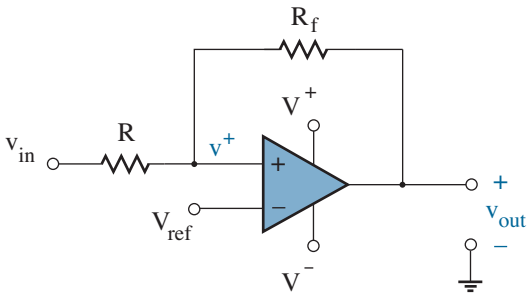


Figure 1.11: Comparator circuit featuring positive feedback.

Suppose v_{in} is sufficiently small to allow $v^+ < V_{ref}$ so that $v_{out} = V^-$. Under what conditions does v_{out} transition to V^+? By superposition,

$$v^+ = v_{in}\left(\frac{R_f}{R+R_f}\right) + V^-\left(\frac{R}{R+R_f}\right), \tag{1.23}$$

and the transition occurs when $v^+ > V_{ref}$. But in terms of v_{in}, we find

$$v_{in} > V_{ref}\left(1+\frac{R}{R_f}\right) - V^-\left(\frac{R}{R_f}\right). \tag{1.24}$$

Voltage V^- is negative, so the second term in Eq. 1.24 shifts the threshold *upwards* in relation to $V_{ref}\,(1 + R/R_f)$.

Now suppose v_{in} is sufficiently large to allow $v^+ > V_{ref}$ and $v_{out} = V^+$. Under what conditions does v_{out} transition to V^-? Again,

$$v^+ = v_{in}\left(\frac{R_f}{R+R_f}\right) + V^+\left(\frac{R}{R+R_f}\right), \tag{1.25}$$

and the transition occurs when $v^+ < V_{ref}$. But in terms of v_{in},

$$v_{in} < V_{ref}\left(1+\frac{R}{R_f}\right) - V^+\left(\frac{R}{R_f}\right). \tag{1.26}$$

Voltage V^+ is positive, so the second term in Eq. 1.26 shifts the threshold *downwards* in relation to $V_{ref}\,(1 + R/R_f)$.

Figure 1.12 shows the transfer characteristic that relates v_{out} to v_{in}. Arrows are needed to indicate upward and downward output transitions, since the v_{out} trajectories for increasing and decreasing v_{in} do not coincide. This non-overlapping behavior is an example of **hysteresis**—the circuit "remembers" the most recent history of its output change.

Noise! When v_{in} becomes just large enough to induce a LOW-to-HIGH upward output transition, any subsequent decrease that is less than some prespecified bound fails to induce a HIGH-to-LOW downward transition. Or now that the output has changed, the effect of superposition has shifted the necessary "undo" threshold to a significantly lower value.

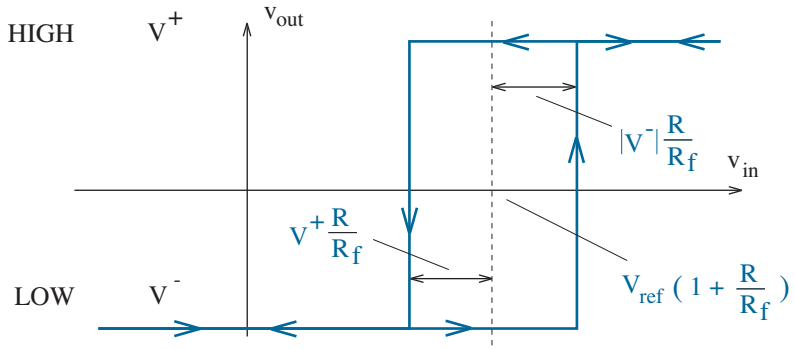


Figure 1.12: Transfer characteristic for the circuit of Fig. 1.11.

The circuit of Fig. 1.11 is called a **Schmitt trigger**. Apart from its use in conjunction with comparators, this electronic marvel enjoys a variety of interesting applications (see Problem 1.54).

Exercise 1.4 In the circuit of Fig. 1.11, $V^+ = +10$ V, $V^- = -10$ V, and $R = 1$ kΩ. Complete the design so that the hysteresis region is 2-V wide and centered at 5 V.

Ans: $R_f = 10$ kΩ, $V_{ref} = 4.55$ V

Exercise 1.5 Consider the circuit of Fig. 1.11, but swap v_{in} and V_{ref}. Determine the output transition requirements. (A sketch may help.)

Ans: LOW-to-HIGH: $v_{in} < V_{ref}\left(\dfrac{R_f}{R+R_f}\right) + V^-\left(\dfrac{R}{R+R_f}\right)$

HIGH-to-LOW: $v_{in} > V_{ref}\left(\dfrac{R_f}{R+R_f}\right) + V^+\left(\dfrac{R}{R+R_f}\right)$

Clouds on the Horizon

The op-amps and comparators presented thus far exhibit ideal performance. Nevertheless, real components are characterized by

- Finite gain, perhaps 100,000 under dc conditions, and diminishing as frequency increases (see Problems 1.55 and 1.56).
- Finite output resistance that reduces output voltage at large currents, notwithstanding external power-supply capability (see Problem 1.57).
- Non-zero currents at the non-inverting (i^+) and inverting (i^-) inputs, typically a few nA or less in older op-amps (see Problem 1.58).

But perhaps the most pernicious dc imperfection is **input offset voltage** —the op-amp or comparator behaves as if a voltage source with value v_{os} is connected in series with the non-inverting input as illustrated in Fig. 1.13.[1] The op-amp output is non-zero when the inputs are connected together, and the comparator output is HIGH when $v^+ + v_{os} > v^-$. While the v_{os} level can be as large as a few mV, one generally finds a means of diminishing its effects either through provisions for offset adjustment or careful design. Figure 1.14 shows typical v_{os} temperature behavior that tends to throw off these mitigating efforts. Look for v_{os} difficulties in Chapter 14.

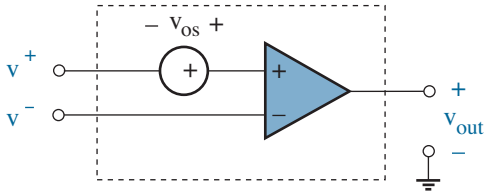


Figure 1.13: Op-amp or comparator with offset voltage.

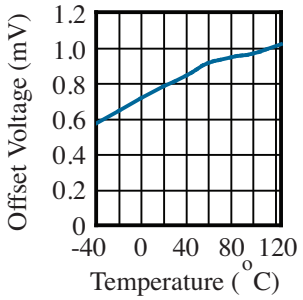


Figure 1.14: Typical input offset voltage vs. temperature for the MAX4400. (Copyright Maxim Integrated Products. Used by permission.)

[1]Strictly speaking, we cite the output offset voltage divided by A_{vd} when $v^+ = v^-$. Some texts define v_{os} (here $-v_{os}$) as the $v^+ - v^-$ difference needed to achieve $v_{out} = 0$.

Apart from dc imperfections, op-amps suffer from dynamic limitations. In the time domain, for example, one typically observes a maximum rate of output voltage change or **slew rate** as shown in Fig. 1.15.

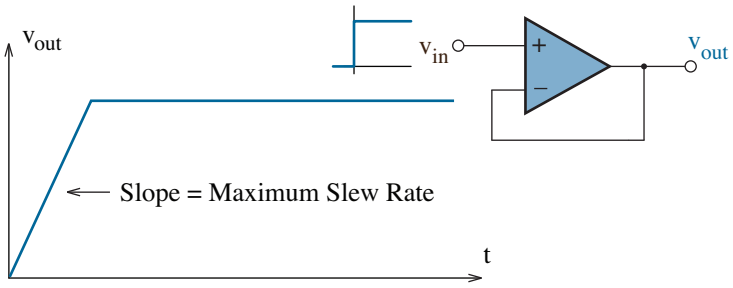


Figure 1.15: Typical step response for a unity-gain op-amp buffer.

In the frequency domain, a sinusoidal op-amp output signal generally demonstrates reduced gain with increasing frequency. Figure 1.16 shows the frequency dependence of voltage gain for a typical non-inverting amplifier. The transresistance of a current-to-voltage converter has similar rolloff.

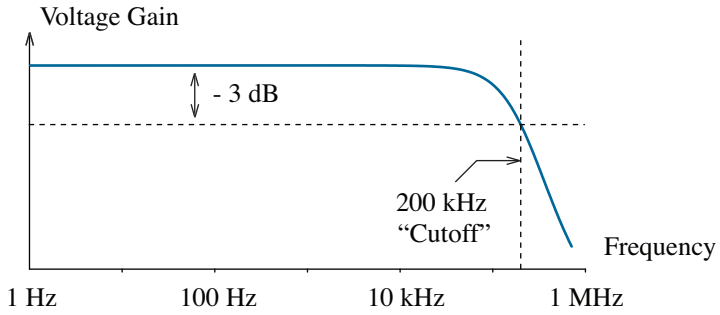


Figure 1.16: Typical gain response for a non-inverting op-amp amplifier.

The op-amp frequency response also introduces a phase shift for A_{vd}, which can transform *stable* negative feedback into effective *unstable* positive feedback when the phase shift is 180° or more $[-\sin(\omega t{+}180°) = +sin(\omega t)]$. Look for feedback and stability considerations in Chapters 11 and 12.

Op-amps and comparators must function despite electrical noise effects at their non-inverting and inverting inputs—recall the disorder in Fig. 1.10. This noise arises from electrical processes that are either internal to the part or external to it as some form of **electromagnetic interference (EMI)**. Finally, op-amp output signals are subject to distortion when the inputs or outputs are disallowed **rail-to-rail** variations within power-supply bounds. Look for these and other signal corruption processes in Chapter 13.

1.2 Digital: A Hierarchy of Connected Gates

A **digital** signal is generally a collection of time-dependent voltages that are individually classified as HIGH or LOW to represent logical states 1 or 0, respectively. This is often convenient for the **binary** coding of certain types of information. For example, the string

1 1 0 0

can reflect

- four conditions in a home security system—front door open (1), back door open (1), south window closed (0), no hallway motion (0),
- the decimal number 12,
- analog 0.75 V ($= 12/16 \times 1$ V on a 1-V maximum scale),

or, when preceded by an additional 1 0 0

- the upper-case "L" on a computer keyboard—as specified by the American Standard Code for Information Exchange (ASCII).

Any string of n "bits" can be viewed in *parallel* as a set of n separate HIGH or LOW node voltages or *serially* as a single node voltage that assumes HIGH or LOW levels during n successive time intervals.

Much of the art of digital logic design rests on principles that are far removed from the techniques of elementary circuit analysis and the scope of this text. So we should avoid presumptuous thoughts of acquired talent as we introduce a small assortment of digital components—it will take a little more than what is offered here to design your next PC. Nevertheless, black-box operating rules tend to support the notion of simple functionality. Indeed, it is hard to imagine a black box that offers a simpler function than that of the **inverter** shown in Fig. 1.17.

Rule: *The inverter output is HIGH if the input is LOW, and conversely.*

HIGH and LOW are typically close to the V^+ supply voltage and ground, respectively. (The V^+ supply is usually denoted as V_{DD} in digital systems.)

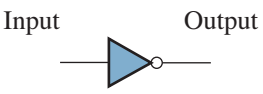


Figure 1.17: Digital inverter. Power and ground connections are not shown.

To justify the preceding black-box rule, we consider the typical inverter transfer characteristic of Fig. 1.18 in which the input voltage varies between the extremes of the output: V_{OL}, the **low-level output voltage**, and V_{OH}, the **high-level output voltage**. Two conditions warrant attention:

- If $v_{in} < V_{IL}$, the **low-level input voltage**, the inverter output is *guaranteed* to be effectively HIGH.
- If $v_{in} > V_{IH}$, the **high-level input voltage**, the inverter output is *guaranteed* to be effectively LOW.

As shown in Fig. 1.18, V_{IL} and V_{IH} correspond to points on the transfer characteristic where the slope is -1. The output voltage is relatively immune to input-voltage changes in regions where the slope is less negative than -1.

Given V_{IL}, we can define a conventional, albeit fuzzy meaning of LOW. Output levels within the **LOW noise margin**

$$NM_L = V_{IL} - V_{OL} \tag{1.27}$$

present a "0" character to the input of another gate despite the corrupting influence of spurious signals. And given V_{IH}, we can do the same for HIGH. Output levels within the **HIGH noise margin**

$$NM_H = V_{OH} - V_{IH} \tag{1.28}$$

present a "1" character to the input of another gate despite the corrupting influence of spurious signals. The **transition region** between V_{IL} and V_{IH} is easily avoided. Noise immunity is a major digital advantage.

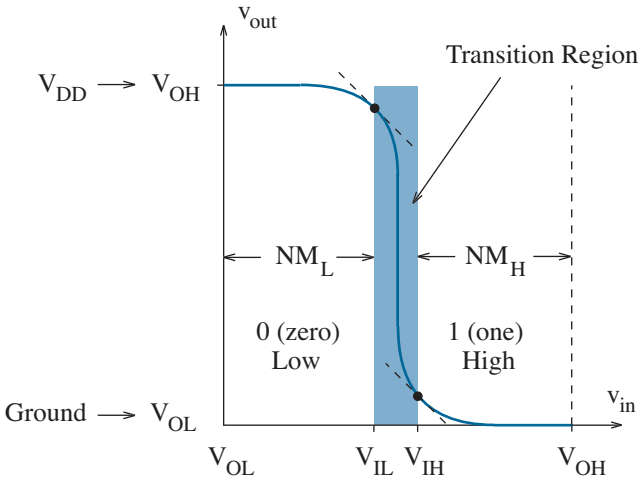


Figure 1.18: Input/output transfer characteristic for a typical inverter. Inputs within the unshaded noise margins reflect valid logic levels.

More sophisticated logical functions are effected using the **AND**, **OR**, **NAND**, and **NOR** "gates" shown in Fig. 1.19. Each gate has multiple—usually 2, 4, or 8—inputs that influence a single output.

Rule: *The AND output is HIGH only if all inputs are HIGH.*

Rule: *The OR output is LOW only if all inputs are LOW.*

NAND (not-AND) and NOR (not-OR) gates function like AND and OR gates, respectively, but with inverted outputs.

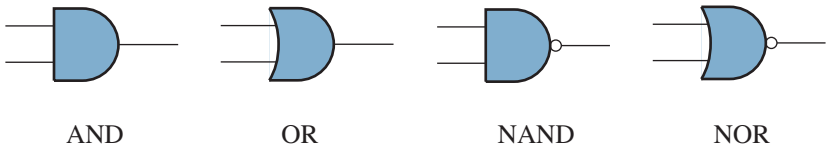


Figure 1.19: AND, OR, NAND, and NOR digital gates.

We shall later discover that the internal circuitry of an inverter serves as the foundation for NAND and NOR gates.

Exercise 1.6 Determine if y is HIGH or LOW in the circuits of Fig. 1.20.

Ans: (a) HIGH (b) LOW (c) LOW (d) HIGH

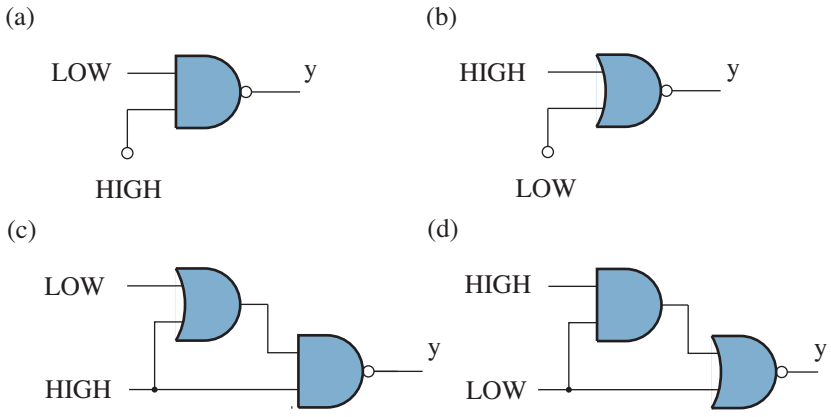


Figure 1.20: Circuits for Exercise 1.6.

An application that underscores the "gate" concept is the **multiplexer** shown in Fig. 1.21. In this circuit, one or the other of the A and B pulse trains is allowed to pass through to the output y—somewhat like action in a railroad freight yard—depending upon the logical state of control input x. For example, if x is made LOW, the output of AND-gate 1 is forced to LOW (so that pulse train A is blocked). Meanwhile, the inverter transforms x into a HIGH level at one of the inputs to AND-gate 2 (so that its output, in turn, assumes HIGHs and LOWs that dynamically follow pulse train B). The OR gate passes this signal on to the output y.

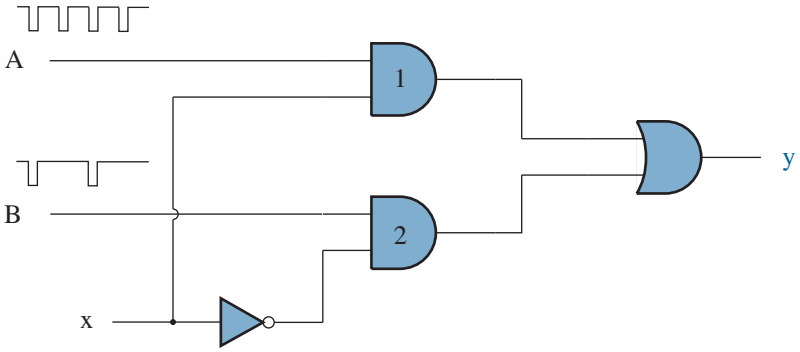


Figure 1.21: Two-input multiplexer.

To divert a pulse train from a single input to one of several outputs, we use a circuit with the form of the **demultiplexer** shown in Fig. 1.22. The pulse train is common to each AND-gate input, but the signal only passes through if the other input is HIGH.

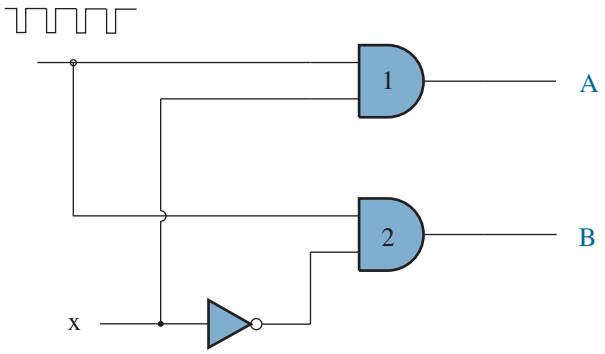


Figure 1.22: Two-output demultiplexer.

When output-to-input feedback is applied to a digital circuit, the result is no less interesting than that for an analog circuit. Consider, for example, the cross-connected NAND gates of Fig. 1.23. You can easily verify that the output levels at Q_1 and Q_2 complement those at the inputs R and S, respectively, when [R, S] are made either [HIGH, LOW] or [LOW, HIGH]. From one perspective, the former input combination "resets" Q_1 LOW, and the latter input combination "sets" Q_1 HIGH. The circuit is called a **bistable RS latch** because the outputs uniquely respond (or "latch" on) to one or the other [HIGH, LOW] [LOW, HIGH] input conditions, *and then remain in the same logical states after both inputs have been made HIGH.* Things are not particularly exciting whenever both inputs are made LOW —both outputs become HIGH—so this condition is generally avoided.

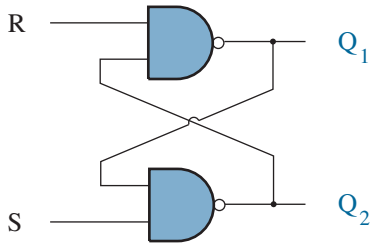


Figure 1.23: Bistable RS latch.

Exercise 1.7 The circuit of Fig. 1.24 has three inputs (R, S, and CK) and one output Q, which has some particular initial state. Determine the Q state *after* a LOW-HIGH-LOW CK-level transition. Consider three [R, S] conditions: (a) [HIGH, HIGH]; (b) [LOW, HIGH]; (c) [HIGH, LOW].

Ans: (a) no change (b) Q HIGH (c) Q LOW

This is the characteristic of an **RS flip-flop** featuring clock (CK) control (for synchronized action in relation to similar circuits elsewhere).

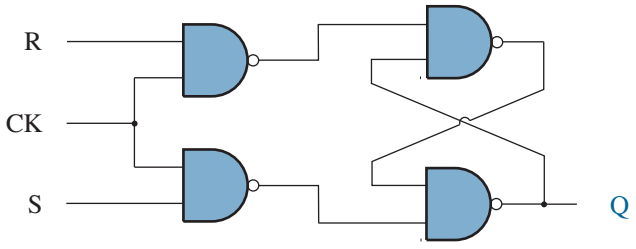


Figure 1.24: Circuit for Exercise 1.7.

The RS flip-flop just examined can be made simpler but more restrictive when an inverter forces the R input to have the logical inverse of the S input as shown for the **D flip-flop** of Fig. 1.25. By convention, D replaces S, and $\bar{Q}$ —read as Q-bar, the logical complement of Q—is a second output.

Rule: *After a complete clock pulse, the Q and $\bar{Q}$ outputs of a D flip-flop assume HIGH and LOW levels, respectively, when the D input is HIGH. In contrast, the Q and $\bar{Q}$ outputs take LOW and HIGH levels, respectively, when D is LOW.*

This black-box behavior represents "delay" (D) action with $D \rightarrow Q$.

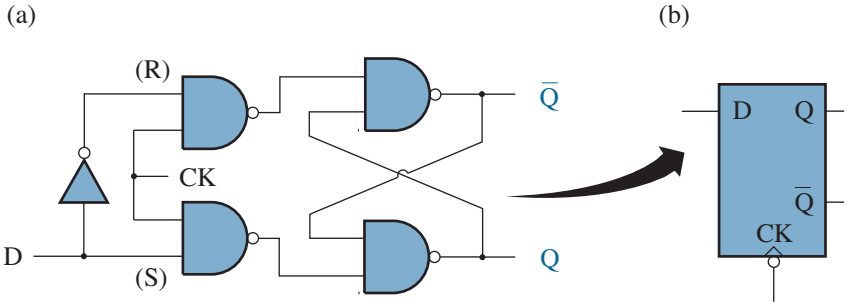


Figure 1.25: D flip-flop: (a) gate-level circuit; (b) black box equivalent.

The D flip-flop becomes the even more restrictive **T flip-flop** of Fig. 1.26 when the $\bar{Q}$ output feeds back to the former input D.

Rule: *After a complete clock pulse, the Q and $\bar{Q}$ outputs of a T flip-flop assume complementary values.*

This black-box behavior represents "toggle" (T) action with $\bar{Q} \rightarrow Q$.

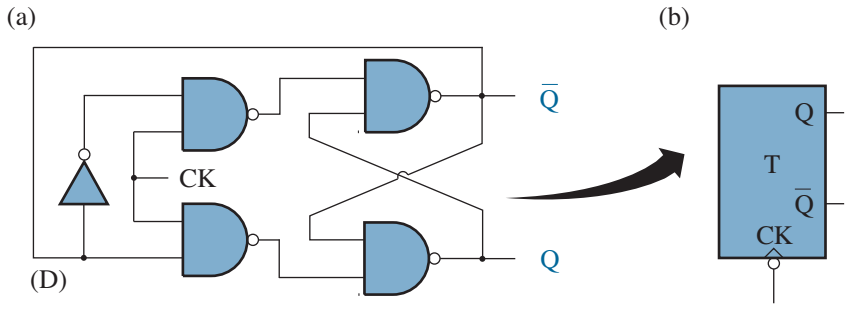


Figure 1.26: T flip-flop: (a) gate-level circuit; (b) black box equivalent. Unmodified, the specific circuit has clock restrictions (see Problem 1.77).

Envision an application in which data of interest are transmitted serially along a single wire with HIGH or LOW values B_0, B_1, B_2, and B_3 at times t_0, t_1, t_2, and t_3 respectively. Our goal is to view this information in parallel after time t_3. So armed with some useful black-box rules, we design a 4-bit **shift register** featuring a chain of four D flip-flops as shown in Fig. 1.27. The Q output of one flip-flop is connected to the D input of another so that flip-flop $n-1$ takes on the state of flip-flop n after one clock pulse. Thus, the register contents are successively B_0xyz, B_1B_0xy, $B_2B_1B_0x$, and $B_3B_2B_1B_0$, where xyz represent old data bits no longer of interest.

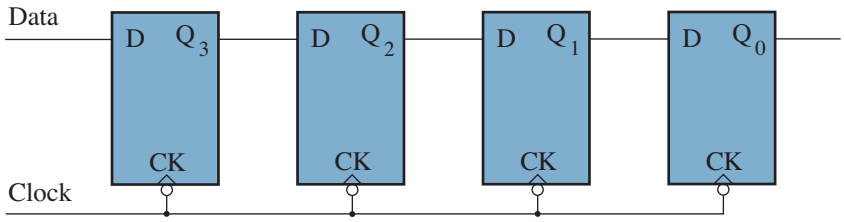


Figure 1.27: Four-bit shift register.

Another application is to provide for parallel observation of the sequence 0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1000, ... This is enough of a pattern to suggest the 4-bit T-flip-fop **counter** design shown in Fig. 1.28. The individual bits are the Q_3, Q_2, Q_1, and Q_0 outputs of the T flip flops. Flip-flop 0 is connected directly to the clock line so that Q_0 changes from 0 (LOW) to 1 (HIGH) with every clock pulse. In turn, this output serves as the clock for flip-flop 1 so that Q_1 changes state every time Q_0 makes a HIGH-to-LOW transition. A similar connection forces Q_2 to change state as Q_1 goes LOW, and so on.

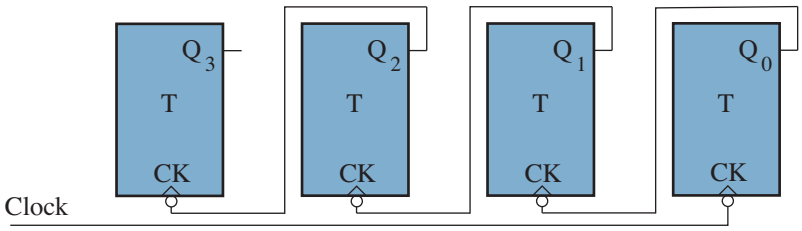


Figure 1.28: Four-bit counter. The count is upwards with each clock pulse.

Shift registers and counters are available to us as special black boxes, which can be used to perform logical operations in a microprocessor, which is an even more specialized black box. Where will it all end?

Diverging Horizons . . .

Not long ago, many electrical and computer engineering students could look forward to a course invariably called "Digital Electronics" that built upon more elementary subjects to relate currents and voltages to ones and zeros. Course alumni are certain to recall happy hours analyzing and designing logic families with names like RTL, DTL, TTL, ECL, I^2L, SFL, nMOS . . . And it was understood that nearly anything was possible once the inverter, NAND, and NOR logic functions had been implemented (see Problem 1.64). Different logic families offered an assortment of benefits and liabilities.

Times change. All of the preceding logic families have become obsolete. The present reigning family, CMOS, demands relatively mundane circuits: One has a collection of paired switches—call them p-switches and n-switches for now—and each switch pair is associated with a controlling input voltage. A HIGH input closes an n-switch with the p-switch open, and a LOW input closes a p-switch with the n-switch open. Figure 1.29 shows the realization of inverter, NOR, and NAND functions. (Take a moment to be convinced.) One has comparable circuit action with paired normally-on and -off relays. But the physical basis and design principles for ultra-small switches is far more complicated. Chapters 2, 5, and 10 address the relevant issues.

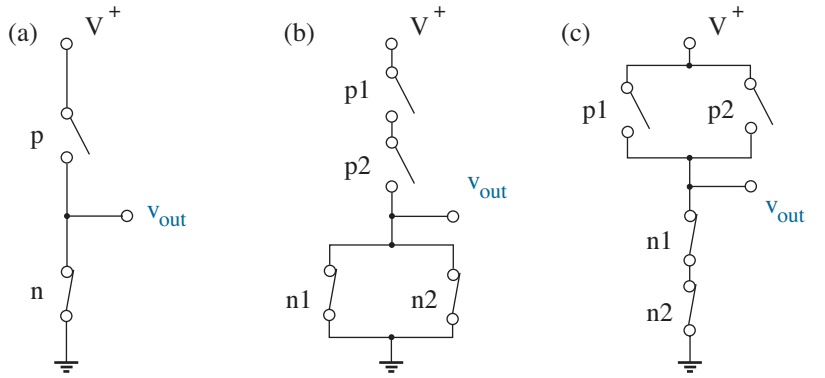


Figure 1.29: Primative CMOS circuits: (a) inverter; (b) NOR; (c) NAND. The p-switches are open when the n-switches are closed, and conversely.

What about Digital Electronics? This course has gradually evolved to something high-tech like "VLSI Design" (as in Very Large Scale Integration). It begins with the necessary support for the CMOS functions of Fig. 1.29 (and no doubt benefits from this or similar texts). Then interests diverge. Given a prerequisite understanding of logic design with just a few gates or flip-flops, how does one optimize, let alone cope with millions of elements? What design architectures are suitable for microprocessors and memory? How does one distribute power and timing signals throughout a network no less complex than New York or Los Angeles? These are all *systems* issues. The answers invite happy hours. But not here.

. . . But Cloudy Nonetheless

Most engineers will be content to design for electrical interactions along the peripheries of complex digital integrated circuits such as a microprocessor. Numerous performance attributes elude the black-box characterization:

- Some CMOS inputs are purely capacitive and look like open circuits. But others have the current-voltage characteristics of Fig. 1.30a when the power is on and Fig. 1.30b when the power is off. What operating standards does this imply? What should be done with unused inputs?
- Most digital CMOS outputs are intended to connect to other inputs. How many connections are feasible? Can outputs connect together?
- Digital circuits absorb and produce data. What are the practical data communication methods, and what speed limitations apply?
- Experiments show that CMOS digital components expend more power as one increases the switching frequency between between logic states. What design factors affect this behavior?

And for those specialized engineers who design within the black box,

- The inverter input/output transfer characteristic of Fig. 1.18 showed equivalent low and high noise margins. How is this condition achieved, and is it necessarily desirable?
- Once an inverter input becomes effectively HIGH, it takes a certain time for the output to become effectively LOW. And it generally takes a different time to realize the opposite transition. Balance is possible. How is this condition achieved, and is it necessarily desirable?
- What *circuit* issues confront the modification of an inverter to achieve NAND, NOR, and beyond?

Look to Chapters 10 and 13 for some answers.

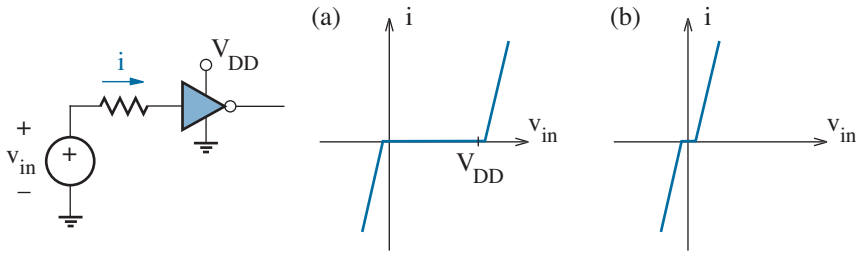


Figure 1.30: Possible input current vs. input voltage for a CMOS inverter: (a) power supply on; (b) power supply off.

1.3 Mixed-Signal Systems: The 555 Timer

Analog or digital? The **555 Timer** has been around since the early 1970s. And even with the occasional new arrival of challengers offering improved performance, it remains a low-cost integrated circuit with popular appeal. Analog *and* digital signals will increasingly play co-operative and equally important roles as complex mixed-signal systems evolve.

In relation to the black box shown in Fig. 1.31, the 555 timer sports:

- Two power connections — V^+ (pin 8) and ground (pin 1).
- Two inputs — The **trigger** (pin 2) and **threshold** (pin 6) are inputs that only have effect when they are made less than or greater than specific reference voltages.
- Two outputs — The **output** (pin 3) and **discharge** (pin 7) assume one of two states: When the output is HIGH (typically $V^+ - 0.9$ V), *the discharge connection appears as an open circuit.* When the output is LOW (typically 0.2 V), *the discharge connection appears as a short circuit to ground.*
- Two special connections — The **reset** (pin 4) forces a LOW output at pin 3 when set to a LOW voltage, and it has no effect when set to a HIGH voltage. The **control** (pin 5) is used to change the values of the reference voltages that govern the behavior of the two inputs. (We shall tend to ignore both special connections.)

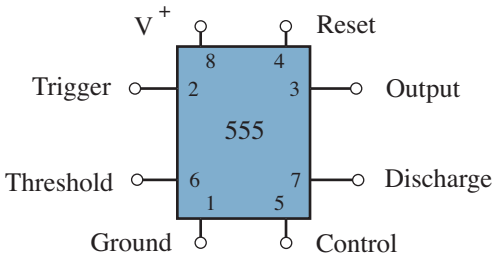


Figure 1.31: Pin designations for the 555 timer.

There are three simple governing rules:

Rule 1: Barring a conflict with Rule 2, *the output goes HIGH and stays there if the trigger voltage is made less than* $(1/3)V^+$.

Rule 2: Barring a conflict with of Rule 1, *the output goes LOW and stays there if the threshold voltage is made greater than* $(2/3)V^+$.

Rule 3: *The input terminal currents are ideally zero.*

Monostable Behavior

What do the 555-timer rules imply? Suppose the initial output, trigger, and threshold voltages are LOW, 6 V, and 0 V, respectively, and let $V^+ = 6$ V. If the trigger is subsequently set to 0 V, which is less than $(1/3)V^+ = 2$ V, Rule 1 tells us that the output will become HIGH and stay there indefinitely (even as the trigger is set back to 6 V shortly afterwards). This is consistent with the trigger and output waveforms shown in Fig. 1.32.

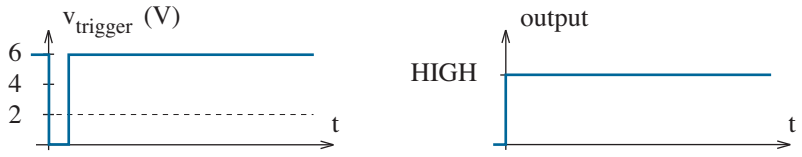


Figure 1.32: 555 trigger and output waveforms.

Nothing very exciting so far. However, we can limit the time duration of the HIGH output condition by taking advantage of Rule 2 —we merely force the *threshold* voltage above $(2/3)V^+ = 4$ V at a desired time following the completion of the trigger pulse. One way to do this is to connect the threshold input to the RC circuit shown in Fig. 1.33a. The initial threshold voltage v_{th} is 0 V, and the threshold terminial draws no current (Rule 3). Thus, at time t,

$$v_{th} = V^+ \left(1 - e^{-t/RC}\right) . \tag{1.29}$$

In turn, $v_{th} = (2/3)V^+$ at time

$$T = RC \ln 3 = 1.1\, RC . \tag{1.30}$$

The consistent threshold and output waveforms appear in Fig. 1.33b.

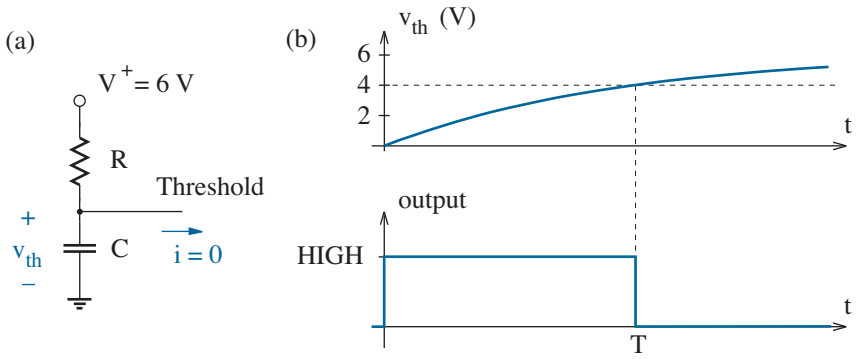


Figure 1.33: (a) 555 threshold circuit; (b) threshold and output waveforms.

Things are looking much better, apart from a minor technical difficulty: How can we ensure that the threshold voltage begins to rise when the 555 output goes HIGH? And how can we ensure that the system produces another output pulse in response to a subsequent trigger signal?

Both problems resolve by tying the 555 discharge to the threshold input. When the output is initially LOW, the discharge appears as a short circuit to ground, and it holds the threshold voltage to an approximate zero level. When the output becomes HIGH, the discharge appears as an open circuit, and the threshold voltage is made free to rise. When the output becomes LOW again, the discharge forces the threshold voltage back near zero.

So now we have a 555 **monostable** or **one-shot** circuit that produces a long output pulse of *fixed* duration in response to a shorter trigger pulse of *arbitrary* duration. Figure 1.34 shows the complete monostable circuit. Note that the reset terminal is tied to V^+, and the control terminal is tied to ground through a 0.01-μF capacitor (to suppress undesired transients).

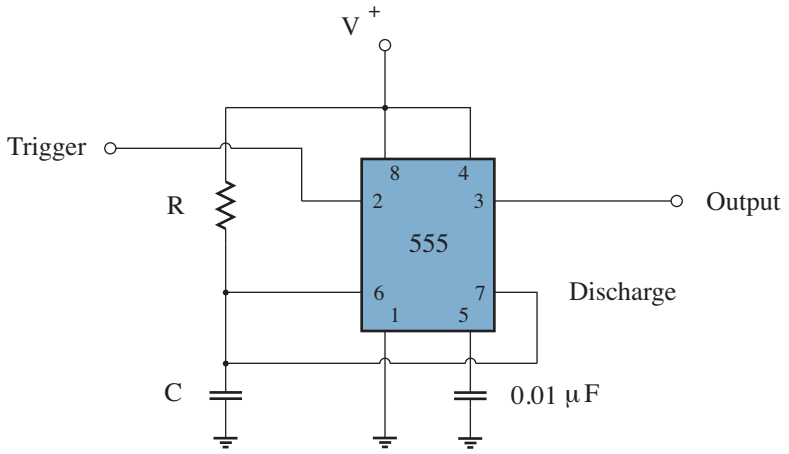


Figure 1.34: 555 monostable circuit.

Exercise 1.8 A 555 monostable circuit is intended to produce a 0.5-s output pulse subject to a design with $C = 0.1$ μF. Determine R.

Ans: $R = 4.5$ MΩ

Exercise 1.9 The capacitor of the preceding exercise discharges through an effective resistance of 1 Ω. Determine the time needed for the threshold voltage to return to 0.2 V from its highest value. Assume $V^+ = 6$ V.

Ans: $t = 0.3$ μs

Astable Behavior

The prospects for another useful 555 circuit will soon become apparent with the help of Fig. 1.35. Here, voltage v_c is $(2/3)V^+$ when the switch is closed at $t = 0$. Our interest is the time at which $v_c = (1/3)V^+$.

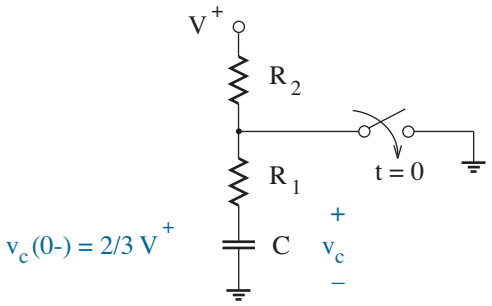


Figure 1.35: RC demonstration circuit.

The capacitor voltage decreases exponentially between initial and final values with time constant R_1C. Specifically,

$$v_c(t) = v_{final} + (v_{initial} - v_{final})\, e^{-t/R_1C} \,. \tag{1.31}$$

So with $v_{initial} = (2/3)V^+$ and $v_{final} = 0$,

$$v_c(t) = \frac{2}{3}\, V^+ e^{-t/R_1C} \,. \tag{1.32}$$

And when $v_c = (1/3)V^+$,

$$t = t_1 = R_1C \,\ln 2 = 0.693\, R_1C \,. \tag{1.33}$$

Now open the switch again at $t' = t - t_1 = 0$. Our new interest is the time at which $v_c = (2/3)V^+$, the initial condition for the preceding process. The capacitor voltage increases exponentially between $v_{initial} = (1/3)V^+$ and $v_{final} = V^+$ with time constant $(R_1 + R_2)C$. Thus, we look to the form of Eq. 1.31 to obtain

$$v_c(t') = V^+ - \frac{2}{3}\, V^+ e^{-t'/(R_1+R_2)C} \,. \tag{1.34}$$

In turn, when $v_c = (2/3)V^+$,

$$t' = t_2 = (R_1 + R_2)C \,\ln 2 = 0.693\,(R_1 + R_2)C \,. \tag{1.35}$$

If the switching cycle repeats indefinitely, the frequency is

$$f = \frac{1}{t_1 + t_2} = \frac{1.443}{(2R_1 + R_2)C} \,. \tag{1.36}$$

Enter the 555 timer. In consideration of Rule 1 and Rule 2, we connect the trigger and threshold inputs to v_c so that the 555 output becomes HIGH when $v_c < (1/3)V^+$ and LOW when $v_c > (2/3)V^+$. The v_c time dependence is not affected (Rule 3). Thus, the LOW and HIGH intervals are t_1 and t_2, respectively.

While the 555 output is LOW (and v_c decreases), the discharge appears as a short circuit to ground—just like the switch. And while the 555 output is HIGH (and v_c increases), the discharge appears as an open circuit—just like the switch. So we can eliminate the switch and, more importantly, sustain the switching cycle by connecting the discharge to the node between R_1 and R_2. Here is yet another triumph for circuit feedback.

Figure 1.36 shows the complete **astable** circuit.

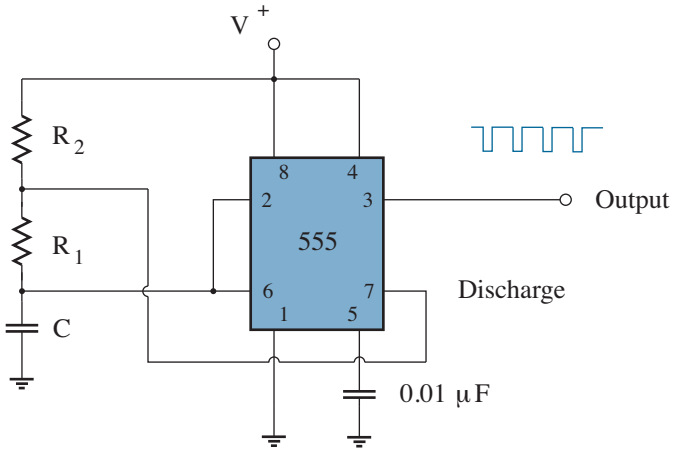


Figure 1.36: 555 astable circuit.

The **duty cycle** of the pulse train produced by a 555 astable circuit is defined as the ratio of the HIGH interval (t_2) to the waveform period ($t_1 + t_2$). Thus, in consideration of Eqs. 1.33 and 1.35,

$$\text{duty cycle} = \frac{R_1 + R_2}{2R_1 + R_2} \times 100\,\% \,. \tag{1.37}$$

If $R_1 \gg R_2$, this approaches 50 %, the duty cycle for a square-wave.

Exercise 1.10 A 555 astable circuit with the form of Fig. 1.36 is intended to produce a 2-kHz pulse train with 80% duty cycle subject to a design with $C = 0.1$ μF. Determine R_1 and R_2.

Ans: $R_1 = 1.4$ kΩ, $R_2 = 4.4$ kΩ

Inside the 555 Black Box

Peel back the cover of a 555 timer, and you will see the assortment of interconnected components and black boxes shown in Fig. 1.37. Abstractly, you find a chain of three equal-value resistors between V^+ and ground, two op-amp-like comparators, an *RS* latch, and an electronic device called a **transistor**—actually an npn bipolar junction transistor or BJT. No doubt you have heard of this last component, as it pervades the popular culture. For the moment, we treat the BJT as an especially fundamental black box that functions like a switch: there is an effective short circuit between the C (collector) and E (emitter) terminals when the B (base) terminal is tied through a resistor to a HIGH voltage level, and there is an open circuit between C and E when B is similarly connected to a LOW voltage level. In practice, the BJT rules are much more complicated.

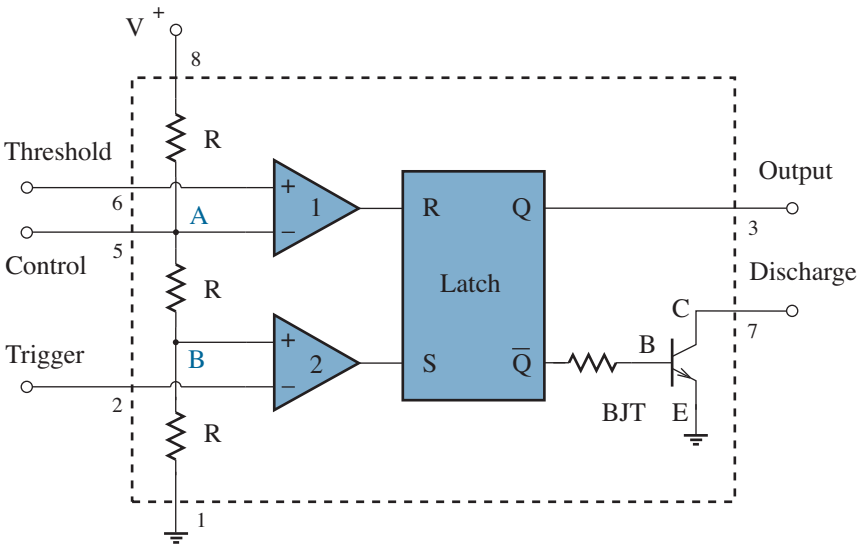


Figure 1.37: Inside the 555 timer.

The new 555 abstraction explains the output and discharge conditions encountered previously. When the external output is HIGH, the internal Q output of the *RS* latch is also HIGH, and its complement $\bar{Q}$ is LOW, which induces the BJT to make the discharge appear as an open circuit. But when the external output is LOW, Q and $\bar{Q}$ are LOW and HIGH, respectively, and the latter induces the BJT to make the discharge appear as a short circuit to ground.

Meanwhile, the internal comparators draw zero input currents (Rule 3). The three-resistor voltage divider is thus made free to establish reference voltages of $(2/3)V^+$ at node A and $(1/3)V^+$ at node B (provided that there is an open connection at the external control terminal). Then we have ...

• **Rule 1:** Barring a conflict with Rule 2 means that the threshold voltage is less than $(2/3)V^+$ so that comparator 1 yields a LOW voltage at the R input to the latch. And when the trigger voltage becomes less than $(1/3)V^+$, comparator 2 yields a HIGH voltage at the S input to the latch. In turn, Q is "set" HIGH.

• **Rule 2:** Barring a conflict with Rule 1 means that the trigger voltage is greater than $(1/3)V^+$ so that comparator 2 yields a LOW voltage at the S input to the latch. And when the threshold voltage becomes greater than $(2/3)V^+$, comparator 1 yields a HIGH voltage at the R input to the latch. In turn, Q is "reset" LOW.

. . . as advertized.

Engineers design *with* integrated circuits—

only a relatively few design integrated circuits.

But woe to the engineer who overlooks the specifics of black-box interiors (see Problem 1.91).

→ Peel back the cover of an op-amp or comparator, and you will see an assortment of interconnected transistors that function much like valves—they pass current, but in an intermediate sense with not just all or nothing. How do they establish a *large* (but not infinite) differential voltage gain? What factors contribute to input offset voltage?

→ Peel back the cover of an RS flip-flop and the several covers of the gates within it, and you will see an assortment of interconnected transistors that function much like switches—shorted when closed, no current when open. How do they recognize and establish particular HIGH and LOW levels? What time constraints apply?

→ Peel back the cover of a (black-box) transistor, and you will find a device structure that is governed by a coterie of material and physical principles. What is the best transistor for valve- or switch-like applications?

Electronics is a discipline with an endless hierarchy of little black boxes. All of these boxes function with individual sets of ideal rules. Nevertheless, it is necessary to ask

When do the ideal black-box rules break down?

Alas, you probably skipped right over the Preface—just like most readers. So it bears repeating that this text concerns the fragility of black-box rules. Try as we may to understand electronics at the highest levels of abstraction, there's no escaping the need to peel back the covers. We are poised to begin by acquiring some fundamental electronic concepts in Chapters 2 - 8.

1.4 Enter the Computer

So far our attention has been limited to simple black-box "models" and related analytical procedures that facilitate pencil-and-paper calculations. Nevertheless, we have set the stage for the gradual appearance of more sophisticated models and a supporting cast of quirky electronic devices. Hand calculations will retain their importance as a means for initiating the design process. Computer simulations will be useful for design *verification*.

Op-Amp SPICE Analysis

As noted in the Preface, our discussion of computer-aided electronic circuit simulation focuses on SPICE,[2] an "industry standard" application that is readily available to students in various forms (HSpice, PSpice, etc.). The general rules and command formats that apply to SPICE simulations appear in the Appendix. Some readers may want to review this material.

When applied to the demonstration circuit of Fig. 1.38, the SPICE code for dc analysis takes the following form:[3]

```
* Op-Amp Demonstration Circuit

Vin     1     0     0
R1      2     0     10k
Rf      2     3     90k
Rin     1     2     1000MEG
Eout    3     0     1     2     1E6

.dc     Vin   -2.5  +2.5  0.01
.probe

.end
```

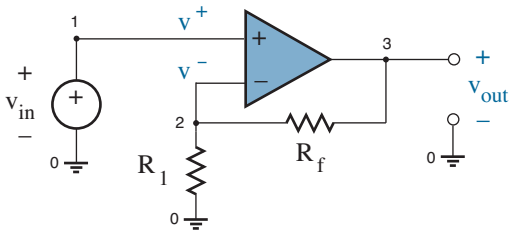


Figure 1.38: Demonstration circuit with node labels for SPICE analysis.

[2]SPICE (Simulation Program with Integrated Circuit Emphasis) is a general-purpose circuit analysis program developed at the University of California, Berkeley in the 1970s. Its progeny enjoy widespread use today.

[3]Popular "schematic-capture" SPICE implementations avoid a code in favor of a computer-generated circuit diagram with related specification and command windows. The circuits used in this text offer little disadvantage to the generic code.

Briefly, the SPICE code consists of four statement groups:

- An arbitrary circuit name (always preceded by an asterisk)
- A circuit netlist —

  In this example, the first three statements apply to components that are external to the op-amp, whereas the fourth and fifth statements make a simple approximation to the ideal op-amp model (see Fig. 1.2b). Resistor Rin ensures that the non-inverting and inverting inputs are always connected to something—SPICE can be unhappy otherwise —and the large resistance value keeps the input currents negligible. Element Eout is a voltage-dependent voltage source that establishes a voltage between nodes 3 and 0 (the op-amp output and ground) that is proportional to the voltage between the input nodes 1 and 2. The proportionality factor of 10^6 is the differential gain (A_{vd}).
- A set of command statements —

  The .dc command generates an input-output transfer characteristic, which is visualized via the .probe command. This simulation sweeps the input (Vin) from -2.5 V to +2.5 V in 0.01-V increments.
- A .end statement (self-explanatory)

Figure 1.39 shows .probe results for the demonstration circuit of Fig. 1.38. As expected, the voltage gain is 10.

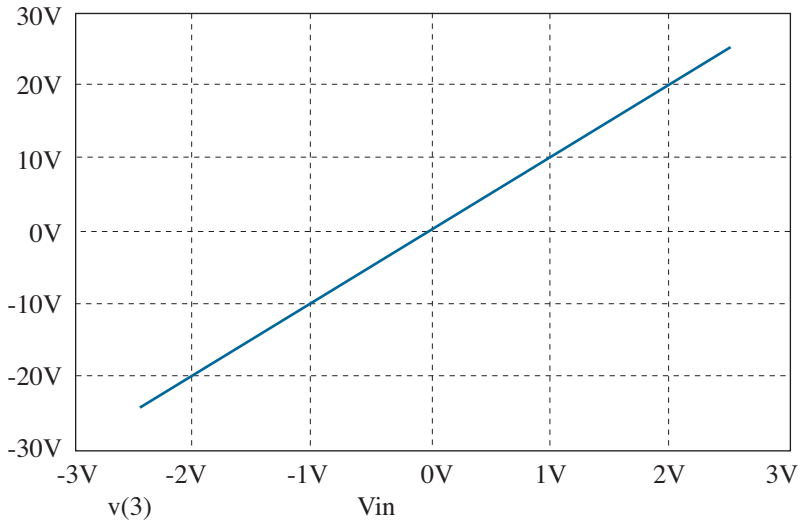


Figure 1.39: .probe results for the demonstration circuit of Fig. 1.38.

Attentive readers ought to be somewhat disturbed by the -25 V to +25 V output variation in the preceding simulation. The output of a real op-amp is limited by the V^+ and V^- power-supply rails, and it is unlikely that the magnitude of either level exceeds 25 V. Lacking any contrary information, computers cheerfully overextend models to provide misleading results.

One way to include the effects of the op-amp power-supply connections is to modify the description of the voltage-dependent voltage source (Eout) as follows:

```
Eout    3    0    TABLE  {V(1,2)} =  (-2.5u,-2.5)  (+2.5u,+2.5)
```

This statement yields a dependent-source output between nodes 3 and 0 (as before). The controlling voltage V(1,2) is the difference between the voltages at nodes 1 and 2. If V(1,2) is -2.5 μV or less, the output is -2.5 V; if V(1,2) is +2.5 μV or greater, the output is +2.5 V. Intermediate V(1,2) values produce interpolated output levels between the ±2.5-V extremes. Note that the TABLE voltage pairs are consistent with $A_{vd} = 10^6$.

As a bonus, the new op-amp model is useful for comparator circuits—the previous model is thoroughly inadequate without power-supply limits. For the case of arbitrary V^+ and V^- comparator levels, the Eout statement can use $(V^-/A_{vd}\ , V^-)$ and $(V^+/A_{vd}\ , V^+)$ as the TABLE voltage pairs. Chapter 4 offers an alternate output limitation procedure.

Figure 1.40 shows the revised simulation results.

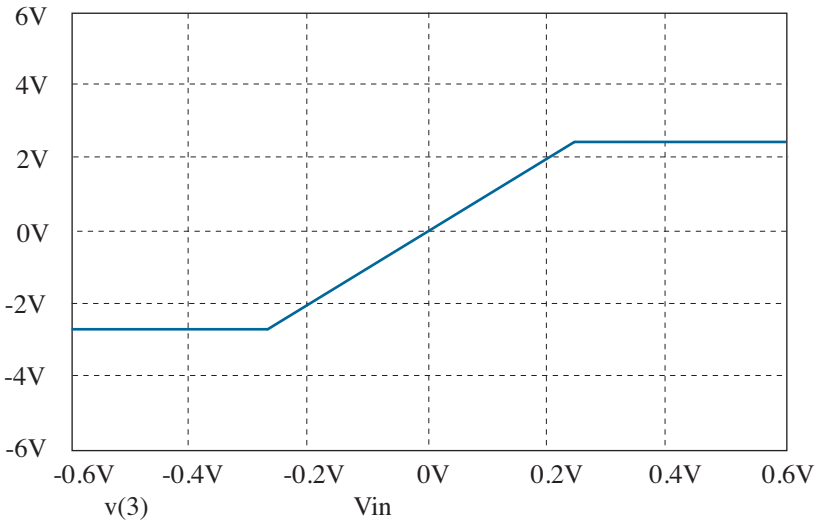


Figure 1.40: .probe results for the demonstration circuit of Fig. 1.38, revised to include the effects of ±2.5-V power supplies.

Subcircuits (Black Boxes)

Popular electronic components such as op-amps are often used repeatedly in the same circuit. Thus, when performing SPICE computer simulations, it is sometimes convenient to describe a black-box model as a **subcircuit**, thereby avoiding cumbersome repetitive netlists.

The following SPICE code features the op-amp model of Fig. 1.41.

```
* Op-Amp Demonstration Circuit

Vin       1      0      0
R1        2      0      10k
Rf        2      3      90k
X1        3      0      1      2      OpAmp

.dc       Vin    -10    +10    0.1
.probe

.subckt   OpAmp         out    com    in_p   in_n
Rin       in_p   in_n   1000MEG
Eout      out    com    TABLE  {V(in_p,in_n)} = (-2.5u,-2.5) (+2.5u,+2.5)
.ends

.end
```

The op-amp now appears as netlist component X1 whose model is described in the OpAmp subcircuit and whose four *external* connections are made in specific order to nodes 3, 0 , 1, and 2. The corresponding node names that are *internal* to the model are specified after the .subckt heading. As shown, these names do not have to be integers. The subcircuit netlist describes the circuit model of Fig. 1.41 in relation to specific internal node names. Statement .ends is the subcircuit delimiter.

"OpAmp" is a subcircuit with only two lines. Expect it to grow.

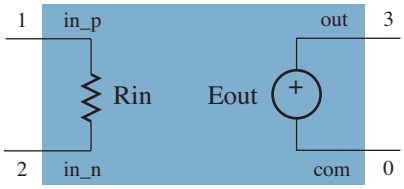


Figure 1.41: "OpAmp" subcircuit model for SPICE.

Common SPICE Statements

The following list of common SPICE statements provides support for the problems at the end of this chapter. Statement formats that accompany electronic devices will be revealed by means of example as the text unfolds.

Components

R1	**a**	**b**	**value**		Resistor R_1	end nodes a and b
C1	**a**	**b**	**value**		Capacitor C_1	end nodes a and b
L1	**a**	**b**	**value**		Inductor L_1	end nodes a and b
V1	**a**	**b**	**value**		dc Voltage Source V_1	a (plus) b (minus)
I1	**a**	**b**	**value**		dc Current Source I_1	current from a to b
V1	**a**	**b**	**ac**	**amplitude**	ac Voltage Source V_1	a (plus) b (minus)
I1	**a**	**b**	**ac**	**amplitude**	ac Current Source I_1	current from a to b

E1 **a** **b** **c** **d** A Voltage-controlled Voltage Source E_1
± output nodes a and b, ± controlling nodes c and d, gain A

G1 **a** **b** **c** **d** G Voltage-controlled Current Source G_1
current from a to b, ± controlling nodes c and d, transconductance G

Time-Dependent Sources

V1 **a** **b** **PULSE** **(IV FV TD TR TF PW PER)**
initial value IV, final value FV, delay time TD, rise time TR, fall time TF, pulse width PW, period PER — pulsed waveform

V1 **a** **b** **SIN** **(VO VA FREQ TD DF PHASE)**
offset voltage VO, peak amplitude VA, frequency FREQ, delay time TD, damping factor DF, phase angle PHASE — sinusoidal waveform

— Similar formats apply for pulsed and sinusoidal current sources (**I1** ...)

Commands

.dc **X** **IV** **FV** Δ
source name X, initial value IV, final value FV, increment Δ — dc sweep

.ac **M** **N** **IF** **FF**
Mode (**LIN** linear or **DEC** by decade) M, points (total or per decade) N, initial frequency IF, final frequency FF — ac sweep

.tran Δ_i **TF** **NP** Δ_{max}
Initial time step Δ_i, final time TF, no-print value (generally set to zero) NP, maximum time step (for precise simulations) Δ_{max}— transient analysis

— See the preceding examples for comment *, .probe, .end, etc.

Concept Summary

Black-box electronics affords simple solutions for many applications.

- Analog functions such as amplification, current-to-voltage conversion, and voltage-to-current conversion are receptive to op-amp designs.
  - The black-box op-amp obeys two rules:
    - Subject to *negative feedback*, the non-inverting and inverting input terminal voltages are equal.
    - The input terminal currents are zero.
  - The op-amp output voltage is bounded by power-supply levels.
  - Black-box non-inverting amplifiers exhibit voltage gains (v_{out}/v_t) that are unrelated to either the load resistance or the Thevenin resistance associated with an input signal source.
  - Black-box current-to-voltage converters exhibit transresistances (v_{out}/i_n) that are unrelated to either the load resistance or the Norton resistance associated with an input signal source.
  - Without feedback, the black-box op-amp acts as a comparator.
    - The comparator output assumes one of the power-supply levels depending on the relative voltages at the inputs.
    - *Positive feedback* can be used to suppress the effects of noise.
- Digital functions such as multiplexing, register shifting, and counting derive from a hierarchy of connected gates.
  - The black-box inverter has an output that is HIGH if the input is LOW, and conversely.
  - The internal circuitry of the inverter serves as the foundation for more complicated NAND and NOR gates.
  - Digital feedback is the basis for latches and flip-flops.
- The 555 timer is a mixed-signal integrated circuit that supports two modes of operation in conjunction with an external RC circuit.
  - In the monostable mode, a short input pulse initiates capacitor charging and a high-level output. The output becomes low again when the capacitor charges to 2/3 of the supply voltage.
  - In the astable mode, feedback causes the capacitor to alternate between upward charging (output high) and downward charging (output low) within the middle-third range of the supply voltage.

Some Thoughts on Design

Many of the end-of-chapter problems require some design effort, so now is as good a time as any to comment on some related philosophical issues.

- Form follows function. Cultivate a *reflexive* understanding of simple "building-block" circuits that implement common system functions. For example, attenuation often calls for a voltage or current divider, two resistors and an op-amp can combine to provide voltage gain, etc.
- Function follows form. In mathematics, the evaluation of integrals is an art because of the need to put the integrand in a *recognizable* form before a series of operations can come together to produce a solution. Circuit design is an art. Develop a "roadmap" with recognizable form.
- Know the rules: two Kirchhoff laws, $v = iR$, $i = Cdv/dt$, $v = Ldi/dt$. All design procedures hinge upon these primary analytical relations. Know the limits: failure mechanisms, power requirements, cost.
- Know the trade. Read the literature to know what just doesn't work. And for practicing engineers, read patent applications to know whether someone else has claims on your prospective design.
- Anything (within reason) is possible unless *you* have proved otherwise. Those who have accumulated "wisdom" as to what cannot be done in the absence of demonstrative evidence are unlikely to be creative. (Alas, this syndrome tends to affect older engineers.)
- Teamwork facilitates design. Nevertheless, broad access to someone else's design approach is guaranteed to minimize your creative effort. Don't go with the crowd. Let it come to you.
- Try to get something on paper immediately. Settle little details later. Forget about analysis when you need to function in a creative mode. Put aside the organizational flow of design when you need to analyze. Design and analysis involve conflicting patterns of thought.
- Design ideas, good or bad, often take time to develop. Unfortunately, bad ideas typically take even more time to die. Muster the intellectual courage to give up on designs that become interminable, unreliable, or hopelessly noncompetitive in terms of cost, maintenance, or safety.
- Your design effort does not end until it has been fully implemented. Layout-associated packaging and other incidentals that surround your circuit are parts of the circuit. Ignore them at your risk.

Keep these thoughts in mind as you progress through the text.

Have fun!

Problems

Section 1.1

Unless otherwise indicated, assume ideal op-amps with ±15-V supplies.

Op-Amp Circuit Analysis

1.1 Consider the circuit of Fig. P1.1.

(a) Determine v_{out} and node voltage v_x.

(b) What is the apparent circuit function, and what modifications (if any) lead to a better design?

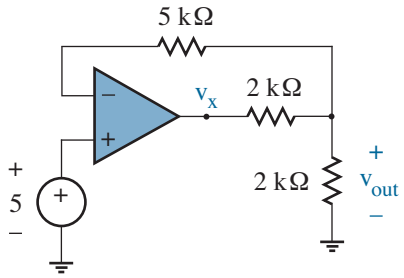


Figure P1.1

1.2 Consider the circuit of Fig. P1.2.

(a) Determine v_{out} and node voltage v_x.

(b) What is the apparent circuit function, and what modifications (if any) lead to a better design?

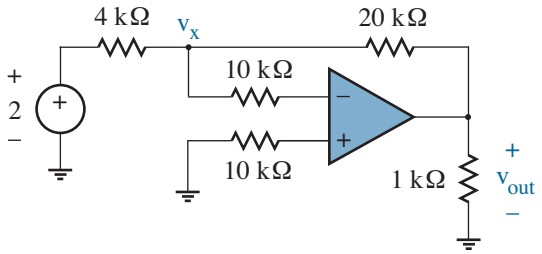


Figure P1.2

1.3 Consider the circuit of Fig. P1.3.

(a) Determine v_{out}.

(b) Determine i_x.

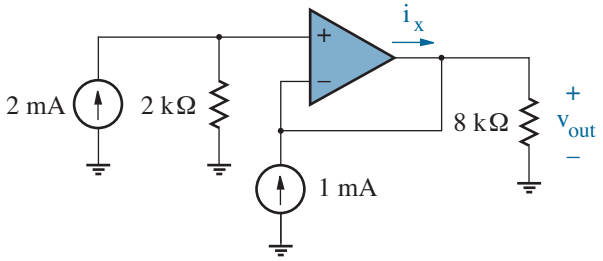


Figure P1.3

1.4 Consider the circuit of Fig. P1.4.

(a) Determine v_{out}.

(b) Determine i_x.

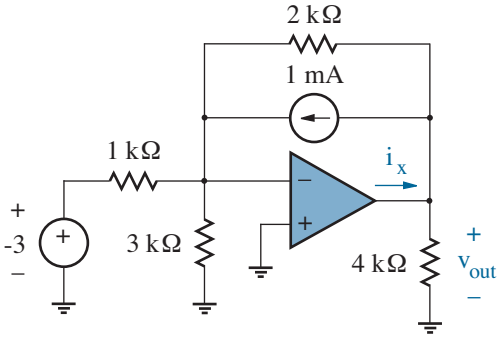


Figure P1.4

1.5 Op-amp circuits can often be analyzed by marking up the circuit diagram with simple observations. Writing equations comes as a last resort.

Consider the circuit of Fig. P1.5.

(a) Find node voltages v_a and v_b in terms of v_{out}.

(b) Find v_{out} in terms of v_{in}.

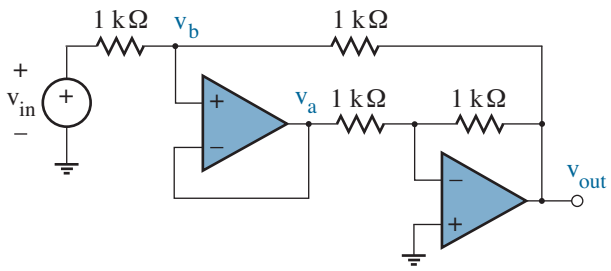


Figure P1.5

1.6 Determine node voltages v_a and v_b in the circuit of Fig. P1.6.

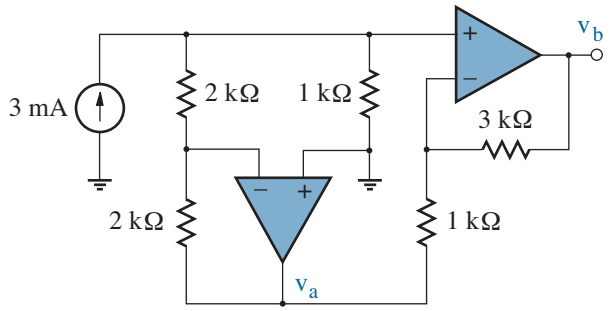


Figure P1.6

1.7 Determine node voltages v_a and v_b in the circuit of Fig. P1.7. Hint: Derive a relation for v_b in terms of v_a, then find v_a.

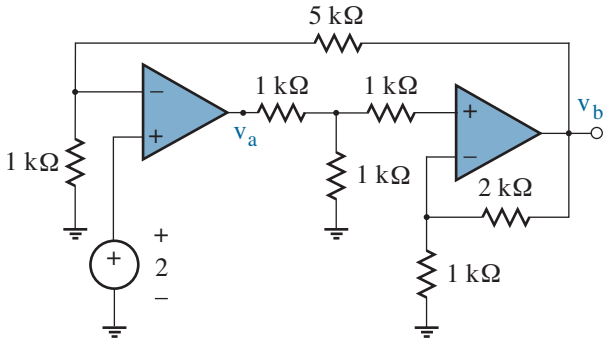


Figure P1.7

1.8 Determine node voltages v_a and v_b in the circuit of Fig. P1.8. Hint: Derive a relation for v_b in terms of v_a, then find v_a.

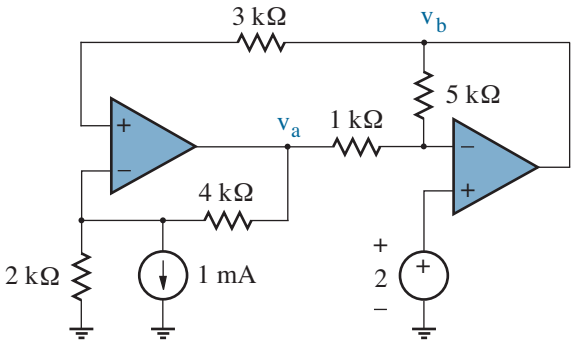


Figure P1.8

1.9 For the circuit of Fig. P1.9, sketch v_{out} vs. v_{in} (-10 V $\leq v_{in} \leq$ +10 V).

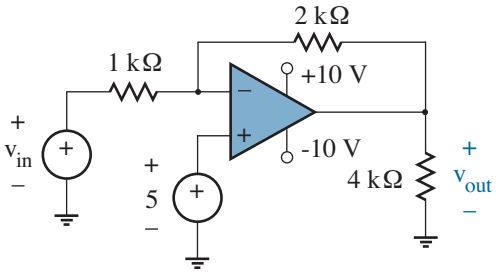


Figure P1.9

1.10 For the circuit of Fig. P1.10, sketch v_{out} vs. v_{in} (-10 V $\leq v_{in} \leq$ +10 V).

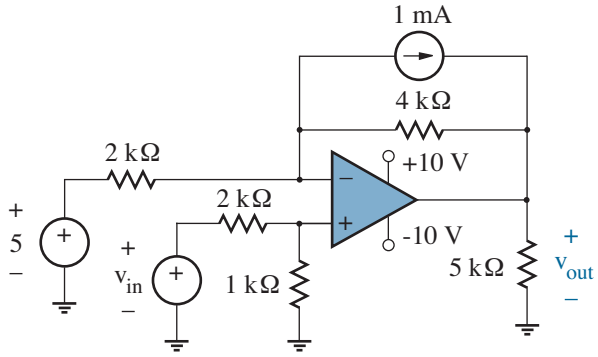


Figure P1.10

1.11 For the circuit of Fig. P1.11, sketch v_{out} vs. v_{in} (-10 V $\leq v_{in} \leq$ +10 V).

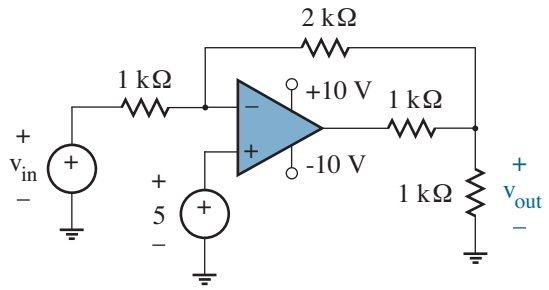


Figure P1.11

1.12 For the circuit of Fig. P1.12, sketch v_{out} vs. v_{in} (-10 V $\leq v_{in} \leq$ +10 V).

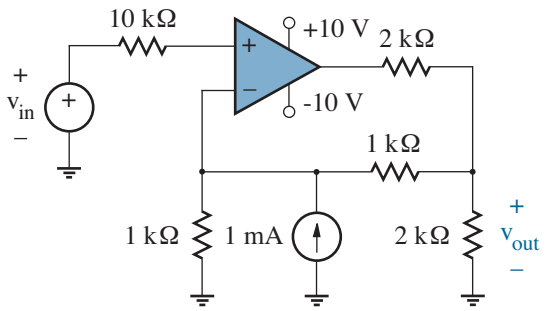


Figure P1.12

1.13 The circuit of Fig. P1.13 features a sinusoidal input $v_{in}(t) = \tilde{v} \sin 100t$. Sketch $v_{out}(t)$ subject to: (a) $\tilde{v} = 10$ mV; (b) $\tilde{v} = 100$ mV.

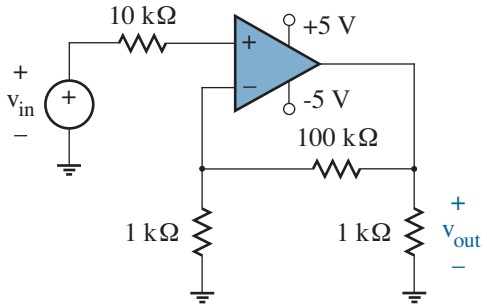


Figure P1.13

1.14 The circuit of Fig. P1.14 features a sinusoidal input $v_{in}(t) = \tilde{v} \sin 100t$. Sketch $v_{out}(t)$ subject to: (a) $\tilde{v} = 0.1$ V; (b) $\tilde{v} = 1.0$ V.

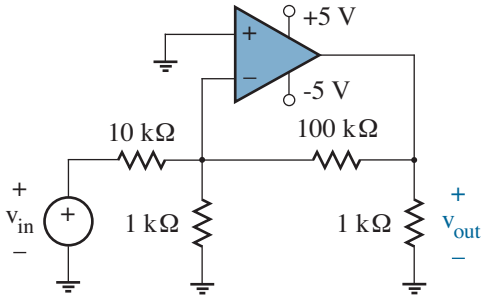


Figure P1.14

Op-Amp Circuit Design

1.15 Design non-inverting op-amp amplifiers with the following voltage gains: 20; 1 V/40 mV; 6 dB. Use 100-kΩ feedback resistors (R_f) for each design.

1.16 A signal varies between -50 mV and +250 mV. Design a non-inverting op-amp amplifier that gives linear voltage gain subject to ± 5-V power supplies. Use a 100-kΩ feedback resistor (R_f).

1.17 The circuit of Fig. P1.17 is intended to have a voltage gain between +4 and +8 depending on the position of the wiper for the 20-kΩ potentiometer. Complete the design.

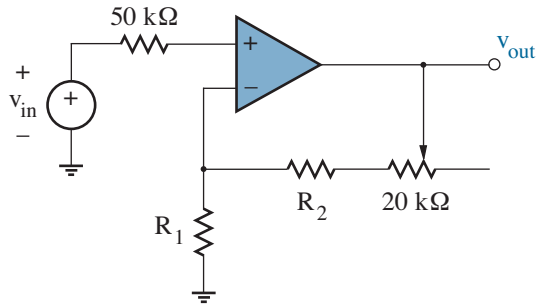


Figure P1.17

1.18 The circuit of Fig. P1.18 is intended to operate such that $v_{out} = 2v_1 + v_2$. Complete the design.

Note: This circuit suffers with v_1 and v_2 not isolated from one another.

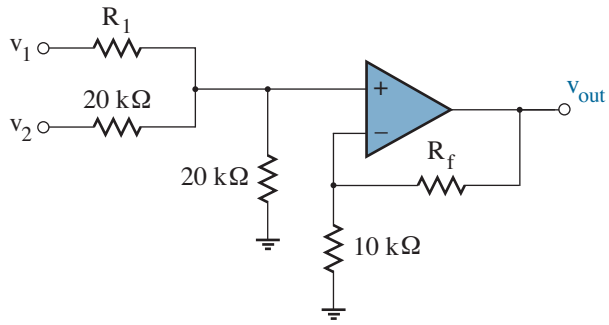


Figure P1.18

1.19 The non-inverting amplifier of Fig. 1.2a has a voltage gain that is necessarily greater than or equal to one. To achieve positive voltage gains that are less than one, a resistive voltage divider and a unity-gain buffer can be combined as shown in Fig. P1.19.

Let $v_{in} = 5$ V and $R_3 = 100\ \Omega$. Complete the design such that:

1. $v_{out} = 3$ V;
2. The power dissipated in the load R_3 is 1000 times greater than that provided by the input source v_{in}. (The power gain is 30 dB.)

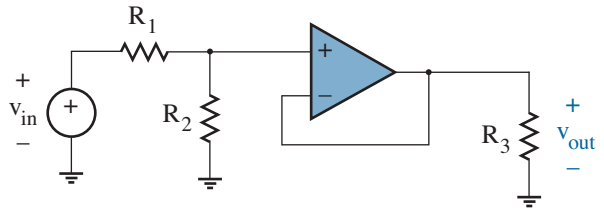


Figure P1.19

1.20

(a) Show that a non-inverting op-amp amplifier has a voltage gain that is inversely proportional to the **feedback factor** relating v^- to v_{out} —less feedback, more gain. The feedback factor for a resistive network is less than or equal to unity, so the corresponding non-inverting voltage gain is greater than or equal to unity.

(b) Given the result of part a, you embark on a non-inverting amplifier design that seeks large gain but with moderate resistor values. Complete the design of Fig. P1.20 so that the voltage gain is +440 subject to $R_1 = 10$ kΩ. Compare with a "standard" two-resistor design.

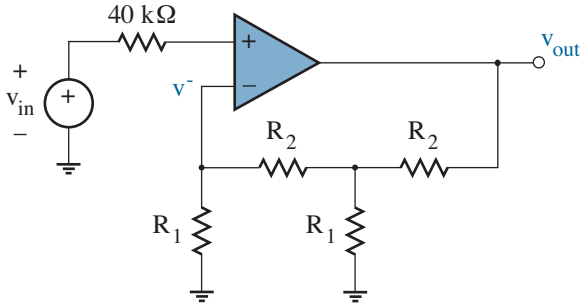


Figure P1.20

1.21 Design inverting op-amp amplifiers with the following voltage gains: -30; -4 V/50 mV; 14 dB. The signal source provides 0.1 mW when $v_{in} = 1$ V.

1.22 A signal varies between -180 mV and +300 mV. Design an inverting op-amp amplifier that provides linear voltage gain subject to ± 5-V power supplies. Use a 100-kΩ feedback resistor (R_f).

1.23 An inverting op-amp amplifier must operate with a signal source with 5-kΩ Thevenin resistance. The feedback resistor is 100 kΩ. Provide a design that yields -10 V for a 2-V Thevenin source voltage.

1.24 The circuit of Fig. P1.24 is intended to have a voltage gain between -10 and -20 depending on the position of the wiper for the 20-kΩ potentiometer. Complete the design.

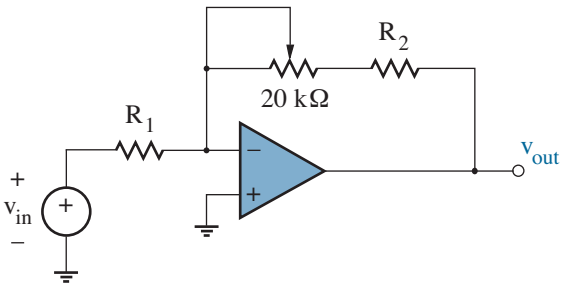


Figure P1.24

1.25

(a) Show that the transresistance of an op-amp current-to-voltage converter is inversely proportional to the **feedback factor** that relates the current i_x at the virtual ground to v_{out}. Thus, less feedback leads to more transresistance.

(b) Given the result of part a, you embark on a converter design that aims for large transresistance but with moderate resistor values. Complete the design of Fig. P1.25 for 1-MΩ transresistance. Compare with a "standard" one-resistor design.

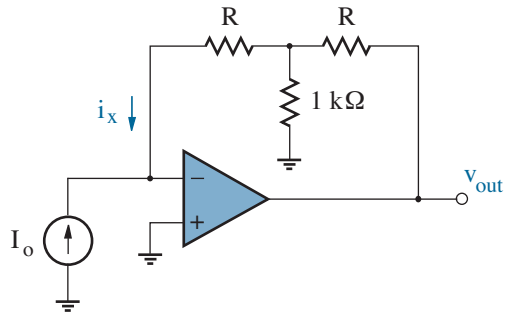


Figure P1.25

1.26 A particular temperature sensor is a current source that produces 300 μA ($T = 0$ °C) and 450 μA ($T = 100$ °C) subject to 100-kΩ Norton resistance. Design an op-amp circuit that converts this current to a voltage output of 0 V and +3 V at $T = 0$ °C and $T = 100$ °C, respectively. Use a voltage source at the non-inverting input to set the "zero" output.

1.27 Design a summing amplifier for which $v_{out} = -(v_1 + 2v_2 + 4v_3)$. Use a 100-kΩ feedback resistor.

1.28 Design a summing amplifier for which $v_{out} = -(\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 + \alpha_4 v_4)$, where $3 < \alpha_i < 5$ are adjustable factors. Use a 100-kΩ feedback resistor.

1.29 Design a summing amplifier for which $v_{out} = -(5v_1 + 2v_2 + 200i_x)$, where i_x has units of mA.

1.30 Design a summing amplifier for which $v_{out} = -5v_1 + 2v_2 - 3v_3 - 200i_x$, where i_x has units of mA. Provide designs with both one and two op-amps.

1.31 Consider the op-amp circuit of Fig. P1.31.

(a) Show that

$$v_{out} = A_{vd}\,(v_1 - v_2) + A_{vc}\left(\frac{v_1 + v_2}{2}\right),$$

where A_{vd} is the differential voltage gain and A_{vc} is the **common-mode voltage gain** that apply to the *complete* circuit. Specify A_{vd} and A_{vc} in terms of the resistor values.

(b) If the circuit is to function as a differential amplifier, the common-mode gain should be zero — the output should depend upon the *difference* between v_1 and v_2, not on the average of these two inputs. Determine the resistor relationship that ensures $A_{vc} = 0$.

(c) Find the differential gain under the constraint of part b, then complete a design so that $A_{vd} = 20$. Let $R_4 = 100$ kΩ.

(d) Determine the common-mode voltage gain that results if $R_4 = 99$ kΩ. Express the result in dB.

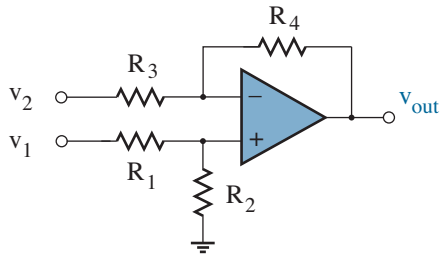


Figure P1.31

1.32 The circuit of Fig. P1.32 is intended to have

$$v_{out} = A_{vd}\,(v_1 - v_2)\ ,$$

where A_{vd} is a differential voltage gain that applies to the *complete* circuit.

(a) Determine the necessary resistor conditions, then complete a design such that $A_{vd} = 200$.

(b) The circuit allows for large A_{vd} without large resistor values. Why?

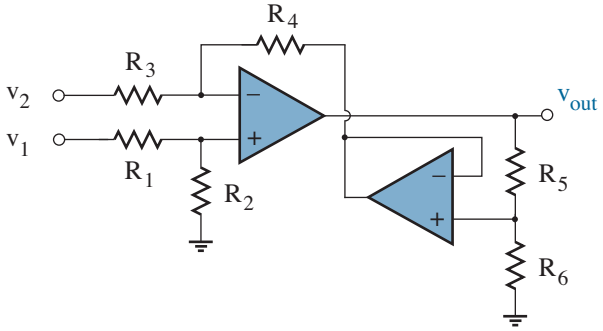


Figure P1.32

1.33 The differential amplifiers featured in Problems 1.31 and 1.32 both suffer from finite input resistance. For example, the input resistance "seen" by voltage source v_1 is $R_1 + R_2$ —hardly infinite in practice. The circuit of Fig. P1.33 presents very large input resistances to v_1 and v_2 since the voltages are applied at separate non-inverting terminals. Determine the resistor conditions such that

$$v_{out} = A_{vd}\,(v_1 - v_2)\ ,$$

then complete a design such that $A_{vd} = 50$.

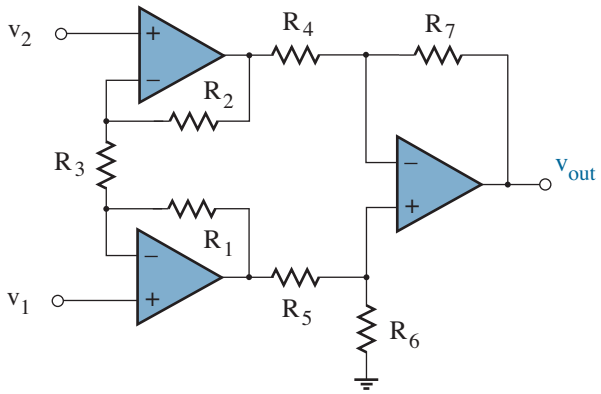


Figure P1.33

1.34 The circuit of Fig. P1.34 is intended to have

$$v_{out} = A_{vd}\,(v_1 - v_2)\ ,$$

where A_{vd} is a differential voltage gain that applies to the *complete* circuit.

(a) Determine the necessary resistor conditions, then complete the design such that $A_{vd} = 50$. Let $R_4 = 100$ kΩ.

(b) Like the circuit of Problem 1.33, this circuit provides high impedance at inputs 1 and 2, but it features only two op-amps. Suppose $v_1 = v_2$. Find the maximum common-mode (average) input signal that ensures $v_{out} = 0$ for ±15-V supply voltages. How does this compare to the circuit of Problem 1.33 with similar supplies?

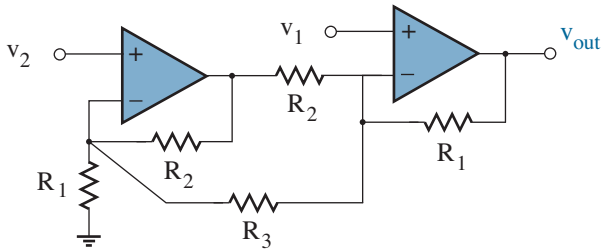


Figure P1.34

1.35 The resistor with value $R + \Delta R$ in the bridge circuit of Fig. P1.35a is a sensor element for which ΔR is proportional to a stimulus (such as force).

(a) Show that v_{out} is a *non-linear* function of ΔR.

(b) Show that v_{out} is *linear* with ΔR in the op-amp circuit of Fig. P1.35b, then complete a design so that $\Delta R = 1\ \Omega$ produces $\Delta v_{out} = 5$ mV.

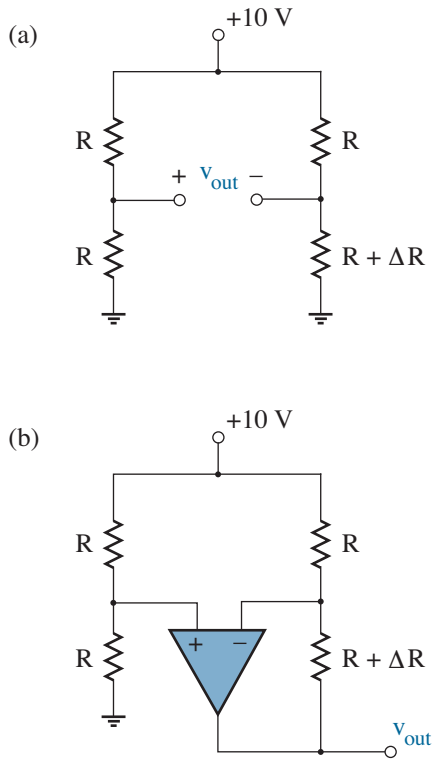


Figure P1.35

1.36 The circuit of Fig. P1.36 serves as a dependent current source for load R_L. Complete the design so that $i_{out} = 1$ mA.

Note: The circuit has practical limitations because R_L can have no connection to ground.

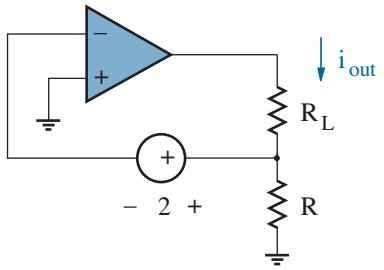


Figure P1.36

1.37 The **Howland current source** of Fig. P1.37 is intended to have

$$i_{out} = G_{md}\,(v_1 - v_2)\ ,$$

where G_{md} is a **differential transconductance** that is load *independent*.

(a) Specify the necessary resistor conditions, then complete the design such that $i_{out} = 1$ mA when $v_1 - v_2 = 1$ V.

(b) Compare the practical aspects of this circuit with that of Fig. 1.8.

Hint: Derive an expression for the voltage at the non-inverting op-amp terminal and ensure that it is proportional to R_L.

Note: This current source has one side of the load safely connected to the ground potential (in contrast to the "floating" load of Problem 1.36).

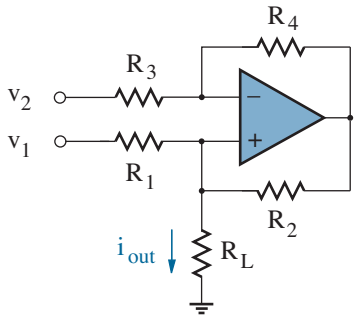


Figure P1.37

1.38 Consider the circuit of Fig. P1.38.

(a) Determine node voltages v_a and v_b in terms of node voltage v_x, then find voltage v' and the output current i_{out}.

(b) Complete a design such that $i_{out} = 1$ mA when $v_1 = -2$ V.

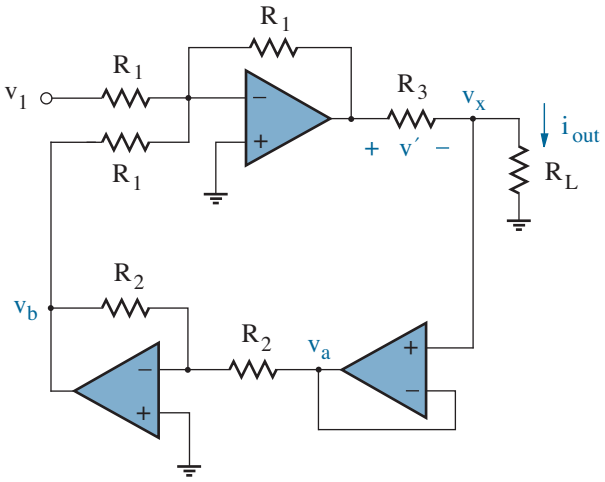


Figure P1.38

Special Operations

1.39 Consider the circuit of Fig. P1.39.

(a) Show that the circuit is an **integrator** with

$$v_{out}(t) = \frac{-1}{RC}\int_{-\infty}^{t} v_{in}(t')\,dt' .$$

(b) Suppose v_{in} is a square wave that alternates between -2 V and +2 V with a 1-ms period, and let $C = 0.1$ μF. Find R such that v_{out} has a 2-V amplitude, then sketch the output waveform.

Chapters 13 and 14 examine applications for the op-amp integrator.

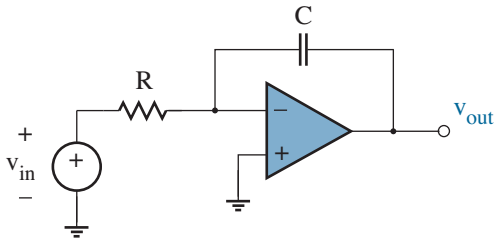


Figure P1.39

1.40 Consider the circuit of Fig. P1.40.

(a) Show that the circuit is a **differentiator** with

$$v_{out}(t) = -RC\,\frac{dv_{in}}{dt} .$$

(b) Suppose $v_{in} = \sin(2000t)$, and let $C = 0.1$ μF. Determine R such that v_{out} has a 5-V amplitude, then sketch the output waveform.

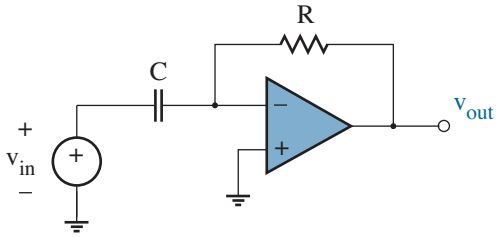


Figure P1.40

1.41 The circuit of Fig. P1.41 has periodic output. Specify the functional form of the output as well as the output frequency.

Note: This "textbook" circuit is not very practical since the output voltage amplitude is undetermined. Extreme v_{out} excursions near the power-supply levels add non-linear distortion to the idealized waveform. (You may want to build the circuit to observe the actual output characteristic.)

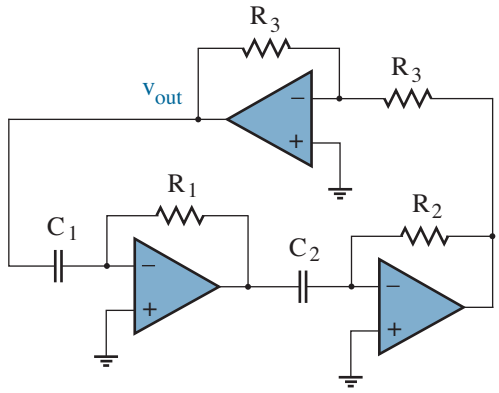


Figure P1.41

1.42 Consider the circuit of Fig. P1.42.

(a) Show that the circuit functions as a **negative impedance converter** with

$$Z_{in} = -\left(\frac{R_1}{R_2}\right) Z\,.$$

(b) Let $R_1 = R_2$, $Z = R$, and connect a capacitor (C) and inductor (L) in series with defined Z_{in}. Describe the behavior of this new circuit.

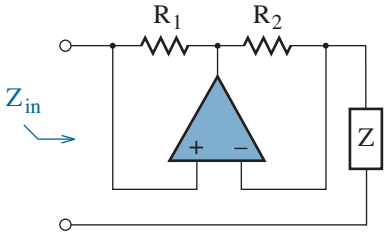


Figure P1.42

1.43 A cascade of two negative impedance converters (see Problem 1.42) yields the **Riordan circuit** of Fig. P1.43.

(a) Show that

$$Z_{in} = \left(\frac{Z_1}{Z_2}\right)\left(\frac{Z_3}{Z_4}\right) Z\,.$$

(b) Show that Z_{in} takes the form of an inductive impedance when the Riordan circuit contains four resistors and one capacitor.

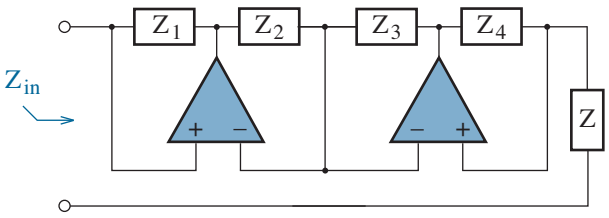


Figure P1.43

1.44 A **gyrator** is a two-port network (Fig. P1.44a) with the terminal relations

$$v_1 = -R_2 i_2\,,$$

$$i_1 = \frac{1}{R_1} v_2\,.$$

(a) Show that the circuit of Fig. P1.44b conforms to the gyrator relations.

(b) Show that the impedance looking into gyrator terminal pair 1 is that of an inductor when the other terminal pair connects to a capacitor.

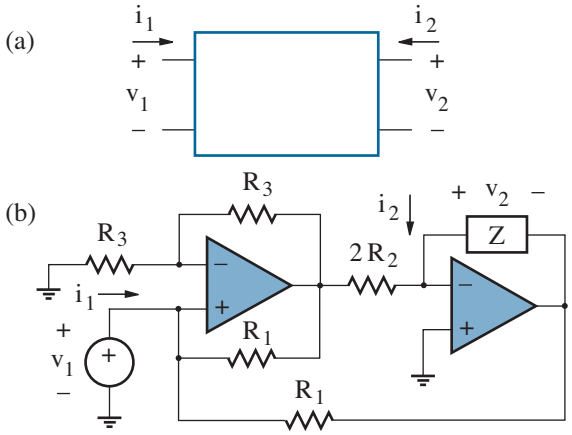


Figure P1.44

Comparators

1.45 Consider the circuit of Fig. P1.45.

(a) Determine v_{out} when $I_x = 0$.

(b) Let I_x increase from zero. What value (if any) induces a comparator output change?

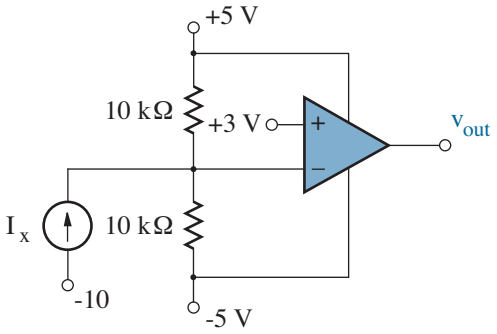


Figure P1.45

1.46 For the circuit of Fig. P1.46, sketch v_{out} vs. I_x (-10 mA $\leq I_x \leq$ +10 mA).

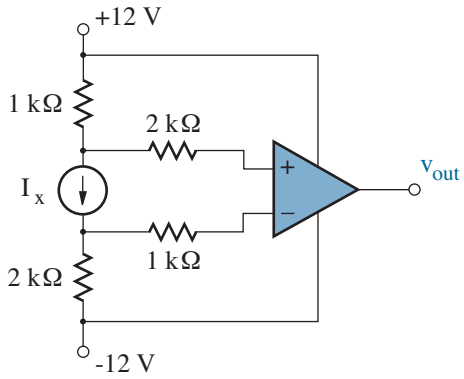


Figure P1.46

1.47 The switch in the circuit of Fig. P1.47 is closed at $t = 0$, having been open for a very long time.

(a) Determine v_{out} at $t = 0+$.

(b) Determine the subsequent time at which v_{out} makes a transition.

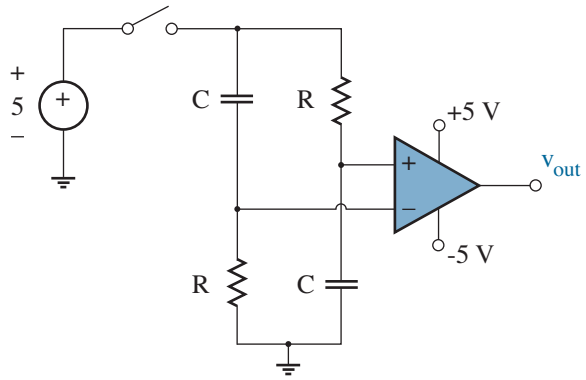


Figure P1.47

1.48 For the circuit of Fig. P1.48, sketch i_{out} vs. v_{in} (-10 V $\leq v_{in} \leq$ +10 V).

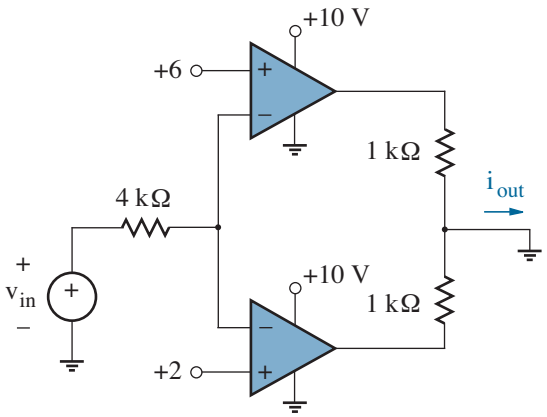


Figure P1.48

1.49 The comparators in the circuit of Fig. P1.49 feature **open-drain** outputs —these reflect special internal circuitry to be examined in later chapters. When comparator outputs tie together as shown, v_{out} is LOW when *either* component demands a LOW output state, and v_{out} is HIGH otherwise.

Design a **window detector** to produce a HIGH v_{out} when an input voltage is between 1.5 V and 2.5 V. Use the circuit of Fig. P1.49 with only three additional resistors (and 1 kΩ for at least one of them).

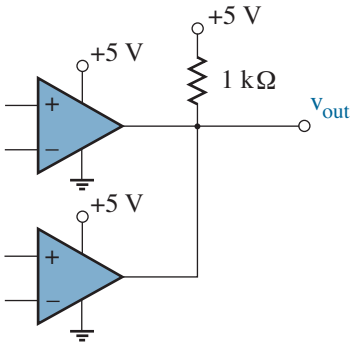


Figure P1.49

1.50 For the circuit of Fig. P1.50, sketch v_{out} vs. v_{in} as the latter increases from -5 to +12 V.

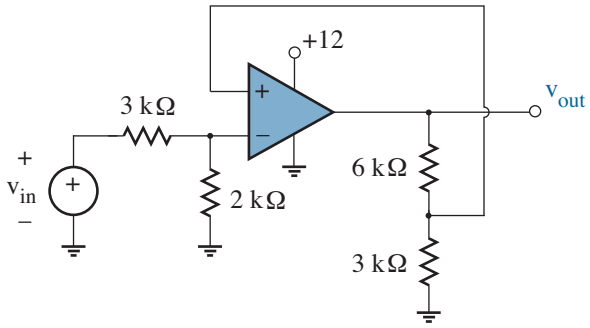


Figure P1.50

1.51 Design a non-inverting Schmitt trigger circuit (Fig. 1.11) with downward and upward transition points at -2 V and 4 V, respectively. Assume ±10-V power supplies, and let $R_f = 1$ kΩ.

1.52 Design an inverting Schmitt trigger circuit (Fig. 1.11, $v_{in} \leftrightarrow V_{ref}$) with downward and upward transition points at 4 V and 2 V, respectively. Assume ±10-V power supplies, and let $R_f = 1$ kΩ.

1.53 Consider the circuit of Fig. P1.53. The switch is open if the comparator output is LOW, and the switch is closed if the comparator output is HIGH. Show that the circuit exhibits output hysteresis and specify the transition points.

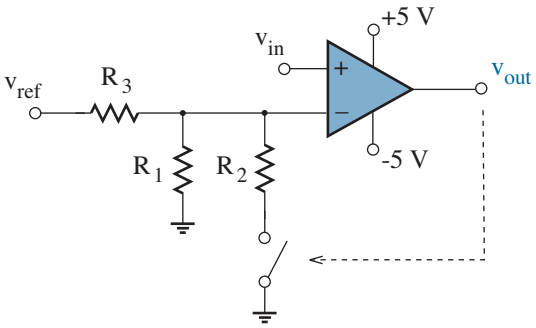


Figure P1.53

1.54 The circuit of Fig. P1.54 is called a **relaxation oscillator**.

(a) Determine v_c at the time of a -5 V to +5 V transition ($t = 0$).

(b) Determine the subsequent time at which v_{out} transitions to -5 V.

(c) Determine the subsequent time at which v_{out} transitions to +5 V.

(d) Complete a design such that the frequency of oscillation is 10 kHz.

(e) Discuss the consequences of unequal R_x values.

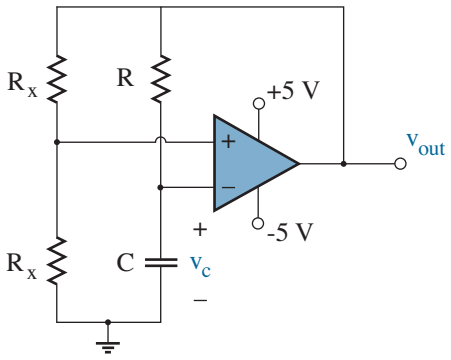


Figure P1.54

Op-Amp Limitations

1.55 The circuit of Fig. P1.55 is intended to have a voltage gain of 1001 if the op-amp is ideal.

(a) Derive an exact expression for the voltage gain in terms of op-amp gain parameter A_{vd}.

(b) Find the voltage gain when $A_{vd} = 100,000$, $A_{vd} = 1000$, $A_{vd} = 10$.

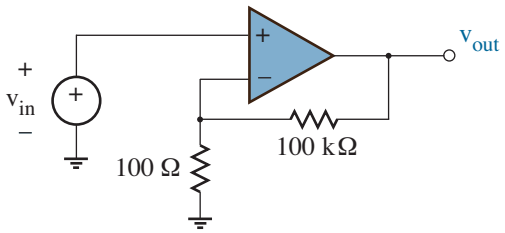


Figure P1.55

1.56 A current-to-voltage converter is intended to have 100-kΩ transresistance if the op-amp is ideal.

(a) Derive an exact expression for the transresistance in terms of op-amp gain parameter A_{vd}.

(b) Find the transresistance when $A_{vd} = 100,000$, $A_{vd} = 1000$, $A_{vd} = 10$.

1.57 A real op-amp features a finite output resistance r_o as shown in the circuit model of Fig. P1.57.

(a) Show that the Thevenin resistance looking back into the output of the non-inverting amplifier (Fig. 1.2a) is given by

$$r_{out} \approx \frac{r_o}{A_{vd}} \left(1 + \frac{R_f}{R_1}\right)$$

subject to very large A_{vd}. Thus, $r_{out} \to 0$ as $A_{vd} \to \infty$.

(b) Consider a non-inverting amplifier with $R_f = 0$, and let $v_{in} = 5$ V. Let $A_{vd} = 200,000$ and $r_o = 50\ \Omega$. Determine the load current if the load resistance maximizes power transfer.

(c) Does the result of part b make sense? Compare with the short-circuit output current for a real op-amp of your choosing. (You can find some data sheets at www.maxim-ic.com.)

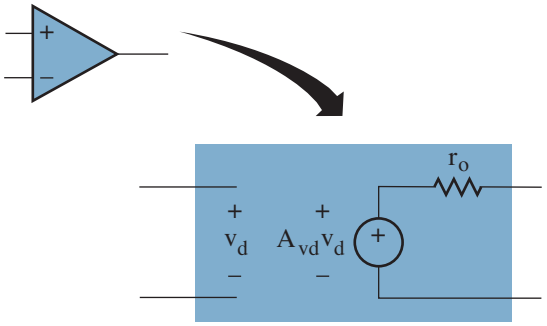


Figure P1.57

1.58 A real op-amp features an undesired input bias current i_{bias} such that i^+ and i^- are non-zero when the op-amp is connected to an external circuit. This behavior is consistent with the circuit model of Fig. P1.58a.

(a) The inverting amplifier of Fig. P1.58b is to be designed to produce $v_{out} = -(R_f/R_1)\, v_{in}$. Determine v_{out} when $i_{bias} \neq 0$.

(b) Determine the value of resistor R that minimizes the i_{bias} error.

(c) In practice, the non-ideal op-amp model also includes an offset current such that $(i^+ - i^-) = i_{offset}$ and $(i^+ + i^-)/2 = i_{bias}$. Show how the circuit model of Fig. P1.58a should be revised to include the effects of i_{offset}.

(d) Let R have the value found in part b. Find v_{out} when $i_{offset} \neq 0$. Suggest a way to minimize the current offset error.

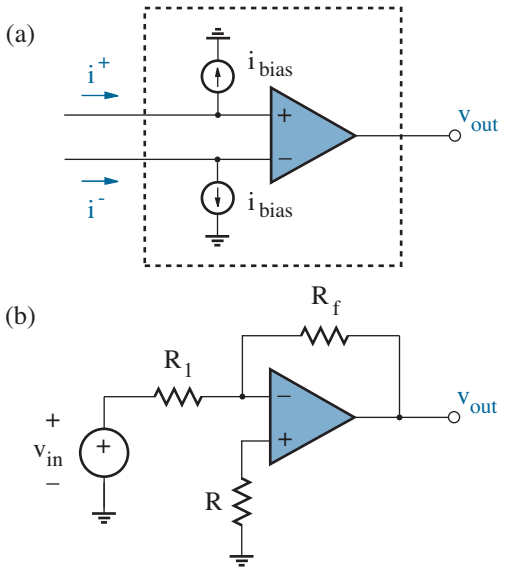


Figure P1.58

1.59 A real op-amp features an undesired input offset voltage v_{os} such that $v_{out} = A_{vd}v_{os}$ when $v^+ = v^-$. This behavior is consistent with the circuit model of Fig. 1.13.

(a) Determine v_{out} for the circuit of Fig. P1.59 if $v_{os} = 2$ mV.

(b) Demonstrate how the circuit of Fig. P1.59 can be modified to eliminate the effect of v_{os} using only a resistor and an adjustable voltage source. (The latter can be obtained from the wiper of a potentiometer with its ends at V^+ and V^-.)

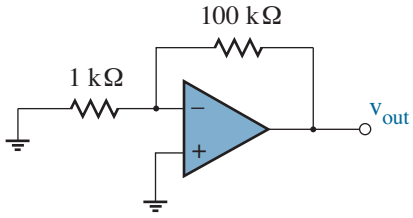


Figure P1.59

1.60 For some reason, you decide to use the op-amp integrator of Fig. P1.60 in a clock circuit. Switch S is opened at noon, and the output voltage reflects the elapsed time with 1 V = 1 hour. At midnight, when $v_{out} = 12$ V, the clock resets by momentarily closing S. The clock resets again the following noon.

(a) Let $R = 10$ MΩ. Determine the necessary value for C.

(b) Suppose the op-amp is non-ideal with a particular offset voltage v_{os} at the non-inverting input (see Fig. 1.13). Find the maximum acceptable v_{os} that limits the clock error to 1s over 12 hours.

(c) Suppose the op-amp is non-ideal with current i_{bias} at the inverting input (see Problem 1.58). Find the maximum acceptable i_{bias} that limits the clock error to 1s over 12 hours.

(d) Consider a non-ideal switch with parallel resistance R_s. Find the minimum acceptable R_s that limits the clock error to 1s over 12 hours.

Note: This problem is dedicated to the memory of Thomas Durgavich, who, as an engineering student, took unusual interest in the practical engineering aspects of a peculiar clock design.

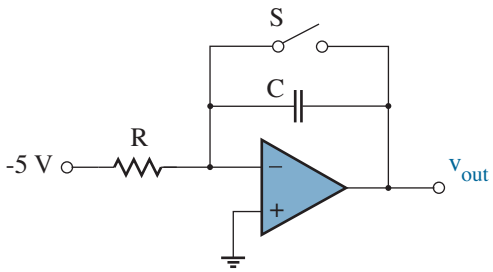


Figure P1.60

1.61 This problem concerns the potential adverse behavior of op-amp circuits that lack bypass capacitors for the V^+ and V^- power supplies.

Consider a non-inverting op-amp amplifier with zero input and (without loss of generality) zero no-load input offset voltage. The op-amp is characterized by a **Power Supply Rejection Ratio (PSRR)** defined

as the ratio of the change in input offset voltage to the change in power-supply voltage that produces it. While counterintuitive, a *decrease* in the actual V^+ supply voltage reflects an *increase* in the input offset voltage as we have defined it. Thus, an upward output excursion that draws current from the V^+ supply imparts the voltage change $-i|Z|$ at the positive supply terminal and $+i|Z|\,PSRR$ for the corresponding input offset voltage source as shown in Fig. P1.61. Parameter Z is the Thevenin impedance associated with the V^+ supply.

(a) Apply a stability argument (as in the discussion following Eq. 1.8) to show that stability requires

$$|Z| < \frac{R_L}{PSRR}\left(\frac{R_1}{R_1 + R_f}\right) .$$

Assume $R_L \ll R_1 + R_f$.

(b) Let $Z = R_T + \jmath\omega L_T$ with $R_T = 0.1\ \Omega$ and $L_T = 100$ nH. Further, assume that the power supply rejection ratio varies with frequency according to the approximate empirical relation

$$PSRR = 10^{-4}\, f/f_o\,,$$

where $f_o = 200$ Hz (see data for the MAX4400). At what frequency is the stability condition for Z violated if the non-inverting op-amp voltage gain is 100? What is the reactive component of Z at this frequency? Express the result in Ω.

(c) Now add capacitor C from node X to ground. Find the C value that reduces the modified $|Z|$ by an order of magnitude (10x) when $V^+ \to 0$.

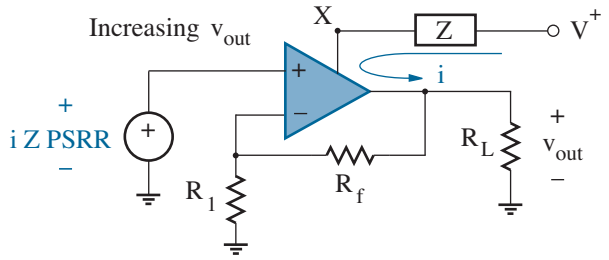


Figure P1.61

Section 1.2

1.62 The circuit of Fig. P1.62 has three digital inputs a, b, and c. Specify a truth table that shows the condition of output y for each input combination.

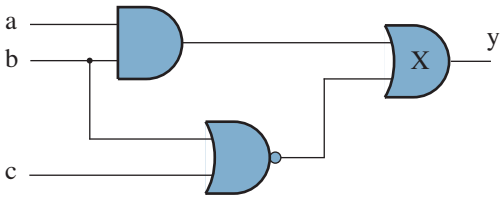


Figure P1.62

1.63 Repeat Problem 1.62, but replace OR gate X with a NAND gate.

1.64 A small circle near the input of a digital gate represents a logical inversion of that input.

(a) Consider AND and OR gates whose inputs are inverted as shown in Figs. P1.64a and P1.64b, respectively. For each case, determine an equivalent gate with non-inverted inputs.

(b) Redesign the circuit of Problem 1.62 so that it features only inverter, NAND, or NOR gates.

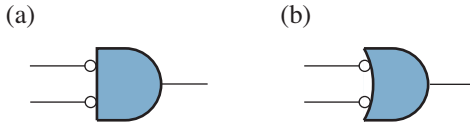


Figure P1.64

1.65 Design digital circuits to realize the following logical functions:

(a) $y = (a + \bar{b}) \cdot c$

(b) $y = a + (b \cdot \bar{c})$

Note: $+$ denotes the OR operation, $\cdot$ denotes the AND operation.

1.66 The **exclusive OR** (XOR) operation conforms to the following truth table:

a	b	$a \otimes b$
0	0	0
0	1	1
1	0	1
1	1	0

Design a digital circuit that implements the XOR function using only inverter, NAND, or NOR gates (see Problem 1.64).

1.67 Sometimes a gate has more inputs than needed.

(a) Show how a one-input inverter function can be realized from: a two-input NAND gate; a two-input NOR gate.

(b) Consider a 4-input NAND gate. Determine the 3-input function that results when: one input is LOW; one input is HIGH.

(c) Consider a 4-input NOR gate. Determine the 3-input function that results when: one input is LOW; one input is HIGH.

1.68 Design a muliplexer/demultiplexer (MUX) that steers one of four pulse trains to one of two outputs.

1.69 Design a muliplexer/demultiplexer (MUX) that steers one of two pulse trains to one of three outputs.

1.70 Design a muliplexer/demultiplexer (MUX) that steers digital input pair a and b into the output pair y and z. In a separate mode of operation, the circuit is required to steer input pair a and c into output pair z and y.

1.71 Consider the cross-coupled NOR gates shown in Fig. P1.71.

(a) Show that the circuit is a bistable RS latch, and specify the R and S inputs to reset and set Q_1.

(b) What input condition has no practical value?

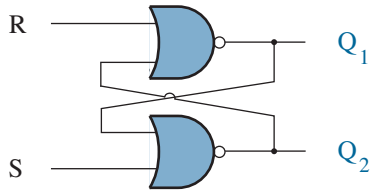


Figure P1.71

1.72 Design an RS flip-flop with cross-coupled NOR gates (see Problem P1.71).

1.73 Consider the **debounce** circuit of Fig. P1.73. After remaining in position 1 for a long period, switch S is suddenly toggled at time $t = 0$. The ground-connected arm ties to position 2 after 5-ms delay. However, mechanical "bounce" causes the arm to break contact at $t = 5.5$ ms and resume contact indefinitely at $t = 6$ ms.

Assume instantaneous gate response, assume zero currents into gate inputs, and let +5 V = HIGH, 0 V = LOW.

(a) Sketch the time dependence of the logic levels at nodes 1 and 2.

(b) Sketch the time dependence of the logic level at output Q.

(c) Describe the (very poor) switch behavior that defeats the intent of the debounce circuit.

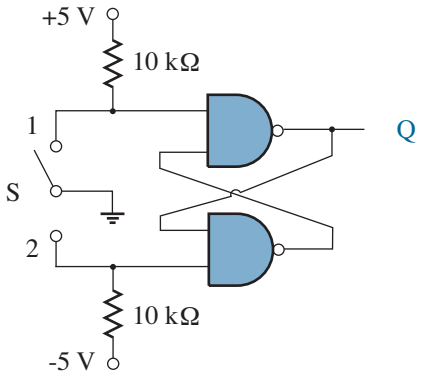


Figure P1.73

1.74 Figure P1.74 shows two RS flip-flops in the **master-slave** configuration.

(a) Sketch the time dependence of the logic levels at Q_1, Q_2, Q_3, and Q_4 when R, S, and clock-pulse CK have the waveforms shown.

(b) Describe the circuit function in relation to the upward clock edge.

(c) Describe the circuit function in relation to the downward clock edge.

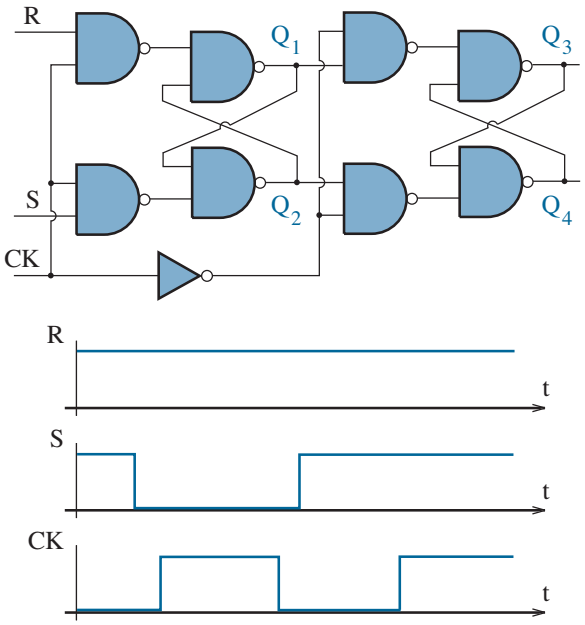


Figure P1.74

1.75 Design an RS flip-flop with two additional inputs $CLEAR$ and $PRESET$. $CLEAR$ has the same function as input R and $PRESET$ has the same function as input S; however, *both* additional inputs force output changes regardless of the clock status. Hint: Apply two 3-input NAND gates.

1.76 The D flip-flop shown in Fig. 1.25 has two cross-connected NAND gates. Show what circuit modifications are necessary to achieve D flip-flop action when the cross-connected elements are NOR gates.

1.77 Suppose each gate in the T flip-flop of Fig. 1.26 has a **propagation delay** of 10 ns —this is the time needed to effect an output change after any input changes state. Assume Q = HIGH and $\overline{Q}$ = LOW for all $t < 0$.

(a) The clock goes HIGH at $t = 0$ for 100 ns. Sketch the time dependence of the Q and $\overline{Q}$ ouputs.

(b) What clock restrictions does the preceding behavior suggest?

(c) Show that the clock restrictions can be resolved when the T flip-flop derives from a master-slave RS flip-flop (see Problem 1.74).

1.78 Figure P1.78 shows a **JK flip-flop**. Show that the circuit function as either a D flip-flop or a T flip-flop subject to proper choices for inputs J and K. Assume short clock pulses.

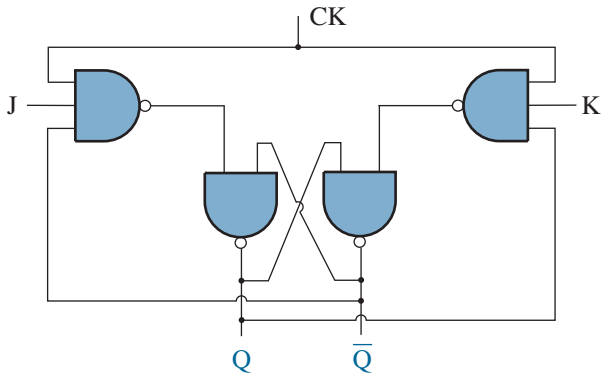


Figure P1.78

1.79 Consider the four-bit shift register shown in Fig. 1.27. Find the sequence of Q_3, Q_2, Q_1, and Q_0 outputs if the "Data" input is:

(a) $\overline{Q_2} + Q_1$

(b) $Q_2 \cdot \overline{Q_1} \cdot Q_0$

Assume that each Q output is initially LOW.

Note: $+$ denotes the OR operation, $\cdot$ denotes the AND operation.

1.80 Design a circuit that loads data into a 4-bit shift register after four clock pulses, then clears the shift register after the fifth clock pulse. The process repeats over five clock pulses. Assume that each flip-flop in the shift register has a CLEAR input that induces 0 content when LOW.

1.81 Determine the action of the circuit of Fig. P1.81 for regular clock pulses.

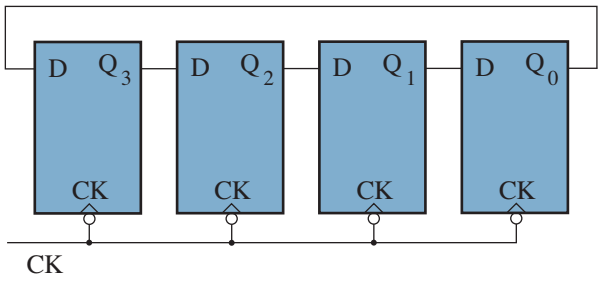


Figure P1.81

1.82 Modify the four-bit counter circuit of Fig. 1.28 to count *downwards* with each clock pulse.

1.83 Design a 4-bit up-down counter that counts upwards when bit X is HIGH and downwards when bit X is LOW.

Section 1.3

1.84 Design a 555 timer circuit that produces a 10-s output pulse in response to a shorter trigger pulse. Use at least one 47-μF capacitor.

1.85 Design a 555 timer circuit that produces an output pulse between 50-ms and 100-ms duration depending on the state of a 20-kΩ potentiometer.

1.86 The monostable or one-shot 555 timer circuit of Fig. 1.34 requires a trigger pulse that is shorter than the intended output pulse, a disadvantage for some applications. Show that the circuit of Fig. P1.86 can accommodate any trigger pulse when R_x and C_x are appropriately chosen.

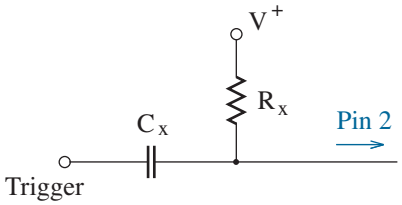


Figure P1.86

1.87 Design a 555 timer circuit that produces a 10-kHz periodic waveform with 60% duty cycle. Use at least one 0.33-μF capacitor.

1.88 Design a circuit that controls windshield wipers. The circuit must produce a 5-V pulse of 1-s duration at 2-s, 4-s, or variable 6-10-s intervals. Assume that the 5-V pulse interacts with other circuitry to move the wipers through one cycle.

1.89 Consider an odd 555 timer whose internal circuitry (Fig. 1.37) features an R-$2R$-R resistor chain. Derive an appropriate design equation for a one-shot application (comparable to Eq. 1.30).

1.90 Consider an odd 555 timer whose internal circuitry (Fig. 1.37) features an R-$2R$-R resistor chain. Derive the appropriate design equations for the frequency and duty cycle of an astable waveform.

1.91 This problem concerns an astable 555 timer whose voltage at the Control terminal (pin 5) is set to some particular value (see Fig. 1.37).

(a) Let the Control voltage equal v_x. Determine the time duration of the LOW portion of the output cycle in terms of R_1, R_2, C, and v_x.

(b) Let the Control voltage equal v_x. Determine the time duration of the HIGH portion of the output cycle in terms of R_1, R_2, C, and v_x.

(c) Determine a consistent frequency for the output waveform.

(d) Suggest an application for a circuit with variable Control voltage.

Section 1.4

1.92 Use SPICE to demonstrate the input-output transfer characteristic for the current-to-voltage converter of Fig. 1.4. Let $R_f = 2$ kΩ, $R_n = 100$ kΩ, and sweep i_n from -10 mA to +10 mA. The simulation should feature an ideal op-amp model with ±15-V power supplies.

1.93 Use SPICE to demonstrate the input-output transfer characteristic for the inverting amplifier of Fig. 1.5. Let $R_1 = 1$ kΩ and $R_f = 10$ kΩ. The ideal op-amp model is subject to ±5-V power supplies.

1.94 Simulate the Problem-1.9 circuit with SPICE.

1.95 Simulate the Problem-1.10 circuit with SPICE.

1.96 Simulate the Problem-1.11 circuit with SPICE.

1.97 Simulate the Problem-1.12 circuit with SPICE.

1.98 Use SPICE to demonstrate the two-op-amp circuit of Fig. 1.8. Let $R_1 = R_2 = R_3 = 1$ kΩ, and let $R_L = 100$ Ω. Examine the output current i_{out} as v_1 varies between -20 V and +20 V. The simulation should have the "OpAmp" subcircuit of Section 1.4.

1.99 Use SPICE to demonstrate the three-op-amp circuit shown in Problem 1.38. Let $R_1 = R_2 = R_3 = 1$ kΩ, and let $R_L = 100$ Ω. Examine the output current i_{out} as v_1 varies between -20 V and +20 V. Apply the "OpAmp" subcircuit of Section 1.4.

1.100 Use SPICE to verify the action of the integrator circuit of Problem 1.39. Let $R = 1$ kΩ, $C = 10$ μF, and $v_{in} = 2$ V. Examine the output time dependence over the interval [0, 100 ms] using the command

```
.tran        1n       100m
```

Assume an ideal op-amp with ±15-V power supplies.

1.101 Use SPICE to demonstrate the differentiator circuit of Problem 1.40. Let $R = 1$ kΩ, $C = 10$ μF, and $v_{in} = 2\cos 20t$ (V). Examine the output time variation on the interval [0, 100 ms] (Problem 1.98). Assume an ideal op-amp with ±15-V power supplies.

1.102 Use SPICE to demonstrate the input-output transfer function for the Schmitt trigger of Fig. 1.11. Let $R = 2$ kΩ, $R_f = 10$ kΩ, and $v_{ref} = 1$ V. The simulation should feature a comparator with HIGH and LOW output levels of 5 V and 0 V (ground), respectively. Assume a differential voltage gain of 50,000 when v_{out} is between these levels.

Note: To observe the output hysteresis, one must perform two input sweeps: the first from a low to high voltage, the second from a high to low voltage.

1.103 This problem concerns an op-amp with some non-ideal features.

(a) Revise the "OpAmp" subcircuit of Section 1.4 to account for the following observations:

  (a) When the external inputs are connected together, the output reacts as it would for a differential input of 0.5 mV.

  (b) Each external input draws 100-pA constant current under all connection conditions.

  (c) The Thevenin resistance at the external output is 100 Ω.

(b) Use the revised model of part a to simulate a non-inverting amplifier with $R_1 = R_f = 10$ MΩ (see Fig. 1.2).

(c) Repeat the simulation, but use the relatively ideal "OpAmp" model. Compare the results.

Perspective: One-Way Current

Imagine you are an electrical engineer in the early days of the profession, say 1880. Yours is a primitive world of knife switches, batteries, resistors, condensers (capacitors) and coils (inductors). Nevertheless, it is an exciting world of telephones, wireless telegraphs, and other emerging technologies that promise revolutionary changes for society—no less so than cell phones and iWhatevers of modern times.

One day, you read a brief report of an unusual two-terminal resistor that has been the subject of recent experiments. The details are obscure: the resistor is formed through the point contact of two materials, and unlike the familiar $i = v/R$ relationship for an ordinary resistor, it has a *non-linear* current-voltage characteristic of the form $i = f(v)$. Your interest is aroused. What possible use could there be for this type of component?

As now, you are thoroughly familiar with the analysis of dc networks. So you ponder the simple circuit shown in Fig. A1. The linear portion to the left of the dashed line is the Thevenin equivalent of a circuit with arbitrary complexity, and the non-linear portion to the right represents the unusual two-terminal device for which $i = f(v)$. You cautiously mark the schematic to distinguish one of the device terminals, and you assume $V_t > 0$.

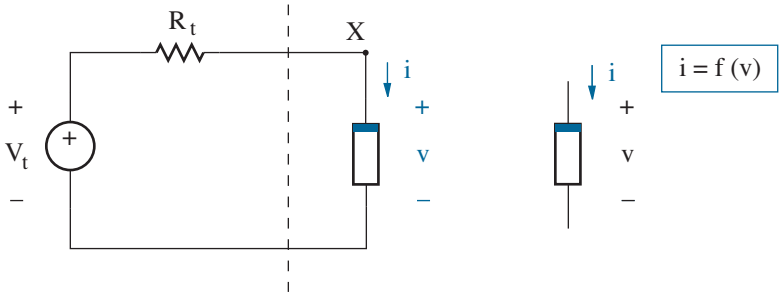


Figure A1: Simple circuit featuring a two-terminal non-linear device with the current-voltage characteristic $i = f(v)$.

You resolve to determine the voltage v. Thus, you express Kirchhoff's Current Law at node X in terms of v. Specifically,

$$\frac{V_t - v}{R_t} = f(v)\,. \tag{A1}$$

However, you note that Eq. A1 does not represent much progress, since it will be difficult to obtain an algebraic solution for v when $f(v)$ is non-trivial. You are somewhat confident that you can obtain a numerical solution through trial and error—high-speed computers are a distant dream in 1880. But the results of this analysis will do little to improve your understanding of circuit performance as V_t or R_t are changed.

Not wanting to give up too easily, you attempt a graphical solution. First, as shown in Fig. A2, you plot the constraint that the non-linear device imposes between the electrical variables i and v:

$$i = f(v)\,. \tag{A2}$$

You are uncertain about the form of this **device characteristic**; however, your report suggests the shape of a boomerang wrapped about the origin. Next, you plot the constraint that the linear portion of the circuit imposes between the *same* electrical variables:

$$i = \frac{V_t - v}{R_t}\,. \tag{A3}$$

Equation A3 defines a straight line, which you call the **load line** (taking the perspective of the non-linear device that is connected to a Thevenin "load"). Finally, you note that both constraints are satisfied at an intersection point. You are working with a dc circuit at rest, so you designate this special point as the **quiescent operating point** or **Q-point**. The solutions for i and v are the Q-point current and voltage co-ordinates, respectively.

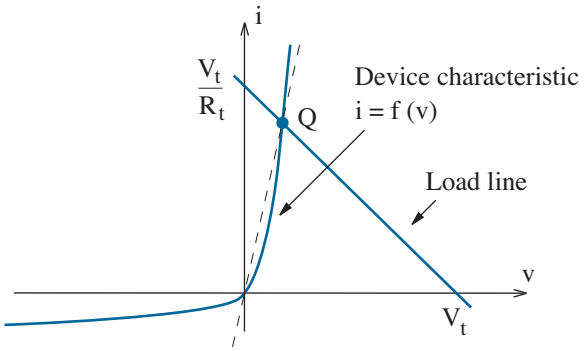


Figure A2: Graphical solution for the circuit of Fig. A1.

The non-linear device does not impress you. If one only wanted to have a particular Q-point, the current-voltage characteristic of a linear resistor (the dashed line in Fig. A2) would have been sufficient.

With the hope that the non-linear device has more to offer, you repeat the analysis of Fig. A1, but you reconfigure the circuit so that the device is upside-down as in Fig. A3a. You maintain the i and v orientations relative to the linear portion of the circuit so that the non-linear device is effectively described by the relation $-i = f(-v)$.

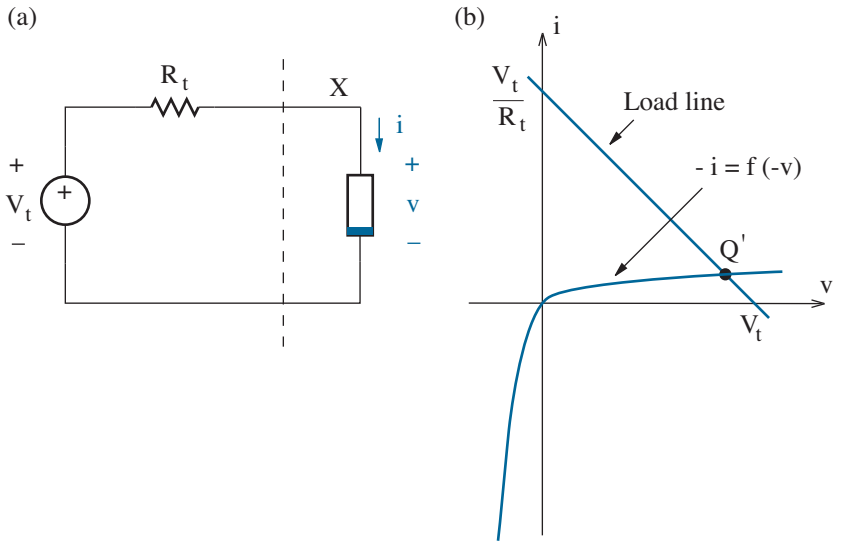


Figure A3: Reconfigured non-linear circuit and its graphical solution.

Figure A3b shows your new graphical solution. It features the same load line as before, but the current-voltage constraint that is imposed by the non-linear device has been modified. Specifically, the device characteristic of Fig. A2 has been rotated around the current axis and then the voltage axis by 180^o to yield a new Q-point (Q′). You note that the original Q-point would be preserved if you had changed the orientation of a linear resistor, since $-i = -v/R$ and $i = v/R$ describe identical characteristic curves.

Thus, you find that a two-terminal non-linear device with a boomerang current-voltage characteristic is potentially useful by virtue of its polarity. One device polarity yields a small voltage and a large current (Q), whereas the other device polarity yields a large voltage and a small current (Q′). Surely some applications must follow.

Back to the present— An electrical device that exhibits the preceding behavior is called a **diode** (Greek: *di*, two + *ōdēs*, path).

In an extreme case, which applies when a polarity change shifts the Q-point between current- and voltage-axis intercepts at the load-line ends, we observe action of an **ideal diode** with the current-voltage characteristic shown in Fig. A4. The diode symbol appears in the figure inset.

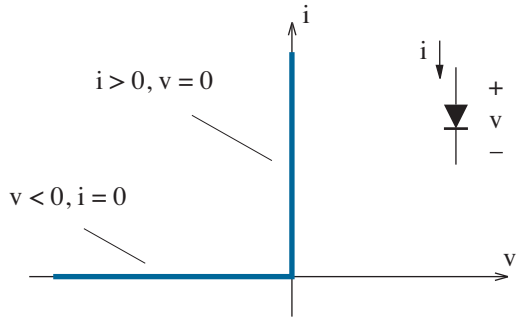


Figure A4: Current-voltage characteristic of an ideal diode.

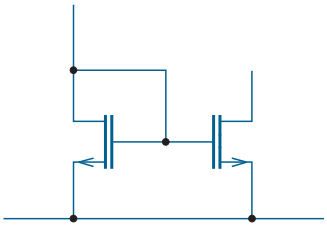
An ideal diode is a one-way street for current. It supports short-circuit passage in one direction (no voltage-drop penalty) and total open-circuit blockage in the other direction (despite arbitrarily large voltage "pressure"). Both conditions are duplicated with a lossless switch. However, as a non-polar device, the switch changes state through the action of a mechanical or electrical stimulus that is independent of its *i-v* terminal connections. In contrast, *the ideal diode changes state through reversal of the voltage polarity applied at its i-v terminal pair.* (Convince yourself that this is true by repeating the graphical analysis of Fig. A1 with $V_t < 0$.)

Back to reality— How can one obtain a physical device that exhibits the ideal diode characteristic? At best, we can only hope to approximate ideal diode behavior, and we consider the features of one approach in Chapter 2. A supporting discussion of semiconductor principles provides the necessary background for more powerful electronic devices in Chapter 5 and beyond. (For some readers, this discussion will be a helpful review.)

Chapter 3 builds upon the informative but awkward graphical process used to demonstrate diode action in a simple circuit. We need to acquire new skills and patterns of thought for the efficient analysis of diode circuits (and non-linear circuits in general). Development and application of device models with varying degrees of complexity is crucial to this endeavor.

Chapter 4 examines diode applications, some old, and some not so old, but all important for discrete and integrated circuits. This will be welcome, as we return to physical principles in Chapters 5 and 6.

Problems relating to this Perspective appear at the end of Chapter 2.



Chapter 2

Physical Foundations

The ideal diode reflects a simple electrical concept—one-way current flow. How can we best exploit the physical world to achieve real diode behavior? And what performance limitations must we endure?

This chapter begins by examining some electrical properties of metals, semiconductors, and insulators. Then we demonstrate the current-voltage characteristic of the pn junction diode, a fundamental device that is formed by joining two different "flavors" of semiconductor. We finish the chapter with a practical discussion of pn junction forms and the relative behavior of metal/semiconductor and optoelectronic diode devices.

Objectives

After studying this chapter and working through the problems, you should be able to perform each of the following:

- Describe the process of electrical conduction in a metal, an intrinsic semiconductor, and an insulator (Section 2.1).
- Contrast n-type and p-type extrinsic semiconductors (Section 2.1).
- Explain the origin of depletion regions and the associated built-in potential barrier in a pn junction diode (Section 2.2).
- Specify the elementary current-voltage characteristic for a pn junction diode, and indicate the important features of forward- and reverse-bias diode operation (Section 2.3).
- Explain how the current-voltage characteristics of a pn junction vary with temperature (Section 2.3).
- Describe a metal-semiconductor Schottky diode (Section 2.4).
- Describe light-emitting diodes and photodiodes (Section 2.4).

2.1 Solid-State Materials

Before we attempt to understand the physical behavior of solid-state diodes (amongst other devices), we must examine the electrical conduction process in three basic material categories: metals, semiconductors, and insulators. Our study requires frequent reference to the periodic table of the elements, so we promptly introduce a portion of it in the distorted form of Fig. 2.1. The sizes of the boxes and the lettering reflect the relative importance of the elements to the electronics industry.

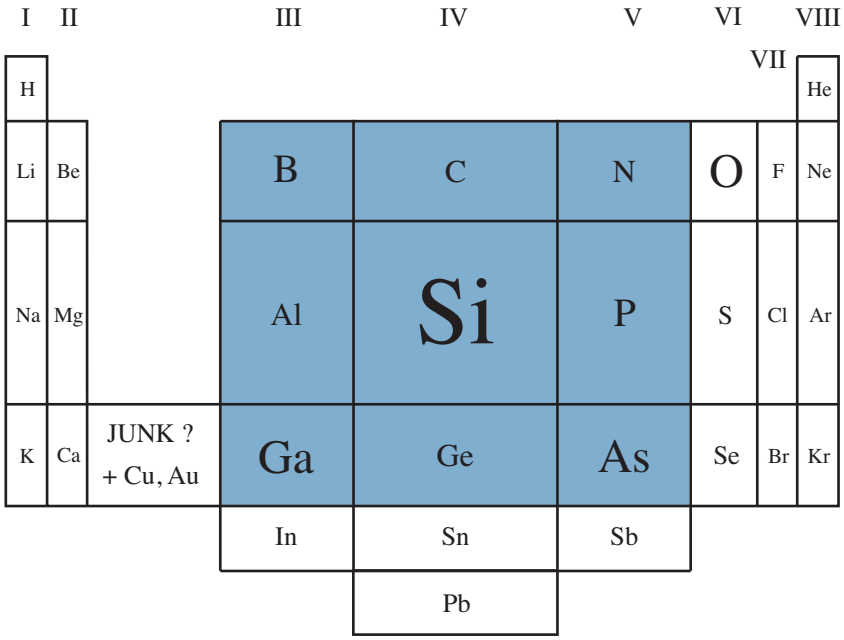


Figure 2.1: Periodic table of the elements for an electronics engineer.

Metals

While highly reactive and otherwise useless in form, sodium (Na) is a metal whose electronic properties are easily understood. On a microscopic scale, sodium ideally consists of a unique and regular arrangement of atoms in a three-dimensional crystal lattice. Figure 2.2a shows the body-centered cubic sodium lattice structure, and Fig. 2.2b shows a simplified projection of this lattice onto a two-dimensional page.

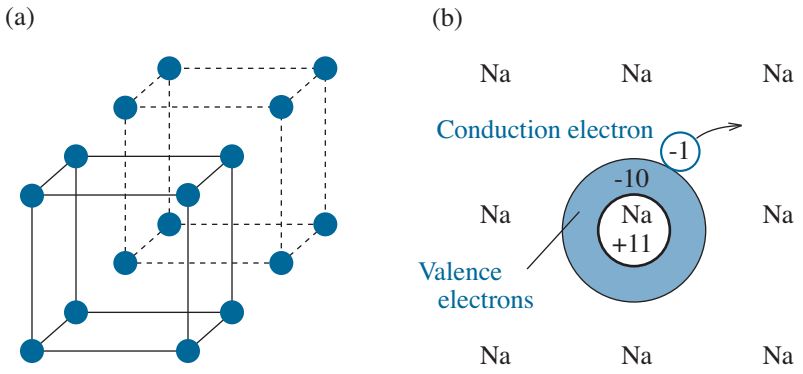


Figure 2.2: (a) Sodium crystal lattice; (b) simplified atomic arrangement. In the body-centered cubic lattice, each atom lies at the center of a cube defined by eight neighboring atoms.

Each lattice site contains a sodium nucleus consisting, in part, of eleven positively charged protons that are externally balanced by an equivalent number of negatively charged electrons to maintain net charge neutrality. Sodium is positioned in Column I of the periodic table, and we recall, from our study of chemistry, that the electronic configurations of the elements in Column VIII have greatest stability. Thus, for each set of eleven electrons, ten **valence electrons** are tightly bound near a particular sodium nucleus (as with the inert element neon), but one **conduction electron** is easily disassociated. In a useful physical model, this extra electron joins an ensemble of similar electrons to form an "electron gas" that permeates a sea of immobile and positively ionized sodium atoms. One can show that the extra or non-localized electrons can be treated as freely moving particles by specifying an effective electron mass, which is generally different from the electron mass in an isolated environment.

If we apply an electric field with magnitude $\mathcal{E}$ in the $+x$ direction, the free electrons respond by accelerating in the $-x$ direction. So for any particular electron, a prospective equation of motion takes the form

$$m_e^* \frac{dv_x}{dt} = -q\,\mathcal{E}\,, \tag{2.1}$$

where v_x is the x-directed velocity, q is the magnitude of the electron charge (1.6×10^{-19} coulomb), and m_e^* is the effective electron mass.

The solution to Eq. 2.1— $v_x = -q\,\mathcal{E}t/m_e^\star$ + an initial x-directed velocity —is hardly worth mentioning because it fails to take account for electron interactions with positively ionized sodium atoms and thermally induced lattice vibrations. As indicated in Fig. 2.3, these random interactions tend to scatter electrons into directions other than that of the electric field. Equation 2.1 only applies at intermediate times between scattering events.

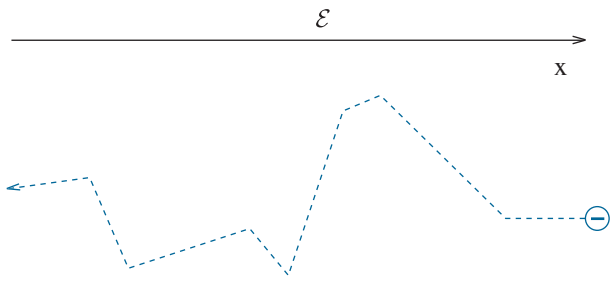


Figure 2.3: Typical electron trajectory in a solid with applied electric field. The scattering is purely random.

It is a hopeless assignment to monitor individual electron trajectories, since details of scattering interactions are unpredictable and the electron density is very large. At best, we define an x-directed **drift velocity** ($\bar{v}_x$) that averages the effects of scattering events. The averaged electron velocity is proportional to the electric field. Specifically,

$$\bar{v}_x = -\mu_e\,\mathcal{E}\,, \tag{2.2}$$

where μ_e is a constant called the electron **mobility**. In the case of sodium, $\mu_e \approx 50$ cm^2/V-s.

It is also hopeless to attempt a direct observation of the drift velocity. Instead, we observe the total electron current density J in response to the applied electric field. By definition, J indicates the charge that crosses a plane per unit area per unit time, and the plane is generally perpendicular to the direction of the electric field. If n is the electron concentration,

$$J = -qn\,\bar{v}_x\,. \tag{2.3}$$

Then with the help of Eq. 2.2, we eliminate $\bar{v}_x$ to obtain

$$J = q\mu_e n\,\mathcal{E} = \sigma\mathcal{E}\,. \tag{2.4}$$

In this expression, σ is the **conductivity** of the metal. For sodium, $n = 2.5 \times 10^{22}$ cm^{-3}, and $\sigma \approx 2 \times 10^5$ (Ω-cm)$^{-1}$.

Example 2.1

Estimate the resistance of a bar of sodium metal with cross-sectional area $A = 0.01$ cm^2 and length $l = 1$ cm.

Solution

Figure 2.4 shows the bar geometry. We apply voltage v, and we measure the current $i = JA$. Then from Eq. 2.4, and assuming $\mathcal{E} = v/l$,

$$i = \frac{\sigma A}{l}\, v = \frac{1}{R}\, v\,. \tag{2.5}$$

This is Ohm's law in terms of resistance R. With $\sigma \approx 2 \times 10^5$ $(\Omega\text{-cm})^{-1}$, the bar resistance is about 5×10^{-4} Ω.

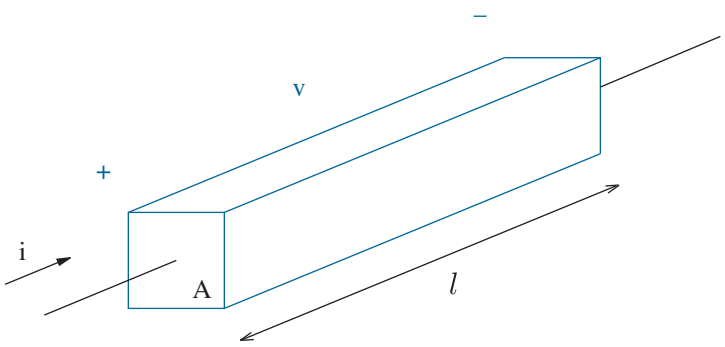


Figure 2.4: Bar geometry for Example 2.1.

Unlike sodium, metals such as aluminum, copper, and gold are useful for solid-state devices and integrated circuits. Their electronic properties are more difficult to model, but they share an inherent high conductivity as a consequence of large free-electron concentrations.

Exercise 2.1 A copper interconnect line within a printed circuit board has 10-cm length, 0.1-cm width, and 20-μm thickness. Copper **resistivity** (reciprocal conductivity) is 1.7×10^{-6} Ω-cm. Determine the resistance of the interconnect line.

Ans: $R = 0.085$ Ω

Semiconductors

We now examine the semiconductor as a material with electronic behavior more versatile than that of a metal. In accordance with the bias of Fig. 2.1, we are primarily concerned with silicon, an **elemental semiconductor**. Other semiconductors are considered in relation to the benchmark of silicon device technology.

Intrinsic Silicon

Figure 2.5a shows the characteristic diamond structure of crystalline silicon in which individual atoms lie at the centers of tetrahedra whose four corners are nearest silicon neighbors. As a Column-IV element in the periodic table, a silicon atom features four outermost electrons in addition to ten tightly bound valence electrons. In contrast to "electron gas" conditions in a metal, each outermost electron is shared with one of the four nearest neighbors to form a set of **covalent bonds** consisting of electron pairs. The sharing is a means of tricking every silicon atom into believing that it has, on average, a stable eighteen-electron configuration of Column-VIII argon. Figure 2.5b shows a simplified two-dimensional projection of silicon covalent bonding arrangements in effect at absolute zero temperature ($T = 0$ K).

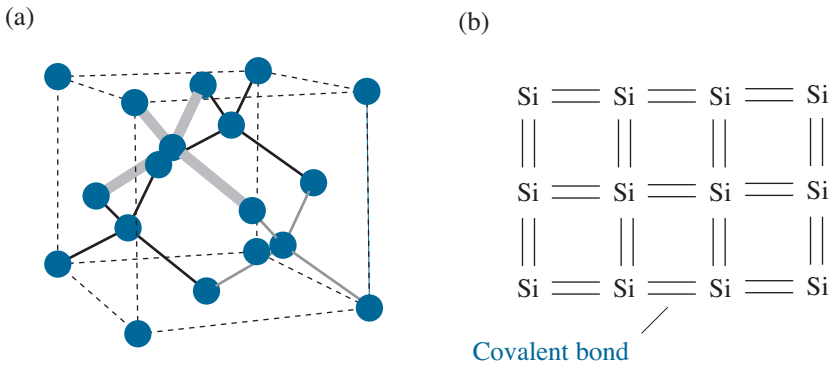


Figure 2.5: (a) Silicon crystal structure; (b) covalent bonding arrangements in effect at absolute zero temperature.

The electrons of Fig. 2.5b are essentially valence electrons even though they are localized to specific bonds rather than a specific silicon nucleus. With no conduction electrons, **intrinsic** (pure) silicon has zero conductivity at absolute zero temperature.

With increasing temperature, however, there is sufficient thermal energy to break some bonds, thereby creating free conduction electrons. As shown in Fig. 2.6 for the case of a single broken bond, the missing bond participant can also move freely through a series of bond rearrangements. This missing electron is called a **hole**.

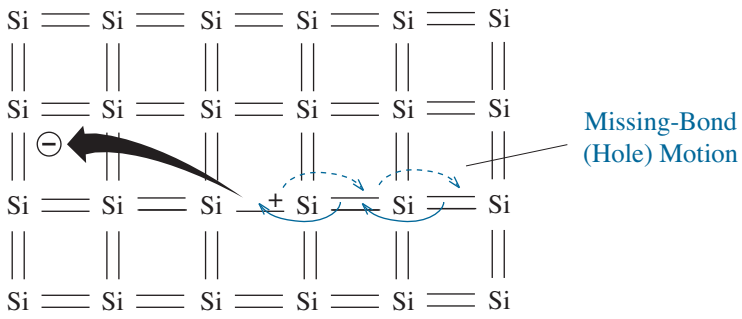


Figure 2.6: Silicon covalent bonding for $T > 0$ K. A broken bond creates an electron and a hole that are free to move independently.

We might better appreciate the electron/hole concept by considering a crude analogy featuring drops of water. Visualize a droplet of water that is falling through air over some short distance. As an observer, you would probably say "a droplet of water is falling." If asked to further elaborate, you might comment on the mass of the water droplet, the downward force of gravity, etc. Now visualize a glass tube that is closed at both ends and completely filled with water except for a single missing droplet, a bubble, at the top end of the tube. Suddenly someone turns the tube upside-down. How would you describe your subsequent observations? It is doubtful that you would say "a large ensemble of water droplets is collectively falling" even though this is the correct description. Instead, you would probably say "a water bubble is rising" since the bubble is the focus of your attention. To account for its rise in the presence of a downward gravitational force, you could describe the bubble in terms of a negative mass.

An electron (droplet) is a real particle, but a hole (bubble) is fictitious. We expect that a hole under the influence of an electric field will move in the opposite direction from that of an electron. Thus, we treat the hole as a free particle having *positive* charge of the same magnitude as the electron charge. The hole effective mass is also positive, but $m_h^* \neq m_e^*$.

The simple theory that led to an expression for conductivity in a metal is easily modified to include hole charge transport. The new conductivity in a semiconductor is

$$\sigma = q\mu_e n + q\mu_h p \,, \tag{2.6}$$

where μ_h is the hole mobility and p is the hole concentration. For the case of intrinsic silicon at room temperature,

$$n = p = n_i \sim 10^{10} \text{ cm}^{-3} \,. \tag{2.7}$$

Nevertheless, intrinsic silicon is a rather poor conductor, since the value of n_i is about twelve orders of magnitude smaller than typical free-electron concentrations in metals.

Fortunately, this is not the whole story.

Extrinsic Silicon

Silicon is made **extrinsic** by introducing certain types of impurities during crystal growth or subsequent device fabrication—these practical details are discussed in the Interlude. Our present task is to consider two cases of impurity **doping** and the modified electrical conductivity that results.

n-type Silicon

Case I introduces an element from Group V of the periodic table such as phosphorus (P). Figure 2.7 shows the bonding arrangements that apply when a phosphorus atom attempts to imitate a silicon atom by taking a substitutional lattice position. Of the five outermost phosphorus electrons, four are shared as covalent-bond constituents with nearby silicon atoms. However, a fifth easily disassociates to become a free conduction electron, a process that leaves behind an immobile and positively ionized impurity. Thus, phosphorus is called a **donor**, and the silicon is said to be **n-type** since negatively charged electrons are the predominant free carriers.

Except for very low temperatures, nearly all of the donors in n-type silicon are ionized and the electron concentration is

$$n = {N_d}^+ \approx N_d \,, \tag{2.8}$$

where N_d is the donor concentration. This can be varied over several orders of magnitude, with 10^{14} cm^{-3} $< N_d <$ 10^{19} cm^{-3} as a practical range of concentration values that allow for effective substitution and ionization.

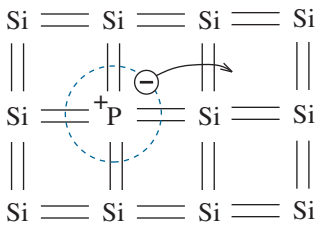


Figure 2.7: Bonding arrangements for substitutional phosphorus in silicon. Phosphorus ionization yields a free electron.

What about holes? Apart from the donor ionization process, which does not involve holes, electron/hole pairs participate in a thermally induced reversible process of the form

$$\text{unbroken bond} \rightleftharpoons \text{electron} + \text{hole} \; .$$

Here, the right and left "reactions" reflect electron/hole **generation** and **recombination**, respectively. The two reaction rates must be balanced in order to prevent a runaway condition in either direction. If we assume that the generation rate is independent of extrinsic doping, and if we assume that the recombination rate is proportional to the product of the electron and hole concentrations (as one would expect for a chemical law of mass action), then np is constant. So with $n = p = n_i$ in an intrinsic semiconductor, we obtain the general relation

$$np = {n_i}^2 \; . \tag{2.9}$$

In turn, for n-type silicon,

$$p = \frac{{n_i}^2}{n} \approx \frac{{n_i}^2}{N_d} \; . \tag{2.10}$$

Thus, the hole concentration is actually depressed well below the intrinsic concentration level, and it is usually negligible.

The preceding results suggest that the conductivity of n-type silicon is given by

$$\sigma = q\mu_e n + q\mu_h p \approx q\mu_e N_d \; . \tag{2.11}$$

The conductivity can be varied over a wide range as N_d is changed.

p-type Silicon

Case II introduces an element from Group III of the periodic table such as boron (B). Figure 2.8 shows the bonding arrangements that apply when a boron atom attempts to mimic a silicon atom as a substitutional impurity. Only three outermost boron electrons can participate in covalent bonds; however, a fourth electron is easily "accepted" from a nearby covalent bond. This is equivalent to the disassociation of a hole, a process that leaves behind an immobile and negatively ionized impurity. Thus, boron is called an **acceptor**, and the silicon is said to be **p-type** since positively charged holes are the predominant free carriers.

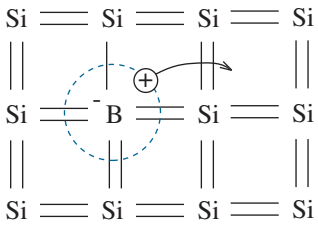


Figure 2.8: Bonding arrangements for substitutional boron in silicon. Boron ionization yields a free hole.

The equations that describe a p-type semiconductor are similar in form to those describing an n-type semiconductor. If most acceptors are ionized, the hole concentration is

$$p = {N_a}^- \approx N_a \,, \tag{2.12}$$

where N_a is the acceptor concentration. This can be varied over several orders of magnitude, with 10^{14} cm^{-3} $< N_a <$ 10^{19} cm^{-3} as a practical range of concentration values. The electron concentration is

$$n = \frac{{n_i}^2}{p} \approx \frac{{n_i}^2}{N_a} \,, \tag{2.13}$$

and it is negligible if N_a is large. Finally, the conductivity of p-type silicon is given by

$$\sigma = q\mu_e n + q\mu_h p \approx q\mu_h N_a \,. \tag{2.14}$$

The conductivity can be varied over a wide range as N_a is changed.

Example 2.2

A silicon bar with cross-sectional area $A = 0.01$ cm^{-2} and length $l = 1$ cm has a phosphorus doping concentration of 2×10^{16} cm^{-3}. The electron and hole mobilities are $\mu_e = 1000$ cm^2/V-s and $\mu_h = 450$ cm^2/V-s, respectively. Determine the electron and hole concentrations, and the bar resistance.

Solution

We assume complete phosphorus ionization so that $n = 2 \times 10^{16}$ cm^{-3}. Then subject to $np = {n_i}^2$, $p = 10^{20}$ cm^{-6} / 2×10^{16} cm^{-3} $= 5 \times 10^3$ cm^{-3}, which is negligible. From Eq. 2.14, $\sigma \approx 1.6 \times 10^{-19}$ C $\times$ 1000 cm^2/V-s $\times$ 2×10^{16} cm^{-3} $= 3.2$ $(\Omega\text{-cm})^{-1}$. Thus, in consideration of Example 2.1,

$$R = \frac{l}{\sigma A} = \frac{1 \text{ cm}}{10^{-2} \text{ cm}^2 \times 3.2 \ (\Omega - \text{cm})^{-1}} \approx 30 \, \Omega \, .$$

Example 2.3

A silicon bar has boron doping with concentration 1.99×10^{16} cm^{-3} and phosphorus **compensation doping** with concentration 2.01×10^{16} cm^{-3}. Determine the electron and hole concentrations.

Solution

We assume complete ionization and uniform distribution of both impurities. And whereas the bar is electrically neutral, the charge density ρ vanishes. Specifically,

$$\rho = p - n + {N_d}^+ - {N_a}^- = 0 \, ,$$

where ${N_d}^+$ is the concentration of fixed ionized phosphorus donors and ${N_a}^-$ is the concentration of fixed ionized boron acceptors. So with $p = {n_i}^2/n$ and some algebra,

$$n^2 - Nn - {n_i}^2 = 0 \, ,$$

where $N = {N_d}^+ - {N_a}^- = 2 \times 10^{14}$ cm^{-3} is a *net* impurity concentration. The necessarily positive solution to the quadratic equation is

$$n = \frac{N + N\sqrt{1 + 4\,{n_i}^2/N^2}}{2} \approx N + \frac{{n_i}^2}{N}$$

(subject to $\sqrt{1+x} \approx 1 + x/2$ when x is small). In turn, $n \approx 2 \times 10^{14}$ cm^{-3} and $p \approx 5 \times 10^5$ cm^{-3}. It is straightforward to show that the slab is n-type with $n \approx N$ when $N \gg n_i$ and p-type with $p \approx -N$ when $-N \gg n_i$. Thus, compensation doping tends to be impractical when N_d and N_a are large since uncertainties in either doping often exceed the desired difference.

Insulators

Materials with atomic arrangements that do not allow free electron or hole conduction are called insulators. Silicon dioxide and, to a lesser extent, silicon nitride are common insulators in semiconductor device technologies. Polymer materials can be useful as insulators between interconnect wires, especially when multiple criss-crossing layers are required.

Energy-Band Models

A more sophisticated description of electrons and holes in a semiconductor employs an energy-band model. At absolute zero temperature, electrons in an intrinsic semiconductor occupy a range of energies in a **valence band**, and a separate **conduction band** is totally empty as shown in Fig. 2.9a. Neither condition allows electronic conduction. The process of breaking a covalent bond moves an electron from the valence to the conduction band. Thus, the **bandgap** or range of forbidden energies between the bands is determined by the energy necessary to break a covalent bond. Figure 2.9b shows the conduction band partially filled with electrons and the valence band partially filled with holes (or empty of electrons) for an intrinsic semiconductor at non-zero temperature. Finally, Figs. 2.9c and 2.9d illustrate donor and acceptor impurity ionization, respectively, in the extrinsic case. The impurities provide discrete electron energy states within the bandgap, and the required ionization energies are typically small when compared to the bandgap energy. For example, the silicon bandgap is 1.12 eV at room temperature,[1] while the average thermal energy of an electron is only about 0.04 eV at room temperature. Thus, impurity ionization is generally needed to obtain significant electron concentrations in the conduction band or hole concentrations in the valence band.

The energy-band model is also convenient for describing materials other than semiconductors. For example, a material featuring a conduction band that is partially filled even at absolute zero temperature is a common metal. This reflects a small bandgap or, as shown in Figure 2.10a, a bandgap that becomes zero from overlapping valence and conduction bands. In contrast, a material featuring a conduction band that is mostly empty at moderate temperatures is an insulator with a large bandgap as shown in Fig. 2.10b. Most semiconductors have intermediate bandgap energies.

Energy-band models provide a theoretical framework for descriptions of device behavior subject to electric fields and other perturbing factors. In particular, the models promote the visualization of carrier distributions and various physical barriers that impede carrier transport. Nevertheless, a satisfactory treatment of band theory goes beyond the scope of this text, and we shall tend to avoid it in future discussion.

[1] The energy unit eV or **electron-volt** is the energy needed to move an electron through a potential difference of 1 volt.

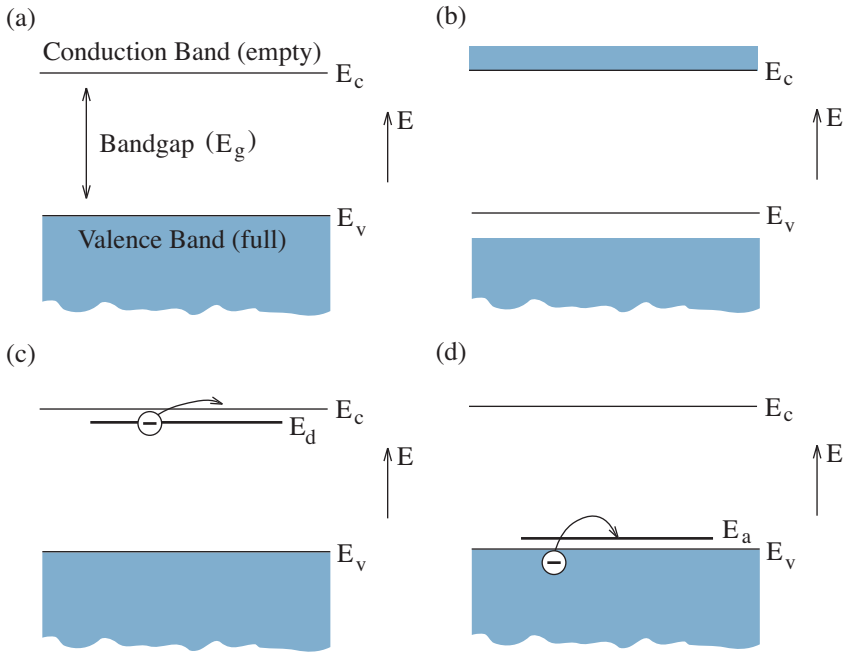


Figure 2.9: Semiconductor energy-band model: (a) $T = 0\ K$; (b) $T > 0\ K$; (c) donor ionization at energy E_d; (d) acceptor ionization at energy E_a.

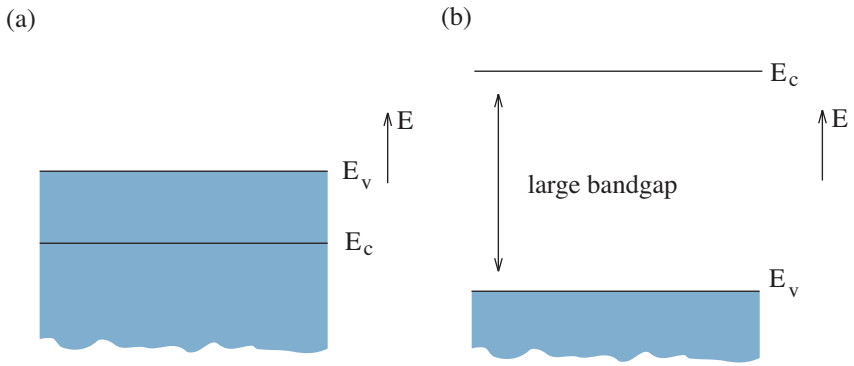


Figure 2.10: Energy-band diagrams: (a) for a metal; (b) for an insulator.

Semiconductors Other Than Silicon

Germanium and carbon (in diamond form) are elemental semiconductors. The former actually enjoyed a period of commercially successful technology development in the 1950s. However, germanium devices tend to suffer from relatively large n_i, and fabrication procedures are difficult and restrictive. We will better appreciate these problems in later chapters.

Most common **compound semiconductors** are formed by combining Group-III and -V elements (as for the case of gallium arsenide, GaAs). Impurity doping processes are complex, since the electrical behavior of extrinsic material depends on the particular element (such as Ga or As) that is substitutionally displaced. Other compound semiconductors include "II-VI" materials (such as CdS) and "IV-VI" materials (such as PbS). When compared with silicon, compound semiconductors generally exhibit much larger electron mobility values for use with very-high-speed devices and special optical properties for optoelectronic devices. Notwithstanding, they are encumbered by a relatively expensive process technology that tends to limit widespread application, especially for consumer electronics.

Some of the prospects for **alloyed semiconductors** can be appreciated with the help of Fig. 2.11, which shows bandgap energy vs. lattice constant (an indication of crystalline density) for several compound semiconductors. For example, the AlAs and GaAs bandgap energies are 2.16 eV and 1.42 eV, respectively, but the lattice constants are nearly the same. Thus, one can form $Al_xGa_{1-x}As$, an alloyed semiconductor whose bandgap depends on x, the fractional aluminum content. Optical properties vary markedly with x. The exotic electronic devices made possible through "bandgap engineering" are also beyond the scope of this text. Nevertheless, Chapter 8 will find use for $Si_{1-x}Ge_x$ to improve the high-speed performance of certain transistors.

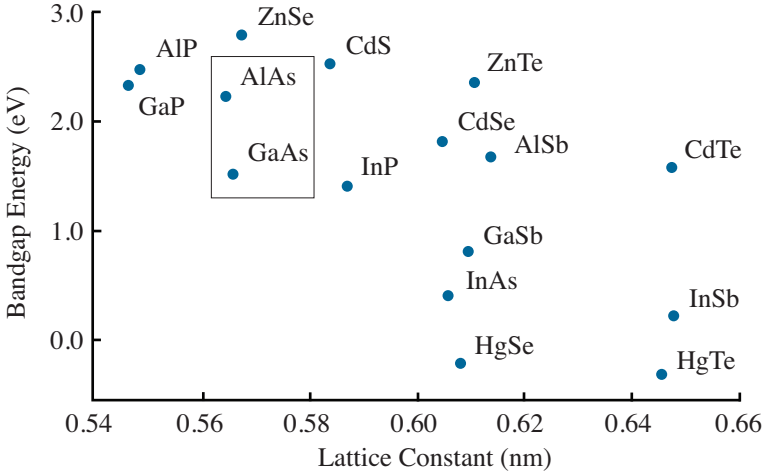


Figure 2.11: Semiconductor bandgap energy vs. lattice constant.

Concept Summary

- The conductivity (σ) relating the current density of charged carriers in a solid to an applied electric field is proportional to the atomic unit of charge (q), a mobility factor (μ), and the density of carriers.

$$\sigma = q\,\mu \times \text{density of carriers}$$

- Metals generally exhibit high conductivities as a consequence of large free-electron concentrations.
- Semiconductors conduct electrical current with two kinds of carrier: negatively charged electrons and positively charged holes.
  - Elemental semiconductors such as Si are crystals of one element.
  - Compound semiconductors such as GaAs or $Al_xGa_{1-x}As$ are crystalline arrangements of two or more elements.
  - Intrinsic (pure) semiconductors contain equal electron and hole concentrations that are typically many orders of magnitude less than the free-electron concentration in a metal.
  - Extrinsic semiconductors contain impurities that boost the concentration of one free carrier at the expense of the other.
    - n-type semiconductors have a free electron concentration (n) that balances an equal concentration of fixed positively ionized *donor* impurities ($N_d{}^+$) under isolated conditions.

$$\sigma \approx q\mu_e n = q\mu_e N_d{}^+$$

    - p-type semiconductors have a free hole concentration (p) that balances an equal concentration of fixed negatively ionized *acceptor* impurities ($N_a{}^-$) under isolated conditions.

$$\sigma \approx q\mu_h p = q\mu_h N_a{}^-$$

    - Extrinsic conductivities are much less than that for a metal, but they can be significant when impurity doping is high.
- Insulators do not contain free carriers and do not conduct current. Thus, they can provide electrical isolation in solid-state circuits.
- The conductance of a uniform solid is proportional to its conductivity and geometric factors. Its resistance is the inverse of the conductance.

2.2 Zero-Bias Diode Conditions

Armed with two distinct types of semiconductor material, we are prepared to examine a fundamental solid-state device that exhibits diode behavior. This is the pn junction diode. Section 2.4 investigates a somewhat different metal/semiconductor diode.

The pn Junction Diode

Consider a structure consisting of two uniformly doped p-type and n-type semiconductor slabs that are joined to form an abrupt **pn junction** along the plane where $x = 0$ as shown in Fig. 2.12. Fabrication procedures appear in the Interlude, when our repertoire of electronic devices is more complete. In this section, we wish to determine the electrical conditions that apply when the pn junction is subject to zero **bias voltage** ($v = 0$). Section 2.3 examines the current-voltage characteristic that results when $v \neq 0$.

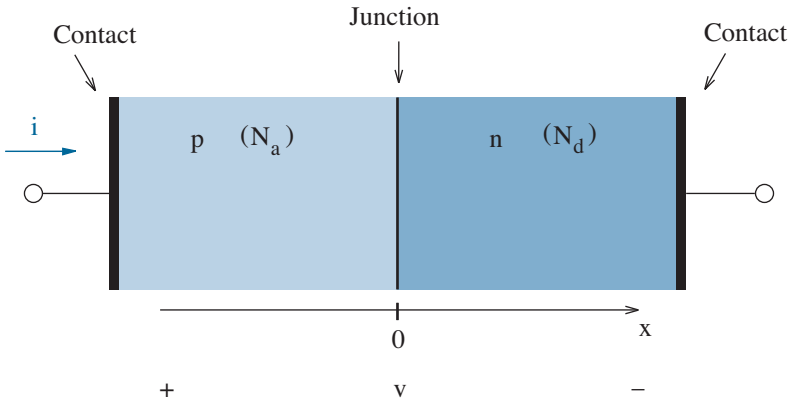


Figure 2.12: Geometry for an abrupt pn junction diode.

Electrostatic Distributions

When separate, the p-type slab contains free holes and immobile negatively ionized acceptor impurities, whereas the n-type slab contains free electrons and immobile positively ionized donor impurities. In each slab, a balance between free and immobile charge preserves electrical neutrality.

When the slabs are joined, free carriers migrate to the junction side for which they are in the minority, an indication of a carrier **diffusion** process that attempts to smooth large concentration gradients. The redistribution of free carriers causes local upsets of the former charge-balance condition. In turn, one observes **depletion regions** of uncovered immobile impurity charge in the vicinity of the pn junction.

A careful analysis reveals a gradual transition from full depletion of carriers at the junction plane to no depletion—and electrical neutrality—far away from the junction plane. However, the proper electrical behavior of a pn junction is demonstrated with nearly the same accuracy by assuming fully depleted regions that end abruptly at $-x_p$ and x_n on the p and n sides, respectively. Figure 2.13 shows the corresponding simple distribution of *net* charge density. Global charge balance is required despite the local upsets, so the shaded rectangles have equal area. Specifically,

$$qN_a x_p = qN_d x_n \, . \tag{2.15}$$

Equation 2.15 predicts that the widest of the two depletion regions extends into the semiconductor with the smallest impurity concentration. This will bear considerable interest when we discuss three-terminal semiconductor devices in Chapters 5 and 6.

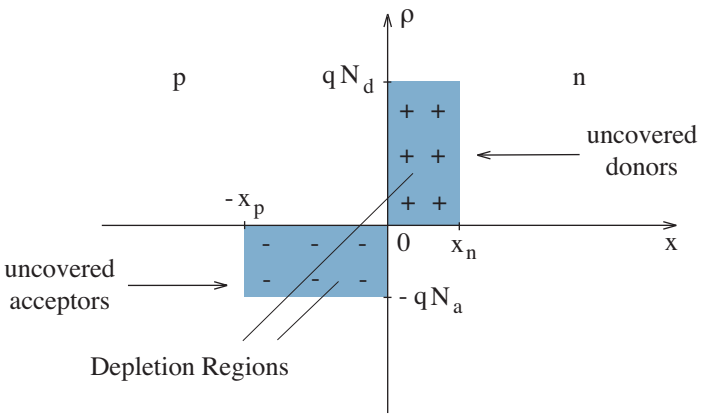


Figure 2.13: Net charge-density distribution near a pn junction.

The presence of depletion charge density suggests that there is a spatial variation of the electric potential, and we determine it with two calculations. First, we use Gauss' law to find the electric field:

$$\epsilon \frac{d\mathcal{E}}{dx} = \rho(x) \, , \tag{2.16}$$

where ϵ is the dielectric permittivity of the semiconductor. The solution to Eq. 2.16 requires $\mathcal{E} = 0$ at $-x_p$, the edge of the left-side neutral region. Meanwhile, the charge-balance condition of Eq. 2.15 ensures $\mathcal{E} = 0$ at x_n, the edge of the right-side neutral region. We pleasantly avoid some tedious algebra by applying a graphical integration procedure in which the electric field at position x is proportional to the shaded area at more negative x in Fig. 2.13, and regions of negative charge density contribute negative area. The results of this integration appear in Fig. 2.14.

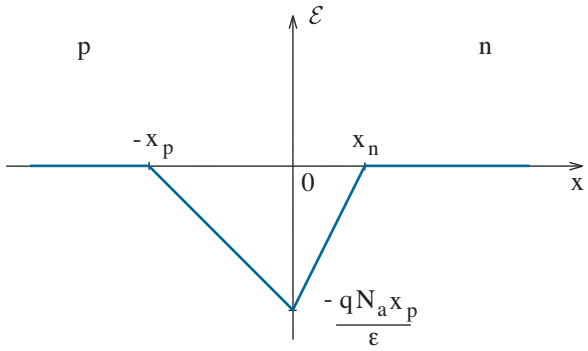


Figure 2.14: Electric field variation near a pn junction (zero bias).

For the second calculation, we recall that the electric potential ϕ is related to the electric field through the relation

$$-\frac{d\phi}{dx} = \mathcal{E}(x)\,. \tag{2.17}$$

So by choosing $\phi(-x_p) = 0$, we can apply another graphical integration procedure to obtain Fig. 2.15. Here, the potential maximum at $x = x_n$ is

$$\phi(x_n) = V_{bi} = \frac{qN_a {x_p}^2}{2\epsilon} + \frac{qN_d {x_n}^2}{2\epsilon}\,. \tag{2.18}$$

The specified assignment for $\phi(x_n)$ features V_{bi} as the **built-in potential**, a reflection of non-zero x_p and x_n when there is zero bias voltage.

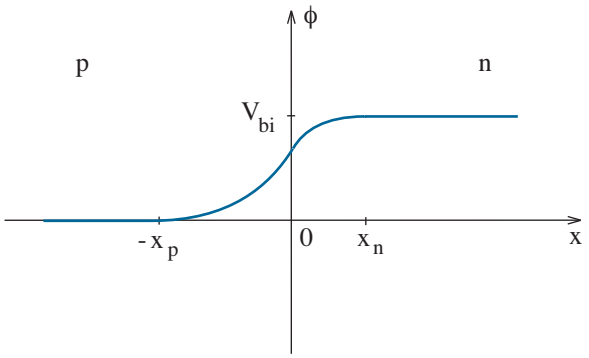


Figure 2.15: Potential variation near a pn junction (zero bias).

In physical terms, the built-in potential ensures a balance when $v = 0$ between carrier diffusion, which results from large concentration gradients, and carrier drift, which results from the built-in electric field of Fig. 2.14. In turn, the balance ensures zero open-circuit junction current.

A particular built-in potential does not imply that the diode functions as an energy source. Zero short-circuit diode current is consistent with zero open-circuit diode voltage, and this is achieved through the added presence of built-in potentials at the far-removed metal/semiconductor contacts on the p and n sides of the junction. We return to the contact problem at the beginning of Section 2.4.

It would now appear that the magnitude of the built-in potential can be determined using Eq. 2.18. Nevertheless, we lack individual expressions for x_p and x_n when $v = 0$ (or otherwise). A separate argument, which we reserve for interested readers, is needed to determine V_{bi} —then we can go back to evaluate the depletion-region widths (Example 2.4, Section 3.4). The separate argument is based on thermodynamics, and it yields

$$V_{bi} = \frac{kT}{q} \ln\left(\frac{N_a N_d}{{n_i}^2}\right) , \tag{2.19}$$

where k is Boltzmann's constant ($k/q = 8.62 \times 10^{-5}$ V per degree K) and T is the absolute temperature. For the case of an asymmetric silicon pn junction with $N_a = 10^{15}$ cm^{-3}, $N_d = 10^{19}$ cm^{-3}, and $n_i = 10^{10}$ cm^{-3}, $V_{bi} = 0.83$ V at room temperature (300 K).

Before continuing, we note that Eq. 2.19 features an important quantity called the **thermal voltage** with value

$$\frac{kT}{q} = 25.9 \text{ mV} \tag{2.20}$$

at 300 K. We will frequently encounter this special voltage in our study of electronics, and we will tend to use the unrounded value (to encourage agreement with SPICE circuit simulations). Some individuals may prefer to use $kT/q = 25$ mV in the spirit of $\pi = 3$.

Exercise 2.2 Determine the thermal voltage at (a) -55 °C; (b) 85 °C

Ans: (a) $kT/q = 18.8$ mV (b) $kT/q = 30.9$ mV

Exercise 2.3 Determine the built-in potential for a silicon pn junction with $N_a = 4 \times 10^{17}$ cm^{-3} and $N_d = 2 \times 10^{15}$ cm^{-3} at 300 K.

Ans: $V_{bi} = 0.77$ V.

The Built-In Potential

We do not want to become bogged down with too much physical detail, but a crude argument concerning the magnitude of the built-in potential will prove to be useful for later discussion.[2] First, a brief digression.

Consider a long upright cylindrical tube with unit cross-sectional area that is closed at the bottom ($z = 0$) and filled with a large but finite number of *ideal* gas molecules as shown in Fig. 2.16. The gas molecules interact with the walls of the cylinder, which are held at constant temperature T, but not with each other. The gas molecules are also subject to a constant downward gravitational force mg, where m is the molecular mass and g is the acceleration due to gravity. It is hopeless to attempt to monitor each gas molecule separately, so we must be content to observe macroscopic system properties such as pressure P and molecular density $\hat{n}$. Both are position dependent, but uniform temperature and system isolation suggest that both are invariant with time and past history. These invariant conditions reflect a system at **equilibrium**.

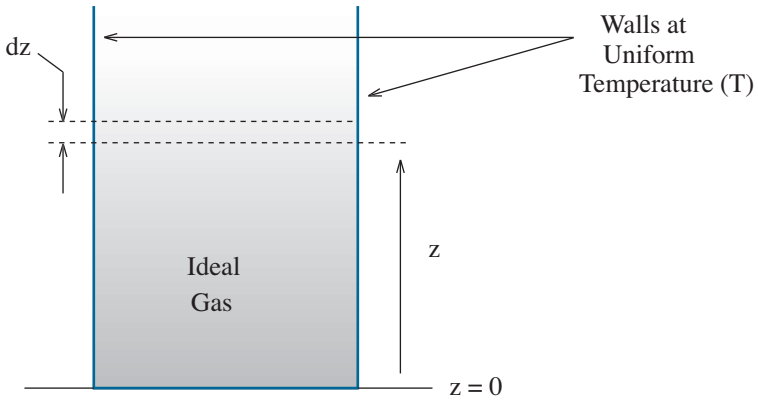


Figure 2.16: Cylindrical tube filled with gas molecules that are subject to a downward gravitational force.

In an ideal gas, pressure, volume, molecular density, and temperature are related. The corresponding gas law can be cast in several ways, and we choose to write it in the form

$$P = \hat{n}\, kT\,, \tag{2.21}$$

[2]We follow a similar argument presented in R. B. Adler, A. C. Smith, and R. L. Longini, *Introduction to Semiconductor Physics*, Semiconductor Electronics Education Committee, Vol. 1, (Wiley, New York, 1964), pp. 104-111.

where k is Boltzmann's constant (now epressed as 8.62×10^{-5} eV/K). If T is constant, then the differential pressure is given by

$$dP = d\hat{n}\, kT\,. \tag{2.22}$$

However, as we move from position z to $z + dz$, the pressure change is the weight per unit area of a thin slice of gas with thickness dz. Thus,

$$dP = -mg\,\hat{n}\,dz\,. \tag{2.23}$$

We determine $\hat{n}(z)$ by eliminating dP from Eqs. 2.22 and 2.23 and then integrating to obtain

$$\hat{n} = \hat{n}_o\, e^{-mgz/kT}\,. \tag{2.24}$$

But $-mgz$ is the potential energy U of a gas molecule in relation to $z = 0$. Equation 2.24 more generally implies

$$\hat{n} = \hat{n}_o\, e^{-U(z)/kT}\,. \tag{2.25}$$

The molecular density of an ideal gas at equilibrium decreases exponentially as potential energy increases. This is a **Boltzmann** density distribution.

Now return to the pn junction diode. Having witnessed the utility of an "electron gas" model in Section 2.1, it seems reasonable to assume that hole and electron concentrations in a diode "container" exhibit an exponential dependence on potential energy provided the pn junction is at equilibrium with $v = 0$ and no hole or electron currents. This is approximately valid when carrier concentrations are dilute (or less than about 10^{19} cm^{-3} for the case of silicon). Large concentrations obey different statistics that incorporate the rules of quantum mechanics, and we defer this more accurate treatment to advanced texts.

In what follows, it is convenient to specify an electric potential ψ that is adjusted to make $\psi = 0$ at the pn junction plane where $n = p = n_i$. (This should not be confused with ϕ, the electric potential of Eq. 2.17 that had a different point of reference.) The potential energy of a hole is $q\psi$, so the hole concentration is given by

$$p = n_i\, e^{-q\psi/kT}\,. \tag{2.26}$$

In contrast, the potential energy of an electron is $-q\psi$, so the electron concentration is given by

$$n = n_i\, e^{\,q\psi/kT}\,. \tag{2.27}$$

Thus, the spatial dependence of p or n is known if $\psi(x)$ can be determined. Note that Eqs. 2.26 and 2.27 maintain the $pn = {n_i}^2$ equilibrium condition.

You might be alarmed to observe that Eqs. 2.26 and 2.27 also predict non-zero hole and electron concentrations in the vicinity of the pn junction, an apparent contradiction of the full-depletion *approximation* that led to Fig. 2.13. A rigorous treatment of the pn junction would have incorporated the potential dependence of the carrier concentrations in the right-hand side of Eq. 2.16, but the extra mathematical effort does not justify the meager improvement in precision or the loss of physical insight.

We calculate the built-in potential by applying Eqs. 2.26 and 2.27 on the p and n sides of the junction, respectively. At $x = -x_p$, where $p \approx N_a$,

$$\psi(-x_p) = -\frac{kT}{q}\ \ln\left(\frac{p}{n_i}\right) \approx -\frac{kT}{q}\ \ln\left(\frac{N_a}{n_i}\right). \tag{2.28}$$

Similarly, at $x = x_n$, where $n \approx N_d$,

$$\psi(x_n) = \frac{kT}{q}\ \ln\left(\frac{n}{n_i}\right) \approx \frac{kT}{q}\ \ln\left(\frac{N_d}{n_i}\right). \tag{2.29}$$

So in agreement with Eq. 2.19, the built-in potential is

$$V_{bi} = \psi(x_n) - \psi(-x_p) = \frac{kT}{q}\ \ln\left(\frac{N_a N_d}{{n_i}^2}\right). \tag{2.30}$$

We assume constant V_{bi} even for small departures from equilibrium ($v \neq 0$).

Example 2.4

A so-called n^+p silicon diode has a large doping concentration (10^{19} cm^{-3}) in the n region and a relatively light doping concentration in the p region. Design it so that the zero-bias p-side depletion width is $|x_p| = 2\ \mu$m.

Solution

We multiply both sides of Eq. 2.18 by $2\epsilon/q{x_p}^2$ to find

$$\frac{2\epsilon V_{bi}}{q{x_p}^2} = N_a + N_d\left(\frac{x_n}{x_p}\right)^2.$$

Then with the help of Eq. 2.15, which relates depletion widths and dopings,

$$\frac{2\epsilon V_{bi}}{q{x_p}^2} = N_a\left(1 + \frac{N_a}{N_d}\right) \approx N_a$$

(since $N_d \gg N_a$). What remains would appear easy, but V_{bi} depends on N_a. So we guess that N_a will be about 10^{15} cm^{-3}. Then with $N_d = 10^{19}$ cm^{-3}, we use Eq. 2.19 (or Eq. 2.30) to obtain $V_{bi} = 0.83$ V. For silicon, we have $\epsilon = 11.8 \times 8.854 \times 10^{-14}$ C/V-cm. In turn, $N_a = 3.3 \times 10^{14}$ cm^{-3}, a result that is close to our initial guess. Nevertheless, we are able to revise V_{bi} to 0.81 V. The second-round N_a design value is 3.2×10^{14} cm^{-3} —good enough.

Concept Summary

Subject to zero applied bias voltage,

- Electrons on the n side of a pn junction tend to diffuse to the p side, leaving behind fixed positively ionized donor impurities.
- Holes on the p side of a pn junction tend to diffuse to the n side, leaving behind fixed negatively ionized acceptor impurities.
- The impurity uncovering thus described constitutes the formation of depletion regions in the vicinity of the pn junction.
  - The depletion regions are void of free carriers.
  - The depletion is effectively abrupt, extending completely to x_n on the n side and $-x_p$ on the p side.
  - The widths of the depletion regions satisfy

$$N_a x_p = N_d x_n$$

    so that the extent of free-charge depletion is greatest on the more lightly doped side of the pn junction.
- The fixed charge in the depletion regions establishes an electric field that opposes the tendency for free carriers to diffuse.
  - The electric field is directed from the n region to the p region.
  - The electric field is maximum at the pn junction plane.
  - The electric field vanishes at x_n and $-x_p$.
- The voltage that is consistent with the opposing electric field is called the built-in potential and is given by

$$V_{bi} = \frac{kT}{q} \ln\left(\frac{N_a N_d}{{n_i}^2}\right) ,$$

  where kT/q is the thermal voltage (25.9 mV at room temperature) and n_i is the carrier concentration for an intrinsic semiconductor. For the case of a silicon diode, V_{bi} is typically about 0.8 V.
- Built-in potentials are created at the metal/semiconductor contacts to the p and n sides of the junction, and they are sufficient to offset V_{bi} so that the short-circuit junction current is zero.

2.3 A Diode Current-Voltage Characteristic

We now apply bias voltage v to our diode, thereby changing $\phi(x_n)$ to $V_{bi}-v$ in Eq. 2.18 and modifying the height of the potential barrier in Fig. 2.15.

Shockley Relation

To reach the opposing n side, holes on the p side of the pn junction must climb a potential hill with height $V_{bi}-v$. And to reach the p side, mole-like electrons must burrow downward by $V_{bi}-v$. Thus, carrier transport across the junction is made easier under **forward bias** when $v>0$ (Fig. 2.17a), and it is more difficult under **reverse bias** when $v<0$ (Fig. 2.17b).

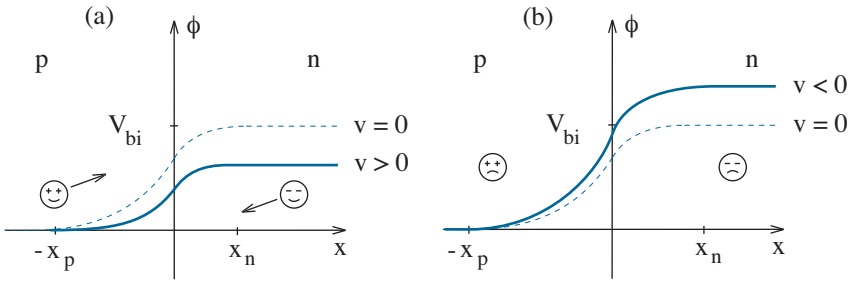


Figure 2.17: Potential "hills" in the vicinity of a semiconductor pn junction: (a) forward bias; (b) reverse bias. Holes climb hills; electrons burrow under.

In a crude sense, the change in the forward current across the junction—holes to the n-side, electrons to the p side—is proportional to the change in the bias voltage and the value of the existing current, an indication of the degree of congestion on the transjunction "freeway." Specifically,

$$\Delta i = (\Delta v/v_t)\, i\,, \tag{2.31}$$

where v_t is constant. The solution to this differential equation implies that the forward current increases *exponentially* with increasing diode voltage. Nevertheless, the net current across the junction must be zero when $v=0$. The elementary current-voltage characteristic that captures the functional behavior of the first condition while satisfying the second has the form

$$i = I_s\left(e^{\,qv/kT}-1\right)\,, \tag{2.32}$$

where kT/q is the previously encountered thermal voltage. When the diode is forward biased, the first parenthetical term provides exponential growth, with the voltage scaled in kT/q units. But when the diode is reverse biased, this term is quickly made negligible, and the current approaches $-I_s$ for $v \ll -kT/q$. Thus, parameter I_s reflects a **reverse saturation current**. Typically, I_s is of the order of 10^{-15} A to 10^{-10} A.

Figure 2.18 shows a plot of Eq. 2.32, the so-called **Shockley relation**. While hardly ideal, the diode-like quality is clearly evident.

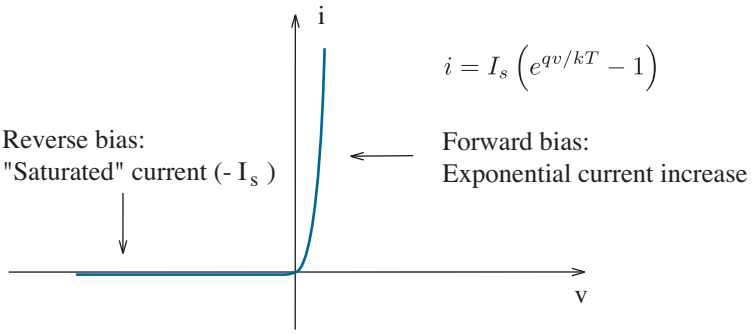


Figure 2.18: Shockley current-voltage relation for a pn junction diode.

For the unconvinced, the following more detailed argument is presented in support of the Shockley relation. Some may wish to skip this discussion, since it assumes familiarity with the concepts used to determine V_{bi}.

We presume that the Boltzmann carrier statistics used to determine the built-in junction potential at equilibrium are also applicable under small departures from equilibrium. Adding bias voltage v reduces the "adjusted" potential $\psi(x_n)$ at the top of the barrier. Specifically (see Eq. 2.29),

$$\psi(x_n) \approx \frac{kT}{q} \ln\left(\frac{N_d}{n_i}\right) - v \,. \tag{2.33}$$

Thus, with the help of Eq. 2.26,

$$p(x_n) = n_i \exp\left[-\ln\left(\frac{N_d}{n_i}\right)\right] e^{\,qv/kT} = \frac{{n_i}^2}{N_d}\, e^{\,qv/kT} \,. \tag{2.34}$$

And whereas the minority hole concentration on the n-side of the junction is ${n_i}^2/N_d$ to begin with, the *excess* hole concentration at x_n is

$$p'(x_n) = \frac{{n_i}^2}{N_d}\left(e^{\,qv/kT} - 1\right) \,. \tag{2.35}$$

Similar arguments applied to electrons—we avoid the intermediate equations—reveal that the *excess* electron concentration at $-x_p$ is

$$n'(-x_p) = \frac{{n_i}^2}{N_a}\left(e^{\,qv/kT} - 1\right) \,. \tag{2.36}$$

The carriers that give rise to Eq. 2.35 or 2.36 are said to have been *injected* into diode regions where they are in the minority. The injected charge level is greatest from the junction side with the largest impurity concentration. For example, injected hole charge dominates when $N_a \gg N_d$.

Equations 2.35 and 2.36 clearly suggest an expected voltage dependence, but we still have to relate $p'(x_n)$ and $n'(-x_p)$ to actual junction current. Consider the distribution of injected minority-carrier holes in the n-side region bounded by x_n (now designated as 0^+ to simplify the mathematics[3]) and the diode contact at W_n. Despite hole injection, the region is said to be **quasi-neutral** if $p' \ll {N_d}^+$ throughout. And with almost no electric field, the region supports hole transport by means of diffusion as opposed to drift. The diffusion process follows Fick's law, which states that a diffusion current density is proportional to the local gradient of the carrier concentration. Specifically, with respect to excess holes,

$$J_h = -qD_h \frac{dp'}{dx} , \tag{2.37}$$

where D_h is the hole **diffusion coefficient**. Two limiting p' profiles result: Figure 2.19a shows an example of the steady-state excess hole distribution in a "short" diode featuring negligible recombination except at the contact. The current density is constant over the quasi-neutral region, so the excess hole concentration falls linearly from its maximum at 0^+ to zero at W_n. Figure 2.19b shows an example of the steady-state excess hole distribution in a "long" diode featuring sufficient recombination over the course of the quasi-neutral region so that nearly all of the excess holes are eliminated before they have a chance to reach the contact. At any point, the gradient of the hole diffusion current density is proportional to recombination rate, which, in turn, is proportional to the local concentration of excess holes. The result is an exponential decrease for $p'(x)$ (see Problem 2.33).

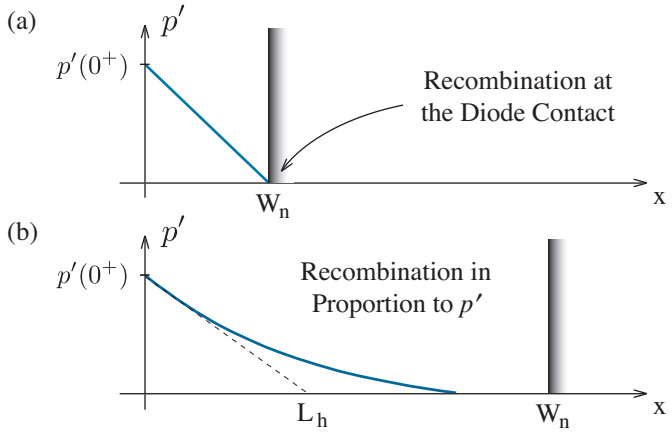


Figure 2.19: Steady-state hole distributions in the "neutral" n-side region: (a) short diode; (b) long diode. Parameter L_h is the hole diffusion length.

[3]Letting $x_n \approx 0^+$, a position slightly greater than zero, is often justified on the much larger geometric scale that features a diode contact at W_n.

The excess hole distribution consistent with Fig. 2.19a is given by

$$p'(x) = p'(0^+)\left[\frac{W-x}{W}\right] . \tag{2.38}$$

And for Fig. 2.19b,

$$p'(x) = p'(0^+)\, e^{-x/L_h} , \tag{2.39}$$

where L_h is the hole **diffusion length**, an indication of the average distance a hole can diffuse before succumbing to recombination with an electron. In either case, we apply Eq. 2.37 at $x = 0^+$ to obtain

$$J_h(0^+) = -qD_h \frac{dp'}{dx}\bigg|_{0^+} = \frac{qD_h\, p'(0^+)}{X_n} , \tag{2.40}$$

where $X_n = W_n$ (Fig. 2.19a), $X_n = L_h$ (Fig. 2.19b), or, more commonly, X_n is an intermediate characteristic length. Arguments in statistical physics that we will not pursue show that the hole diffusion coefficient is related to the hole mobility through the **Einstein relation**:

$$D_h = \mu_h \left(\frac{kT}{q}\right) . \tag{2.41}$$

Thus, in consideration of Eq. 2.35,

$$J_h(0^+) = \frac{kT\mu_h}{X_n}\frac{{n_i}^2}{N_d}\left(e^{\,qv/kT} - 1\right) . \tag{2.42}$$

Similar physical processes and governing equations apply to the injection of excess electrons into the quasi-neutral p-side region bounded by $-x_n$ (now 0^-) and the diode contact at $-W_p$. In turn, one eventually obtains

$$J_e(0^-) = \frac{kT\mu_e}{X_p}\frac{{n_i}^2}{N_a}\left(e^{\,qv/kT} - 1\right) , \tag{2.43}$$

where X_p is intermediate to W_p and L_e, the electron diffusion length.

Unfortunately the 0^+ (x_n) and 0^- ($-x_p$) positions lie at opposite ends of the depletion regions. Yet if we assume that n-side electrons enter the depletion regions at 0^+ and then rapidly sweep across to 0^- in response to a strong electric field, there is little opportunity for recombination, and

$$J_e(0^-) \approx J_e(0^+) . \tag{2.44}$$

Then with $i = A\,[J_h(0^+) + J_e(0^+)]$,

$$i = \underbrace{kT\,{n_i}^2 A\left[\frac{\mu_h}{X_n N_d} + \frac{\mu_e}{X_p N_a}\right]}_{I_s}\left(e^{\,qv/kT} - 1\right) . \tag{2.45}$$

This is the Shockley relation with an explicit I_s.

Ideality Factor

The elementary diode theory ignores some physical processes that alter the functional form of Eq. 2.32 (as opposed to I_s). One pair of processes, which are assisted by semiconductor "trap" imperfections that capture and release carriers, concerns the hole/electron generation and recombination rates within the junction depletion regions. We dare not try to derive the associated current-voltage dependence, even for highly motivated readers, but we accept it with the form

$$i = I_{s2} \left[e^{\, qv/2kT} - 1 \right] . \tag{2.46}$$

The complete pn junction diode characteristic is now

$$i = I_{s1} \left[e^{\, qv/kT} - 1 \right] + I_{s2} \left[e^{\, qv/2kT} - 1 \right] . \tag{2.47}$$

However, in anticipation of forward bias conditions, it is usually desirable to specify an *empirical* current-voltage characteristic:

$$i = I_s \left[e^{\, qv/nkT} - 1 \right] , \tag{2.48}$$

where n is an **ideality factor** (not confused with electron concentration) that depends on the relative values of I_{s1} and I_{s2} in Eq. 2.47, and $1 \leq n \leq 2$.

Integrated diodes usually have n so close to one that it can be neglected. Not so for discrete diodes, which tend to have widely varying n values for differently manufactured parts (together with a rarity of pertinent data). So-called "low-leakage" diodes usually exhibit low-range n values, whereas "high-speed" diodes are often characterized by mid-range n values.

Exercise 2.4 A silicon pn diode features $I_s = 5 \times 10^{-13}$ A and $n = 1.3$. Determine i when: (a) $v = 0.5$ V; (b) $v = 0.7$ V.

Ans: (a) $i = 1.4$ μA (b) $i = 0.53$ mA

Exercise 2.5 A silicon pn diode features $I_s = 2 \times 10^{-14}$ A and $n = 1.1$. Determine v when: (a) $i = 0.1$ mA; (b) $i = 10$ mA.

Ans: (a) $v = 0.64$ V (b) $v = 0.77$ V

Exercise 2.6 A particular circuit application will not tolerate more than 20 mV of change in the diode forward voltage when the current is doubled. Find the restriction on n.

Ans: $n < 1.11$

Diode Parameter Measurements

As a consequence of the substantial difference between forward- and reverse-bias diode behavior, we will find that a rudimentary understanding of diode circuits does not require the knowledge of either I_s or n. Just as well, since manufacturers almost never bother with these parameters on a data sheet. However, when accurate (and especially computer) solutions are desired, we can find n by means of a simple measurement. If the applied forward bias voltage is large compared with several kT/q,

$$i \approx I_s \, e^{\,qv/nkT} \, . \tag{2.49}$$

Thus,

$$\log i = \left(\ln I_s + \frac{qV}{nkT} \right) / \ln 10 \, . \tag{2.50}$$

In turn,

$$\frac{\Delta v}{\Delta \log i} = 2.3 \, n \left(\frac{kT}{q} \right) \, . \tag{2.51}$$

Equation 2.51 shows that the diode voltage changes by 2.3 nkT/q for every order-of-magnitude (10X) change in current. Once n has been determined, I_s follows using Eq. 2.49 in conjunction with forward i–v data.

Forward-Bias Complications

When i becomes very large, the forward diode characteristic is complicated by two additional processes: The effects of **high-level injection** are first observed when the steady-state concentration of carriers entering the more lightly doped side of the junction becomes comparable to the concentration of fixed and similarly charged impurities in the relevant depletion region. This distorts the net charge-density distribution of Fig. 2.13, and

$$i \sim I_s{}' \, e^{\,qv/2kT} . \tag{2.52}$$

The current-voltage characteristic is further modified when the voltage drop across the small but finite resistance R_s of the neutral p and n regions causes the voltage contribution to barrier lowering in Fig. 2.15 to be somewhat less than the voltage applied at the diode contact terminals. Specifically,

$$v = v_{\text{applied}} - iR_s . \tag{2.53}$$

Under these conditions, diode current is "limited" by a resistive bottleneck —passage across the junction is relatively easy—and it approaches a slower linear rate of increase. Equation 2.48 is generally satisfactory at low and moderate current levels, and both high-current effects are accommodated through an effective R_s series resistance.

Reverse-Bias Behavior

While somewhat inappropriate, Eq. 2.48 is occasionally utilized to describe reverse-bias conditions (as opposed to Eq. 2.47). However, most diodes also pass reverse leakage current in response to physical mechanisms that are difficult to describe with accuracy. Thus, one commonly finds a constant reverse current that is larger than I_s and typically a few nA.

When the reverse bias voltage is made very large, the electric field within the depletion region can accelerate holes or electrons to energies that are sufficient to create a hole/electron pair through **impact ionization**. One of the carriers from the new pair can subsequently produce another hole/electron pair, and so on. Subject to appropriate runaway conditions, the pn diode suffers **avalanche breakdown** with large negative current. Large breakdown values of the order of hundreds of volts are often available. Nevertheless, Chapter 4 takes advantage of a low-value **Zener breakdown** process with unique circuit opportunities.

Reverse breakdown voltage is dependent on the impurity concentration profile used to form the pn junction, and it is particularly dependent on three-dimensional geometric effects, such as junction curvature. These were avoided in our one-dimensional analysis, but they are of crucial importance in the field of power semiconductor electronics. In general, asymmetric p^+n or n^+p junctions are favored. The breakdown voltage at $-V_{BV}$ is roughly proportional to the inverse of the impurity concentration on the lightly doped side of the junction, and the breakdown voltage increases when the impurity profile is made less abrupt. Sharp junction corners are undesirable since they promote large electric fields.

Figure 2.20 summarizes the forward- and reverse-bias current-voltage characteristics of a real pn junction diode.

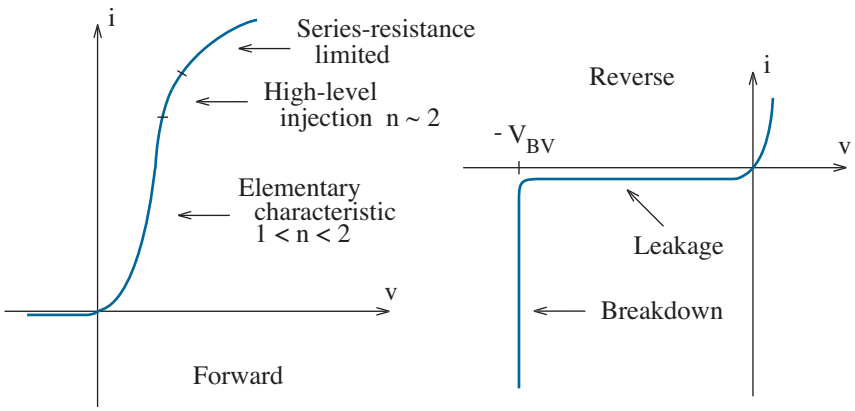


Figure 2.20: Current-voltage characteristics of a real pn junction diode.

Temperature Effects

We limit our discussion of temperature effects to forward bias conditions. Reverse-bias temperature effects are more complex, and the current levels to which they apply are small if the junction is not in breakdown.

In Eq. 2.48, there is an obvious exponential temperature dependence within the bracketed expression. Moreover, there is a hidden temperature dependence through the I_s factor, which is proportional to ${n_i}^2$. The latter quantity has the form

$$ {n_i}^2 \sim T^3 \, e^{-E_g/kT} \, , \tag{2.54} $$

where E_g is the semiconductor **bandgap energy**. Then by neglecting the T^3 term in comparison to the more dominant exponential term in Eq. 2.54, we find that the approximate forward diode current is given by

$$ i \approx I_o \exp\left(\frac{-E_g}{kT}\right) \exp\left(\frac{qv}{nkT}\right) \, , \tag{2.55} $$

where I_o is a constant (with weak temperature dependence).

We consider two cases.

• For the case of constant voltage, Eq. 2.55 predicts that temperature changes yield nearly *exponential* changes in current. This is best expressed as a relation between fractional variations:

$$ \frac{\Delta i}{i} = \frac{1}{i} \left.\frac{\partial i}{\partial T}\right|_v \Delta T = \left(\frac{nE_g/q - v}{nkT/q}\right) \frac{\Delta T}{T} \, . \tag{2.56} $$

Suppose $v = 0.7$ V. Then in silicon, for which $E_g/q = 1.12$ V, $\Delta i/i = 16.2 \, \Delta T/T$ at room temperature (300 K) if $n = 1$.

• For the case of constant current, we rearrange Eq. 2.55 to find that temperature changes yield nearly *linear* changes in voltage:

$$ \Delta v = \left.\frac{\partial v}{\partial T}\right|_i \Delta T = -\left(nE_g/q - v\right) \frac{\Delta T}{T} \, . \tag{2.57} $$

In silicon, with i such that $v = 0.7$ V, $\Delta v = -0.42 \, \Delta T/T$ (volts) if $n = 1$. Note the *negative* variation.

Exercise 2.7 A silicon pn diode with $n = 1.4$ is operated at 17 °C such that $v = 0.78$ V when $i = 100$ mA. If i is held constant, determine the temperature change that yields $v = 0.84$ V.

Ans: $\Delta T = -22$ °C

Concept Summary

Subject to non-zero applied bias voltages,

- Diodes in forward bias ($v > 0$) exhibit narrowed depletion regions in the vicinity of the pn junction, a lowered built-in potential barrier, and favorable conditions for hole and electron currents.
- Diodes in reverse bias ($v < 0$) exhibit widened depletion regions in the vicinity of the pn junction, a raised built-in potential barrier, and unfavorable conditions for hole and electron currents.
- The diode current-voltage characteristic takes the form

$$i = I_s \left[e^{qv/nkT} - 1 \right] ,$$

  where I_s is the reverse saturation current (typically $\sim 10^{-14}$ A) and n is the ideality factor ($1 \leq n \leq 2$). Good diodes have $n \approx 1$.
- Forward-bias diode current:
  - Increases exponentially with voltage scaled in units of nkT/q;
  - Features a dominant constituent carrier (hole or electron) that is determined by the more heavily doped side of the pn junction;
  - Is limited by the voltage across a series junction resistance.
- Forward-bias diode voltage:
  - Increases logarithmically with current;
  - Increases by 2.3 nkT/q for every 10X change in current.
- Reverse-bias diode current:
  - Approaches $-I_s$ in the theoretical limit;
  - Typically has "leakage" contributions of the order of 1 nA;
  - Becomes large and negative when the reverse voltage is sufficient to promote junction breakdown ($v < -V_{BV}$).
- As temperature varies:
  - Diodes under constant voltage yield exponential current change;
  - Diodes under constant current yield nearly linear and negative voltage change.

2.4 Assorted Diode Forms

Discrete Si pn diodes are found in packages varying in size, current-handling capability, and thermal accommodation. Figure 2.21 shows some examples. The p and n sides are called the **anode** and **cathode**, respectively, and the latter is typically marked. Large-current diodes are often called **rectifiers**. Although somewhat rare, Ge pn diodes offer relatively low forward voltages.

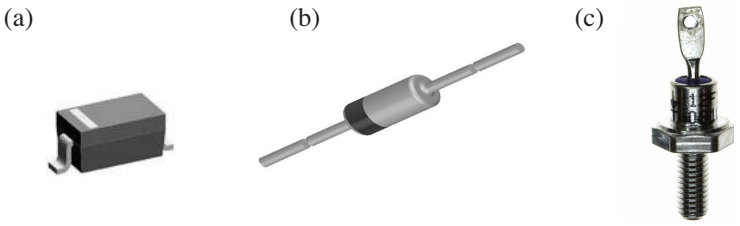


Figure 2.21: Discrete pn diodes: (a) surface mount; (b) axial lead; (c) stud. Images courtesy of Vishay Intertechnology, Inc.

Integrated pn diodes can take the p^+n or n^+p form shown in Fig. 2.22 depending upon the type of the semiconductor substrate, which serves as a common circuit node. Later chapters consider less restricted forms.

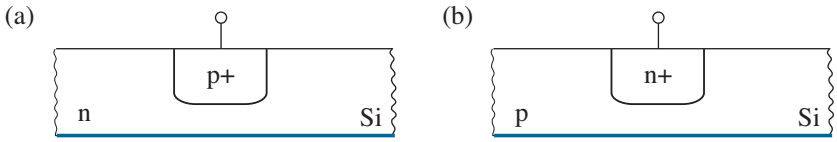


Figure 2.22: Integrated pn diodes: (a) p^+n; (b) n^+p.

Integrated diodes can provide electrical isolation for other circuit functions. For example, the integrated structures in Fig. 2.23 operate as resistors if the connecting node voltages ensure that the pn junction is reverse biased. Geometric factors and the semiconductor resistivity between the heavily doped end regions determine the resistance value (see Problem 2.47).

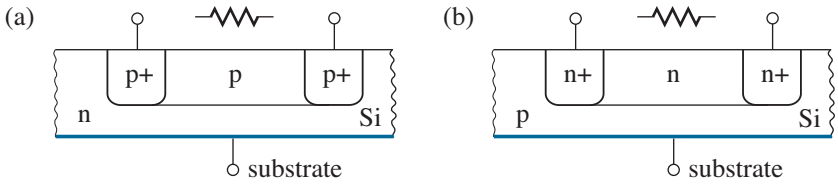


Figure 2.23: Integrated-circuit resistors featuring pn junction isolation. The substrate node is (a) the most positive or (b) the most negative.

Metal/Semiconductor Diodes and Ohmic Contacts

We have argued the necessity of built-in potentials at metal-semiconductor contacts that offset the junction V_{bi} to force zero open-circuit diode voltage. Two effects apply for metal contacts to n-type semiconductors in which the zero-bias built-in voltage is that of the metal relative to the semiconductor. Similar considerations apply for metal contacts to p-type semiconductors.

When the built-in voltage is negative, electrons in the semiconductor are repelled from the contact junction. As a result, the repulsion establishes a potential barrier that is increased or decreased with an applied voltage. The characteristic current-voltage relation is similar to Eq. 2.41:

$$i = I_s \left[e^{\,qv/nkT} - 1 \right] . \tag{2.58}$$

Thus, we have a **rectifying contact** or **Schottky diode** that controls majority-carrier (electron) flow. Shottky diodes tend to have significantly larger I_s values than their pn counterparts when they are comparably sized so that the forward voltage is generally smaller at any particular current. A tradeoff is relatively poor reverse leakage current and breakdown voltage.

When the built-in voltage is positive, electrons in the semiconductor are attracted to the contact junction, and the semiconductor is **accumulated**. Although a potential barrier now exists within the metal, the associated depletion region is extremely thin as a consequence of the metal's very large electron concentration. Thin barriers offer little resistance to carrier flow. Thus, the contact is **ohmic** with a linear current-voltage relationship.

The advantageous effects of thin depletion barriers suggests that ohmic contacts can be formed when the built-in voltage is negative by increasing the donor concentration in the semiconductor. Ohmic contacts are usually made to n^+ (or p^+) semiconductor layers. Nearby "high-low" n^+-n (p^+-p) junctions are always ohmic. Figure 2.24 illustrates these concepts.

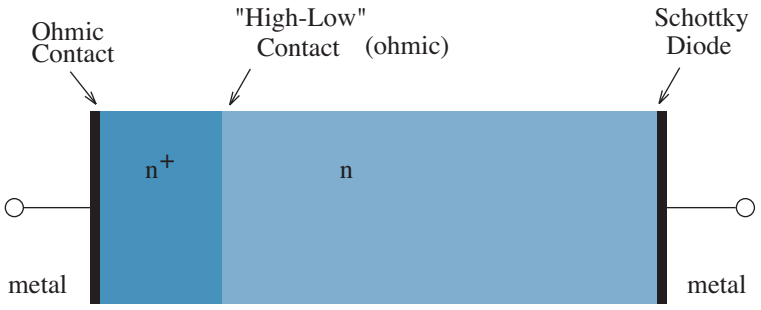


Figure 2.24: Ohmic and non-ohmic (rectifying) contacts.

Light-Emitting Diodes

Light-emitting diodes or **LEDs** are simply pn junctions that radiate light when operated under *forward bias*. But there are restrictions.

In a compound semiconductor, a photon of light is produced when a hole or electron suffers recombination following injection into "enemy" territory. The emitted radiation has wavelength

$$\lambda = \frac{hc}{E_g}\,, \tag{2.59}$$

where E_g is the semiconductor **bandgap energy**, h is Planck's constant (4.13×10^{-15} eV-s), and c is the speed of light in vacuum (3.00×10^{10} cm/s). Material systems have been designed to provide a variety of spectral choices (red, green, blue, etc.).

Carrier recombination in an elemental semiconductor is a roundabout process that does not produce radiation. Thus, there is no silicon LED.

The intensity of light that emerges from an LED is proportional to the forward current, which is often limited with the help of a series resistor. The small n_i and I_s values typical of compound semiconductors make the forward voltage large in comparison to that for a silicon pn diode subject to the same bias current.

LEDs are commonly applied in optical isolation circuits, alphanumeric displays, traffic signals, and lighting. Devices are often specially packaged as in Fig. 2.25 to promote the escape of light.

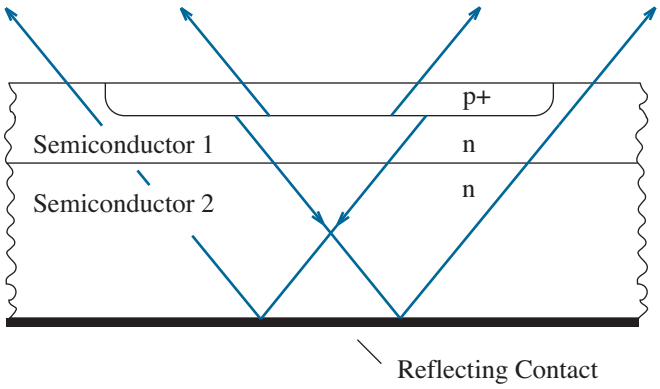


Figure 2.25: Typical LED package that facilitates the escape of light. With bandgaps such that $E_{g2} > E_{g1}$, semiconductor 2 is transparent to the radiation generated in semiconductor 1.

Laser Diodes

Laser diodes are semiconductor devices that exhibit Light Amplification through Stimulated Emission of Radiation. The light-emitting processes are similar to those for LEDs. However, the light is amplified as it travels back and forth in a resonant cavity that is bounded by two reflective surfaces. The escaping radiation is **coherent** with light waves all in phase.

Laser diodes serve as transmitters in fiber-optic communication systems (amongst other applications). Current-voltage characteristics have typical diode form. The optical output power P_o has an approximate linear relation to diode current beyond some threshold current I_{th} as shown in Fig. 2.26.

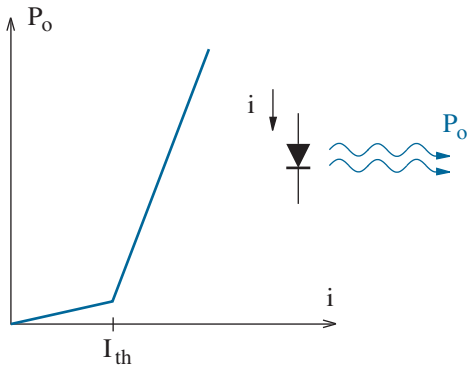


Figure 2.26: Laser-diode optical output power vs. diode current.

Photodiodes

Photodiodes are *reverse-biased* pn junctions that produce reverse current in proportion to the intensity of incident radiation as shown in Fig. 2.27. These devices typically serve as receivers in optical communication systems.

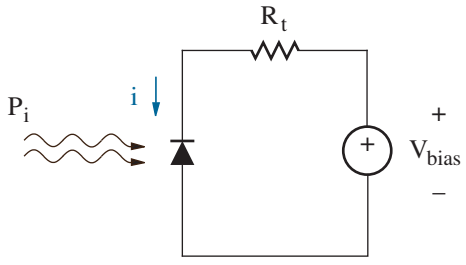


Figure 2.27: Photodiode detector. Incident radiation yields reverse current in proportion to power P_i. V_{bias} is less than the diode breakdown voltage.

The photodiode detection process begins with an absorption of a photon. If the absorption takes place within one of the diode depletion regions, the resulting electron/hole pair separate and move in opposite directions — the electron moves to the neutral n region, the hole moves to the neutral p region—through the action of a large electric field near the junction plane. Special doping promotes large-volume absorption (see Problem 2.49).

Unlike the light-emitting diode, a photodiode can be made from either a compound or an elemental semiconductor. The only restriction concerns the semiconductor bandgap energy, which must be less than the photon energy ($E = hc/\lambda$) that is to be absorbed.

Photodiodes are typically content to operate with a modest reverse bias voltage that is less than that producing junction breakdown. In contrast, an **avalanche photodiode** is generally biased just below breakdown so that optically generated holes and electrons can acquire sufficient energies to create additional hole/electron pairs by means of the impact ionization process shown in Fig. 2.28. Obtaining many collected carriers for only one incident photon provides beneficial optical gain in a low-intensity system. The downside is persistent current in response to a short optical pulse.

Of course, the most common photodiode available is the **solar cell**, which converts solar energy into electrical energy (see Problem 2.50).

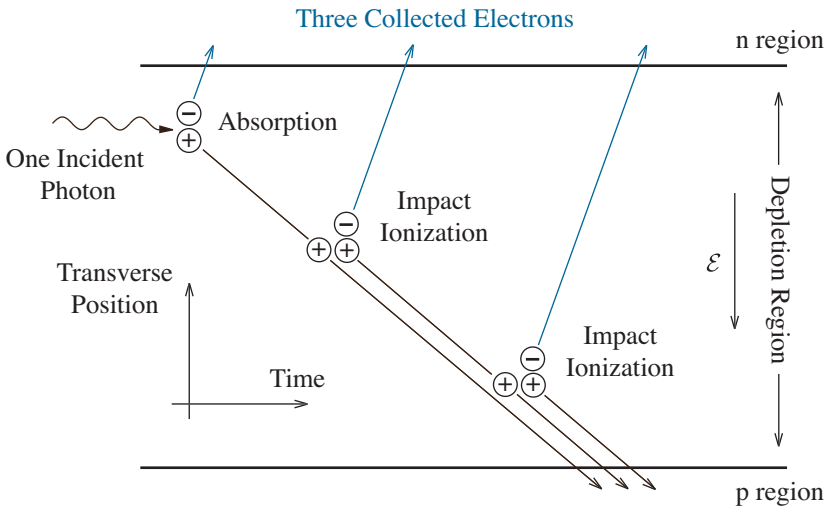


Figure 2.28: Carrier multiplication process in an avalanche photodiode. Energetic holes and electrons exhibit differing probabilities of producing additional hole/electon pairs. Holes are dominant in the process depicted.

Concept Summary

Of the many solid-state diode varieties,

- Silicon pn junction diodes are:
  - Available as discrete components in packages designed to accommodate specific ranges of current and temperature;
  - Easily incorporated into integrated circuits, either as actual diode elements or as isolated pockets for other devices;
  - Benchmark devices to which other diodes are compared.
- Germanium pn junction diodes, although rare, have relatively large I_s values for a small forward voltage.
- Metal-semiconductor or Schottky diodes have:
  - Relatively large I_s values for a small forward voltage;
  - Relatively large reverse leakage current and relatively small breakdown voltage;
  - Ohmic behavior in the limit of very high semiconductor doping.
- Light-emitting diodes (LEDs)
  - Operate under forward bias,
  - Have relatively small I_s values for a large forward voltage,
  - Produce radiation with intensity proportional to diode current,
  - Are available with specific output colors (red, green, blue, etc.),
  - Are available as laser diodes when specially fabricated.
- Photodiodes
  - Operate under reverse bias,
  - Produce current in proportion to incident radiation intensity,
  - Are particularly sensitive when biased just below breakdown,
  - Are commonly available as solar cells.

Problems

Perspective (Preceding Chapter 2)

2.1 The circuit of Fig. P2.1 has a non-linear device for which $i = \sqrt{v}$. Determine v by graphical analysis.

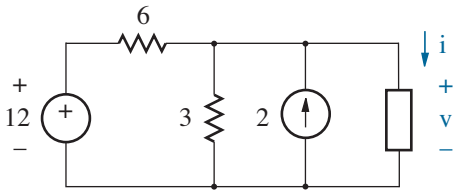


Figure P2.1

2.2 Repeat Problem 2.1, but let $i = 2\sqrt{v}$ and reduce the 2-A source to 1 A.

2.3 The circuit of Fig. P2.3 has a non-linear device for which $i = 2\,v^{3/2}$ (with i in mA). Determine v by graphical analysis.

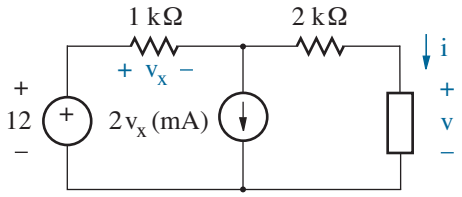


Figure P2.3

2.4 Repeat Problem 2.3 but let v_x appear across the 2-kΩ resistor (with positive node to the left).

2.5 A 10-V source with Thevenin resistance R_t connects to a non-linear device for which $i = 2(v-1)^2$ (with i in mA) subject to $v \geq 1$ V. The device has zero current otherwise. Use a graphical procedure to determine R_t such that $i = 4$ mA.

2.6 A current source I_n with 1-kΩ Norton resistance connects to a non-linear device for which $i = \sqrt{v-2}$ (with i in mA) subject to $v \geq 2$ V. The device has zero current otherwise. Use a graphical procedure to determine I_n such that $i = 8$ mA.

2.7 The circuit of Fig. P2.7 has a non-linear resistor for which

$$v = \begin{cases} ikR & i \geq 0 \\ iR/k & i < 0\,. \end{cases}$$

In this expression, k is a dimensionless constant that is subject to $0 \leq k \leq 1$. The old-fashioned electromechanical ammeter measures time-averaged current

$$< i > = \frac{1}{T}\int_0^t i(t)\,dt\,,$$

where $T = 2\pi/\omega$ is the period of the ac signal.

(a) Determine $< i >$ when $k = 1$ (the limiting case for a linear resistor).

(b) Determine $< i >$ when $k = 0$ (the limiting case for an ideal diode).

(c) Use graphical analysis to explain qualitatively the variation in $< i >$ as k is varied between 0 and 1 (the "diode" becomes less and less ideal).

(d) Determine k such that $< i > = \tilde{v}/10R$.

Note: The disparity of the results for parts a and b reflects diode function as a **detector** of ac amplitude. This is an important application (particularly for the early days of radio).

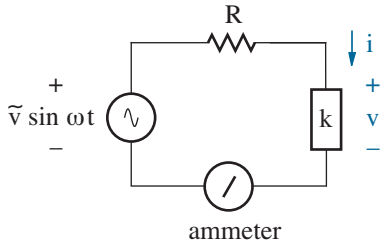


Figure P2.7

Section 2.1

2.8 In the study of chemistry, one learns that the number of molecules in one mole of any substance is given by Avagadro's number (6.02×10^{23}).

(a) Look up the density of gold. Then estimate the free electron concentration in gold, assuming one free electron per atom.

(b) The conductivity of gold is 4.1×10^5 $(\Omega\text{-cm})^{-1}$. Estimate the electron mobility in gold.

2.9 The resistivity of copper is 1.7×10^{-6} Ω-cm. Determine the resistance of 100 feet of AWG #22 copper wire with 0.645-mm diameter.

2.10 In the design of a particular integrated circuit, the aluminum interconnect lines are constrained to have 1-μm thickness. Any line is subject to certain failure (by becoming an open circuit) if it regularly carries a current density that exceeds 5×10^5 A / cm^2. Careful simulations indicate that currents as high as 15 mA can be expected throughout the circuit.

(a) Determine the minimum acceptable line width.

(b) Aluminum conductivity is 3.5×10^5 $(\Omega\text{-cm})^{-1}$. Determine the worst-case voltage drop across a line with 1-mm length.

2.11 An *intrinsic* silicon semiconductor has the bar geometry shown in Fig. 2.4. When the bar is exposed to light, holes and electrons are generated in pairs at a greater-than-equilibrium rate. (For simplicity, assume uniform generation throughout the bar.) In steady state, when generation and recombination rates balance, the *excess* carrier concentrations are $\Delta n = \Delta p = \gamma I$, where I is the light intensity and γ is a constant.

(a) Show that the fractional change in conductance G is given by

$$\frac{\Delta G}{G} = \frac{\gamma I}{n_i} \, .$$

(b) The fractional conductance change is maximized if the semiconductor slab is slightly *extrinsic*. Determine the necessary doping condition and the corresponding conductance change.

(c) The slab is used as a **photoconductor** in the light detection circuit of Fig. P2.11. Given G, what R value ensures that output voltage v has maximum sensitivity to conductance changes?

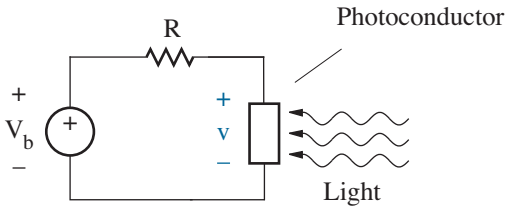


Figure P2.11

2.12 A silicon slab is doped with an arsenic concentration of 2×10^{17} cm^{-3}. Determine the conductivity type, and the electron and hole concentrations.

2.13 A silicon slab is doped with indium so that the concentration of minority carriers is 2.5×10^4 cm^{-3}. Determine the conductivity type, and the electron and hole concentrations.

Note: Amongst other reasons, indium is seldom used as a dopant in silicon due to limited solid solubility.

2.14 Find the electron and hole concentrations in a semiconductor with $N_d{}^+ \approx N_d = 2.2 \times 10^{14}$ cm^{-3} and $N_a{}^- \approx N_a = 2.0 \times 10^{14}$ cm^{-3}.

2.15 The intrinsic carrier concentration in a semiconductor varies with temperature according to the relation (Eq. 2.54)

$$n_i{}^2 \sim T^3 \, e^{-E_g/kT} \, ,$$

where E_g is the semiconductor bandgap energy and k is Boltzmann's constant (8.62×10^{-5} eV/K).

(a) Consider an extrinsic silicon slab for which the donor concentration is 10^{15} cm^{-3}. Estimate the temperature at which the hole concentration is 10% of the electron concentration. Assume $n_i = 1.6 \times 10^{10}$ cm^{-3} at 300 K and $E_g = 1.12$ eV.

(b) Repeat part a for extrinsic germanium with a donor concentration of 10^{15} cm^{-3}. Assume $n_i = 2.4 \times 10^{13}$ cm^{-3} at 300 K and $E_g = 0.66$ eV.

"There's nothing wrong with germanium. It's just that room temperature is too high." —C. Mueller

2.16 A silicon slab has a phosphorus doping concentration of 8×10^{16} cm^{-3}. Find the electron and hole concentrations at -55 °C and 125 °C (the limits of the "military" temperature range). (See Problem 2.13.)

2.17 A p-type silicon slab with uniform doping concentration N_a is oriented so that current i flows in the $+x$ direction as shown in Fig. P2.17. A magnetic field of magnitude B is applied in the $+y$ direction.

(a) Show that the magnetic field induces a voltage of magnitude

$$v_H = \frac{iB}{qN_a w}.$$

This is a consequence of the **Hall effect**.

(b) What is the Hall voltage if the slab is n-type with doping concentration N_d?

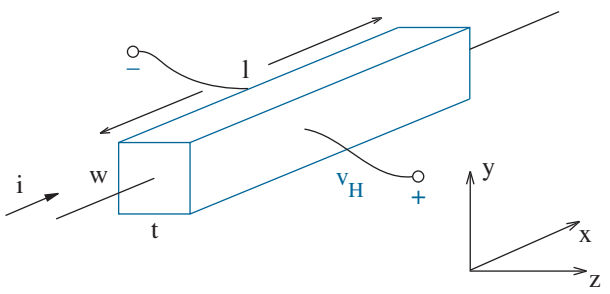


Figure P2.17

2.18 The A1301 continuous-time linear Hall-effect IC sensor features a sensitivity of 2.5 mV/Gauss. Look up the data sheet at www.allegromicro.com. While the details of the manufacturing process are proprietary, it is reasonable to assume a silicon-based technology, perhaps with $w \approx 0.5$ μm (Fig. P2.17). Discuss the design features that you would impose on the Hall device and other circuitry to achieve the cited sensitivity. Assume $l, t \geq 0.5$ μm.

2.19 Experiments indicate the following empirical expressions for electron and hole mobility in silicon as a function of impurity concentration (cm^{-3}):

$$\mu_e = 69 + \frac{1345}{1 + \left(\dfrac{N_d}{9.2 \times 10^{16}}\right)^{0.71}} \text{ cm/V s},$$

$$\mu_h = 45 + \frac{426}{1 + \left(\dfrac{N_a}{2.2 \times 10^{17}}\right)^{0.72}} \text{ cm/V s}.$$

The mobility reduction with increased doping concentration reflects a greater degree of carrier scattering from ionized impurities.

(a) Use a spreadsheet program to make a graph of resistivity vs. impurity concentration for both n- and p-type extrinsic silicon.

(b) A silicon wafer exhibits a resistivity of 5 Ω-cm. Find the electron and hole concentrations for n- and p-type material.

2.20 A silicon IC resistive layer has 1-μm thickness. Use the carrier mobility expressions of Problem 2.16 to estimate the uniform donor concentration needed to obtain an n-type sheet resistance of 500 Ω/square (a square region).

2.21 The electron and hole mobilities in silicon vary with absolute temperature like $T^{-2.4}$ and $T^{-2.2}$, respectively. A p-type silicon resistor has 1-kΩ value at room temperature (300 K). Find the resistance at -40 °C and 85 °C (the limits of the "industrial" temperature range).

Section 2.2

2.22 A symmetric *linearly graded* pn junction has the doping concentration profile

$$N_d - N_a = N_o \frac{x}{x_o}$$

for $|x| < x_o$. The doping is constant for $|x| > x_o$.

(a) Sketch the spatial distributions of net charge density, electric field, and electric potential, and label any important features. Make use of the depletion approximation, and assume that x_o is much larger than either depletion-layer width.

(b) Determine an expression for x_o such that $x_p = x_n = x_o$ when $v = 0$.

2.23 A p-type silicon bar with uniform doping concentration $N_a = 10^{15}$ cm^{-3} is bisected by a positive charge sheet at $x = 0$ with density Q coulombs/cm^2 as shown in Fig. P2.23. The fixed positive charge repels mobile holes, thereby establishing a potential barrier between the left and right sides of the slab. What value for Q yields a barrier height of 0.5 V under zero-bias conditions?

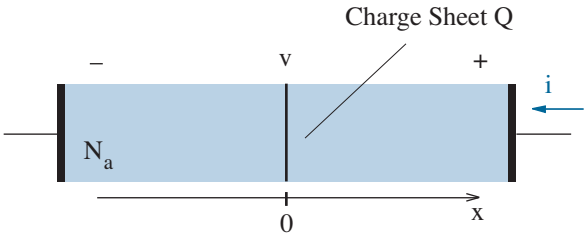


Figure P2.23

2.24 Consider a silicon pn diode with n- and p-side impurity concentrations $N_d = 2 \times 10^{15}$ cm^{-3} and $N_a = 5 \times 10^{15}$ cm^{-3}, respectively.

(a) Determine the built-in potential, the maximum zero-bias electric field, and the larger of the two depletion-layer widths.

(b) Repeat the calculations for a germanium diode ($n_i = 2.4 \times 10^{13}$ cm^{-3}, $\epsilon_R = 16.0$).

(c) Repeat the calculations for a GaAs diode ($n_i = 2.1 \times 10^{6}$ cm^{-3}, $\epsilon_R = 12.9$).

(Parameter ϵ_R is the relative dielectric constant.)

2.25 In the hot-probe measurement of Fig. P2.25, free carriers in the semiconductor diffuse along the direction of the thermal gradient established by the hot solder-gun tip.

(a) Describe the process that ensures zero open-circuit current.

(b) What is the polarity of the open-circuit voltage for a p-type semiconductor? Explain.

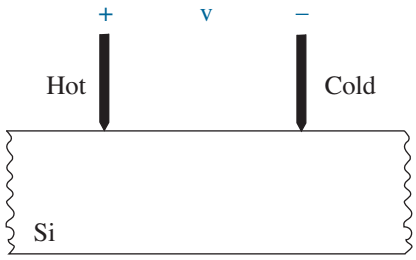


Figure P2.25

Section 2.3

2.26 Under forward-bias conditions, it is often convenient to neglect the -1 term that appears in the equation for diode current as a function of voltage. Determine the forward voltage that produces a 5-% error when this approximation is used for $n = 1$.

2.27 A pn junction diode yields the following forward i–v data:

i (mA)	v (volts)	Temperature = 290 K
0.1	0.660	
1.0	0.738	
10.	0.816	

Determine parameters I_s and n.

2.28 A pn junction diode yields the following forward i–v data:

i (mA)	v (volts)	Temperature = 320 K
0.1	0.713	
1.0	0.798	
10.	0.909	

Determine parameters I_s and n.

2.29 Specify the n-side hole and p-side electron concentrations at the depletion-region edges for a pn junction diode under large reverse bias ($v \ll -kT/q$), then explain the voltage dependence of the reverse junction current. Assume that the n- and p-side doping concentrations are N_d and N_a, respectively.

2.30 Show that the maximum electric field in the vicinity of a reverse-biased pn junction is primarily dependent upon the impurity concentration in the more lightly doped junction region.

2.31 An abrupt pn junction is doped with $N_a = N_d = N$ on the n and p sides. The junction undergoes reverse breakdown if the maximum electric field exceeds 10^6 V/cm. Determine N so that the diode breakdown voltage is 10 V.

Note: The large electric field in this **Zener breakdown** process tears apart covalent bonds, and the resulting electron/hole pairs sweep across the junction as reverse current. Junctions with p^+n^+ character are poor candidates for large breakdown voltage.

2.32 Consider a cross-sectional semiconductor slice of width Δx at position x as shown in Fig. P2.32. The hole current density that enters from the left is $J_h(x)$, and the hole current density that leaves to the right is $J_h(x + \Delta x)$.

(a) Let $p(x)$ represent the slice hole concentration. Show that

$$\frac{\partial p}{\partial t} = \frac{-1}{q}\frac{\partial J_h}{\partial x} + G - R\,,$$

where G and R are the hole generation and recombination rates per unit volume in the slice.

(b) In a region with a hole concentration gradient, holes flow by diffusion. Then $J_h = -qD_h\,\partial p/\partial x$, where D_h is the diffusion coefficient for holes. Subject to $G = R = 0$, show that the hole concentration satisfies an equation of the form

$$\frac{\partial p}{\partial t} = D_h\,\frac{\partial^2 p}{\partial x^2}\,.$$

(c) A semiconductor region ($0 \le x \le W_n$) features $p'(0+) = p_o$ (excess holes are injected from the region $x < 0$) and $p'(W_n) = 0$ (excess holes die at a contact through electron recombination). Show that for steady-state conditions in which $\partial p/\partial t = 0$,

$$J_h(0+) = \frac{qD_h p_o}{W_n}\,.$$

Assume $\mathcal{E} \approx 0$ (approximate neutrality).

Note: The conditions of part c apply to a "short" diode with a very small distance between the edge of the n-side depletion region (here located at $x = 0+$) and the n-side contact.

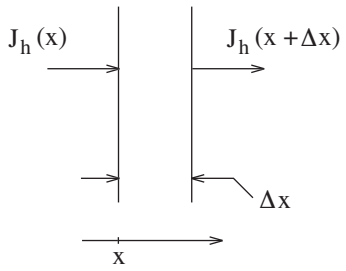


Figure P2.32

2.33 In a so-called "long" diode, holes diffusing away from the n-side depletion-layer edge are likely to recombine with electrons before reaching a contact. The recombination rate is proportional to the concentration of excess holes. Specifically, $R = p'/\tau_h$, where τ_h is the hole **lifetime**.

(a) In consideration of Problem 2.32, show that

$$\frac{\partial p}{\partial t} = D_h \frac{\partial^2 p}{\partial x^2} - \frac{p'}{\tau_h} .$$

(b) Subject to steady-state $p'(0+) = p_o$ and $p' \to 0$ as $x \to \infty$, show that

$$p'(x) = p_o \, e^{-x/L_h} ,$$

and give a relation for L_h, the **diffusion length** for holes.

(c) Finally, show that

$$J_h(0+) = \frac{qD_h p_o}{L_h} .$$

2.34 Specify the n-side doping concentration that is required for a p^+n silicon diode with $I_s = 10^{-13}$ A. Assume area $A = 10^{-4}$ cm^2, $\mu_h = 600$ cm^2/V-s, and $W_n = 1$ μm.

2.35 Determine I_s for an n^+p germanium diode with $A = 10^{-4}$ cm^2, $\mu_e = 3900$ cm^2/V-s, $W_p = 25$ μm, and $N_a = 2 \times 10^{15}$ cm^{-3}.

2.36 In the discussion leading to the Shockley relation (Eq. 2.32), it was argued that a forward bias voltage increases the minority-carrier concentration by a factor of $\exp(qv/kT)$ at depletion-region edges. Consider a pn diode with $N_a = 2 \times 10^{15}$ cm^{-3} and $N_d = 5 \times 10^{18}$ cm^{-3}. Estimate the forward voltage that corresponds to the onset of high-level injection.

2.37 A silicon pn diode with $n = 1.1$ is operated at 17 °C such that $v = 0.7$ V when $i = 10$ mA. If v is held constant, determine i when $T = 25$ °C.

2.38 A pn junction diode passes reverse current at different temperatures as in Fig. P2.38. Show that the temperature variation takes the form

$$\frac{\Delta i}{i} = \frac{\Delta T}{T_o} .$$

Then estimate T_o (using data with $v = -200$ V).

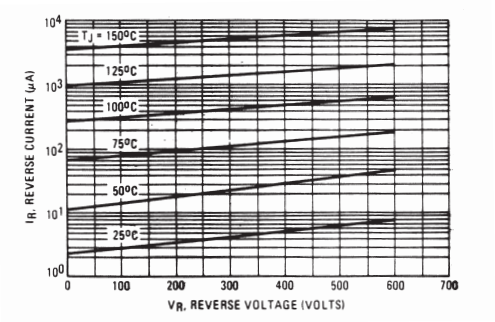


Figure P2.38

Diode Circuit Problems

2.39 Consider a diode with $I_s = 10^{-14}$ A, and $n = 1$.

(a) Find i when $v = 2$ V (at room temperature). Does this make sense?

(b) The same diode has series resistance $R_s = 0.1$ Ω. Find the current i when $v = 2$ V, and specify the division of applied voltage between the diode and the series resistance.

2.40 A pn diode with $I_s = 5 \times 10^{-14}$ A and $n = 1.2$ connects in parallel with a 1-kΩ resistor. A set of six of these combinations are connected in series across a source with all diodes in forward bias.

(a) Determine the source current for the case of a 5-V source.

(b) Determine the source voltage for the case of a 5-mA source.

2.41 One way to measure temperature electronically is to measure a diode voltage at a modest current i_x, say 0.1 mA, and then again at $10\,i_x$.

(a) Determine the absolute temperature consistent with a 60-mV voltage difference. Assume $n = 1$.

(b) Design an op-amp circuit that produces a voltage output proportional to temperature (in °C) divided by 10. Be sure to include actual current sources in your design.

2.42 Two diodes with $n = 1$ have saturation currents I_a and I_b. The diodes are connected in parallel as shown in Fig. P2.42. Determine i_1 and i_2.

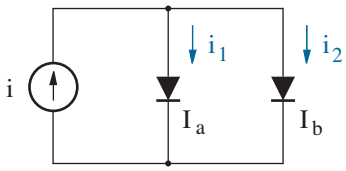


Figure P2.42

2.43 Two diodes with $n = 1$ have saturation currents I_a and I_b. The diodes are connected in series as shown in Fig. P2.43. Determine v_1 and v_2.

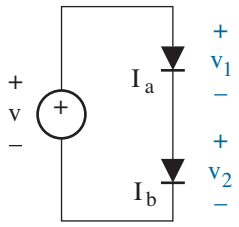


Figure P2.43

2.44 The circuit of Fig. P2.44 features an ideal op-amp and a diode for which

$$i = I_s \left[e^{\,qv/kT} - 1 \right],$$

and the circuit operates subject to the restriction that $v_{in} < 0$. Determine an *approximate* expression that relates v_{out} to v_{in}. Assume $v_{out} \gg kT/q$.

2.45 The circuit of Fig. P2.45 features an ideal op-amp and a diode for which

$$i = I_s \left[e^{\,qv/kT} - 1 \right],$$

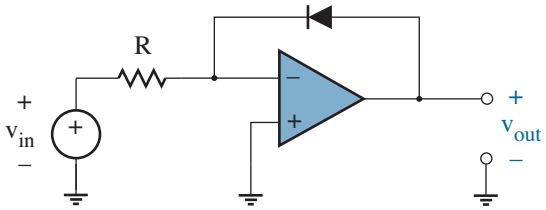


Figure P2.44

and the circuit operates subject to the restriction that $v_{in} > 0$. Determine an *approximate* expression that relates v_{out} to v_{in}. Assume $v_{in} \gg kT/q$.

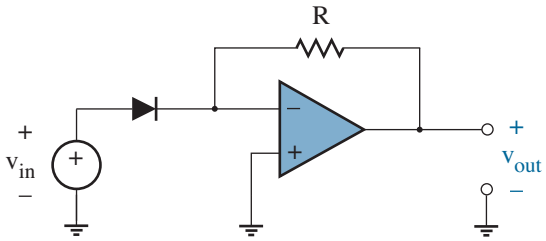


Figure P2.45

2.46 Consider the circuit of Fig. P2.46.

(a) Show that the circuit functions as an analog divider, and specify the input/output relation. Assume identical diodes for which

$$i = I_s \left[e^{\,qv/kT} - 1 \right].$$

Use reasonable approximations.

(b) Design an analog multiplier circuit. Specify any restrictions that apply to the inputs.

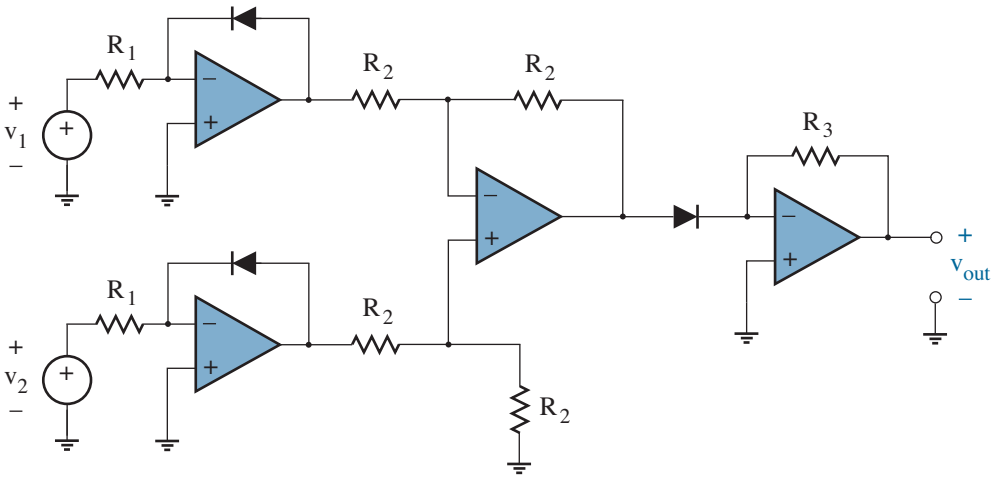


Figure P2.46

Section 2.4

2.47 Use the electron mobility data of Problem 2.19 to design a 2-kΩ integrated resistor of the form in Fig. 2.23a. Assume a 10-μm spacing between the heavily doped regions and a 0.5-μm junction depth.

2.48 Repeat Problem 2.47 subject to a 10-kΩ resistor value and p-type doping (Fig. 2.23b).

2.49 Figure P2.49 shows a p-i-n silicon diode that is used to detect optical signals. The 20-μm-long "i" region is intrinsic. For this problem, assume a lightly doped n region with $N_{di} = 10^{13}$ cm^{-3}. The p$^+$ and n end regions feature doping concentrations $N_a = 10^{18}$ cm^{-3} and $N_d = 10^{16}$ cm^{-3}, respectively.

(a) Determine the reverse bias voltage required to fully deplete the "i" region.

(b) Estimate the electric field in the "i" region when the applied reverse bias voltage is -20 V.

(c) Suppose a hole-electron pair is generated at the center of the "i" region as a consequence of incident radiation. What happens next?

(d) Let $\mu_h = 600$ cm^2/V-s and $\mu_e = 1500$ cm^2/V-s in the "i" region. Estimate the characteristic times for the events of part (c).

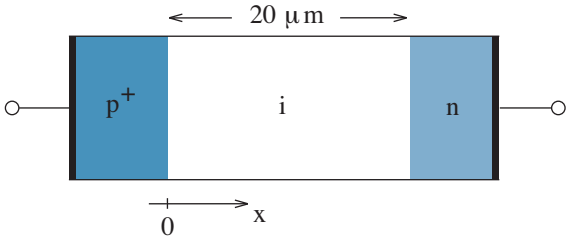


Figure P2.49

2.50 Figure P2.50a shows a pn junction that operates as a **solar cell** when illuminated with light. A simple equivalent circuit (Figure P2.50b) consists of a diode for which

$$i' = I_s \left[e^{\,qv/kT} - 1 \right],$$

and a parallel-connected current source whose value I_L is proportional to the light intensity.

(a) Determine the open-circuit voltage v_{oc} and short-circuit current i_{sc}. (Assume $v_{oc} \gg kT/q$.)

(b) If the solar cell connects to a load resistance R, show a graphical solution for voltage v. Hint: Consider the characteristic curve that describes the entire solar-cell model.

(c) Show that maximum power transfer to R implies a voltage v such that

$$\left(1 + \frac{qv}{kT}\right) e^{\,qv/kT} = \left(1 + \frac{I_L}{I_s}\right).$$

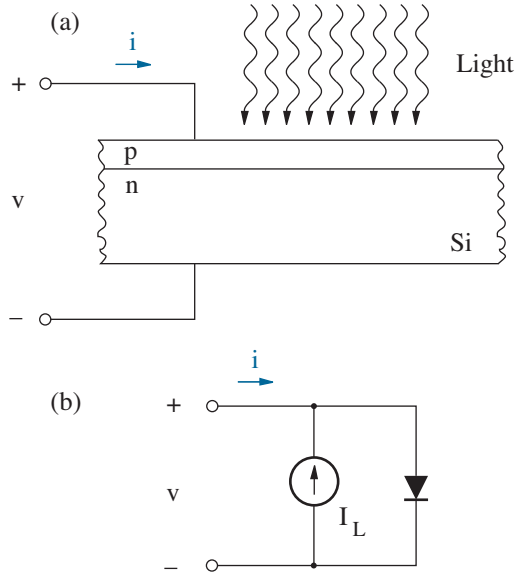
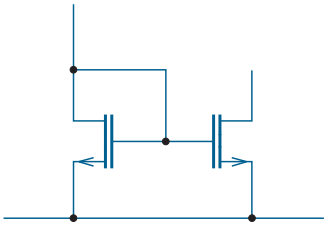


Figure P2.50



Chapter 3

Diode Circuit Analysis and Model Development

The graphical analytical procedure introduced in the Perspective is useful for understanding the function of non-linear circuits. But it is cumbersome, and it fails whenever a circuit contains more than one non-linear element. More versatile analytical tools are needed.

This chapter develops some formal methods for obtaining solutions to diode circuits with arbitrary complexity. We begin by finding "zero-order" solutions that assume ideal diode behavior. Then we obtain improved first- and second-order solutions that use diode **models** with increasing degrees of sophistication. For most students, the application of these methods will entail new patterns of thought that will prove to have enduring importance, especially with regard to design.

We conclude the chapter by discussing SPICE diode circuit solutions under static and transient conditions. For the latter analysis, we consider the physical basis of non-linear diode capacitance.

Objectives

After studying this chapter and working through the problems, you should be able to perform each of the following:

- Determine a zero-order solution to a diode circuit using the method of assumed diode states or the method of breakpoints (Section 3.1).
- Determine a first- or second-order solution to a diode circuit using an appropriate diode model (Section 3.2).
- Analyze a diode circuit using SPICE (Section 3.4).
- Define depletion and diffusion pn junction capacitance (Section 3.4).

3.1 Zero-Order Analysis

Consider the circuit of Fig. 3.1 in which the anode (triangle) and cathode sides of a diode are connected to node voltage v_a and ground, respectively. A zero-order solution for current i assumes that the diode is ideal, and we examine two methods for obtaining it.

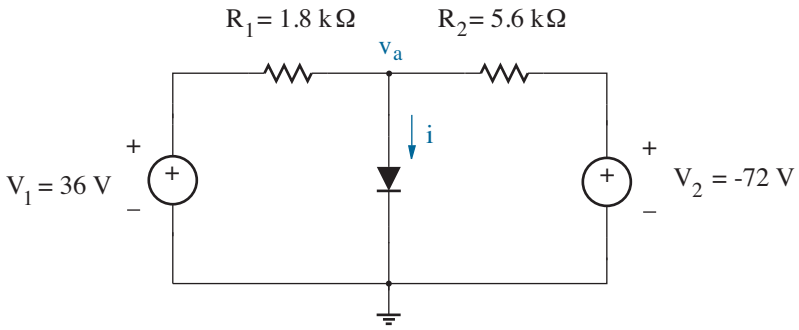


Figure 3.1: Demonstration circuit for diode analysis.

The Method of Assumed Diode States

The method of assumed diode states is a pompous way of describing a procedure in which we *guess* whether each diode in a circuit is either "on" (a short circuit) or "off" (an open circuit). In Fig. 3.1, the diode is shared between a pair of "Thevenized" circuits, and the 36-V source on the left appears to be far less significant than the -72-V source on the right. Thus, it seems reasonable to guess that the diode is "off". And although trivial, the solution $i = 0$ also seems reasonable.

— A job well done. Guessing is so easy.

But wait! Look at the node voltage v_a. Since the diode is "off", we can remove it from the circuit. Then by superposition,

$$v_a = 36\left(\frac{5.6}{1.8+5.6}\right) + (-72)\left(\frac{1.8}{1.8+5.6}\right) = 9.730 \text{ V}. \tag{3.1}$$

However, $v_a > 0$ is inconsistent with the diode "off" state. This difficulty can only be resolved by assuming the diode is "on" so that $v_a = 0$. Thus,

$$i = \frac{36}{1.8} + \frac{-72}{5.6} = 7.143 \text{ mA}\,, \tag{3.2}$$

and the result is consistent with the diode "on" state with $i > 0$.

Asked to characterize the preceding analytical method, your impression might fall somewhere between "not straightforward" and "handwaving." But careful consideration reveals a general pattern of thought summarized in the flow-chart diagram of Fig. 3.2. We began by making a set of circuit operating assumptions concerning the behavior of every non-linear device. Next, we chose an appropriate device model—in this case, a short circuit or an open circuit—that reduced the network to a form in which only linear elements are present. Then we solved the reduced circuit using standard techniques of linear analysis. Finally, and most importantly, *we checked to ensure that our solution was consistent with initial operating assumptions.* At first, it was not. Nevertheless, a set of revised behavioral conditions led to a satisfactory solution, and only then we were able to stop.

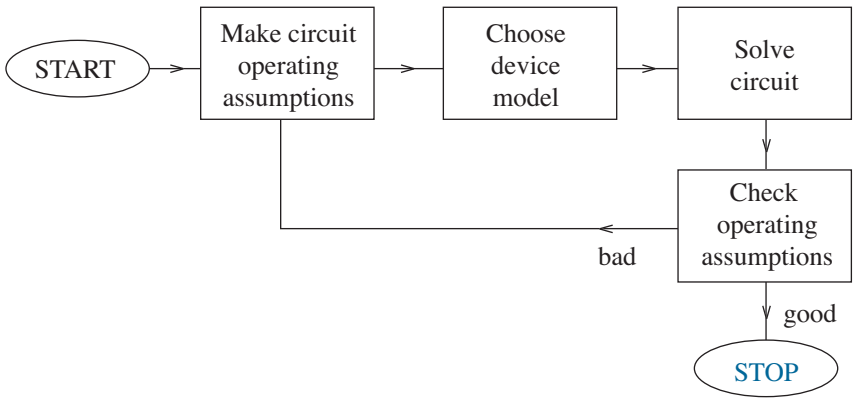


Figure 3.2: Flow-chart summary of a general analytical procedure.

We will repeatedly follow similar patterns of thought when considering more complicated non-linear circuits throughout other sections of this text. *The necessity of the solution consistency check cannot be overemphasized.*

Diode Consistency Checks

If the diode is "off" with zero current,
the diode voltage is *negative*.

If the diode is "on" with zero voltage,
the diode current is *positive*.

Exercise 3.1 For each of the circuits in Fig. 3.3, use the method of assumed diode states to determine the current i as indicated, then specify the applicable diode condition(s).

Ans: (a) $i = 1$ mA, D_1 "on" (b) $i = 3$ mA, D_1 "off"

(c) $i = 6$ mA, D_1 "on" (d) $i = 5$ mA, D_1 "on", D_2 "off"

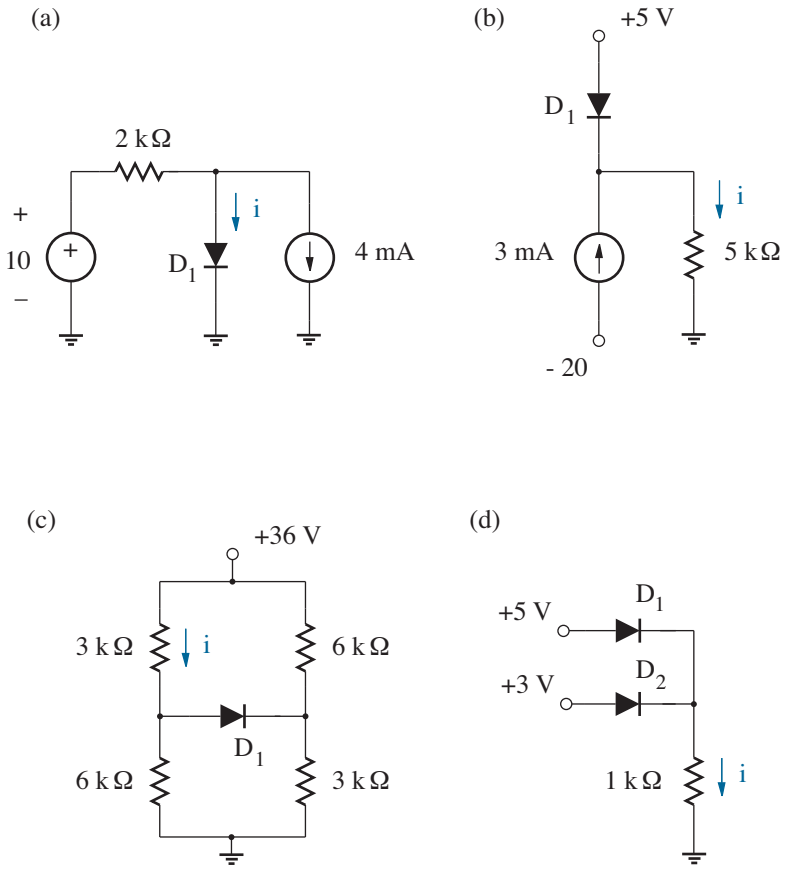


Figure 3.3: Circuits for Exercise 3.1.

The Method of Breakpoints

A second technique that is useful for the analysis of ideal diode circuits is called the method of breakpoints. It avoids guessing by determining circuit conditions, usually in terms of one of the source variables, that cause each diode to change between "on" and "off" states. For any particular diode, the change occurs when current and voltage are *both* zero. This corresponds to the point of slope discontinuity, or **breakpoint**, in the diode's current-voltage characteristic.

The method of breakpoints is easily applied to the circuit of Fig. 3.1. As a first breakpoint condition, we set $v_a = 0$, and we write an expression for the diode current i in terms of the source variable V_1. This yields

$$i\,(\text{mA}) = \frac{V_1}{1.8} + \frac{-72}{5.6}\,. \tag{3.3}$$

The other breakpoint condition requires $i = 0$, so we obtain $V_1 = 23.14$ V. A little thought reveals that for $V_1 > 23.14$ V, the diode is "on", and for $V_1 < 23.14$ V, the diode is "off". Thus, with $V_1 = 36$ V in the particular circuit of Fig. 3.1, the diode is "on", $v_a = 0$, and $i = 7.143$ mA.

Note that the voltage drop across any resistor in series with a diode is zero at breakpoint (as a consequence of zero diode current).

Exercise 3.2 Determine the values of source voltage V_s such that the diodes are at breakpoint in the circuits drawn below.

Ans: (a) $V_s = 9$ V (b) $V_s = -33$ V

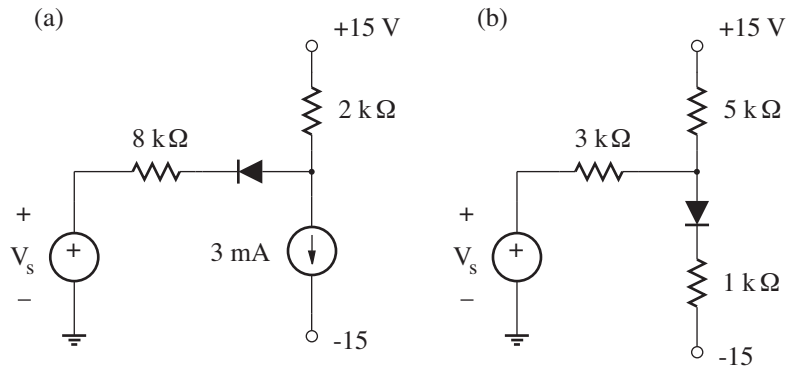


Figure 3.4: Circuits for Exercise 3.2.

Example 3.1

For the circuit of Fig. 3.5, determine the current i as a function of source voltage V_s. Both diodes are ideal.

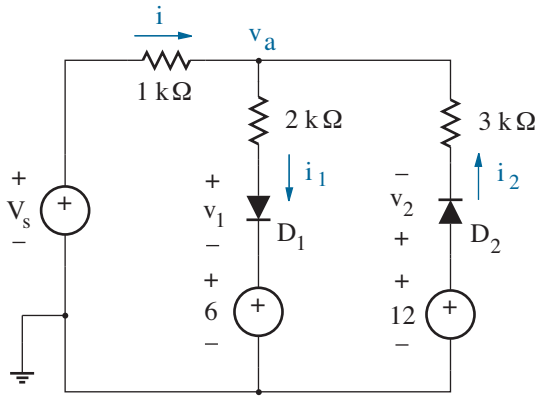


Figure 3.5: Circuit for Example 3.1

Solution

First, we need to determine the conditions that make diodes D_1 and D_2 either "on" or "off." We assume a particular state for one diode, and we apply the breakpoint method to learn about the other. In general, a circuit containing n diodes requires $n-1$ operating assumptions.

Suppose diode D_1 is "off." This removes D_1, so we also eliminate the dangling 2-kΩ resistor and 6-V source from the circuit. Then the breakpoint for diode D_2 occurs when $V_s = 12$ V —v_2 and i_2 are both zero, and D_2 is "on" when $V_s < 12$ V. Similarly, with D_2 "off", the breakpoint for diode D_1 occurs when $V_s = 6$ V, and D_1 is "on" when $V_s > 6$ V.

These results appear reasonable. But if D_1 is "on" when $V_s > 6$ V, it is also "on" when $V_s = 12$ V, which contradicts the assumption made when determining the breakpoint for D_2. The same problem applies to diode D_2 at the breakpoint for D_1. Thus, our initial circuit assumptions were poor, but we have been saved through a careful consistency check. Have patience! Our guesses will improve with practice.

Now assume that diode D_1 is "on". At the breakpoint for D_2, node-voltage $v_a = 12$ V (since $v_2 = 0$, $i_2 = 0$), and by superposition,

$$v_a = 12 = V_s \left(\frac{2}{1+2}\right) + 6 \left(\frac{1}{1+2}\right) . \tag{3.4}$$

(Note that we ignore the 12-V source, since $i_2 = 0$.) Thus, $V_s = 15$ V at breakpoint, and diode D_2 is "on" when $V_s < 15$ V. Similarly, assume that

D_2 is "on". At the breakpoint for D_1, $v_a = 6$ V, and by superposition,

$$v_a = 6 = V_s \left(\frac{3}{1+3}\right) + 12 \left(\frac{1}{1+3}\right) . \tag{3.5}$$

Thus, $V_s = 4$ V at the breakpoint, and diode D_1 is "on" when $V_s > 4$ V. These results are more than reasonable; they satisfy the consistency check.

Having determined the diode states for different ranges of V_s, we specify i as a function of V_s. For $V_s > 15$ V, D_1 is "on", D_2 is "off", and we find

$$i = \frac{V_s - 6}{3} \text{ mA} . \tag{3.6}$$

But for $V_s < 4$ V, D_1 is "off" and D_2 is "on". So now

$$i = \frac{V_s - 12}{4} \text{ mA} . \tag{3.7}$$

Finally, for $4 < V_s < 15$ V, D_1 and D_2 are both "on", and we obtain

$$i = \frac{5\,V_s - 42}{11} \text{ mA} . \tag{3.8}$$

This latter result can be found by superposition, but it is much easier to determine the equation of the line that connects the endpoints of the line segments specified by Eqs. 3.6 and 3.7. Figure 3.6 shows a plot of i vs V_s. Note the slope discontinuities that occur at the diode breakpoints.

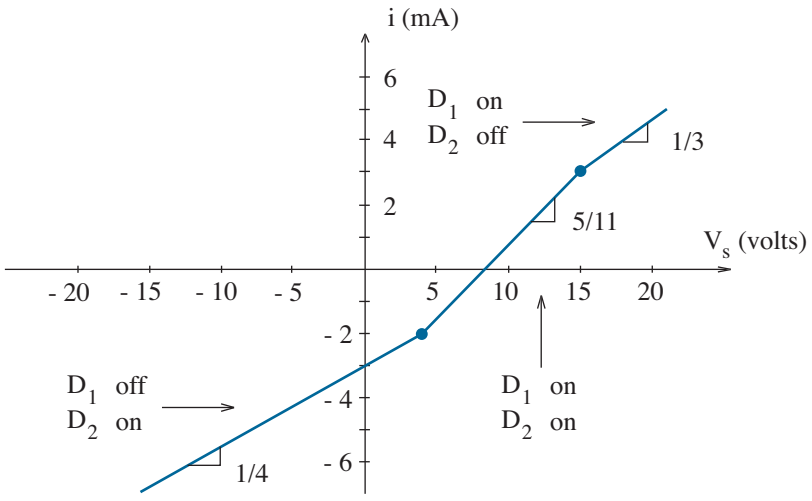


Figure 3.6: Solution for Example 3.1

Example 3.2

Determine i_x in the circuit of Fig. 3.7. Both diodes are ideal.

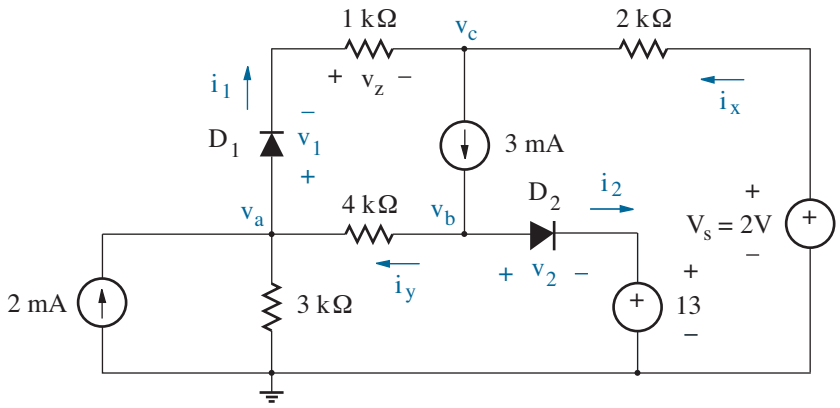


Figure 3.7: Circuit for Example 3.2

Solution

We determine the breakpoints for diodes D_1 and D_2 in terms of source V_s. When D_1 is at breakpoint, $v_1 = 0$ and $i_1 = 0$. And when D_2 is "on", $v_2 = 0$ so that node $v_b = 13$ V. The node equation for v_a now takes the form

$$2\text{ mA} + \frac{0 - v_a}{3\text{ k}\Omega} + \frac{13 - v_a}{4\text{ k}\Omega} = 0\,.$$

Thus, $v_a = 9$ V, $i_y = (13 - 9)/4$ kΩ $= 1$ mA, and $i_2 = 2$ mA. This i_2 value is positive and consistent with the D_2 "on" state. (We leave it as an exercise to show that D_2 cannot be "off" when D_1 operates at breakpoint.) Then with $v_1 = 0$ and $v_z = 0$ (since $i_1 = 0$), $v_c = 9$ V. The $i_1 = 0$ condition requires $i_x = 3$ mA, so $V_s = 9 + 3\text{ mA} \times 2\text{ k}\Omega = 15$ V. Increasing V_s tends to increase the voltage at the D_1 cathode, and D_1 is "off" for $V_s > 15$ V.

When D_2 is at breakpoint, $v_2 = 0$ and $i_2 = 0$. In turn, node $v_b = 13$ V. And with $i_y = 3$ mA, $v_a = 13 - 3\text{ mA} \times 4\text{ k}\Omega = 1$ V. Then an application of Kirchhoff's Current Law at v_a reveals $i_1 = 14/3$ mA —D_1 is clearly "on". Voltage v_c is $1 - 14/3\text{ mA} \times 1\text{ k}\Omega = -11/3$ V, and with $i_x = -5/3$ mA, $V_s = -11/3 - 5/3\text{ mA} \times 2\text{ k}\Omega = -7$ V. Increasing V_s tends to increase i_x in a manner that D_2 can absorb. Thus, D_2 is "on" for $V_s > -7$ V.

Whereas $V_s = 2$ V, our remaining task is to find i_x when D_1 is "on" and D_2 is "on" so that both function as effective short circuits. A quick node-voltage analysis—no need for details—shows that $v_c = 8/33 = 0.242$ V. Then $i_x = (2 - 8/33)/2\text{ k}\Omega = 0.879$ mA.

We are developing some patterns of thought with long-term benefits.

Example 3.3

Design the circuit of Fig. 3.8 so that $i_2 = 1$ mA. Both diodes are ideal.

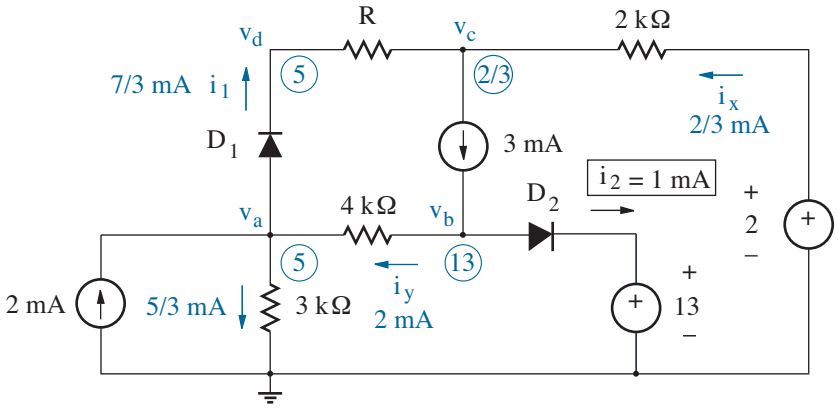


Figure 3.8: Circuit for Example 3.3

Solution

We mark up the circuit diagram with the results of some trivial calculations starting with the desired i_2. The goal is to find a consistent value for R.

- With $i_2 = 1$ mA, D_2 is in forward bias, and $v_b = 13$ V.
- The same i_2 requirement also has $i_y = 3\,\text{mA} - 1\,\text{mA} = 2$ mA.
- In turn, $v_a = 13 - 2\,\text{mA} \times 4\,\text{k}\Omega = 5$ V.
- The ground current through the 3-kΩ resistor is $5\,\text{V}/\,3\,\text{k}\Omega = 5/3$ mA.
- With $i_1 = 2\,\text{mA} + 2\,\text{mA} - 5/3\,\text{mA} = 7/3\,\text{mA}$, D_1 is in forward bias, and $v_d = 5$ V.
- Meanwhile, on the other side of R, $i_x = 3\,\text{mA} - 7/3\,\text{mA} = 2/3\,\text{mA}$.
- In turn, $v_c = 2 - 2/3\,\text{mA} \times 2\,\text{k}\Omega = 2/3$ V.
- Finally, $R = (5\,\text{V} - 2/3\,\text{V})/(7/3\,\text{mA}) = 1.857\,\text{k}\Omega$.

Marking up a circuit diagram following the application of KCL, KVL, and Ohm's Law (while maintaining consistent non-linear device behavior) is a powerful design practice that we will often employ throughout this text. Solving sets of complicated equations is seldom necessary.

3.2 Diode Models for Higher-Order Solutions

Once we have obtained a zero-order solution to a diode circuit, we can seek first- and second-order solutions that have successively higher degrees of precision by increasing the sophistication of the diode model. In each case, the diode model uses a combination of *linear* circuit elements in order to *approximate* non-linear device behavior (usually over a restricted range of voltage or current). After model substitution, a circuit solution is obtained using familiar techniques of linear analysis.

Figure 3.9 shows a useful set of linearized diode models. Each row in the figure corresponds to an order of sophistication, and the left and right columns correspond to forward- and reverse-bias conditions, respectively. We are already acquainted with the zero-order short- and open-circuit diode models that appear in the first row.

Consider a forward-biased diode. The first-order diode model derives by drawing a straight vertical line that tends to coincide with the steep portion of the diode current-voltage characteristic, and the point where this line intersects the voltage axis is denoted by V_f. Thus, a voltage source with value V_f yields a reasonable approximation to diode behavior if the forward diode current is "large." Apart from model failure, "small" currents suggest a nearly "off" condition, so V_f is often called the diode **turn-on voltage**. The "large" and "small" current distinction is qualitative, and a precise formula for V_f is unavailable. However, a commonly used rule-of-thumb is $V_f \sim 0.7$ V for a typical silicon diode when i is of the order of a few mA.

It is important to note that a real voltage source with value $V_f > 0$ exhibits a current-voltage characteristic that extends into the region of the $i-v$ plane where $i < 0$. This does not concern us since we are (habitually?) prepared to disallow diode circuit solutions that go beyond the limits of model validity. *A forward-bias diode model must always yield a circuit solution that is consistent with positive diode current.*

The second-order model for a forward-biased diode is similar to the first-order model except for the inclusion of a series resistance so that the straight-line approximation to the steep portion of the diode current-voltage characteristic has finite slope. This resistance is denoted by R_f, and it is typically of the order of a few ohms or less. Note that the value of V_f featured in the second-order model is not necessarily the same as (and is usually slightly less than) the value of V_f in the first-order model.

Exercise 3.3 A silicon pn diode features $I_s = 5 \times 10^{-13}$ A and $n = 1.2$. Estimate V_f if the diode operates over a 1 - 10 mA range.

Ans: $V_f = 0.72$ V

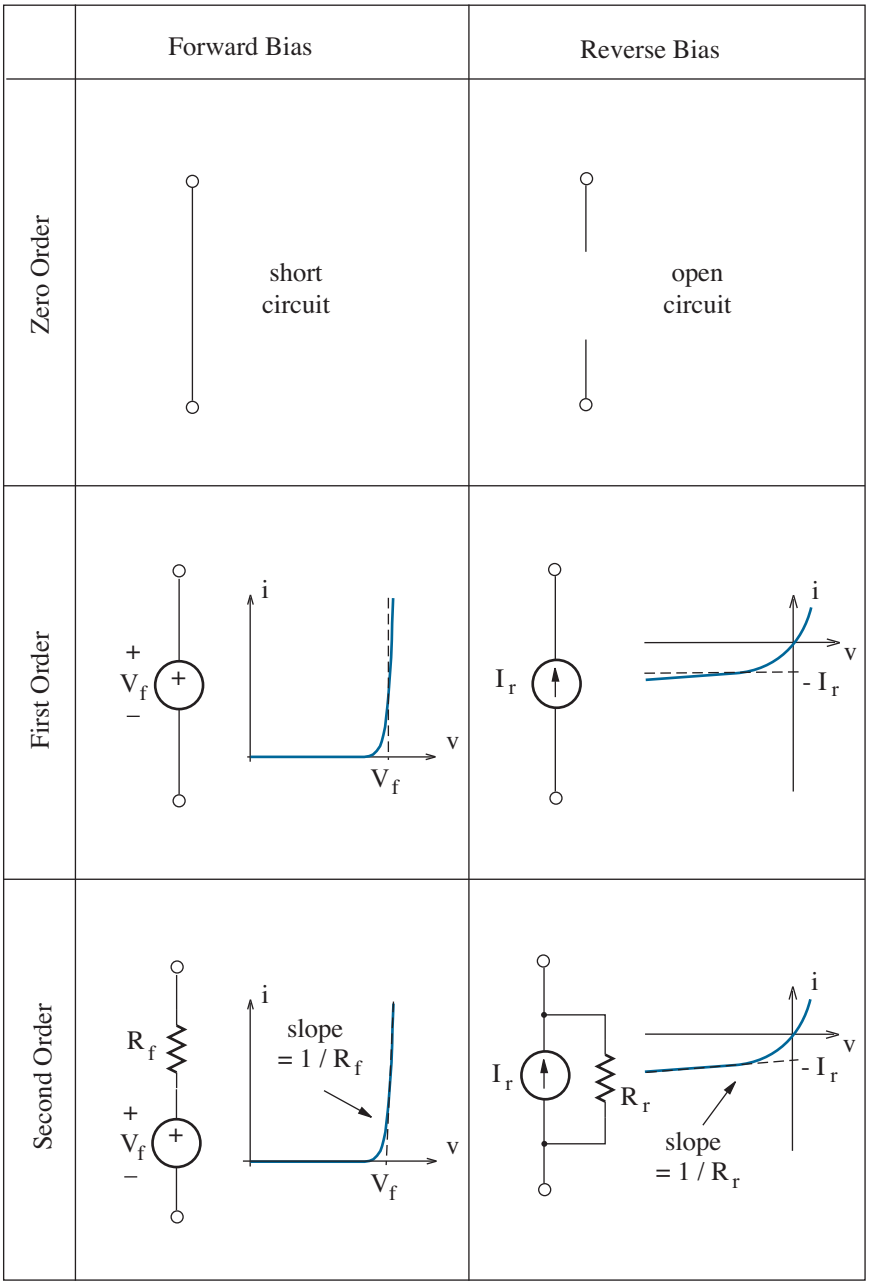


Figure 3.9: Linearized diode models for dc analysis.

Now consider a reverse-biased diode *that is not in breakdown.* The first-order diode model derives by drawing a straight horizontal line that tends to coincide with the flat portion of the diode current-voltage characteristic, and the point where this line intersects the current axis is denoted by $-I_r$. Thus, a reverse-directed current source with value I_r yields a reasonable approximation to diode behavior if the magnitude of the reverse diode voltage is "large" (or at least much larger than the thermal voltage kT/q). Typically, I_r is less than 1 μA.

As indicated previously, it is important to remember the limits of model validity. A real current source with value $-I_r < 0$ exhibits a current-voltage characteristic that extends into the region of the $i - v$ plane where $v > 0$. However, *a reverse-bias diode model must always yield a circuit solution that is consistent with negative diode voltage.*

The second-order model for a reverse-biased diode that is not in breakdown is similar to the first-order model except for the inclusion of a parallel resistance so that the straight-line approximation to the flat portion of the diode current-voltage characteristic has non-zero slope. This resistance is denoted by R_r, and it is typically of the order of tens of MΩ or greater. Note that the value of I_r in the second-order model is not necessarily the same as (and is usually slightly less than) the value of I_r featured in the first-order model.

Model parameters V_f, R_f, I_r, and R_r lack special physical significance, and they are best determined by measurement. Nevertheless, V_f, which is by far the most troublesome of these parameters, is proportional to $\ln(i/I_s)$. In turn, I_s is proportional to ${n_i}^2$, which is exponentially dependent on the semiconductor bandgap energy (see Eq. 2.54). Diodes from semiconductors with relatively small bandgaps, such as germanium, feature the lowest V_f values at a given current level. Schottky (metal/semiconductor) diodes also tend to exhibit low V_f values.

Once again, we consider the circuit of Fig. 3.1. The zero-order solution indicated that the diode is forward biased, so we seek a first-order solution by replacing the diode with the first-order forward-bias model. No problem. But what value should be used for V_f?

For the moment, we will try to be as precise as possible, and we will abandon rules of thumb such as $V_f \sim 0.7$ V in favor of a measurement. Since the zero-order solution (with a new notation) is $i^{(0)} = 7.143$ mA, it is reasonable to measure the diode current-voltage characteristic over a 0 - 10 mA range. And for diode xyz, we might estimate $V_f = 0.74$ V by drawing a vertical dashed line as shown in Fig. 3.10a. The remaining analysis is straightforward. Specifically,

$$i^{(1)} = \frac{36 - 0.74}{1.8} + \frac{-72 - 0.74}{5.6} = 6.599 \text{ mA}\,, \tag{3.9}$$

which is consistent with the anticipated 0 - 10 mA diode current range.

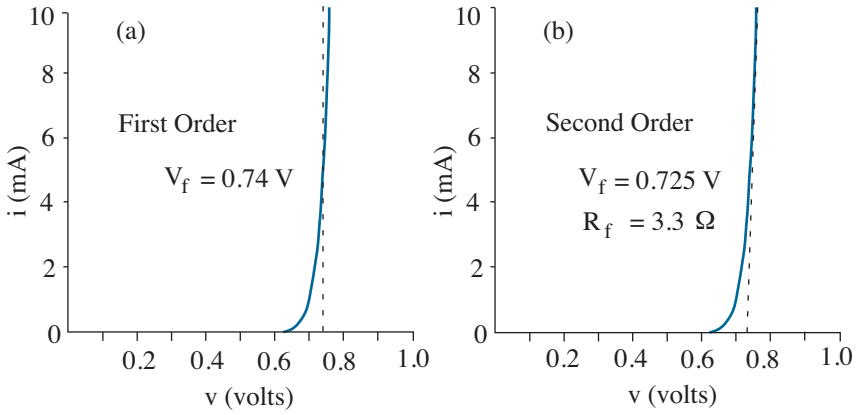


Figure 3.10: Model parameter measurements for diode xyz.

The importance of selecting a reasonable current range when measuring forward-bias model parameters is illustrated in Fig. 3.11. Here, we have current-voltage data for the same diode xyz with 0 - 10 μA and 0 - 10 A current scales. The first-order V_f values are significantly different.

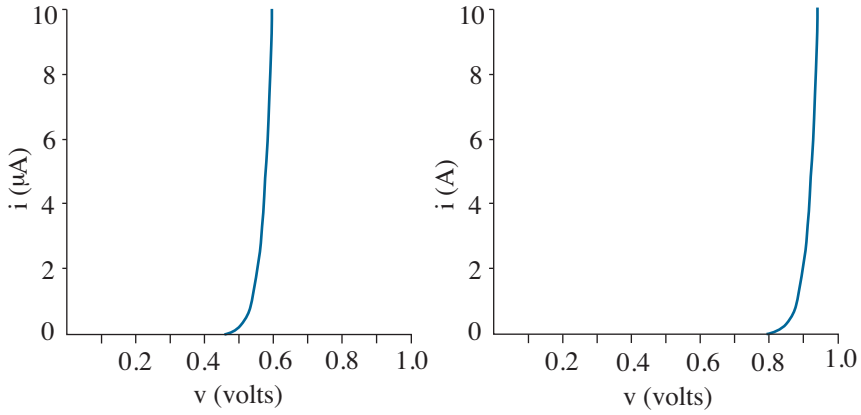


Figure 3.11: Forward-bias data for diode xyz over different current ranges.

One finds a second-order solution following the first-order procedure. We use the second-order forward-bias model, and we estimate $V_f = 0.725$ V and $R_f = 3.3\ \Omega$ by drawing a slightly tilted dashed line close to the diode's measured current-voltage characteristic as shown in Fig. 3.10b. The circuit analysis is more complicated with resistor R_f, yet we eventually obtain

$$i^{(2)} = 6.595 \text{ mA}. \tag{3.10}$$

In the absence of a third-order forward-bias diode model, this represents the best solution that can be obtained using linear analysis.

Exercise 6.4 Consider the circuit of Fig. 3.1, but let $V_2 = -60$ V. Determine zero-order, first-order, and second-order solutions for current i (using the V_f and R_f data from Fig. 3.10).

Ans: $i^{(0)} = 9.286$ mA, $i^{(1)} = 8.742$ mA, $i^{(2)} = 8.732$ mA

A postscript to the preceding discussion requires an honest confession. The "data" for diode xyz in Figs. 3.10 and 3.11 actually correspond to a "textbook" diode for which

$$i = I_s \left[e^{qv/nkT} - 1 \right] \tag{3.11}$$

with $I_s = 2 \times 10^{-15}$ A, $n = 1$, and $kT/q = 25.9$ mV. So the approximate zero-, first-, and second-order solutions can be compared with an "exact" solution that is determined by iteration. The following algorithm is a sure treat for the average computer.

Round 1: Let $v = 0$. Store the circuit constraint:

$$i = \frac{36 - v}{1.8} + \frac{-72 - v}{5.6} \, . \tag{3.12}$$

Then find $i = 7.143$ mA. This is the same as the programmer's $i^{(0)}$.

Round 2: Given the current from Round 1, calculate a new value for diode voltage v using an expression of the form

$$v \approx n \left(\frac{kT}{q} \right) \ln \left(\frac{i}{I_s} \right) \, . \tag{3.13}$$

Thus, $i = 7.143$ mA corresponds to $v = 0.7486$ V. In turn, go back to Eq. 3.12 to find $i = 6.593$ mA.

Round 3: Given the current from Round 2, Eq. 3.13 yields $v = 0.7465$ V. In turn, use Eq. 3.12 again to find $i = 6.595$ mA.

Round 4: Given the current from Round 3, Eq. 3.13 yields $v = 0.7465$ V. In turn, use Eq. 3.12 again to find $i = 6.595$ mA.

Observation: After the preceding round, the diode current has converged in the sense that there is no change—the programmer is content with only four significant figures of accuracy. Make that dull-witted person happy by spitting out $i = 6.595$ mA—the same result as $i^{(2)}$. End of program. Exhibit smug satisfaction (based on superior numerical capability).

Are you now impressed with our second-order solution? Read on.

3.3 The Perils of Precision . . .

You may have noticed that the preceding sections specified solutions to the demonstration circuit of Fig. 3.1 with the shameless precision of four significant figures. Back in the "good old days" of slide-rule calculations, one seldom tried to specify numerical results to more than two significant figures of accuracy because slide-rule scales are difficult to read. However, modern times permit the use of electronic calculators, and even the most primitive of these could have easily determined that $i^{(2)} = 6.5946388$ mA. This uncomfortable eight-significant-figure result begs to be rounded off. But to what extent?

The circuit of Fig. 3.1 offers a clue since none of its component values is specified to more than two significant figures. One expects less precision for a circuit containing a 5.6-kΩ resistor than for a similar circuit containing a 5.600-kΩ resistor. Yet degrees of component uncertainty are not judged by significant figures alone. A more useful measure is percentage **tolerance**, which is provided by a manufacturer as a means of ensuring a "uniform" product despite irregularities in the manufacturing process.

Suppose you own a carbon-valley company that makes 5.6-kΩ resistors. Your customers have agreed that they will buy your product at a particular unit price if each resistor value is 5.6 kΩ ± 5 %. Fortunately, your product **yield** is 90 %, or equivalently, 90 % of all resistors that are intended to have a 5.6-kΩ value fall within the specified tolerance range. You cannot sell the remaining 10 % of your product, but you still make a tidy profit.

Now suppose you decide to sell only "precision" resistors with 5.6 kΩ ± 0.5 % value. If you are constrained to use the same manufacturing process as before, would you sell the precision resistors at the same unit price? No! Reduced tolerance implies reduced product yield, perhaps only 10 % here, so a much higher price is needed in order to maintain an acceptable profit. Some fussy customers will be prepared to accept the higher price, but there will be few of them.

Most circuit designers avoid the use of costly precision components if they can achieve desired circuit performance to within an acceptable error. Of course, it is wrong to assume that all components in a circuit have equal influence on every voltage or current. Clever design practice often reduces circuit sensitivity to component characteristics that are particularly uncertain or susceptible to temperature change.

For the inverse analytical process, *a circuit solution need be no more precise than the margin of error introduced by component tolerance factors.* We are often hard pressed to apply this as a rule, since tolerance data is not always specified and, as noted previously, uncertain components may have unequal influence. Instead, we express circuit solutions that seem comfortably precise in terms of our acquired experience.

Returning to the circuit of Fig. 3.1, we let R_1 and R_2 each have typical ± 5 % tolerance. Table 3.1 indicates the zero-, first-, and second-order solutions for the diode current when the resistors have "as-specified" and "worst-case" values. Both V_1 and V_2 are fixed.

Table 3.1:

R_1 (kΩ)	R_2 (kΩ)	$i^{(0)}$ (mA)	$i^{(1)}$ (mA)	$i^{(2)}$ (mA)
1.80	5.60	7.143	6.599	6.595
1.71	5.88	8.808	8.249	8.240
1.89	5.32	5.514	4.983	4.982

"Worst-case" solutions differ from "as-specified" solutions by about 25 %, so four significant figures of accuracy is clearly excessive.

We can make no claim that the preceding example is representative of all other electronic circuits, and we particularly do not wish to imply that resistance variations are more important than specification variations for other components. Notwithstanding, we now adopt the following practice: *Circuit solutions will be rounded to two significant figures of accuracy after calculations featuring three significant figures of accuracy.* And even after rounding, we should be suspicious of our result until we have considered sensitivity to component variations (on a circuit-by-circuit basis).

. . . And the Choice of an "Acceptable" Model

Given the availability of zero-, first-, and second-order diode models, which should be used to find an "acceptable" solution to some arbitrary circuit? No failsafe rules are available. But in light of our comments on precision, we offer the following *general* guidelines:

- Zero-order diode models are used to demonstrate circuit function.
- A first-order diode model should be considered when the Thevenin equivalent (forward) voltage at the diode terminals is less than about 10 V or when the Norton equivalent (reverse) current at the diode terminals is less than about 10 μA.
- Second-order diode models are seldom necessary. Exceptions may apply when the zero-order forward-bias diode current or the zero-order reverse-bias diode voltage is very large (usually several amperes or several hundred volts, respectively).

When applying these guidelines, there is no substitute for experience.

3.4 Computer Solutions

To some extent the analytical methods considered thus far reflect one of our inherent weaknesses. We are prone to isolation with no more than pencil and paper. And under these circumstances, we welcome approximate diode models and circuit solutions that require more conceptual thought than toilsome grind. But the ubiquitous personal computer is no mere mortal—it thrives on incessant number crunching. We need not hesitate to provide the computer with an all-purpose non-linear diode model that is open to improvement as the physical understanding of the device progresses.

.model Statement and Support

When applied to the demonstration circuit of Fig. 3.12, the SPICE code takes the following form:

```
* Diode Demonstration Circuit

V1          1           0           36
V2          3           0           -72
R1          1           2           1.8K
R2          2           3           5.6K
D1          2           0           TestDiode

.model      TestDiode   D           ( IS = 2f )

.op

.end
```

Here, the third statement group is a .model specification that indicates the characteristics of a non-linear diode, specifically, saturation current I_s. The .op command statement requests the diode operating point.

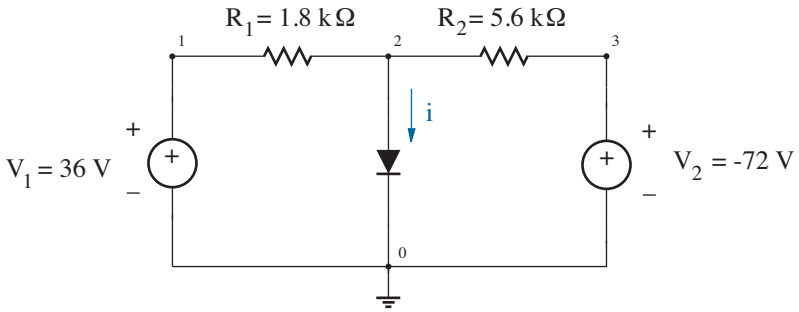


Figure 3.12: Demonstration circuit with node labels for SPICE analysis.

Here is a typical portion of the corresponding output file . . .

```
*Diode Demonstration Circuit

****     Diode MODEL PARAMETERS

*****************************************************************

                    TestDiode
          IS        2.000000E-15


****     SMALL SIGNAL BIAS SOLUTION
                    TEMPERATURE = 27.000 DEG C

*****************************************************************

NODE     VOLTAGE  NODE     VOLTAGE  NODE     VOLTAGE

(1)      36.0000  (2)      .7455    (3)      -72.0000

VOLTAGE SOURCE CURRENTS

NAME                CURRENT

V1                  -1.959E-02
V2                   1.299E-02

TOTAL POWER DISSIPATION     1.64E+00 WATTS


****     OPERATING POINT INFORMATION
                    TEMPERATURE = 27.000 DEG C

*****************************************************************

**** DIODES

NAME                D1
MODEL               TestDiode
ID                  6.60E-03
VD                  7.46E-01
```

. . . and the solution for the diode current (ID) is in good agreement with the results of Section 3.2.

Painless procedure? Yes. But the computer solution is dependent upon —and is obviously no better than—the choice of a particular diode model. This is set by specifying keyword *changes* (in relation to the default values) in the .model statement. Some important keywords appear in Table 3.2. All apply under room-temperature conditions.

Table 3.2:

Symbol	SPICE keyword	Parameter Name	Default value	Unit
I_s	IS	Saturation current	10^{-14}	A
n	N	Ideality factor	1	
R_s	RS	Parasitic resistance	0	Ω
BV	BV	Reverse breakdown voltage	infinite	V
IBV	IBV	Reverse breakdown current	10^{-10}	A
E_g	EG	Bandgap energy	1.11	eV
$C_d(0)$	CJO	Zero-bias capacitance	0	F
V_{bi}	VJ	Built-in potential	1	V
m	M	Grading coefficient	0.5	
τ_d	TT	Transit time	0	sec

The first three table entries, which merit greatest attention, apply to the current-voltage expressions of Eqs. 2.48 and 2.53 subject to $v > -5kT/q$. Each parameter is obtained from a measurement of $\log i$ vs v for $v > 0$. In the preceding example, diode xyz was described by Eq. 3.11 with $I_s = 2 \times 10^{-14}$ A, $n = 1$, and zero R_s series resistance. Thus, only IS was revised in the .model statement.

Sometimes it is desirable to simulate circuits containing *ideal* diodes. This is easily accomplished by setting the ideality factor to a very small (but non-physical) value such as 0.01.

The next two table entries apply to the point of reverse breakdown in the current-voltage characteristic of Fig. 2.19. For $v <$ -BV, the model assumes that the reverse breakdown current increases exponentially with $-v + \text{BV}$ in kT/q units. The new parameter IBV denotes the magnitude of the reverse diode current at the onset of breakdown.

Parameter E_g denotes the semiconductor bandgap energy, with default value applicable to silicon. For germanium and GaAs pn diodes, E_g is 0.68 and 1.42 eV, respectively. The bandgap energy is used in conjunction with Eqs. 2.45 and 2.54 to calculate variations of parameter I_s with temperature. Temperature changes are made with the .temp statement.

Diode Capacitance

In this and other chapters, we consider physical phenomena that limit the time response of an electronic device in conjunction with SPICE discussion. Dynamic circuit analysis is usually reserved for the computer.

As the diode voltage is changed, we must either supply or remove free-carrier charge in order to accommodate changes in the depletion-region widths or changes associated with the level of forward injection current. By definition, the variation of charge with voltage is capacitance.

Consider the p-side depletion region with doping concentration N_a and depth x_p (see Fig. 2.13). The uncovered depletion charge is

$$Q_d = -qAN_a x_p \,, \tag{3.14}$$

where A is the junction area. Chapter-2 equations for charge balance (2.15) and built-in potential V_{bi} (2.18) are solved simultaneously for x_p to obtain

$$x_p = \sqrt{\frac{2\epsilon}{q}\,\frac{N_d/N_a}{N_d+N_a}\,(V_{bi}-v)}\,. \tag{3.15}$$

Thus,

$$C_d = \frac{\partial Q_d}{\partial v} = \frac{\partial Q_d}{\partial x_p}\,\frac{\partial x_p}{\partial v} = C_d(0)\left[1-\frac{v}{V_{bi}}\right]^{-1/2}, \tag{3.16}$$

where $C_d(0)$, the junction capacitance at zero bias voltage, is given by

$$C_d(0) = A\sqrt{\frac{\epsilon\, q}{2V_{bi}}\,\frac{N_d N_a}{N_d+N_a}}\,. \tag{3.17}$$

Equation 3.16 represents a **depletion capacitance**, which corresponds to a parallel-plate capacitor with plate separation $w = x_p + x_n$. Specifically,

$$C_d = \frac{\epsilon A}{w}\,, \tag{3.18}$$

where

$$w = \sqrt{\frac{2\epsilon}{q}\,\frac{N_d+N_a}{N_d N_a}\,(V_{bi}-v)}\,. \tag{3.19}$$

SPICE parameters CJO and VJ in Table 3.2 correspond to $C_d(0)$ and V_{bi}, respectively. Parameter -M ($-m$) replaces the -1/2 power factor in Eq. 3.16. In general, $0.33 < m < 0.5$, with $m = 0.5$ for an abrupt pn junction and $m = 0.33$ for a linearly graded pn junction (with gradual p-to-n transition). We will tend to let $m = 0.5$, the default value.

Exercise 3.4 A silicon pn diode with $V_{bi} = 0.75$ V features a depletion capacitance of 40 pF when $v = -6$ V. Determine $C_d(0)$.

Ans: 120 pF

Apart from the depletion capacitance, a forward-biased diode exhibits additive "storage" or **diffusion capacitance** related to the distribution of injected minority-carrier charge on each side of the junction (see Fig. 3.13). In the steady state (for which time derivatives vanish), we assume that the total injected charge Q_s is proportional to the diode current. Thus,

$$Q_s = \tau_d \, i \, , \tag{3.20}$$

and

$$C_s = \frac{\partial Q_s}{\partial v} = \tau_d \, \frac{\partial i}{\partial v} \, . \tag{3.21}$$

So in view of the diode current-voltage relation (Eq. 2.48), we have

$$C_s = \frac{\tau_d}{nkT/q} \, I_s \, e^{\, qv/kT} \approx \frac{i \, \tau_d}{nkT/q} \, . \tag{3.22}$$

Parameter τ_d appears as SPICE keyword TT in Table 3.2.

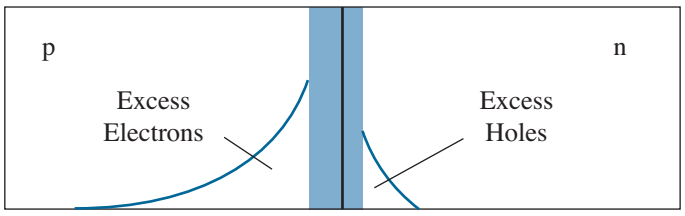


Figure 3.13: Minority-carrier charge near a forward-biased pn-junction. The junction depletion regions are shaded.

The physical interpretation of τ_d is dependent on the minority-carrier injection conditions and the widths of the diode neutral regions. When the neutral-region widths are made short to reduce parasitic series resistance, τ_d denotes an average **transit time** for carriers to diffuse to one of the diode contacts (where they die through recombination). In a "long" diode, however, most of the minority carriers recombine before reaching a contact, and τ_d denotes an average **minority-carrier lifetime**. We need not worry about these distinctions, since τ_d is generally a measured quantity.

To demonstrate the effects of parameter τ_d, we use SPICE to simulate the circuit of Fig. 3.14 (without bothering to list the appropriate code). Voltage source V_s abruptly switches between -10 V and +10 V at $t = 0$. The diode features $I_s = 2 \times 10^{-14}$ A and $\tau_d = 100$ ns.

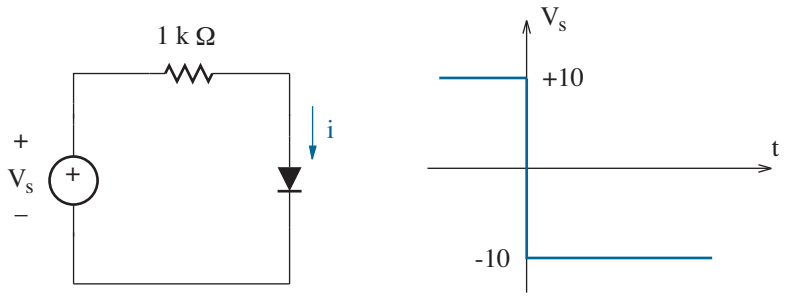


Figure 3.14: Demonstration circuit for diode transient response.

Figure 3.15 shows .probe time dependence of diode current i. For $t < 0$, the diode current is 9.305 mA, which is consistent with a forward diode voltage of about 0.7 V. At $t = 0$, the diode current suddenly changes to -10.67 mA, which is consistent with the diode voltage remaining steady. This negative current persists for 60 ns. Then the diode current rapidly approaches zero over the course of another 3 ns.

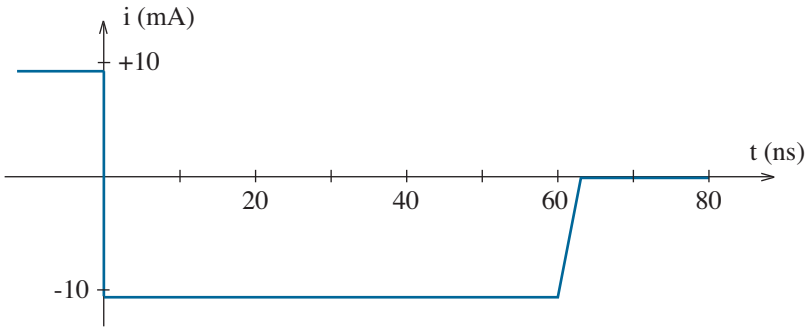


Figure 3.15: SPICE simulation results for the circuit of Fig. 3.14.

How can we explain the unexpected *negative* transient diode current? Once established under forward-bias conditions, the diffusion charge does not instantaneously disappear when the diode is subjected to reverse bias. And the presence of excessive minority-carrier concentrations at the edges of the depletion regions imposes a constraint of positive junction voltage—recall the Boltzmann relation between carrier concentration and potential. Negative diode current reflects the removal of diffusion charge.

A quantitative explanation begins with the **charge control** relation

$$i = \frac{Q_s}{\tau_d} + \frac{dQ_s}{dt} \,, \tag{3.23}$$

which states that the diode current features two components: one that is proportional to Q_s (to support the minority-carrier recombination process) and another related to changes in Q_s that add or remove diffusion charge. Note that the charge control relation reduces to Eq. 3.20 in the steady-state. For $t < 0$, let $i = I_F$. The initial charge store is thus

$$Q_s(0) = \tau_d \, I_F \,. \tag{3.24}$$

Subsequently, for $0 < t < t_s$, we observe constant $i = -I_R$. In turn,

$$-I_R = \frac{Q_s}{\tau_d} + \frac{dQ_s}{dt} \,. \tag{3.25}$$

Subject to the initial $Q_s(0)$ condition, the solution is

$$Q_s = \tau_d \, (I_F + I_R) \, e^{-t/\tau_d} - \tau_d \, I_R \,. \tag{3.26}$$

We solve for t_s, the time at which $Q_s = 0$, to obtain

$$t_s = \tau_d \, \ln \left(1 + \frac{I_F}{I_R} \right) \tag{3.27}$$

as the diode **storage time**. Equation 3.26 has no physical basis for $t > t_s$. With $t_s = 60$ ns, $I_F = 9.305$ mA, and $I_R = 10.67$ mA, we find $\tau_d = 96$ ns, not far from the 100-ns value used in the simulation.

In fairness, we note that the preceding experiment was "rigged" since SPICE uses the Q_s charge-control relation to determine transient response. Actual variations in the distribution of diffusion charge and the currents thus obtained are relatively difficult to model and support larger τ_d values. Nevertheless, the measurement procedure implied by Eq. 3.27 is adequate to obtain an effective τ_d that is useful for reasonably accurate simulations.

Two major options are available for improving diode transient response: One can use a **fast-recovery** pn diode that is fabricated with additional impurities such as gold to assist minority-carrier recombination (lower τ_d). However, the improvement in switching speed is achieved at the expense of increased reverse-bias leakage current and increased ideality factor ($n \sim 2$). Both conditions reflect a larger value of I_{s2} in Eq. 2.47. One can also use a Schottky diode, which does not feature minority-carrier diffusion charge. Of course, neither option eliminates the effect of depletion capacitance, which prolongs the post-storage-time decay in Fig. 3.15.

Concept Summary

Diodes exhibit very different circuit behavior in relation to a breakpoint for which $v = 0$ and $i = 0$.

- The forward-bias state has $i > 0$; the reverse-bias state has $v < 0$.
- The diode state can be assumed, or it can be determined in relation to a calculated breakpoint.
- Linear diode i-v models of varying order support hand calculations.
  - For forward bias,
    - Zero order: short circuit.
    - First order: voltage source with value V_f (turn-on voltage). Typically, $V_f \sim 0.7$ V when i is in the mA range.
  - For reverse bias,
    - Zero order: open circuit.
    - First order: current source with value $-I_r$ (small leakage).
  - Second-order models are available (Fig 3.9) but seldom used.
- Non-linear diode i-v models support computer calculations.
  - SPICE .model parameters are available (Table 3.2).
  - Some models have implicit temperature dependence.
- The diode state must be consistent with the circuit solution.
- Circuit solutions with exceptional precision are often overwhelmed by statistical variations that relate to uncertain components.
- Non-linear diode capacitance complicates time-dependent circuits.
  - A reverse-biased diode has depletion capacitance of the form

$$C_d = C_d(0)\left[1 - \frac{v}{V_{bi}}\right]^{-m},$$

    where $C_d(0)$ is the capacitance at zero bias, V_{bi} is the built-in junction potential, and m is typically 0.5. Depletion capacitance decreases as the reverse voltage becomes more negative.
  - The forward-bias capacitance includes a "diffusion" capacitance, which is proportional to the diode current.

Problems

Section 3.1

Assume ideal diodes unless otherwise indicated.

3.1 In the circuit of Fig. P3.1, find V_1 so that the diode is at breakpoint.

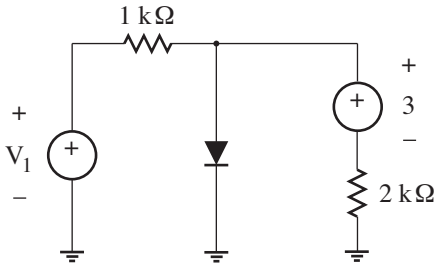


Figure P3.1

3.2 In the circuit of Fig. P3.2, find I_o so that the diode is at breakpoint.

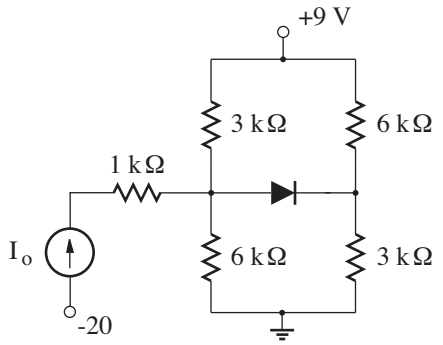


Figure P3.2

3.3 In the circuit of Fig. P3.3, find R so that the diode is at breakpoint.

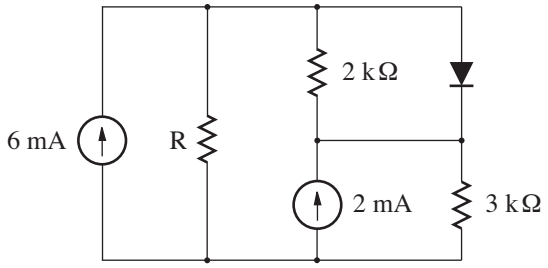


Figure P3.3

3.4 In the circuit of Fig. P3.4, find R so that the diode is at breakpoint

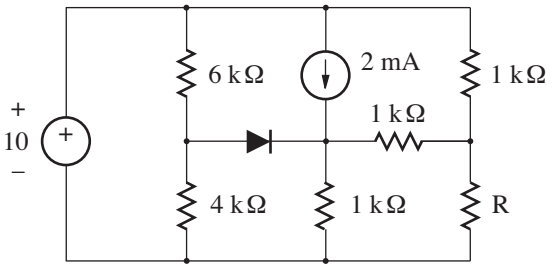


Figure P3.4

3.5 Determine v in the circuit of Fig. P3.5.

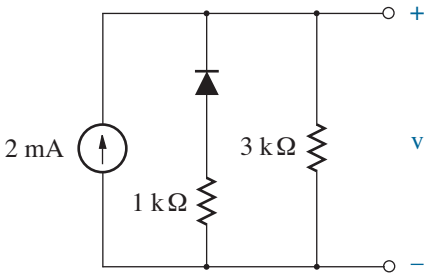


Figure P3.5

3.6 Determine i In the circuit of Fig. P3.6.

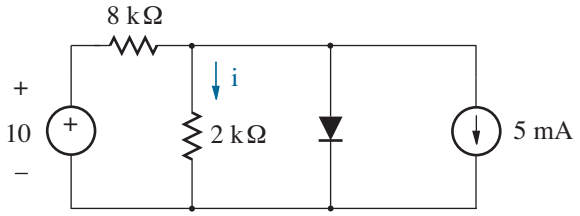


Figure P3.6

3.7 Sketch v_{out} vs. v_{in} $(-20 \leq v_{in} \leq 20$ V) for the circuit of Fig. P3.7.

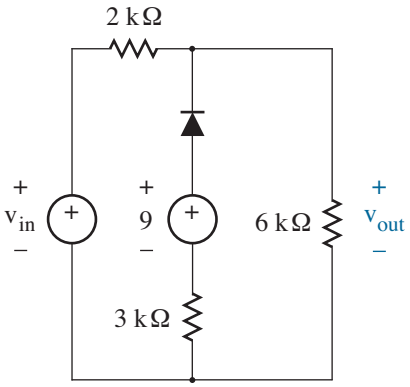


Figure P3.7

3.8 Sketch v_{out} vs. i $(-20 \leq i \leq 20$ mA) for the circuit of Fig. P3.8.

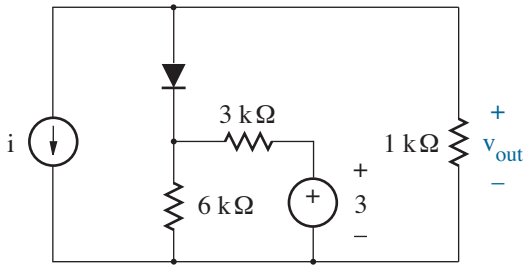


Figure P3.8

3.9 In the circuit of Fig. P3.9, find I_o so that D_1 is at breakpoint, then do the same for D_2.

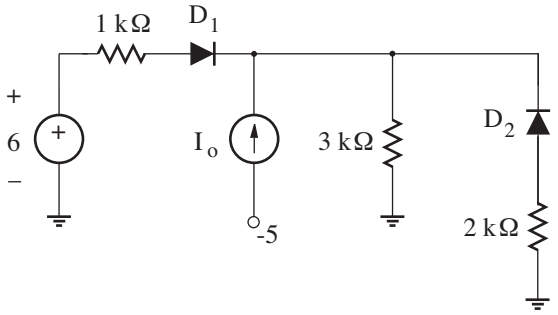


Figure P3.9

3.10 In the circuit of Fig. P3.10, find I_o so that D_1 is at breakpoint, then do the same for D_2.

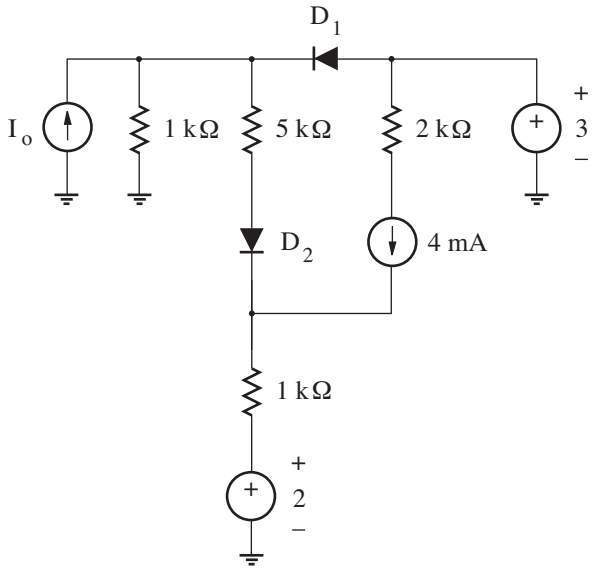


Figure P3.10

3.11 In the circuit of Fig. P3.11, find V_s so that D_1 is at breakpoint, then do the same for D_2.

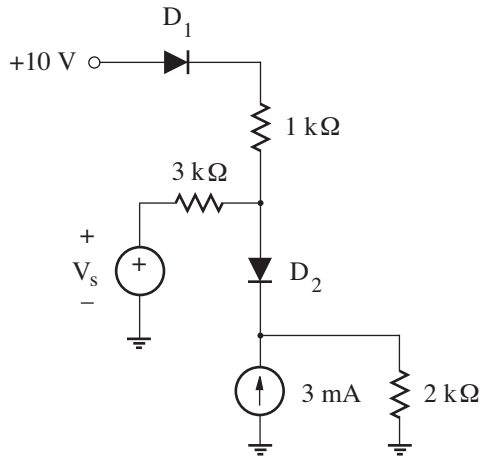


Figure P3.11

3.12 In the circuit of Fig. P3.12, find I_o so that D_1 is at breakpoint, then do the same for D_2.

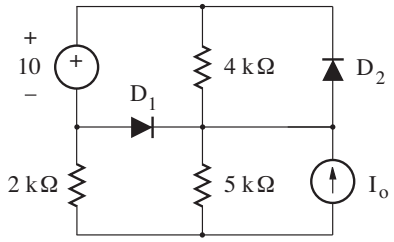


Figure P3.12

3.13 Determine i in the circuit of Fig. P3.13.

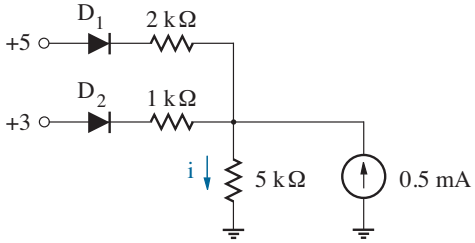


Figure P3.13

3.14 Determine i in the circuit of Fig. P3.14.

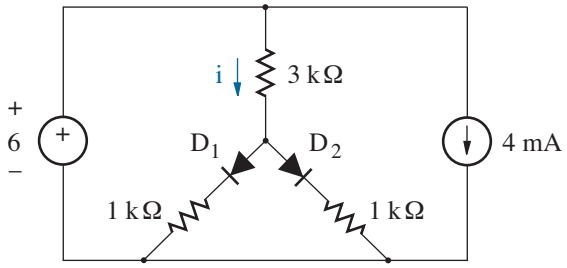


Figure P3.14

3.15 Determine i in the circuit of Fig. P3.15.

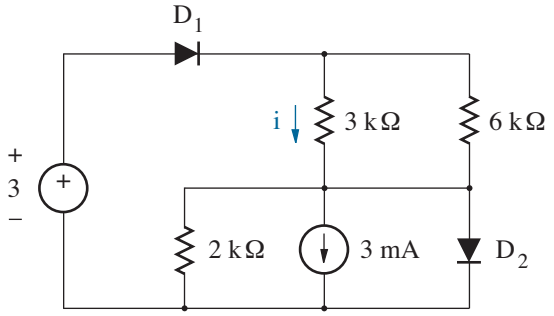


Figure P3.15

3.16 Determine i in the circuit of Fig. P3.16.

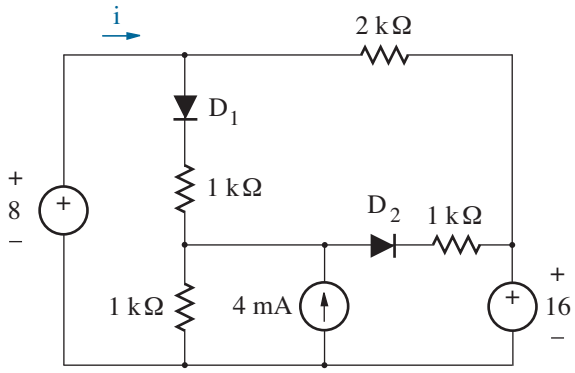


Figure P3.16

3.17 Sketch $v_{out}(t)$ for the circuit of Fig. P3.17 if $v_{in}(t) = 10 \sin \omega t$.

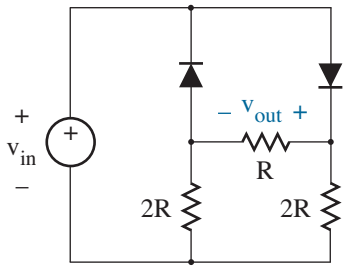


Figure P3.17

3.18 Sketch $v_{out}(t)$ for the circuit of Fig. P3.18 if $v_{in}(t) = 10 \sin \omega t$.

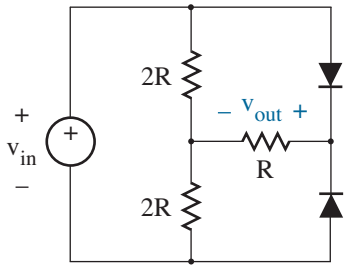


Figure P3.18

3.19 Sketch i vs. v $(-10 \leq v \leq 10$ V) for the circuit of Fig. P3.19.

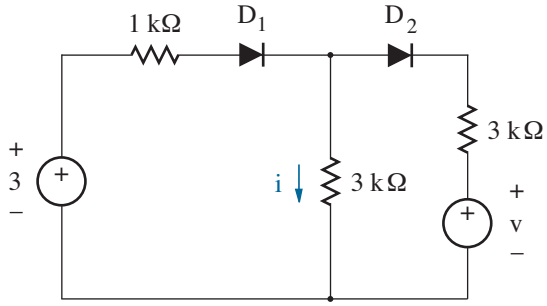


Figure P3.19

3.20 Sketch i vs. v $(-10 \leq v \leq 10$ V) for the circuit of Fig. P3.20.

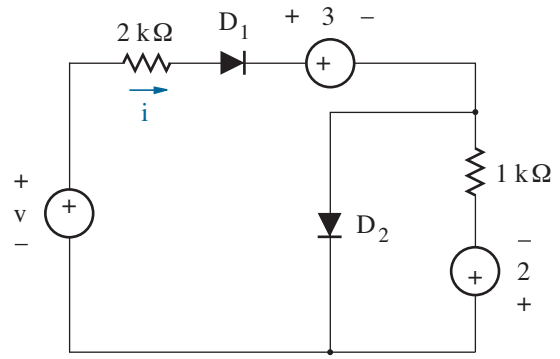


Figure P3.20

3.21 The op-amp in the circuit of Fig. P3.21 is ideal. Determine V_s so that D_1 is at breakpoint, then do the same for D_2.

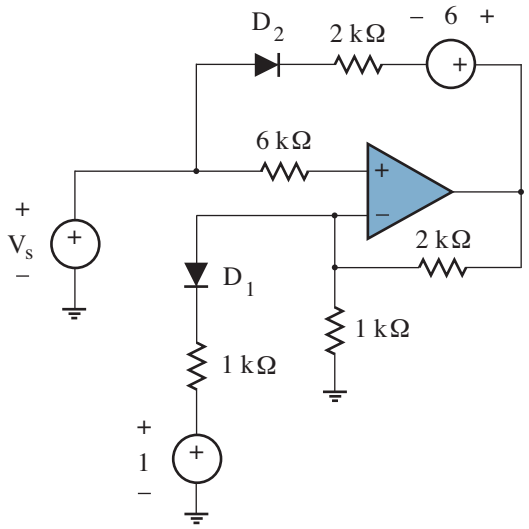


Figure P3.21

3.22 The op-amp in the circuit of Fig. P3.22 is ideal. Sketch the circuit transfer characteristic v_{out} vs. v_{in} $(-10 \leq v_{in} \leq 10$ V).

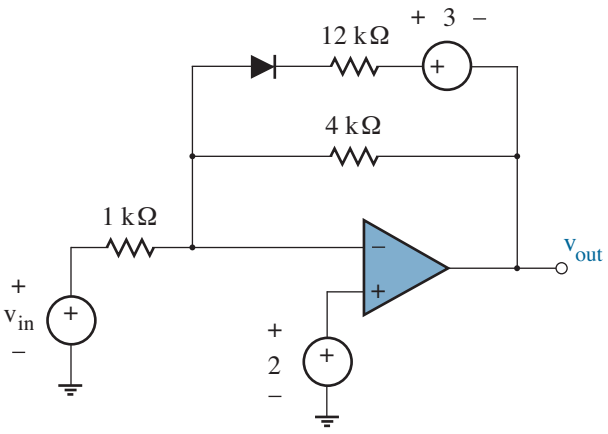


Figure P3.22

Sections 3.2 and 3.3

3.23 Consider the circuit of Fig. P3.23.

(a) Determine a zero-order solution for v.

(b) Determine a first-order solution for v if $V_f = 0.77$ V and $I_r = 5$ μA.

(c) Determine a second-order solution for v if $V_f = 0.75$ V, $R_f = 5$ Ω, $I_r = 4$ μA and $R_r = 1$ MΩ.

(d) Which solution is adequate if the resistors have 5 % tolerance, 0.5 % tolerance?

3.24 Repeat Problem 3.23, but change the orientation of the diode.

3.25 Use an iterative procedure to determine v in the circuit of Fig. 3.23.

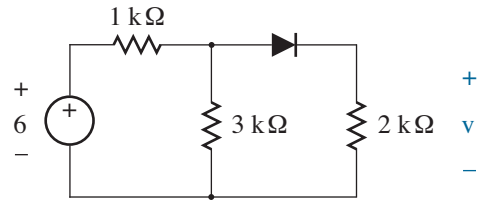


Figure P3.23

3.26 Consider the circuit of Fig. P3.26.

(a) Determine a zero-order solution for v.

(b) Determine a first-order solution for v if $V_f = 0.72$ V and $I_r = 10$ μA.

(c) Determine a second-order solution for v if $V_f = 0.70$ V, $R_f = 8$ Ω, $I_r = 8$ μA and $R_r = 2$ MΩ.

(d) Which solution is adequate if the resistors have 5 % tolerance, 0.5 % tolerance?

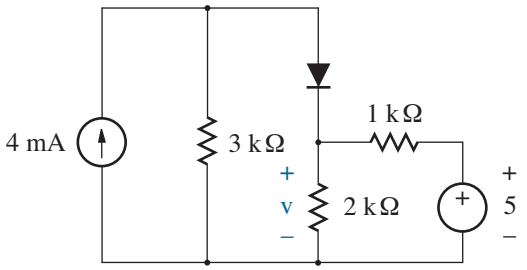


Figure P3.26

3.27 Repeat Problem 3.26, but change the orientation of the diode.

3.28 Use an iterative procedure to determine v in the circuit of Fig. 3.26.

Section 3.4

SPICE Analysis

3.29 Consider the circuit of Fig. 3.5. Use SPICE to verify the results of Fig. 3.6.

3.30 Consider the circuit of Fig. 3.1.

(a) Use SPICE to determine i if the diode features $I_s = 2 \times 10^{-15}$ A.

(b) Repeat part a, but increase R_1 by 2% and decrease R_2 by 2%.

(c) Repeat part a, but decrease R_1 by 2% and increase R_2 by 2%.

3.31

(a) Use SPICE to show the result for Problem 3.8.

(b) Repeat part a, but let the diode feature $I_s = 5 \times 10^{-14}$ A.

3.32

(a) Use SPICE to show the result for Problem 3.12.

(b) Repeat part a, but let the diodes feature $I_s = 5 \times 10^{-14}$ A.

3.33

(a) Use SPICE to show the result for Problem 3.16.

(b) Repeat part a, but let the diodes feature $I_s = 5 \times 10^{-14}$ A.

3.34

(a) Use SPICE to show the result for Problem 3.22.

(b) Repeat part a, but let the diodes feature $I_s = 5 \times 10^{-14}$ A.

3.35 Diodes D_1 and D_2 in the circuit of Fig. P3.35 are characterized by $I_s = 2 \times 10^{-15}$ A, $n = 1.6$ and $I_s = 6 \times 10^{-15}$ A, $n = 1.2$, respectively. Use SPICE to determine R such that $v = 0$.

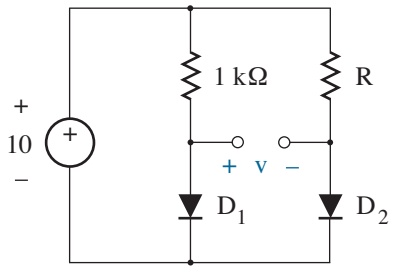


Figure P3.35

3.36 The diodes in the circuit of Fig. P3.36 feature $I_s = 2 \times 10^{-15}$ A and $n = 1.2$. Use SPICE to plot v vs. temperature in the range $-40 \leq T \leq 85$ °C.

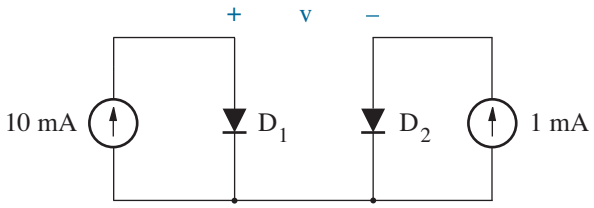


Figure P3.36

3.37 Consider the circuit of Fig. P3.37.

(a) Use SPICE to determine i if the diodes are identical with $I_s = 5 \times 10^{-15}$ A.

(b) Repeat part a, but assume an ambient temperature of 40 °C.

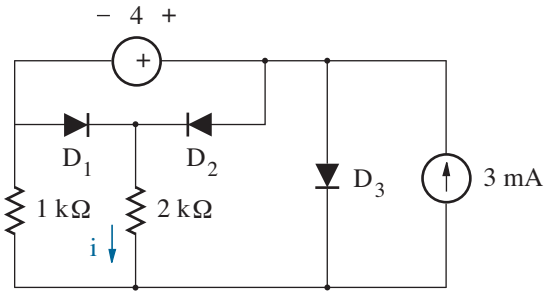


Figure P3.37

3.38 The diodes in the circuit of Fig. P3.38 feature $I_s = 5\times10^{-12}$ A, $n = 1.1$, and $R_s = 1\ \Omega$. The diodes experience breakdown at 50-V reverse bias, and the reverse current is 10 μA at the onset of breakdown. Use SPICE to plot v_{out} subject to $0 \le v_{in} \le 100$ V.

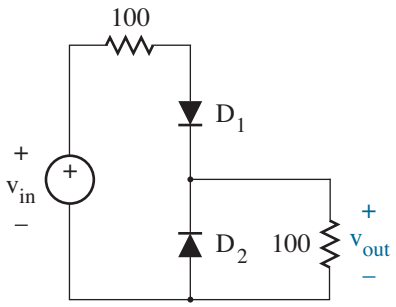


Figure P3.38

Diode Capacitance

3.39 The reverse bias across a pn junction diode is changed from -5 to -10 V. If $V_{bi} = 0.75$ V, by what factor does the depletion capacitance change?

3.40 A symmetric pn junction is *linearly graded*, specifically

$$N_d - N_a = N_o \frac{x}{x_o}$$

for $|x| \le x_o$. The semiconductor doping is constant for $|x| > x_o$.

Show that the depletion capacitance varies according to the relation

$$C_d = C_d(0)\left[1 - \frac{v}{V_{bi}}\right]^{-1/3}.$$

3.41 A pn diode with uniform n- and p-side doping exhibits the following capacitance-voltage data:

Reverse Bias (V)	Capacitance (pF)
0.0	10.00
0.2	9.35
0.4	8.85
0.8	8.12
1.2	7.60
1.6	7.19
2.0	6.87
4.0	5.84
8.0	4.87
12.0	4.35

Determine the SPICE parameters CJO, VJ, and M.

Hint: First, find VJ so that a log-log plot of C vs. $(1 - v/\text{VJ})$ is a straight line.

3.42 The circuit of Fig. P3.42 exploits the voltage-dependent capacitance of a diode to tune a resonator. Inductor L_2 is large ($L_2 \gg L_1$) so that control source V_s is isolated from high-frequency ac signals, and capacitor C_2 is large ($C_2 \gg C_1$) so that dc current does not flow through inductor L_1. Thus, the resonant frequency is approximately $1/\sqrt{L_1 C_1}$. The diode is characterized by $C_d(0) = 500$ pF and $V_{bi} = 0.75$ V. Determine the range of V_s that tunes the resonant circuit over the FM band (88 to 108 MHz).

Note: A diode used as a voltage-dependent capacitor is called a **varactor**.

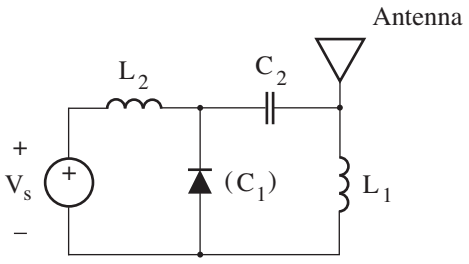
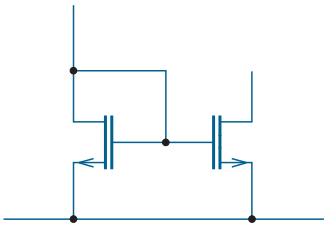


Figure P3.42



Chapter 4

Diode Applications

As electronics engineers, we expect to be rewarded for the implementation of our own ideas as opposed to the analysis of circuits designed by others. Thus, we should be eager to apply the physical principles and analytical skills that have been developed in relation to the diode in Chapters 2 and 3. In this spirit, we prepare to consider several diode circuit categories that will serve as building blocks for design. These include: power conversion, voltage regulation, signal conditioning and clamping, and optoelectronics.

Objectives

After studying this chapter and working through the problems, you should be able to perform each of the following:

- Design a half- or full-wave rectifier circuit with capacitive filtering, and specify the necessary peak voltage and peak current diode ratings (Section 4.1).
- Design a simple buck or boost dc-to-dc converter (Section 4.1).
- Design a Zener diode voltage regulator, taking care to permit worst-case operating conditions (Section 4.2).
- Explain the behavior of a bandgap reference circuit (Section 4.2).
- Outline circuit conditions that promote positive or negative waveform clipping in relation to a particular reference voltage (Section 4.3).
- Outline circuit conditions that promote positive or negative waveform clamping in relation to a particular reference voltage (Section 4.3).
- Describe the features of integrated circuits used for digital or analog optocoupling between incompatible environments (Section 4.4).

4.1 Power Conversion

The necessity of efficient electrical power transfer over long distances and changes of voltage amplitude upon delivery restrains the use of large-scale dc power distribution systems. As a result, we are conditioned to view the ubiquitous 60-Hz ac wall outlet as an all-purpose electrical power source. Yet numerous applications, particularly those with electronic components, require dc voltages for proper function. When batteries are inconvenient—they are often bulky and tend to wear down—we seek a means for deriving dc power from an ac source.

Ac-to-dc conversion usually employs a two-step process:

- **Rectification** alters an ac voltage waveform so that its instantaneous value is never less than zero.
- **Filtering** extracts a dc voltage from a positive waveform by removing unwanted time-varying components.

Rectification and filtering processes are old (and perhaps a little bit boring). Nevertheless, they form the basis for modern power-supply circuits.

Basic Rectification Circuits

The single-diode configuration shown in Fig. 4.1a enjoys a distinction as the simplest rectifier circuit with load R. For analysis, we assume that the diode is ideal, and $v_{in}(t) = \tilde{v} \sin \omega t$. When the instantaneous value of v_{in} is positive, the diode is forward biased, $v = 0$, and $v_{out} = v_{in}$. However, when the instantaneous value of v_{in} is negative, the diode is reverse biased, $i = 0$, and $v_{out} = 0$. The output voltage waveform that appears in Fig. 4.1b consists of half of the input voltage waveform, so the circuit is called a **half-wave rectifier**.

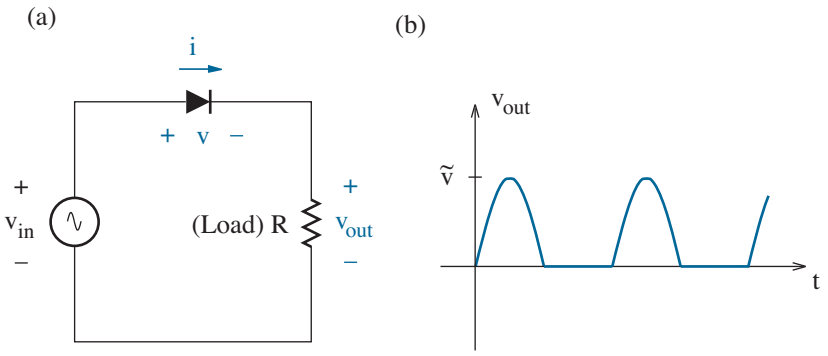


Figure 4.1: Half-wave rectifier circuit and typical output waveform.

Unfortunately, the half-wave rectifier circuit fails to deliver half of the available input power to the load. This inefficiency can be corrected through the use of a **full-wave rectifier** circuit in which the negative portions of the input voltage waveform are converted into positive load voltages.

A common full-wave rectifier circuit consists of four diodes in the "bridge" configuration of one-way current streets shown in Fig. 4.2a. The solid and dashed arrows show the direction of current flow for positive and negative instantaneous v_{in} values, respectively. In either case, current through load resistance R always flows from left to right, and v_{out} is positive with the time dependence shown in Fig. 4.2b subject to $v_{in}(t) = \tilde{v} \sin \omega t$, as before. The time average of the output waveform is twice that of Fig. 4.1b.

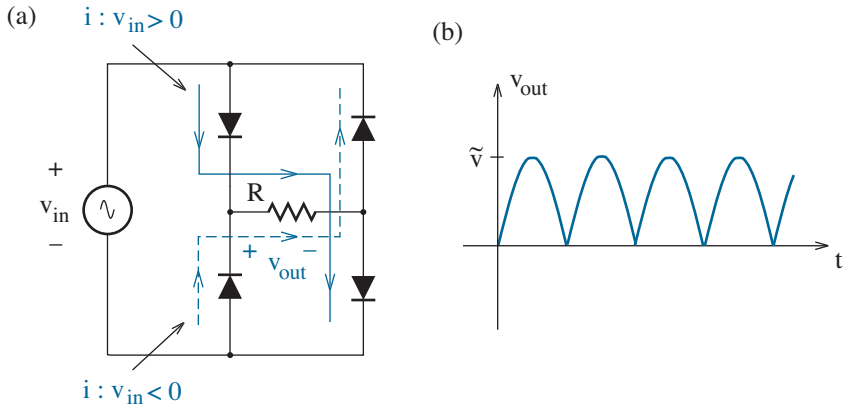


Figure 4.2: Full-wave rectifier circuit and typical output waveform.

Four full-wave-rectifier diodes are often specially packaged (Fig. 4.3).

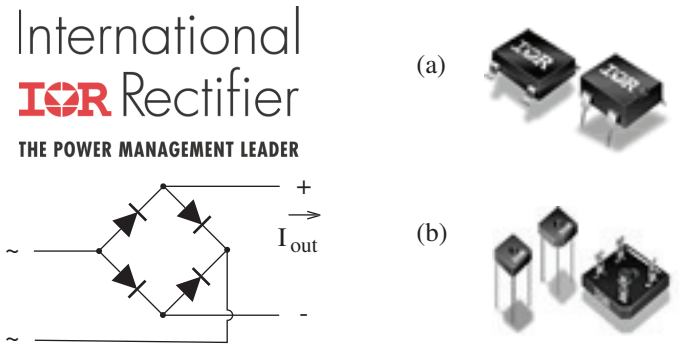


Figure 4.3: Full-wave-rectifier packages: (a) dual in-line; (b) square bridge. Courtesy International Rectifier Corporation.

Practical systems generally require that one side of the load resistance connect to a "ground" potential to guard against electric shock hazards. This can present a problem if one side of the input voltage source is also at ground potential, as for the case of electrical outlets throughout the home. Simultaneous load/source ground connections are allowed in the half-wave rectifier circuit of Fig. 4.1a. But the load must "float" when one side of the source is at ground potential in the full-wave rectifier circuit of Fig. 4.2a. Otherwise, one of the diodes is subject to a short circuit.

The four-diode full-wave rectifier circuit of Fig. 4.4a avoids ground faults by including a transformer with 1:1 turns ratio to provide isolation between the load and the source. The two-diode full-wave rectifier circuit of Fig. 4.4b provides another option if the isolation transformer features a 1:2 turns ratio and a "grounded" center tap. Here, the primary voltage v_{in} appears across both halves of the transformer secondary circuit with the polarities shown, and the diodes ensure that the current through load resistance R always flows towards ground for positive or negative v_{in}.

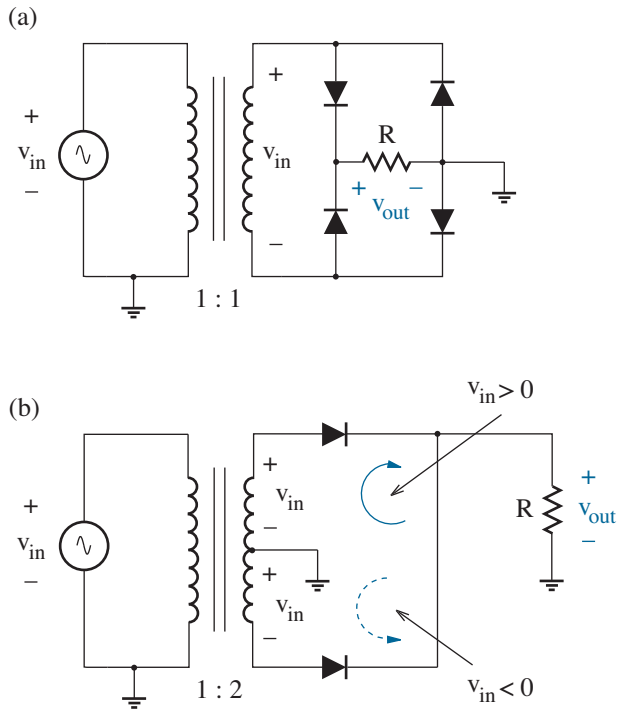


Figure 4.4: Full-wave rectifier circuits with load/source isolation.

Filtering

In each of the preceding rectifier circuits, the output voltage waveform has ac and dc components. If a constant voltage is desired, as for a dc supply, the ac components need to be suppressed through the action of a filter. The suppression can be nearly complete, but never 100 %. In what follows, we consider capacitive, inductive, and hybrid filtering techniques.

Capacitive Filtering

The simplest filter is realized by connecting a capacitor C in parallel with the load resistance R as shown in Fig. 4.5 for the case of a half-wave rectifier. The capacitor receives charge when the diode is subject to forward bias, and charging is nearly instantaneous for small Thevenin source resistance. Thus, v_{out} is quick to track v_{in} to a maximum value as the latter increases. Once v_{in} peaks and then begins to fall, the diode is reverse biased, and the capacitor discharges through R with time constant $\tau = RC$. Nevertheless, v_{out} experiences little change if $\tau \gg T$, where $T = 2\pi/\omega$ is the time period of the input voltage waveform.

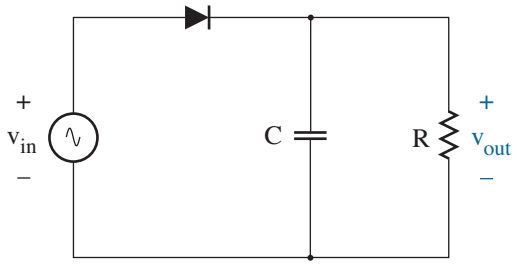


Figure 4.5: RC rectification filtering.

In the limit as $C \to \infty$, v_{out} is *held* at $v_{max} = \tilde{v}$, which is not what one would expect from a low-pass filter. A true low-pass filter applied to the half-wave-rectified waveform of Fig. 4.1 removes ac components and passes the dc time-averaged component: $< v_{in} >= \tilde{v}/\pi$. In following convention, we use the term "filter" somewhat loosely. A "smoothing" circuit might be a better description. The circuit also acts as a **peak detector**.

Rather than worry about semantics, we must develop a design procedure for selecting a suitable filter/smoothing capacitor. A naive approach is to find a capacitor that is extraordinarily large. But large capacitance implies large physical size and high cost. We seek a modest capacitance value based on some form of design compromise.

Figure 4.6 shows filtered and unfiltered ($C = 0$) output voltage waveforms for the circuit of Fig. 4.5, which is subject to initial rest for $t < 0$. The unfiltered output voltage reaches its first two maximum values (v_{max}) at times $t = T/4$ and $t = 5T/4$. The filtered output voltage is similar when $0 < t < T/4$, but it slowly decays to a first minimum value (v_{min}) during the time interval $T/4 < t < 5T/4 - \delta$. Then it is restored to a maximum value after the subsequent time interval of duration δ. We assume that the RC decay constant is large when compared with T so that δ is very small.

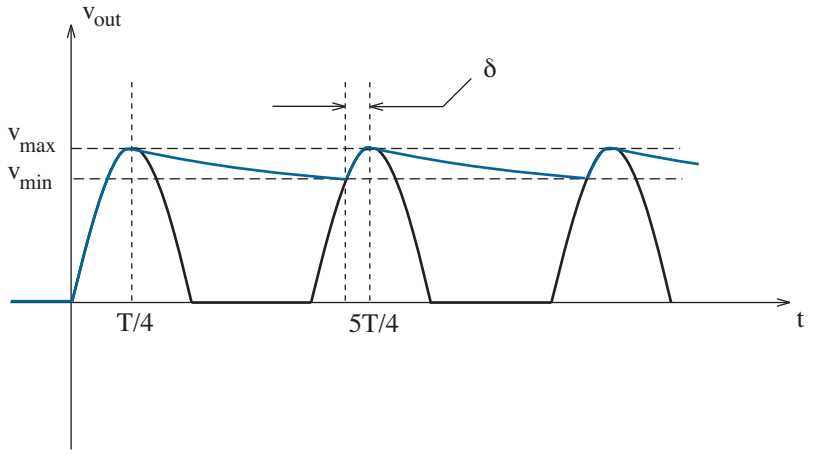


Figure 4.6: Filtered and unfiltered rectifier output voltage waveforms.

Now examine the functional form of the filtered rectifier output voltage. For $T/4 < t < 5T/4 - \delta$,

$$v_{out}(t) = v_{max} \exp\left[\frac{-(t - T/4)}{RC}\right] . \tag{4.1}$$

And assuming that δ is very small,

$$v_{min} = v_{out}(5T/4 - \delta) \approx v_{out}(5T/4) = v_{max} \exp\left[\frac{-T}{RC}\right] . \tag{4.2}$$

We would like to avoid the exponential term in Eq. 4.2, so we expand it into a Taylor series and drop second- and higher-order terms to obtain

$$v_{min} \approx v_{max}\left(1 - \frac{T}{RC}\right) . \tag{4.3}$$

The **ripple voltage** is defined as $v_{max} - v_{min}$. Thus,

$$v_{ripple} = v_{max}\left(\frac{T}{RC}\right) = \tilde{v}\left(\frac{T}{RC}\right) . \tag{4.4}$$

A generic measure of effective waveform filtering is the **ripple factor**, which is defined as

$$r = \frac{\text{rms value of ac waveform component}}{\text{average value of waveform}} . \tag{4.5}$$

However, it is sometimes convenient to specify a **percent ripple factor**, which is simply the ripple factor multiplied by 100.

The filtered waveform in Fig. 4.6 has an average value of $v_{max} - v_{ripple}/2$, and its ac component is very similar to a sawtooth waveform that moves between $\pm v_{ripple}/2$ (relative to the average value) over period T. In turn, we grind through some tedious algebra to find

$$r = \frac{v_{ripple}/2\sqrt{3}}{v_{max} - v_{ripple}/2} \approx \frac{T}{2\sqrt{3}RC} . \tag{4.6}$$

Ripple factors of 0.01 or less are desirable for most applications.

For the case of full-wave rectification, similar derivations yield

$$v_{ripple} \approx v_{max} \left(\frac{T}{2RC}\right) = \tilde{v} \left(\frac{T}{2RC}\right) \tag{4.7}$$

and

$$r \approx \frac{T}{4\sqrt{3}RC} . \tag{4.8}$$

The full-wave-rectified ripple voltage and ripple factor are each lower than their half-wave-rectified counterparts by a factor of two.

Exercise 4.1 A half-wave rectifier circuit with the form of Fig. 4.1 has $R = 1$ kΩ, $C = 2$ mF, and a 60-Hz sinusoidal input with 50-V amplitude. Determine the ripple voltage at the output.

Ans: $v_{ripple} = 0.42$ V

Exercise 4.2 A full-wave rectifier circuit with the form of Fig. 4.2 has $R = 500\ \Omega$ and a 60-Hz 115-V (rms) sinusoidal input. Determine C so that the ripple factor is 0.01 at the output.

Ans: $C = 480\ \mu$F

Exercise 4.3 What percent ripple factor corresponds to a ripple voltage that is 5 % of the maximum output voltage?

Ans: $r = 1.44$ %

It is important to note that the filter capacitor receives charge during short time intervals of duration δ. Thus, very large diode currents may be required if the charge needed to restore the output voltage is substantial. The maximum diode current is given by

$$i_{max} = C \left. \frac{dv_{out}}{dt} \right|_{5T/4\ -\delta} = C\, v_{max}\, \omega \cos\left(\pi/2 - \omega\delta\right) . \tag{4.9}$$

For convenience we rewrite this expression in terms of a **conduction angle**

$$\phi = \omega\, \delta \tag{4.10}$$

so that

$$i_{max} = C\, v_{max}\, \omega \sin\phi\, . \tag{4.11}$$

However (see Problem 4.2),

$$\cos\phi = 1 - \frac{v_{ripple}}{v_{max}}\, . \tag{4.12}$$

Equations 4.11 and 4.12 determine the **peak diode current** for the case of half- or full-wave rectification.

Apart from the peak diode current, it is often necessary to determine the **peak inverse voltage** that a diode must sustain under reverse bias. This is seldom difficult, since the reverse diode current is nearly zero.

Exercise 4.4 Consider the half-wave rectifier circuit of Exercise 4.1. Determine the conduction angle, the peak diode current, and the peak inverse voltage that the diode must sustain.

Ans: $\phi = 7.43\,^\circ$, $i_{max} = 4.9$ A, PIV $= 100$ V

Example 4.1

Design an RC-filtered power supply that can safely deliver up to 250 mA at 50 V from a 60-Hz 115-V (rms) ac source. An acceptable design will feature a ripple voltage of 0.5 V or less.

Solution

The 50-V output specification requires a reduction of the rms source voltage with a transformer. Full-wave rectification minimizes the filter capacitance, a four-diode bridge configuration avoids a transformer center tap, and a fuse gives protection to the ac line source. Finally, a useful safety feature allows the filter capacitor to discharge when the load resistance is disconnected. These considerations suggest the trial circuit of Fig. 4.7.

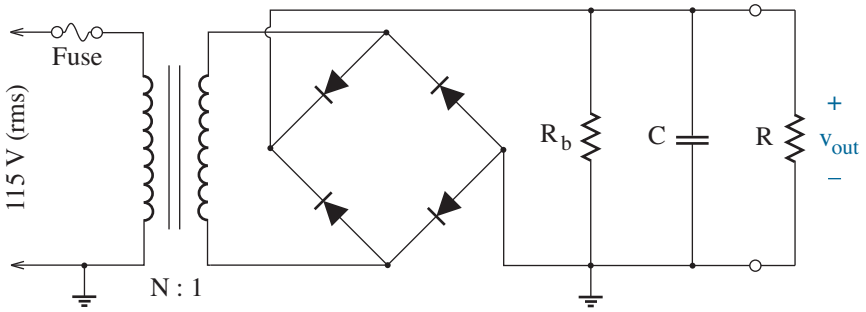


Figure 4.7: RC-filtered power-supply circuit for Example 4.1.

To ensure a design that meets specifications, we consider the effect of a *minimum* load resistance with $R = 50$ V / 250 mA $= 200\ \Omega$. Resistor R_b, which discharges C in the absence of load R, should be made very large. We arbitrarily set $R_b = 1\ \text{M}\Omega$ (and ignore it hereafter). To determine C, we use Eq. 4.7 with $v_{ripple} = 0.5$ V, $v_{max} = 50$ V, and $T = 1/60 = 16.7$ ms. The equivalent resistance "seen" by the capacitor is $R_{eq} = R \parallel R_b \approx R$ —*the capacitor cannot discharge back through the diode rectification matrix*, so the small resistance of the transformer secondary winding does not enter the calculation for C. With these factors in hand, we obtain $C = 4.2$ mF. But we choose $C = 4.7$ mF, since this is a standard capacitance value.

The maximum voltage at the transformer primary is $115\sqrt{2} = 163$ V, and we require a 50-V maximum voltage from the transformer secondary. This is achieved if the transformer has a 3.25:1 turns ratio.

Our last design step determines the diode power-handling specifications. We use Eqs. 4.11 and 4.12 to find the peak diode current, and a consistent diode rating should be greater than 12.5 A. The diodes must also sustain a peak inverse voltage that is larger than 50 V.

How good is our design? Ripple of about 0.5 V is not particularly small. It can be reduced by increasing C. But if we use Eq. 4.7 when choosing C, and if we use this relation in Eq. 4.11 to eliminate C, we find

$$i_{max} = \frac{T\omega}{2R}\,\frac{{v_{max}}^2}{v_{ripple}}\,\sin\phi \approx \frac{\pi\sqrt{2}}{R}\,\frac{{v_{max}}^{3/2}}{{v_{ripple}}^{1/2}}\,. \tag{4.13}$$

So attempts to lower the ripple voltage lead to larger peak diode current. We also note that the peak diode current scales as v_{max} to the 3/2 power. The compromise between ripple voltage and peak diode current becomes less severe for designs that demand lower output voltage.

Inductive and Hybrid Filtering

Although somewhat outdated, inductive filtering is a conditioning process that establishes concepts for switching power supplies considered shortly.

Figure 4.8 shows a prospective filter that uses a **choke** or series-connected inductor to resist changes in load current (and load voltage) for the case of half-wave rectification. The RL combination suggests a true low-pass filter. As in previous discussion, the diode is ideal.

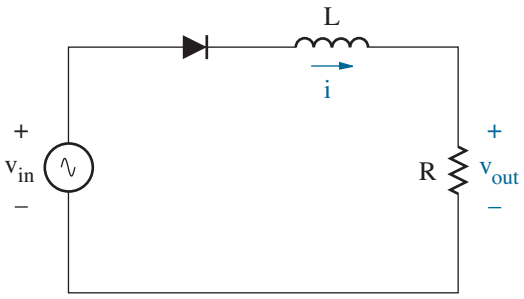


Figure 4.8: RL rectification filtering.

When we considered the RC "filter" circuit of Fig. 4.5, we were able to deduce the time dependence of the filtered output voltage by inspection. Unfortunately, there is no obvious output behavior for the RL filter circuit. Instead, we begin our analysis by assuming $i = 0$ for all $t < 0$ (initial rest), and we let $v_{in}(t) = \tilde{v} \sin \omega t$ for $t \geq 0$. In turn, the diode is forward biased at $t = 0+$, and the subsequent diode current is described by a differential equation of the form

$$v_{in} = L \frac{di}{dt} + iR \,. \tag{4.14}$$

This has the solution

$$i = \frac{\tilde{v}/R}{\sqrt{1 + \tan^2 \psi}} \left[\sin(\omega t - \psi) + \exp\left(\frac{-\omega t}{\tan \psi} \right) \sin \psi \right] , \tag{4.15}$$

where

$$\tan \psi = \frac{\omega L}{R} \,. \tag{4.16}$$

The load voltage is simply $v_{out} = iR$. However, the solution is only valid if $i > 0$. In the alternative, the diode must be reverse biased—it cannot pass negative current—and v_{out} is zero. Figure 4.9 shows the load-voltage waveform (normalized by $\tilde{v}$) for several different values of $\omega L/R$.

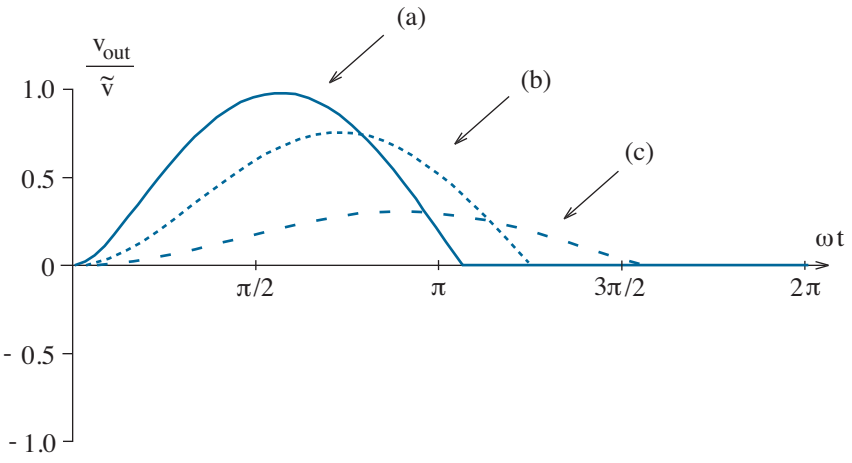


Figure 4.9: Normalized RL-filtered load-voltage waveforms:
(a) $\omega L/R = 0.2$; (b) $\omega L/R = 1.0$; (c) $\omega L/R = 5.0$.

Each curve in Fig. 4.9 commences zero load voltage at the **cutout time** that corresponds to the solution of

$$\sin(\omega t - \psi) + \exp\left(\frac{-\omega t}{\tan\psi}\right) \sin\psi = 0\,, \tag{4.17}$$

and the load voltage remains at zero until the condition $v_{in} > 0$ returns the diode to a forward-bias state. Thereafter, the output waveform is repeated, since it derives from the zero-current initial condition previously engaged. The cutout time interval can be reduced by increasing L, but this is achieved at the expense of voltage amplitude. We conclude that the RL-filtered half-wave rectifier is not a practical circuit.

An inspection of Fig. 4.9 suggests that an RL-filtered full-wave rectifier circuit will never experience cutout, since the inductor current is positive at $\omega t = \pi$ for all values of $\omega L/R$. The diodes in the full-wave rectifier circuits of Fig. 4.4 may alternate between "on" and "off" states, but positive load current is continually delivered. Thus, the circuit can be analyzed assuming sinusoidal steady-state conditions. For this endeavor, we find it helpful to decompose the unfiltered full-wave-rectified waveform into a Fourier series. Specifically,

$$v_{in}(t) = \frac{2\tilde{v}}{\pi} - \frac{4\tilde{v}}{3\pi} \cos 2\omega t + \text{higher-order harmonics}\,. \tag{4.18}$$

The first term is the desired dc output voltage.

When we apply the RL filter transfer characteristic to Eq. 4.18, we find

$$v_{out}(t) \approx \frac{2\tilde{v}}{\pi} - \frac{4\tilde{v}/3\pi}{\sqrt{1+\tan^2\psi}}\cos(2\omega t - \psi)\,, \tag{4.19}$$

where

$$\tan\psi = \frac{2\omega L}{R}\,. \tag{4.20}$$

The ripple voltage is twice the amplitude of the ac component in Eq. 4.19. And the ripple factor is

$$r = \frac{2}{3\sqrt{2}}\left[1+\left(\frac{2\omega L}{R}\right)^2\right]^{-1/2}. \tag{4.21}$$

Subject to $2\omega L/R \gg 1$, this reduces to

$$r \approx \frac{1}{3\sqrt{2}}\,\frac{R}{\omega L}\,. \tag{4.22}$$

Unlike the case of capacitive filtering, the ripple component is smallest when R is small or the load current is large.

Note that the approximate ripple factor is $r \approx 2/3\sqrt{2} = 0.471$ in the worst-case (unfiltered) limit as $L \to 0$. The exact ripple factor is $r = 0.483$. Thus, we justifiably neglected the higher-order harmonic terms in Eq. 4.18.

Example 4.2

Design an RL-filtered power supply that delivers 10 A at 50 V from a 60-Hz 115-V (rms) ac source. An acceptable design will feature a ripple voltage of 0.5 V or less.

Solution

We use the circuit of Fig. 4.8, with v_{in} replaced by a transformer followed by the four-diode bridge configuration. Full-wave rectification is required to avoid cutout problems. The load resistance is $R = 50\ \text{V}/10\ \text{A} = 5\ \Omega$.

In consideration of the dc term in Eq. 4.19, we find $\tilde{v} = 25\pi = 78.5$ V. Thus, the N:1 transformer stepdown factor has $N = 115\sqrt{2}/78.5 = 2.1$. Then by setting the ac amplitude in Eq. 4.19 to one-half the ripple voltage, we find $\tan\psi = 133$. In turn, $L = 0.88$ H (from Eq. 4.20). The maximum current rating for the diodes and the choke must obviously exceed 10 A.

How good is our design? A 0.88-H choke that can accommodate 10 A is physically large and probably expensive—you are challenged to find one. This design will never fly.

Having failed miserably with the design of Example 4.2, we are ready to examine a hybrid filter circuit that features an inductor and a capacitor as shown in Fig. 4.10. We use full-wave rectification so that the inductor current can remain positive with no cutout.

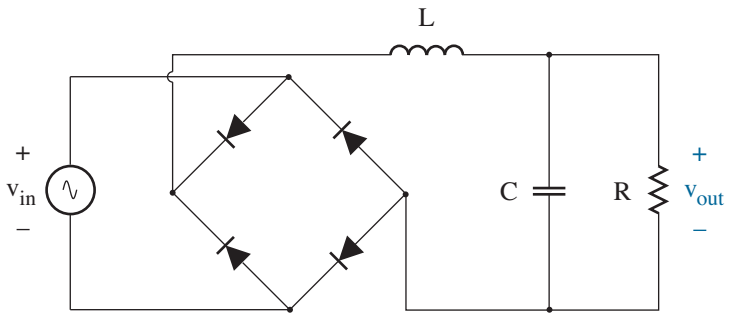


Figure 4.10: LC (hybrid) rectification filtering.

Capacitor C is intended to shunt ac current past R. Thus, we expect impedance values such that

$$\frac{1}{\omega C} \ll R\,. \tag{4.23}$$

The impedance formed by the parallel combination of C and R is now

$$Z = \frac{1}{\jmath\omega C} \parallel R \approx \frac{1}{\jmath\omega C}\,, \tag{4.24}$$

and voltage division between impedances $\jmath\omega L$ and Z yields an approximate *low-pass* filter characteristic of the form

$$H(\jmath\omega) = \frac{v_{out}}{v_{in}} \approx \frac{1}{1-\omega^2 LC}\,. \tag{4.25}$$

Nevertheless, we anticipate $\omega^2 LC \gg 1$ (in order to exploit the full effect of ω^2 in the denominator) so that

$$H(\jmath\omega) \approx \frac{-1}{\omega^2 LC}\,. \tag{4.26}$$

In turn, with $v_{in}(t)$ given by Eq. 4.18,

$$v_{out}(t) \approx \frac{2\tilde{v}}{\pi} + \frac{\tilde{v}}{3\pi\omega^2 LC}\cos 2\omega t\,. \tag{4.27}$$

The ripple voltage is twice the amplitude of the ac component in Eq. 4.27. And the ripple factor is

$$r \approx \frac{1}{6\sqrt{2}\omega^2 LC}\,. \tag{4.28}$$

Caution: The preceding analysis suggests that the LC filter circuit can function without difficulty in the limit as $R \to \infty$. Indeed, this helps the design condition of Eq. 4.23. Yet the LC filter "cuts out" if R is too large, and the analysis fails.

To avoid cutout, the instantaneous inductor current must be *positive* at all times. This current features a time-averaged (dc) component, which must exceed the amplitude of the superimposed fluctuating (ac) component. The former is

$$< i > = \frac{2\tilde{v}}{\pi R} \tag{4.29}$$

(since the dc capacitor current is zero), and the latter is

$$\tilde{i} = C \left. \frac{dv_{out}}{dt} \right|_{\max} = \frac{2\tilde{v}}{3\pi\omega L} \tag{4.30}$$

(subject to Eq. 4.27 and *small* ac current through R). Thus, we require

$$R < 3\omega L \, . \tag{4.31}$$

Practical designs tend to minimize L.

Example 4.3

Design an LC-filtered power supply that delivers 10 A at 50 V from a 60-Hz 115-V (rms) ac source. An acceptable design will feature a ripple voltage of 0.5 V or less.

Solution

We use the circuit of Fig. 4.10 subject to a load resistance $R = 50$ V/10 A $= 5\ \Omega$. With $\tilde{v} = 25\pi = 78.5$ V, the transformer stepdown factor is 2.1:1. In consideration of Eq. 4.31, we design for minimum $L > R/3\omega = 4.4$ mH —large inductors are harder to find than large capacitors—but we accept a standard $L = 4.7$ mH. Then we set the ac amplitude in Eq. 4.27 to one-half the ripple voltage to find $C = 50$ mF. However, we adopt a 68-mF standard. (The capacitance value is consistent with the design condition of Eq. 4.23). The maximum current rating for the diodes and choke must exceed 10 A.

How good is our design? A 4.7-mH choke is more appealing than the 880-mH choke of Example 4.2. The capacitor is large, so perhaps we should increase L and decrease C. However, we can do even better if we demand a higher ac frequency, say 400 Hz. In this case, the revised L and C values are 0.68 mH and 10 mF, respectively. The 400-Hz specification is standard in aircraft (to ensure lightweight circuit components, at least in years past).

DC-to-DC Conversion

Some modern power supplies use switching processes to convert a dc input to a dc output with a different voltage level. The input can be a true dc voltage (as from a battery) or a "rough" dc voltage that is achieved after ac rectification and modest filtering. We consider two basic converter circuits.

Buck Converter

Figure 4.11 shows the **buck converter** circuit in which switch S closes and opens regularly with time period T. The durations of the closed and open intervals are DT and $(1-D)T$, respectively, where D is the duty cycle.

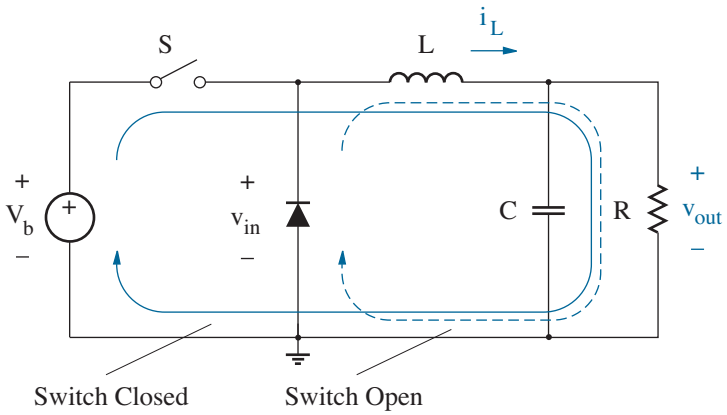


Figure 4.11: Buck converter.

For simplicity, we assume the diode is ideal. When the switch is closed, $v_{in} = V_b$, the diode is reverse biased, and current flows clockwise along an outer path as indicated with the solid arrow. When the switch is opened, the inductor attempts to maintain constant current, and the diode is happy to oblige by becoming forward biased, thereby completing an inner current path as indicated with the dashed arrow. The diode action is made possible by *decreasing* inductor current so that

$$v_{in} = v_{out} + L\,\frac{di_L}{dt} \tag{4.32}$$

is reduced to zero. Indeed, v_{in} would have been reduced to a large negative value had it not been **clamped** to a fixed level—here ground minus zero—that is consistent with the maximum forward diode voltage. If the diode remains in forward bias to allow $i_L > 0$ during the "open" time interval, the circuit functions in the **continuous mode**, and v_{in} exhibits the time dependence shown in Fig. 4.12. The inductor current returns to its peak value at the end of the subsequent "closed" time interval.

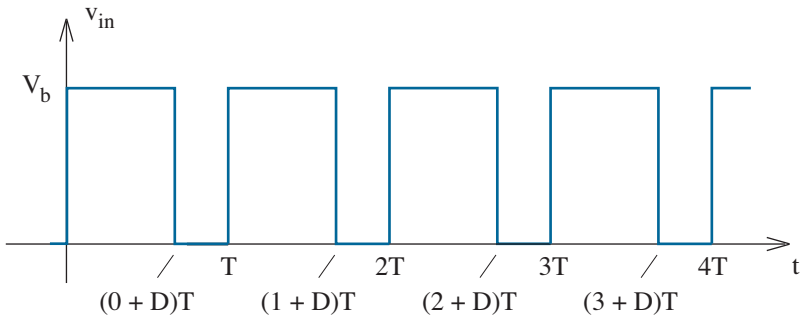


Figure 4.12: v_{in} time dependence for the buck converter shown in Fig. 4.11. The circuit operates in the continuous mode.

In what follows, it will be helpful to separate every voltage or current into time-averaged (dc) and changing (ac) components. For example,

$$v_{in} = < v_{in} > + \Delta v_{in}\,(t)\,, \tag{4.33}$$

where

$$< v_{in} > = \frac{1}{T}\int_0^T v_{in}(t)\,dt\,. \tag{4.34}$$

We also recall two simple rules:

In the steady-state, time-averaged inductor voltages are zero.

In the steady-state, time-averaged capacitor currents are zero.

These rules will save considerable effort.

Determining $< v_{out} >$ is now straightforward. We determine the v_{in} time average of Fig. 4.12 and apply the first rule to find

$$< v_{in} > = DV_b = < v_{out} > + 0\,. \tag{4.35}$$

Then with $0 < D < 1$, *the dc output from the buck converter is less than input voltage* V_b.

Determining $\Delta v_{out}\,(t)$ takes a little more work. However, we note that the inductor tends to block ac current, and, with favorable current division, the capacitor tends to absorb ac current that the inductor fails to suppress. The latter condition implies a design for which

$$RC \gg T\,. \tag{4.36}$$

In turn, we have an LC low-pass filter that makes $\Delta v_{out}\,(t)$ very small. *The buck converter output voltage is almost exclusively dc.*

As a first step toward determining the actual ripple voltage, we need to investigate the time dependence of the inductor current. This effort will also establish the maximum current requirements for the diode and switch, and it will further provide a design condition that ensures circuit operation in the continuous mode.

With the help of Kirchhoff's Current Law and the second rule concerning time-averaged quantities, we find

$$< i_L > = \frac{< v_{out} >}{R} + 0 = \frac{DV_b}{R} . \tag{4.37}$$

Next, we observe that the time-dependent i_L component satisfies

$$L \frac{d}{dt} \Delta i_L(t) = \Delta v_{in}(t) - \Delta v_{out}(t) . \tag{4.38}$$

So with $\Delta v_{in}(t)$ *constant* and *positive* over the time interval $[\,0, DT\,]$, and with $\Delta v_{out}(t)$ nearly zero (subject to the action of a well-designed filter), the inductor current ramps linearly from a minimum value $i_{L,min}$ at $t = 0$ to a maximum value $i_{L,max}$ at $t = DT$. Then with $\Delta v_{in}(t)$ *constant* and *negative* over the time interval $[\,DT, T\,]$, and with $\Delta v_{out}(t)$ remaining zero, the inductor current ramps linearly from $i_{L,max}$ back to $i_{L,min}$ at $t = T$. The cycle repeats during successive time periods.

An up-and-down ramp with the time average given by Eq. 4.37 requires

$$\frac{1}{2} \left(i_{L,max} + i_{L,min} \right) = \frac{DV_b}{R} . \tag{4.39}$$

—Pause to convince yourself that this is true. You may want to consider the geometry that applies to a rough sketch.

We can also integrate Eq. 4.38 over the time interval $[\,0, DT\,]$ for which $\Delta v_{in} = V_b - DV_b$. Specifically,

$$i_{L,max} - i_{L,min} = \frac{V_b(1-D)DT}{L} . \tag{4.40}$$

—Again, pause to verify this result.

In turn, we solve Eqs. 4.39 and 4.40 simultaneously to obtain

$$i_{L,min} = \frac{DV_b}{R} - \frac{V_b(1-D)DT}{2L} \tag{4.41}$$

and

$$i_{L,max} = \frac{DV_b}{R} + \frac{V_b(1-D)DT}{2L} . \tag{4.42}$$

The minimum inductor current defined in Eq. 4.41 must remain positive —negative values are inconsistent with diode orientation. Thus, to ensure circuit operation in the continuous mode,

$$L > \frac{V_b(1-D)DT}{2 \times \text{dc load current}} \,. \tag{4.43}$$

The maximum inductor current defined in Eq. 4.42 is also the maximum current for the diode and switch.

Having determined the inductor current, which is supplied to the RC parallel combination, we are ready to find $\Delta v_{out}(t)$ and the ripple voltage. Provided $RC \gg T$, the result for the latter is (see Problem 4.17)

$$v_{ripple} \approx i_{L,p-p} \left(\frac{1}{\omega C}\right) \frac{(2-D)\pi}{6} \,, \tag{4.44}$$

where $i_{L,p-p}$ is the peak-to-peak inductor current with value

$$i_{L,p-p} = \frac{V_b(1-D)DT}{L} \,, \tag{4.45}$$

and $\omega = 2\pi/T$. The ripple voltage appears to vanish as $\omega C \to \infty$.

We digress for a moment to observe that the buck converter circuit of Fig. 4.11 fails to include parasitic resistance in *series* with the capacitor. This impedance is generally called **equivalent series resistance** or **ESR**. Many capacitor manufacturers specify ESR directly, while others specify ESR indirectly in terms of a **dissipation factor** DF (sometimes $\tan\delta$) that is equal to the ratio of the resistive and reactive components of the capacitor impedance at a particular angular frequency. Specifically,

$$\text{DF} = \omega\, C \times \text{ESR} \,. \tag{4.46}$$

Either way, one finds typical ESR values of about 0.2 Ω or less for high-quality aluminum electrolytic capacitors that are designed for switching power supplies. ESR increases significantly with decreasing temperature. ESR effects are reduced by using a capacitor with a large working voltage —possibly much larger than the intended operating voltage—but with the added increase of physical size and cost. And the effect of total ESR is reduced by using two half-value capacitors in parallel.

Back to the buck converter. When ESR is included in the circuit model, the revised ripple voltage is

$$v_{ripple} \approx i_{L,p-p} \times \text{ESR} \tag{4.47}$$

in the limit as $\omega C \to \infty$. No vanishing here.

Example 4.4

Design a buck converter that delivers 0.5 A at 9 V from a 12-V dc source. An acceptable design will feature a ripple voltage of 0.25 V or less.

Solution

We use the buck circuit of Fig. 4.11. Under the specified output conditions, the load resistance is $R = 9\text{ V} / 0.5\text{ A} = 18\ \Omega$, and the required duty cycle is $D = 9/12 = 0.75$. Our buck converter is expected to join a product line that features 25-kHz switching frequencies (to allow small-size L and C components), so we accept $T = 40\ \mu$s as a system constraint.

To ensure the continuous mode of converter operation, we require $L > 90\ \mu$H (Eq. 4.43). Thus, we feel safe with $L = 120\ \mu$H, a standard inductor value that is available from one of our suppliers. The consistent peak-to-peak and maximum inductor currents are 0.75 A (Eq. 4.45) and 0.875 A (Eq. 4.42), respectively. We demand comfortable 2-A component ratings. And we choose a Schottky diode to minimize associated turn-on voltage.

The XYZ series of aluminum electrolytic capacitors features maximum ESR values of 0.2 Ω. So not wanting to make C too large (and expensive), we opt for the same $1/\omega C$ to obtain $C \approx 33\ \mu$F —a standard XYZ value. The equivalent capacitor impedance is now $0.2\ (1 + \jmath)\ \Omega$, and the ripple voltage is approximately $0.2 \times \sqrt{2}\ \Omega \times 0.75\text{ A} = 0.21\text{ V}$ —just under spec.

How good is our design? We check the converter operation using the following SPICE code, which features a voltage-dependent switch:

```
* Buck Converter Test

Vb      1       0       12
S       1       2       5       0       Switch
Vr      5       0       PULSE   (0 5 0 0 0 30u 40u)
D       0       2       Diode
L       2       3       120u
R-esr   3       4       0.2
C       4       0       33u
R       3       0       18

.model  Switch  VSWITCH  ( RON=0.1, ROFF=100MEG,
+                          VON=4.0, VOFF=1.0 )
.model  Diode   D        ( IS = 0.1n )

.tran   0.1u    2m
.probe

.end
```

Figure 4.13 shows the .probe plot that corresponds to the node-3 voltage. After an initial transient, the output becomes nearly steady at 8.85 V — only slightly less than the intended 9-V level because of the non-zero forward diode voltage. Subject to $4RC \gg L/R$, the transient consists of oscillations with period $2\pi\sqrt{LC}$ (0.4 ms) and an exponentially decaying amplitude with time constant $2RC$ (1.2-ms). The 15-V overshoot shortly after $t = 0$ may cause problems in the absence of special start-up procedures.

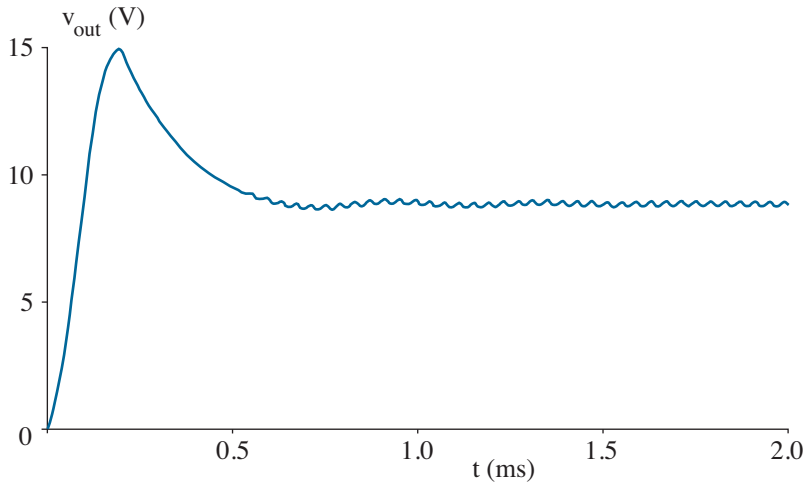


Figure 4.13: Buck converter output voltage for the design of Example 4.4.

To determine the ripple voltage, one needs to zoom in on a portion of the output voltage near the end of the 2-ms time interval. The finding is $v_{ripple} = 0.18$ V, which is in good agreement with our hand calculations. Ripple is reduced with smaller ESR or smaller $i_{L,p-p}$ (through larger L).

A separate .probe plot (not shown) confirms the triangular waveform for the inductor current after transients have settled. The same type of waveform is also observed across R-esr, the parasitic capacitor resistance. In turn, we expect power dissipation within the capacitor, which may cause overheating and failure. The rms inductor current has value

$$i_{L,rms} = \frac{i_{L,p-p}}{2\sqrt{3}} , \tag{4.48}$$

or 0.217 A for this design, so the power dissipation is a meager 9.4 mW. Nevertheless, we are wise to worry about lifetime-limiting thermal effects.

Of course, the real test of our buck converter comes when we build it. For this effort, we require the help of a colleague who has read Chapter 5 (or Chapter 6). The controlled switch is beyond our talent for now.

Boost Converter

Figure 4.14 shows the **boost converter**, which has the same components as the buck converter, but with different positions. The durations of the closed and open switching intervals are DT and $(1-D)T$, respectively, where D is the duty cycle and T is the time period.

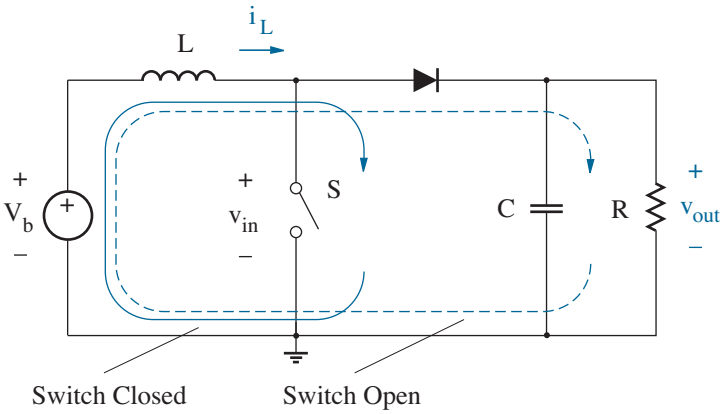


Figure 4.14: Boost converter.

For simplicity, we assume an ideal diode. When the switch is closed, $v_{in} = 0$, and the inductor current increases from a minimum level along a short inner path as indicated with the solid arrow. Meanwhile, the diode is reverse biased so that v_{out} remains nearly constant—any deviation from a previously established positive steady-state value requires the capacitor to discharge through load R, a process that is relatively slow if $RC \gg T$ (just like the smoothing observed in the RC-filtered half-wave rectifier). When the switch is opened, the inductor attempts to maintain continuous current along an outer path as indicated with the dashed arrow. This is not difficult if $v_{out} < V_b$ —the diode assumes forward bias while i_L increases (albeit more slowly). But even if $v_{out} > V_b$, there is the prospect to put the diode in forward bias—the inductor current *decreases* in this case, thereby generating a positive $-L\, di_L/dt$ **flyback voltage** so that

$$v_{in} = V_b - L\,\frac{di_L}{dt} \tag{4.49}$$

is sufficiently large to reach the diode breakpoint. If $i_L > 0$ during the "open" time interval, v_{in} is clamped to $< v_{out} >$, and the circuit functions in the continuous mode with the v_{in} time dependence shown in Fig. 4.15. Under steady-state, inductor energy that is gained (through increasing i_L) balances capacitor energy that is lost (through decreasing v_{out}) during the subsequent "closed" time interval.

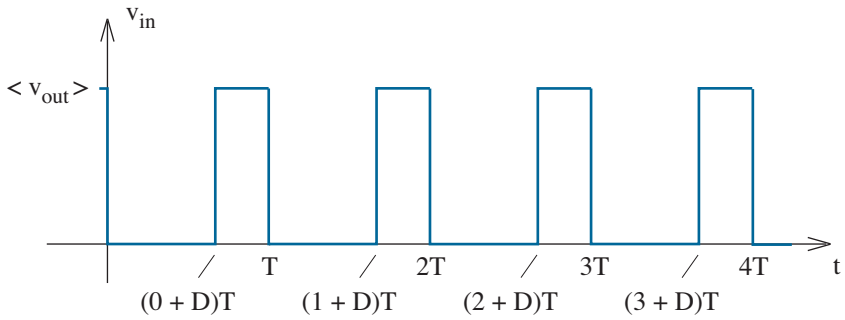


Figure 4.15: v_{in} time dependence for the boost converter shown in Fig. 4.14. The circuit operates in the continuous mode.

To determine the time-averaged (dc) output voltage, we recall that the steady-state time-averaged inductor voltage is zero. Thus,

$$< v_{in} > = (1 - D) < v_{out} > = V_b - \ 0 \, , \tag{4.50}$$

or

$$< v_{out} > = \frac{V_b}{1 - D} \, . \tag{4.51}$$

With $0 < D < 1$, *the dc output from the boost converter is greater than input voltage* V_b. This is a consequence of the flyback voltage.

Apart from a different discharge period (DT), the analysis of the boost-converter ripple voltage is the same as that for the RC-filtered half-wave rectifier (if we neglect the influence from parasitic ESR). The result is

$$v_{ripple} = \frac{V_b}{1 - D} \left(\frac{DT}{RC} \right) \, . \tag{4.52}$$

Equations 4.51 and 4.52 are key to a successful boost converter design. Nevertheless, we still need to know the maximum and minimum inductor currents to allow for satisfactory component ratings and to determine a design condition that ensures circuit operation in the continuous mode. Assuming steady-state conditions, we integrate Eq. 4.49 over time interval $[\, 0, DT \,]$ to find

$$i_{L,max} - i_{L,min} = \frac{V_b DT}{L} \, . \tag{4.53}$$

Then we equate the inductor energy that is gained during this time period to the capacitor energy that is lost (see Problem 4.21). In turn,

$$\frac{1}{2} \left(\, i_{L,max} + i_{L,min} \, \right) = \frac{V_b}{R(1 - D)^2} \, . \tag{4.54}$$

Finally, we solve Eqs. 4.53 and 4.54 simultaneously to obtain

$$i_{L,min} = \frac{V_b}{R(1-D)^2} - \frac{V_b DT}{2L} \tag{4.55}$$

and

$$i_{L,max} = \frac{V_b}{R(1-D)^2} + \frac{V_b DT}{2L} \,. \tag{4.56}$$

The minimum current defined in Eq. 4.55 must remain positive to ensure circuit operation in the continuous mode. Thus,

$$L > \frac{V_b(1-D)DT}{2 \times \text{dc load current}} \,, \tag{4.57}$$

which happens to be the same design condition used for the buck converter. The maximum inductor current defined in Eq. 4.56 is also the maximum current for the diode and switch.

Example 4.5

Design a boost converter that delivers 0.3 A at 15 V from a 6-V dc source. An acceptable design will feature a ripple voltage of 0.1 V or less.

Solution

We use the boost circuit of Fig. 4.14. Under the specified output conditions, the load resistance is $R = 15$ V / 0.3 A $= 50\ \Omega$, and the required duty cycle is $D = 1 - 6/15 = 0.6$. As in Example 4.4, the converter is expected to join a product line that features 25-kHz switching frequencies, so we accept $T = 40\ \mu$s as a system constraint.

To ensure the continuous mode of converter operation, we require $L >$ 96 μH (Eq. 4.57). In turn, we select $L = 120\ \mu$H, a supplier's standard. The consistent maximum inductor current is 1.35 A (Eq. 4.56), so 2-A component ratings should be acceptable.

Our last step is to choose a capacitor value that achieves the desired 0.1-V ripple voltage. With this goal, we use Eq. 4.52 to find $C = 72\ \mu$F. However, we let $C = 82\ \mu$F, since we are restricted to a set of standard component values. Perhaps the low-cost ABC capacitor series will do.

How good is our design? We simulate the converter operation using a SPICE code similar to that for Example 4.4. After a fascinating transient (see Problem 4.22), the output settles to 15 V minus one forward diode turn-on voltage, and the ripple is right on target at 85 mV. But when we include 0.5-Ω ESR with our cheap capacitor, the output has 0.59-V spikes. So much for neglecting ESR. This design needs ESR less than about 0.1 Ω.

Converter Control

Modern dc-to-dc converters used in cellphones, laptop computers, and other portable electronic equipment are usually paired with integrated circuits that regulate an output voltage through a particular feedback mechanism. The switch is sometimes included in the integrated package for applications with modest current demands, but the inductor is invariably external.

Figure 4.16 shows a simple **pulse-width modulation** (PWM) process for buck converter control. A scaled output voltage connects at the FB pin by means of a voltage divider, and the reduced output is subtracted from an internal reference V_{ref}. In turn, the difference voltage is amplified by a gain factor K to produce an "error" voltage

$$v_e = K(V_{ref} - f\, v_{out})\,, \tag{4.58}$$

where $f = R_1/(R_1 + R_f)$. The error voltage and a triangular waveform v_r with peak-to-peak amplitude V_m and period T are applied to a comparator that causes switch S to connect LX to V_o over time intervals when $v_e > v_r$. Thus, in consideration of the geometry of comparison in the figure inset, abnormally large or small output voltages tend to decrease or increase the duty cycle, respectively. The corrective action stabilizes the output.

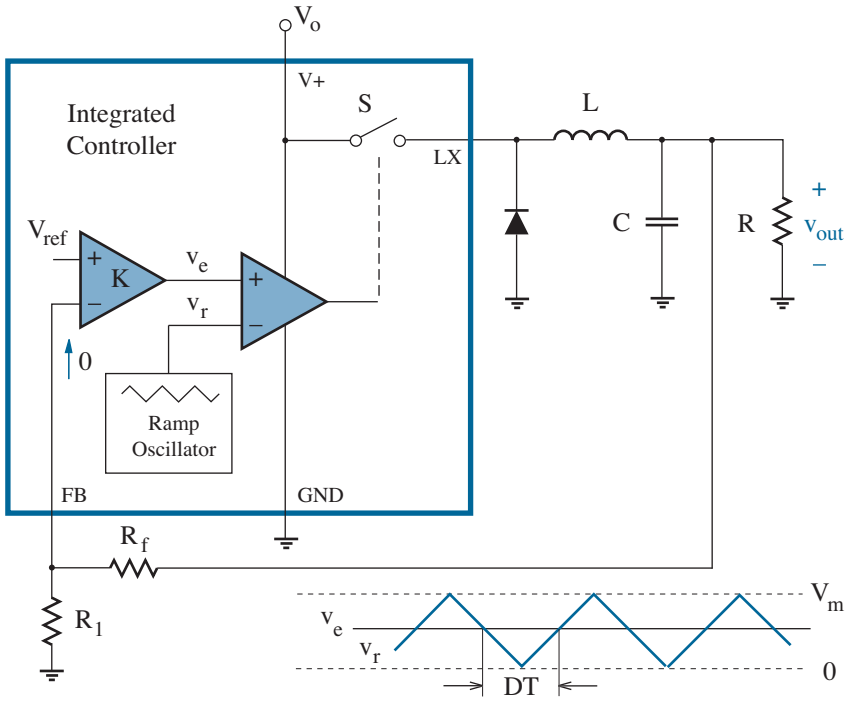


Figure 4.16: Pulse-width-modulation circuit for buck converter control.

The preceding control process is called "voltage mode" because the duty cycle is directly related to the output voltage. An alternative current-mode control process monitors the inductor current by measuring the voltage v_s across a small resistance in series with the switch. The result is compared with the error voltage defined in Eq. 4.58, and the switch is made to close when $v_s > v_e$ as shown in Fig. 4.17. Thus, the duty cycle reduces indirectly when an overly large output voltage decreases v_e.

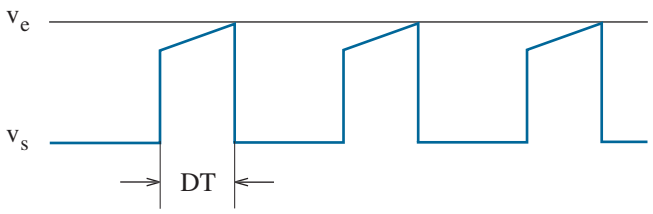


Figure 4.17: Current-mode control for a buck converter.

Current-mode control has several benefits: It prevents large current that can potentially damage the circuit. It avoids inductor **saturation** effects that limit capability for energy storage (see Problem 4.26). And it supports favorable dynamic feedback performance (a topic reserved for Chapter 12). A downside concerns the influence of abrupt inductor-current transients, especially when duty cycles exceed 0.5 (see Problem 4.27).

The **efficiency** of a dc-to-dc converter is defined as the ratio of power consumed by the load and power provided by the dc supply. High efficiency is required for portable-power applications so as to maximize battery life. Thus, the power dissipated by a diode with modest forward voltage is often reduced using a **synchronous rectifier** in the form of a switch (Fig. 4.18) that opens and closes at appropriate times as determined by the controller. The actions of switch S and the synchronous rectifier are both achieved at the expense of energy, so controller efficiencies are often improved by deliberately skipping switching cycles when load currents are very small. Unfortunately, the irregularities of skip-mode switching "noise" can cause problems in certain types of peripheral circuits.

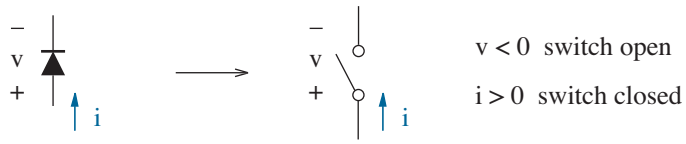


Figure 4.18: Synchronous rectifier for improved converter efficiency.

4.2 Voltage Regulation

In the preceding section, we examined a class of circuits that produce dc voltages with a small, but never fully suppressed superimposed ac ripple. We now consider the process of **voltage regulation** in which post-filter ripple is further reduced (possibly from levels that otherwise seem high).

Suppose you require a circuit element that has the appearance of a 5.6-V source over a limited range of current. You could use a 5.6-V battery. You could use a well-filtered 5.6-V power supply. And you could even use the set of eight series-connected forward-biased diodes shown in Fig. 4.19. Recall that the first-order forward-bias diode model consists of a voltage source with value $V_f \approx 0.7$ V when the diode current is of the order of a few mA. So eight first-order diodes with shared current contribute equally to yield a 5.6-V total drop *that is nearly independent of R and source* V_s. The circuit of Fig. 4.19 may seem peculiar. However, we will later see that integrated circuits occasionally feature one or two forward-biased diodes to provide current-insensitive references of about 0.7 or 1.4 V, respectively.

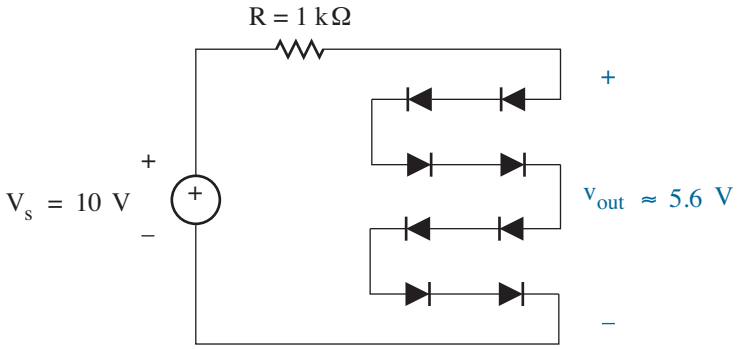


Figure 4.19: Eight-diode circuit that achieves a 5.6-V output.

There is yet another way. You could use a single reverse-biased diode that features a relatively small but carefully "designed" breakdown voltage. This device is called a **Zener diode** (in recognition of one of two physical mechanisms that can induce reverse breakdown).[1]

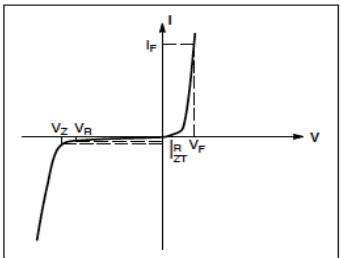
Figure 4.20 shows the current-voltage behavior of a typical Zener diode. Manufacturer's data that apply to different points along the reverse portion of the characteristic curve are listed for specific Zener diodes, and some of the electrical test conditions are described in the footnotes of the data sheet. A Zener diode with 5.6-V breakdown is the MMSZ5232.

[1] In contrast to avalanche breakdown via carrier multiplication, Zener breakdown occurs when mid-junction electric field strength is sufficient to tear apart covalent bonds.

MMSZ52xxxT1G Series, SZMMSZ52xxxT1G Series

ELECTRICAL CHARACTERISTICS (T_A = 25°C unless otherwise noted, V_F = 0.95 V Max. @ I_F = 10 mA)

Symbol	Parameter
V_Z	Reverse Zener Voltage @ I_{ZT}
I_{ZT}	Reverse Current
Z_{ZT}	Maximum Zener Impedance @ I_{ZT}
I_{ZK}	Reverse Current
Z_{ZK}	Maximum Zener Impedance @ I_{ZK}
I_R	Reverse Leakage Current @ V_R
V_R	Reverse Voltage
I_F	Forward Current
V_F	Forward Voltage @ I_F



Zener Voltage Regulator

5% TOLERANCE FG ELECTRICAL CHARACTERISTICS (T_A = 25°C unless otherwise noted, V_F = 0.9 V Max. @ I_F = 10 mA)

Device*	Device Marking	Zener Voltage (Notes 3 and 4) V_Z (Volts) Min	Nom	Max	@ I_{ZT} mA	Zener Impedance (Note 5) Z_{ZT} @ I_{ZT} Ω	Z_{ZK} @ I_{ZK} Ω	mA	Leakage Current I_R @ V_R µA	Volts
MMSZ5221BT1G	C1	2.28	2.4	2.52	20	30	1200	0.25	100	1
MMSZ5222BT1G	C2	2.38	2.5	2.63	20	30	1250	0.25	100	1
MMSZ5223BT1G	C3	2.57	2.7	2.84	20	30	1300	0.25	75	1
MMSZ5224BT1G	C4	2.66	2.8	2.94	20	30	1400	0.25	75	1
MMSZ5225BT1G	C5	2.85	3.0	3.15	20	29	1600	0.25	50	1
MMSZ5226BT1G	D1	3.14	3.3	3.47	20	28	1600	0.25	25	1
MMSZ5227BT1G	D2	3.42	3.6	3.78	20	24	1700	0.25	15	1
MMSZ5228BT1G	D3	3.71	3.9	4.10	20	23	1900	0.25	10	1
MMSZ5229BT1G	D4	4.09	4.3	4.52	20	22	2000	0.25	5	1
MMSZ5230BT1G	D5	4.47	4.7	4.94	20	19	1900	0.25	5	2
MMSZ5231BT1G	E1	4.85	5.1	5.36	20	17	1600	0.25	5	2
MMSZ5232BT1G	E2	5.32	5.6	5.88	20	11	1600	0.25	5	3
MMSZ5233BT1G	E3	5.70	6.0	6.30	20	7	1600	0.25	5	3.5
MMSZ5234BT1G	E4	5.89	6.2	6.51	20	7	1000	0.25	5	4
MMSZ5235BT1G	E5	6.46	6.8	7.14	20	5	750	0.25	3	5
MMSZ5236BT1G	F1	7.13	7.5	7.88	20	6	500	0.25	3	6
MMSZ5237BT1G	F2	7.79	8.2	8.61	20	8	500	0.25	3	6.5
MMSZ5238BT1G	F3	8.27	8.7	9.14	20	8	600	0.25	3	6.5
MMSZ5239BT1G	F4	8.65	9.1	9.56	20	10	600	0.25	3	7
MMSZ5240BT1G	F5	9.50	10	10.50	20	17	600	0.25	3	8
MMSZ5241BT1G	H1	10.45	11	11.55	20	22	600	0.25	2	8.4
MMSZ5242BT1G/T3G	H2	11.40	12	12.60	20	30	600	0.25	1	9.1
MMSZ5243BT1G	H3	12.35	13	13.65	9.5	13	600	0.25	0.5	9.9
MMSZ5244BT1G	H4	13.30	14	14.70	9.0	15	600	0.25	0.1	10
MMSZ5245BT1G	H5	14.25	15	15.75	8.5	16	600	0.25	0.1	11
MMSZ5246BT1G	J1	15.20	16	16.80	7.8	17	600	0.25	0.1	12
MMSZ5247BT1G	J2	16.15	17	17.85	7.4	19	600	0.25	0.1	13
MMSZ5248BT1G	J3	17.10	18	18.90	7.0	21	600	0.25	0.1	14
MMSZ5249BT1G	J4	18.05	19	19.95	6.6	23	600	0.25	0.1	14
MMSZ5250BT1G	J5	19.00	20	21.00	6.2	25	600	0.25	0.1	15
MMSZ5251BT1G	K1	20.90	22	23.10	5.6	29	600	0.25	0.1	17
MMSZ5252BT1G	K2	22.80	24	25.20	5.2	33	600	0.25	0.1	18
MMSZ5253BT1G	K3	23.75	25	26.25	5.0	35	600	0.25	0.1	19
MMSZ5254BT1G	K4	25.65	27	28.35	4.6	41	600	0.25	0.1	21
MMSZ5255BT1G	K5	26.60	28	29.40	4.5	44	600	0.25	0.1	21
MMSZ5256BT1G	M1	28.50	30	31.50	4.2	49	600	0.25	0.1	23
MMSZ5257BT1G	M2	31.35	33	34.65	3.8	58	700	0.25	0.1	25
MMSZ5258BT1G/T3G	M3	34.20	36	37.80	3.4	70	700	0.25	0.1	27
MMSZ5259BT1G	M4	37.05	39	40.95	3.2	80	800	0.25	0.1	30
MMSZ5260BT1G	M5	40.85	43	45.15	3.0	93	900	0.25	0.1	33
MMSZ5261BT1G	N1	44.65	47	49.35	2.7	105	1000	0.25	0.1	36
MMSZ5262BT1G	N2	48.45	51	53.55	2.5	125	1100	0.25	0.1	39
MMSZ5263BT1G	N3	53.20	56	58.80	2.2	150	1300	0.25	0.1	43
MMSZ5264BT1G	N4	57.00	60	63.00	2.1	170	1400	0.25	0.1	46
MMSZ5265BT1G	N5	58.90	62	65.10	2.0	185	1400	0.25	0.1	47
MMSZ5266BT1G	P1	64.60	68	71.40	1.8	230	1600	0.25	0.1	52
MMSZ5267BT1G	P2	71.25	75	78.75	1.7	270	1700	0.25	0.1	56
MMSZ5268BT1G	P3	77.90	82	86.10	1.5	330	2000	0.25	0.1	62
MMSZ5269BT1G	P4	82.65	87	91.35	1.4	370	2200	0.25	0.1	68
MMSZ5270BT1G	P5	86.45	91	95.55	1.4	400	2300	0.25	0.1	69
MMSZ5272BT1G	R2	104.5	110	115.5	1.1	750	3000	0.25	0.1	84

3. "B" Suffix Type numbers shown have a standard tolerance of ±5% on the nominal Zener voltages.
4. Nominal Zener voltage is measured with the device junction in thermal equilibrium at T_L = 30°C ±1°C.
5. Z_{ZT} and Z_{ZK} are measured by dividing the AC voltage drop across the device by the ac current applied. The specified limits are for $I_{Z(AC)}$ = 0.1 $I_{Z(dc)}$ with the AC frequency = 1 kHz.

*Include SZ-prefix devices where applicable.

Figure 4.20: Manufacturer's data for a series of 500-mW Zener diodes. Used with permission from SCILLC dba ON Semiconductor.

Consider the circuit of Fig. 4.21a. We wish to find the voltage v across the reverse-connected Zener diode. (Note the Zener diode symbol.)

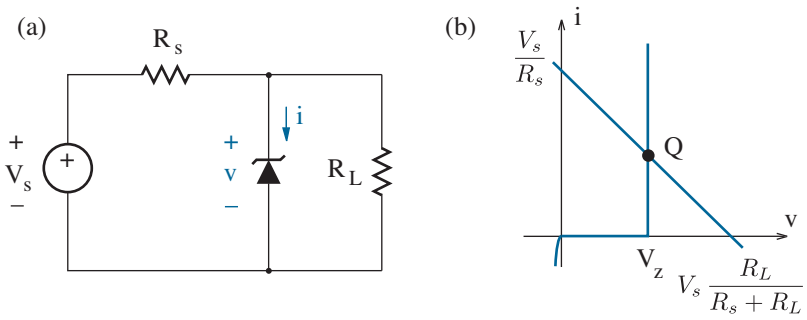


Figure 4.21: Simple Zener diode circuit and its graphical solution.

We determine a graphical solution by plotting the characteristic curve of a reverse-connected Zener diode and a load line as shown in Fig. 4.21b, where variables i and v denote *reverse* current and voltage, respectively. The resulting Q-point features a voltage component that is fixed at V_z, the magnitude of the Zener breakdown voltage, provided

$$V_s \left(\frac{R_L}{R_s + R_L} \right) > V_z \, . \tag{4.59}$$

Thus, the circuit functions as a **voltage regulator**.

A non-graphical form of circuit solution is particularly straightforward. We assume that the Zener diode is deliberately reverse biased beyond its breakdown value—why would one use it otherwise? Then we replace the Zener diode with a first-order breakdown model that consists of a voltage source of value V_z. In turn, $v = V_z$. But before we accept this easy solution, we need to make sure that the reverse Zener-diode current is greater than the reverse current corresponding to the "knee" in the characteristic curve, a special current denoted by i_{ZT} in Fig. 4.20.[2] So we check our result by formulating a Zener breakdown requirement:

$$i = \frac{V_s - V_z}{R_s} - \frac{V_z}{R_L} > i_{ZT} \, . \tag{4.60}$$

If the inequality holds true, the "breakdown" solution is acceptable. If not, a different diode model is necessary. Most likely, a zero-order reverse-bias model that consists of an open circuit would be acceptable.

A slightly better solution is obtained using a second-order breakdown model that consists of a voltage source of value V_z and a series resistance of value R_z. On the data sheet of Fig. 4.20, R_z is denoted as Z_{ZT}.

[2] A separate set of graphs on the MMSZ52-series data sheet show that Zener reverse voltages are nearly constant for reverse currents as low as 0.01 mA. The i_{ZT} values specified in Fig. 4.20 are conservatively large. A practical i_{ZT} value is typically 1 mA.

Exercise 4.5 Show that Eqs. 4.59 and 4.60 are equivalent if $i_{ZT} = 0$.

Exercise 4.6 A voltage regulator circuit with the form of Fig. 4.21 has $V_s = 15$ V, $R_s = 100\ \Omega$, $R_L = 500\ \Omega$, and a MMSZ5240 Zener diode. Determine the reverse current through the Zener diode.

Ans: 30 mA

In what follows, we examine an antiquated Zener-diode regulator design to establish vulnerabilities that lead to a modern perspective.

Example 4.6

Design a Zener-regulated power supply that produces a steady 15-V output from a 15-V (rms) 60-Hz ac source that is full-wave rectified and RC filtered to within a 10-% ripple voltage. The minimum load resistance is 1 kΩ.

Solution

The trial circuit of Fig. 4.22 features a four-diode full-wave rectifier, and the circuit to the right of the filter capacitor is similar to that in Fig. 4.21. For now, we assume capacitor C can be chosen so that $21.2 > V_s > 19.1$ V (in agreement with the specified source conditions).

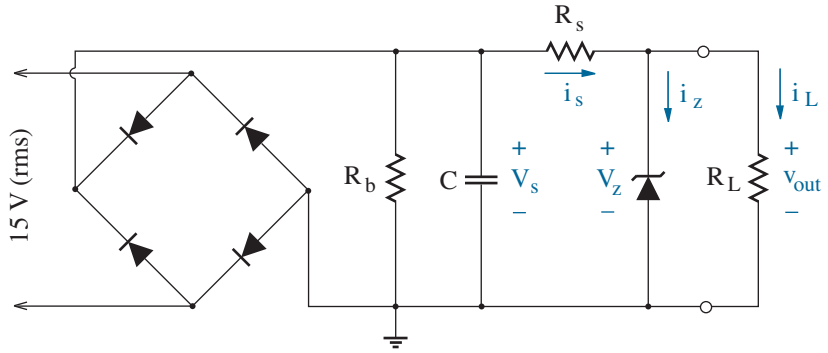


Figure 4.22: Regulated power-supply circuit for Example 4.6.

We begin by selecting a Zener diode with $V_z = 15$ V, since $v_{out} = V_z$. If component selection is limited to the Zener diodes described in Fig. 4.20, the MMSZ5245 is the proper choice. The data sheet shows that the V_z value for a randomly chosen diode can vary between 14.25 V and 15.75 V. Nevertheless, we assume that the MMSZ5245 Zener diode can be selected from a large group and pretested to certify an actual 15-V V_z specification.

In order to ensure breakdown, the Zener diode must be "fed" beyond a minimum reverse knee current, or $i_z > i_{ZT}$. As noted previously, practical i_{ZT} values are typically of the order of 1 mA. Thus, we choose

$$i_{z,min} = 1 \text{ mA} . \tag{4.61}$$

Note that i_z is equal to the *difference* between two widely varying currents: i_s (due to a fluctuating source voltage) and i_L (due to an uncertain R_L). As such, i_z is a particularly sensitive current. Designs that make heroic attempts to force precarious minimum currents are often unreliable.

What conditions yield minimum i_z? One condition is minimum V_s. With V_z fixed, this leads to minimum source current (i_s) that is supplied from the input. The other condition is minimum R_L —the load shares limited source current with the Zener diode, and a minimum value leads to *maximum* current demand. Thus,

$$i_{z,min} = \frac{V_{s,min} - V_z}{R_s} - \frac{V_z}{R_{L,min}} . \tag{4.62}$$

We solve for R_s to find $R_s = 256\ \Omega$, but we choose a standard $R_s = 220\ \Omega$. Larger R_s values starve the Zener diode of current under worst-case load conditions, and smaller R_s values lead to wasteful power dissipation.

A similar argument applied to $i_{z,max}$ yields the expression

$$i_{z,max} = \frac{V_{s,max} - V_z}{R_s} - \frac{V_z}{R_{L,max}} . \tag{4.63}$$

So with R_s already known and $R_{L,max} \to \infty$, we obtain $i_{z,max} = 28.2$ mA. The maximum instantaneous power absorbed by the Zener diode is

$$P_z = i_{z,max} V_z = 423 \text{ mW} , \tag{4.64}$$

and this is less than the 500-mW maximum power rating that is specified for Fig. 4.20. (A realistic regulated power-supply design would probably expect significantly lower load resistance. Consequently, one would require smaller R_s and a higher Zener-diode power rating. The Zener diodes of Fig. 4.20 are intended for low-power applications.)

The circuit is completed using the RC-filter design procedures outlined in Example 4.1, and we do not repeat them here. However, we note that when calculating C, the equivalent resistance "seen" by the capacitor is $R = R_s \parallel R_b \approx R_s$ (if $R_b = 1\ \text{M}\Omega$). This assumes a first-order Zener-diode breakdown model consisting of a voltage source. When turned off for the equivalent resistance calculation, the voltage source becomes a short circuit. Thus, with $v_{ripple}/v_{max} = 0.1$ and $T = 1/60 = 16.7$ ms, $C \approx 470\ \mu$F. Capacitor C would be needlessly large had v_{ripple}/v_{max} been smaller.

How good is our design? A measure of success is **percent regulation**, which is defined as

$$\text{Percent regulation} = \frac{v_{\text{no load}} - v_{\text{full load}}}{v_{\text{full load}}} \times 100\% \,. \tag{4.65}$$

The design suggests a constant output voltage subject to maximum and minimum load conditions, so the percent regulation is zero—a good result. But before we celebrate, we should remember that we have neglected R_z, the series resistance appearing in the second-order Zener breakdown model. In the worst case, this is 16 Ω for the MMSZ5245 device.

Now let $V_s = 21.2$ V (the maximum source value). When we substitute the second-order Zener breakdown model into the trial circuit of Fig. 4.22, we obtain the subcircuit shown in Fig. 4.23. A quick solution yields $v_{out} = 15.194$ V and $v_{out} = 15.420$ V when $R_L = 1$ kΩ and $R_L \to \infty$, respectively. The new percentage regulation is 1.5 % —still not bad. But the percentage regulation could have been much worse if the minimum load resistance had been smaller, thereby subjecting the Zener diode to greater current change.

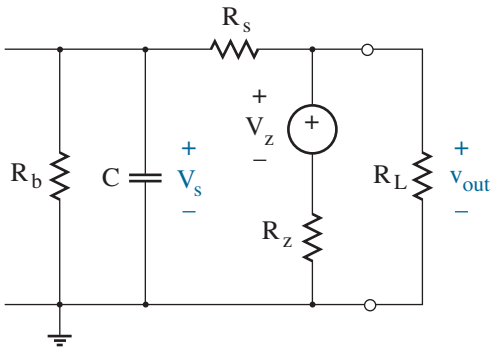


Figure 4.23: Power supply subcircuit for Example 4.6. The Zener diode has been replaced by its second-order breakdown model.

The inclusion of $R_z \neq 0$ in the subcircuit of Fig. 4.23 suggests a revised calculation of capacitance C. With $V_z \to 0$, the new equivalent resistance that the capacitor "sees" is

$$R = [R_s + (R_z \parallel R_L)] \parallel R_b \,. \tag{4.66}$$

We assume the worst-case (small-R) load resistance $R_L = 1$ kΩ. In turn, for $R_z = 16$ Ω, $R = 236$ Ω and $C \approx 470$ μF (a standard value) as before.

Finally, we note that our design can be improved by lowering the rms source voltage so that less power is wasted in R_s. The minimum difference between V_s and v_{out} that preserves the regulator function is called the **drop-out voltage**.

Example 4.7

Use SPICE to evaluate the effect of Zener resistance on the output voltage of the regulated power-supply of Example 4.6.

Solution

We simulate the test circuit shown Fig. 4.24. Here, we decompose V_s into two sources: V_1 is held constant at 19.1 V, and V_2 (the ripple component) varies over the range $0 < V_2 < 2.1$ V —time dependence is not important. The Zener diode is represented by an equivalent **macromodel** within the dashed box. Diode D_1 (with $I_s = 2 \times 10^{15}$ A) functions as an "ordinary" diode with infinite breakdown voltage. And the three-element combination in parallel with D_1 simulates breakdown when $v_{out} > V_z$. For this purpose, we make D_2 ideal (N = 0.01).

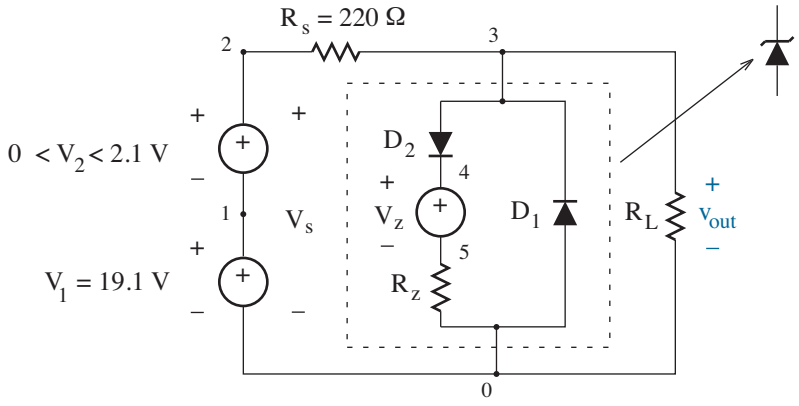


Figure 4.24: SPICE test circuit for Example 4.7.

The following SPICE code is appropriate for a one-time simulation. Frequent use of the macromodel warrants an archival subcircuit (.subckt).

```
* Test Circuit for Example 4.7

V1          1           0           19.1
V2          2           1           0
Rs          3           2           220
RL          3           0           1K

*Zener Diode Macromodel
D1          0           3           DMOD1
D2          3           4           DMOD2
```

```
Vz          4           5           15
Rz          5           0           16

.model      DMOD1       D           ( IS = 2f )
.model      DMOD2       D           ( N = 0.01 )

.dc         V2          0           2.1         0.05
.probe

.end
```

Figure 4.25 shows simulation results for the two extreme load conditions ($R_L = 1$ kΩ and $R_L \to \infty$) with $R_z = 32$, 16, and 8 Ω.

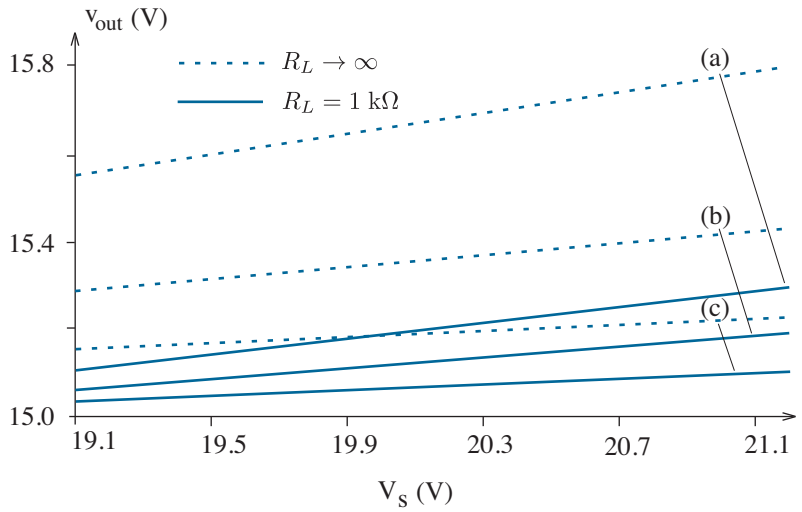


Figure 4.25: Output voltage variations for the test circuit of Example 4.7: (a) $R_z = 32$ Ω; (b) $R_z = 16$ Ω; (c) $R_z = 8$ Ω.

Although quite hard to see for this example (because R_z is relatively small), greater output voltage variations are observed when the load resistance is very large so that Zener-diode current fluctuations have greater magnitude. Variations decrease as R_z decreases. The middle case ($R_z = 16$ Ω), which applies to the design of Example 4.6, indicates a 140-mV output variation. This is a significant improvement over the previous 2.1-V ripple in V_s, but it is substantial nonetheless. We will need to lower the rms source voltage or find a Zener diode with lower R_z if we wish to reduce the ripple voltage further while using the circuit of Fig. 4.22.

Zener Reference Circuits

In recognition of the R_z problems discovered in Examples 4.6 and 4.7, (and in recognition of ever-continuing electronic progress), we now focus on the Zener diode as a **voltage reference** that establishes a particular voltage level, leaving the high-current rough-and-tumble aspects of voltage regulation to other components. Consider the circuit of Fig. 4.26 in which a Zener diode is connected to an op-amp whose V^+ and V^- supplies are the unregulated input (v_{in}) and ground, respectively. We assume that the op-amp is ideal—the real op-amp needs provision for large output current.

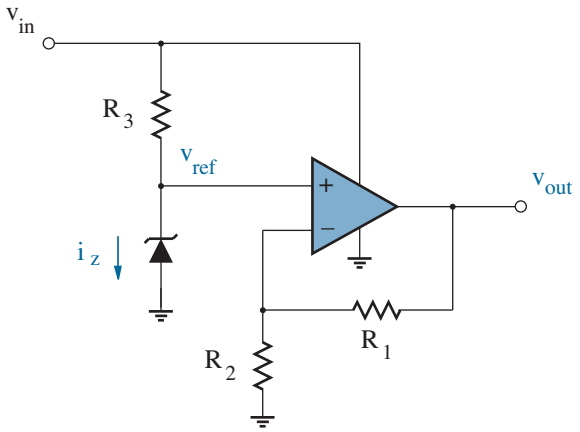


Figure 4.26: Op-amp voltage regulator with Zener reference.

The node voltage at the non-inverting op-amp terminal is $v_{ref} = V_z$, the Zener breakdown voltage, if i_z is greater than the Zener "knee" current. Negative feedback makes this voltage the same as the node voltage at the inverting op-amp terminal, so (after rearrangement)

$$v_{out} = v_{ref}\left(1 + \frac{R_1}{R_2}\right) . \tag{4.67}$$

To regulate v_{out} at 15 V (as in Example 4.6), we arbitrarily choose $V_z = 6$ V, $R_1 = 3$ kΩ, and $R_2 = 2$ kΩ. Then to establish $i_z \sim 1$ mA under worst-case $v_{in} = 19.1$ V, we choose $R_3 = (19.1 - 6)/1 = 13$ kΩ. Now let $R_z = 16$ Ω. When $v_{in} = 19.1$ V, $v_{ref} = 6 + (19.1 - 6)/13 \times 0.016 = 6.016$ V, and $v_{out} = 2.5\, v_{ref} = 15.040$ V. Similarly, when $v_{in} = 21.2$ V, $v_{ref} = 6.019$ V, and $v_{out} = 15.047$ V. Thus, the output voltage variation is a mere 7 mV, which is 20 times less than that observed in Example 4.7. We can do even better if we increase R_3 so that i_z is closer to the actual knee current.

The circuit of Fig. 4.26 is the basis for the (National Semiconductor) LM723 integrated-circuit voltage regulator, an early IC in this category.

Voltage regulation is by no means the only electronic application that requires precision voltage references. Just one example is the process of **analog-to-digital conversion**, in which analog signals are judged greater or less than sets of specific voltage references—see Chapter 14 for details—and individual bits are made either HIGH or LOW accordingly.

A good Zener voltage reference needs small impedance (to avoid voltage changes in the face of changing current) and small temperature coefficient (to avoid standardization errors in changing environments). Figure 4.27 shows impedance and temperature data for the ON Semiconductor MMSZ52 series of Zener diodes. Note the sharp minimum impedance (of about 3 Ω) when $V_z = 6$ V, $i_z = 5$ mA. The corresponding temperature coefficient is positive and above the minimum. However, the coefficient can be reduced by adding a series-connected diode in *forward* bias (see Problem 4.38).

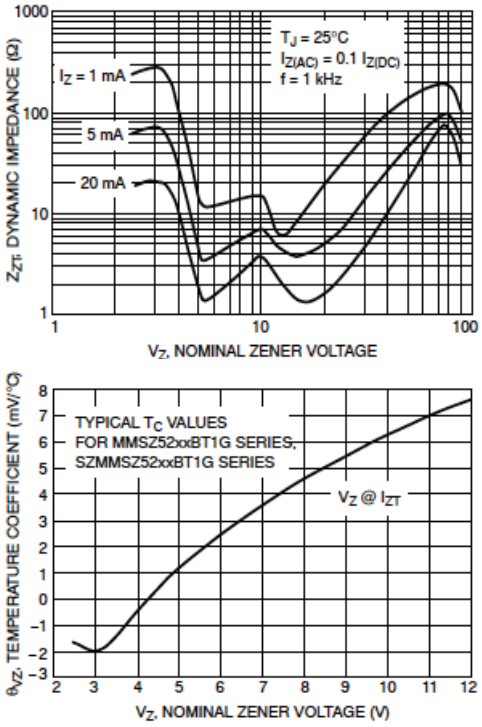


Figure 4.27: Zener diode impedance (Z_z) and temperature characteristics. Used with permission from SCILLC dba ON Semiconductor.

So it appears that one can select a Zener diode with specifications that are nearly optimum for reference stability. And once chosen, this particular diode can be used to establish an arbitrary voltage reference level using a circuit similar to that in Fig. 4.26. One need only adjust R_1 and R_2.

Bandgap Reference Circuits

One of the prime features of integrated circuit technology is the ability to fabricate diodes and other components that are virtually indentical, having been made side by side in the same semiconductor substrate. And when the forward voltages across two identical pn junctions are properly maintained, one can obtain a reference voltage that is nearly temperature independent and roughly equal to the semiconductor bandgap energy. This is called a **bandgap reference**, the modern alternative to a Zener reference.

To understand how the bandgap reference voltage is obtained, consider the demonstration circuit of Fig. 4.28. Simpler circuits without an op-amp are available, but they are beyond our present analytical skills.

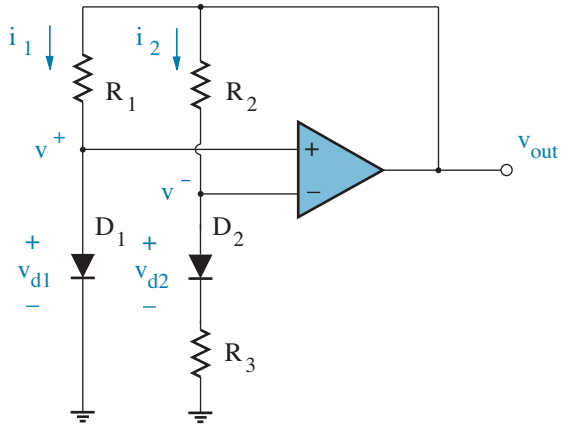


Figure 4.28: Demonstration circuit for a bandgap reference voltage.

The inverting and non-inverting op-amp terminals have equal voltages as a consequence of negative feedback. Thus, with R_3 in series with D_2 and no resistance in series with D_1, we find a rare configuration that does not allow the assignment of equal forward diode voltages of about 0.7 V. Instead, each forward diode voltage must be expressed in terms of the forward current using the fundamental diode characteristic. We assume that both voltages are much larger than kT/q, and we assume unity ideality factors ($n = 1$). Then at the non-inverting op-amp terminal,

$$v^+ \approx \frac{kT}{q} \ln\left(\frac{i_1}{I_{s1}}\right) . \tag{4.68}$$

And at the inverting terminal,

$$v^- \approx \frac{kT}{q} \ln\left(\frac{i_2}{I_{s2}}\right) + i_2 R_3 . \tag{4.69}$$

Identical diodes feature $I_{s1} = I_{s2}$. So with $v^+ = v^-$, and a little algebra

$$\frac{kT}{q} \ln\left(\frac{i_1}{i_2}\right) = i_2 R_3 \, . \tag{4.70}$$

However, the $v^+ = v^-$ condition also implies

$$v_{out} - i_1 R_1 = v_{out} - i_2 R_2 \tag{4.71}$$

so that $i_1/i_2 = R_2/R_1$. We solve Eqs. 4.70 and 4.71 simultaneously for i_1 to find a proportional-to-absolute temperature (PTAT) curent:

$$i_1 = \frac{1}{R_3}\left(\frac{R_2}{R_1}\right)\frac{kT}{q}\ln\left(\frac{R_2}{R_1}\right) \, . \tag{4.72}$$

In turn,

$$v_{out} = v_{d1} + \left(\frac{R_2}{R_3}\right)\frac{kT}{q}\ln\left(\frac{R_2}{R_1}\right) \, . \tag{4.73}$$

Any voltage of this form can be adjusted to a temperature-insensitive value that is close to the bandgap potential.

How so? We recall (from Chapter 2) that the temperature dependence of the diode current is found through the relation

$$i \approx I_o \exp\left(\frac{-E_g}{kT}\right)\exp\left(\frac{qv}{kT}\right) \, , \tag{4.74}$$

where I_o is a constant with weak temperature dependence and E_g is the bandgap energy (1.12 eV for silicon). A typical I_o current is 5×10^4 A. Now let $v = v_{d1}$, solve for v_{d1}, and substitute in Eq. 4.73. This yields

$$v_{out} = \frac{E_g}{q} + \frac{kT}{q}\ln\left(\frac{i_1}{I_o}\right) + \left(\frac{R_2}{R_3}\right)\frac{kT}{q}\ln\left(\frac{R_2}{R_1}\right) \, . \tag{4.75}$$

Suppose, for the moment, i_1 is temperature independent. With $i_1 \ll I_o$, the second term in Eq. 4.75 is negative and subject to cancellation under an appropriate condition for the third term. Specifically,

$$\left(\frac{R_2}{R_1}\right)^{R_2/R_3} = \frac{I_o}{i_1} \, . \tag{4.76}$$

We are left with v_{out} at the bandgap voltage (1.12 V).

In practice, there is a need to consider the relatively weak temperature variations in I_o and E_g. And, for this example, i_1 is proportional to absolute temperature. Thus, the reference voltage at some nominal temperature T_o is obtained by expanding the second term in Eq. 4.75 in powers of $(T - T_o)$ and adjusting the third term to cancel first-order (linear) variations in T, leaving v_{out} slightly higher than the bandgap voltage (see Problem 4.40). Similar procedures can be used to cancel higher-order variations in T.

Choosing a Voltage Reference

Voltage references have a simple functionality, but many practical concerns tend to complicate the selection process.

- The particular output voltage is obviously the prime selection factor. References with 1.25-V, 2.5-V, 3-V, or 5-V outputs are most common. For analog-to-digital and digital-to-analog conversion applications, 2.048-V or 4.096-V outputs conveniently reflect 2^{11} or $2^{12} \times 1$ mV.
- PC-board considerations determine the integrated-circuit package. Mechanical stress effects can cause problems for very small packages.
- Well-designed voltage references are insensitive to temperature.
  - Notwithstanding the specification of a temperature coefficient in ppm/°C, the manufacturer's data sheet should show output voltage vs. temperature over the range -40 °C $< T <$ +125 °C. Variations typically display a broad peak at room temperature.
  - Even at constant temperature, the output may show gradual fluctuations over hundreds of hours. Look for typical data.
  - Check for hysteresis effects that can produce an output change following a complete temperature cycle.
- Voltage references have input power requirements:
  - Wide variations in supply voltage should not affect the output as long as it exceeds V_{ref} by a small, well-defined drop-out voltage.
  - Supply current should be modest, especially for portable circuits. Examine supply-current variations with temperature.
  - Some voltage references have "shutdown" capability to conserve power when a system enters a standby mode.
- Voltage references have output power limitations:
  - The manufacturer's data sheet will typically show the variation of output voltage with current that is "sourced" or provided to another circuit. Source currents of tens of of μA are prevalent. Watch out for potential problems beyond 1 mA.
  - It is less common to find the variation of output voltage with current that is "sinked" or drawn from another circuit.
- Look for a demonstration of transient effects such as the change in output voltage following start-up or a sudden change in load.
- The output should exhibit minimal electrical noise (see Chapter 13), and it should reject any ac signals that may be superimposed on the input supply voltage.
- Seek the lowest cost that satisfies reasonable design requirements.

4.3 Signal Conditioning and Clamping

Time-dependent signals often fail to live up to our demanding expectations. So we beat them into palatable forms by removing something here or adding substance there during encounters with special analog conditioning circuits. (More powerful digital signal processing methods are typically expensive.) This section concerns diode-assisted non-linear processes.

Passive Clippers

As would-be signal sculptors, two basic non-linear tools are at our disposal. These are the positive and negative **clipper** circuits, which remove portions of the input waveform through a passive downscale or diminutive process. Clipper or **limiter** circuits have the half-wave-rectifier form, but the output voltage of interest no longer appears across the load resistance. In contrast, we focus on the voltage across the diode and series-connected voltage source (actually a reference voltage) with value V_b.

The positive clipper circuit shown in Fig. 4.29a features an ideal diode. When $v_{in} < V_b$, the diode operates in reverse bias, $i = 0$, and $v_{out} = v_{in}$. However, when $v_{in} > V_b$, the diode is forward biased, $i \neq 0$, and $v_{out} = V_b$ —the voltage that exceeds V_b appears across resistance R. Figure 4.29b shows the applicable input/output **transfer characteristic**, and Fig. 4.29c shows the output waveform that results when $v_{in}(t) = \tilde{v} \sin \omega t$, $(\tilde{v} > V_b)$. The "positive" (upper) portion of the input waveform is removed (clipped) in relation to voltage V_b.

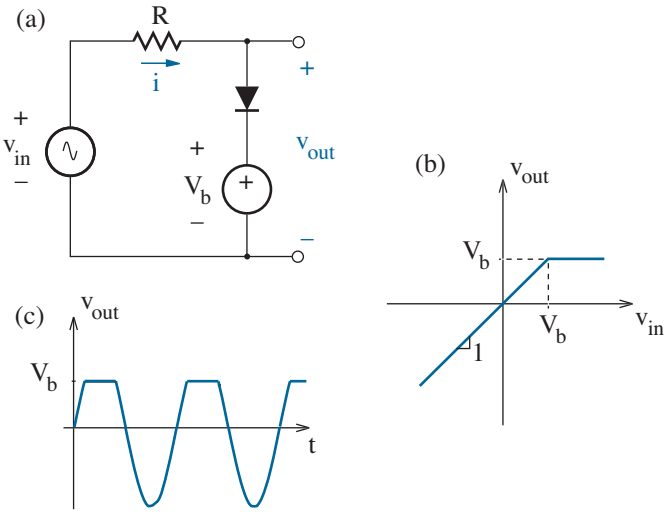


Figure 4.29: Positive clipper circuit, transfer characteristic, and sample output when $v_{in}(t) = \tilde{v} \sin \omega t$.

A negative clipper circuit derives from the positive case by reversing the diode polarity as shown in Fig. 4.30a. Similar arguments concerning forward- and reverse-bias diode conditions lead to the input/output transfer characteristic of Fig. 4.30b, and the particular output waveform of Fig. 4.30c results when $v_{in}(t) = \tilde{v}\sin\omega t$ ($\tilde{v} > V_b$). The "negative" (lower) portion of the input waveform is removed (clipped) in relation to voltage V_b .

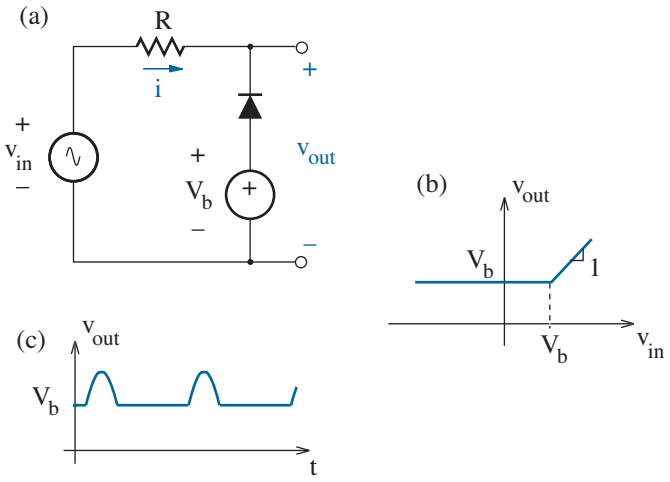


Figure 4.30: Negative clipper circuit, transfer characteristic, and sample output when $v_{in}(t) = \tilde{v}\sin\omega t$.

When signals are unruly, positive and negative clippers often provide a simple means for circuit protection. For example, in the circuit of Fig. 4.31, diodes D_1 and D_2 provide positive clipping action in relation to +5 V and negative clipping action in relation to ground, respectively. Thus, abnormal out-of-bound input voltages cannot damage the subsequent sensitive circuit.

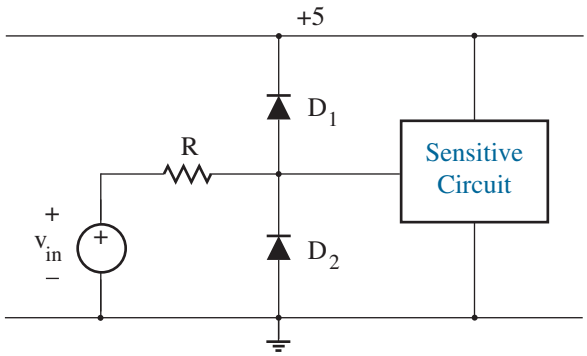


Figure 4.31: Positive and negative clipping for input protection.

More sophisticated waveform "chisels" can be realized by inserting a resistor between the diode and the reference voltage source in either the positive or negative clipper circuit. Figure 4.32a shows the "sophisticated" positive clipper circuit. When $v_{in} < V_b$, the diode is reverse biased, and $v_{out} = v_{in}$, as before. However, when $v_{in} > V_b$, the diode is forward biased, and by superposition

$$v_{out} = V_b + \frac{R_1}{R_1 + R_0}\,(v_{in} - V_b)\,. \tag{4.77}$$

Figure 4.32a also shows the corresponding transfer characteristic. Portions of an input waveform that are greater (more positive) than reference voltage V_b are downscaled (diminished) by the factor $R_1/(R_1 + R_0)$. In the limit as $R_1 \to 0$, a zero scale factor constitutes total removal, and we obtain the simple positive clipper circuit.

The "sophisticated" negative clipper circuit and corresponding transfer characteristic appear in Fig. 4.32b. Portions of an input waveform that are less (more negative) than reference voltage V_b are downscaled (diminished) by the factor $R_1/(R_1 + R_0)$. In the limit as $R_1 \to 0$, a zero scale factor constitutes total removal, and we obtain the simple negative clipper circuit.

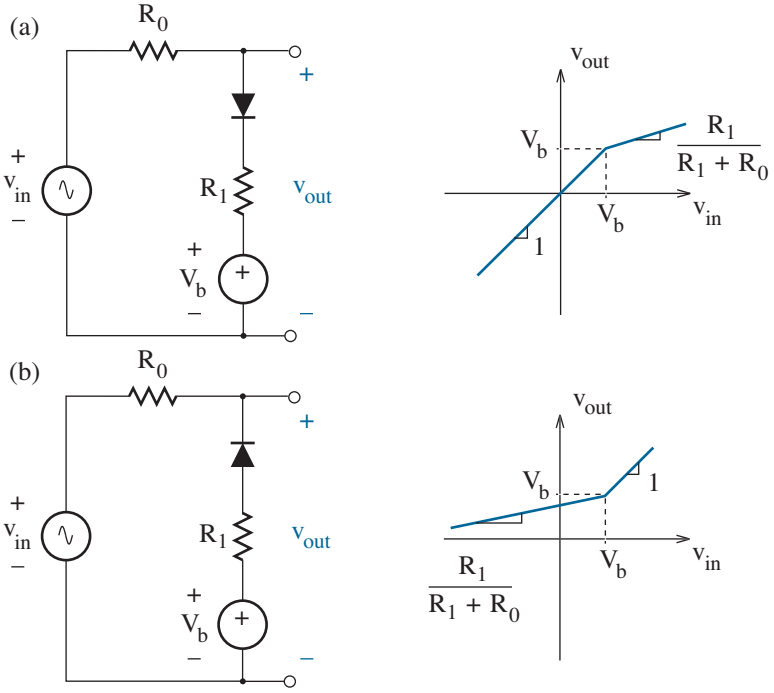


Figure 4.32: "Sophisticated" clipper circuits and transfer characteristics: (a) positive; (b) negative.

Example 4.8

Design a circuit with the non-linear transfer characteristic $v_{out} = \sqrt{v_{in}}$. The circuit corrects for a sensor output with square-law dependence.

Solution

We approximate the actual square-root transfer characteristic with three straight-line segments as shown in Fig. 4.33a. The first and second segments have slopes $m_1 = 1$ and $m_2 = 1/3$, respectively. Together, these suggest the action of a positive sophisticated clipper circuit. Then in consideration of a third segment with slope $m_3 = 1/5$, we appear to have a second positive sophisticated clipper circuit that is cascaded against the output of the first. We thus obtain the trial circuit of Fig. 4.33b, and the design is complete if acceptable values for V_{b1}, V_{b2}, R_1, and R_2 can be determined.

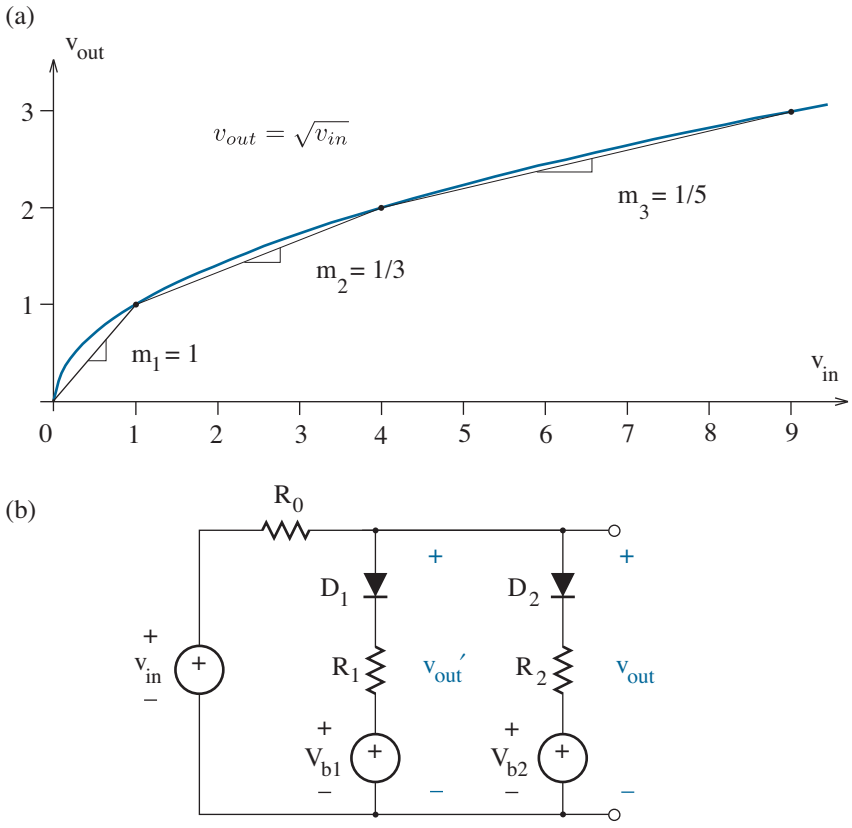


Figure 4.33: Square-root transfer characteristic: (a) approximate behavior consisting of straight-line segments; (b) trial "sophisticated" clipper circuit that implements the approximate transfer characteristic.

We match the transfer characteristic of the left half of the trial circuit to the first two straight-line segments of the desired transfer characteristic by making $V_{b1} = 1$ V and $R_1 = R_0/2$. The right half of the trial circuit remains inactive provided $v_{out}{}'$, the intermediate output voltage, remains less than V_{b2}, and in this case $v_{out}{}' = v_{out}$. That was easy.

Now consider the third straight-line segment, which begins at $v_{in} = 4$ V. Should we make $V_{b2} = 4$ V? No! The second sophisticated clipper reacts to $v_{out}{}'$ as its input. And when $v_{in} = 4$ V, $v_{out}{}' = 2$ V. Thus, $V_{b2} = 2$ V. To determine R_2, we realize that both diodes are "on" over the course of the third segment. Then by superposition,

$$v_{out} = v_{in}\,\frac{R_1 \parallel R_2}{R_0 + R_1 \parallel R_2} + V_{b1}[\qquad\qquad] + V_{b2}[\qquad\qquad]\,. \tag{4.78}$$

In this expression, the coefficients that multiply V_{b1} and V_{b2} are irrelevant. We need only recognize that the coefficient that multiplies v_{in} is the slope of the third straight-line segment. It soon follows that $R_1 \parallel R_2 = R_0/4$, and $R_2 = R_0/2$. We set $R_0 = 2$ kΩ to limit currents to a few mA or less. Then $R_1 = 1$ kΩ and $R_2 = 1$ kΩ.

How good is our design? We may want to improve functional accuracy by approximating the actual transfer characteristic with more than three straight-line segments. And we should be considerate of forward voltage drops across non-ideal diodes.

Given the latter concern, we engage SPICE to verify our design when the diodes are ideal. Then we repeat the simulation for the case of non-ideal diodes with $I_s = 1 \times 10^{-14}$ A. (The format of an appropriate SPICE code is similar to that used in Example 4.7, so we do not bother to list it here.) Figure 4.34 shows the results. The effect of the diode turn-on voltage is to shift the transfer characteristic upward by about 0.5 V. For improvement, we need to adjust V_{b1} and V_{b2} or find a diode with smaller turn-on voltage.

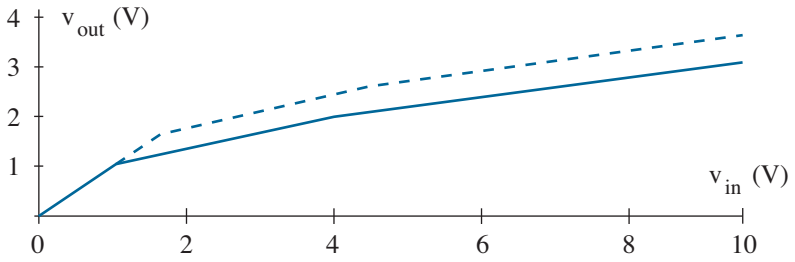


Figure 4.34: SPICE transfer characteristic for Example 4.8. The solid curve corresponds to ideal diodes. The dashed curve reflects $I_s = 1 \times 10^{-14}$ A.

Active Waveshaping

The preceding conditioning circuits are *passive* —waveform modifications are achieved by downscale processing. In contrast, upscale (augmentative) processes require *active* waveshaping circuits whereby selected waveform features are amplified. These are easily realized with operational amplifiers.

Consider the op-amp circuit of Fig. 4.35. For the moment, we assume that the subcircuit containing diode D_2 is not connected.

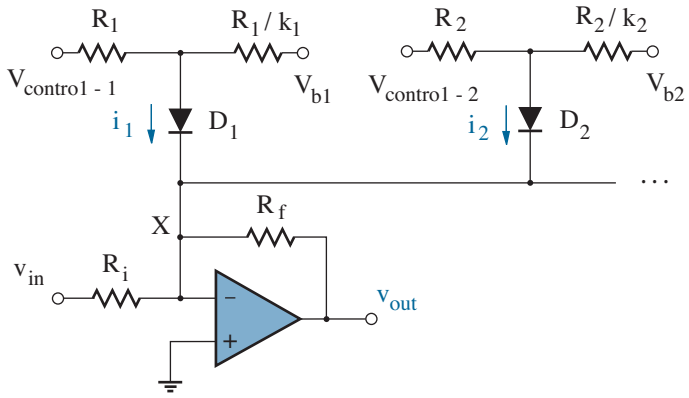


Figure 4.35: Active waveshaping circuit.

For some range of voltage $V_{\text{control}-1}$ relative to V_{b1}, D_1 is reverse biased. We are left with a simple linear inverting amplifier for which

$$v_{out} = m\, v_{in}\,, \tag{4.79}$$

where m is the slope of the transfer characteristic. In this case,

$$m = \frac{\Delta v_{out}}{\Delta v_{in}} = -\frac{R_f}{R_i}\,. \tag{4.80}$$

Equation 4.80 must change when diode D_1 is forward biased. We determine the breakpoint for D_1 by first noting that the op-amp inverting input node (X) is at virtual ground. Then at breakpoint,

$$i_1 = \frac{V_{\text{control}-1}}{R_1} + \frac{V_{b1}}{R_1/k_1} = 0\,. \tag{4.81}$$

It follows that diode D_1 is forward biased when

$$V_{\text{control}-1} > -k_1 V_{b1} \quad \text{(“positive” diode orientation).} \tag{4.82}$$

We note that if D_1 is reverse connected, it is forward biased when

$$V_{\text{control}-1} < -k_1 V_{b1} \quad \text{(“negative” diode orientation).} \tag{4.83}$$

Now determine the change in the slope of the transfer characteristic when D_1 is forward biased and connected as in Fig. 4.35. (Different results are expected when D_1 is reverse connected.) We examine two cases.

Case 1: $V_{\text{control}-1} = v_{out}$. From Kirchhoff's Current Law at node X,

$$\frac{v_{out}}{R_f} + \frac{v_{in}}{R_i} + \frac{V_{b1}}{R_1/k_1} + \frac{v_{out}}{R_1} = 0\,. \tag{4.84}$$

Thus,

$$m = \frac{\Delta v_{out}}{\Delta v_{in}} = -\frac{(R_f \parallel R_1)}{R_i}\,. \tag{4.85}$$

We combine this with earlier results to obtain the transfer characteristic shown in Fig. 4.36a. The breakpoint is controlled by v_{out} relative to V_{b1}, and the magnitude of m is *decreased* for v_{out} greater than breakpoint. Note that the transfer characteristic does not pass through the origin. Opposite behavior is observed when the diode is reverse connected to yield the transfer characteristic of Fig. 4.36b.

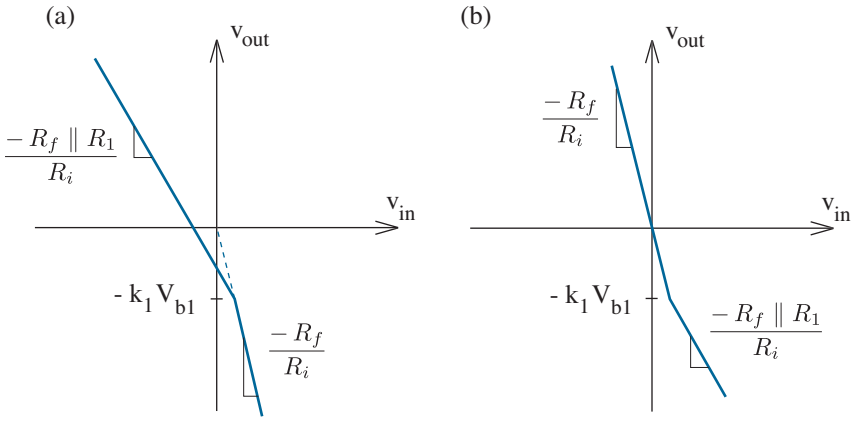


Figure 4.36: (a) Transfer characteristic for the active waveshaping circuit of Fig. 4.35 with $V_{\text{control}-1} = v_{out}$; (b) Transfer characteristic for the same circuit with reverse diode orientation.

Case 2: $V_{\text{control}-1} = v_{in}$. From Kirchhoff's Current Law at node X,

$$\frac{v_{out}}{R_f} + \frac{v_{in}}{R_i} + \frac{V_{b1}}{R_1/k_1} + \frac{v_{in}}{R_1} = 0\,. \tag{4.86}$$

Thus,

$$m = \frac{\Delta v_{out}}{\Delta v_{in}} = -\frac{R_f}{(R_i \parallel R_1)}\,. \tag{4.87}$$

We combine this with earlier results to obtain the transfer characteristic shown in Fig. 4.37a. The breakpoint is controlled by v_{in} relative to V_{b1}, and the magnitude of m is *increased* for v_{in} greater than breakpoint. Note that the transfer characteristic does not pass through the origin. Opposite behavior is observed when the diode is reverse connected to yield the transfer characteristic of Fig. 4.37b.

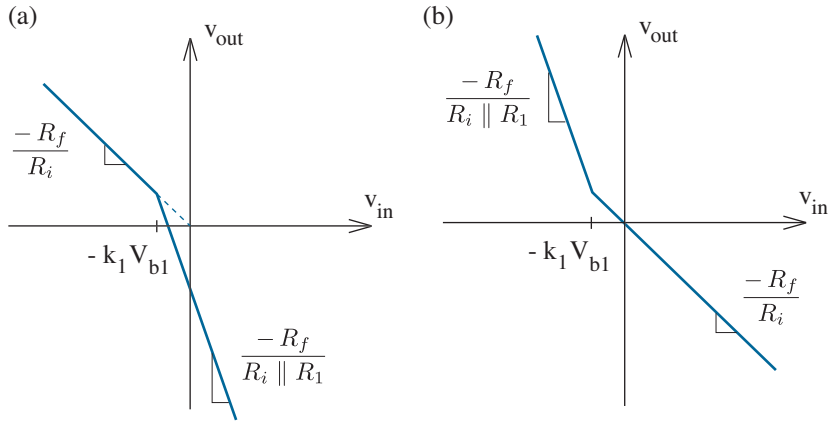


Figure 4.37: (a) Transfer characteristic for the active waveshaping circuit of Fig. 4.35 with $V_{\text{control}-1} = v_{in}$; (b) Transfer characteristic for the same circuit with reverse diode orientation.

For either of these two cases, v_{out} is independent of loading conditions. Arbitrary resistance can be inserted between v_{out} and ground without changing the ideal transfer characteristic. In contrast, passive waveshaping circuits considered previously are load dependent.

If the subcircuit that contains diode D_2 is connected to the op-amp inverting input terminal as suggested in Fig. 4.35, the analysis of the new circuit behavior is straightforward. We simply add another breakpoint that is governed by an equation similar to Eq. 4.82 (or Eq. 4.83 for reverse D_2). The slope of the transfer characteristic must be determined for all D_1 and D_2 bias states. Suppose, for example, D_1 and D_2 are both forward biased (and positively oriented as shown in the figure). If the circuit is configured so that $V_{\text{control}-1} = V_{\text{control}-2} = v_{out}$, we can write an equation similar to Eq. 4.84 but with extra terms to obtain

$$m = \frac{\Delta v_{out}}{\Delta v_{in}} = -\frac{(R_f \parallel R_1 \parallel R_2)}{R_i} \,. \tag{4.88}$$

In view of the many degrees of freedom implied by the circuit of Fig. 4.35 and the ease of adding diode subcircuits to increase the crucial breakpoints, the possibilities for transfer-characteristic design are limitless.

Waveform Clamping

Apart from an added capacitor, the circuit of Fig. 4.38 is identical to the positive clipper considered previously. But when excited with a periodic input voltage waveform, the circuit yields a significantly different response subject to $RC \gg T$, where T is the waveform period. For reasons that will soon become clear, the altered behavior is indicative of a positive **clamp**.

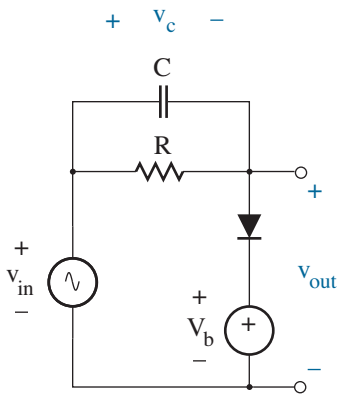


Figure 4.38: Positive clamp circuit.

We find $v_{out}(t)$ subject to the following conditions: the diode is ideal, $v_{in} = \tilde{v} \sin \omega t$ turns on at $t = 0$, $v_{in,max} = \tilde{v} > V_b > 0$, $v_c = 0$ for $t < 0$, and $RC \to \infty$. The solution is obtained by plotting the v_{in}, v_c, and v_{out} waveforms in succession as shown in Fig. 4.39.

Let t' denote the time when $v_{in} = V_b$. For $0 < t < t'$, the diode is reverse biased. So in the absence of charging current, v_c remains zero, and $v_{out} = v_{in}$. For $t' < t < T/4$, the diode is forward biased, and $v_{out} = V_b$. Meanwhile, the presence of charging current increases v_c with v_{in} until $v_c\,(T/4) = \tilde{v} - V_b$. The time $t = T/4$ marks the conclusion of a "transient" interval that establishes *constant* v_c —v_{in} never increases beyond $v_{in,max}$ and the capacitor never discharges ($RC \to \infty$). Thus, we have

$$v_{out} = v_{in} - v_c = \tilde{v}\,(\sin \omega t - 1) + V_b\,. \tag{4.89}$$

For $t > T/4$, the output voltage waveform has the same shape and peak-to-peak difference as the input voltage waveform, but *its positive peak is clamped (held) at* V_b. This reference voltage can assume any value that is less than $v_{in,max}$ (even though $V_b < 0$ has a modified transient response). Conversely, the diode is under reverse bias, the capacitor does not charge, and $v_{out} = v_{in}$ for all time.

Note that the diode approaches but never quite reaches a forward-bias state at times $t = T/4 + mT$, where m is an integer. This is not the case when RC is finite. If the capacitor voltage decays during time period T, it must be restored to $\tilde{v} - V_b$ over short time intervals of duration δ prior to $t = T/4 + mT$. (See Fig. 4.6.) During these intervals, $v_{out} = V_b$.

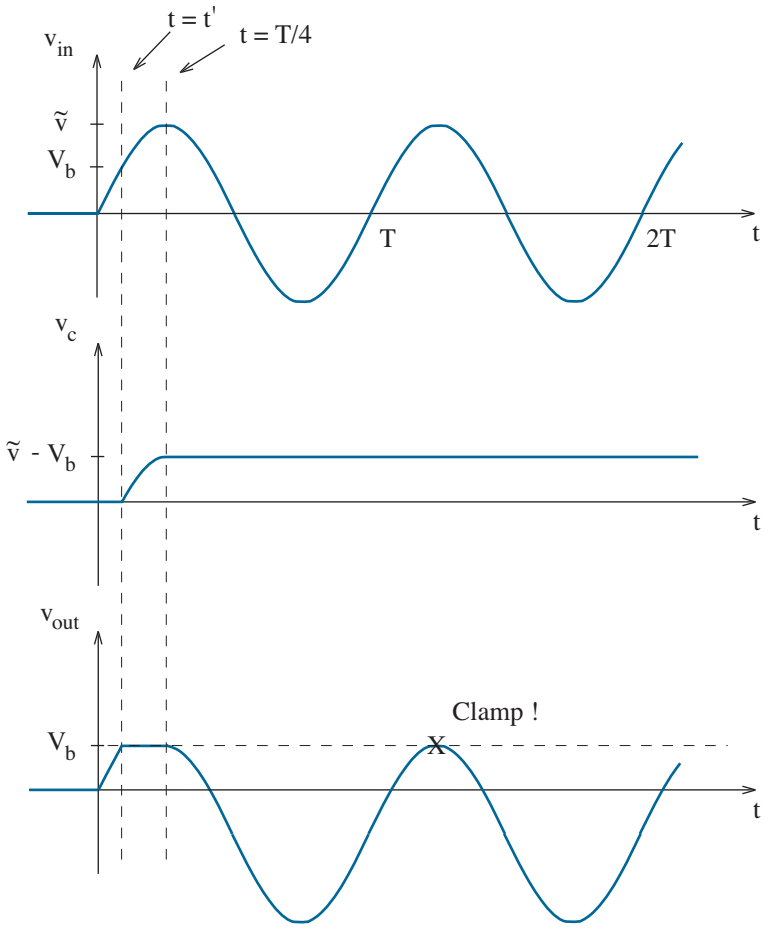


Figure 4.39: Waveforms for the positive clamp circuit.

When designing the clamp circuit, Eq. 4.4 can be used to determine a capacitor value that maintains nearly constant v_c during a waveform period. However, it is often convenient simply to remove R from the circuit so that $R \to \infty$. This makes C small and somewhat arbitrary.

The negative clamp circuit is derived from the positive clamp circuit by reversing the diode polarity as shown in Fig. 4.40.

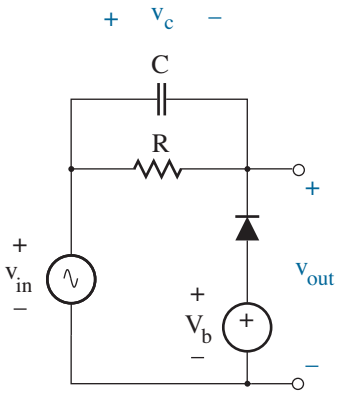


Figure 4.40: Negative clamp circuit.

We determine $v_{out}(t)$ subject to the conditions cited for the positive clamp circuit, except $v_{in,min} = -\tilde{v} < V_b < 0$ (to allow $v_c = 0$ for $t < 0$). As before, the solution is obtained by plotting v_{in}, v_c, and v_{out} waveforms in succession as shown in Fig. 4.41.

Let t' denote the time when $v_{in} = V_b$. For $0 < t < t'$, the diode is reverse biased. So in the absence of charging current, v_c remains zero, and $v_{out} = v_{in}$. For $t' < t < 3T/4$, the diode is forward biased, and $v_{out} = V_b$. Meanwhile, the presence of charging current decreases v_c with v_{in} until $v_c\,(3T/4) = -\tilde{v} - V_b$. Time $t = 3T/4$ is the end of a "transient" interval. Subsequently, v_c remains constant, since v_{in} never decreases below $v_{in,min}$ and the capacitor never discharges ($RC \to \infty$). Thus, we have

$$v_{out} = v_{in} - v_c = \tilde{v}\,(\sin\omega t + 1) + V_b\,. \tag{4.90}$$

For $t > 3T/4$, the output voltage waveform has the same shape and peak-to-peak difference as the input voltage waveform, but *its negative peak is clamped (held) at* V_b. This reference voltage can assume any value that is greater than $v_{in,min}$ (even though $V_b > 0$ has a modified transient response). Conversely, the diode is under reverse bias, the capacitor does not charge, and $v_{out} = v_{in}$ for all time.

Note that the diode approaches but never quite reaches a forward-bias state at times $t = 3T/4 + mT$, where m is an integer. This is not the case when RC is finite. If the capacitor voltage decays during time period T, it must be restored to $-\tilde{v} - V_b$ over short time intervals of duration δ prior to $t = 3T/4 + mT$. During these intervals, $v_{out} = V_b$.

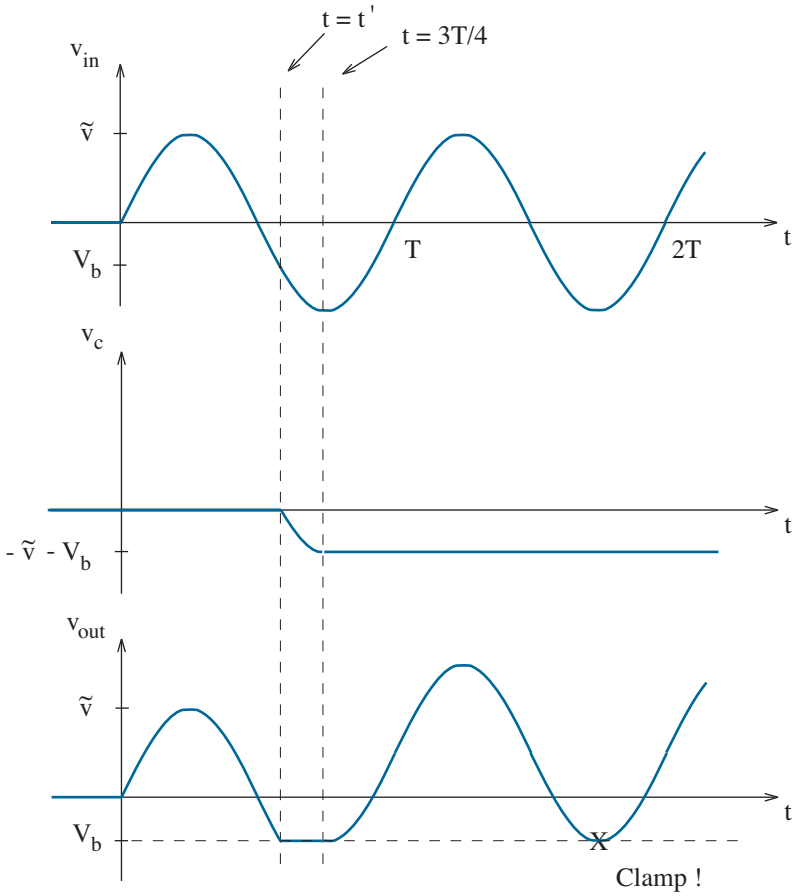


Figure 4.41: Waveforms for the negative clamp circuit.

Figure 4.42 depicts a classic clamping application for signal processing. Suppose one has a time-dependent signal (Fig. 4.42a) whose instantaneous value in relation to zero (ground) is an indication of some physical attribute. Further suppose that system considerations require **capacitive coupling** for the signal to enter a processing module—the capacitor acts as a short circuit for the ac signal component, but it blocks the dc signal component. Unfortunately, the "zero" dc level for the signal thus derived (Fig. 4.42b) reflects a relatively slow-moving time average of essentially random changes. Thus, the signal calibration is lost. To correct the problem, one can arrange for the original signal to be interspersed with short pulses that extend to zero or some other *minimum* reference voltage (Fig. 4.42c). Upon entering the processing module, the ac component of this signal has unspecified time average as before. However, the pulse portions of the signal can be tied to zero with a negative clamp circuit to restore proper calibration (Fig. 4.42d). Such clamping is common for RGB and other forms of video signals.

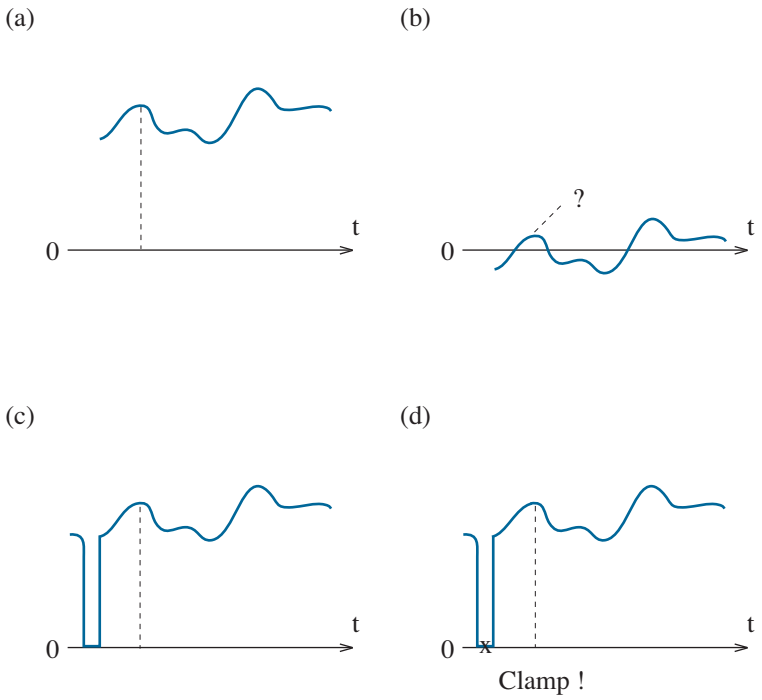


Figure 4.42: Motivating waveforms for signal processing with diode clamps: (a) calibrated signal relative to ground; (b) time-averaged (ac) signal with lost calibration; (c) calibrated signal with added pulse to reference voltage; (d) time-averaged signal with clamp to reference voltage.

4.4 Optoelectronic Circuits

We conclude this chapter with some optoelectronic diode applications.

LED and Laser-Diode Drivers

Whereas the intensity of light from an LED or semiconductor laser diode is linearly related to forward current, a current source is generally necessary. In trivial form, this is merely a voltage source and a current-limiting resistor R_s as shown in Fig. 4.43. The turn-on voltage for a typical $Al_xGa_{1-x}As$ LED is about 1.2 V, so the associated current is

$$i_f \approx \frac{V_s - 1.2}{R_s} , \tag{4.91}$$

subject to $i_f > 0$. No further complexity is warranted for on/off indicators.

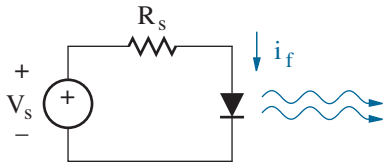


Figure 4.43: Simple circuit for LED or laser-diode output control.

An improved current-control circuit that accounts for manufacturing- or temperature-induced variations in diode forward voltage can be realized with the help of feedback as shown in Fig. 4.44. Here, an op-amp produces diode current i_f, which induces a voltage across a *small* series-connected resistor R_{sense}. This voltage is multiplied by factor A_{vd} though the action of a differential amplifier, and the result is fed back to the inverting terminal of the op-amp. Equality with the non-inverting terminal voltage V_s requires

$$i_f = \frac{V_s}{A_{vd}\, R_{sense}} . \tag{4.92}$$

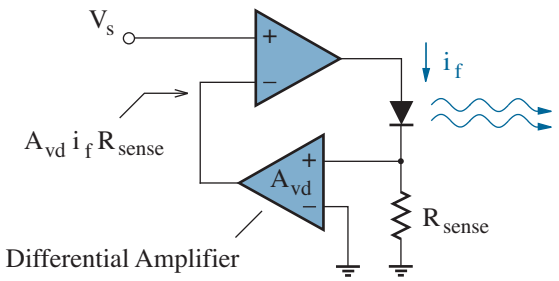


Figure 4.44: Improved circuit for LED or laser-diode output control.

The circuit of Fig. 4.44 is the basis for controllers such as the MAX16803. Resistor R_{sense} sets the maximum LED current, an "Enable" pin provides on-off control, and a similar "Dim" pin allows for perceived LED brightness control in proportion to the duty cycle of a supplied rectangular waveform.

Some LED controllers accept moderate-valued resistors for setting the output current. In Fig. 4.45, for example, a programming resistor R_p sets current i_{ref}, which is amplified by a factor A_i and transferred to the LED. A digitally controlled potentiometer is sometimes used in place of R_p.

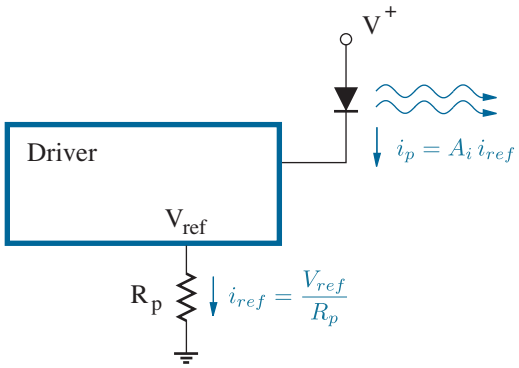


Figure 4.45: LED or laser-diode output control with moderate R_p.

Photodetection

The photodetection process used to receive optical information requres a reverse-biased photodiode and the current-to-voltage converter of Fig. 4.46. Subject to photocurrent i_p, the output response is

$$v_{out} = -i_p\, R_f\ . \tag{4.93}$$

Look to Chapter 12 for issues that limit dynamic op-amp performance.

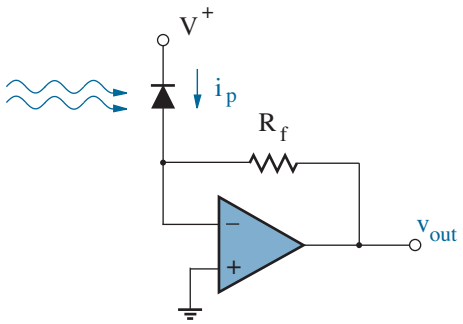


Figure 4.46: Photodetector and current-to-voltage converter.

Optical Communications

Laser diodes and photodetectors work together as transmitters and receivers in optical communication systems. Fiber-optic systems (as for the internet) and handheld optical communication devices (as for a TV remote control) engage complex protocols that go beyond the scope of this text. A simpler system conveys information over a short distance so that the transmitter and receiver circuit environments are electrically isolated from each another. This is often desirable when one of them "floats" at a high voltage.

Figure 4.47 shows the salient features of a digital **optocoupler** such as the Fairchild FOD8001. When v_{in} is LOW, the output of the front-end inverter is HIGH so that D_1 emits light. Absorption by D_2 produces photocurrent i_p, and the op-amp current-to-voltage converter forces

$$v_x = V_{ref1} - i_p R_f < V_{ref2} \tag{4.94}$$

(subject to $V_{ref1} > V_{ref2}$). Thus, the comparator action makes v_{out} LOW. Conversely, when v_{in} is HIGH, D_1 is off, D_2 does not absorb any light, and

$$v_x = V_{ref1} > V_{ref2}\,. \tag{4.95}$$

In turn, the comparator action makes v_{out} HIGH. The overall behavior is the same as that of a digital buffer.

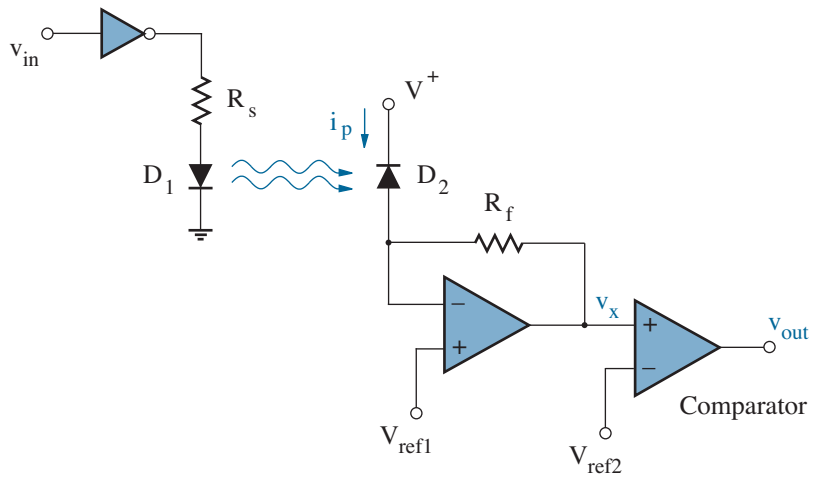


Figure 4.47: Basic circuitry in a digital optocoupler such as the FOD8001. The portions containing D_1 and D_2 are electrically isolated.

The realization of an analog optocoupler is relatively difficult because the input and output need to be proportional to one another as opposed to both on or both off. The problem is resolved with an integrated optocoupler (such as the Vishay IL300) that promotes feedback as shown in Fig. 4.48. Positive v_{in} at the non-inverting input to op-amp X induces current i_f — the character of the relationship is unimportant—so that D_x emits light. A fraction K_1 of this radiation is absorbed by D_1 to yield photocurrent i_{p1}, which is fed back to op-amp X to force

$$i_{p1}R_1 = v_{in} \, . \tag{4.96}$$

Meanwhile, a fraction K_2 of the D_x radiation is absorbed by D_2 to yield photocurrent i_{p2}, which flows through R_2 to produce

$$i_{p2}R_2 = v_{out} \tag{4.97}$$

at the output of the op-amp buffer. In turn, we divide Eq. 4.97 by Eq. 4.96 and rearrange to find

$$v_{out} = \left(\frac{i_{p2}}{i_{p1}}\right)\frac{R_2}{R_1}\,v_{in} = \left(\frac{K_2}{K_1}\right)\frac{R_2}{R_1}\,v_{in} \tag{4.98}$$

The ratio of K_2 to K_1 is a predetermined constant, so the output and input are proportional as required.

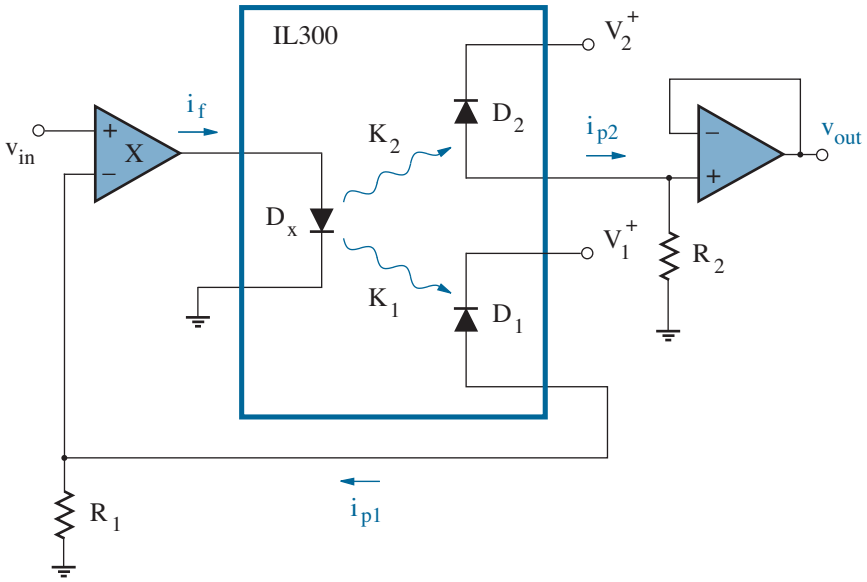


Figure 4.48: Basic circuitry in an analog optocoupler such as the IL300. The portion containing D_2 is electrically isolated from everything else.

Concept Summary

- Rectifier circuits eliminate (half-wave) or make positive (full-wave) the negative portions of an ac waveform to yield a non-zero dc average.
  - A dc voltage equal to the peak value of a rectified ac waveform derives from a single capacitor in parallel with the resistive load. A small ripple voltage favors small load current.
  - A dc voltage equal to the time average of a rectified ac waveform dervies from an LC filter positioned in front of the resistive load. A small ripple voltage favors large load current.
- dc-to-dc converters achieve modified dc voltages from a battery input.
  - Step-down or buck converters generate a rectangular waveform that switches between the dc input and zero with duty cycle D. An LC filter extracts a dc time average in proportion to D.
  - Step-up or boost converters switch from inductor energy storage to release with duty cycle D. A capacitor sustains the peak of a rectified rectangular waveform that varies inversely with $1 - D$.
  - The preceding converters have common design features:
    - Inductors must be sufficiently large to ensure their operation with continuous non-zero current.
    - Capacitors require small equivalent series resistance so that internal current fluctuations do not promote ripple voltage.
    - High switching frequencies ensure small-value components.
- Zener diodes impart voltage regulation through precise breakdown.
  - The Zener diode absorbs more or less reverse current, as needed, to accommodate varying input and load conditions.
  - In modern applications, the Zener diode simply provides a stable voltage reference, leaving an op-amp to supply the load current. Bandgap voltage references provide even better stabiity.
- Clipping diode circuits are used to remove the more positive or more negative portions of a waveform in relation to a reference voltage.
- Clamping diode circuits are used to affix the most positive or most negative value of a waveform to the level of a reference voltage.
- Optoelectronic circuits can support digital or analog communication between systems requiring electrical isolation.

Problems

Section 4.1

4.1 Starting with the current-voltage characteristic for a capacitor C, show that for the case of full-wave rectification and RC filtering

$$v_{ripple} \approx \frac{i_{Load}}{2fC},$$

where i_{Load} is load current and f is the frequency of the ac input.

4.2 Show that Eq. 4.12

$$\cos\phi = 1 - \frac{v_{ripple}}{v_{max}}$$

applies to both half-wave and full-wave rectification.

4.3 The circuit of Fig. P4.3 features an ideal diode and 0.6-V output ripple. The ac power source (V_s) produces 18 V rms at 200 Hz.

(a) Determine the maximum instantaneous load current (through R).

(b) Determine the time duration of the intervals of forward diode bias.

(c) Find the maximum instantaneous diode current.

(d) Find the diode peak reverse voltage.

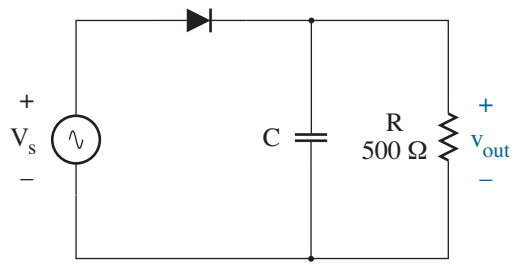


Figure P4.3

4.4 A particular circuit with RC filtering rectifies a 50-Hz voltage waveform. Two conditions are known:

- The capacitor value is 10 mF.
- The current to the parallel RC combination has the time dependence shown in Fig. P4.4.

(a) Does the circuit feature half-wave or full-wave rectification?

(b) Determine the ripple voltage.

(c) Find the maximum instantaneous load current.

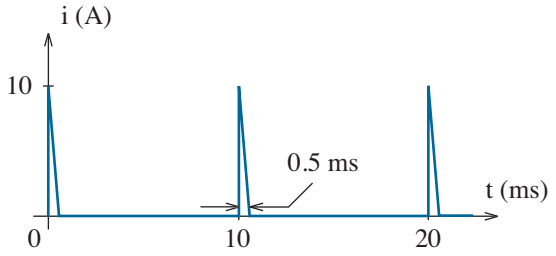


Figure P4.4

4.5 Design a capacitively filtered half-wave rectifier circuit to deliver 50 mA to a 400-Ω load resistance with 2-% ripple voltage. Specify the diode ratings. Assume a 60-Hz 115-V (rms) power source.

4.6 Repeat Problem 4.5 with full-wave rectification.

4.7 Design a capacitively filtered half-wave rectifier circuit to deliver 100 mA to a 100-Ω load resistance with 5-% ripple voltage. Specify the diode ratings. Assume a 60-Hz 115-V (rms) power source.

4.8 Repeat Problem 4.7 with full-wave rectification.

4.9 The 4.7-mF (electrolytic) capacitor selected for the RC-filtered power supply of Example 4.1 has a 75-mΩ parasitic equivalent series resistance (ESR). Estimate the power dissipated in the capacitor when $R_L = 200\ \Omega$ (minimum load resistance).

4.10 This problem features a diode as a **detector**. In a simple **amplitude modulation** (AM) process, a received radio signal has the form

$$v_r(t) = \tilde{v}_m(1 + m \sin \omega_m t) \sin \omega_c t \, ,$$

where ω_c is the "carrier" angular frequency (to which the radio is tuned), ω_m is the "modulating" angular frequency of the information signal of interest, and constant m is the **modulation index** $(0 \le m \le 1)$. In what follows, let ω_c and ω_m correspond to 10-kHz and 1-kHz frequencies, respectively. In practice, ω_c is significantly larger, and ω_m is part of a spectrum with particular $\tilde{v}_m$ weightings.

(a) Use SPICE to plot $v_r(t)$ over two ω_m periods. Assume that the signal is applied to a 1-kΩ load, and let $m = 1$.

(b) Use SPICE to demonstrate the load voltage that results when the preceding circuit is modified to provide ideal half-wave rectification.

(c) A good AM receiver will produce an output that approximates $m\,\tilde{v}_m \sin \omega_m t$. Use SPICE to explore this prospect when a capacitor C is placed in parallel with the load of the preceding circuit. Let $C = 1$ μF, then seek an "optimum" value. Discuss your results for $m = 1$ and $m = 0.5$.

(d) Show that least distortion results when

$$RC \le \frac{\sqrt{1 - m^2}}{m\,\omega_m} \, .$$

In turn, practical AM modulation has $m < 1$. Hint: Consider the slope of the output voltage in relation to the factor that modulates $\sin \omega_m t$.

Note: The diode's role as a detector is arguably the most important factor in its historical development. (See also Problem 2.7.)

4.11 In the circuit of Fig. P4.11, the end connections of the transformer primary are alternately switched to ground while the center tap connects to (dc) V_o. A full-wave-rectified signal is desired at load R.

(a) Design an appropriate secondary circuit that features 2 diodes.

(b) Design an appropriate secondary circuit that features 4 diodes.

(c) How do the output amplitudes compare?

Note: The MAX845 provides low-power switching.

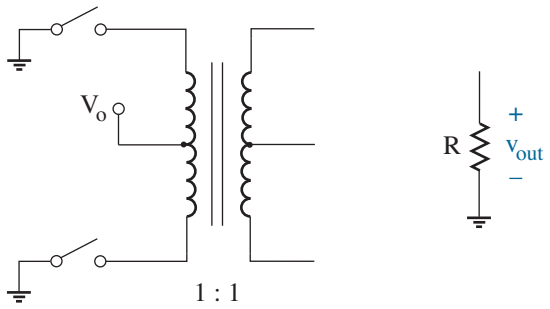


Figure P4.11

4.12 The three-phase circuit of Fig. P4.12 features ideal diodes and source voltages $V_a = \tilde{v} \sin(\omega t)$, $V_b = \tilde{v} \sin(\omega t - 120°)$, and $V_c = \tilde{v} \sin(\omega t - 240°)$.

Sketch $v_{out}(t)$, and give an expression for the ripple voltage when the load has parallel capacitance C.

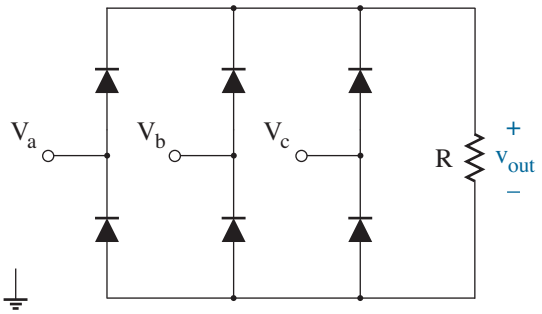


Figure P4.12

4.13 Design an LC-filtered full-wave rectifier circuit to deliver 2 A to a 20-Ω load with 2-% ripple (max), and specify the diode ratings. Assume a 60-Hz 115-V (rms) power source. Use SPICE to verify the design.

4.14 Consider the LCR circuit of Fig. P4.14.

(a) Show that the circuit acts as a low-pass filter with an exact transfer characteristic of the form

$$H(s) = \frac{v_{out}}{v_{in}} = \frac{\omega_o^2}{s^2 + s\omega_o/Q + \omega_o^2} ,$$

and give expressions for ω_o and Q.

(b) Let $L = 10$ μH. Complete a design with $f_o = \omega_o/2\pi = 100$ kHz and $Q = 5$.

(c) Use SPICE to simulate the ac response between 1 kHz and 10 MHz.

(d) Use SPICE to simulate the transient response to a 1-V step input.

(e) Show how the transient response changes if L/R and RC increase separately by 25 %.

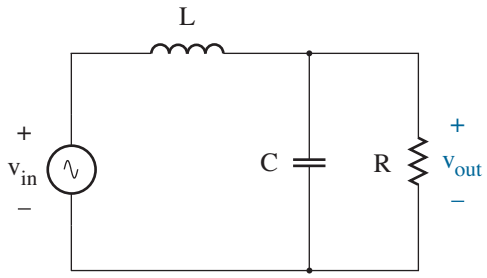


Figure P4.14

4.15 Design a buck converter that delivers 300 mA at 3.3 V from a 9-V dc source with a ripple voltage of 0.1 V or less. Assume 100-kHz switching frequency. Be sure to specify the capacitor ESR.

4.16 Design a buck converter that delivers 200 mA at 6.2 V from a 9-V dc source with a ripple voltage of 0.2 V or less. Assume 180-kHz switching frequency. Be sure to specify the capacitor ESR.

4.17 Consider an "ideal" buck converter circuit whose capacitor has zero ESR. Determine an expression for the output ripple voltage.

Hint: Consider the RC parallel impedance.

4.18 Consider a buck converter that operates in the *discontinuous* mode with the inductor current falling to zero over a portion of the switching period. Let L have 1/2 the value specified by Eq. 4.43.

(a) Sketch v_{in} as a function of time t (see Fig. 4.12).

(b) Determine an expression for $< v_{out} >$.

(c) Use SPICE to examine v_{out} given a 200-kHz switching frequency, $D = 0.5$, $V_o = 6$ V, and 50-mA load current. Assume $C = 1$ μF.

(d) Use SPICE to examine the inductor current for the conditions of part c.

(e) Discuss the efficiency of the discontinuous mode in relation to the continuous mode of operation.

Note: The $< v_{out} >$ expression is more complicated than that for the case of continuous inductor current. Nevertheless, appropriate feedback can be used to control $< v_{out} >$ through duty-cycle variations.

4.19 Design a boost converter that delivers 80 mA at 10 V from a 6-V dc source with a ripple voltage of 0.25 V or less. Assume 100-kHz switching frequency. Be sure to specify the capacitor ESR.

4.20 Design a boost converter that delivers 50 mA at 15 V from a 6-V dc source with a ripple voltage of 0.1 V or less. Assume 220-kHz switching frequency. Be sure to specify the capacitor ESR.

4.21 Verify Eq. 4.54:

$$\frac{1}{2}\left(i_{L,max} + i_{L,min} \right) = \frac{V_o}{R(1-D)^2} .$$

Hint: Examine the power dissipated in the load in relation to the rate of energy that is supplied to it.

4.22 Consider the boost converter of Example 4.5.

(a) Use SPICE to simulate the transient response, and explain the results.

(b) Show the change effected when L is doubled.

4.23

(a) Design a buck converter that delivers 400 mA at 5.8 V from a 9-V dc source. Use the MAX887 controller at 300 kHz. Specify pin connections.

(b) Use the appropriate typical operating characteristic supplied by the manufacturer to estimate the operating efficiency of your design.

(c) All of the external components for your design should be surface-mount parts for a PC board. Provide component specifications and ordering information from Digikey (www.digikey.com), and indicate total cost (assuming component purchases in lots of 1000).

4.24

(a) Design a boost converter that delivers 60 mA at 8 V from a 4.5-V dc source. Use the MAX1790 controller at 640 kHz. Specify pin connections.

(b) Use the appropriate typical operating characteristic supplied by the manufacturer to estimate the operating efficiency of your design.

(c) Perform a cost analysis (as in Problem 4.23).

4.25 Consider the PWM process for buck converter control in Fig. 4.16.

(a) Determine an expression for the duty cycle D in terms of the error voltage v_e and the ramp amplitude V_m.

(b) Subject to $v_{out} = DV_o$, find expressions for v_{out} and v_e in the limit of large gain-factor K.

4.26 Figure P4.26 shows the magnetic flux linkage λ that applies to an inductor with current i and a particular core material. The inductor voltage is

$$v = \frac{d\lambda}{dt}\,.$$

Show that stored magnetic energy is proportional to the shaded area in the figure (and thus the effective inductance is limited at large currents).

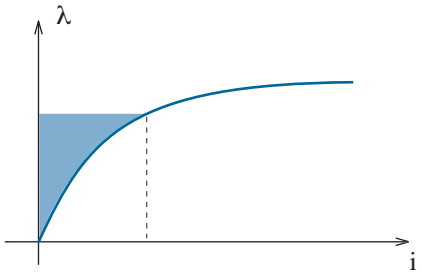


Figure P4.26

4.27 Consider the current-mode PWM process for buck converter control.

(a) Suppose the inductor-current-related v_s waveform undergoes an abrupt perturbation by Δv_s as in Fig. P4.27a. Show (graphically) that the system eventually returns to normal operation if $D < 0.5$ and that it is otherwise unstable.

(b) Figure P4.27b shows an example of **slope compensation** in which a small negative ramp adds to the error voltage v_e to prevent the instability of part a. Find the minimum magnitude of the ramp slope that is required for some $D > 0.5$.

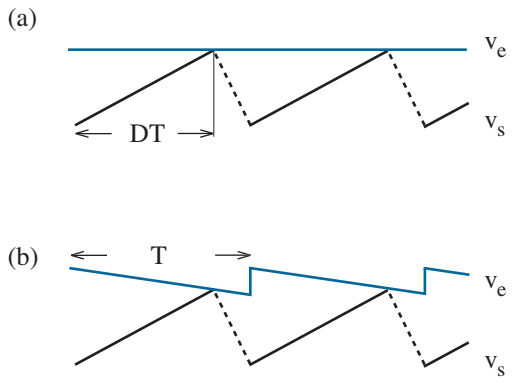


Figure P4.27

4.28 Figure P4.28 shows the approximate time dependence of the current and voltage for a switch used in a buck or boost converter. As the switch closes, the current ramps between zero and I_x, and the voltage ramps between V_x and zero. As the switch opens, the current and voltage ramp in reverse. Both ramps occur over time intervals $[0, T_s]$.

(a) Show that the time-averaged power dissipated in the switch is

$$< P > = \frac{V_x I_x}{3} \left(\frac{T_s}{T} \right) ,$$

where T is the converter switching period.

(b) Discuss the result of part a in terms of selecting an optimum switching frequency.

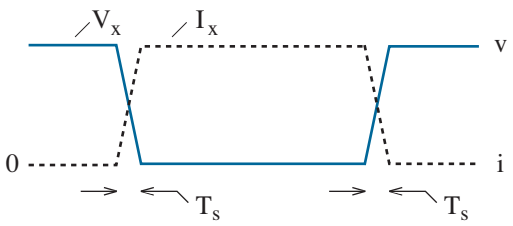


Figure P4.28

4.29 Figure P4.29 shows a simple **buck-boost** dc-to-dc converter. The switch operates with period T and duty cycle D. Inductor current is continuous, and the diode is ideal.

(a) Make a sketch of v_{in} vs. time. Assume constant $v_{out} \approx < v_{out} >$.

(b) Show that

$$< v_{out} > = \frac{-DV_o}{1-D} .$$

Hint: Consider the value of the time-averaged inductor voltage.

(c) Determine the inductor current over the course of one time period.

(d) Determine a design condition for continuous inductor current.

(e) Specify design criteria for capacitor C.

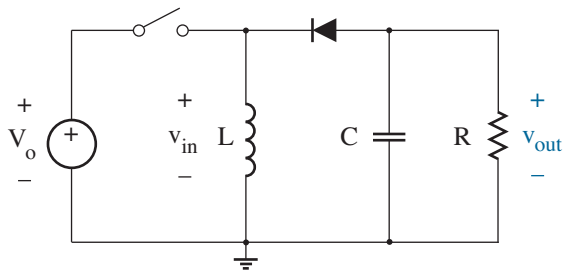


Figure P4.29

4.30 A **SEPIC** (Single-Ended Primary Inductance Converter) circuit has the form shown in Fig. P4.30. The switch operates with time period T and duty cycle D, the inductor currents are continuous, and the diode is ideal.

(a) Determine the time average of voltage v_{c1}.

(b) Make a sketch of v_x vs. time. Assume constant $v_{out} \approx < v_{out} >$.

(c) Show that

$$< v_{out} > = \frac{-DV_o}{1-D} .$$

(d) Determine design conditions that promote continuous L_1 and L_2 inductor currents.

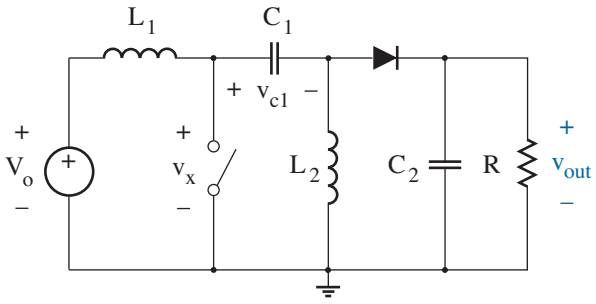


Figure P4.30

Section 4.2

4.31 The circuit of Fig. P4.31 has a Zener diode with 5-V breakdown, 1.5-mA "knee" current, and 1-W power rating.

(a) Determine the minimum V_o value that ensures perpetual Zener breakdown.

(b) If V_o has the value of part a, find the maximum *instantaneous* power that the diode absorbs.

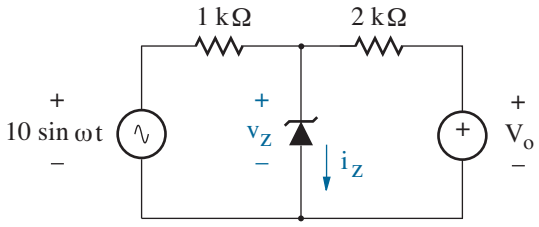


Figure P4.31

4.32 The circuit of Fig. P4.32 has a Zener diode with 12-V breakdown, 3-mA "knee" current, and 1-W power rating.

(a) Determine the maximum I_o value that ensures perpetual Zener breakdown.

(b) If I_o has the value of part (a), find the maximum *instantaneous* power that the diode absorbs.

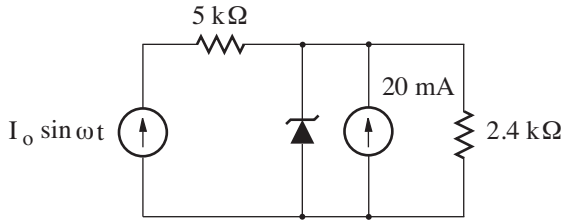


Figure P4.32

4.33 The circuit of Fig. P4.33 has a Zener diode with 9-V breakdown, 6-mA "knee" current, and 1-W power rating. Asuume 500 Ω $< R <$ 1500 Ω.

(a) Determine R_s such that $v_{out} = 9$ V for any R in the specified range.

(b) Determine the ripple voltage if $R_z = 25$ Ω.

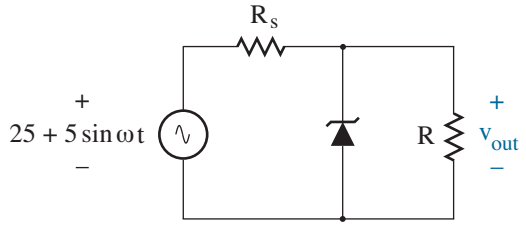


Figure P4.33

4.34 The Zener diode in the circuit of Fig. P4.34 has 6-V reverse breakdown. Find R such that $i = 1$ mA.

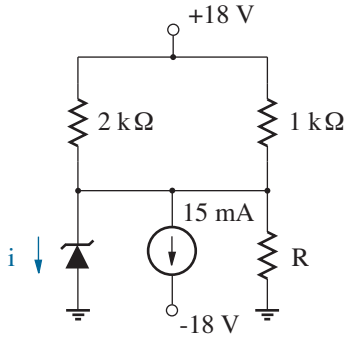


Figure P4.34

4.35 Design a Zener-regulated power supply to provide a steady 12-V output from a 12-V (rms) 60-Hz ac source that is half-wave rectified and RC filtered to within a 5-% ripple voltage. The power-supply load–current requirement ranges from 20 to 50 mA. Specify percent regulation if $R_z = 20$ Ω.

4.36 Design a Zener-regulated power supply to provide a steady 24-V output from a 30-V (rms) 60-Hz ac source that is full-wave rectified and RC filtered to within a 15-% ripple voltage. The power-supply load-current requirement ranges from 5 to 20 mA. Specify percent regulation if $R_z = 80\ \Omega$.

4.37 The circuit of Fig. P4.31 has $V_o = 50$ V.

(a) Use SPICE to evaluate i_z.

(b) Use SPICE to evaluate v_z if $R_z = 10\ \Omega$.

4.38 A 10-V Zener diode absorbs 5 mA in breakdown with no parallel load and -40 °C $\leq T \leq$ +85 °C.

(a) Use the data of Fig. 4.27 to develop a SPICE macromodel with consistent temperature behavior for V_z. Demonstrate your results.

(b) Repeat the preceding simulation, but include a forward-biased pn-junction diode in series with the Zener macromodel. Let $I_s = 1 \times 10^{-14}$ A. Explain the observed behavior.

(c) Find a value for V_z that yields a reference voltage with minimum temperature variation when the Zener diode is used as in part b.

4.39 Design a Zener reference like that of Fig. 4.26 subject to $v_{out} = 12$ V, $i_z = 1$ mA, and $v_{in} = 15$ V. Consider designs with $V_z = 4$ V and $V_z = 6$ V in conjunction with the data of Fig. 4.27. Simulate with SPICE, and specify Δv_{out} if v_{in} increases by 100 mV.

4.40 The diode currents in the bandgap reference circuit of Fig. 4.28 exhibit the following dependence on voltage and temperature:

$$i = I_o{}' \left(\frac{T}{T_o}\right)^{\eta} \exp\left(\frac{-E_g}{kT}\right) \exp\left(\frac{qV}{kT}\right),$$

where $\eta = 3.6$ and $E_g = 1.12$ eV. Let $T_o = 300$ K.

(a) Determine $I_o{}'$ if $I_s = 2 \times 10^{-16}$ A.

(b) Determine an expression for v_{d1}, then expand it as a Taylor series in powers of $(T - T_o)$ up to second order. Regroup terms in powers of T.

(c) Let $R_2 = 5$ kΩ and $R_3 = 1$ kΩ. Determine R_1 such that v_{out} has no linear variation with T.

(d) Find the consistent bandgap reference voltage.

(e) Use SPICE to plot the bandgap reference voltage from -40 °C to +85 °C. Use IS = 0.2f, XTI = 3.6, and EG = 1.12 in the .model statement.

4.41 Show that the circuit of Fig. 4.28 can function as a bandgap reference if $R_1 = R_2$ and the diodes have different areas with $A_1/A_2 = X$. How should X compare to unity if the circuit is to provide temperature compensation?

4.42 The bandgap reference circuit of Fig. P4.42 contains a PTAT current (i_1) and a PTAT^2 current (i_x) proportional to absolute temperature squared. The latter is available in an integrated circuit.

(a) Find an expression for the reference voltage v_{out} in terms of v_{d1}, i_1, and i_x.

(b) Discuss a design procedure that minimizes the temperature dependence of the reference voltage to second order in T.

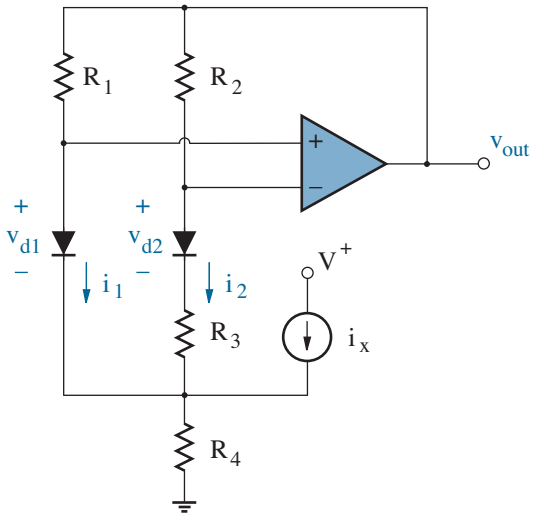


Figure P4.42

Section 4.3

4.43 Use appropriate clipping circuits to develop an op-amp model for SPICE so that the output voltage can never extend beyond ±2.5 V (see Fig. 1.40). The differential gain is 10^6. Show the effectiveness of the model in a simple comparator circuit.

4.44 Design a circuit that yields the transfer characteristic of Fig. P4.44. Assume ideal diode behavior.

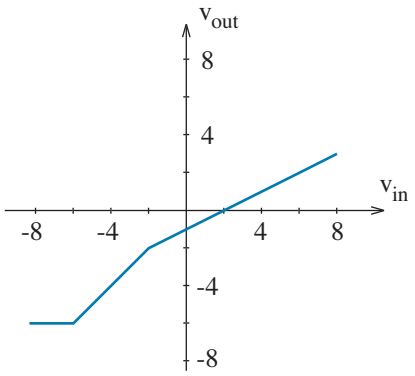


Figure P4.44

4.45 Design a circuit that yields the transfer characteristic of Fig. P4.45. Assume ideal diode behavior.

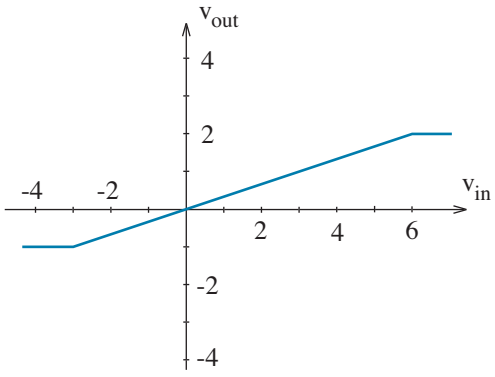


Figure P4.45

4.46 Design a three-diode circuit that yields $v_{out} \approx {v_{in}}^{2/3}$ for $0 < v_{out} < 10$ V. Assume ideal diodes.

4.47 In the circuit of Fig. P4.47, Zener diodes D_1 and D_2 exhibit 6-V and 8-V breakdown, respectively. Sketch v_{out} vs. v_{in} subject to $-10 < v_{in} < 10$ V.

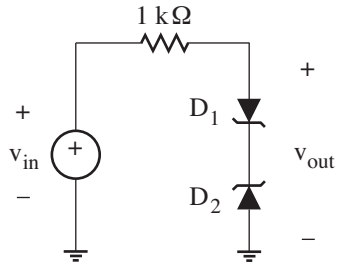


Figure P4.47

4.48 Design a circuit that yields the transfer characteristic of Fig. P4.48. Assume ideal diode behavior.

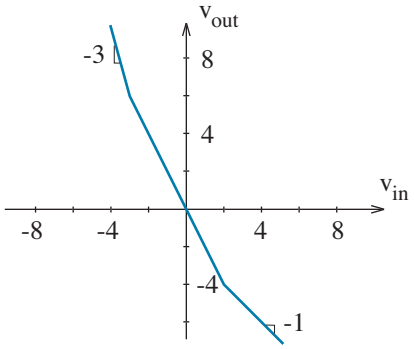


Figure P4.48

4.49 Design a circuit that yields the transfer characteristic of Fig. P4.49. Assume ideal diode behavior.

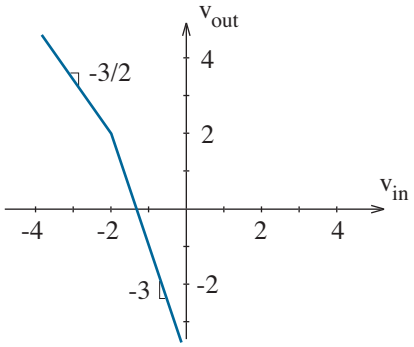


Figure P4.49

4.50 Design a three-diode circuit that yields $v_{out} \approx {v_{in}}^{3/2}$ for $0 < v_{out} < 10$ V. Assume ideal diodes.

4.51 The circuit within the dashed lines in Fig. P4.51 uses op-amp feedback to implement a "superdiode" with near-ideal behavior.

(a) Show that the voltage v is given by

$$v = \frac{1}{A_{vd}} \frac{kT}{q} \ln\left(\frac{i}{I_s}\right),$$

where A_{vd} is the differential voltage gain for the op-amp and I_s is the diode saturaton current.

(b) Use SPICE to demonstrate superdiode action.

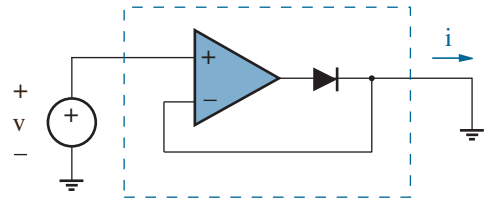


Figure P4.51

4.52 Figure P4.52 shows a popular "fuzz" circuit that adds distortion to the output signal from an electric guitar. The diodes feature $I_s = 10^{-14}$ A. Use SPICE to show distortion effects on sinusoidal inputs with 40-mV, 80-mV, and 120-mV amplitudes. Explain your results. Hint: Use a .tran simulation.

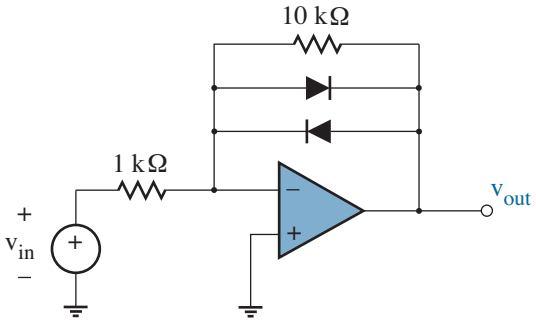


Figure P4.52

4.53 In the circuit of Fig. P4.53, the ordinary diodes and the op-amp are ideal, and the Zener diode has $V_z = 8$ V.

Sketch the output voltage waveform if v_{in} takes the form of a triangle wave that varies between ±5 V.

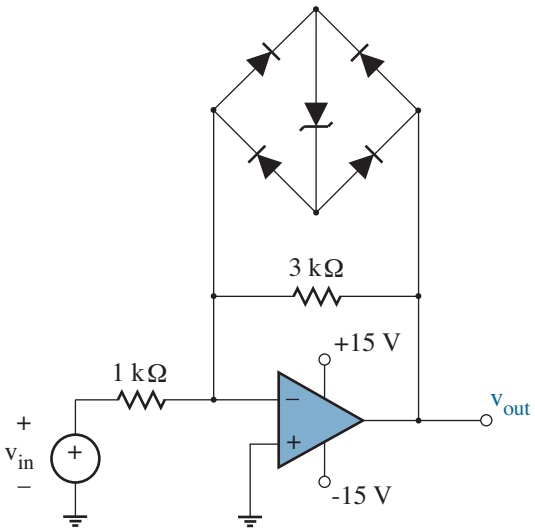


Figure P4.53

4.54 You have a function generator that can produce a square wave that moves between 0 and +5 V at varying frequency. Nevertheless, you require a single *pulse* that transitions from +5 V to 0 and back with varying frequency and a duration less than 10 μs.

Design a circuit that produces the pulse waveform using a +5-V power supply, a 100-nF capacitor, and other circuit elements as needed. Verify your design with SPICE.

4.55 A 15-V power supply operates a particular load resistance, and a 12-V backup battery operates the same load under conditions of ac power failure. Design a simple circuit that meets these objectives.

4.56 The circuit of Fig. P4.56 features a so-called **snubber** diode that protects the switch from an arc that would otherwise result when abruptly cutting off current from a large inductive load.

Let $L = 100$ mH subject to 0.5-Ω series resistance. For the diode, $I_s = 0.1$ pA and $C_j(0) = 200$ pF.

(a) Use SPICE to observe the inductor current after opening the switch.

(b) Use SPICE to observe the inductor current after closing the switch.

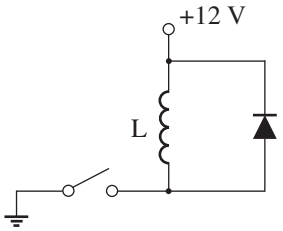


Figure P4.56

4.57 For the circuit of Fig. P4.57, sketch $v_{out}(t)$ if $v_{in}(t) = 5\sin(\omega t)$. Assume steady-state conditions. The diode is ideal.

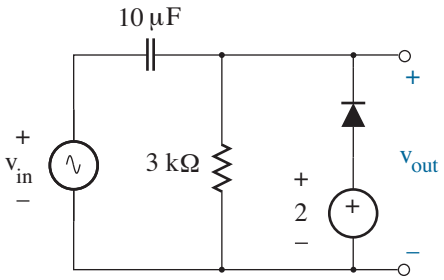


Figure P4.57

4.58 Use a clamp circuit and appropriate "filtering" to design a **voltage doubler** that gives a constant 30-V output from a sinusoidal input signal with 15-V amplitude at 60 Hz. The minimum-R load is 100 kΩ. Use SPICE to demonstrate your design.

4.59 The circuit of Fig. P4.59 has two ideal diodes. Sketch $v_1(t)$ and $v_2(t)$ when $v_{in}(t) = 10\sin(100t)$. Assume steady-state conditions.

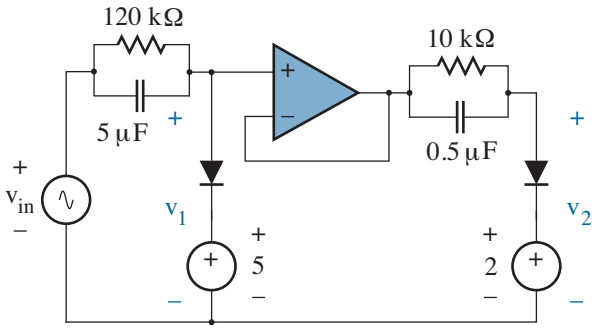


Figure P4.59

4.60 Consider the 4-ideal-diode circuit of Fig. P4.60. Sketch the output waveform if $v_{in}(t) = 10\sin(100t)$. The capacitors are uncharged for $t < 0$.

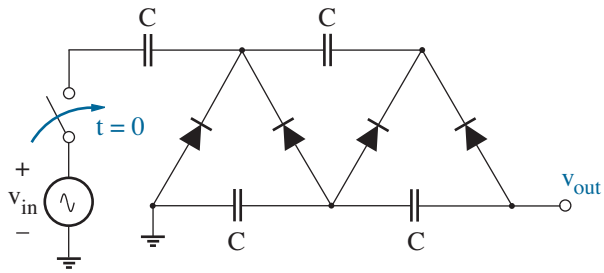


Figure P4.60

4.61 Consider the video waveform of Fig. 4.42c. After ac capacitive coupling, the pulse portion of the signal is known to extend exclusively below -2 V.

(a) Describe a circuit that clamps the waveform to zero using only a capacitor, a comparator, and a switch —no diode(s). The integrated-circuit switch is closed when the comparator output is HIGH, and it is open otherwise.

(b) Discuss the advantages and disadvantages of the preceding design in relation to a diode clamp.

Section 4.4

4.62 Download the data sheet for the MAX16803 (www.maxim-ic.com), then design an automotive LED tail-light controller with the following features:

1. The tail light has maximum brightness when the brake pedal is depressed.
2. The tail light has 60% brightness when the headlights are on.
3. The tail light is otherwise off.

Hint: A 555 or similar IC timer may be helpful.

4.63 The op-amp in the photodetector circuit of Fig. 4.47 is non-ideal with input offset voltage v_{os}. Determine the apparent "dark" photocurrent when no light is present.

4.64 The op-amp in the photodetector circuit of Fig. 4.47 is non-ideal with input currents i_{bias}. Determine the apparent "dark" photocurrent when no light is present.

4.65 Design an IL300 optocoupler (www.vishay.com) with unity voltage gain.

Perspective:
Current Control Valves

The previous Perspective featured the ideal diode as a two-terminal nonlinear circuit element that establishes a Q-point at the zero-voltage or zero-current end position along a particular load line. When the load line is fixed, the Q-point is changed by removing the diode from its circuit environment and then reconnecting it with inverted orientation, as indicated in Fig. B1. A cumbersome *mechanical* process is hardly worthy of an electrical engineer. So for most applications, the diode position is fixed. The Q-point is changed by shifting the load line through variations of the Thevenin voltage V_t.

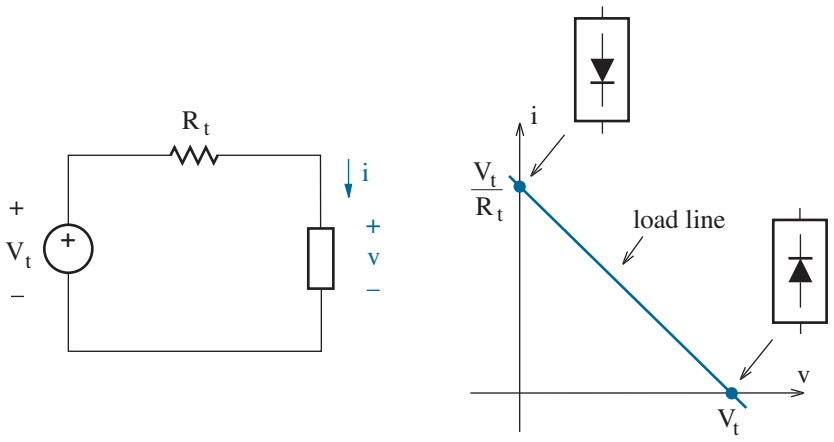


Figure B1: Q-point positions on a fixed load line subject to two ideal-diode orientations in a simple circuit.

If we wish to achieve *electrical* control of the Q-point position on a fixed load line, we need to implement a new type of circuit element in which the current-voltage characteristic relative to two "mainline" terminals depends upon the electrical action at a third (and possibly a fourth) control terminal. Such is the fundamental objective of electronics.

As an example, consider the graphical analysis of the circuit of Fig. B2. Here, the mainline terminals of a new device are labeled **source** and **drain**, and a single control terminal is identified as the **gate**. (These designations imply *electron* flow from source to drain through the gate, yielding positive drain current i_d under a condition of positive drain-to-source voltage v_{ds}.) A constant-V^+ Thevenin subcircuit, which is connected to the source and drain terminals, imposes a fixed load-line constraint between i_d and v_{ds}. Meanwhile, a separate, variable-V_g Thevenin subcircuit induces a distinct gate current i_g or gate-to-source voltage v_{gs}. In turn, the device selects an accommodated constraint from a *family* of characteristic curves that relate i_d and v_{ds}. Figure B3 shows the graphical results of this process.

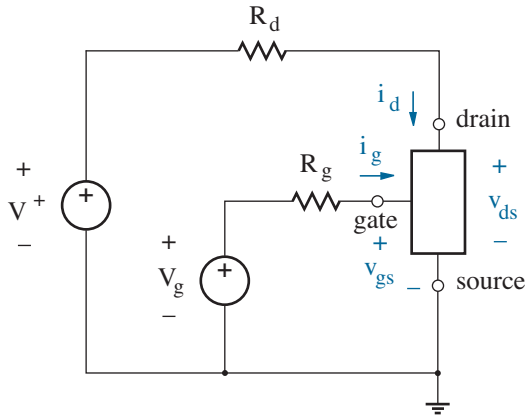


Figure B2: Demonstration circuit for controlled drain current.

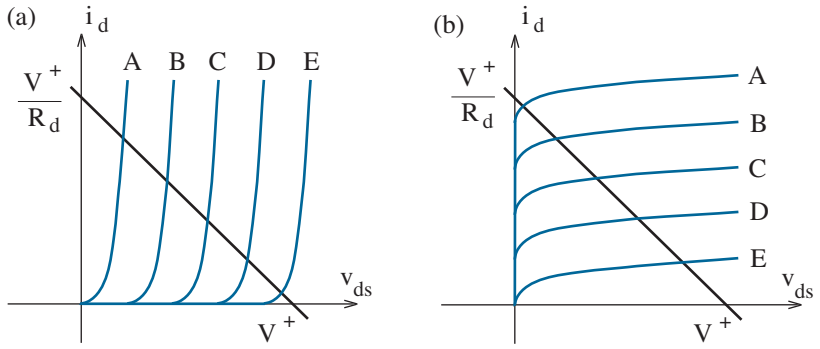


Figure B3: Graphical solutions for the circuit of Fig. B2. The characteristic curves are labelled A through E for five distinct i_g or v_{gs} valve-control levels: (a) dependent voltage-source control; (b) dependent current-source control. The Q-point is the point of intersection with the load line.

The graphical circuit solutions shown in Fig. B3 have two different sets of hypothetical device characteristic curves that suggest the action of a dependent voltage source (Fig. B3a) or dependent current source (Fig. B3b). There are limitless other possibilities. But in each case, the form of Q-point control indicates valve-like regulation of drain current i_d. The valve is wide open at the zero-voltage load-line extreme, and it is completely closed at the zero-current load-line extreme. These conditions apply to an ideal switch. An ideal valve allows any intermediate drain-current condition.

Before we demand some particular form of valve control, we need to be aware of a practical limitation: Modern solid-state **transistor** valves tend to exhibit the characteristic behavior shown in Fig. B3b. Thus, we examine the transistor as a physical approximation to a dependent current source. In contrast, outdated electronic devices, such as the vacuum-tube triode (a British "valve"), tend to exhibit the characteristic behavior of Fig. B3a. This once motivated a different mode of study.

Numerous physical processes have been exploited to yield transistors with imperfect realization of valve action.

Imperfect! Just what are we looking for?

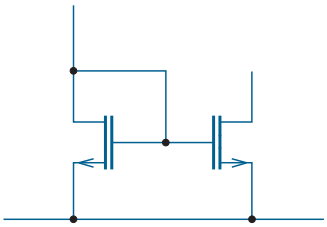
Some answers to this question will emerge as we examine the conflicting demands of modern circuit applications beginning in Chapter 7 and beyond. For now, we need to understand the structure and operation of two types of transistor that have earned large portions of ever-shifting electronic "turf." Specifically, we consider:

- The MOS field-effect transistor (Chapter 5);
- The bipolar junction transistor (Chapter 6).

Less common transistors such as the MESFET are left to other texts.

The physical models offered in Chapters 5 and 6 are somewhat complex. And even at this level, they are inadequate to cover every behavioral detail, especially when dimensions are very small. Nevertheless, the models will prove useful in support of pencil-and-paper design calculations that place theoretical circuit operation within the "ballpark" of actual performance. Both chapters offer selected circuit applications to stimulate interest.

> Warning: Two conflicting sets of compelling arguments suggest optimum discussion when one type of transistor is treated first. Chapters 5 and 6 are independent. Read them in either order.



Chapter 5

The MOS Field-Effect Transistor

The MOS field-effect transistor or **MOSFET** derives its name from a metal / silicon dioxide / silicon substructure used in a primitive technology to regulate source-to-drain carrier flow. Conceptually oldest of transistors that will concern us, the MOSFET was patented by J. E. Lilienfeld in 1930. But nearly forty years of sporadic development would be required before it could claim commercial success. Today, the MOSFET leads the way toward the most compact integrated-circuit applications.

We begin by examining the physical principles of MOSFET operation. Our goal is to derive static (time-independent) interterminal relationships and identify key device parameters that are easily measured. These lead to large-signal models for the analysis of some elementary circuits.

Objectives

After studying this chapter and working through the problems, you should be able to perform each of the following:

- Describe accumulation, depletion, and inversion conditions in relation to an MOS capacitor (Section 5.1).
- Determine the "on" requirements for enhancement- and depletion-mode MOSFETs with n and p channels (Section 5.1).
- Identify the parameters $K'W/L$ and V_T. Then sketch the MOSFET characteristic curves that derive from them (Section 5.1).
- Specify the "on" resistance of a MOSFET switch (Section 5.2).
- Analyze simple dc MOSFET circuits using an appropriate large-signal model or SPICE (Sections 5.3 and 5.4).

5.1 Operating Principles

For the moment, our objective is to construct a switch in a silicon substrate. The familiar pn junction is a first building block of interest, since it allows the formation of electron- or hole-rich contacts that are electrically isolated from one another when the switch is open. Figure 5.1 shows n^+ **source** and **drain** regions for two prospective switches that share a p-type substrate (as for an integrated circuit). Isolation is assured if the substrate or **body** potential is never greater than the node voltage at any source or drain.

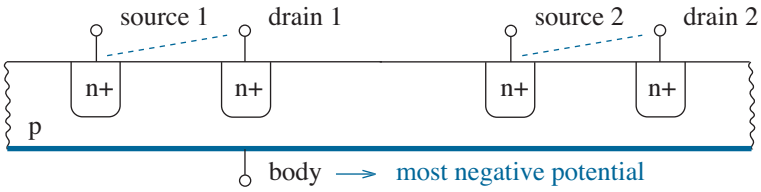


Figure 5.1: n^+ switch contacts in a p-type substrate.

Alternatively, we can make p^+ source and drain regions as shown in Fig. 5.2. Here, the body potential must never be less than any other node voltage.

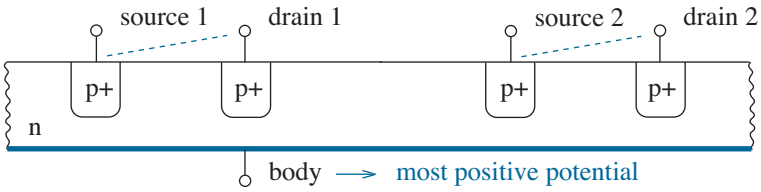


Figure 5.2: p^+ switch contacts in an n-type substrate.

We can even have coexisting n^+ and p^+ contact pairs as shown in Fig. 5.3.

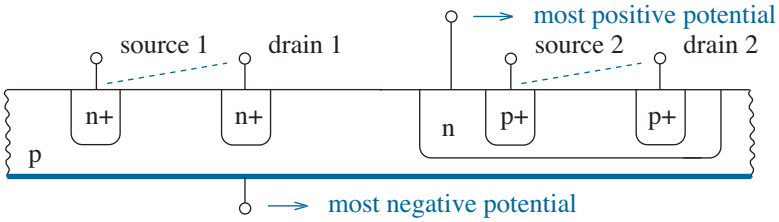


Figure 5.3: n^+ contacts with p^+ contacts in an adjacent n-well.

But all of these prospective switches are perpetually open.

To close a switch, we require a different type of building block that forms (and later removes) an n- or p-type electrical bridge between contacts.

The MOS Capacitor

In our quest for electronic building blocks, we examine a simple device that exhibits a new basis for conductivity control. This is the MOS capacitor, shown with typical form in Fig. 5.4. Here, the substrate is p-type silicon with doping concentration N_a, the insulator is silicon dioxide (SiO_2) with thickness t_{ox}, and the metal plate is aluminum (Al). To test this device, we apply voltage v_{gb} between the upper **gate** and lower body terminals. What do we observe?

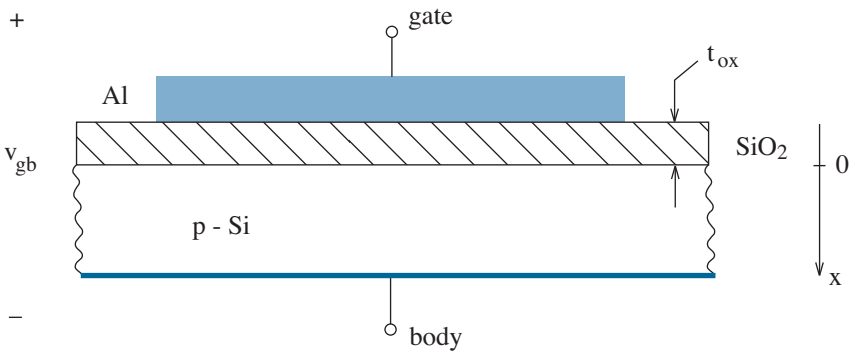


Figure 5.4: Typical MOS capacitor structure.

Not much—at least at first glance. The SiO_2 insulator beneath the gate simply blocks any dc gate-terminal current, just like an ordinary capacitor. Yet, on second thought, we witness some unusual behavior because of the presence of the v_{gb}-induced electric field. If $v_{gb} < 0$, holes are attracted to the insulator/substrate interface, and the interface is **accumulated** with excess holes. This bears little interest, since the interface is already p-type. On the other hand, if $v_{gb} > 0$, holes are repelled from the insulator/substrate interface, and the substrate near the interface becomes locally **depleted**. By adjusting positive v_{gb}, we can control the extent of the depletion region. Meanwhile, $v_{gb} > 0$ tends to attract what few electrons are around to the insulator/substrate interface—the population rises in the manner of Eq. 2.27. Indeed, if v_{gb} is greater than a certain positive **threshold voltage** (V_T), the interface electron concentration is greater than that for substrate holes, and the interface is **inverted** or effectively n-type. Both the depletion and especially the inversion conditions can be exploited for the realization of transistor switching action.

We note the opposite behavior when the MOS capacitor features an n-type substrate. The unexciting accumulation condition applies for $v_{gb} > 0$ (electrons are attracted to the insulator/substrate interface). The depletion and inversion conditions apply for $v_{gb} < 0$ and $v_{gb} < V_T < 0$, respectively (electrons are repelled from the interface, holes are attracted there).

Caution: The preceding discussion assumed that the electric field in the vicinity of the insulator/substrate interface has zero value when $v_{gb} = 0$. In practice, different electrical properties of the gate and body materials as well as the presence of charge at the generally imperfect interface give rise to a non-zero electric field when $v_{gb} = 0$. To achieve zero electric field, we require $v_{gb} = V_{fb}$, where V_{fb} is the characteristic **flat-band voltage** for the MOS capacitor. Cautions aside, the MOS threshold voltage retains its physical significance, as it includes the effect of the flat-band voltage.

Exercise 5.1 The MOS capacitors of Fig. 5.5 all feature $|V_T| = 2$ V. Specify whether the capacitors are under accumulation, depletion without inversion, or inversion. Assume $V_{fb} = 0$.

Ans: (a) depletion without inversion (d) inversion
(b) accumulation (e) inversion
(c) accumulation (f) depletion without inversion

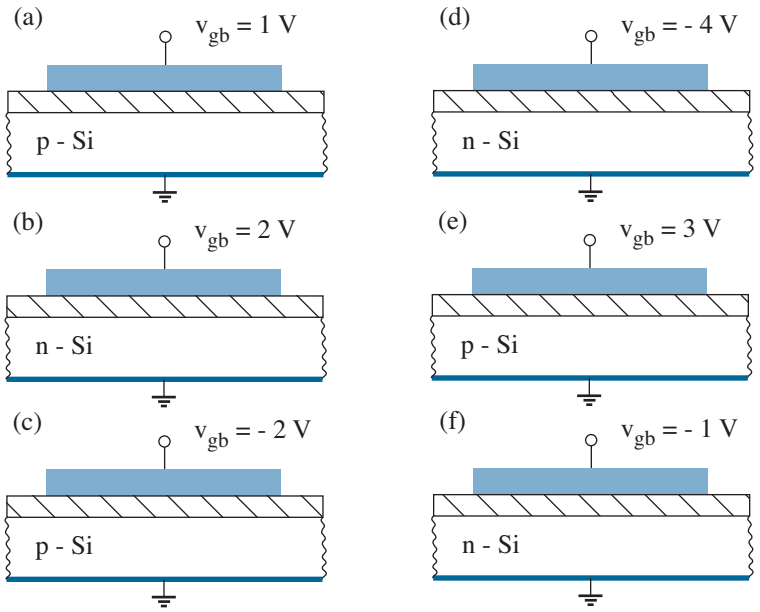


Figure 5.5: MOS capacitors for Exercise 5.1.

The chapter Appendix provides a more quantitative description of the MOS capacitor and the factors that affect the threshold voltage.

The MOSFET Family

We are now prepared to examine specific MOSFET devices, which are made available in several different flavors to support a wide range of applications.

Enhancement-Mode MOSFET

The n-channel **enhancement-mode** MOSFET derives from a p-substrate MOS capacitor and a pair of n^+ source and drain regions near the gate as shown in Fig. 5.6a. An enhancement designation requires a *positive* threshold voltage V_T (now referenced to the source). Thus, under "normal" conditions when $v_{gs} = 0$, the source and drain are electrically isolated by an intervening p region, and the transistor is "off" (with $i_d = 0$ for $v_{ds} > 0$). However, under conditions when $v_{gs} > V_T$, an electron-rich inversion layer is present at the insulator/substrate interface beneath the gate, and the transistor is "on" (with $i_d > 0$ for $v_{ds} > 0$). There is no dc gate current.

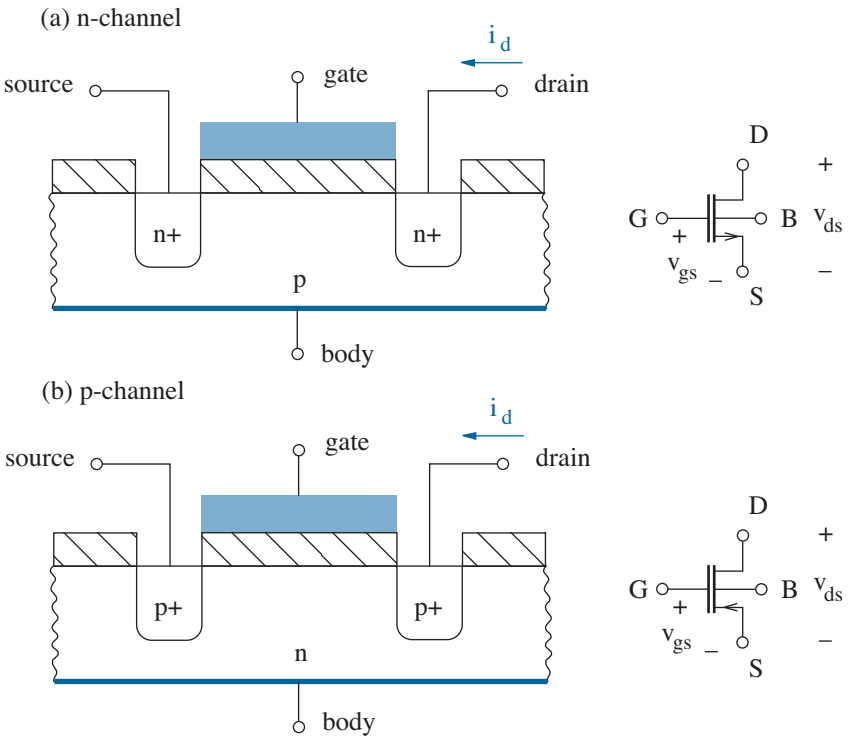


Figure 5.6: Enhancement-mode MOSFET structures and corresponding device symbols: (a) n-channel; (b) p-channel.

Similarly, a p-channel enhancement-mode MOSFET derives from an n-substrate MOS capacitor and a pair of p^+ source and drain regions near the gate as shown in Fig. 5.6b. This device features a *negative* threshold voltage, and it is normally "off." To achieve the "on" condition, we require $v_{gs} < V_T$ (with $i_d < 0$ for $v_{ds} < 0$).

Our preferred symbols for the enhancement-mode MOSFET suggest a gate that is insulated from the substrate, and both feature a distinguishing directed arrow along the source terminal connection.[1] The chosen symbols properly convey the MOSFET as a four-terminal device—the drain current i_d is derived from a particular set of v_{gs}, v_{ds}, and v_{bs} values. Nevertheless, we will tend not to depict the body terminal. Apart from reducing circuit-schematic clutter, this acknowledges a common situation in which $v_{bs} = 0$ (just about always for discrete MOSFETs, and often in integrated circuits). An implied body-to-source connection simplifies MOSFET circuit analysis, and it facilitates comparisons with three-terminal transistors yet examined. Chapters 9 and 10 reveal integrated-circuit applications with non-zero v_{bs}. The performance adjustments will not be welcome.

> When considering MOSFET characteristic relations, we focus on the n-channel device configuration. We need say little about its p-channel counterpart, except to require similar principles of operation when *all* terminal currents and voltages are altered by a change in sign. Familiarity easily derives from practice.

As noted, the n-channel enhancement-mode MOSFET controls source-to-drain flow of electrons under conditions when $v_{gs} > V_T$ and $v_{ds} > 0$ to produce $i_d > 0$. If v_{ds} is small, the electron concentration in the inversion layer is nearly proportional to $v_{gs} - V_T$, and it is almost uniform between the source and drain. Thus, we observe a v_{gs}-dependent channel conductance—the larger the better for a good switch.

Now consider the MOSFET behavior as v_{ds} increases. Fig. 5.7 shows channel conditions that apply for three different v_{ds} values and $v_{gs} > V_T$. As v_{ds} becomes moderately large, the electron concentration in the inversion layer decreases as one moves from the source to the drain (Fig. 5.7a). This is apparent from the relation $v_{gd} = v_{gs} - v_{ds}$. The tapered electron concentration implies reduced channel conductance, so i_d increases with v_{ds} at a sub-linear rate. At larger $v_{ds} = v_{gs} - V_T$ (such that $v_{gd} = V_T$), the inversion-layer electron concentration is essentially reduced to zero at point X next to the drain (Fig. 5.7b). This so-called "pinch-off" condition corresponds to the onset of drain-current **saturation** for which $i_d = i_{d,sat}$ and $v_{ds} = v_{ds,sat}$. Further increases in v_{ds} simply move the pinch-off point X closer to the source (Fig. 5.7c), and i_d remains at $i_{d,sat}$.

[1] The MOSFET symbols in this text differ from certain presumptuous "standards." You will tend to find different symbols in use as you vary your technical environment.

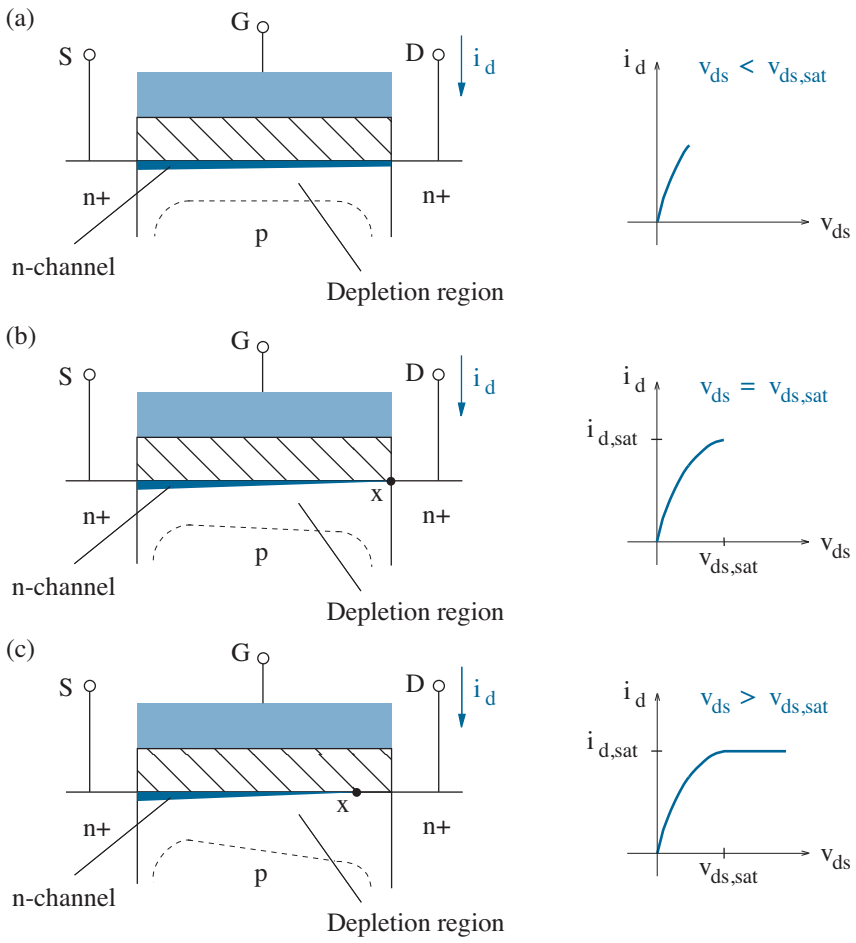


Figure 5.7: n-channel enhancement-mode MOSFET cross-sections and current-voltage characteristics for $v_{gs} > V_T$ and variable v_{ds}.

It may seem odd to observe finite saturated drain current under channel pinch-off conditions. Nevertheless, the potential at point X is $v_{ds,sat}$, so the unblocked channel to the left of X sustains an electric field that injects electrons into the right-side depletion region. And once injected, the electrons are swept to the drain by a strong internal electric field.

As $v_{gs} - V_T$ becomes larger, the overall electron charge in the inversion layer increases prior to v_{ds}-induced tapering. Thus, the onset of saturation becomes more difficult, and the drain current within saturation increases. In this respect, the MOSFET looks like a v_{gs}-dependent current source (with constant i_d provided v_{ds} is sufficiently large to ensure saturation). Here is the more general "valve" action that we are looking for.

MOSFET Current-Voltage Characteristics

A quantitative derivation of the current-voltage characteristics for the n-channel enhancement-mode MOSFET assumes the geometry of Fig. 5.8. We make the x direction perpendicular to the insulator/substrate interface subject to $x = 0$ at the interface. The y direction is parallel to the interface, with $y = 0$ at the source and $y = L$ at the drain. The channel width is W. Further complications notwithstanding, we let $v_{bs} = 0$.

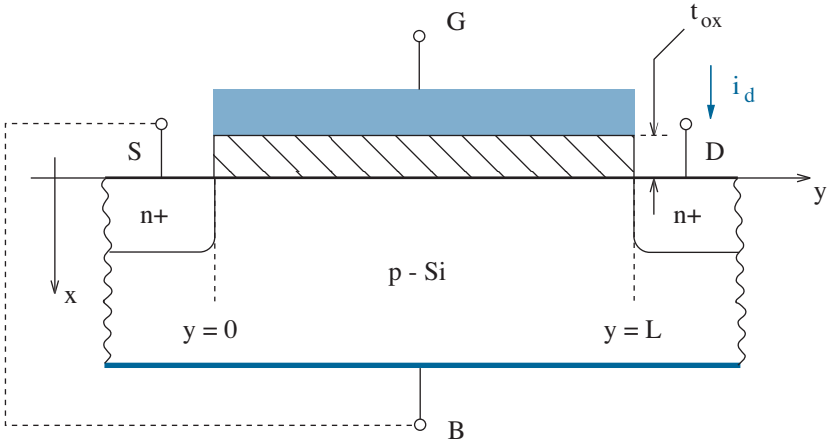


Figure 5.8: n-channel enhancement-mode MOSFET geometry for analysis. The channel region extends into the page by distance W.

In what follows, we assume a **long-channel** device for which the x-directed gate-induced electric field is significantly larger than the y-directed v_{ds}-induced electric field. Further, we assume a fixed level of depletion-layer charge when determining the gate-induced charge in the inversion layer—this yields a modest departure from more exact expressions that are better reserved for computer simulations (see Problem 5.91).

The y-directed electron current density at any point along the channel is given by

$$J_y = -\sigma \mathcal{E}_y = q\mu_e n \frac{d\psi}{dy} \,, \tag{5.1}$$

where ψ is the electric potential along the channel in relation to the source. (The negative sign is consistent with positive current in the $-y$ direction.) We use this expression to obtain the total electron current. Specifically,

$$i_d = q\mu_e W \left[\int_0^\infty n(x,y)\, dx \right] \frac{d\psi}{dy} = q\mu_e W N_I(y) \frac{d\psi}{dy} \,. \tag{5.2}$$

Here, $N_I(y)$ is the areal (per-unit-area) inversion-layer electron density.

Our next step is to integrate Eq. 5.2 over the interval $[0, L]$, subject to a gradual variation of the integrand. The result is

$$\int_0^L i_d \, dy = q \int_{\psi(0)}^{\psi(L)} \mu_e W N_I(\psi) \, d\psi \, . \tag{5.3}$$

But with constant i_d along the channel region, the left-hand side of Eq. 5.3 is simply $i_d L$. The right-hand side of Eq. 5.3 gives us a bit more difficulty. From Kirchhoff's Voltage Law, the potential difference between the gate and position y along the channel is $v_{gy} = v_{gs} - \psi$. Thus, we propose

$$qN_I(\psi) \approx C_{ox}(v_{gs} - \psi - V_T) \, , \tag{5.4}$$

where $C_{ox} = \epsilon_{ox}/t_{ox}$ for an SiO_2 insulator with dielectric permittivity ϵ_{ox} —see the Appendix for justification. In turn, with $\psi(0) = 0$, $\psi(L) = v_{ds}$, and the assumption of constant μ_e along the channel,

$$i_d = \frac{1}{2} K' \frac{W}{L} [\, 2(v_{gs} - V_T)v_{ds} - {v_{ds}}^2 \,] \, , \tag{5.5}$$

where

$$K' = \mu_e C_{ox} \, . \tag{5.6}$$

Equation 5.5 reflects the **resistive** or **triode** mode of MOSFET operation. At the onset of saturation, $\partial i_d / \partial v_{ds} = 0$. So $v_{ds} = v_{ds,sat} = v_{gs} - V_T$, and

$$i_d = i_{d,sat} = \frac{1}{2} K' \frac{W}{L} (v_{gs} - V_T)^2 \, . \tag{5.7}$$

The drain current does not increase beyond $i_{d,sat}$ for $v_{ds} > v_{ds,sat}$.

Readers who have skipped over some of the preceding discussion will appreciate the following summary of the current-voltage characteristics for the n-channel enhancement-mode MOSFET:

$$i_d = \begin{cases} 0 & \text{for } v_{gs} - V_T \leq 0 \, ; \quad (5.8)\text{a} \\ \dfrac{1}{2} K' \dfrac{W}{L} [2(v_{gs} - V_T)v_{ds} - {v_{ds}}^2] & \text{for } 0 < v_{ds} \leq v_{gs} - V_T \, ; \quad (5.8)\text{b} \\ \dfrac{1}{2} K' \dfrac{W}{L} (v_{gs} - V_T)^2 & \text{for } 0 < v_{gs} - V_T \leq v_{ds} \, . \quad (5.8)\text{c} \end{cases}$$

In these expressions, K' is the MOSFET **transconductance parameter**, and W and L are the MOSFET channel width and length, respectively. The latter two parameters are easily adjusted in integrated circuit designs. Thus, they are set apart from K', which is dependent upon relatively rigid details of the MOSFET fabrication process.

Depletion-Mode MOSFET

Although becoming increasingly rare, a second pair of constituents reveals the rich diversity of the MOSFET family. These are the n-and p-channel **depletion-mode** MOSFETs shown in Figs. 5.9a and 5.9b, respectively. Both are similar to their enhancement-mode counterparts except for the presence of a conducting channel between source and drain when $v_{gs} = 0$. (The channel is introduced during the MOSFET fabrication process by inserting a *thin* layer of impurity charge with appropriate conductivity type.) Thus, depletion-mode MOSFETs are normally "on." To underscore this mode of operation, our preferred device symbols contain an extra solid line between the source and drain.

The depletion designation requires *negative* threshold voltage for an n-channel MOSFET and *positive* threshold voltage for a p-channel MOSFET. Apart from these parametric differences, both depletion-mode devices have current-voltage characteristics with the same functional forms as those that have been derived for enhancement-mode transistors.

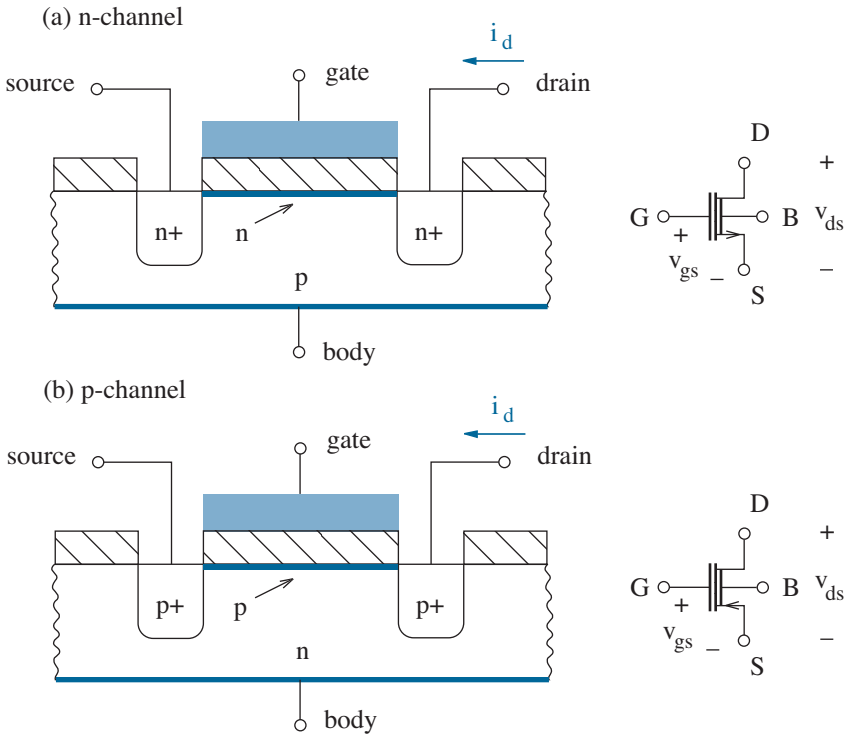


Figure 5.9: Depletion-mode MOSFET structures and their corresponding device symbols: (a) n-channel; (b) p-channel.

Example 5.1

An n-channel MOSFET features $K'W/L = 2.0$ mA/V^2, and $V_T = 1.5$ V. Sketch the characteristic curves.

Solution

We construct two parallel graphs as shown in Fig. 5.10. On the left, we plot the saturation current $i_{d,sat}$ as a function of v_{gs} using Eq. 5.8c. This curve is the right-half portion of a parabola that intercepts the v_{gs} axis at V_T. On the right, we prepare i_d and v_{ds} axes with the same current and voltage scales used for the left-side graph. These axes will embrace the MOSFET characteristic curves.

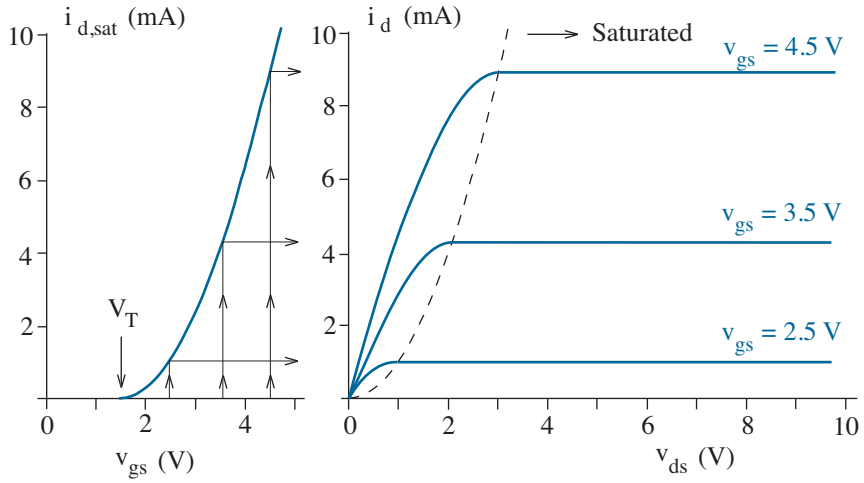


Figure 5.10: Construction of the MOSFET characteristic curves.

The boundary between non-saturated and saturated operating regions in the right-side graph is determined from the relation $v_{ds,sat} = v_{gs} - V_T$. Some careful thinking reveals that this boundary is a half parabola with the same shape as the curve in the left-side graph, but it is shifted to the left by V_T so that it intersects the origin.

Next, we consider the saturated portions of the MOSFET characteristic curves, which lie to the right of the $v_{ds,sat}$ boundary. For each of several v_{gs} specifications, we determine an associated $i_{d,sat}$ value with the aid of the left-side graph. Then we draw a flat curve segment.

Our last step is to consider the non-saturated portions of the MOSFET characteristic curves, which extend as straight lines from the origin and then gradually decrease in slope until they intersect the $v_{ds,sat}$ boundary. A crude free-hand sketch will generally suffice.

Example 5.2

Use the MOSFET test circuit of Fig. 5.11 to measure $K'W/L$ and V_T.

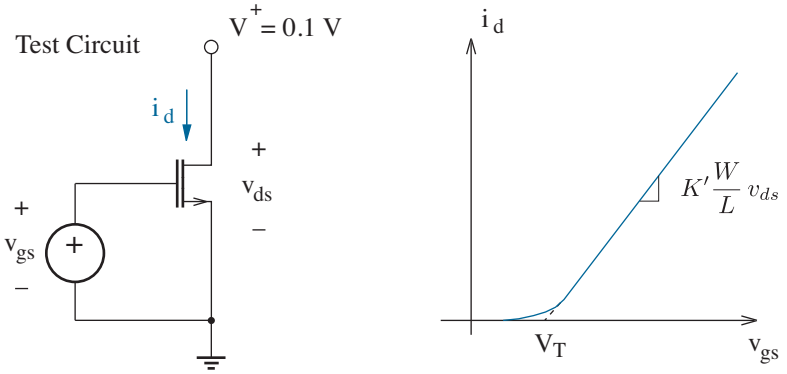


Figure 5.11: Measurement procedure for $K'W/L$ and V_T.

Solution

With small $v_{ds} = 0.1$ V, the MOSFET is quick to establish the resistive mode of operation as v_{gs} sweeps beyond V_T. The drain current is

$$i_d = \frac{1}{2}K'\frac{W}{L}[2(v_{gs} - V_T)v_{ds} - {v_{ds}}^2] .$$

And subject to $v_{ds} \ll v_{gs} - V_T$,

$$i_d \approx K'\frac{W}{L}(v_{gs} - V_T)v_{ds} .$$

Thus, the slope of the linear portion of the i_d vs. v_{gs} curve is $K'(W/L)v_{ds}$, and the extrapolated intercept with the v_{gs} axis is V_T.

Consider some actual measurement data:

v_{gs} (V)	v_{ds} (V)	i_d (mA)
1.0	0.1	0.072
1.5	0.1	0.192
2.0	0.1	0.312
2.5	0.1	0.432
3.0	0.1	0.552

After plotting these data (to verify their nearly linear character), we find $K'W/L \approx (0.552 - 0.072)/(2 \times 0.1) = 2.4$ mA/V^2. In turn, at $v_{gs} = 3$ V, $(3 - V_T) \approx 0.552/(2.4 \times 0.1) = 2.3$ V so that $V_T \approx 0.7$ V.

Exercise 5.2 The MOSFETs of Fig. 5.12 feature $K' = 50\ \mu\text{A/V}^2$, $L = 2\ \mu\text{m}$, and $|V_T| = 1$ V. Determine the consistent channel width W for the specified circuit conditions. Be sure to check for saturation.

Ans: (a) $W = 80\ \mu\text{m}$ (b) $W = 100\ \mu\text{m}$ (c) $W = 40\ \mu\text{m}$ (d) $W = 160\ \mu\text{m}$

(e) $W = 20\ \mu\text{m}$ (f) $W = 120\ \mu\text{m}$

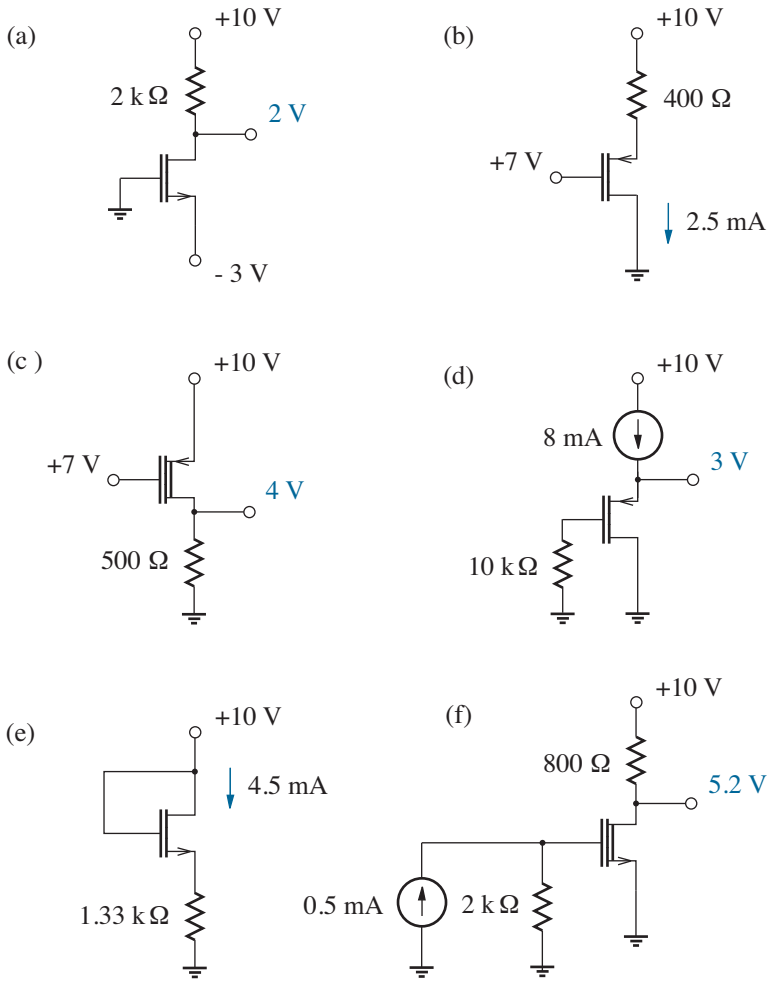


Figure 5.12: Circuits for Exercise 5.2.

Concept Summary

The MOSFET is a four-terminal electronic device.

- Current flows from *source* to *drain*.
  - n-channel MOSFETs have n^+ terminals in a p-type substrate.
    - Electron flow reflects positive inward drain current.
  - p-channel MOSFETs have p^+ terminals in an n-type substrate.
    - Hole flow reflects negative inward drain current.
- An electrically isolated *gate* terminal is used to induce an electric field that can either promote or suppress a conducting channel between the source and drain in relation to a threshold voltage V_T.
  - n-channel MOSFETs are "on" for $v_{gs} > V_T$.
    - Enhancement-mode devices are normally off with $V_T > 0$.
    - Depletion-mode devices are normally on with $V_T < 0$.
    - The MOSFET is "saturated" (valve-like) for $v_{ds} > v_{gs} - V_T$.
  - p-channel MOSFETs are "on" for $v_{gs} < V_T$.
    - Enhancement-mode devices are normally off with $V_T < 0$.
    - Depletion-mode devices are normally on with $V_T > 0$.
    - The MOSFET is "saturated" (valve-like) for $v_{ds} < v_{gs} - V_T$.
  - The dc gate current is always zero.
- A *body* or *substrate* terminal ensures zero- or reverse-bias conditions for the source and drain pn junctions.
  - n-channel MOSFETs require the body voltage to be at the most negative potential in a circuit.
  - p-channel MOSFETs require the body voltage to be at the most positive potential in a circuit.
  - The dc body current is zero under proper operating conditions.
- The MOSFET drain current is
  - Proportional to a device transconductance parameter K', which, in turn, is proportional to the channel carrier mobility;
  - Proportional to a pre-specified geometric aspect ratio W/L.

5.2 Large-Signal Models

In this section, we relate n- and p-channel MOSFET operating principles to large-signal lumped-element models that can be used for circuit analysis. The models are static—dynamic capacitive effects are examined elsewhere.

> Caution: As we introduce the MOSFET models, we should not attempt to commit them to memory. Rather, we accept them for reference with an understanding of their physical origin. Familiarity will come with practice in later chapters.

Models for a MOSFET Switch

For an n-channel MOSFET with $v_{gs} < V_T$, the switch is "off" with $i_d = 0$. The gate current is also zero, so we obtain the trivial MOSFET model of Fig. 5.13 in which the device terminals are nothing more than open circuits.

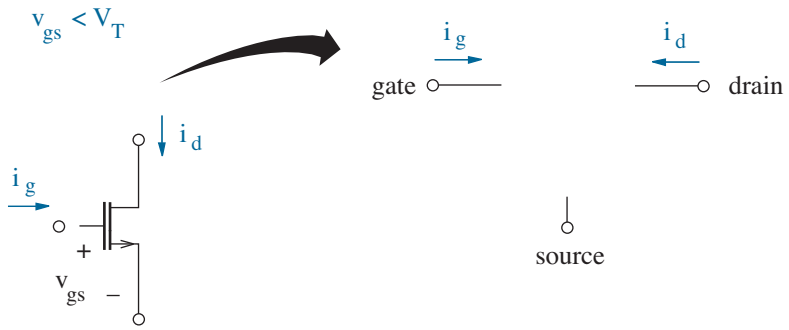


Figure 5.13: MOSFET switch model for the "off" state.

The "on" state presents a greater challenge. Since we expect *small* v_{ds} but possibly *large* v_{gs}, it seems reasonable to assume that $v_{ds} \ll v_{gs} - V_T$. In turn, the MOSFET is not in saturation. The alternative is the resistive mode of operation with

$$i_d = \frac{1}{2} K' \frac{W}{L} \left[2(v_{gs} - V_T) v_{ds} - {v_{ds}}^2 \right] . \tag{5.9}$$

—Some mess. But ignoring the relatively insignificant ${v_{ds}}^2$ term, we have $i_d \approx v_{ds}/r_{ds}$, where

$${r_{ds}}^{-1} = K' \frac{W}{L} (v_{gs} - V_T) . \tag{5.10}$$

Thus with zero gate current, the resulting MOSFET circuit model has the simple form shown in Fig. 5.14. Subject to $K'W/L = 10\ \text{mA/V}^2$, $V_T = 1$ V, and $v_{gs} = 5$ V, the MOSFET "on" resistance is $r_{ds} = 25\ \Omega$.

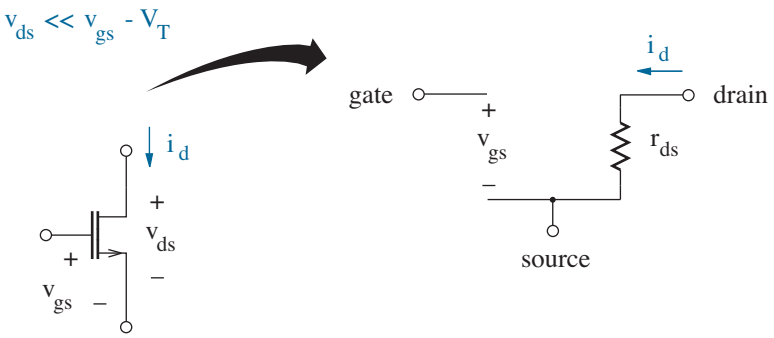


Figure 5.14: MOSFET switch model for the "on" state.

A more sophisticated on-state model could include r_d and r_s parasitic resistance in series with the drain and source, respectively. Nevertheless, these modifications typically fail to make up for the inaccuracy that arises from neglect of the ${v_{ds}}^2$ term in Eq. 5.9. Improved circuit precision is best reserved for SPICE simulations.

The p-channel MOSFET is "off" for $v_{gs} > V_T$ and "on" for $v_{gs} < V_T$. In the latter case, the device functions like its n-channel counterpart with currents and voltages altered by sign changes in the governing equations. The device parameters remain unchanged, except for V_T, which is now $-V_T$ (to ensure the proper sense of threshold: $-v_{gs} = -V_T$). Thus, we find

$$-i_d = \frac{1}{2} K' \frac{W}{L} \left[2(-v_{gs} + V_T)(-v_{ds}) - (-v_{ds})^2 \right] . \tag{5.11}$$

We subsequently drop the $(-v_{ds})^2$ term and let $-i_d \approx -v_{ds}/r_{ds}$ to obtain

$${r_{ds}}^{-1} = K' \frac{W}{L} (-v_{gs} + V_T) . \tag{5.12}$$

The resistive on-state condition requires $v_{gs} - V_T < v_{ds} < 0$.

Exercise 5.3 A particular n-channel MOSFET has "on" resistance values of 80 Ω and 30 Ω when $v_{gs} = 3$ V and $v_{gs} = 6$ V, respectively. Determine r_{ds} when $v_{gs} = 12$ V.

Ans: $r_{ds} = 13\ \Omega$

Models for a MOSFET Valve

The first-order model for the n-channel MOSFET in *saturation* (Fig. 5.15) features an open circuit at the gate terminal and a single v_{gs}-dependent current source with the character of Eq. 5.8c. The model assumes $v_{bs} = 0$. Four-terminal models with $v_{bs} \neq 0$ are reserved for Chapter 9 and beyond.

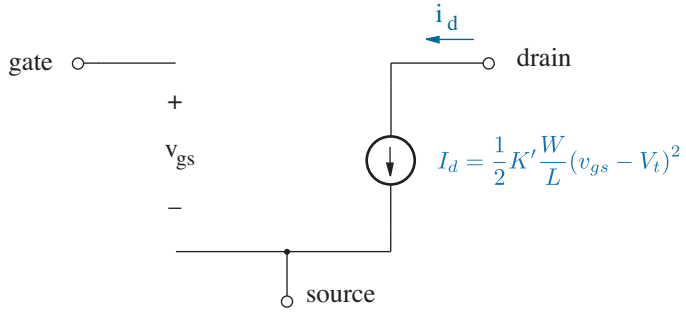


Figure 5.15: Three-terminal large-signal "valve" model for the n-channel MOSFET in saturation.

When designing analog MOSFET circuits, it is often desirable to find the gate-to-source voltage consistent with a given saturation drain current. Specifically, for the model of Fig. 5.15,

$$v_{gs} = V_T + \sqrt{\frac{2\, i_d}{K'(W/L)}}\,. \tag{5.13}$$

The second term on the right-hand side of Eq. 5.13 is sometimes called the **overdrive gate voltage**, a measure of how small v_{ds} can become while preserving MOSFET saturation. Low overdrive values require large W/L.

As noted previously, the p-channel MOSFET functions like its n-channel counterpart, but all terminal currents and voltages are altered by sign changes in the governing relations, and $V_T \rightarrow -V_T$. In place of Eq. 5.8c, we have

$$-i_d = \frac{1}{2}K'\frac{W}{L}(-v_{gs} + V_T)^2\,. \tag{5.14}$$

or

$$-i_d = \frac{1}{2}K'\frac{W}{L}(v_{gs} - V_T)^2\,. \tag{5.15}$$

Thus, the p-channel large-signal model has a form similar to Fig. 5.15. However, the orientation of the dependent current source is reversed.

Pending Complications

The preceding MOSFET models will prove their value for first-order circuit analysis and as initiators for practical design. Nevertheless, the models fail to account for complications involving short channel lengths or non-zero v_{bs}. We can look forward to the following in later chapters —

When we examine analog integrated circuits in Chapter 9, we will find that a relatively moderate short-channel effect produces a saturation drain current with the n-channel form

$$i_{d,sat} \approx \frac{1}{2}K'\frac{W}{L}(v_{gs} - V_T)^2\,(1 + \lambda v_{ds})\,, \tag{5.16}$$

where λ is constant. (The moderate influence on digital circuits is benign.) Apart from the departure from an exclusively v_{gs}-dependent current source, the effect will lead to opportunities for solid-state loading without resistors. Some analog integrated circuits will also involve non-zero v_{bs} and a shift in threshold voltage. For the n-channel case, the so-called **body effect** yields

$$V_T = V_{T0} + \gamma\,(\sqrt{2\phi_f - v_{bs}} - \sqrt{2\phi_f})\,, \tag{5.17}$$

where V_{T0} is the threshold voltage for $v_{bs} = 0$, and γ and $2\phi_f$ are constants. This reduces the saturation drain current subject to constant v_{gs}. Finally, we will find that MOSFETs have **subthreshold current** for $|v_{gs}| < |V_T|$. Specifically, for n-channel devices,

$$i_d \approx I_s \exp\left[\frac{-q\,(V_T - v_{gs})}{nkT}\right]\,, \tag{5.18}$$

where I_s and n are constants. Equation 5.18 indicates a gradual turn-off condition that is best avoided.

When we examine digital integrated circuits in Chapter 10, we will find that very short channel lengths promote electric fields that are sufficiently high to no longer support proportional carrier velocities in terms of mobility. The resulting **velocity saturation** reduces the drain current. Specifically, for relatively vulnerable n-channel devices,

$$i_{d,sat} \approx \frac{1}{2}K'\frac{W}{L}(v_{gs} - V_T)^2\left[1 + \frac{0.8(v_{gs} - V_T)}{\mathcal{E}_c L}\right]^{-1}\,, \tag{5.19}$$

where $\mathcal{E}_c$ is a critical value for the electric field. The drain current is now more linearly dependent on $v_{gs} - V_T$ in the limit when $v_{gs} - V_T \gg \mathcal{E}_c L$.

Our study plan is to become familiar with first-order MOSFET models and then delve into complications (and physical origins) as we need them. Each will tend to support hand calculations and expanded insight regarding circuit operation and design. Complications beyond those listed above tend to be more obscure and are better reserved for SPICE computer simulations. Remember the perils of precision.

Concept Summary

MOSFET circuit models support switch- or valve-like operation.

- A MOSFET switch appears as a resistor r_{ds} between drain and source. The gate terminal is set aside as a means for control.
  - For n-channel devices,
    - The switch is off for $v_{gs} < V_T$,
    - The switch is otherwise on with ${r_{ds}}^{-1} = K' \frac{W}{L}(v_{gs} - V_T)$.
  - For p-channel devices,
    - The switch is off for $v_{gs} > V_T$,
    - The switch is otherwise on with ${r_{ds}}^{-1} = K' \frac{W}{L}(-v_{gs} + V_T)$.
  - The model assumes a consistent $|v_{ds}| < |v_{gs} - V_T|$.
- A MOSFET valve appears as a dependent current source I_d between drain and source. The gate terminal is set aside as a means for control.
  - For n-channel devices,
    - The valve is closed for $v_{gs} < V_T$,
    - The valve is otherwise open with $I_d = \frac{1}{2} K' \frac{W}{L}(v_{gs} - V_T)^2$.
  - For p-channel devices,
    - The valve is closed for $v_{gs} > V_T$,
    - The valve is otherwise open with $-I_d = \frac{1}{2} K' \frac{W}{L}(v_{gs} - V_T)^2$.
  - The model assumes a consistent $|v_{ds}| > |v_{gs} - V_T|$.
- The body terminal has no first-order influence when properly biased.
- The preceding elementary models have pending complications that necessitate second-order hand or computer analysis:
  - The valve current has v_{ds} dependence,
  - The threshold voltage depends on the body-to-source voltage,
  - The subthreshold current is small but not zero,
  - The effective channel mobility degrades via velocity saturation.
- Numerous other complications are generally reserved for SPICE.

5.3 Introductory MOSFET Circuits

In this section, we highlight the MOSFET as a switch of choice for modern circuit applications. Our consideration of valve action is relatively brief—the modest goal provides a reference for circuit comparisons in Chapter 6. We postpone more thorough treatment to Chapter 7.

The MOSFET switch has three major applications:

- Digital logic circuits and memory
- Analog switching (for signal processing and micropower circuits)
- Power conditioning and control

Of these, the first application is the most important and thus worthy of special consideration (Chapter 10). The others are examined here.

Digital Logic

The main objective of most digital logic circuits is to have HIGH and LOW output states that are dependent upon similarly quantized input states, either past or present. For the moment, suppose we are content to use a single input to switch between a LOW state at ground and a HIGH state at some V^+ supply voltage. A simple design option uses a single switch that ties either directly to ground (Fig. 5.16a) or V^+ (Fig. 5.16b) so that one ideally observes the corresponding logic level when the switch is closed. When the switch is open, the output is effectively tied to the opposite logic level through a "load" element such as a resistor that features zero voltage drop under zero-current conditions.

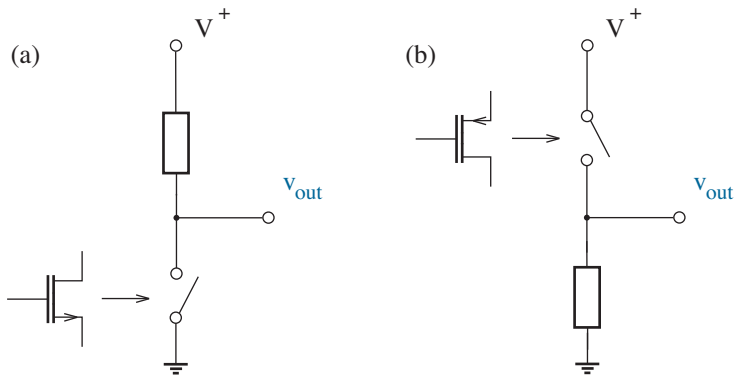


Figure 5.16: Elementary digital switching circuits. The n- and p-channel MOSFET switches have source connections at ground and V^+, respectively.

Although the circuits of Fig. 5.16 have the desired output dependence, there is still a need to establish real switches with appropriate input control. As part of a complex integrated system, the circuits react to the outputs of similar circuits. Thus, the only input control levels are ground and V^+. And for now, the voltage-controlled electronic switches are limited to either n- or p-channel enhancement-mode MOSFETs that are "off" for $v_{gs} = 0$. Two connection rules apply:

- The MOSFET is oriented so that positive current flows in the same direction as the arrow along the source terminal.
- The MOSFET source is tied to a fixed node voltage so that changes in v_{gs} exclusively reflect changes in V_g, the node voltage at the gate.

Accordingly, the circuit of Fig. 5.16a requires an n-channel MOSFET with source connected to ground. The switch is on for $v_{gs} = V_g - 0 > V_T > 0$. In contrast, the circuit of Fig. 5.16b requires a p-channel MOSFET with source connected to V^+. The switch is on for $v_{gs} = V_g - V^+ < V_T < 0$. For either circuit, a particular input level yields the opposite output level. This is the function of a digital **inverter**.

How good is our design? The r_{ds} "on"-resistance value for the n-channel MOSFET is given by Eq. 5.10, and although hopefully small, it is not zero. So to achieve a LOW v_{out} value close to ground, the load must either be a large resistance or some other device that is capable of sustaining a large voltage when current is small. And even if we can find an acceptable load, the inverter always dissipates power when the MOSFET switch is closed. This implies a practical limit on the number of MOSFET-and-load inverters or related logic elements that can be used in a very large integrated circuit.

Example 5.3

The inverter circuit of Fig. 5.16a features $V^+ = 5$ V, a load resistor R, and an n-channel MOSFET with $K'W/L = 0.5$ mA/V^2 and $V_T = 1$ V. Determine R so that the LOW output level is 0.2 V.

Solution

A LOW output implies a HIGH input of 5.0 V. Thus with $v_{gs} = 5.0$ V and $v_{ds} = 0.2$ V, Eq. 5.9 yields $i_d = 0.39$ mA as drain current shared with R. In turn, $R = (5.0 - 0.2)/0.39 = 12.3$ kΩ.

An alternative approximate solution uses Eq. 5.10 to obtain $r_{ds} = 500\ \Omega$ as the "on" resistance for the MOSFET switch. The output voltage is obtained from an elementary divider relation: $v_{out} = V^+\ r_{ds} / (r_{ds} + R)$. So with $v_{out} = 0.2$ V, we find $R = 12.0$ kΩ, which is in good agreement.

You may have felt that the critique of the single-switch inverter designs of Fig. 5.16 was suggestive of an inadequate technology. And you were right. Nevertheless, the design shortcomings suggest a means for improvement. If we require a high-resistance load, why not use an open switch? Further, if we require that the load make an "effective" tie to an opposite logic level on some occasions, why not connect directly by means of a closed switch? In other words, why not use two switches instead of one?

Figure 5.17 shows a two-switch design in which the output is connected to a LOW or HIGH level when one switch is closed and the other is open. This behavior is observed when the lower and upper switches are replaced by n- and p-channel enhancement-mode MOSFETs, respectively. The two gate terminals are tied together. Thus, when the n-channel device is "on" with $V_g > V_{Tn}$, the p-channel device is "off" —provided $V_g - V^+ > V_{Tp}$ (as expected, since HIGH $V_g = V^+$). When the p-channel device is "on" with $V_g - V^+ < V_{Tp}$, the n-channel device is "off" —provided $V_g < V_{Tn}$ (as expected, since LOW $V_g = 0$). The circuit is our first example of **CMOS** in which complementary MOS transistors work together to achieve functionality that is improved over that for only one type of transistor. CMOS technology is the basis for most complex digital circuits in use today.

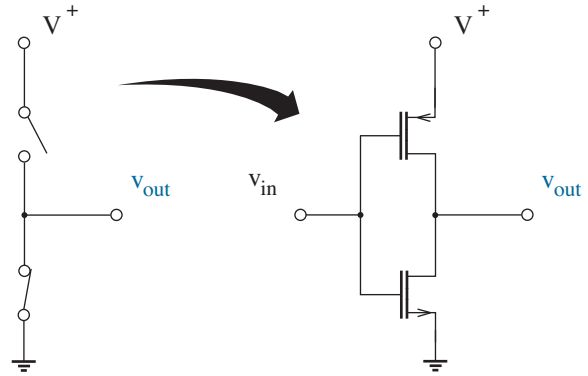
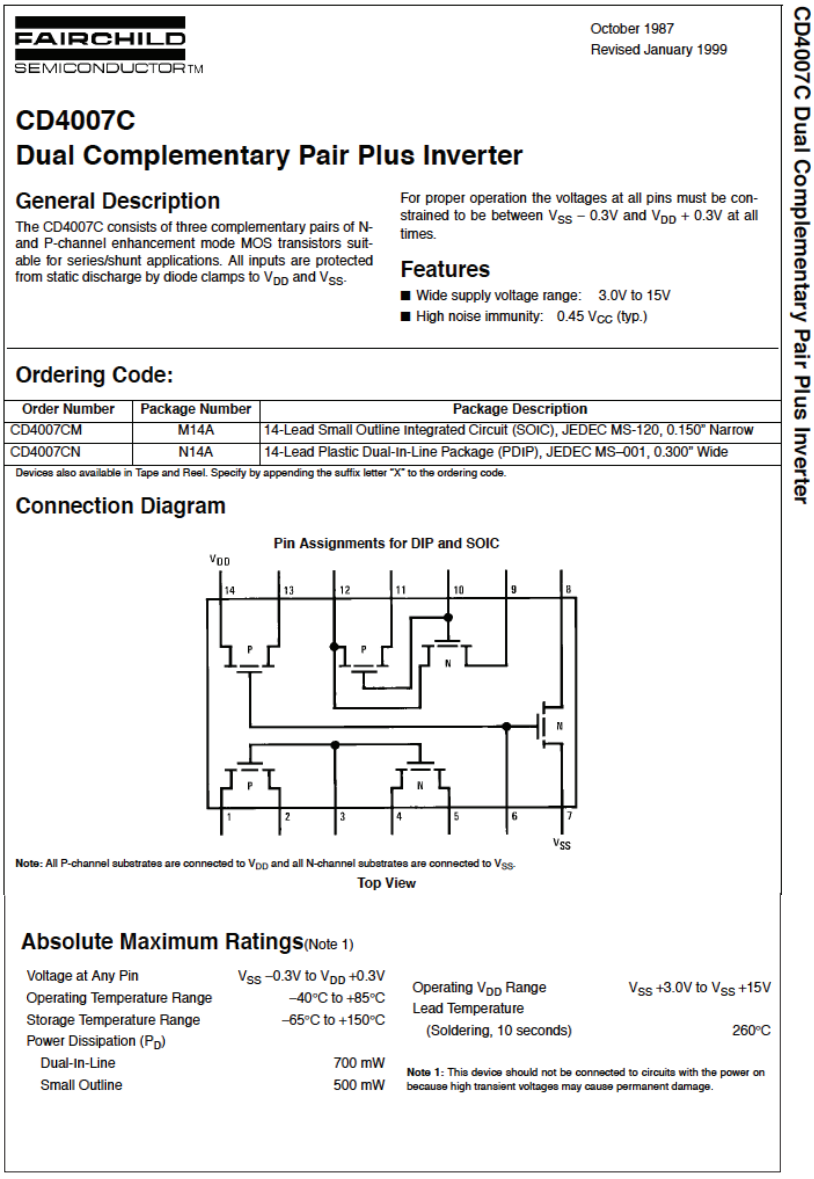


Figure 5.17: CMOS inverter. One switch is closed when the other is open.

The most obvious advantage of the CMOS digital inverter is *zero* power dissipation when the input is LOW or HIGH so that one of the MOSFET switches does not pass current. Power is only expended during output-level transitions (through dynamic processes examined in Chapter 9).

Another advantage of the CMOS inverter concerns transition speeds. Inverter outputs typically "drive" capacitor-like inputs of similar circuits. The capacitance charges and discharges during HIGH-to-LOW and LOW-to-HIGH transitions, respectively, and *both* processes in CMOS are limited by a low-resistance MOSFET switch. Compare this with the circuits of Fig. 5.16 where one transition is prolonged by a high-resistance load.

The CD4000 series of integrated small-scale CMOS logic builds upon the hierarchy described in Chapter 1 —Chapter 10 offers some circuit details— and it is still available despite the trend towards ever increasing complexity. Figure 5.18 shows part of the data sheet for the CD4007 CMOS inverter, which remains popular for student experiments.



FAIRCHILD SEMICONDUCTOR™

October 1987
Revised January 1999

CD4007C Dual Complementary Pair Plus Inverter

CD4007C
Dual Complementary Pair Plus Inverter

General Description

The CD4007C consists of three complementary pairs of N- and P-channel enhancement mode MOS transistors suitable for series/shunt applications. All inputs are protected from static discharge by diode clamps to V_{DD} and V_{SS}.

For proper operation the voltages at all pins must be constrained to be between V_{SS} – 0.3V and V_{DD} + 0.3V at all times.

Features

- Wide supply voltage range: 3.0V to 15V
- High noise immunity: 0.45 V_{CC} (typ.)

Ordering Code:

Order Number	Package Number	Package Description
CD4007CM	M14A	14-Lead Small Outline Integrated Circuit (SOIC), JEDEC MS-120, 0.150" Narrow
CD4007CN	N14A	14-Lead Plastic Dual-In-Line Package (PDIP), JEDEC MS–001, 0.300" Wide

Devices also available in Tape and Reel. Specify by appending the suffix letter "X" to the ordering code.

Connection Diagram

Pin Assignments for DIP and SOIC

Note: All P-channel substrates are connected to V_{DD} and all N-channel substrates are connected to V_{SS}.

Top View

Absolute Maximum Ratings (Note 1)

Voltage at Any Pin	V_{SS} –0.3V to V_{DD} +0.3V
Operating Temperature Range	–40°C to +85°C
Storage Temperature Range	–65°C to +150°C
Power Dissipation (P_D)	
Dual-In-Line	700 mW
Small Outline	500 mW
Operating V_{DD} Range	V_{SS} +3.0V to V_{SS} +15V
Lead Temperature (Soldering, 10 seconds)	260°C

Note 1: This device should not be connected to circuits with the power on because high transient voltages may cause permanent damage.

Figure 5.18: CD4007 data sheet. Courtesy of Fairchild Semiconductor.

Analog Switching

In the general case, an analog switch is used to join or disconnect to a node that has any voltage within a continuous range, not just V^+ or ground. The general switch is also **bidirectional**; current flows in either direction as determined by conditions in the surrounding circuit.

A major consequence of the preceding requirements is that either side of an n-channel MOSFET must be free to assume the role of source or drain depending on the direction of current flow—positive current flows from drain to source—despite any particular orientation of the device symbol.[2] As always, the MOSFET body (substrate) must be separately connected to the most *negative* available voltage to ensure that both sides of the switch are isolated from the body by means of reverse-biased pn junctions. Meanwhile, the MOSFET gate is connected to the most *positive* available voltage when the switch is on, and it is connected to the most negative available voltage when the switch is off. Figure 5.19 shows the on-state terminal connections. Note the source relationships to current flow.

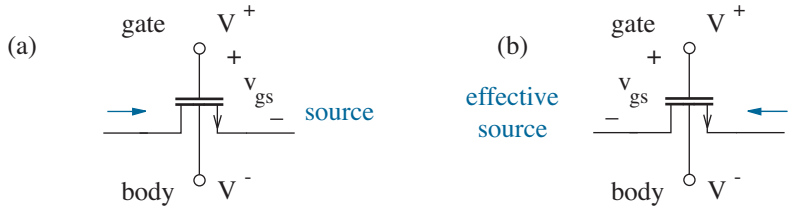


Figure 5.19: Terminal connections for a general n-channel analog switch: (a) subject to rightward current flow; (b) subject to leftward current flow. In either case, the actual drain and source are consistent with $v_{ds} \geq 0$.

Now consider an n-channel analog switch subject to the "on" state and with modest passthrough current so that source and drain are nearly at the same potential. The gate voltage is fixed at V^+ while the source voltage tracks an analog signal bounded by V^+ and V^- upper and lower limits. Equation 5.10 suggests that the switch has minimum "on" resistance when the source voltage is minimum so that v_{gs} is maximum. As the source voltage increases, r_{ds} also increases. And when the source voltage has increased to the extent that $v_{gs} < V_T$, the switch loses its "on" state entirely ($r_{ds} \to \infty$). In turn, the n-channel analog switch is best used for *low-side* switching when signals are near V^- or, better yet, one side of the switch is fixed at V^-.

[2]Some texts avoid this difficulty—at the expense of confusion elsewhere—by using a different MOSFET device symbol that fails to distinguish between source and drain. The everpresent body terminal features an inward arrow for n-channel devices and an outward arrow for p-channel devices.

Take a moment to contrast the preceding discussion in the context of a p-channel analog switch. Here, the MOSFET body is separately connected to the most *positive* available voltage to ensure source and drain isolation. Meanwhile, the MOSFET gate is connected to the most *negative* available voltage when the switch is on, and it connects to the most positive available voltage when the switch is off. The switch functions with minimum "on" resistance when the source voltage is maximum, and r_{ds} increases as the source voltage decreases. In turn, the p-channel analog switch is best used for *high-side* switching when signals are near V^+ or, better yet, one side of the switch is fixed at V^+.

Given the separate limitations of n- and p-channel analog switches,

Why not use two switches instead of one?

This design tactic was certainly successful for digital switching.

All-purpose analog switches are made by having two devices operate in parallel, together with an inverter to allow a single gate-controlling voltage. Figure 5.20 shows this complementary MOS or CMOS analog switch with individual and parallel r_{ds} curves. Subject to matched $K'W/L$ values, Eqs. 5.10 and 5.12 combine to suggest constant r_{ds} over the range of v_s:

$$r_{ds}{}^{-1}{}_{\text{Total}} = K'\frac{W}{L}\left[\,(V^+ - V^-) - (V_{T1} - V_{T2})\,\right]. \tag{5.20}$$

In a digital circuit, the same combination is called a **transmission gate**.

CMOS analog switches are now standard issue for general applications. Apart from embedded use in large integrated circuits, these switches are packaged as "discrete" IC components. Single-pole single-throw (SPST), double-pole single-throw (DPST), single-pole double-throw (SPDT), and double-pole double-throw (DPDT) options are all available.

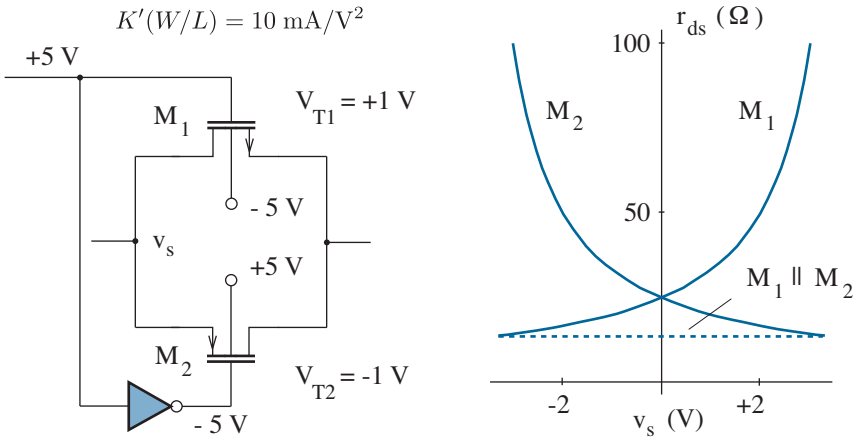


Figure 5.20: CMOS analog switch with n- and p-channel r_{ds} characteristics.

Figure 5.21 shows a partial data sheet for the MAX4706 (normally on) and MAX4707 (normally off) analog CMOS switches. The key specification that invites their choice over competing parts is the 2-Ω "on" resistance. Other important criteria include the dc off-state "leakage" current (1 nA), power-supply current (0.02 μA), off-state isolation (-67 dB $= 4.47 \times 10^{-4}$, the off-state attenuation factor for a signal on one side of the switch with respect to the other), and -3 dB bandwidth (190 MHz, the frequency beyond which a signal passing through the on-state switch is attenuated by more than a factor of $1/\sqrt{2}$). Chapter 8 explores the ac behavior in greater detail.

MAXIM

Low-Voltage, 2 Ω, SPST, CMOS Analog Switches

MAX4706/MAX4707

General Description

The MAX4706/MAX4707 single-pole/single-throw (SPST) switches operate from a single 1.8V to 5.5V supply. The MAX4706 is a normally closed (NC) switch and the MAX4707 is the normally open (NO) version. These switches provide 2Ω on-resistance (R_{ON}) and 0.6 Ω R_{ON} flatness with a +2.7V supply. These devices typically consume only 0.02µA of quiescent current, making them suitable for use in low-power, portable applications. The MAX4706/MAX4707 feature low-leakage currents over the entire temperature range, TTL/CMOS-compatible digital logic, and excellent AC characteristics.

The MAX4706/MAX4707 are offered in small 5-pin and 6-pin SC70 and 6-pin µDFN packages.

Applications

Battery-Operated Equipment
Audio Signal Routing
Low-Voltage Data-Acquisition Systems
Sample-and-Hold Circuits
Communications Circuits
Relay Replacement

Features

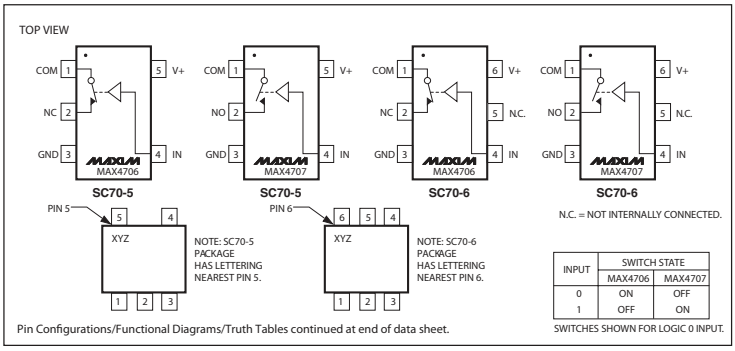
- **Guaranteed R_{ON} 2Ω max (+2.7V Supply)**
- **Guaranteed R_{ON} Flatness: 0.6Ω (+2.7V Supply)**
- **Excellent AC Characteristics**
  High Off-Isolation: -67dB at 1MHz
  -3dB Bandwidth: 190MHz
- **0.013% Total Harmonic Distortion**
- **Low Supply Current: 0.02µA**
- **Tiny SC70 and µDFN Packages**
- **1.8V to 5.5V Single-Supply Operation**

Ordering Information

PART	TEMP RANGE	PIN-PACKAGE	TOP MARK
MAX4706EXK-T	-40°C to +85°C	5 SC70-5	AFO
MAX4706EXT-T	-40°C to +85°C	6 SC70-6	ABS
MAX4706ELT-T*	-40°C to +85°C	6 µDFN-6	—
MAX4707EXK-T	-40°C to +85°C	5 SC70-5	AFP
MAX4707EXT-T	-40°C to +85°C	6 SC70-6	ABT
MAX4707ELT-T*	-40°C to +85°C	6 µDFN-6	—

*Future product—contact factory for availability.

Pin Configurations/Functional Diagrams/Truth Tables



Pin Configurations/Functional Diagrams/Truth Tables continued at end of data sheet.

MAXIM Maxim Integrated Products 1

Figure 5.21: MAX4706 data sheet. Copyright Maxim Integrated Products. Used by permission.

In what follows, we examine two basic analog switching applications. Some others have already been encountered in Chapter 4.

Exhibit 1

For our first example, we consider the op-amp circuit shown in Fig. 5.22. Here, a CMOS analog switch is controlled by voltage source V_g (not shown), and the op-amp functions as a unity-gain buffer so that v_{out} follows the capacitor voltage despite connections to v_{out} that would otherwise affect the capacitor. If V_g is initially LOW, the switch is off, and the capacitor is isolated from v_{in}. When V_g becomes HIGH at $t = 0+$, the switch turns on, and the capacitor voltage assumes the value of v_{in}. Finally, when V_g returns LOW at $t = \tau$, the switch is off, and the capacitor voltage remains at $v_{in}(\tau)$. Thus, we have a **sample-and-hold** circuit (perhaps for the front end of an analog-to-digital converter).

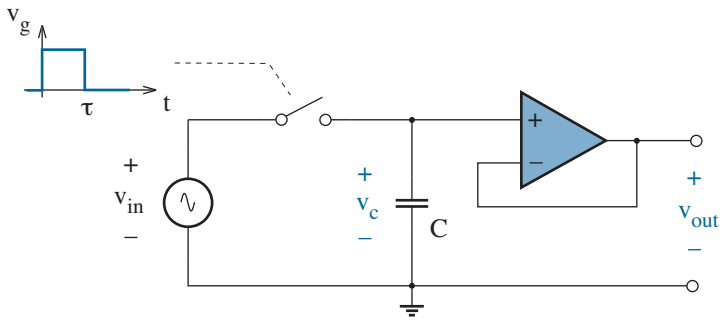


Figure 5.22: Sample-and-hold circuit.

The sample period τ must be sufficiently short to ensure a true snapshot of v_{in}. In turn, the switch must have an r_{ds} "on" resistance that is sufficiently small to ensure rapid settling of capacitor voltage. Specifically,

$$r_{ds}C \ll \tau \,. \tag{5.21}$$

These design rules usually force a small value for C.

Sample-and-hold circuits suffer from non-ideal switch behavior other than non-zero r_{ds}. In particular, on-state MOSFET switch constituents have channel charge with value

$$Q_c \approx \pm\, WLC_{ox}(v_{gs} - V_T)\,, \tag{5.22}$$

where the + and - signs apply to p- and n-channel devices, respectively. When the MOSFET is turned off, some of this charge moves to the source and the remainder moves to the drain—typically subject to an even split. If the capacitor in the circuit of Fig. 5.22 receives a fraction of this charge, the sampled voltage will experience proportionate error (see Problem 5.32).

Exhibit 2

A second example concerns the routing of high-frequency video signals. At low frequencies, the "standard" SPST analog switches of Fig. 5.23 have well-defined on and off states. But at high frequencies, say 10 MHz or more, the off states are blurred by the actions of parasitic shunt capacitors that couple input to output. In the particular circuit at hand, a small fraction of signal A combines with signal B subject to switch-1 open, switch-2 closed. This is an unacceptable corruption of the multiplexed video output.

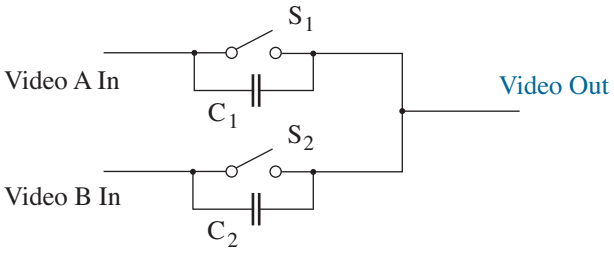


Figure 5.23: Analog multiplexer with parasitic shunt capacitance.

Figure 5.24 indicates how a combination of three analog switches can be used in place of a standard multiplexer switch to alleviate video crosstalk. The individual switches operate as follows for the two "T-switch" states:

T-Switch State	S_x	S_y	S_z
ON	Closed	Open	Closed
OFF	Open	Closed	Open

Thus, in the on state, a video signal propagates through switches S_x and S_z with just a negligible amount of attenuation due to the C_y path to ground. In the off state, the signal propagation path through S_x and S_z is removed, and the unintentional path through C_x is diverted to ground with S_y closed. The off-state isolation is about a hundredfold better than standard.

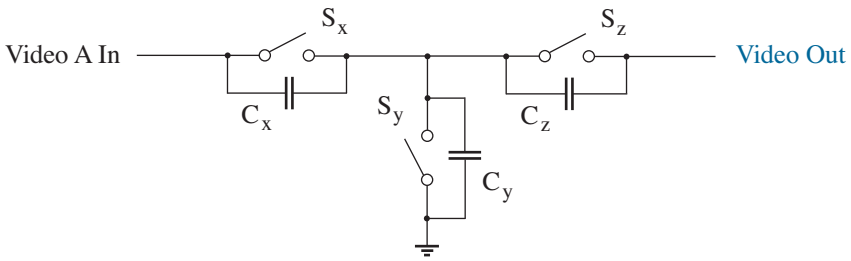


Figure 5.24: T-switch configuration for high-frequency applications.

Power Conditioning

Power conditioning applications that require switches include:

- dc-to-dc conversion
- dc-to-ac conversion
- ac-to-ac conversion

The dc-to-dc power conditioning process has been discussed in Chapter 4 (together with ac-to-dc conversion, which does not use switching devices). In what follows, we examine a variation on the dc-to-dc conversion process as well as a simple example of dc-to-ac conversion. Some aspects of ac-to-ac conversion are the topic of Problem 5.39.

First, a brief digression.

When very small r_{ds} values are needed, as for high-current applications, one generally employs a **power MOSFET**. This device is designed to have very large W/L while maintaining very small parasitic drain and source resistances (outside the realm of v_{gs} control). The geometry can be unusual. For example, the typical n-channel power MOSFET configuration shown in Fig. 5.25 features a bottom-side drain contact so that on-state current flow proceeds laterally from the source across a short inversion layer and then downward to the drain. The metal contact to the source also makes contact to the body. Thus, one finds a parasitic diode between the drain and source. This diode is seldom problematic because it is reverse biased when $v_{ds} > 0$. Indeed, the diode can be a benefit in some circuits.

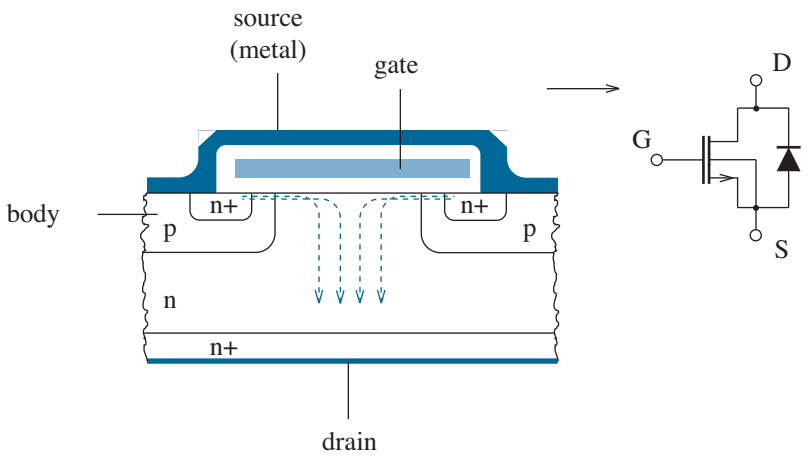


Figure 5.25: Typical power MOSFET and corresponding equivalent circuit. Dashed arrows show the direction of on-state electron flow.

Exhibit 3

The circuit of Fig. 5.26 contains two capacitors, a load, and two pairs of power switches (S_1, S_2) and (S_3, S_4) that are operated in antiparallel—one pair is on while the other is off. Both pairs are clocked at frequency f. What happens at the output?

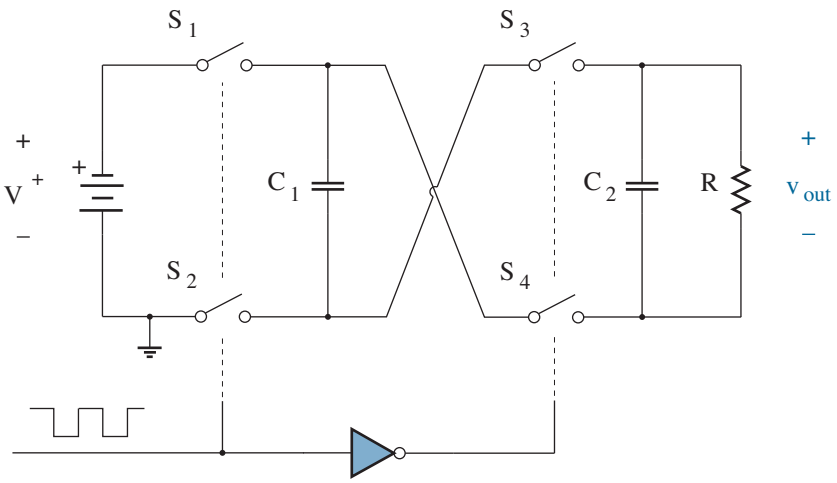


Figure 5.26: Charge-pump micropower circuit.

As a "flying" capacitor, C_1 is alternately charged to the voltage V^+, then reconnected across C_2 *but with a change in orientation.* The effect is that of a **charge pump** that supplies power to the load with *inverted* voltage polarity. Assume $RC_2 \gg 1/f$ so that the output voltage v_{out} is nearly constant once C_1 has been disconnected. After one switching cycle, the net charge transfer to the RC_2 combination is $\Delta Q = C_1(-V^+ - v_{out})$, and in the steady state, the average rate of charge transfer is current $\Delta Q \times f$. Time-averaged (dc) current is not absorbed into capacitor C_2. Thus,

$$C_1(-V^+ - v_{out})f \approx \frac{v_{out}}{R} \tag{5.23}$$

and

$$v_{out} \approx \frac{-V^+ RC_1 f}{1 + RC_1 f}\,. \tag{5.24}$$

In the design limit where $RC_1 f \gg 1$, $v_{out} \approx -V^+$. So it appears that we have the means of providing a negative supply voltage for an op-amp in a single-battery-powered circuit. Just try some other way.

Exhibit 4

The circuit of Fig. 5.27 contains a battery, a load, and two pairs of power switches (S_1, S_2) and (S_3, S_4) that are operated in antiparallel—one pair is on while the other is off. Both pairs are clocked at frequency f. What happens at the output?

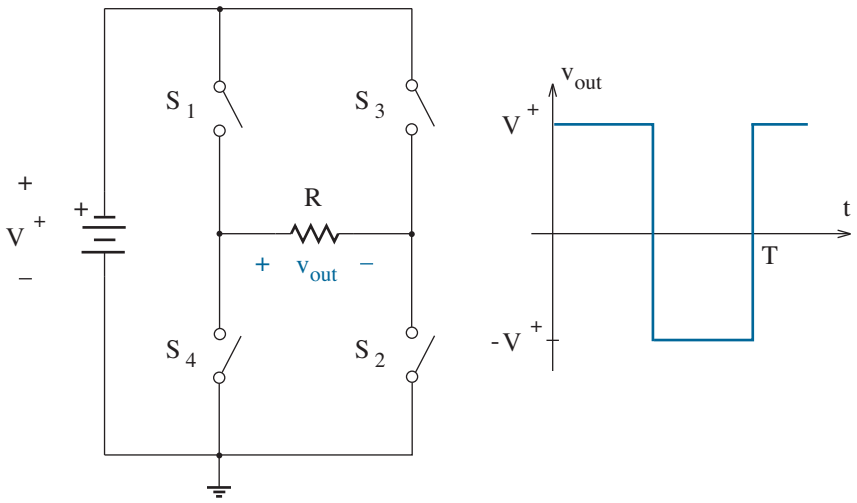


Figure 5.27: "Inverter" circuit for dc-to-ac power conversion.

When switches S_1 and S_2 are closed, the load is connected to the battery such that $v_{out} = V^+$. And when switches S_3 and S_4 are closed, the load reconnects such that $v_{out} = -V^+$. This square wave has period $T = 1/f$. A power circuit that produces an ac waveform from a dc source such as a car battery or a solar panel is called an **inverter**.

Further analysis shows that the output voltage is a superposition of sinusoidal signals of the form

$$v_{out} = \sum_{n,odd} \frac{4V^+}{n\pi} \sin n\omega_o t \,, \tag{5.25}$$

where $\omega_o = 2\pi f$. In turn, the power delivered in the n-th harmonic is

$$P_n = \frac{1}{R}\left(\frac{4V^+}{n\pi}\right)^2 . \tag{5.26}$$

Thus, about 81 % of the available power is delivered in the first harmonic. Significant improvement can be obtained by keeping all switches open for a short time interval near every half cycle (see Problem 5.36).

Valve Action

Before closing this section, we need practice with MOSFET circuits that function with valve-like action (saturation). The experience will provide a basis for comparison with the bipolar junction transistor of Chapter 6.

Example 5.4

The circuit of Fig. 5.28 contains two identical n-channel enhancement-mode MOSFETs for which $K'(W/L) = 2$ mA/V^2 and $V_T = 1$ V. Complete the design such that $i_2 = 4$ mA.

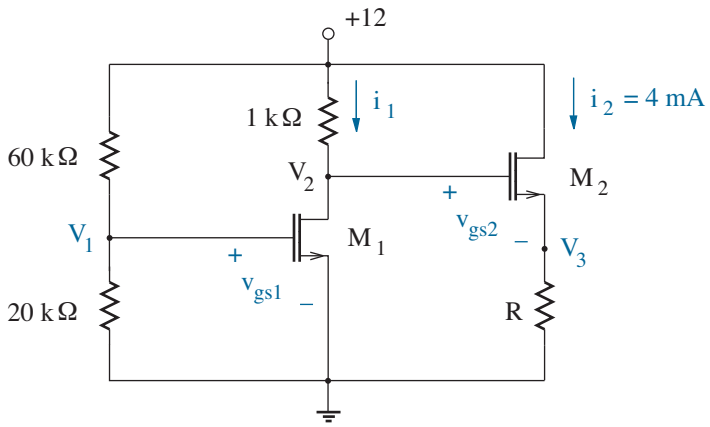


Figure 5.28: Circuit for Example 5.4.

Solution

We fearlessly mark up the circuit diagram by identifying as many currents and node voltages as possible using simple relations. All of the calculations assume that the MOSFETs are in saturation, and gate currents are zero. In order of interest, we have:

- $V_1 = 12 \times 20\,/(20 + 60) = 3$ V $= v_{gs1}$.
- $i_1 = \dfrac{1}{2}K'\dfrac{W}{L}(3-1)^2 = 4$ mA. Thus, $V_2 = v_{ds1} = 12 - 1 \times 4 = 8$ V.
- If $i_2 = 4$ mA, $v_{gs2} = 3$ V (like v_{gs1} for M_1).
- $V_3 = V_2 - 3 = 5$ V. In turn, $R = 5\,/\,4 = 1.25$ kΩ.
- $v_{ds1} = 8$ V $> v_{gs1} - V_T = 2$ V, $v_{ds2} = 12 - V_3 = 7$ V $> v_{gs2} - V_T = 2$ V. (So M_1 and M_2 are both in saturation, as assumed.)

Example 5.5

The MOSFETs in the circuit of Fig. 5.29 have $K'(W/L) = 2$ mA/V^2 and $|V_T| = 1$ V. Complete the design such that $v_{out} = 0$ V.

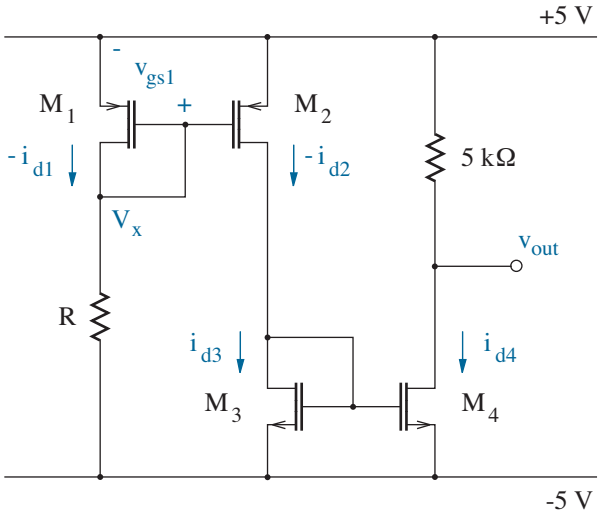


Figure 5.29: Circuit for Example 5.5.

Solution

Mark up the circuit diagram. With $v_{out} = 0$ V, $i_{d4} = (5 - 0)/5 = 1$ mA. Then with M_3 and M_4 having the same gate-to-source voltages, $i_{d3} = i_{d4}$ subject to saturation conditions. Similarly, $-i_{d1} = -i_{d2}$ for M_1 and M_2. Thus, $-i_{d1} = 1$ mA, since $-i_{d2} = i_{d3}$. In turn, we apply Eq. 5.15:

$$1 \text{ mA} = \frac{1}{2}(2 \text{ mA})\,(v_{gs1} + 1)^2$$

so that $v_{gs1} = 0$ V or $v_{gs1} = -2$ V. Only the latter solution is valid with $v_{gs} < V_T$ for the p-channel MOSFET. We now have $V_x = 5 + v_{gs1} = 3$ V. Finally, $R = [3 - (-5)]\,/1 = 8$ kΩ.

We leave it as an exercise to show that every MOSFET is in saturation and that M_1 and M_3 are necessarily saturated by virtue of $v_{gs} = v_{ds}$ and their enhancement-mode status.

The MOSFET combinations that promote $-i_{d2} = -i_{d1}$ and $i_{d4} = i_{d3}$ demonstrate the separate reflective actions of a **current mirror**, a common current-manipulating feature in analog integrated circuits such as op-amps. See Chapters 7 and 10 for further discussion.

Exercise 5.4 Design the circuits of Fig. 5.30 to achieve the specified values for v_{out}. Assume $K'(W/L) = 4$ mA/V^2 and $|V_T| = 2$ V.

Ans: (a) $R = 500\ \Omega$ (b) $R = 2$ kΩ

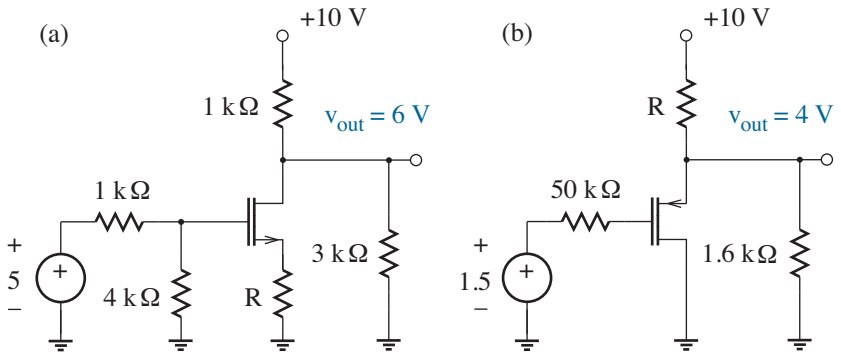


Figure 5.30: Circuits for Exercise 5.4.

Exercise 5.5 Specify R such that $v_{out} = 10$ V in the circuits of Fig. 5.31. Assume $K'(W/L) = 1$ mA/V^2 and $|V_T| = 1$ V for both MOSFETs.

Ans: (a) $R = 2.9$ kΩ (b) $R = 10$ kΩ

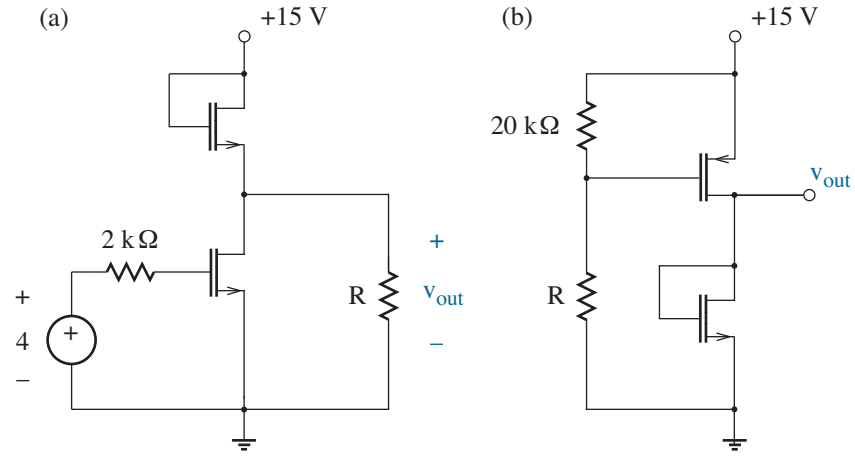


Figure 5.31: Circuits for Exercise 5.5.

Concept Summary

For switch applications with a single MOSFET

- An n-channel enhancement-mode device
  - Turns on and off when its gate connects to the highest and lowest circuit potentials, respectively.
  - Defines its source such that $v_{ds} \geq 0$.
  - Performs best when its source is at the lowest circuit potential.
  - Has increasing r_{ds} "on" resistance as source potential increases.
- A p-channel enhancement-mode device
  - Turns on and off when its gate connects to the lowest and highest circuit potentials, respectively.
  - Defines its source such that $v_{ds} \leq 0$.
  - Performs best when its source is at the highest circuit potential.
  - Has increasing r_{ds} "on" resistance as source potential decreases.

Switch applications that combine the preceding MOSFETs

- Use series-connected devices with shared gate connections so that one transistor functions as a switch (when on) while the other transistor functions as a load (when off) given HIGH and LOW gate voltages. (This is the basis for the CMOS digital inverter.)
- Use parallel-connected devices with the gate of one HIGH while the other is LOW so that both transistors operate in the on or off state. (This is the basis for the CMOS analog switch or transmission gate.)
  - The r_{ds} "on" resistance is approximately constant regardless of the voltage on either side of the switch.
  - The switch stores channel charge when turned on and releases channel charge when turned off, a problem for some applications.
  - A set of three CMOS analog switches can be used to improve off-state isolation by having one of them divert signals to ground.

Valve applications have yet to be explored in detail.

5.4 SPICE Analysis

SPICE achieves the representation of a MOSFET using a .model statement of the form:

.model < name > NMOS [model parameters]

or

.model < name > PMOS [model parameters]

for n-channel and p-channel MOSFETs, respectively. The static .model MOSFET parameters mentioned in Sections 5.1 and 5.2 appear in Table 5.1. The parameters that apply to pending complications are highlighted in blue.

Table 5.1: Elementary Static MOSFET SPICE Parameters

Symbol	SPICE Keyword	Parameter Name	Default Value	Unit
K'	KP	Transconductance parameter	2×10^{-5}	A/V^2
V_{T0}	VTO	(Zero-v_{bs}) Threshold voltage	0	V
λ	LAMBDA	Channel-length modulation	0	V^{-1}
γ	GAMMA	Body-effect parameter	0	$V^{1/2}$
$2\phi_f$	PHI	Surface potential (at inversion)	0.6	V
W	W	Channel width	†	meter
L	L	Channel length	†	meter

† If W and L are not specified, $W/L = 1$

Note: When dimensions vary between devices, as in an integrated circuit, it is convenient to specify W and L values outside the .model statement. For example,

M1 2 1 0 0 MOSFET W=10u L=2u

describes a MOSFET for which $W/L = 5$.

If MOSFET process details are known—an unlikely situation for mere mortal application engineers— TOX (m), UO (cm^2/V-s), and NSUB (cm^{-3}) can be specified for insulator thickness, channel-carrier surface mobility, and substrate doping concentration, respectively. In turn, SPICE will use these parameters to calculate KP, GAMMA, and PHI.

Parameters RS, RD, RG, and RB model parasitic terminal resistances that are in series with the source (r_s), drain (r_d), gate (r_g), and body (r_b).

Example 5.6

An n-channel enhancement-mode MOSFET features $K'W/L = 2$ mA/V^2 and $V_T = 1.5$ V. Use SPICE to simulate the MOSFET test circuit shown in Fig. 5.32a, then plot $\sqrt{i_d}$ vs V_{ds}.

Solution

We use the following SPICE code:

```
* MOSFET Test Circuit

Vds     1       0       1
M1      1       1       0       0       MOSFET

.model  MOSFET  NMOS    (KP = 2m, VTO = 1.5)

.dc     Vds     0       5       0.01
.probe

.end
```

The MOSFET nodes are specified in the order of drain, gate, source, and body connections. The .dc statement sweeps V_{ds} from 0 V to 5 V.

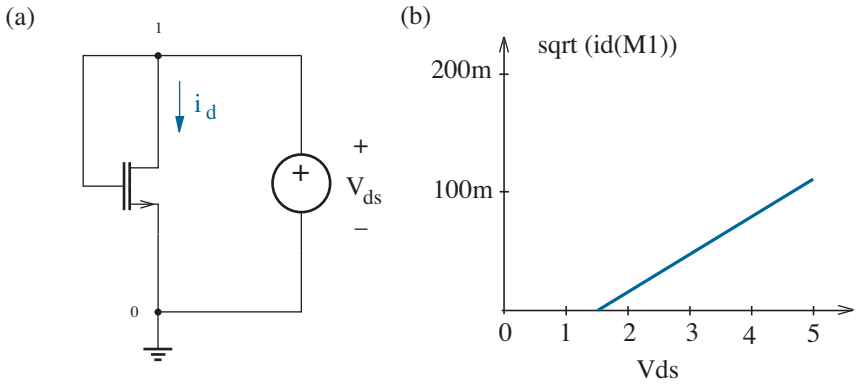


Figure 5.32: (a) MOSFET circuit for Example 5.6; (b) .probe plot.

Figure 5.32b shows the result when .probe is used to plot sqrt(id(M1)) following circuit simulation. The curve is a straight line that intersects the V_{ds} axis at 1.5 V. Thus, we observe an alternative measurement procedure for determining the threshold voltage for an *enhancement-mode* MOSFET. It is easily demonstrated that the slope of the curve is $\sqrt{K'W/2L}$.

MOSFET Capacitance

We will use SPICE to determine the speed of analog and digital circuits, and this will require the specification of several MOSFET capacitive effects. MOSFET capacitance between the gate and source (C_{gs}) and capacitance between the gate and drain (C_{gd}) have two components:

$$C_{gs} = C_{gs}\,(\text{overlap}) + C_{gs}\,(\text{channel}) \tag{5.27}$$

and

$$C_{gd} = C_{gd}\,(\text{overlap}) + C_{gd}\,(\text{channel})\,. \tag{5.28}$$

The first component is a constant parasitic capacitance that reflects the extent of overlap between the gate and the MOSFET $n+$ end regions. Figure 5.33 shows the related geometry. Design values for width and length are W and L, respectively. Notwithstanding, the $n+$ regions tend to spread laterally by diffusion during the MOSFET fabrication process so that the actual channel length is $L_{eff} = L - 2L_d$ (and treated as smaller L, for now). The overlap capacitances are given by

$$C_{gs}\,(\text{overlap}) = CGSO \cdot W \tag{5.29}$$

and

$$C_{gd}\,(\text{overlap}) = CGDO \cdot W\,. \tag{5.30}$$

SPICE parameters CGSO and CGDO are specified in F/m. The extent of the overlap is typically 0.2 μm or less in an ordinary MOSFET process.

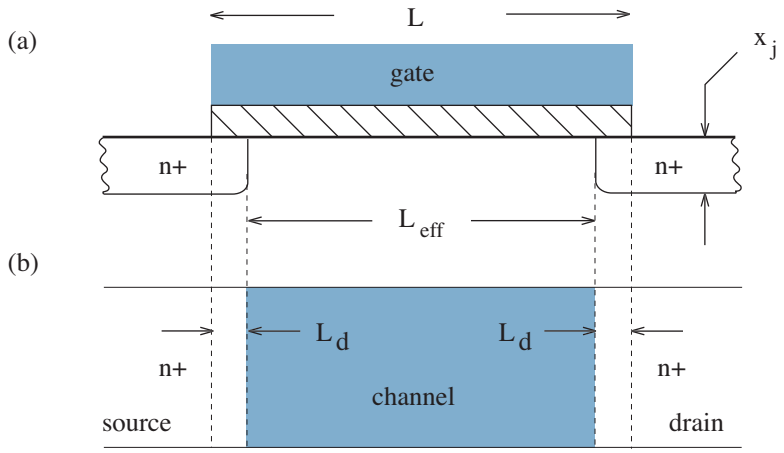


Figure 5.33: MOSFET overlap geometry: (a) side view; (b) top view.

The second component in C_{gs} or C_{gd} is a voltage-dependent capacitance that reflects the influence of terminal voltages on MOSFET channel charge. When v_{ds} is small (as for the resistive mode), the total gate capacitance—$C_{ox}WL$ —is almost evenly split between C_{gs} (channel) and C_{gd} (channel). As v_{ds} increases, the drain voltage gradually loses channel-charge influence, and C_{gd} (channel) decreases. If v_{ds} is sufficiently large to ensure saturation, the drain-voltage influence is lost entirely, and C_{gd} (channel) = 0.

A simple calculation provides the C_{gs} channel component at saturation. With reference to Fig. 5.8 and Eq. 5.4, the differential charge $|dQ|$ located between channel position y and $y + dy$ is given by

$$|dQ| = WC_{ox}\,(v_{gs} - \psi - V_T)\,dy\,. \tag{5.31}$$

And given Eqs. 5.3 and 5.4, we infer

$$dy = C_{ox}\frac{\mu_e W}{i_d}\,(v_{gs} - \psi - V_T)\,d\psi\,. \tag{5.32}$$

Thus,

$$|Q| = \frac{(WC_{ox})^2\mu_e}{i_d}\int_0^{v_{ds}}(v_{gs} - \psi - V_T)^2\,d\psi\,. \tag{5.33}$$

The upper limit of integration is $v_{gs} - V_T$ at the onset of saturation, so

$$|Q| = \frac{1}{3}\,\frac{(WC_{ox})^2\mu_e}{i_d}\,(v_{gs} - V_T)^3\,. \tag{5.34}$$

Next, we substitute Eq. 5.7 for i_d at saturation to obtain

$$|Q| = \frac{2}{3}WL\,C_{ox}\,(v_{gs} - V_T)\,. \tag{5.35}$$

In turn,

$$C_{gs}\ (\text{channel}) = \frac{\partial|\mathrm{Q}|}{\partial \mathrm{v_{gs}}} = \frac{2}{3}\mathrm{WL\,C_{ox}}\,. \tag{5.36}$$

The interested reader is referred to the bibliographical essay for the details of C_{gs} (channel) and C_{gd} (channel) behavior in the resistive mode.

Both C_{gs} and C_{gd} channel-charge components require the inclusion of SPICE parameter TOX to support the related calculation of $C_{ox} = \epsilon_{ox}/t_{ox}$. A reasonable guess must suffice if the insulator thickness is not known.

Apart from C_{gs} and C_{gd}, a three-terminal MOSFET also includes C_{ds}, the capacitance between drain and source. This capacitance is always small as a consequence of the relatively large drain/source separation distance. Thus, SPICE models do not provide for a C_{ds} parameter.

In our discussion of static current-voltage relationships, we treated the MOSFET as a three-terminal device. The fourth body terminal does not pass current—it changes the threshold voltage when $v_{bs} \neq 0$ (Chapter 9). This limited role does not apply under high-frequency dynamic conditions.

Figure 5.34 shows the complete dynamic four-terminal MOSFET model. Three new capacitive components can pass ac current to the body terminal, so each must be properly characterized for accurate SPICE simulations.

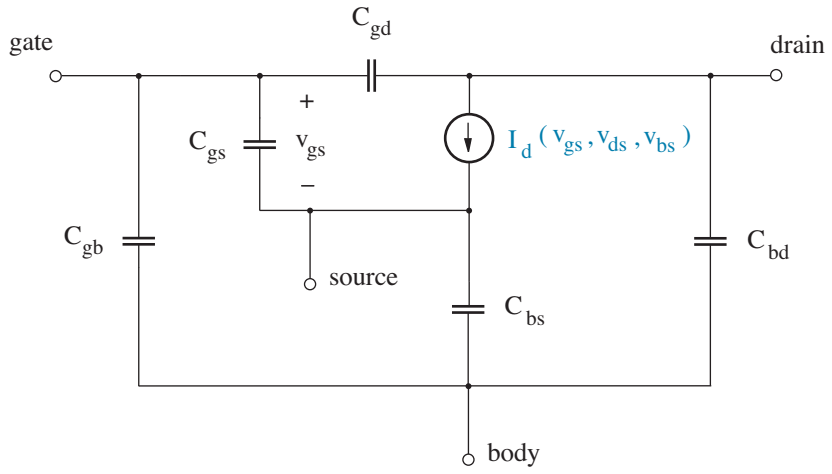


Figure 5.34: Four-terminal MOSFET model.

The body-to-source (C_{bs}) and the body-to-drain (C_{bd}) capacitances are both associated with pn junctions that isolate the MOSFET end regions from the substrate. Both components are bias dependent as demonstrated in Chapters 2 and 3, and both can be specified in one of two ways.

Method 1:

Total capacitance is known for zero-bias conditions. In turn,

$$C_{bs} = CBS \left(1 - \frac{v_{bs}}{PB}\right)^{-MJ} \tag{5.37}$$

and

$$C_{bd} = CBD \left(1 - \frac{v_{bd}}{PB}\right)^{-MJ}. \tag{5.38}$$

Here, CBS and CBD are the zero-bias capacitance values, PB is the junction built-in voltage, and MJ is the junction grading coefficient. The latter two SPICE parameters often assume defaults of 0.8 V and 0.5, respectively.

Method 2:

The MOSFET is one of many with different sizes but similar fabrication (as expected for integrated circuits). For this case, individual MOSFETs feature SPICE statements of the form

```
M1    3     2     1     0     MOSFET   W=4u    L=2u
+     AS=50p      PS=30u      AD=50p   PD=30u
```

Here, geometric parameters AS and AP are the source and drain areas (m^2), and PS and PD are the source and drain perimeters (m) —see Fig. 5.35.

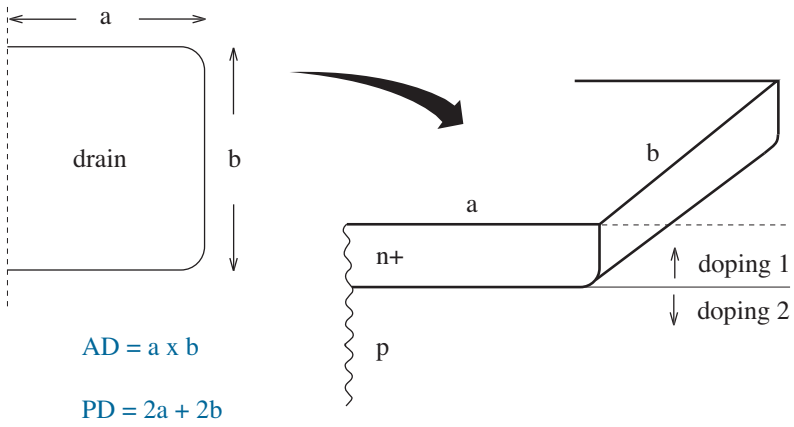


Figure 5.35: MOSFET drain geometry used to determine C_{bd}.

The C_{bs} and C_{bd} capacitance values are calculated from the relations

$$C_{bs} = CJ \cdot AS \left(1 - \frac{v_{bs}}{PB}\right)^{-MJ} + CJSW \cdot PS \left(1 - \frac{v_{bs}}{PB}\right)^{-MJSW} \quad (5.39)$$

and

$$C_{bd} = CJ \cdot AD \left(1 - \frac{v_{bd}}{PB}\right)^{-MJ} + CJSW \cdot PD \left(1 - \frac{v_{bd}}{PB}\right)^{-MJSW} . \quad (5.40)$$

SPICE parameters CJ and CJSW are the areal junction capacitance (F/m^2) and the per-unit-legth sidewall junction capacitance (F/m), respectively. The latter needs to be specified, since the doping concentration along the sides of the source or drain pn junction is generally different from the doping elsewhere in the substrate. The default sidewall junction grading coefficient is 0.33 to accommodate a linearly graded junction profile.

The gate-to-body capacitance (C_{gb}) has two components:

$$C_{gb} = C_{gb}\,(\text{overlap}) + C_{gb}\,(\text{channel})\,. \tag{5.41}$$

The first component is a constant parasitic capacitance that reflects overlap between the gate and the substrate outside the MOSFET channel region. In some cases, this includes capacitive coupling from gate-connected parts such as metal wire runs—the alternative is to lump interconnect parasitics into a single capacitance that is made separate from the MOSFET model. The overlap capacitance is assumed to be proportional to the device width. Specifically,

$$C_{gb}\,(\text{overlap}) = CGBO \cdot W\,. \tag{5.42}$$

SPICE parameter CGBO is specified in F/m.

The second component is a voltage-dependent capacitance that reflects coupling to the substrate region directly beneath the gate. For n-channel MOSFETs, the capacitance is *zero* when $v_{gs} > V_T$ —the inversion layer acts as a conductive shield. In contrast, the capacitance is

$$C_{gb}\,(\text{channel}) = WL\,C_{ox} \tag{5.43}$$

when v_{gs} is less than the flat-band voltage such that the insulator/substrate interface is under accumulation. For model simplicity, C_{gb} (channel) varies linearly between zero and WLC_{ox} at intermediate subthreshold voltages:

$$C_{gb}\,(\text{channel}) = WL\,C_{ox}\,\frac{V_T - v_{gs}}{2\phi_f}\,, \tag{5.44}$$

where $2\phi_f$ (PHI) is the change in surface potential that achieves inversion relative to flat-band conditions. Note that C_{gb} (channel) is always zero if TOX is not specified.

Exercise 5.6 A MOSFET with 50-nm insulator thickness is designed to have $W = 15$ μm and $L = 2$ μm. Nevertheless, the effective channel length is only 1.7 μm after fabrication. Determine $CGSO$ and $CGDO$.

Ans: $CGSO = CGDO = 1.04 \times 10^{-10}$ F/m

Exercise 5.7 A MOSFET features n+ source and drain regions with 0.5-μm junction depth and 5×10^{19} cm^{-3} doping concentration. The p concentrations in the source-drain peripheral regions and the substrate are 5×10^{16} cm^{-3} and 2×10^{15} cm^{-3}, respectively. Determine CJ and $CJSW$.

Ans: $CJ = 7.65 \times 10^{-4}$ F/m^2 $\quad CJSW = 3.27 \times 10^{-10}$ F/m

Capacitance Summary

Readers will appreciate Tables 5.2 and 5.3, which summarize the SPICE capacitance parameters and the related capacitance expressions.

Table 5.2: Elementary Dynamic MOSFET SPICE Parameters

Symbol	SPICE Keyword	Parameter Name	Default Value	Unit
C_{gso}	CGSO	Gate/source overlap cap.	0	F/m
C_{gdo}	CGDO	Gate/drain overlap cap.	0	F/m
C_{gbo}	CGBO	Gate/body overlap cap.	0	F/m
t_{ox}	TOX	Gate insulator thickness	∞	m
L_d	LD	Lateral diffusion	0	m
C_{bs}	CBS	Zero-bias body/source cap.	0	F
C_{bd}	CBD	Zero-bias body/drain cap.	0	F
C_j	CJ	Zero-bias areal junction cap.	0	F/m^2
M_j	MJ	Areal junction grading coeff.	0.5	
$C_{j,sw}$	CJSW	Zero-bias side-wall cap.	0	F/m
$M_{j,sw}$	MJSW	Side-wall grading coeff.	0.33	
ϕ_j	PB	Built-in junction voltage	0.8	V

Table 5.3: Inter-terminal Capacitance Expressions (Excluding Overlap)

	Gate	Drain	Body
Source	0 (sub) 1/2 C_g (res) 2/3 C_g (sat)	0	Eq. 5.37 or Eq. 5.39
Gate	—	0 (sub) 1/2 C_g (res) 0 (sat)	Eq. 5.44 ($v_{gs} < V_T$) 0 ($v_{gs} \geq V_T$)
Drain		—	Eq. 5.38 or Eq. 5.40

$C_g = WLC_{ox}$.

sub → Sub-threshold, res → Resistive mode, sat → Saturation mode

Appendix: MOS Charge Control

A quantitative one-dimensional analysis of the MOS depletion condition begins (quite naturally) with the depletion approximation first encountered in Chapter 2. We assume that the substrate is totally devoid of holes over the interval $0 < x < w$, but it is neutral for $x \geq w$. This reflects the net charge-density ρ distribution shown in Fig. 5.36a. Next, we use Gauss' law to determine a consistent electric field $\mathcal{E}$. Specifically,

$$\epsilon_{Si} \frac{d\mathcal{E}}{dx} = \rho(x) \,, \tag{5.45}$$

where ϵ_{Si} is the dielectric permittivity of silicon. We (graphically) integrate Eq. 5.45 to obtain Fig. 5.36b with the required zero electric field at $x = w$. Finally, we determine a consistent electric potential ϕ using the relation

$$-\frac{d\phi}{dx} = \mathcal{E}(x) \,, \tag{5.46}$$

which we (graphically) integrate to obtain the distribution shown in Fig. 5.36c. The integration is subject to zero reference potential at $x = w$.

As a result of the preceding calculations, we find that the substrate surface potential ϕ_o is given by

$$\phi_o = \frac{qN_a w^2}{2\epsilon_{Si}} \,. \tag{5.47}$$

And when this expression is rearranged to give w in terms of ϕ_o, we have

$$w = \sqrt{\frac{2\epsilon_{Si}}{qN_a}\,\phi_o} \,. \tag{5.48}$$

Nevertheless, some further manipulation offers a more convenient relation of the form

$$w = \sqrt{2}\, L_d \sqrt{\frac{q\phi_o}{kT}} \,, \tag{5.49}$$

where kT/q is the familiar thermal voltage (25.9 mV at room temperature) and

$$L_d = \sqrt{\frac{\epsilon_{Si} kT}{q^2 N_a}} \tag{5.50}$$

is the **extrinsic Debye length** (typically about 0.13 μm for silicon with $N_a = 10^{15}$ cm^{-3} and $\epsilon_{Si} = 11.8 \times 8.854 \times 10^{-14}$ C/V-cm). A more careful analysis of the MOS capacitor shows that the edge of the depletion layer is actually somewhat smeared, and the transition between total depletion and neutrality develops over several extrinsic Debye lengths. Notwithstanding, the averaging effects of two integrations (one for $\mathcal{E}$, and the other for ϕ) diminish the prospects for improvement over the depletion approximation. The same smearing applies at the depletion-layer edges near a pn junction.

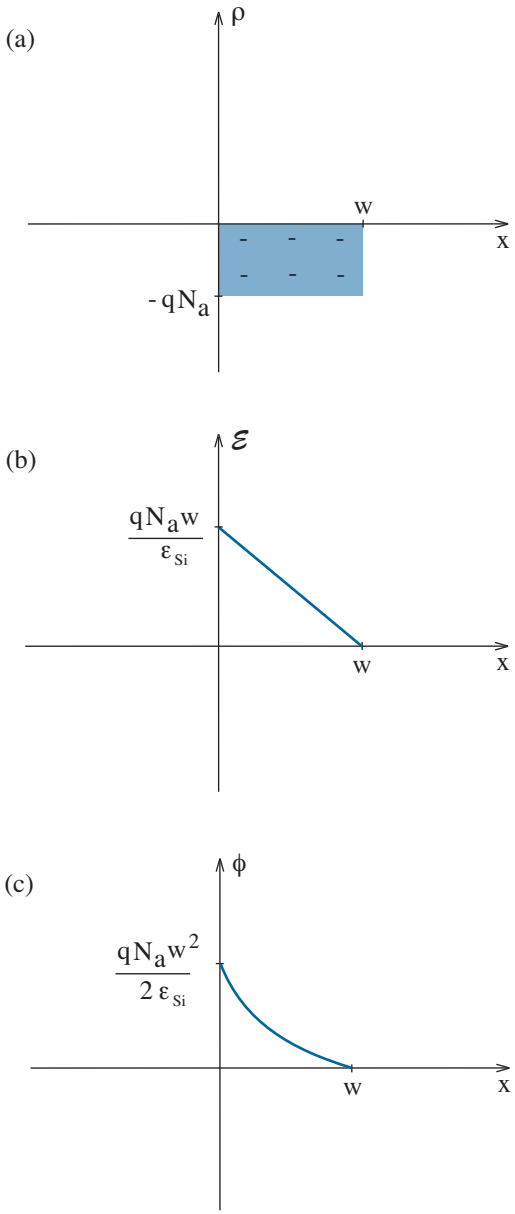


Figure 5.36: Distributions in the substrate portion of the MOS capacitor subject to positive surface potential and hole depletion without inversion: (a) net charge density; (b) electric field; (c) electric potential. The position $x = 0$ corresponds to the insulator/substrate interface (see Fig. 5.4).

Unfortunately, we are not quite finished, since we do not have a "handle" on potential ϕ_o. We need to relate it to v_{gb}, which is directly under control. Since the electric displacement is continuous at $x = 0$, we have

$$\epsilon_{ox}\,\mathcal{E}_{\rm ox}|_{0-} = \epsilon_{Si}\,\mathcal{E}_{\rm Si}|_{0+}\,, \tag{5.51}$$

where ϵ_{ox} is the dielectric permittivity of silicon dioxide. Then with the help of Fig. 5.36b, and assuming $V_{fb} = 0$, we rewrite Eq. 5.51 in the form

$$\epsilon_{ox}\left(\frac{v_{gb} - \phi_o}{t_{ox}}\right) = qN_a w\,. \tag{5.52}$$

Note that the right-hand side of this expression is simply the negative of the total areal (per-unit-area) depletion charge in the substrate ($-Q_d$).

For convenience, we designate $C_{ox} = \epsilon_{ox}/t_{ox}$ and $C_{fb} = \epsilon_{Si}/L_d$ as areal insulator and **flat-band** capacitances, respectively. (In the latter case, a rigorous analysis shows that $C_{fb} = -\Delta Q_d/\Delta\phi_o$ in the limit as $\phi_o \to 0$.) When these definitions apply in conjunction with Eq. 5.49, we transform Eq. 5.52 as follows:

$$C_{ox}\,(v_{gb} - \phi_o) = qN_a\sqrt{2}\,L_d\sqrt{\frac{q\phi_o}{kT}} = \sqrt{2}\,\left(\frac{qN_a{L_d}^2}{\epsilon_{Si}}\right)\,C_{fb}\sqrt{\frac{q\phi_o}{kT}}\,. \tag{5.53}$$

We apply Eq. 5.50 to eliminate ${L_d}^2$. Then solving for v_{gb}, we have

$$v_{gb} = \phi_o + \gamma\sqrt{\phi_o}\,, \tag{5.54}$$

where γ is the **body-effect parameter** given by

$$\gamma = \sqrt{2}\,\frac{C_{fb}}{C_{ox}}\sqrt{\frac{kT}{q}}\,. \tag{5.55}$$

This parameter can be measured (Problem 5.92), and it has broad influence. Typically $0.2 < \gamma < 0.8$ $\mathrm{V}^{1/2}$.

Equation 5.54 will soon capture our interest beyond the task at hand, but we still need to solve it for ϕ_o. The result is

$$\frac{4\phi_o}{\gamma^2} = \left[\sqrt{1 + \frac{4v_{gb}}{\gamma^2}} - 1\right]^2\,. \tag{5.56}$$

When writing Eq. 5.56, we have chosen to normalize ϕ_o and v_{gb} by the factor $\gamma^2/4$ to emphasize a general relation between two dimensionless variables. One can show that the relation is almost linear for typically large $4v_{gb}/\gamma^2$. Nevertheless, there are no wide-ranging approximations that simplify hand calculations requiring v_{gb} for a particular depletion depth w.

Our ability to control the substrate depletion depth is potentially useful. But how do we achieve inversion at the insulator/substrate interface?

To answer this question, we apply the Boltzmann relations that were used in Chapter 2 to determine the built-in potential across a pn junction. Our immediate objective is to specify the hole and electron concentrations at the insulator/substrate interface in terms of ϕ_o, the surface potential.

If the gate current is zero and v_{gb} is constant, the MOS capacitor is at equilibrium, and the hole and electron concentrations in the semiconductor substrate are given by

$$p = n_i \, e^{-q\psi/kT} \tag{5.57}$$

and

$$n = n_i \, e^{\,q\psi/kT}, \tag{5.58}$$

where $\psi = \phi + \phi'$, and ϕ' is a suitable constant. We use Eq. 5.57 (Eq. 5.58) to determine the ratio between the hole (electron) concentrations at the insulator/substrate interface and at the edge of the depletion layer.

For holes, with $\phi(0) = \phi_o$ and $\phi(w) = 0$, we have

$$\frac{p(0)}{p(w)} = \frac{n_i \, \exp\left[-q(\phi_o + \phi')/kT\right]}{n_i \, \exp\left[-q(0 \, + \phi')/kT\right]} = e^{-q\phi_o/kT}. \tag{5.59}$$

And since $p(w) = N_a$,

$$p(0) = N_a \, e^{-q\phi_o/kT}. \tag{5.60}$$

Thus, the hole concentration at the insulator/substrate interface becomes negligibly small when $\phi_o \gg kT/q$ (as expected under depletion conditions).

What about electrons? From Eq. 5.58,

$$\frac{n(0)}{n(w)} = \frac{n_i \, \exp\left[q(\phi_o + \phi')/kT\right]}{n_i \, \exp\left[q(0 \, + \phi')/kT\right]} = e^{\,q\phi_o/kT}. \tag{5.61}$$

And with $n(w) = {n_i}^2/N_a$,

$$n(0) = \frac{{n_i}^2}{N_a} \, e^{\,q\phi_o/kT}. \tag{5.62}$$

Thus, the electron concentration close to the insulator/substrate interface increases *exponentially* with increasing surface potential in units of kT/q. This ultimately produces the anticipated inversion condition in which the electron concentration exceeds the concentration of fixed ionized acceptors. The thin interfacial region is effectively n-type with $n(0) > N_a$.

Threshold Voltage

Having defined the $n(0)$ condition for inversion, we apply it to Eq. 5.62 to find the corresponding range of substrate surface potential ϕ_o. Specifically,

$$\phi_o > 2\left(\frac{kT}{q}\right)\ln\frac{N_a}{n_i} = 2\phi_f\ , \tag{5.63}$$

or $\phi_o > 0.59$ V for a p-type silicon substrate with uniform $N_a = 10^{15}$ cm^{-3}. However, we need to use Eq. 5.54 to find the gate-to-body voltage at the onset of inversion. The resulting threshold voltage is

$$V_T = 2\phi_f + \gamma\sqrt{2\phi_f}\ . \tag{5.64}$$

But in practice, V_T must be modified as follows:

$$V_T = 2\phi_f + \gamma\sqrt{2\phi_f} + V_{fb}\ , \tag{5.65}$$

where V_{fb} is the previously encountered MOS flat-band voltage given by

$$V_{fb} = \phi_{ms} - \frac{Q_i}{C_{ox}}\ . \tag{5.66}$$

The first term on the right-hand side of Eq. 5.66 is the metal-semiconductor **work-function difference**, which accounts for the difference in energy needed to remove an electron from an isolated slab of metal as opposed to an isolated slab of semiconductor. Positive ϕ_{ms} implies the presence of a built-in electric field that opposes the tendency for electrons to move from the semiconductor to the metal when the two materials are in close proximity. The inversion condition requires that the field be offset. For the case of aluminum gate material and a p-type silicon substrate with $N_a = 10^{15}$ cm^{-3}, $\phi_{ms} \approx -0.68$ V. However, modern MOS configurations feature polycrystalline silicon gate material and

$$\phi_{ms} = \mp 0.56\ \text{V}\ \mp\left(\frac{kT}{q}\right)\ln\frac{N}{n_i}\ . \tag{5.67}$$

Here, the first sign is - for an n^+ gate and + for a p^+ gate. The second sign is - for a p-type substrate and + for an n-type substrate. Equation 5.67 yields $\phi_{ms} \approx -0.85$ V for an n^+ gate and p-type $N = N_a = 10^{15}$ cm^{-3}.

The second term on the right-hand side of Eq. 5.66 is proportional to Q_i, the effective Si/SiO_2 interface charge density arising from imperfect oxide growth or external contaminants (especially sodium). Positive Q_i attracts electrons to the interface and *reduces* threshold voltage. The need to control Q_i is one of several reasons why semiconductor fabrication facilities require ultra-clean environments. Typically, $Q_i/q < 5 \times 10^{10}$ cm^{-2} for a "good" MOS process.

Both ϕ_{ms} and Q_i tend to be uncertain, even under the best conditions, so V_T is generally a measured quantity.

As v_{gb} is increased beyond V_T, the areal inversion-layer charge density $(-qN_I)$ increases exponentially with increasing ϕ_o, as suggested by Eq. 5.62. Conversely, the total areal depletion charge $(-Q_d)$ increases more slowly in proportion to $\sqrt{\phi_o}$, as suggested by Eq. 5.49. Thus, for $v_{gb} > V_T$, changes in v_{gb} primarily yield changes in the inversion layer. If we assume "zero" inversion-layer charge density when $v_{gb} = V_T$, then

$$qN_I \approx C_{ox}(v_{gb} - V_T)\,. \tag{5.68}$$

A consistent surface potential increases logarithmically with increasing v_{gb}, but the variation is so slow that we consider ϕ_o effectively "pinned" at $2\phi_f$.

After all this work, we must confess that the simple MOS capacitor is not a particularly useful device when acting alone. It merely displays a total capacitance given by the series combination of a fixed insulator capacitance and a voltage-dependent depletion (substrate) capacitance as in Fig. 5.37. The total capacitance value is

$$C_{\text{Total}} = \left[\frac{1}{C_{ox}} + \frac{1}{C_s}\right]^{-1}. \tag{5.69}$$

Notwithstanding, the capacitance-voltage characteristics of an isolated MOS capacitor can provide diagnostic information such as insulator thickness, substrate doping concentration, and interface charge density that is helpful for evaluating MOSFET fabrication runs (see Problems 5.88 and 5.89).

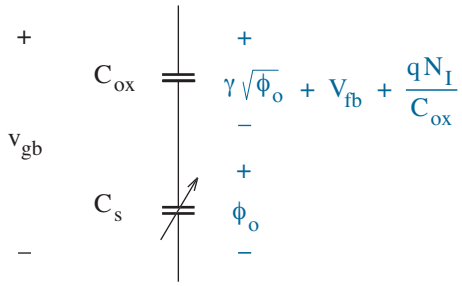


Figure 5.37: Equivalent circuit for an MOS capacitor. The inversion-layer charge density $-qN_I$ is consistent with Eq. 5.68 subject to $\phi_o \approx 2\phi_f$.

As part of a system, the MOS capacitor is the building block we require to provide an electrical bridge between switch contacts for transistor action. This functionality was the subject of Section 5.1. The MOS capacitor also serves as a building block for **dynamic random-access memory** (dRAM) (see Chapter 10) and for **charge-coupled devices** (see Problem 5.90).

Example 5.7

An MOS capacitor featuring $t_{ox} = 50$ nm and $N_a = 5 \times 10^{15}$ cm^{-3} is biased so that $\phi_o = 0.60$ V. Determine the extent of the depletion region, the electron concentration at the substrate surface, the surface condition, and the consistent gate-to-body voltage. Assume $V_{fb} = 0$.

Solution

We start with Eq. 5.50 for the extrinsic Debye length L_d. Specifically,

$$L_d = \sqrt{\frac{11.8 \times 8.854 \times 10^{-14} \text{ C/V–cm} \times 0.0259 \text{ V}}{1.602 \times 10^{-19} \text{ C} \times 5 \times 10^{15} \text{ cm}^{-3}}} = 5.81 \times 10^{-6} \text{ cm} .$$

Then Eq. 5.49 yields $w = \sqrt{2}\, L_d \sqrt{q\phi_o/kT} = 0.396$ μm.

The electron concentration at the substrate surface follows Eq. 5.62:

$$n(0) = \frac{10^{20} \text{ cm}^{-6}}{5 \times 10^{15} \text{ cm}^{-3}}\, e^{\,0.60 \text{ V}/0.0259 \text{ V}} = 2.30 \times 10^{14} \text{ cm}^{-3} .$$

This is less than the doping concentration N_a, so the surface is not inverted.

With $\epsilon_{ox} = 3.9 \times 8.854 \times 10^{-14}$ F/cm, $C_{ox} = \epsilon_{ox}/t_{ox} = 6.91 \times 10^{-8}$ F/cm^2 and $C_{fb} = \epsilon_{Si}/L_d = 1.80 \times 10^{-7}$ F/cm^2. When applied to Eq. 5.55, these values yield $\gamma = 0.593$ V$^{1/2}$. In turn, from Eq. 5.54, $V_{gb} = 1.06$ V. Warning: Eq. 5.54 does not apply if an inversion layer is present.

Example 5.8

Estimate the threshold voltage for an MOS capacitor with $t_{ox} = 25$ nm, $N_a = 2 \times 10^{16}$ cm^{-3}, and $Q_i/q = 5 \times 10^{10}$ cm^{-2}. Assume that the gate is fabricated using n$^+$ polycrystalline silicon.

Solution

From Eq. 5.63, the substrate surface potential required for inversion is

$$2\phi_f = 2 \times 0.0259 \text{ V} \times \ln\left(\frac{2 \times 10^{16} \text{ cm}^{-3}}{10^{10} \text{ cm}^{-3}}\right) = 0.75 \text{ V} .$$

To find the consistent voltage sustained across the insulator, we require γ. So with $C_{ox} = \epsilon_{ox}/t_{ox} = 1.38 \times 10^{-7}$ F/cm^2 and $C_{fb} = \epsilon_{Si}/L_d = 3.60 \times 10^{-7}$ F/cm^2 ($L_d = 2.91 \times 10^{-6}$ cm), $\gamma = 0.594$ V$^{1/2}$. In turn, $\gamma\sqrt{2\phi_f} = 0.51$ V. We calculate the flat-band voltage by determining contributions from the work-function difference and the presence of fixed interface charge (Q_i). The former is $\phi_{ms} = -0.94$ V (Eq. 5.67 with two negative signs), and the latter is $-Q_i/C_{ox} = -0.06$ V. Thus, $V_{fb} = -1.00$ V. Finally, from Eq. 5.65,

$$V_T = 0.75 \text{ V} + 0.51 \text{ V} - 1.00 \text{ V} = 0.26 \text{ V} .$$

Problems

Section 5.1

Elementary MOSFET Characteristics

Readers who are wary of p-channel MOSFETs may wish to reflect upon pertinent discussion in Section 5.2 before confronting some of these problems.

5.1 Which of the following MOSFET descriptions do not apply to a real device?

(a) p-channel, enhancement-mode, $V_T = 1$ V.

(b) p-channel, depletion-mode, $V_T = 1$ V.

(c) n-channel, enhancement-mode, $V_T = 1$ V.

(d) n-channel, depletion-mode, $V_T = 1$ V.

5.2 Complete the following Table that applies to a set of n-channel MOSFETs.

i_d (mA)	v_{gs} (V)	v_{ds} (V)	K' (mA/V^2)	W/L	V_T (V)
	2.0	5.0	0.05	40	1.0
2	4.0	2.0	0.05		1.0
3	3.0	4.0	0.10	60	
4		3.0	0.08	25	0.5
5	6.5	10.0		8	1.5
6	6.0		0.04	50	1.0

5.3 Complete the following Table that applies to a set of n-channel MOSFETs.

i_d (mA)	v_{gs} (V)	v_{ds} (V)	K' (mA/V^2)	W/L	V_T (V)
4.5	2.3	1.8		80	1.4
2.0		3.9	0.06	56	0.7
0.5	4.7	3.7	0.08		0.9
3.4	3.2		0.05	24	0.5
	5.8	9.6	0.11	66	1.1
1.4	7.2	6.1	0.07	22	

5.4 Complete the following Table that applies to a set of p-channel MOSFETs.

i_d (mA)	v_{gs} (V)	v_{ds} (V)	K' (mA/V^2)	W/L	V_T (V)
-7.4		-4.2	0.06	51	-0.6
-1.6	-2.5	-0.5		32	-0.9
	-3.3	-2.8	0.09	86	-0.5
-4.6	-4.1		0.05	14	-0.3
-3.7	-2.6	-3.9	0.07		-1.1
-0.1	-0.9	-1.5	0.08	63	

5.5 Complete the following Table and identify the type of MOSFET.

i_d (mA)	v_{gs} (V)	v_{ds} (V)	K' (mA/V^2)	W/L	V_T (V)
5.1	1.5	3.2	0.06	35	
1.4	3.2		0.07	24	0.5
	0.5	5.1	0.05	42	0.8
-2.8	-2.5	-2.1		16	-0.4
-0.4		-1.7	0.04	57	0.7
2.3	-0.5	4.3	0.12		-1.2

5.6 Complete the following Table and identify the type of MOSFET.

i_d (mA)	v_{gs} (V)	v_{ds} (V)	K' (mA/V^2)	W/L	V_T (V)
	-0.4	-2.6	0.04	37	-0.5
5.0	0.5	2.8	0.09		-0.8
1.4		3.5	0.05	47	0.7
-8.3	-3.7	-6.2		12	1.1
-3.6	-2.9		0.08	58	-0.4
7.7	2.1	3.2	0.07	25	

5.7 An n-channel MOSFET ($W/L = 25$, $V_T = 1$ V, $t_{ox} = 40$ nm) operates with $i_d = 1$ mA, $v_{ds} = 5$ V, and $v_{gs} = 2$ V. Find the electron channel mobility.

5.8 A p-channel MOSFET ($W/L = 50$, $V_T = -1$ V, $\mu_h = 400$ cm^2/V-s) operates with $i_d = -0.5$ mA, $v_{ds} = -5$ V, and $v_{gs} = -2$ V. Find the thickness of the SiO_2 insulator.

5.9 The MOSFET shown in Fig. P5.9 has $K_n' = 50\ \mu\text{A/V}^2$ and $V_{Tn} = 1$ V. Complete the design so that $v_{out} = 3$ V.

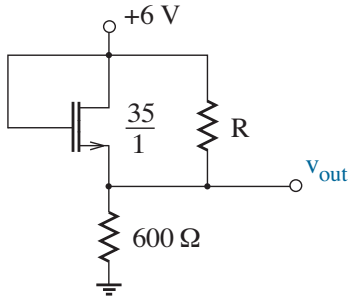


Figure P5.9

5.10 The MOSFET shown in Fig. P5.10 has $K_p' = 20\ \mu\text{A/V}^2$ and $V_{Tp} = -1$ V. Complete the design so that $v_{out} = 2$ V.

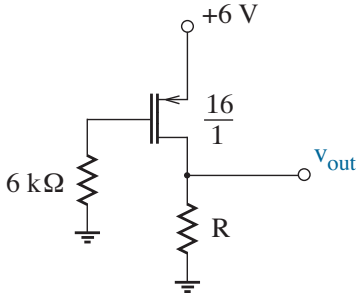


Figure P5.10

5.11 The MOSFET shown in Fig. P5.11 has $K_n' = 50\ \mu\text{A/V}^2$ and $V_{Tn} = 1$ V. Complete the design so that v_{out} bisects the minimum ensuring MOSFET saturation and the power-supply maximum.

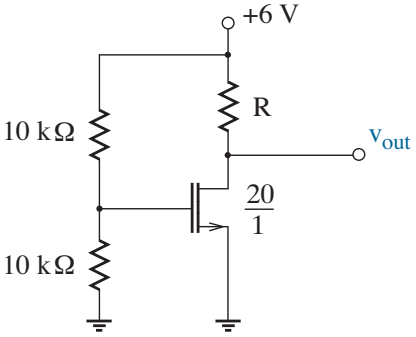


Figure P5.11

5.12 The MOSFET shown in Fig. P5.12 has $K_p' = 20\ \mu\text{A/V}^2$ and $V_{Tp} = -1$ V. Complete the design so that v_{out} bisects the maximum ensuring MOSFET saturation and the ground minimum.

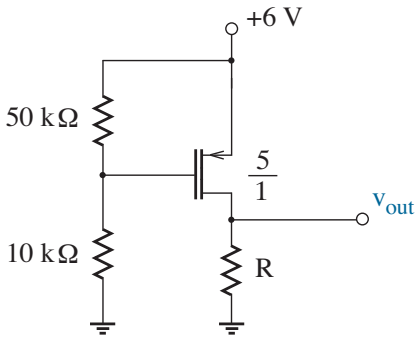


Figure P5.12

5.13 An n-channel enhancement-mode MOSFET has $K_n' = 50\ \mu\text{A/V}^2$, $W/L = 10$, and $V_{Tn} = 1$ V. Sketch the characteristic curves.

5.14 A p-channel enhancement-mode MOSFET has $K_p' = 20\ \mu\text{A/V}^2$, $W/L = 8$, and $V_{Tp} = -2$ V. Sketch the characteristic curves.

5.15 An n-channel depletion-mode MOSFET has $K_n' = 50\ \mu\text{A/V}^2$, $W/L = 40$, and $V_{Tn} = -1$ V. Sketch the characteristic curves.

Parameter Measurements

5.16 A MOSFET exhibits the following data with $v_{ds} = 0.1$ V and $v_{bs} = 0$. Find $K'W/L$ and V_T.

i_d (mA)	v_{gs} (V)
0.036	1.0
0.153	1.5
0.266	2.0
0.377	2.5
0.484	3.0
0.588	3.5

5.17 A MOSFET exhibits the following data with $v_{ds} = -0.1$ V and $v_{bs} = 0$. Find $K'W/L$ and V_T.

i_d (mA)	v_{gs} (V)
-0.216	-1.0
-0.522	-1.5
-0.825	-2.0
-1.126	-2.5
-1.423	-3.0
-1.717	-3.5

5.18 A MOSFET exhibits the following data with $v_{ds} = 5.0$ V and $v_{bs} = 0$. Determine $K'W/L$ and V_T. (Note the saturation mode of operation.)

i_d (mA)	v_{gs} (V)
0.000	1.0
0.225	1.5
1.600	2.0
4.225	2.5
8.100	3.0
13.23	3.5

5.19 The "mask defined" MOSFET length typically reflects the span of the gate material between source and drain. The actual channel length is shorter since the n^+ (or p^+) source and drain regions can diffuse beneath the gate during fabrication (see Fig. P5.19). Consider a set of MOSFETs for which $W = 40$ μm. The measured $K'W/L$ values are 2.700, 1.662, 1.200, and 0.939 mA/V^2 for mask-defined lengths of 1.0, 1.5, 2.0, and 2.5 μm, respectively. The devices are otherwise identical. Find the actual channel lengths.

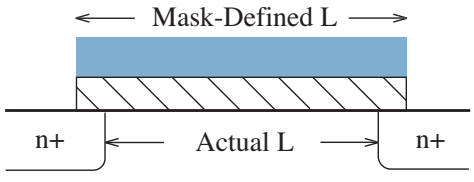


Figure P5.19

Section 5.2

5.20 An n-channel MOSFET operates in the resistive mode with $K'W/L = 6$ mA/V^2 and $V_T = 0.8$ V.

(a) Determine r_{ds} if $v_{gs} = 3$ V.

(b) Find the v_{gs} that halves the preceding r_{ds}.

5.21 An n-channel MOSFET operates in the resistive mode with $K' = 56$ μA/V^2 and $V_T = 0.6$ V. The channel length is 1.2 μm. Specify W/L so that $r_{ds} = 20$ Ω when $v_{gs} = 5$ V.

5.22 A p-channel MOSFET operates in the resistive mode with $K'W/L = 8$ mA/V^2 and $V_T = -0.8$ V.

(a) Determine r_{ds} if $v_{gs} = -8$ V.

(b) Find the v_{gs} that doubles the preceding r_{ds}.

5.23 A p-channel MOSFET operates in the resistive mode with $K' = 22$ μA/V^2 and $V_T = -0.6$ V. The channel length is 1.8 μm. Specify W/L so that $r_{ds} = 50$ Ω when $v_{gs} = -5$ V.

5.24 Figure P5.24 shows r_{ds} values plotted against L (the mask-defined length, Problem 5.19) for three different MOSFETs ($W = 80$ μm). The two series of measurements feature $v_{gs} = 4$ V and $v_{gs} = 8$ V. Indicate how the device model of Fig. 5.14 should be modified to accommodate these data.

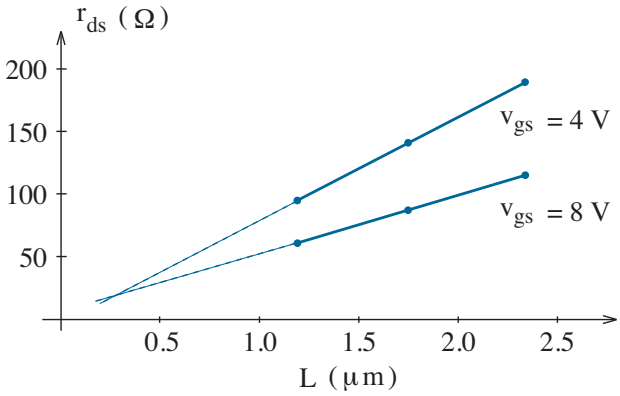


Figure P5.24

Section 5.3

MOSFETs as Switches

5.25 The circuit of Fig. P5.25 is intended to operate as a digital inverter. Complete the design so that $v_{out} = 0.2$ V (LOW) with 0.5-mW power dissipation when $v_{in} = 3.3$ V (HIGH). The MOSFET features $K_n' = 50\ \mu\text{A/V}^2$ and $V_{Tn} = 0.6$ V. Assume that the transistor has a particular r_{ds} in the resistive mode.

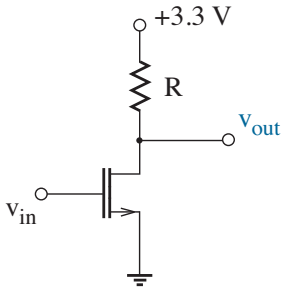


Figure P5.25

5.26 The circuit of Fig. P5.26 is intended to operate as a digital inverter. Complete the design so that $v_{out} = 3.1$ V (HIGH) with 0.5-mW power dissipation when $v_{in} = 0$ V (LOW). The MOSFET features $K_p' = 20\ \mu\text{A/V}^2$ and $V_{Tp} = -0.6$ V. Assume that the transistor has a particular r_{ds} in the resistive mode.

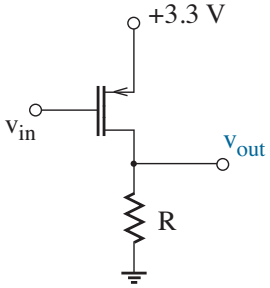


Figure P5.26

5.27 Figure P5.27 shows an **nMOS inverter** with an enhancement-mode MOSFET switch (M_1) and a depletion-mode MOSFET load (M_2)—now obsolete. Assume $K_n' = 50\ \mu\text{A/V}^2$, and $V_{Tn} = \pm 0.5$ V.

(a) Specify W/L values so that when v_{in} is HIGH (5 V), $v_{out} = 0.2$ V (LOW) and the total power dissipation is 0.5 mW. Assume that M_1 has a particular r_{ds} in the resistive mode.

(b) Subject to the design of part a, estimate the times needed to transition from HIGH to 2.5 V and from LOW to 2.5 V when the input makes an abrupt level transition. Neglect MOSFET capacitance in relation to the 10-pF load.

Note: nMOS inverters are easy to fabricate with only one extra processing step needed to transform M_2 from enhancement- to depletion-mode operation. The highly favorable power requirements for CMOS have more than made up for the relative complexity of the associated fabrication process.

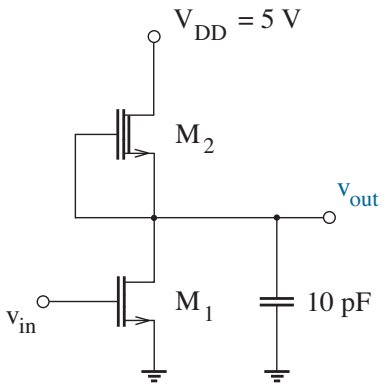


Figure P5.27

5.28 The n- and p-channel MOSFETs in the CMOS inverter of Fig. P5.28 feature $K_n' = 50\ \mu\text{A/V}^2$, $V_{Tn} = 0.6$, and $K_p' = 20\ \mu\text{A/V}^2$, $V_{Tp} = -0.6$ V. The HIGH level is 3.3 V, the LOW level is 0 V.

(a) Determine W/L for the n-channel MOSFET so that when v_{out} is HIGH, the time needed to transition to $V_{DD}/2$ (1.65 V) is 10 ns following an abrupt LOW-to-HIGH input transition. For a rough design, assume that M_1 has constant r_{ds} at $t = 0+$. Neglect MOSFET capacitance in relation to the 10-pF load.

(b) Determine W/L for the p-channel MOSFET so that when v_{out} is LOW, the time needed to transition to $V_{DD}/2$ (1.65 V) is 10 ns following an abrupt HIGH-to-LOW input transition. For a rough design, assume that M_2 has constant r_{ds} at $t = 0+$. Neglect MOSFET capacitance in relation to the 10-pF load.

(c) Compare the ratio of the results of parts a and b with the ratio of K_n' and K_p'.

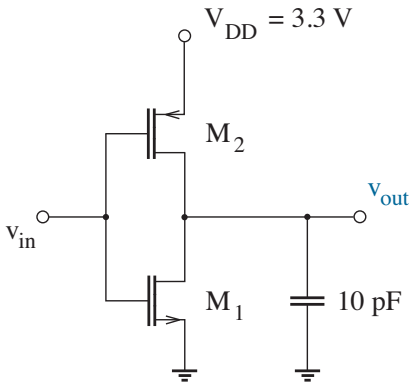


Figure P5.28

5.29 The circuit of Fig. P5.29 is to function as a variable voltage attenuator. Complete the design to achieve an attenuation factor of 0.1 for $v_x = 5$ V. The MOSFET has $K_n' = 50\ \mu\text{A/V}^2$, $V_{Tn} = 0.6$ V.

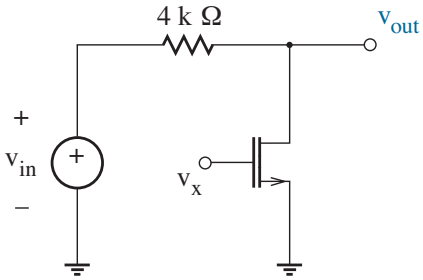


Figure P5.29

5.30 The circuit of Fig. P5.30 is to function as a variable voltage attenuator. Complete the design to achieve an attenuation factor of 0.2 for $v_x = 1$ V. The MOSFET has $K_p' = 20\ \mu\text{A/V}^2$, $V_{Tp} = -0.6$ V.

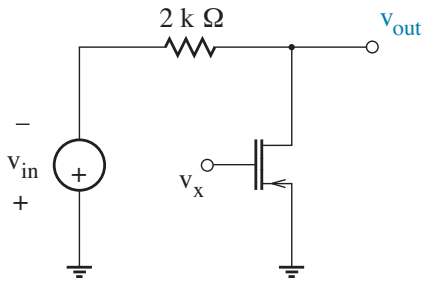


Figure P5.30

5.31 Figure P5.31 shows a simple sample-and-hold circuit with a single n-channel MOSFET for which $K_n' = 50\ \mu\text{A/V}^2$ and $V_{Tn} = 0.6$ V. The input voltage ranges from zero to 1.024 V. The switch on-state gate voltage is +5 V.

(a) Use the r_{ds} switch model to implement a rough design for W/L so that the worst-case change in capacitor voltage occurs with less than 1 % error over 10 ns. Neglect MOSFET input capacitance.

(b) Suppose the MOSFET is off, and $v_x = 0.5$ V. Subject to the design of part a, estimate the voltage acquisition error that results when the switch connects to $v_{in} = 0.6$ V over 10 ns.

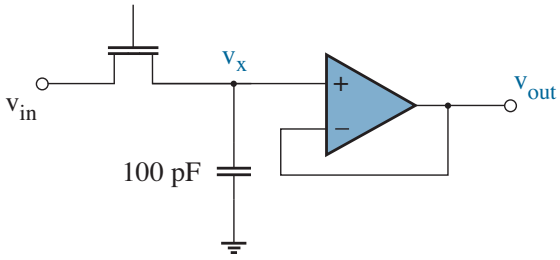


Figure P5.31

5.32 The MOSFET in the sample-and-hold circuit of Fig. P5.32 has $K_n' = 65\ \mu\text{A/V}^2$, $V_{Tn} = 0.6$ V, $C_{ox} = 80\ \text{nF/cm}^2$, $W = 20\ \mu$m, and $L = 1.2\ \mu$m. During sampling, the MOSFET gate voltage is +5 V. After a sample, roughly half of the MOSFET channel charge is absorbed by the capacitor, the rest transfers to the input source. Let $v_{in} = 1$ V. Determine the error in the sampled voltage.

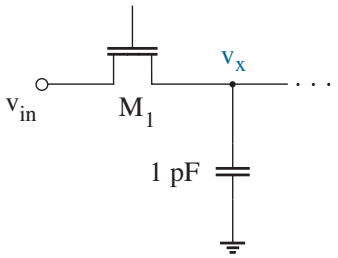


Figure P5.32

5.33 The sample-and-hold of Fig. P5.33 is similar to that of Fig. P5.32 except for M_2, a "dummy" switch connected in parallel with a short circuit. The M_1 and M_2 gates connect to 5-V non-overlapping clock signals: ϕ is HIGH when $\overline{\phi}$ is LOW, and conversely. Let M_1 have the specifications in Problem 5.32, and let M_2 have the same features apart from W_2. Find W_2 so that the sampling error is minimized.

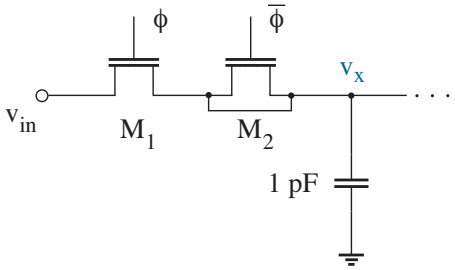


Figure P5.33

5.34 Use MOSFET switches and 10-μF capacitors to design a charge-pump circuit that delivers 6 V to a resistive load from a 3-V dc supply. Assume a 100-kHz switching frequency. Specify the maximum load current that is consistent with an output ripple voltage of 0.1 V.

5.35 Use MOSFET switches and 3.3-μF capacitors to design a charge-pump circuit that delivers 3 V to a resistive load from a 9-V dc supply. Assume a 100-kHz switching frequency. Specify the maximum load current that is consistent with an output ripple voltage of 0.2 V.

5.36 Consider a 4-switch dc-to-ac inverter circuit like that shown in Fig. 5.25. The voltage observed across load R has the time dependence of Fig. P5.36 over one period T.

(a) Sketch the consistent time behavior for switches S_1 through S_4.

(b) Show that the power delivered to the load in the n-th harmonic is

$$P_n = \frac{1}{R}\left(\frac{4V^+}{n\pi}\right)^2 \cos^2 n\alpha\ .$$

(c) Specify the fraction of the total available power delivered in the first harmonic subject to $P_3 = 0$.

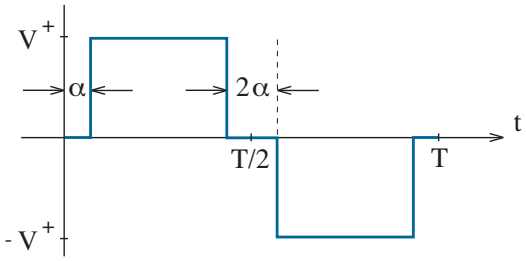


Figure P5.36

5.37 Consider a 4-switch dc-to-ac inverter circuit like that shown in Fig. 5.25, but include an inductor L in series with R.

(a) Sketch the implementation of the circuit with power MOSFETs.

(b) Show that continuous inductor current implies maximum and minimum switch currents

$$I_{max} = -I_{min} = \frac{V^+}{R}\left[\frac{1 - e^{-T/2\tau}}{1 + e^{-T/2\tau}}\right] ,$$

where T is the switching period and $\tau = L/R$.

(c) Whereas I_{min} is negative, the switches must be designed to conduct current in either direction. Show how the circuit of part a is modified with a set of diodes to accommodate this behavior. (As noted in the text, these diodes are generally available as parasitic elements.)

5.38 Consider the 2-switch dc-to-ac inverter circuit shown in Fig. P5.38 with $C = 100\ \mu$F and $R = 100\ \Omega$. The switches alternately open and close with period T and 50% duty cycle.

(a) Show that in the steady state, the load voltage varies over the range $V_m/2 \pm \Delta v$ where

$$\Delta v = \frac{V_m}{2} \sinh\left(\frac{T}{8RC}\right) .$$

Hint: Consider steady-state values for v_1 and v_2 at the beginnings and ends of each half cycle.

(b) Let $V_m = 20$ V. Find the switching frequency consistent with $\Delta v = 0.25$ V.

(c) Estimate the power delivered to the load in the first harmonic subject to the design of part b.

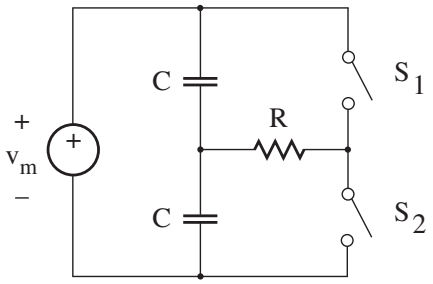


Figure P5.38

5.39 Figure P5.39 shows an ac-to-ac converter as for a light-dimming controller. The input has the form $v_{in}(t) = V_m \sin \omega t$. The switch is open over angular intervals $0 \le \theta \le \alpha$ and $2\pi - \alpha \le \theta \le 2\pi$, and it is closed otherwise.

(a) Sketch the output waveform for load R.

(b) Show that the rms voltage across R is given by

$$\frac{V_m}{\sqrt{2}} \sqrt{1 - \frac{\alpha}{\pi} - \frac{\sin 2\alpha}{2\pi}} .$$

(c) Show how the circuit can be implemented using power MOSFETs.

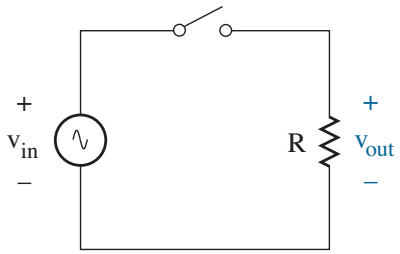


Figure P5.39

MOSFETs as Valves

5.40 The MOSFET in Fig. P5.40 has $K'W/L = 4$ mA/V^2 and $V_T = 1$ V. Design for $v_{out} = 4$ V.

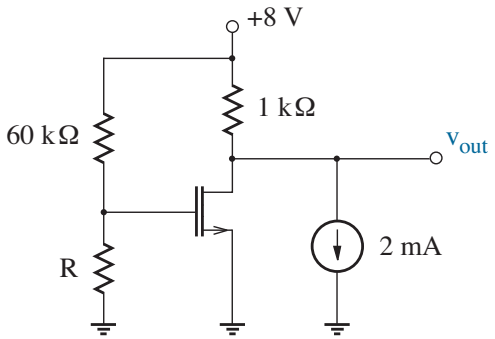


Figure P5.40

5.41 The MOSFET in Fig. P5.41 has $K'W/L = 1$ mA/V^2 and $V_T = -1$ V. Design for $v_{out} = 4$ V.

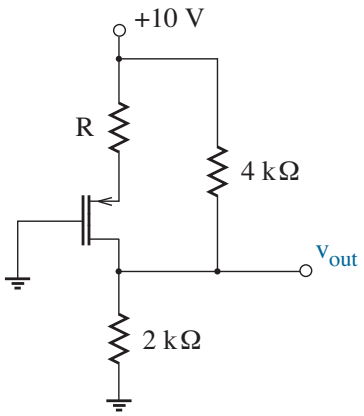


Figure P5.41

5.42 The MOSFETs in Fig. P5.42 have $K'W/L = 0.5$ mA/V^2 and $V_T = 1$ V. Design for $v_{out} = 2$ V.

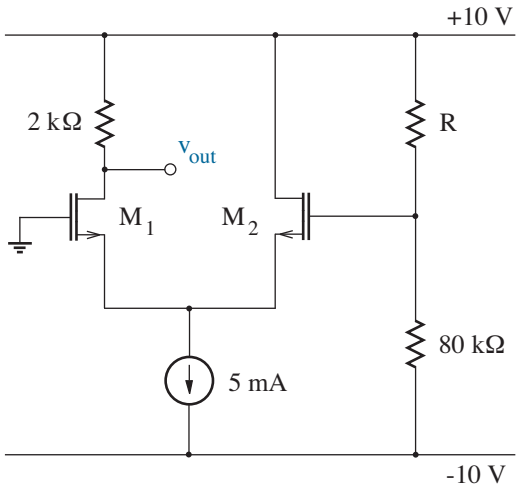


Figure P5.42

5.43 The MOSFETs in Fig. P5.43 have $K'W/L = 1$ mA/V^2 and $V_T = 1$ V. Design for $v_{out} = 8$ V.

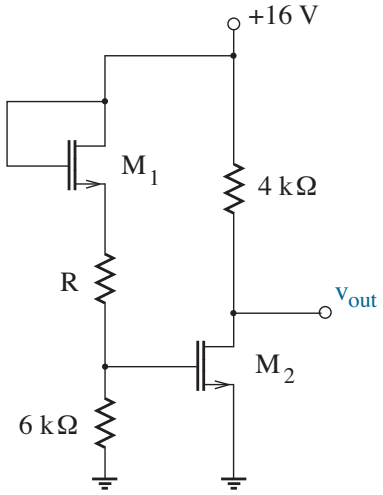


Figure P5.43

5.44 The MOSFETs in Fig. P5.44 have $K'W/L = 2$ mA/V^2 and $V_T = -1$ V. Design for $v_{out} = 2$ V.

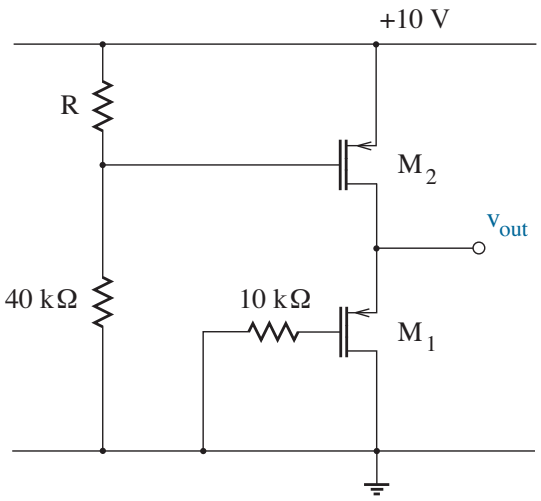


Figure P5.44

5.45 The n- and p-channel MOSFETs in Fig. P5.45 feature $K_n'W/L = 2$ mA/V^2, $V_{Tn} = 1$ V, and $K_p'W/L = 0.25$ mA/V^2, $V_{Tp} = -1$ V, respectively. Design for $v_{out} = 3$ V.

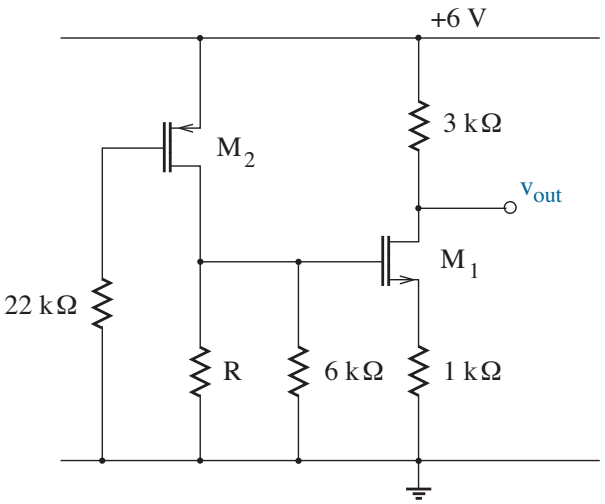


Figure P5.45

5.46 The MOSFETs in Fig. P5.46 have $K'W/L = 2$ mA/V^2 and $V_T = 1$ V. Design for $v_{out} = 6$ V.

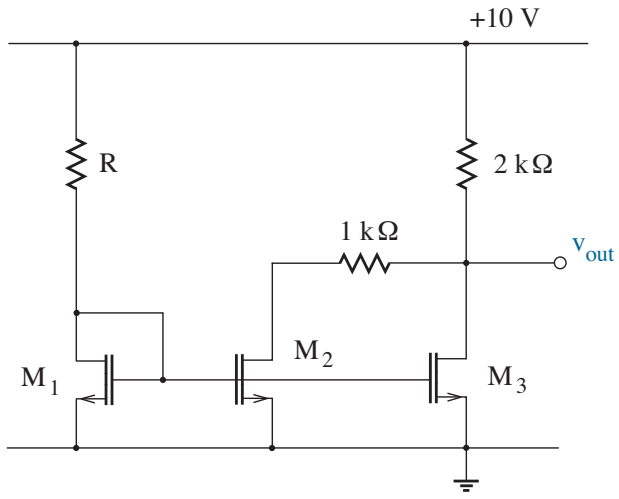


Figure P5.46

5.47 The MOSFETs in Fig. P5.47 have $K_n' = 50$ μA/V^2 and $V_T = \pm 0.5$ V. Specify W/L values for $v_{out} = 2.5$ V and 1-mW total dissipated power.

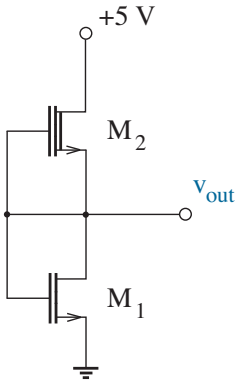


Figure P5.47

5.48 The MOSFETs in Fig. P5.48 have $K_n' = 50$ μA/V^2 and $V_T = \pm 0.5$ V. Specify W/L values for $v_{out} = 0$ V and 7-mW power dissipation in the 2-mA current source.

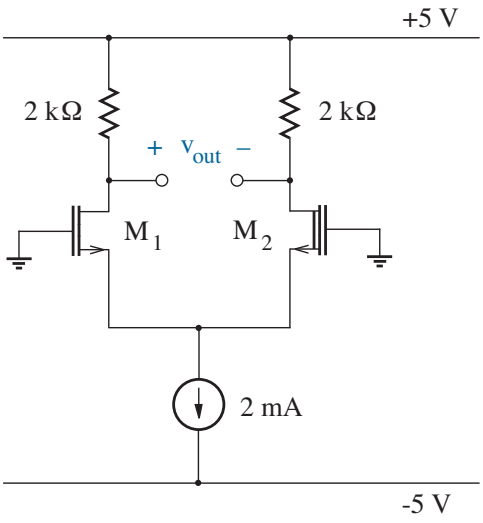


Figure P5.48

Section 5.4

SPICE Analysis: Switching Circuits

Use SPICE parameters as needed from Table 5.1.

5.49 Consider the inverter circuit of Fig. P5.25 with $R = 1$ kΩ and $W/L = 20$.

(a) Use SPICE to plot v_{out} vs. v_{in} over the range $0 \leq v_{in} \leq 3.3$ V.

(b) Let $v_{in} = 3.3$ V. Compare the LOW-level output as determined with SPICE to that obtained when the MOSFET has a particular r_{ds}.

5.50 Consider the inverter circuit of Fig. P5.25 with $R = 1$ kΩ and $W/L = 20$, and include a 10-pF capacitor connected between the output and ground. Neglect any other capacitance, and let $V_{DD} = 5$ V.

(a) Use SPICE to determine the time needed to transition the output to $V_{DD}/2$ following an abrupt HIGH-to-LOW transition at the input.

(b) Use SPICE to determine the time needed to transition the output to $V_{DD}/2$ following an abrupt LOW-to-HIGH transition at the input.

(c) Repeat part b, but model the MOSFET as a voltage-controlled switch in series with the r_{ds} value that is applicable when the input is HIGH (see Example 5.3). Compare transition times.

5.51 Repeat the inverter design of Problem 5.28, but use SPICE to find the W/L values.

5.52 Consider the CMOS analog switch of Fig. 5.19.

(a) Use SPICE to verify the r_{ds} vs. v_s behavior (-3 V $\leq v_s \leq +3$ V) with M_2 absent.

Hint: Connect a variable voltage source to v_s, connect a 1-mV source across the switch, and examine the ratio of the latter source voltage to the current that it produces.

(b) Repeat part a, but with M_1 absent.

5.53 A sample-and-hold circuit featuring a 0.1-pF capacitor connects to an input signal source with 50-Ω Thevenin resistance.

(a) Use SPICE to simulate the sampling process for $v_{in} = 1$ V with the MOSFET parameters of Problem 5.32. (You will need to specify a value for TOX that is consistent with C_{ox}.)

(b) Show what happens to the sampling error as the pulsed gate control voltage becomes less abrupt.

5.54 Use SPICE to simulate the video multiplexing circuits of Figs. 5.21 and 5.22 subject to:

- Switches are parallel-connected n- and p-channel MOSFETs, $K'W/L = 10$ mA/V^2, $V_T = \pm 0.5$ V.
- Switch control voltages are ±5 V.
- Shunt switch capacitance is 1 pF.
- MOSFET sources and drains feature 0.1-pF capacitance to ground.
- Video signal sources have 0.5-V peak-to-peak amplitude at 400 MHz and 75-Ω Thevenin R_t.
- The video output connects to a 75-Ω load.

Examine the off-state isolation and on-state signal loss in each case.

5.55 Use SPICE to simulate the charge-pump of Fig. 5.24 with the following:

- Switches are parallel-connected n- and p-channel MOSFETs, $K'W/L = 50$ mA/V^2, $V_T = \pm 0.5$ V.
- Switch control voltages are ±5 V.
- $C_1 = C_2 = 0.1$ μF, $R = 200$ Ω, $V^+ = 5$ V.

Examine the output at switching frequency $f = 0.1/RC$, $1/RC$, and $10/RC$.

5.56 Use SPICE to simulate the inverter of Fig. 5.25 subject to the following:

- Switches are either n- or p-channel MOSFETs.

- $K'W/L = 100$ mA/V^2 and $V_T = \pm 0.5$ V.
- Switch control voltages are ± 10 V, and the switching frequency is 10 kHz.
- $R = 50\ \Omega$, $V^+ = 10$ V.

(a) Examine the voltage across the resistive load.

(b) Repeat part a, but add 1mH in series with R.

(c) Repeat part b, but add appropriate diodes (IS = 10p).

5.57 The boost converter of Fig. P5.57 features $L = 120\ \mu$H, $C = 100\ \mu$F, and $R = 50\ \Omega$. The switching frequency is 50 kHz subject to a 60-% duty cycle, the capacitor has 0.1-Ω ESR, the diode has IS = 10n.

(a) Calculate the maximum ideal switch current.

(b) The switch is actually an n-channel MOSFET for which $K_n' = 50\ \mu$A/V^2 and $V_{Tn} = 0.5$ V. Determine W/L so that the anticipated on-state switch voltage is 0.1 V. Assume $v_{gs} = 6$ V when the switch is on.

(c) Simulate the converter with SPICE subject to the design of part b. How does the maximum output voltage compare with expectations?

(d) Suggest a simple way to obtain the theoretical output voltage despite the switch limitations, then demonstrate with SPICE.

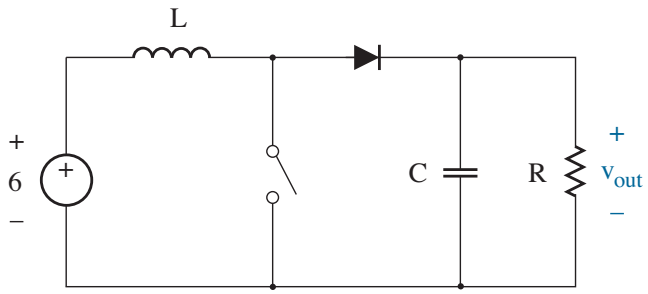


Figure P5.57

5.58 The buck converter of Fig. P5.58 features $L = 120\ \mu$H, $C = 47\ \mu$F, and $R = 20\ \Omega$. The switching frequency is 80 kHz subject to a 50-% duty cycle, the capacitor has 0.1-Ω ESR, the diode has IS = 10n.

(a) Calculate the maximum current in the switch when it is ideal.

(b) The switch is actually a p-channel MOSFET with $K_p' = 20\ \mu$A/V^2 and $V_{Tp} = -0.5$ V. Find W/L so that the anticipated on-state switch voltage is 0.1 V. Assume $v_{gs} = -12$ V when the switch is on.

(c) Simulate the converter with SPICE subject to the design of part b. How does the maximum output voltage compare with expectations?

(d) Determine the maximum diode current for the simulation of part c.

(e) Replace the diode with an n-channel MOSFET that will function as a synchronous rectifier. For this device, $K_n' = 50\ \mu$A/V^2, $V_{Tn} = 0.5$ V. Find W/L so that the anticipated on-state rectifier voltage is 0.1 V. Assume $v_{gs} = 12$ V when the rectifier is on.

(f) Repeat the converter simulation using the synchronous rectifier. Ensure that the MOSFETs are never "on" simultaneously.

(g) Suggest a way to obtain the theoretical output voltage despite the MOSFET limitations, then demonstrate with SPICE.

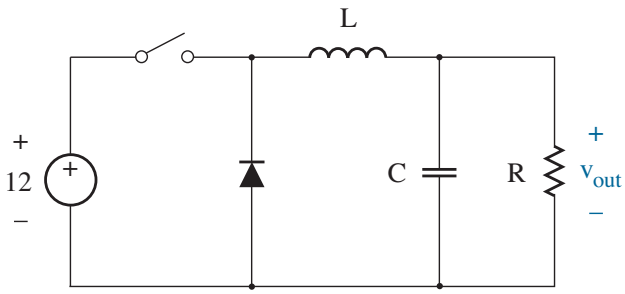


Figure P5.58

5.59 Figure P5.59 shows a so-called "string" 2-bit digital-to-analog converter circuit. The MOSFET switches are characterized by $K_n' = 50\ \mu\text{A/V}^2$, $V_{Tn} = 0.6$ V, the same W/L, and separate 0.1-pF capacitances to ground at the source and drain ends. Any switch is closed when the bit symbol next to it is HIGH at 5 V. Assume $R = 1$ kΩ. The output node connects to $C = 5$ pF.

(a) The converter has been operating in the $\{b_1\, b_0\} = \{1\,1\}$ state for an extended period. Use SPICE to determine a value of W/L that ensures less than 0.25-V error within 20 ns of an abrupt $\{1\,1\} \rightarrow \{0\,1\}$ transition.

(b) The converter has been operating in the $\{b_1\, b_0\} = \{0\,1\}$ state for an extended period. Given the W/L value determined in part b, use SPICE to specify the output error within 20 ns of an abrupt $\{0\,1\} \rightarrow \{1\,1\}$ transition.

(d) In view of the preceding results, discuss how the MOSFETs should be sized to obtain the same output error for the various transitions.

(d) Show how to implement a 3-bit string converter. What are the speed implications as the number of bits increases?

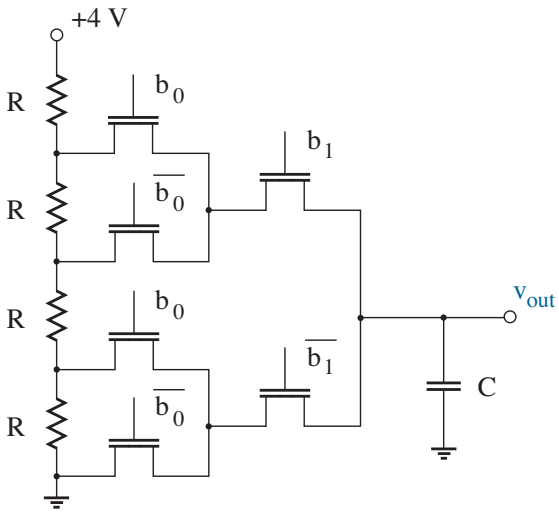


Figure P5.59

SPICE Analysis: "Valve" Circuits

Use SPICE parameters as needed from Table 5.1.

5.60 Use SPICE to verify the results of Example 5.4.

5.61 The circuit of Example 5.4 has $R = 12$ kΩ.

(a) Use SPICE to determine v_{out}.

(b) Repeat the simulation of part a including 10-Ω parasitic drain and source resistances.

5.62 Repeat Problem 5.60, but sweep the ambient temperature over the range -40 °C $\leq T \leq$ +85 °C.

5.63 Consider the circuit of Example 5.5.

(a) Use SPICE to find v_{out} when $R = 8$ kΩ.

(b) Repeat part a, but let $(W/L)_2 = 1.1\,(W/L)_1$ and $(W/L)_4 = 1.1\,(W/L)_3$. Compare the results with hand calculations.

(c) Repeat part a, but use ±4-V power supplies. Compare the results with hand calculations.

(d) Challenge! Repeat part a, but assume that the MOSFETs operate subject to the drain-current complication of Eq. 5.16 with $\lambda = 0.1\ \text{V}^{-1}$. Explain the results.

5.64 Use SPICE to verify the results of Exercise 5.5.

5.65 Use SPICE to simulate the circuit of Problem 5.42 by connecting a voltage source to the M_2 gate and sweeping it over the range -10 V $< v_{g2} <$ 10 V. (Remove the 80-kΩ and R resistors.) Observe the voltage v_{out}.

5.66 The circuit of Problem 5.44 has $R = 5$ kΩ.

(a) Use SPICE to determine v_{out} and the MOSFET drain currents.

(b) Challenge! Repeat part a, but assume that the MOSFETs operate subject to the drain-current complication of Eq. 5.16 with $\lambda = 0.1\ \text{V}^{-1}$. Explain the results.

MOSFET Capacitance

5.67 The MOSFETs in the circuit of Example 5.4 have $W = 50$ μm, $L = 1.25$ μm, $t_{ox} = 80$ nm, and CBS = CBD = 0.5 pF. The MOSFET gates have 0.1-μm source and drain overlap. Find C_{gs}, C_{gd}, C_{bs}, and C_{bd} subject to the applicable circuit conditions.

5.68 The MOSFETs in the circuit of Exercise 5.5 have $W = 40$ μm, $L = 1.0$ μm, $t_{ox} = 65$ nm, and CBS = CBD = 0.6 pF. The MOSFET gates have 0.1-μm source and drain overlap. Find C_{gs}, C_{gd}, C_{bs}, and C_{bd} subject to the applicable circuit conditions.

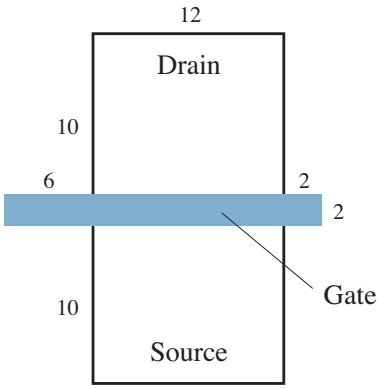
5.69 Repeat Problem 5.67, but assume the source and drain regions measure 50 μm by 50 μm with 1-μm junction depth. Assume CJ = 0.25 fF/μm^2 and CJSW = 0.32 fF/μm.

5.70 Repeat Problem 5.68, but assume both source and drain regions measure 40 μm by 40 μm with 1-μm junction depth. Assume CJ = 0.30 fF/μm^2 and CJSW = 0.38 fF/μm.

5.71 A MOSFET with the geometry shown in Fig. P5.71 has the following SPICE parameters:

Parameter	Value
CJ	1.2m
PB	0.95
MJ	0.45
CJSW	410p
PBSW	1.05
MJSW	0.36
CGDO	620p
CGSO	620p
CGBO	35p
TOX	29n
LD	0.45u

The transistor operates with $v_{gs} = 1.2$ V, $v_{bs} = 0$ V, and $v_{bd} = -3.2$ V. Find C_{gs}, C_{gd}, C_{gb}, C_{bs}, and C_{bd}.



All top-view dimensions in μm

Figure P5.71

5.72 The power MOSFET in the circuit of Fig. P5.72 has the following: $K_n' = 54$ μA/V^2, $V_{Tn} = 0.75$ V, $W = 620$ μm, $L = 1$ μm, TOX = 75 nm, XJ = 1.2 μm, CJ = 0.22 fF/μm^2, CJSW = 0.28 fF/μm. The source region measures 120 μm by 620 μm, and the drain region measures 200 μm by 620 μm. The parasitic diode has IS = 1p and CJO = 2 pF. Capacitances $C_g = 0.8$ pF and $C_d = 1.2$ pF reflect the MOSFET packaging. Use SPICE to determine the time to establish 1.6 A of load current when v_{in} steps from 0 to +10 V.

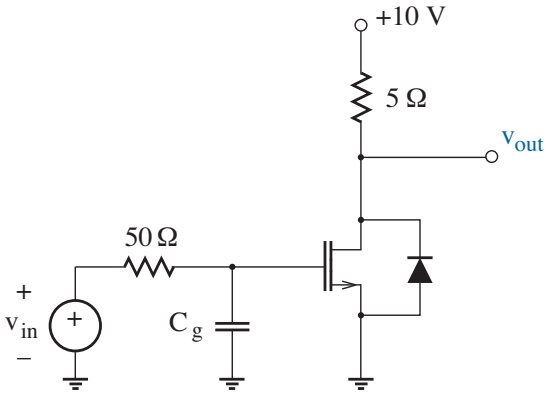


Figure P5.72

Appendix

5.73 Figure P5.73 has two MOS capacitors subject to zero flatband voltage. One of them features $V_T = -1$ V and the other features $V_T = 2$ V. Indicate the ranges of V_x that place each device in the following operating modes: accumulation, depletion without inversion, and inversion.

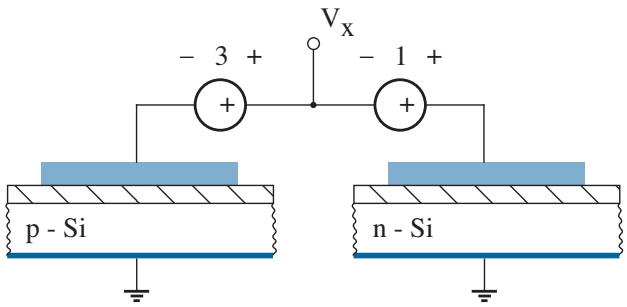


Figure P5.73

5.74 An MOS capacitor exhibits a depletion depth of 0.5 μm when the surface potential of the p-type substrate is 0.5 V. The SiO_2 thickness is 60 nm, and the flat-band voltage is zero.

(a) Determine the substrate doping concentration assuming no inversion, then find $2\phi_f$ to verify the assumption.

(b) Determine the body-effect parameter γ, then find the gate-to-body voltage that produces the specified substrate surface potential.

5.75 Design a p-substrate MOS capacitor with $C_{ox} = 10$ nF/cm^2 and $\gamma = 0.4$ V$^{1/2}$.

5.76 An MOS capacitor has the following structure: n^+ polycrystalline silicon gate, SiO_2 insulator with 50-nm thickness, p-type Si substrate with uniform doping ($N_a = 5\times10^{15}$ cm^{-3}). The interface between the insulator and substrate can be represented by a charge sheet with density $Q_i/q = 8 \times 10^{10}$ cm^{-2}. Determine the threshold voltage.

5.77 Repeat Problem 5.76 for the case of an n-type Si substrate.

5.78 Repeat Problem 5.76 for the case of a p^+ polycrystalline silicon gate.

Note: p^+ gates are generally avoided in MOS devices since boron in the gate tends to diffuse to the silicon substrate during high-temperature processing steps. The diffused boron alters the threshold voltage.

5.79 Consider Eq. 5.37, which relates $\phi_o' = 4\phi_o/\gamma^2$ to $v_{gb}' = 4v_{gb}/\gamma^2$.

(a) Find v_{gb}' at the onset of inversion for an MOS capacitor with $2\phi_f = 0.8$ V and $\gamma = 0.5$ V$^{1/2}$. Assume zero flat-band voltage.

(b) Use a computer application (or other procedure) to graph Eq. 5.37 as v_{gb}' increases from zero to the value determined in part *a*.

(c) Show that the slope of the ϕ_o' curve approaches unity for large values of v_{gb}', then compare with the actual slope at the onset of inversion.

(d) Specify how $d\phi_o'/dv_{gb}'$ relates to $d\phi_o/dv_{gb}$.

5.80 Figure P5.80 shows a cross section of an integrated circuit at an early stage of its fabrication. "Active" regions A and B are to support n-channel MOSFET switches. The intervening "field" region with SiO_2 thickness t_{ox}' provides electrical isolation between A and B despite the later addition of an overlying Al interconnect at 5 V. Find t_{ox}' so that $V_T = 10$ V in the field region (to suppress inversion). Let $N_a = 10^{15}$ cm^{-3} for the p-type silicon substrate, and assume a worst-case $Q_i/q = 8 \times 10^{10}$ cm^{-2}.

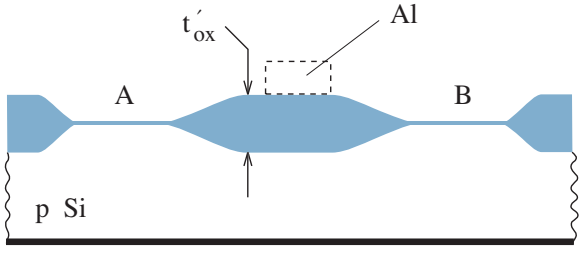


Figure P5.80

5.81 Consider an MOS capacitor with charge sheet $Q(x)$ (C/cm^2) embedded in the gate dielectric at position x ($-t_{ox} \leq x \leq 0$) as shown in Fig. P5.81. Assume $\phi_{ms} = 0$ and $t_{ox} = 50$ nm.

(a) Apply a graphical analysis (as in Fig. 5.34) to sketch the electric field vs. position for $-t_{ox} \leq x \leq w$.

(b) Show that the charge sheet induces a shift in the gate-to-body voltage needed to deplete the semiconductor to depth w. Specifically,

$$\Delta v_{gb} = \frac{-Q}{C_{ox}} \left(1 + \frac{\zeta}{t_{ox}} \right) .$$

This implies the same shift in threshold voltage.

(c) A gate dielectric that is permeated with uniform charge density Q' (C/cm^3) can be viewed as a superposition of many individual charge sheets. Find the resulting shift in threshold voltage.

(d) A sloppy technician (too lazy to wear gloves) leaves a fingerprint over an array of MOS capacitors after inspecting them during fabrication. Fingerprint residue contains about 0.1 μg/cm^2 sodium (and other chemicals) after moisture has evaporated. Suppose 0.01 % of this salt eventually assumes uniform distribution in the gate dielectric as Na^+. Determine the shift in the threshold voltage.

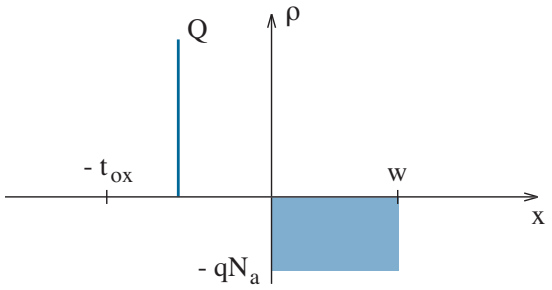


Figure P5.81

5.82 The insulator thickness for an MOS capacitor scales by a factor α. *To first order*, how should the substrate doping scale so that the threshold voltage is unchanged? Assume constant flat-band voltage.

5.83 Design an MOS capacitor so that $V_T = 0.6$ V given a p-type Si substrate with $N_a = 2 \times 10^{16}$ cm^{-3}. The expected Q_i/q is 5×10^{10} cm^{-2}.

5.84 This problem concerns the relationship between the surface potential ϕ_o and gate bias v_{gb} in an MOS capacitor subject to inversion with $\phi_o > 2\phi_f$. Assume that the p-type silicon substrate with doping concentration N_a is depleted to depth w. The charge density in the depletion region is

$$\rho = -qN_a - q\frac{{n_i}^2}{N_a} e^{\,q\phi/kT} ,$$

where ϕ is the electric potential. The first term in this expression represents fixed ionized acceptors. The second term represents free electrons under equilibrium conditions (see Eq. 5.61).

(a) Starting with Poisson's equation

$$\frac{d^2\phi}{dx^2} = -\frac{\rho}{\epsilon_{Si}} ,$$

show that

$$\left.\frac{d\phi}{dx}\right|_0 = -\sqrt{\frac{2qN_a}{\epsilon_{Si}}} \left[\phi_o + \frac{kT}{q} \left(\frac{n_i}{N_a} \right)^2 \left(e^{\,q\phi_o/kT} - 1 \right) \right]^{1/2} .$$

Hint: After substituting for ρ, multiply Poisson's equation by $d\phi/dx$, express both sides as exact differentials, and integrate over $[0, w]$.

(b) Establish continuity of electric displacement at $x = 0$ to show that

$$v_{gb} = \phi_o + \gamma \left[\phi_o + \frac{kT}{q} \left(\frac{n_i}{N_a} \right)^2 \left(e^{\,q\phi_o/kT} - 1 \right) \right]^{1/2} .$$

(c) Plot the preceding relation to demonstrate the weak variation of ϕ_o beyond the condition for inversion. Let $N_a = 10^{15}$ cm^{-3}.

5.85 The total ϕ_o-induced substrate charge in an MOS capacitor is the negative of the electric displacement at $x = 0$.

(a) Use the results of Problem 5.84 to show that the areal (cm^{-2}) electron density in an inversion layer is given by $N_I = \sqrt{2} N_a L_d F$, where

$$F = \left\{ \left[\frac{q\phi_o}{kT} + \left(\frac{n_i}{N_a} \right)^2 \left(e^{\, q\phi_o/kT} - 1 \right) \right]^{1/2} - \left[\frac{q\phi_o}{kT} \right]^{1/2} \right\}.$$

Assume that the inversion charge is localized to a sheet at $x = 0+$.

(b) Plot the natural logarithm of N_I over the range $20 \leq q\phi_o/kT \leq 35$ subject to $N_a = 10^{15}$ cm^{-3}. Find points where $\phi_o = 2\phi_f$ and $N_I = N_a w$. Discuss the character of the N_I variation under weak-inversion ($N_I < N_a w$) and strong-inversion ($N_I > N_a w$) conditions.

5.86 This problem examines the physical significance of Debye length L_d. Consider a semiconductor slab with uniform acceptor doping N_a that is subject to a charge sheet Q (C/cm^2) embedded at $x = 0$. The Q magnitude is sufficient to produce modest hole redistribution as a function of x.

(a) Apply the concentration of ionized acceptors and the concentration of holes at equilibrium (Eq. 5.59) to Poisson's equation (Problem 5.84) to show that the electric potential satisfies

$$\phi = A\, e^{\,-x/L_d}$$

in the region where $x > 0$. Hint: Make a Taylor-series expansion of the expression for the charge density ρ subject to $\phi \ll kT/q$.

(b) Similarly, show that for $x < 0$,

$$\phi = B\, e^{\,x/L_d}\,.$$

(c) Consider conditions at $x = 0$ to find A and B.

The derived potential variations for $x > 0$ and $x < 0$ suggest that free-carrier perturbations tend to smooth out over a few Debye lengths.

5.87 Use the equilibrium concentrations of ionized acceptors and holes (Eq. 5.59) in Poisson's equation (Problem 5.84) to show that $C_{fb} = \epsilon_{Si}/L_d$ for an MOS capacitor in the limit as $\phi_o \to 0$.

Hint: Find $d\phi/dx$ at $x = 0$ following the procedure in Problem 5.84, then make a Taylor-series expansion in powers of ϕ_o.

5.88 MOS capacitors are useful for the characterization of integrated circuits. Figure P5.88 shows capacitance-voltage data for an MOS structure with area $A = 10^{-3}$ cm^2. The measurement conditions are consistent with a circuit model that features an insulator and a depletion capacitance in series.

(a) Specify whether the silicon substrate is n-type or p-type.

(b) Determine the SiO_2 insulator thickness.

(c) Determine the substrate doping concentration.

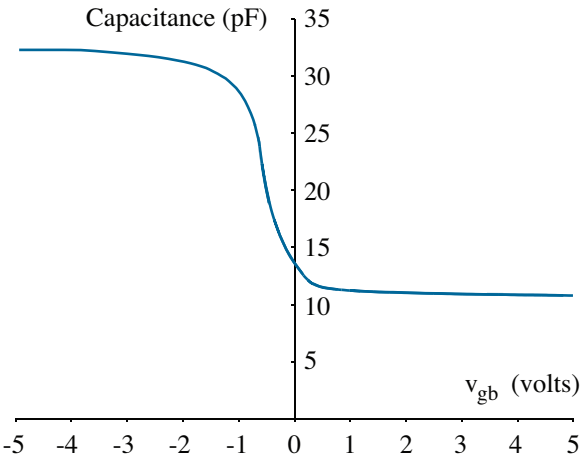


Figure P5.88

5.89 This problem concerns an analytical method that extracts flat-band voltage for an MOS capacitor from capacitance-voltage data. Consider an MOS capacitor in strong accumulation (negative v_{gb} for a p-type substrate with doping N_a).

(a) In consideration of the potential dependence of free holes at equilibrium (Eq. 5.59) and the procedure in Problem 5.84 for the solution of Poisson's equation, show that

$$\left.\frac{d\phi}{dx}\right|_0 = \sqrt{2}\left(\frac{kT/q}{L_d}\right) e^{\,-q\phi_o/2kT}\,.$$

(b) Apply the preceding result to determine Q_s, the substrate accumulation charge density. Then show that

$$Q_s = C_s\left(\frac{2kT}{q}\right)$$

where $C_s = -dQ_s/d\phi_o$ is the areal substrate capacitance (F/cm^2).

(c) Let C be the measured areal MOS capacitance. Show that subject to $V_{fb} = 0$, $C/C_{ox} = 0.95$ corresponds to $v_{gb} \approx -1$ V. In turn,

$$v_{gb}|_{\,C/C_{ox}=0.95} + 1\text{ V} = V_{fb}\,.$$

Assume small ϕ_o. (Given an n-type substrate, subtract 1 V from the designated v_{gb}.)

(d) Determine the flat-band voltage that applies to the capacitance-voltage data in Fig. P5.88, then find Q_i/q. Assume an n$^+$ polysilicon gate.

5.90 Figure P5.90 shows a set of closely spaced MOS capacitors that are intended to function as charge-coupled devices (CCDs). Each gate is n$^+$ polysilicon, the SiO_2 insulator thickness is 60 nm, and the doping concentration in the p substrate is $N_a = 10^{15}$ cm^{-3}. For simplicity, assume $V_{fb} = 0$ for each MOS device. The individual v_{gb} values are all initially zero.

(a) Let v_{gb3} pulse from 0 to 5 V. Find the depletion depth w that results at $t = 0+$ if the inversion layer is absent—it takes time to form.

(b) Determine inversion-layer surface density N_I, once present.

(c) One can show (although not here) that non-equilibrium $np \ll {n_i}^2$ results in the generation of hole-electron pairs with a rate $G = n_i/2\tau$, where τ is a characteristic time. Make a rough estimate of the time to form the inversion layer of part b if $\tau = 50$ μs.

Note: The establishment time is significantly reduced if there is an electron-rich (n$^+$) region in close proximity.

(d) Let the structure be illuminated at intensity $I = 10$ mW/cm^2 (1/10 sun). The energy needed to generate a hole-electron pair is 1.12 eV. Estimate the time to establish one half of the inversion layer of part b. (Thus, the inversion charge relates to illumination if it is produced in a short time relative to the result of part c.)

(e) Suggest a way to change the various v_{gb} values so that the charge packet moves rightward with entire localization beneath gate 4 after it is established beneath gate 3. The charge-transfer process is a means for moving one pixel of video information to the end of a CCD chain where it can be detected as part of a larger image.

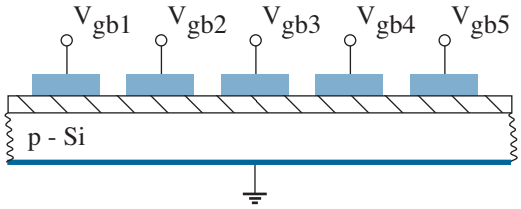


Figure P5.90

5.91 The current-voltage characteristics derived for a long n-channel MOSFET assumed a constant level of depletion-layer charge subject to a surface potential at $\phi = 2\phi_f$ with no other influence from drain to source. This problem relaxes that assumption to obtain a more accurate result.

(a) The source-referenced potential in the channel effectively adds to the surface potential that determines the depletion-layer width. Show that

$$qN_I = C_{ox}[\, v_{gs} - \psi - 2\phi_f - \gamma\sqrt{\psi + 2\phi_f} - V_{fb}\,]\,.$$

Note that this expression reduces to Eq. 5.67 when the fourth term is constant with $\psi = 0$.

(b) Show that the non-saturation drain current is given by

$$i_d = \frac{1}{2}K'\frac{W}{L}\left\{2(v_{gs} - 2\phi_f - V_{fb})v_{ds} - {v_{ds}}^2 - \frac{4}{3}\gamma\left[(v_{ds} + 2\phi_f)^{3/2} - (2\phi_f)^{3/2}\right]\right\}.$$

(c) Show that the preceding expression reduces to the form in Eq. 5.8b when v_{ds} is small. Hint: Expand the terms that are proportional to γ in a Taylor series around $v_{ds} = 0$.

(d) Let $N_a = 5 \times 10^{15}$ cm^{-3}, $t_{ox} = 50$ nm, $\mu_e = 800$ cm^2/V-s, $V_{fb} = 0$, and $W/L = 10$. Plot the elementary (Eq. 5.8b) and revised drain currents vs. v_{ds} for $v_{gs} = 2$ V. Compare the results.

Note: You are hardly alone if you are confused by two co-exisiting potentials: The gate-induced surface potential ϕ yields inversion at the source ($\phi \approx 2\phi_f$). The channel potential ψ provides for inversion-layer variations with ψ set to zero at the source. To clarify, energy-band models are needed to describe electron conditions in regions that extend from the source. This higher-level treatment is available in advanced texts on device physics.

5.92 One way to determine parameter γ is to plot the measured threshold voltage against the quantity $\sqrt{2\phi_f - v_{bs}} - \sqrt{2\phi_f}$ for a particular MOSFET.

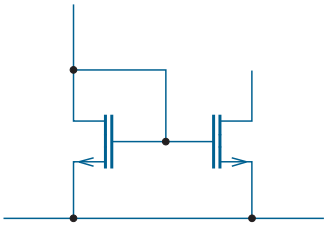
Consider the following data:

V_T (V)	v_{bs} (V)
0.720	0.0
1.421	-2.0
1.871	-4.0
2.230	-6.0
2.538	-8.0
2.813	-10.0

(a) Guess for $2\phi_f$, then plot the data as described. Repeat until a straight line has been obtained.

(b) Determine the applicable γ.

(Chapter 9 explores this behavior in greater detail.)



Chapter 6

The Bipolar Junction Transistor

Since its introduction in 1948, the **Bipolar Junction Transistor** or **BJT** has driven the field of solid-state electronics more than any other transistor. "Valve" or switch action is effected by a current, and the control relationship is nearly linear over a broad range.

Once again, our goal is to derive static (time-independent) terminal relationships and identify key device parameters that are easily measured. In turn, we use the terminal relationships to develop large-signal circuit models that support the analysis of some elmentary circuit examples.

Objectives

After studying this chapter and working through the problems, you should be able to perform each of the following:

- Describe the four modes of BJT operation and the associated terminal relations (Section 6.1).
- Contrast the electrical behavior of npn and pnp devices (Section 6.1).
- Measure β_F using a simple circuit or a curve-tracer display of BJT characteristic curves (Section 6.1).
- Describe the Ebers-Moll BJT model and the large-signal forward-active models that are derived from it (Section 6.2).
- Analyze simple dc BJT circuits using an appropriate large-signal model or SPICE (Sections 6.3 and 6.4).

6.1 Operating Principles

A BJT is constructed from two back-to-back pn junctions with **npn** or **pnp** orientation as shown in Figs. 6.1a and 6.1b, respectively. In either case, the ends of this device are designated as **emitter** and **collector**, and the center is designated as the **base**. These labels are somewhat unfortunate—one could argue in favor of the analogous and more general source, drain, and gate nomenclature used in conjunction with the MOSFET terminals. However, we yield to stubborn convention, hoping that you will appreciate the similarity as circuit patterns begin to emerge.

The npn and pnp transistor symbols differ by the direction of a small arrow near the emitter terminal. For reasons that will soon become clear, we define positive collector and base currents as flowing inward, but we define positive emitter current as flowing outward (unlike some other texts).

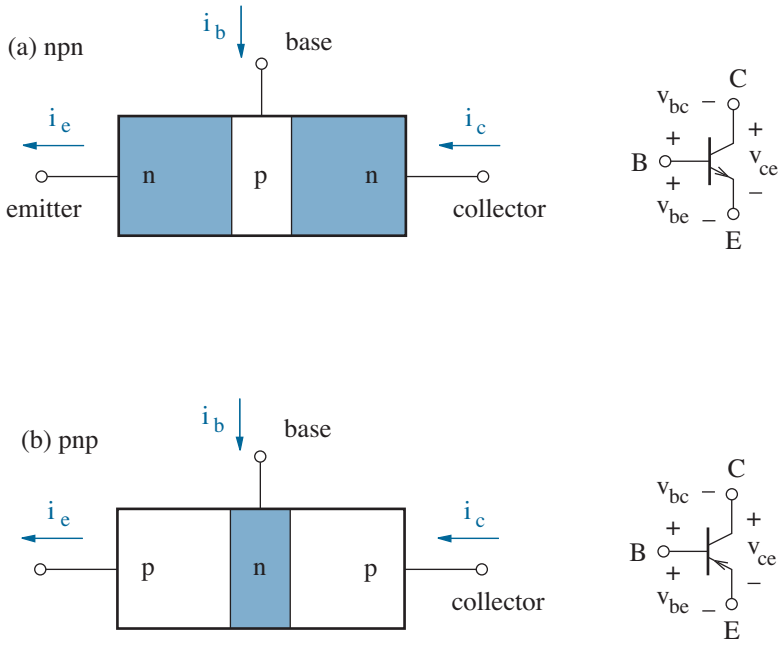


Figure 6.1: BJT npn and pnp configurations, and corresponding symbols.

> In what follows, we focus on the npn BJT device configuration. The pnp transistor exhibits similar behavior, except *all* terminal currents and voltages are altered by a change in sign.

Common-Base Characteristics

Physical descriptions of BJT operation are particularly simple when the base terminal is shared between a left-side input and a right-side output as shown in the circuit of Fig. 6.2a. In this **common-base** configuration,[1] v_{be} and v_{bc} control separate pn junctions, so we expect a different mode of BJT operation for each of *four* combinations of junction bias polarities. The modes correspond to four quadrants in the v_{be}-v_{bc} diagram of Fig. 6.2b, and we consider them in sequence.

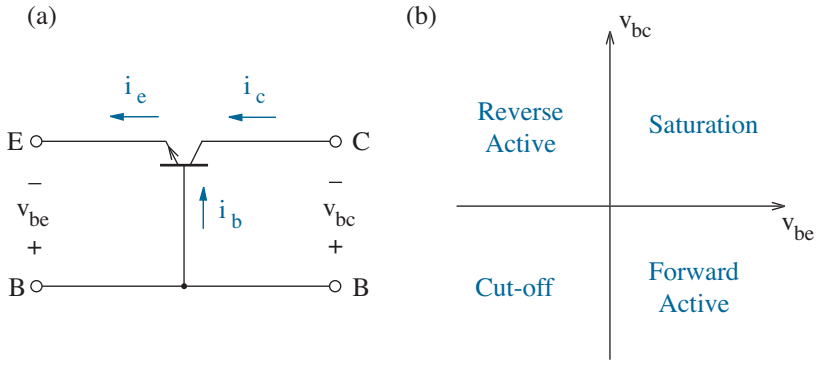


Figure 6.2: Common-base configuration for an npn BJT, and the junction biasing conditions that govern four modes of operation.

• Forward Active Mode

In this the most important mode of BJT operation, the base-emitter junction is *forward biased* and the base-collector junction is *reverse biased*. A simple $i-v$ analysis seems rather straightforward. For the forward-biased base-emitter junction,

$$i_e = I_{ES}\left[e^{qv_{be}/kT} - 1\right], \tag{6.1}$$

where I_{ES} is the base-emitter diode saturation current. If $v_{be} \gg kT/q$, i_e is very large. And for the reverse-biased base-collector junction,

$$i_c = -I_{CS}\left[e^{qv_{bc}/kT} - 1\right], \tag{6.2}$$

[1]The common-base configuration was natural for primitive BJTs, which featured a pedestal-mounted semiconductor—a true base—and closely spaced point-contact base-emitter and base-collector junctions at the surface.

where I_{CS} is the base-collector diode saturation current. If $v_{bc} \ll -kT/q$, i_c is very small. Finally, Kirchhoff's Current Law requires a base current, which is the difference between the two junction currents:

$$i_b = i_e - i_c \,. \tag{6.3}$$

How unexciting! The governing equations hardly indicate useful behavior. Why would anyone take interest in such a device?

Fortunately, the forward active mode has more to offer if we insist on two conditions that pertain to BJT structure.

The first condition acknowledges the ideal n-type emitter as a one-way dispatcher for controlled electron transit to the *collector*. Thus, the total current flowing across the base-emitter junction should primarily consist of electron current, since the absorption of holes into the emitter from the base is counterproductive. For a crude analysis, we quantify the degree of electron **injection** by considering an expression for the electron current component in a pn junction:

$$i_{\text{electron}} \sim \frac{{n_i}^2}{N_a} \left[e^{\,qv/kT} - 1 \right], \tag{6.4}$$

where ${n_i}^2/N_a$ is the equilibrium electron concentration on the p side of the junction with acceptor impurity concentration N_a. We also consider an expression for the hole current component:

$$i_{\text{hole}} \sim \frac{{n_i}^2}{N_d} \left[e^{\,qv/kT} - 1 \right], \tag{6.5}$$

where ${n_i}^2/N_d$ is the equilibrium hole concentration on the n side of the junction with donor impurity concentration N_d. These Chapter-2 relations are used to determine an **emitter defect** δ_e, which we define as the ratio of the undesired hole current component to the total junction current. Specifically,

$$\delta_e = \frac{i_{\text{hole}}}{i_{\text{hole}} + i_{\text{electron}}} \sim \frac{1}{1 + \dfrac{N_d}{N_a}} \,. \tag{6.6}$$

And with $i_{\text{hole}} + i_{\text{electron}} = i_e$,

$$i_{\text{electron}} = (1 - \delta_e)\, i_e \,. \tag{6.7}$$

This is the injected electron current. To make it large, we require δ_e small, so in general (see Problem 6.3 and Chapter 8 for alternative approaches) we make the BJT emitter more heavily doped than the base ($N_d \gg N_a$). The emitter current i_e is given by Eq. 6.1 as before.

The second condition acknowledges the perilous journey encountered by each injected electron that diffuses across enemy base territory. As minority carriers therein, the electrons are subjected to merciless slaughter through recombination with majority-carrier holes. We quantify this condition by means of a **base defect** δ_b, which is defined as the ratio of electron current lost through recombination processes to the total injected electron current. Thus, the probability of "safe" base transport is proportional to $(1 - \delta_b)$, and the collected electron current is given by

$$i_{\text{collected}} = (1 - \delta_b)\, i_{\text{electron}} = (1 - \delta_b)(1 - \delta_e)\, i_e \,. \tag{6.8}$$

For an effective collection process, we require δ_b small, so we make the BJT base width very narrow.

If we ignore the reverse-bias base-collector junction currents that are independent of the emitter-to-collector electron transport process, then

$$i_{\text{collected}} = i_c = \alpha_F\, i_e \,, \tag{6.9}$$

where $\alpha_F = (1 - \delta_e)(1 - \delta_b)$ is the **forward common-base large-signal current gain**. Note that α_F approaches unity if δ_e and δ_b are both small. Typically, α_F is 0.98 or more for a "good" BJT. Base current now consists of an extremely small hole component that is back-injected into the emitter (due to non-zero δ_e) and a small hole component that feeds recombination (due to non-zero δ_b). These hole currents and the various electron current components are indicated in the flow diagram of Fig. 6.3.

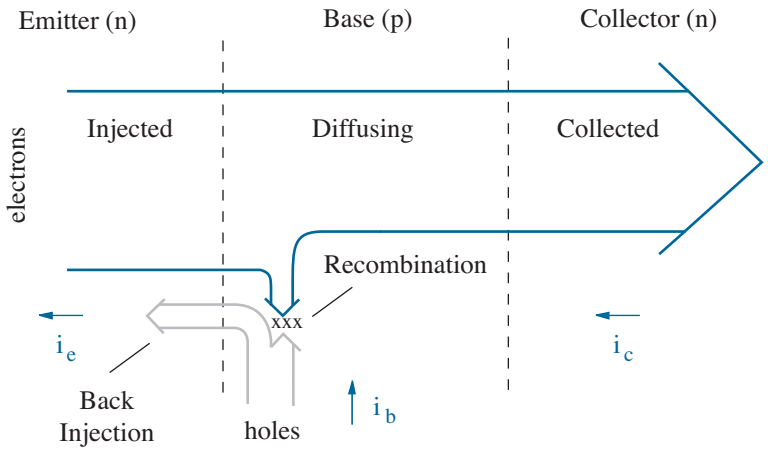


Figure 6.3: BJT (npn) current components in the forward active mode. Injected electrons diffuse across the base. Hole current is relatively small.

Now we observe the following BJT behavior: The emitter current is controlled by v_{be}, and small changes in v_{be} produce large changes in i_e as a consequence of the the exponential factor in Eq. 6.1. However, the same control indirectly applies to collector current, which is proportional to i_e. Meanwhile, forward v_{be} is somewhat small, typically of the order of 0.7 V, but reverse v_{bc} can be quite large as long as it is less than the junction breakdown voltage. This implies power gain in the common-base circuit.

Figure 6.4 shows a set of i_c vs. $-v_{bc}$ characteristic curves that describe the forward active mode. The curves are flat over a broad range of *positive* $-v_{bc}$ values, and they are unequally spaced with respect to v_{be}.

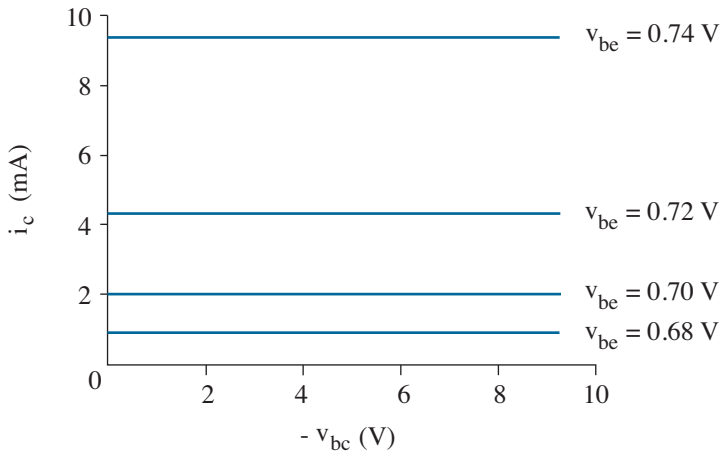


Figure 6.4: Typical BJT (npn) common-base characteristic curves for the forward active mode of operation. The index variable is v_{be}.

• Reverse Active Mode

In this mode of BJT operation, the base-emitter junction is *reverse biased* and the base-collector pn junction is *forward biased*. Thus, we have the same behavior as in the forward active mode, but the roles of the emitter and the collector are interchanged. A parallel analysis reveals

$$i_e = \alpha_R \, i_c \, , \tag{6.10}$$

where i_c is given by Eq. 6.2, and α_R is the **reverse common-base large-signal current gain**. For reasons that we have yet to discuss, the doping concentration in the collector is generally made small compared to the doping concentration in the base. This and often unfavorable geometric effects tend to reduce the collector-to-base electron injection efficiency so that α_R is significantly less than α_F. Typically, α_R is 0.85 or less.

• Saturation Mode

In this mode of BJT operation, *both* the base-emitter and base-collector junctions are *forward* biased. To a zero-order approximation, the collector, base, and emitter terminals appear to be shorted together.

• Cutoff Mode

In this mode of BJT operation, *both* the base-emitter and base-collector junctions are *reverse* biased. To a zero-order approximation, the collector, base, and emitter terminals appear to be open connections.

The saturation and cutoff modes warrant particular interest when the BJT is used as a controlled switch. We consider both modes with greater care in Section 6.2.

Exercise 6.1 For each BJT configuration in Fig. 6.5, specify whether the transistor is in the forward active, reverse active, saturation, or cutoff mode of operation.

Be careful. It may be helpful to reconstruct Fig. 6.2 for a pnp device.

Ans: (a) cutoff (b) forward active (c) saturation (d) reverse active

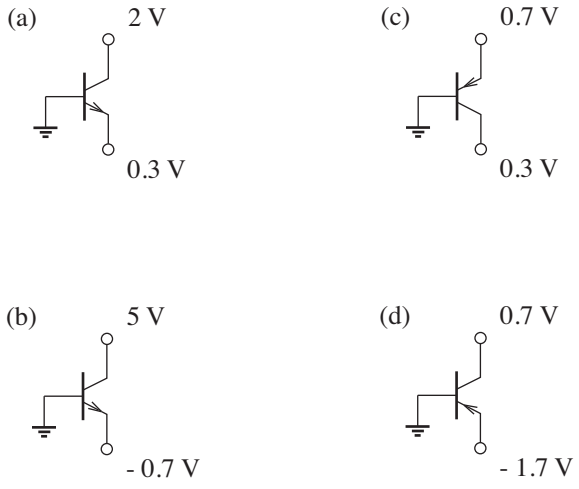


Figure 6.5: BJT configurations for Exercise 6.1.

Common-Emitter Characteristics

The common-base BJT configuration has historical significance; however, we are usually more interested in the common-emitter BJT configuration and the associated i_c vs v_{ce} characteristics. This configuration applies to a "valve" current-control circuit—see Fig. B2 in the Perspective discussion—and it facilitates direct comparison with other types of transistors.

Despite the configuration change, we still want to maximize the degree of collector-current control by operating the BJT in the forward active mode with $v_{be} > 0$ and $v_{bc} < 0$. Both conditions are consistent with a broad range of positive v_{ce} values, since $v_{ce} = -v_{bc} + v_{be}$. To determine the forward-active portions of the characteristic curves, we can use Eqs. 6.1 and 6.9 to find i_c in terms of the controlling variable v_{be}. We expect the curve segments to be flat and unequally spaced as in Fig. 6.4.

Before we examine some actual common-emitter characteristic curves, it will be a relief to note the simplicity of using the base current i_b as an index variable (in place of v_{be}). From Kirchhoff's Current Law and Eq. 6.9, we have

$$i_c + i_b = i_e = \frac{i_c}{\alpha_F} \, . \tag{6.11}$$

Then in terms of i_b, we obtain

$$i_c = \beta_F \, i_b \tag{6.12}$$

and

$$i_e = (\beta_F + 1) \, i_b \, , \tag{6.13}$$

where

$$\beta_F = \frac{\alpha_F}{1 - \alpha_F} \tag{6.14}$$

is the **forward common-emitter large-signal current gain**. The β_F parameter can be rather large, typically $50 - 200$, if α_F is close to unity. Equations 6.12 and 6.13 are fundamental BJT terminal relationships.

Figure 6.6 shows index-modified common-emitter characteristic curves. As expected, the curves are predominantly flat, but they are *equally spaced* for constant β_F—a particularly useful feature for transistor circuits that must process large signal amplitudes. Nevertheless, when v_{ce} is small such that v_{bc} is no longer negative, the curves drop down to the origin as an indication of BJT saturation. The effective saturation region is determined by the relation $0 < v_{ce} < v_{ce,sat}$, where

$$v_{ce,sat} \sim 0.2 \, \text{V}. \tag{6.15}$$

Finally, we observe a small but significant cutoff region along the v_{ce} axis where $i_c = 0$ for $i_b = 0$.

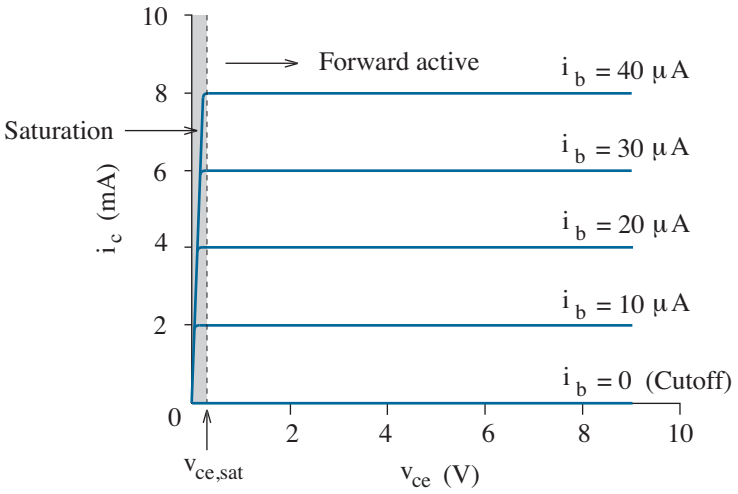


Figure 6.6: Typical set of BJT (npn) common-emitter characteristic curves. The index variable is i_b, and $\beta_F = 200$.

Exercise 6.2 A BJT with $\beta_F = 150$ operates in the forward active mode. Determine the unknown terminal currents under the following conditions: (a) $i_b = 30$ μA; (b) $i_e = 2.7$ mA.

Ans: (a) $i_c = 4.5$ mA, $i_e = 4.53$ mA (b) $i_c = 2.68$ mA, $i_b = 17.9$ μA

A figure similar to Fig. 6.6 contains a mixture of the *reverse active*, saturation, and cutoff BJT operating modes in which $v_{ce} \leq 0$ and $i_c \leq 0$. For the reverse active mode, we still have $i_b > 0$. However,

$$-i_c = (\beta_R + 1)\, i_b \tag{6.16}$$

and

$$-i_e = \beta_R\, i_b\,, \tag{6.17}$$

where

$$\beta_R = \frac{\alpha_R}{1 - \alpha_R} \tag{6.18}$$

is the **reverse common-emitter large-signal current gain**. The β_R parameter is much smaller than β_F, typically 5 or less, since $\alpha_R \ll \alpha_F$. This underscores the influence of BJT emitter/collector asymmetry.

Example 6.1

Both of the circuits in Fig. 6.7 feature a BJT with $\beta_F = 100$ and $\beta_R = 4$. Find v_{ce} in each case.

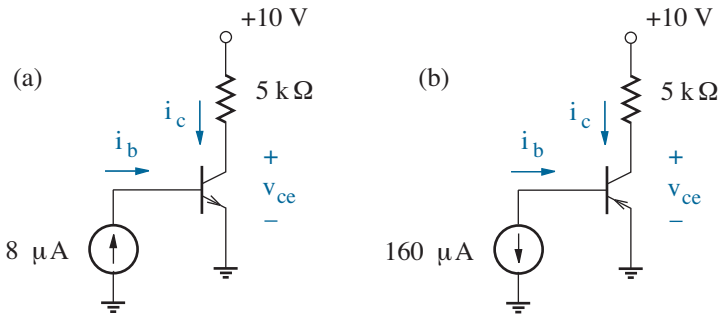


Figure 6.7: Circuits for Example 6.1.

Solution

With $i_c \geq 0$ —positive current tends to flow from high to low potentials— and $v_{ce} \geq 0$ —large i_c never forces v_{ce} sign reversal—the npn BJT of Fig. 6.7a operates in either the forward active, saturation, or cutoff mode (as in Fig. 6.6). We immediately rule out the cutoff mode, since $i_b > 0$. Then assuming the forward active mode, we easily apply Eq. 6.12 to find $i_c = \beta_F i_b = 0.8$ mA. In turn, $v_{ce} = 10 - 5i_c = 6$ V. Finally, we note that v_{ce} is sufficiently large ($\gg v_{ce,sat} \sim 0.2$ V) to justify our assumption.

The circuit of Fig. 6.7b also requires $i_c \geq 0$ and $v_{ce} \geq 0$. Nevertheless, as a consequence of the change-of-sign rule, the pnp BJT demands $i_c \leq 0$ and $v_{ce} \leq 0$ to be subject to similar forward active, saturation, or cutoff operating options that were available to the npn device. In the alternative, the pnp BJT operates in the *reverse active*, saturation, or cutoff mode. Again, we rule out the cutoff mode, since $i_b < 0$. (Note the sign change.) Then assuming the reverse active mode, we easily apply Eq. 6.16 to find $i_c = -(\beta_R + 1)i_b = 0.8$ mA. In turn, $v_{ce} = 10 - 5i_c = 6$ V. Finally, we note that v_{ce} is sufficiently large ($\gg v_{ce,sat} \sim 0.2$ V) to justify our assumption. This is not a very good circuit—the base current that yields the 0.8-mA current through the 5-kΩ resistor would have been much smaller if the pnp collector and emitter terminals had been interchanged. The forward active mode is generally preferable to the reverse active mode.

Exercise 6.3 The BJTs in the circuits of Fig. 6.8 feature $\beta_F = 100$ and $\beta_R = 4$. Design each circuit so that v has the value indicated.

Ans: (a) $R = 2.4$ kΩ (b) $R = 1.8$ kΩ

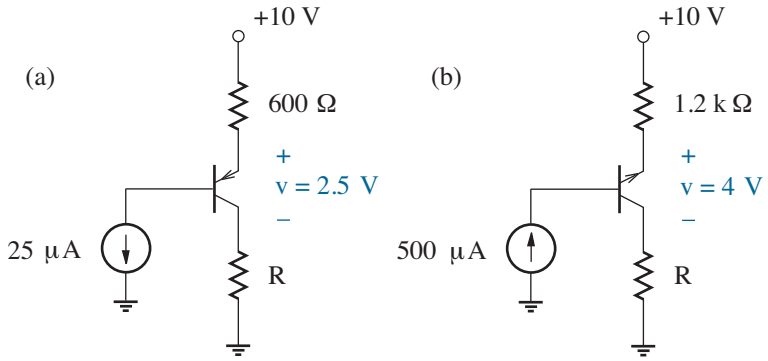


Figure 6.8: Circuits for Exercise 6.3.

Exercise 6.4 The BJTs in the circuits of Fig. 6.9 feature $\beta_F = 100$, $\beta_R = 4$, and $|v_{ce,sat}| = 0.2$ V. For each circuit, determine the range of I that corresponds to BJT cutoff and saturation.

Ans: (a) cutoff: $I = 0$ saturation: $I > 26$ μA

(b) cutoff: $I = 0$ saturation: $I < -39$ μA

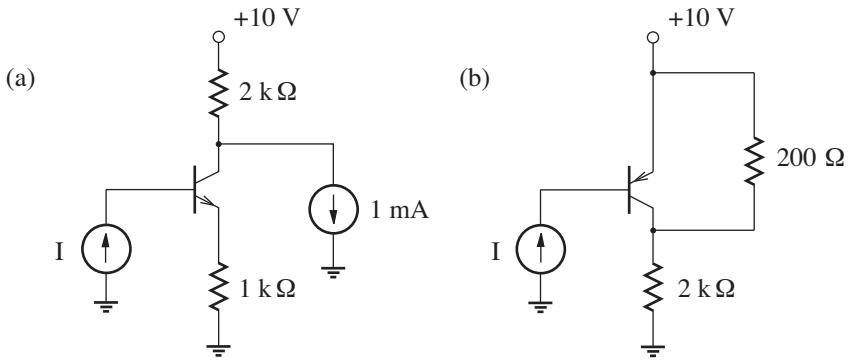


Figure 6.9: Circuits for Exercise 6.4.

Example 6.2

Figure 6.10 shows **Darlington-pair** circuit configurations that feature two npn or two pnp BJTs. The npn configuration is operated with $v_{be}{}' > 0$ and $v_{bc}{}' < 0$ such that Q_1 and Q_2 are both in the forward active mode. Determine $i_c{}'$ and $i_e{}'$ in terms of $i_b{}'$.

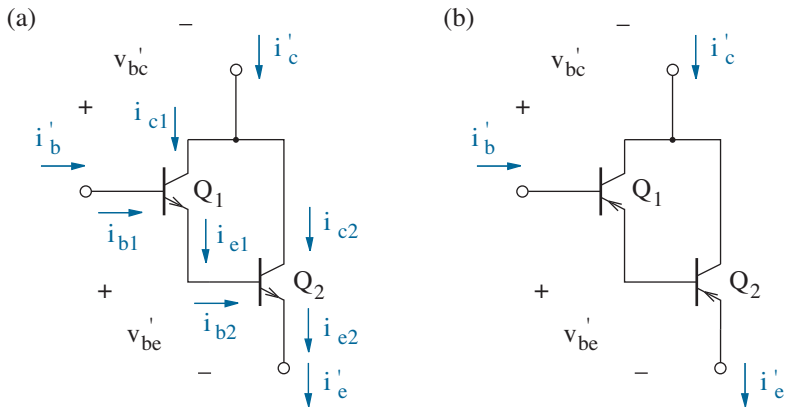


Figure 6.10: BJT Darlington pair configurations: (a) npn; (b) pnp.

Solution

We use Eq. 6.13 to obtain $i_{e1} = (\beta_{F1} + 1)i_{b1} = (\beta_{F1} + 1)i_b{}'$ as the emitter current for Q_1. This is also i_{b2}, the base current for Q_2. We apply Eq. 6.13 again to obtain $i_{e2} = (\beta_{F2} + 1)i_{b2}$. In turn,

$$i_e{}' = (\beta_{F2} + 1)(\beta_{F1} + 1)\, i_b{}' \,.$$

To determine $i_c{}'$, we use Eq. 6.12 to obtain $i_{c1} = \beta_{F1} i_{b1} = \beta_{F1} i_b{}'$ as the collector current for Q_1. Similarly, we use Eq. 6.12 and the earlier result for i_{b2} to obtain $i_{c2} = \beta_{F2} i_{b2} = \beta_{F2}(\beta_{F1} + 1) i_b{}'$ as the collector current for Q_2. Then with $i_c{}' = i_{c1} + i_{c2}$,

$$i_c{}' = (\beta_{F1}\beta_{F2} + \beta_{F1} + \beta_{F2})\, i_b{}' \,.$$

This last expression simplifies to $i_c{}' \approx \beta_{F1}\beta_{F2} i_b{}'$ when β_{F1} and β_{F2} are both large. Thus, the npn Darlington pair behaves like a "super-transistor" with an effective $\beta_F \approx \beta_{F1}\beta_{F2}$, which is easily 10,000 or more.

The forward active pnp results are the same. Darlington pairs are useful in circuits that can only tolerate extremely small base control currents.

Elementary Parameter Measurements

A β_F measurement for an npn transistor uses the test circuit of Fig. 6.11. One increases V_{bb} until the measured collector current is the order of 1 mA (or some particular value of interest). Then with measured i_b, $\beta_F = i_c/i_b$. The 5-V collector voltage is sufficient to avoid saturation, and the 100-kΩ base resistance is typically consistent with modest requirements for V_{bb}. The β_R measurement is similar. However, collector and emitter terminals are reversed, and the practical base resistance is about 100 times smaller. We leave it as an exercise to work out test conditions for a pnp device.

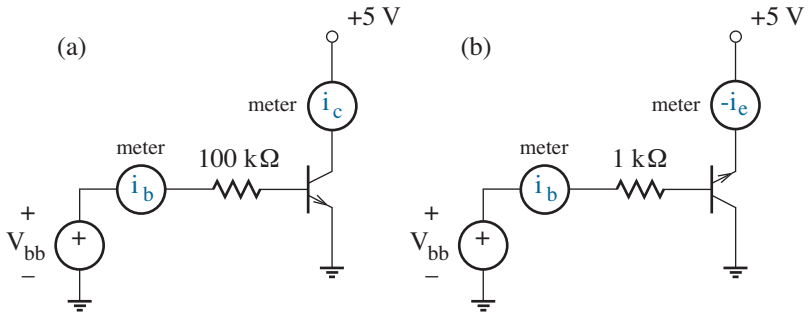


Figure 6.11: Test circuits for current-gain measurement: (a) β_F; (b) β_R.

A so-called curve tracer is a special electronic instrument that is used to display a family of transistor characteristic curves. Vertical and horizontal deflections on the instrument's CRT screen are proportional to i_c and v_{ce}, respectively. Thus, one obtains a single characteristic curve by rapidly and repeatedly sweeping v_{ce} between zero and some maximum. Simultaneous curves are obtained by stepping through a periodic sequence of i_b values. Usually $i_b = Ni_{step}$, where N is an integer and i_{step} is the step current.

Figure 6.12 shows curve-tracer displays for a randomly selected 2N3904 npn BJT with normal and inverted (collector $\leftrightarrow$ emitter) orientations that reflect forward- and reverse-mode characteristics, respectively. In each case, the vertical current scale (1 mA per division) has been adjusted to cover an intended i_c operating range (10 mA), and i_{step} has been adjusted to provide a moderately large family of curves within that range.

By definition, β_F is the i_c/i_b ratio near $v_{ce} = 0$. Thus, for the forward characteristics, $\beta_F \approx 230$, and for the reverse characteristics, $\beta_R \approx 3.5$. Caution: The lower right-hand corner of the top display suggests a "beta" factor in relation to the spacing between neighboring characteristic curves. *This is not* β_F. Instead, this is $\beta_o = \Delta i_c/\Delta i_b$, an ac BJT parameter that we have yet to consider (in Chapter 7).

Note that the BJT characteristic curves are not flat. They are tilted over the "active" range of operation, particularly for the inverted orientation. The physical origins and circuit ramifications of this second-order behavior is a topic for Chapter 9.

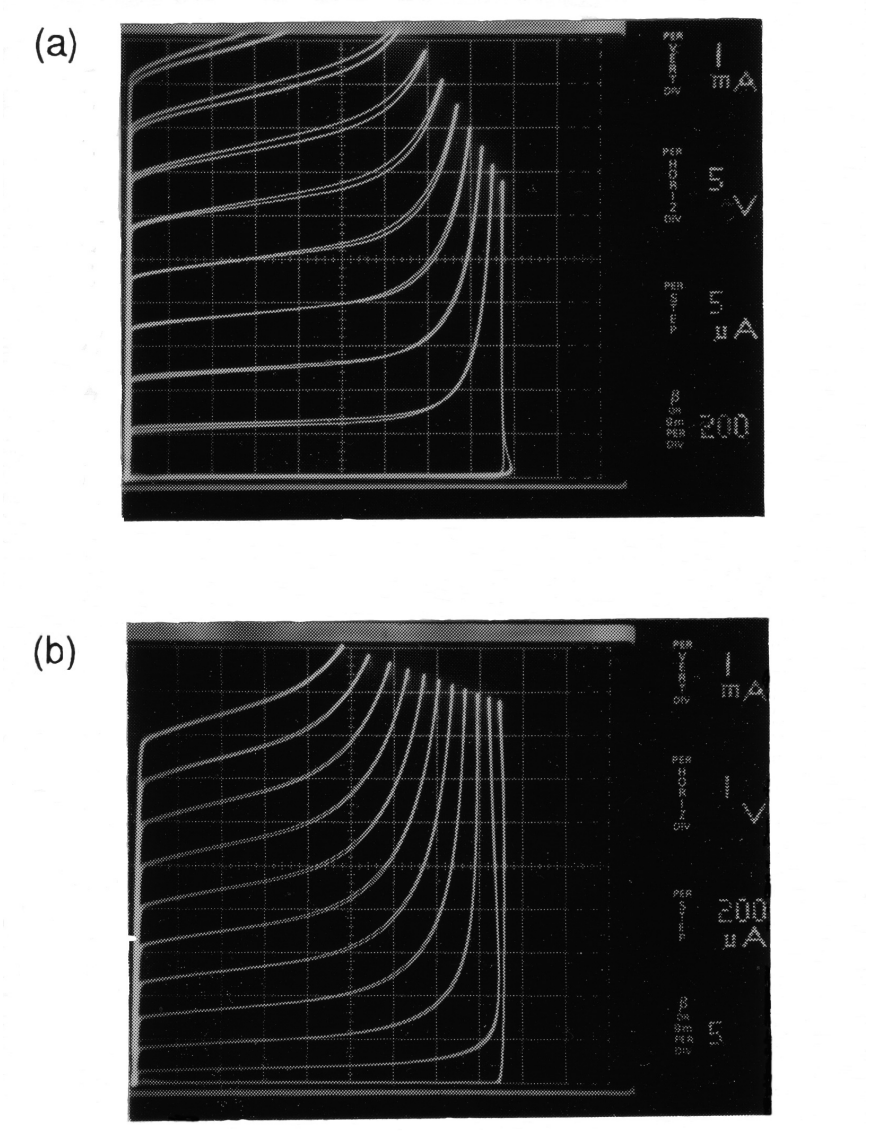


Figure 6.12: Curve-tracer displays for a 2N3904 (npn) BJT: (a) normal, (b) inverted orientation. Note the difference of the horizontal scales.

Concept Summary

The BJT is a three-terminal electronic device.

- It features two back-to-back pn junctions (npn or pnp) with junction regions sequentially designated as emitter, base, and collector.
- By convention, positive collector and base currents flow inward, and positive emitter current flows outward.
- There are four operating modes:
  - In the forward active mode, the base-emitter junction is forward biased and the base-collector junction is reverse biased.
    - Carriers are injected from the emitter to the base, and most of them diffuse across the base to allow collector absorption.
    - The collector terminal current is proportional to the base terminal current: $i_c = \beta_F \, i_b$ (with β_F typically 50 - 200).
    - A consistent condition has v_{ce} in excess of about 0.2 V.
    - The forward active mode supports valve-like behavior.
  - In the reverse active mode, the base-collector junction is forward biased and the base-emitter junction is reverse biased.
    - Carriers are injected from the collector to the base, and some of them diffuse across the base to allow emitter absorption.
    - The emitter terminal current is proportional to the base terminal current: $-i_e = \beta_R \, i_b$ (with β_R typically 1 - 5).
    - A consistent condition has v_{ec} in excess of about 0.2 V.
    - The reverse active mode is generally avoided.
  - In the saturation mode, both pn junctions are forward biased.
    - To zero order, the BJT terminals are shorted together.
    - The saturation mode reflects an on-state switch.
  - In the cutoff mode, both pn junctions are reverse biased.
    - To zero order, the BJT terminals are are open connections.
    - The cutoff mode reflects an off-state switch.
- Good BJTs typically have heavily doped emitter regions, moderately doped and thin base regions, and lightly doped collector regions.
- When combined as a Darlington pair, two BJTs behave like a single device with an effective β_F value of 10,000 or more.

6.2 Large-Signal Models

This section relates BJT operating principles to large-signal models that can be used for circuit analysis. We begin discussion with a general model. Then we use it to derive simpler models for valve and switching action.

The Ebers-Moll Model

The four operating modes that relate to different combinations of junction biasing in the BJT are incorporated in the **Ebers-Moll** model of Fig. 6.13. The diode and the dependent current source on the left side of the model circuit account for forward electron transport, and similar components on the right side of the model circuit account for reverse electron transport. Terminal currents are given by

$$i_c = \alpha_F I_F - I_R \,, \tag{6.19a}$$
$$i_e = I_F - \alpha_R I_R \,, \tag{6.19b}$$
$$i_b = (1-\alpha_F) I_F + (1-\alpha_R) I_R \,, \tag{6.19c}$$

where

$$I_F = I_{ES}\,[e^{\,qv_{be}/kT} - 1] \tag{6.20}$$

and

$$I_R = I_{CS}\,[e^{\,qv_{bc}/kT} - 1] \,. \tag{6.21}$$

We omit the diode ideality factors n_F and n_R, which are typically unity.

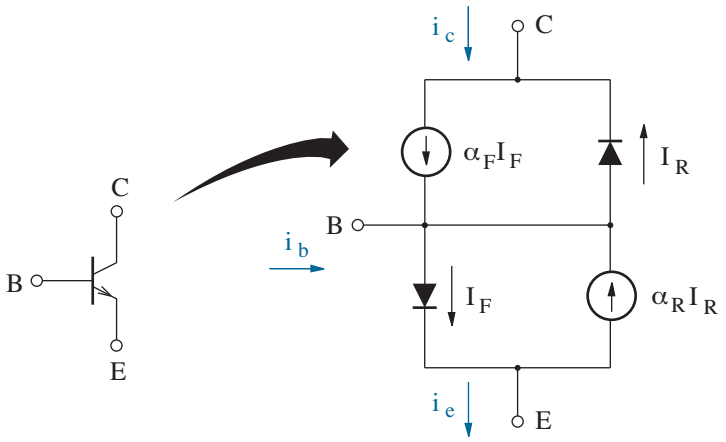


Figure 6.13: Ebers-Moll BJT (npn) model.

Unfortunately, the Ebers-Moll model of Fig. 6.13 does not have a form that is convenient for common-emitter circuit analysis. Thus, we seek to correct this difficulty by subjecting the model to a circuit transformation. Those readers who wish to forego the details of the transformation may skip to the end result, which is specified in Fig. 6.14.

As a first step, we recall (from Section 2.3) the extent of the electric potential variation across depletion regions in the vicinity of a pn junction. Given bias v, the potential changes from zero on the p side of the junction to $V_{bi} - v$ on the n side of the junction, where V_{bi} is the built-in potential. Then assuming a Boltzmann electron distribution, even under departures from equilibrium conditions (as long as they are small), we specify the ratio between the electron concentrations on the p and n sides of the junction through the relation

$$\frac{n_p}{n_n} = \exp\left[\frac{-q(V_{bi} - v)}{kT}\right] . \tag{6.22}$$

At equilibrium, when $v = 0$,

$$\frac{n_{po}}{n_{no}} = e^{\,-qV_{bi}/kT} . \tag{6.23}$$

So if $n_n \approx n_{no}$ on the n side of the junction for all $v \neq 0$, we substitute Eq. 6.23 into a modified Eq. 6.22 to obtain

$$n_p = n_{po}\, e^{\,qv/kT} . \tag{6.24}$$

The preceding result can be used to determine I_{common}, the electron current component common to the BJT emitter and collector currents. The electron concentration on the p side of the base-emitter junction is

$$n_{b,e} = n_{po}\, e^{\,qv_{be}/kT} , \tag{6.25}$$

and the electron concentration on the p side of the base-collector junction is

$$n_{b,c} = n_{po}\, e^{\,qv_{bc}/kT} , \tag{6.26}$$

where $n_{po} = {n_i}^2/N_a$ for the case of uniform base doping concentration N_a. As noted earlier, injected electrons tend to diffuse across the neutral base. The rate of diffusion is proportional to the electron concentration gradient, and the carrier profile is nearly linear (assuming negligible recombination within the base). Thus, with a gradient proportional to $n_{b,e} - n_{b,c}$, we find

$$I_{\text{common}} = I_s \left(e^{\,qv_{be}/kT} - e^{\,qv_{bc}/kT}\right) . \tag{6.27}$$

In this expression, I_s is the BJT **saturation current**.

With the help of Eq. 6.27, we now rewrite Eq. 6.19a in the form

$$\begin{aligned} i_c &= I_{\text{common}} + (\alpha_F I_{ES} - I_s)\, e^{\,qv_{be}/kT} \\ &\quad + I_s \left(e^{\,qv_{bc}/kT} - \frac{\alpha_F I_{ES}}{I_s} \right) - I_{CS} \left(e^{\,qv_{bc}/kT} - 1 \right) . \end{aligned} \tag{6.28}$$

We eliminate the second term on the right-hand side of this expression by requiring $\alpha_F I_{ES} = I_s$. If we also require $\alpha_R I_{CS} = I_s$, we have

$$i_c = I_{\text{common}} - \frac{I_s}{\beta_R} \left(e^{\,qv_{bc}/kT} - 1 \right) , \tag{6.29}$$

where $\beta_R = \alpha_R/(1-\alpha_R)$. Similar manipulations of Eq. 6.19b yield

$$i_e = I_{\text{common}} + \frac{I_s}{\beta_F} \left(e^{\,qv_{be}/kT} - 1 \right) , \tag{6.30}$$

where $\beta_F = \alpha_F/(1-\alpha_F)$.

At this juncture, you may suspect that we are guilty of cheating after we have independently required $\alpha_F I_{ES} = I_s$ and $\alpha_R I_{CS} = I_s$. Nevertheless, one can show that $\alpha_F I_{ES} = \alpha_R I_{CS}$ as a consequence of a **reciprocity** condition that applies when the BJT is treated as a two-port network in the limit of very small v_{be} and v_{bc} (see Problem 6.17).

The Ebers-Moll transformation is complete when we synthesize a simple circuit that exhibits the terminal relations specified by Eqs. 6.29 and 6.30. This yields the **hybrid-π** configuration shown in Fig. 6.14.

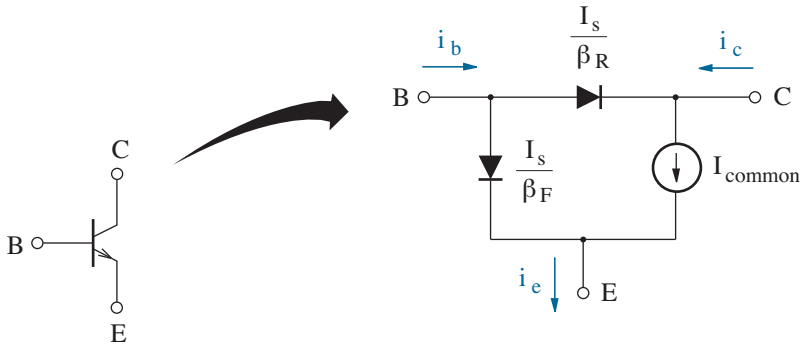


Figure 6.14: Hybrid-π representation of the Ebers-Moll BJT (npn) model. The base-emitter and base-collector diodes feature I_s/β_F and I_s/β_R saturation currents, respectively, and I_{common} is given by Eq. 6.27.

Equations 6.27, 6.29, and 6.30 predict some interesting behavior.

Example 6.3

Use the hybrid-π Ebers-Moll equations to determine two-terminal current-voltage characteristics (i_e vs v_{be}) for each of the npn BJT configurations shown in Fig. 6.15.

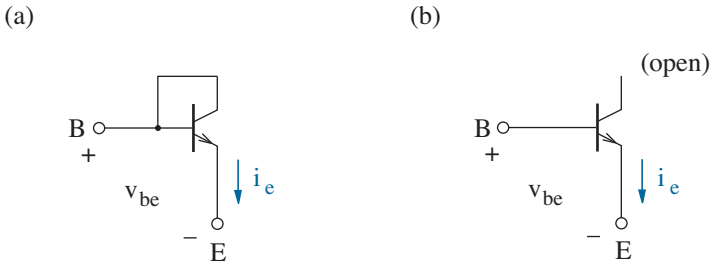


Figure 6.15: BJT configurations for Example 6.2.

Solution

In Fig. 6.15a, the BJT collector and base terminals are tied together so that $v_{bc} = 0$. Then from Eqs. 6.27 and 6.30,

$$i_e = I_s \left(1 + \frac{1}{\beta_F}\right) \left(e^{\,qv_{be}/kT} - 1\right) . \tag{6.31}$$

This is a familiar diode characteristic.

The second BJT configuration in Fig. 6.15b features an open collector terminal so that $i_c = 0$. Then in consideration of Eqs. 6.27 and 6.29,

$$I_s\, e^{\,qv_{bc}/kT} = I_s \left(1 + \frac{1}{\beta_R}\right)^{-1} \left(e^{\,qv_{be}/kT} + \frac{1}{\beta_R}\right) . \tag{6.32}$$

We insert this back into Eq. 6.30 and rearrange to obtain

$$i_e = I_s \left(\frac{1}{1+\beta_R} + \frac{1}{\beta_F}\right) \left(e^{\,qv_{be}/kT} - 1\right) . \tag{6.33}$$

This also has the form of a diode characteristic, but the effective saturation current is lower than that for the previous case. Thus, the configuration of Fig. 6.15a is preferred when a BJT is used in place of a diode (as in certain integrated circuits that more easily accommodate the former device).

Models for a BJT Valve

For many circuit applications, it is convenient to describe the BJT in terms of a simple large-signal model that is restricted to the forward active mode. The first-order forward active model shown in Fig. 6.16 derives from the hybrid-π representation of the Ebers-Moll model. Equations 6.27 and 6.29 determine the collector current. However, we drop any exponential terms containing v_{bc}, since $v_{bc} \ll -kT/q$. In turn,

$$i_c \approx I_s \left(e^{\,qv_{be}/kT} + \frac{1}{\beta_R} \right) . \tag{6.34}$$

Similarly, for the emitter current (Eqs. 6.27 and 6.30),

$$i_e \approx I_s \, e^{\,qv_{be}/kT} + \frac{I_s}{\beta_F} \left(e^{\,qv_{be}/kT} - 1 \right) . \tag{6.35}$$

And with $i_b = i_e - i_c$,

$$i_b \approx \frac{I_s}{\beta_F} \, e^{\,qv_{be}/kT} - \frac{I_s}{\beta_F} - \frac{I_s}{\beta_R} . \tag{6.36}$$

A further approximation assumes that the exponential terms in all of these expressions are dominant, since $v_{be} \gg kT/q$. Thus we have

$$i_c \approx \beta_F \, i_b \, , \tag{6.37}$$

in agreement with our discussion of the forward active mode in Section 6.1. A cumulative effect of both approximations is to replace the reverse-biased base-collector diode with an open circuit so that i_c reflects the action of a single dependent current source (I_c). Finally, $v_{ce} > v_{ce,sat} \sim +0.2$ V.

As noted previously, the pnp BJT functions like its npn counterpart, but all terminal currents and voltages are altered by sign changes in the governing relations. The device parameters remain positive. Thus, the pnp large-signal model is similar to that of Fig. 6.16. However, the base-emitter diode has reverse orientation, and $v_{ce} < v_{ce,sat} \sim -0.2$ V.

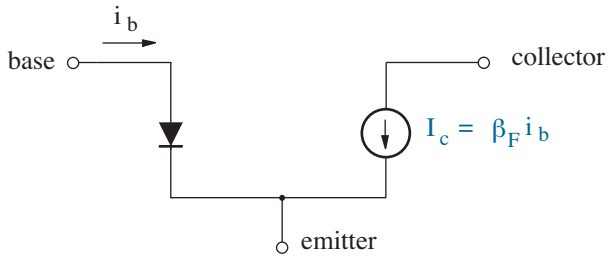


Figure 6.16: Large-signal model for the npn BJT in forward active mode. The pnp "valve" model is similar, apart from the diode orientation.

Beta Variations

As a basis for further discussion, Fig. 6.17 shows typical β_F variations with collector current i_c for the 2N3904 (npn) BJT at different temperatures. The curves are normalized with respect to the maximum β_F value at room temperature (25 °C). Note that the manufacturer has specified β_F as h_{FE}. Chapter 7 reveals the reason for this notation.

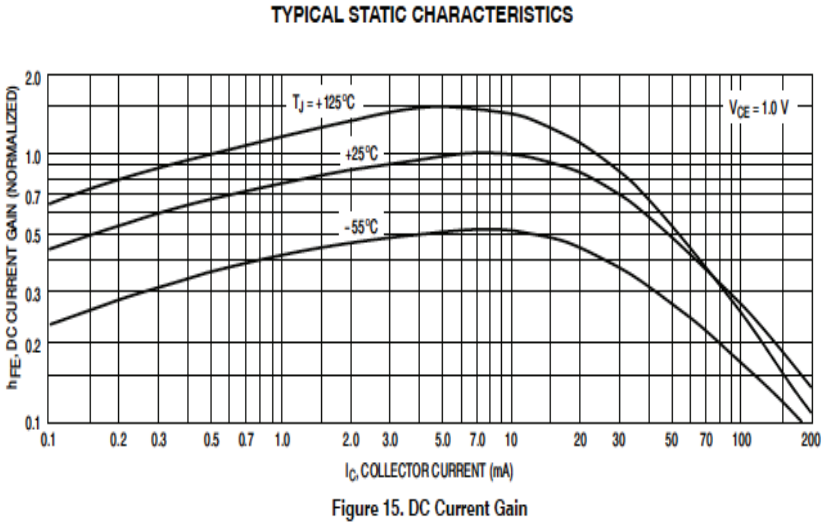


Figure 6.17: Normalized β_F variations with collector current for a 2N3904 (npn) BJT at different temperatures. Used with permission from SCILLC dba ON Semiconductor.

• Current Dependence

The β_F variation with collector current is easily understood with the help of Fig. 6.18, which shows $\ln i_c$ and $\ln i_b$ vs. qv_{be}/kT for the forward active mode of BJT operation. In this semi-logarithmic **Gummel plot**, β_F is proportional to the separation between the two curves.[2] Typically, the separation is nearly constant for moderate i_c, but it decreases for large and small i_c values (as in the case of the 2N3904 data).

We assume that the forward active mode ensures $i_c = \alpha_F i_e$ with α_F close to unity. For small and moderate qv_{be}/kT, the base-emitter junction exhibits ideal behavior, the junction saturation current is I_s (as in Eq. 6.27), and $\ln i_c$ increases with unity slope. However, for large qv_{be}/kT, the base-emitter junction is subjected to **high-level injection** conditions, and $\ln i_c$ increases with slope 1/2. The breakpoint between these two limiting slopes occurs at the **forward knee current** I_{KF}.

[2]The Gummel plot is one feature of the Gummel-Poon BJT model.

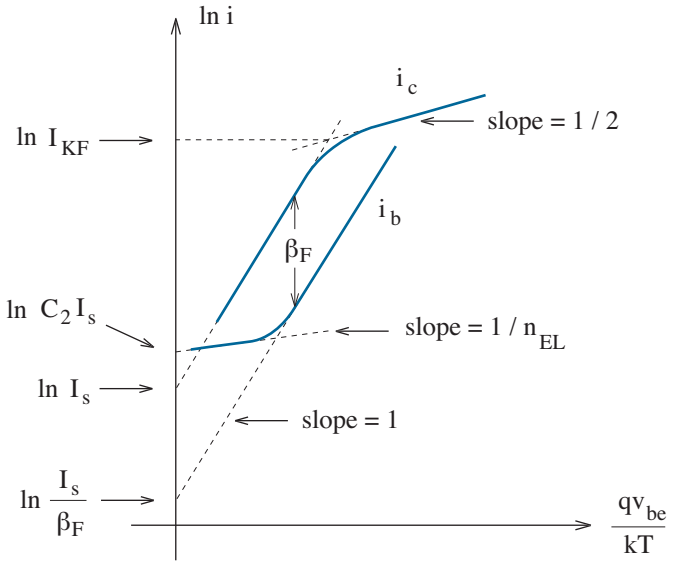


Figure 6.18: Typical Gummel plot of $\ln i_c$ and $\ln i_b$ vs qv_{be}/kT.

In view of our analysis of BJT behavior in the forward active mode, we expect to have $i_b = i_c/\beta_F$. So for moderate qv_{be}/kT, the $\ln i_b$ curve parallels the unity-slope $\ln i_c$ curve in Fig. 6.18. For large qv_{be}/kT, the $\ln i_b$ curve continues to increase with unity slope. However, for small qv_{be}/kT, we observe a more gradual decrease that accounts for previously neglected base recombination and base-to-emitter leakage effects. In the limit as $qv_{be}/kT \to 0$, the $\ln i_b$ curve approaches $\ln C_2 I_s$ with slope $1/n_{EL} \approx 1/2$. Both $C_2 I_s$ and n_{EL} are empirical constants that must be measured.

The same form of Gummel plot applies to β_R for the reverse active mode of BJT operation. Apart from I_s, the relevant new parameters are I_{KR}, $C_4 I_s$, and n_{CL}.

• Temperature Dependence

As indicated in Fig. 6.17, β_F increases monotonically with temperature at any particular collector current. This can be described using an empirical relation of the form

$$\frac{\beta_F(T_2)}{\beta_F(T_1)} = \left(\frac{T_2}{T_1}\right)^{X_{TB}}, \tag{6.38}$$

where T is the absolute temperature in degrees Kelvin and $X_{TB} > 0$. For the data of Fig. 6.17, let $T_1 = 25\ ^oC = 298$ K and $T_2 = 125\ ^oC = 423$ K. Then with $\beta_F(T_2)/\beta_F(T_1) = 1.5$ (at $i_c = 6$ mA), $X_{TB} \approx 1.15$. Typically, $1 < X_{TB} < 2.5$. The same temperature dependence applies to β_R.

Pending Complications

The behavior described by Eqs. 6.12 and 6.13 for the forward active mode of BJT operation no longer holds when v_{ce} is arbitrarily large. In particular, the base-collector junction suffers avalanche breakdown when its reverse bias voltage exceeds some limit, and the collector current increases sharply. (The breakdown is clearly evident in the curve-tracer display of Fig. 6.12.) BJTs with very thin metallurgical base widths exhibit similar behavior at the onset of **punchthrough**, a condition in which the base-side edges of the base-emitter and base-collector junction depletion regions come together to yield a fully-depleted base but an ineffective barrier to carrier crossings. Maximum v_{ce} values are provided on manufacturers' data sheets.

Parasitic resistive elements in series with the BJT terminal connections are appropriately labeled r_c, r_b, and r_e as shown in the model of Fig. 6.19. Of these, the base resistance r_b has the largest value (primarily a result of unfavorable geometry) and especially significant influence (in terms of BJT frequency response, which is considered in Chapter 8).

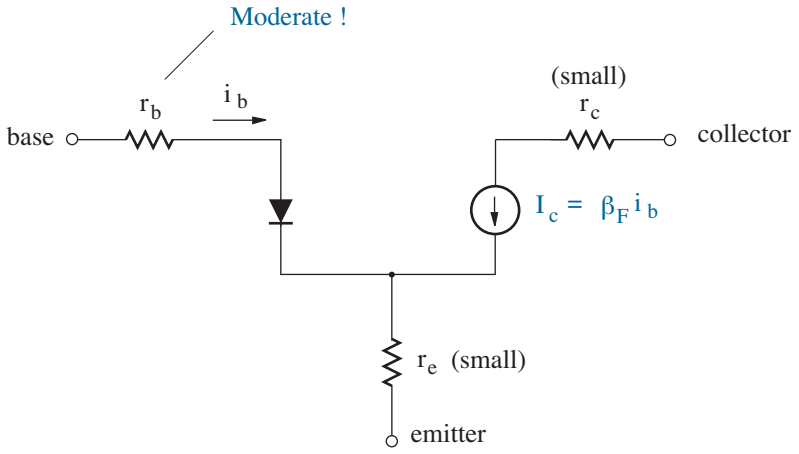


Figure 6.19: BJT model with parasitic resistive elements.

A final complication is deferred to Chapter 9, where we will find that v_{ce}-induced variations in the effective (non-depleted) base width lead to a forward active collector current of the form

$$i_c \approx \beta_F \, i_b \left(1 + \frac{v_{ce}}{V_A}\right) , \tag{6.39}$$

where V_A is constant. This second-order effect parallels one for MOSFETs, and it reflects a departure from an exclusively i_b-dependent current source.

Models for a BJT Switch

The BJT exhibits an "off" state with $i_c = 0$ when it is operated in cutoff such that base-emitter and base-collector junctions are *both* reverse biased. This reflects the trivial open-circuit model of Fig. 6.20 —neither junction passes significant reverse current.

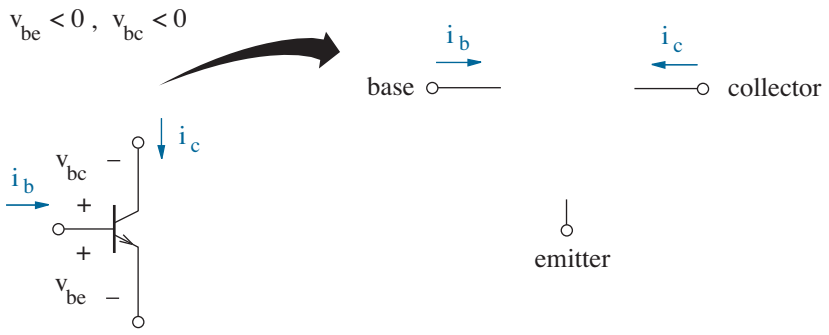


Figure 6.20: BJT model for operation in the cutoff mode.

Caution!

The term "cutoff" is sometimes used rather loosely to describe a BJT whose terminal currents are *approximately* consistent with the model of Fig. 6.20. For example, suppose we have an npn transistor with $v_{be} = 0.3$ V and $v_{bc} = -4.7$ V. Strictly speaking, this device is operating in the forward active mode, and the collector current is given by

$$i_c \approx I_s \, e^{\,qv_{be}/kT} \, . \tag{6.40}$$

But taking $I_s = 10^{-16}$ A (the SPICE default value), we find $i_c \approx 10$ pA, which is essentially zero. The base current is smaller by a factor of $1/\beta_F$. Thus, we are ill prepared to argue much difference between actual conditions and cutoff ($i_c = i_b = 0$).

In general, one can consider a BJT to be effectively "cutoff" when i_c is less than some reasonably small $i_{c,min}$ (say 1 μA). In terms of the base-emitter voltage, this corresponds to $v_{be} < v_{be,on}$, where

$$v_{be,on} \approx \frac{kT}{q} \, \ln\left(\frac{i_{c,min}}{I_s}\right) \, . \tag{6.41}$$

Typically, $v_{be,on}$ is slightly less than the familiar 0.7-V benchmark.

The BJT exhibits an "on" state with minimum $v_{ce} = v_{ce,sat}$ when it is operated in saturation such that base-emitter and base-collector junctions are *both* forward biased. These junction conditions suggest a zero-order saturation model in which the collector, base, and emitter terminals are shorted together. However, a better first-order model shown in Fig. 6.21 recognizes an approximate forward drop of 0.7 V across the base-emitter junction and $v_{ce} = v_{ce,sat} \sim 0.2$ V (as for typical BJT characteristic curves).

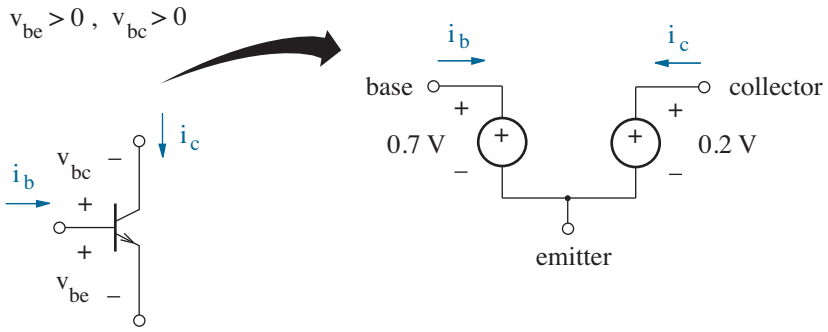


Figure 6.21: BJT model for operation in the saturation mode (npn device). The pnp model is similar, but the voltage sources are reversed.

With v_{ce} stuck at 0.2 V, the saturation model leads to a *circuit-defined* maximum collector current that is *smaller* than the transistor-defined $i_c = \beta_F\, i_b$ collector current otherwise established in the forward active mode. Thus for consistency with BJT saturation, we require

$$i_c < \beta_F\, i_b\,. \tag{6.42}$$

Nevertheless, two equivalent tests for BJT saturation are often convenient. The first test looks for a **forced common-emitter current gain** (β_{forced}) that is less than the common-emitter current gain observed in the forward active mode. Specifically,

$$\beta_{\text{forced}} = \frac{i_c}{i_b} < \beta_F\,. \tag{6.43}$$

The second test looks for a positive **overdrive base current** (i_{bs}) that exceeds the base current necessary to sustain the forward active mode. Specifically,

$$i_{bs} = i_b - \frac{i_c}{\beta_F} > 0\,. \tag{6.44}$$

Any one of these tests is conclusive.

Saturation Voltage

Thus far, we have used $v_{ce,sat} \approx 0.2$ V as the benchmark for saturation. We now consider the inverse problem—what is the value of $v_{ce,sat}$ if the BJT terminal currents are known?

Quantitative analysis begins with the Ebers-Moll large-signal equations. For any mode of BJT operation, we have

$$i_c = \alpha_F I_F - I_R, \tag{6.45a}$$
$$i_e = I_F - \alpha_R I_R, \tag{6.45b}$$
$$i_b = (1-\alpha_F) I_F + (1-\alpha_R) I_R, \tag{6.45c}$$

where

$$I_F = I_{ES}\left[e^{\,qv_{be}/kT} - 1\right] \tag{6.46}$$

and

$$I_R = I_{CS}\left[e^{\,qv_{bc}/kT} - 1\right]. \tag{6.47}$$

We determine v_{be} by eliminating I_R from Eqs. 6.45a and 6.45b. Thus,

$$i_e - \alpha_R\, i_c = I_{ES}\left[e^{\,qv_{be}/kT} - 1\right](1-\alpha_F\alpha_R). \tag{6.48}$$

And with $v_{be} \gg kT/q$ for the saturation mode,

$$v_{be} \approx \frac{kT}{q}\,\ln\left[\frac{i_e - \alpha_R\, i_c}{I_{ES}(1-\alpha_F\alpha_R)}\right]. \tag{6.49}$$

Similarly, we determine v_{bc} by eliminating I_F from Eqs. 6.45a and 6.45b. Thus,

$$\alpha_F\, i_e - i_c = I_{CS}\left[e^{\,qv_{bc}/kT} - 1\right](1-\alpha_F\alpha_R). \tag{6.50}$$

And with $v_{bc} \gg kT/q$ for the saturation mode,

$$v_{bc} \approx \frac{kT}{q}\,\ln\left[\frac{\alpha_F\, i_e - i_c}{I_{CS}(1-\alpha_F\alpha_R)}\right]. \tag{6.51}$$

Kirchhoff's voltage law requires

$$v_{ce} = v_{be} - v_{bc}. \tag{6.52}$$

So the saturation voltage is given by

$$v_{ce,sat} = \frac{kT}{q}\,\ln\left[\left(\frac{I_{CS}}{I_{ES}}\right)\frac{i_e - \alpha_R\, i_c}{\alpha_F\, i_e - i_c}\right]. \tag{6.53}$$

Equation 6.53 simplifies by requiring $I_{CS}/I_{ES} = \alpha_F/\alpha_R$ —this reciprocity condition was used to derive Eqs. 6.29 and 6.30 for i_c and i_e, respectively. Further, with $i_e = i_b + i_c$,

$$v_{ce,sat} = \frac{kT}{q} \ln\left[\left(\frac{\alpha_F}{\alpha_R}\right) \frac{i_b + (1-\alpha_R)i_c}{\alpha_F\, i_b - (1-\alpha_F)i_c}\right] . \tag{6.54}$$

However, $\alpha_F/(1-\alpha_F) = \beta_F$, $\alpha_R/(1-\alpha_R) = \beta_R$, and $i_c/i_b = \beta_{\text{forced}}$. In turn, our final result takes the form

$$v_{ce,sat} = \frac{kT}{q} \ln\left[\left(\frac{\beta_F}{\beta_R}\right) \frac{(1+\beta_R) + \beta_{\text{forced}}}{\beta_F - \beta_{\text{forced}}}\right] . \tag{6.55}$$

Consider a typical estimate for $v_{ce,sat}$. Suppose $\beta_F = 100$, $\beta_R = 5$, and $\beta_{\text{forced}} = 10$. Then $v_{ce,sat} = kT/q \ln 3.6 = 0.033$ V, an order of magnitude less than the 0.2-V saturation value used throughout previous discussion. What is going on?

Equation 6.55 represents an **intrinsic saturation voltage**, whereas the actual (measured) saturation voltage must include ohmic voltage drops across any parasitic collector resistance (r_c) and emitter resistance (r_e) that are *external* to the Ebers-Moll model of Fig. 6.13. Generally, BJT structural designs yield $r_c \gg r_e$. So ignore r_e, and let $r_c = 20\ \Omega$. Further, let $i_c = 10$ mA (apart from β_F, β_R, and β_{forced} values already specified). The new v_{ce} saturation value is $0.033 + 0.200 = 0.233$ V, which is close to the 0.2-V rule-of-thumb. To measure r_c, one considers the slope of the edge of the saturation region (see Problem 6.28).

The preceding results do not affect relatively crude hand calculations—any deviation from $v_{ce,sat} \approx 0.2$ V is an invitation for perilous precision. However, when performing a computer simulation of a circuit with a BJT at or near saturation, it is sometimes desirable to include an estimate for r_c in order to avoid unrealistic values for v_{ce}.

Table 6.1 shows computer simulation results—the circuit details need not concern us here—for five different sets of β_F, β_R, and r_c parameters. Note the relatively minor changes in i_c.

Table 6.1: Parameters for various saturation voltages (v_{ce}).

β_F	β_R	r_c (Ω)	i_b (mA)	i_c (mA)	v_{ce} (V)
100	1	0	4.26	4.97	0.030
100	0.2	0	4.27	4.94	0.064
100	5	0	4.24	4.99	0.010
100	0.2	20	4.27	4.84	0.161
100	5	20	4.24	4.89	0.107

Concept Summary

BJT circuit models support valve- or switch-like operation.

- A BJT valve appears as a dependent current source I_c between the collector and emitter. The current through a diode connected between the base and emitter is the means for control.
  - For npn devices,
    - The valve is closed for a reverse-biased diode ($v_{be} < 0$),
    - The valve is otherwise open with $I_c = \beta_F\, i_b$ ($i_b > 0$).
  - For pnp devices,
    - The valve is closed for a reverse-biased diode ($v_{be} > 0$),
    - The valve is otherwise open with $I_c = \beta_F\, i_b$ ($i_b < 0$).
  - The model assumes a consistent $|v_{ce}| \gtrapprox 0.2$ V.
- The preceding elementary model has complications that necessitate second-order hand or computer analysis:
  - The β_F factor varies with valve current,
  - The valve current is susceptible to collector-base junction breakdown and fully depleted base conditions (punchthrough),
  - BJTs typically have moderate parasitic base resistance,
  - The valve current has v_{ce} dependence.
- Numerous other complications are generally reserved for SPICE.
- A BJT switch has two distinct models:
  - If the base-emitter and -collector junctions are reverse biased, all three BJT terminals are isolated (switch off).
  - If the base-emitter and -collector junctions are forward biased,
    - $v_{be} \approx 0.7$ V and $v_{ce} \approx 0.2$ V (npn devices, switch on).
    - $v_{eb} \approx 0.7$ V and $v_{ec} \approx 0.2$ V (pnp devices, switch on).
    - The on-state switch satisfies any one of three tests:
      - $|i_c| < \beta_F\, |i_b|$,
      - $\beta_{\text{forced}} = |i_c/i_b| < \beta_F$,
      - $i_{bs} = i_b - i_c/\beta_F > 0$ (npn) or $i_{bs} = i_b - i_c/\beta_F < 0$ (pnp) (non-zero overdrive base current).
    - The $|v_{ce}|$ value is affected by parasitic collector resistance.

6.3 Introductory BJT Circuits

In this section, we consider the dc analysis of some simple BJT circuits, particularly those that take special advantage of valvelike device behavior. BJTs have non-zero base current, so our task is moderately challenging.

Elementary Analysis

Consider the BJT circuit shown in Fig. 6.22a. When confronted with a particular set of component and voltage-source values, we have no way of knowing whether the circuit has been intelligently designed so that the BJT provides a desirable form of collector-current control. For optimum npn "valve" action, we hope for a design that ensures the forward active mode with positive base to-emitter voltage and negative base-to-collector voltage. So we begin our analysis by assuming that the circuit's creator had such considerations in mind—someday you may think this is hopelessly naïve—and we replace the BJT with a first-order large-signal forward active model as shown in Fig. 6.22b.

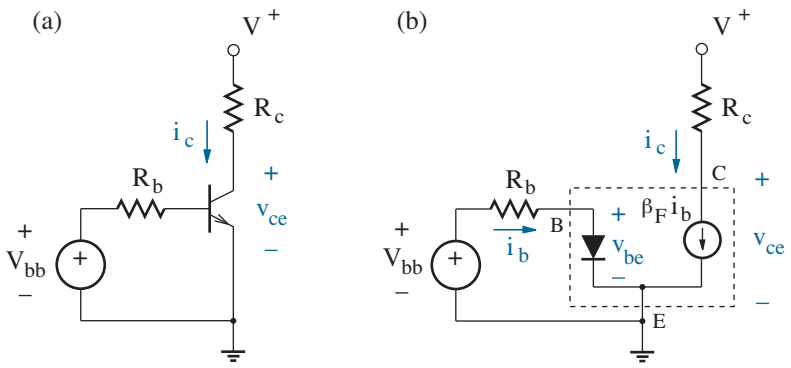


Figure 6.22: A simple BJT npn circuit and its large-signal equivalent under forward-active conditions. The base current is not directly specified.

Since the collector current is effectively controlled by the base current, we seek i_b by applying Kirchhoff's voltage law to the base-emitter loop. Specifically,

$$V_{bb} = i_b R_b + v_{be} \, . \tag{6.56}$$

This is straightforward. But now we need to find v_{be} for an imposing diode, certainly a less-than-pleasant prospect at first thought.

As discomfort fades into calm deliberation, we realize that the diode state is known—its "on" condition reflects the BJT forward active mode. So we simply replace the diode with another first-order model consisting of a single voltage source with value V_f. When estimating V_f, we are careful to remember that the forward voltage for a base-emitter diode is effectively related to the collector current and not the base current, which is smaller. Thus, for the case of a silicon BJT with collector current in the mA range, $v_{be} = V_f \sim 0.7$ V and

$$i_b = \frac{V_{bb} - 0.7}{R_b} \, . \tag{6.57}$$

In turn,

$$i_c = \beta_F \, i_b = \beta_F \, \frac{V_{bb} - 0.7}{R_b} \tag{6.58}$$

and

$$v_{ce} = V^+ - i_c R_c = V^+ - \beta_F \, \frac{V_{bb} - 0.7}{R_b} \, R_c \, . \tag{6.59}$$

A first check for solution consistency verifies the forward active mode. Hopefully, $i_c > 0$ (no cutoff) and $v_{ce} > v_{ce,sat} \sim 0.2$ V (no saturation). A second check verifies the estimate for v_{be}, given this first result for i_c. Except for gross estimation errors, we seldom bother with iterative circuit solutions that provide only slightly improved accuracy. Leave it to SPICE.

If one envisions the emitter terminal as a pair of "feet", the npn BJT likes to stand upright with its feet at or near the ground rail. In contrast, the pnp BJT is like a bat that likes to hang upside-down with its feet at or near the V^+ power-supply rail. Figure 6.23 shows a typical example.

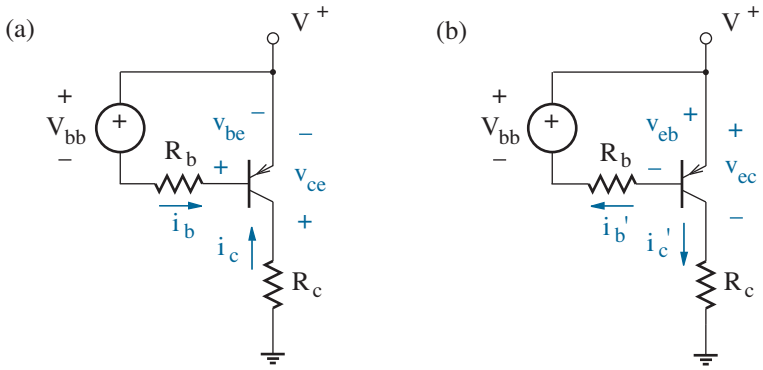


Figure 6.23: BJT pnp circuit (with two different sets of labels).

If the pnp terminal variables assume the conventional orientations shown in Fig. 6.23a, all have negative values within the BJT forward active mode. To demonstrate, we apply Kirchhoff's voltage law to the base-emitter loop:

$$V_{bb} = -v_{be} - i_b R_b \, . \tag{6.60}$$

We replace the BJT with a first-order large-signal forward active model, and we recall that for the pnp device, this is the same as the npn model, except for the orientation of the base-emitter diode. It is reasonable to expect that $-i_c$ lies in the mA range. So with $v_{be} = -V_f \sim -0.7$ V,

$$i_b = -\frac{V_{bb} - 0.7}{R_b} \, . \tag{6.61}$$

In turn,

$$i_c = \beta_F i_b = -\beta_F \, \frac{V_{bb} - 0.7}{R_b} \tag{6.62}$$

and

$$v_{ce} = -V^+ - i_c R_c = -V^+ + \beta_F \, \frac{V_{bb} - 0.7}{R_b} \, R_c \, . \tag{6.63}$$

We make a final check to verify the BJT forward active mode of operation. Hopefully, $i_c < 0$ (no cutoff) and $v_{ce} < v_{ce,sat} \sim -0.2$ V (no saturation).

For upbeat individuals who hate negative results, the same pnp circuit can be analyzed in terms of a set of non-conventional terminal variables as shown in Fig. 6.23b. It is easy to demonstrate that these variables all have positive values when the circuit operates in the forward active mode. Mark-up-the-diagram solutions usually feature positive results.

In the preceding examples, Eqs. 6.58 (6.62) and 6.59 (6.63) specify the BJT operating point (Q-point) for i_c and v_{ce}, respectively. Subject to cutoff and saturation restrictions, $|i_c|$ can be adjusted such that

$$0 < |i_c| < \frac{V^+ - |v_{ce,sat}|}{R_c} \, . \tag{6.64}$$

The limits in this inequality roughly correspond to the load-line extremes. Observe that the non-saturated collector current does not depend on R_c. Thus, Figs. 6.22 and 6.23 feature the BJT as a *current source* with a value governed by V_{bb}, R_b, and β_F. Yet again, we have physically realized a circuit element that we have often used since our early days of network analysis. The npn current source pulls current down through a load that is tied to the V^+ power-supply rail. In contrast, the pnp current source pushes current down into a load that is tied to the ground rail.

Exercise 6.7 The BJTs in the circuits of Fig. 6.24 feature β_F as given. Determine v_{out} in each case.

Ans: (a) $v_{out} = 1.3$ V (b) $v_{out} = 7.1$ V

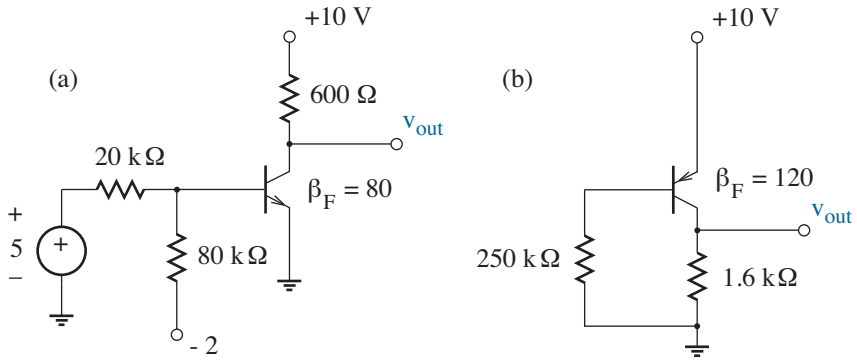


Figure 6.24: Circuits for Exercise 6.7.

Exercise 6.8 The BJTs in the circuits of Fig. 6.25 feature β_F as given. Design each circuit to achieve the specified value for v_{out}.

Ans: (a) $R = 430$ kΩ (b) $R = 170$ kΩ

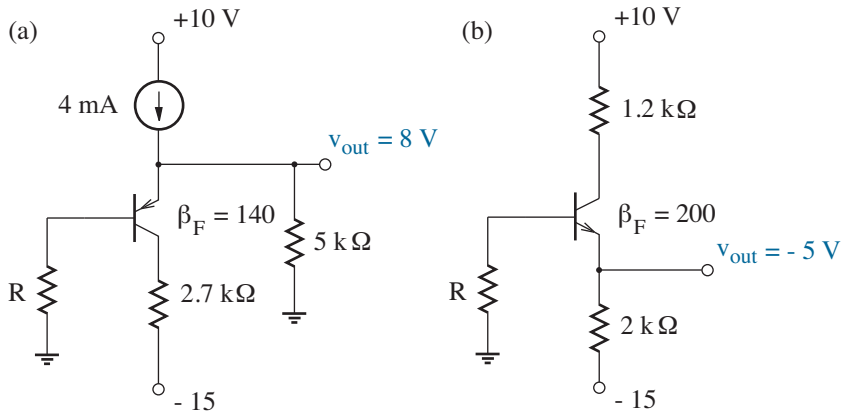


Figure 6.25: Circuits for Exercise 6.8.

Example 6.4

For the circuit of Fig. 6.26, determine R such that $v_{out} = 10$ V.

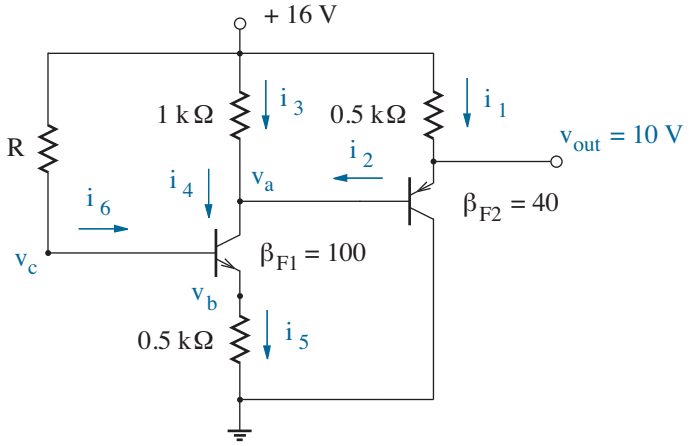


Figure 6.26: Circuit for Example 6.4.

Solution

We solve this problem by identifying as many currents and node voltages as possible while fearlessly marking up the circuit diagram with the results. In order of interest, we have:

- $i_1 = (16 - 10) \,/\, 0.5 = 12$ mA.
- $i_2 = i_1 \,/\, (\beta_{F2} + 1) \approx 0.3$ mA. (Note the direction of i_2.)
- $v_a = 10 - 0.7 = 9.3$ V. (Note the polarity of the 0.7-V drop.)
- $i_3 = (16 - 9.3) \,/\, 1 = 6.7$ mA.
- $i_4 = 6.7 + 0.3 = 7.0$ mA $\approx i_5$ (since β_{F1} is large).
- $v_b = 7.0 \times 0.5 = 3.5$ V.
- $v_c = 3.5 + 0.7 = 4.2$ V.
- $i_6 = i_4 \,/\, \beta_{F1} = 0.07$ mA.
- Finally, $R = (16 - 4.2) \,/\, 0.07 \approx 170\,\text{k}\Omega$.

None of these individual calculations are particularly difficult.

Power Amplification

A special BJT "valve" application is the process of power amplification, not necessarily from voltage gain—we reserve that topic for Chapter 7—but also through increased drive current. Consider the circuits of Fig. 6.27:

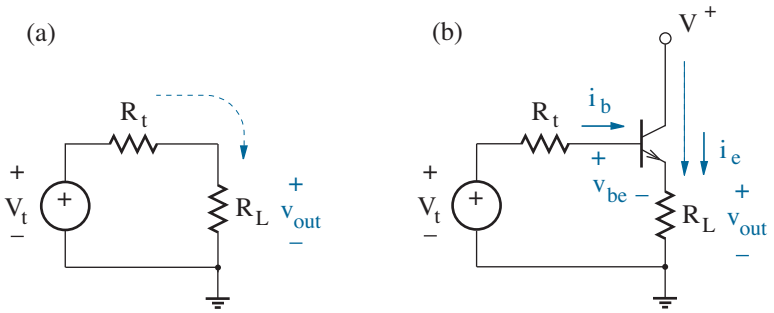


Figure 6.27: Thevenin signal source with two types of interface to load R_L. The dashed arrows indicate the primary paths for load current.

The circuit of Fig. 6.27a is simply a voltage divider. Thus,

$$v_{out} = V_t \, \frac{R_L}{R_t + R_L} \, , \tag{6.65}$$

and $v_{out} \approx V_t$ if $R_t \ll R_L$. The load current is all supplied from source V_t. Now examine the circuit of Fig. 6.27b. We assume that the BJT operates in the forward active mode subject to $v_{be} \sim 0.7$ V and $i_e = (\beta_F + 1) i_b$. Then we apply KVL to the base-emitter loop. Specifically,

$$V_t = i_b R_t + 0.7 + (\beta_F + 1)\, i_b R_L \, . \tag{6.66}$$

We solve this equation for i_b. In turn,

$$v_{out} = (\beta_F + 1)\, i_b R_L = \frac{(\beta_F + 1) R_L}{R_t + (\beta_F + 1) R_L} \, (V_t - 0.7) \, , \tag{6.67}$$

and $v_{out} \approx (V_t - 0.7)$ if $R_t \ll (\beta_F + 1) R_L$. Thus, the effective load resistance from the Thevenin source perspective has increased by a factor of $\beta_F + 1$. Most of the load current is now supplied from the V^+ source (as indicated). The Thevenin source provides current i_b and the load receives $(\beta_F + 1)\, i_b$, so the current gain is $\beta_F + 1$. The signal and load voltages are comparable (apart from a 0.7-V offset), so the power gain approaches $\beta_F + 1$.

What we have just witnessed is the action of an **emitter follower**—the BJT emitter "follows" the signal source if the drop across R_t is small. Or at least it tries. The emitter node voltage can never become negative without the BJT entering effective cutoff beforehand. At the other extreme, the emitter node voltage can never exceed $V^+ - v_{ce,sat}$ without saturation. Thus, a sinusoidal source yields output waveforms like those of Fig. 6.28. The second case (Fig. 6.28b) is clearly better. Only the *upper* half of the input is made available at the output, but at least it is not clipped.

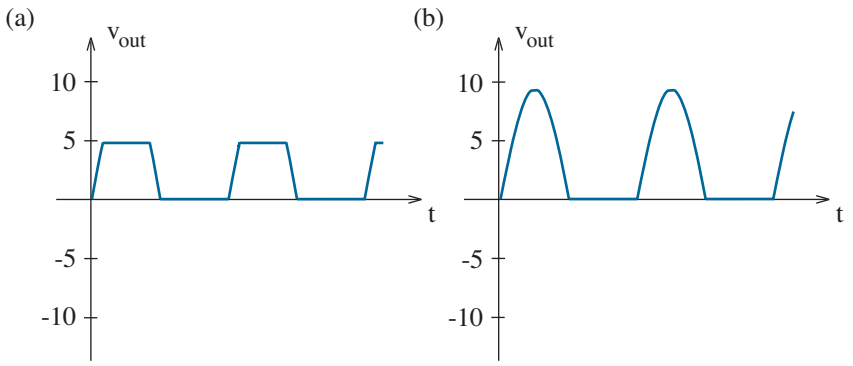


Figure 6.28: Emitter follower responses to $V_t = 10 \sin \omega t$ (see Fig. 6.27) with $R_t = R_L = 1$ kΩ and $\beta_F = 200$: (a) $V^+ = 5$ V; (b) $V^+ = 15$ V.

Figure 6.29 shows an emitter-follower circuit with a pnp transistor and a negative supply voltage. You should not be surprized that only the *lower* half of the input is made available at the output in the absence of clipping.

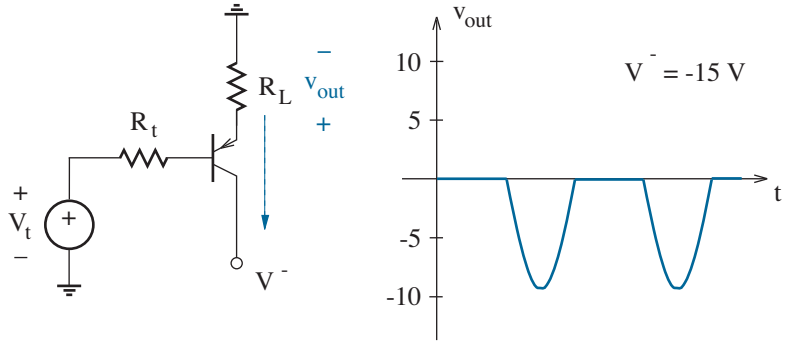


Figure 6.29: Emitter follower circuit and typical response for a pnp BJT. The circuit parameters are similar to those for Fig. 6.28b.

So how do we achieve power gain and a complete sinusoidal output?

You guessed it. One simply configures the npn and pnp emitter follower circuits to work together as shown in Fig. 6.30. During positive half cycles, the npn transistor pushes current into the load from the V^+ supply while the pnp device is idle (cutoff). And during negative half cycles, the pnp transistor pulls current from the load to the V^- supply while the npn device is idle (cutoff). This is the action of a **push-pull** power amplifier.

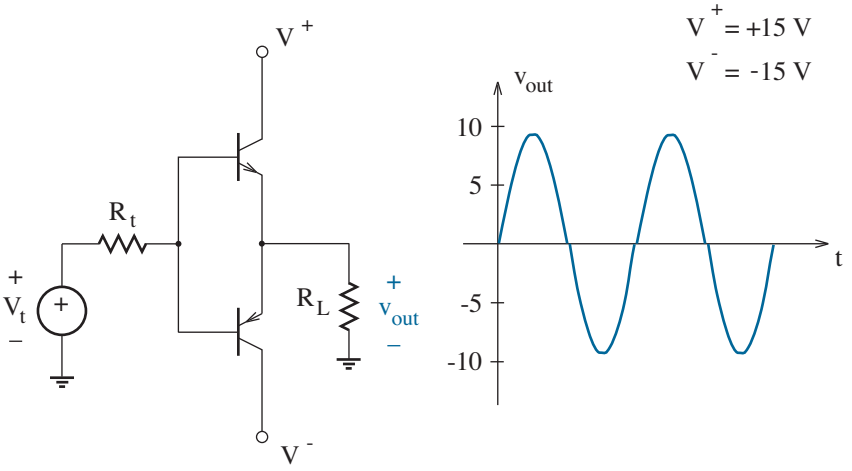


Figure 6.30: Push-pull power amplifier and typical response.

The push-pull output looks promising. Yet closer inspection reveals that both transistors are cutoff within about 0.7 V of each V_t zero crossing. (see Fig. 6.31). This unwanted behavior is called **crossover distortion**. What can we do to avoid it?

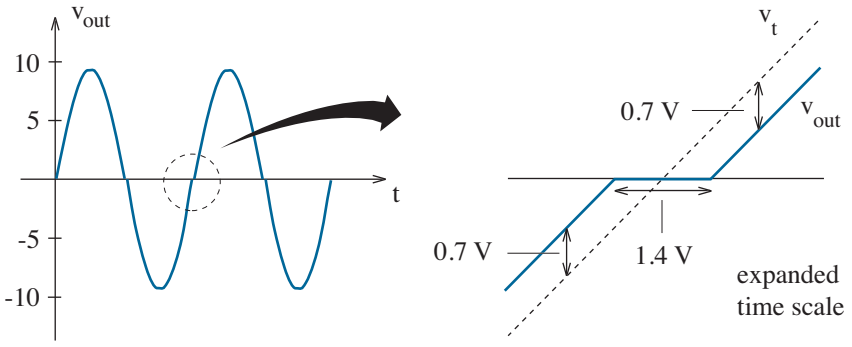


Figure 6.31: Push-pull crossover distortion.

The key to straightening the push-pull output is a more crooked input. For instance, we let $V_t \to V_t + 0.7$ V for positive half cycles so that $v_{out} \approx (V_t + 0.7) - 0.7 = V_t$. Similarly, we let $V_t \to V_t - 0.7$ V for negative half cycles so that $v_{out} \approx (V_t - 0.7) + 0.7 = V_t$. The 0.7-V counteroffsets come from forward-biased diode "batteries" as shown in Fig. 6.32.

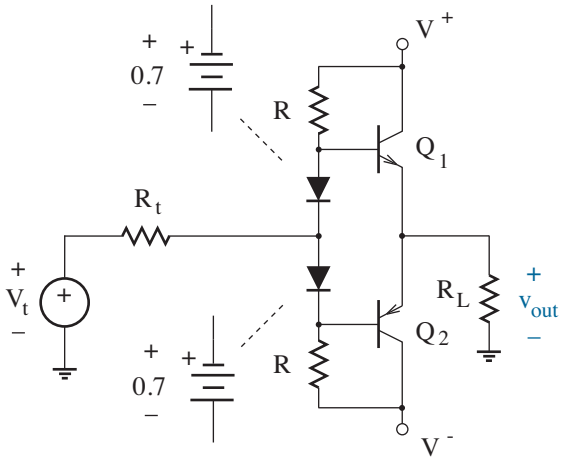


Figure 6.32: Push-pull power amplifier with diode counteroffsets.

Alternatively, we can choose to excite the separate npn and pnp current drivers with complementary emitter follower circuits as shown in Fig. 6.33. This has the added benefit of much reduced V_t source current.

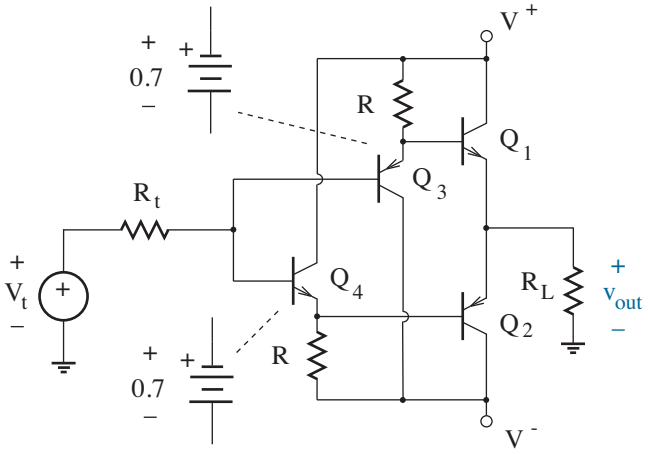


Figure 6.33: Push-pull power amplifier with emitter-follower counteroffsets.

No doubt you are now eager to trash part of your stereo system in favor of the power amplifier of Fig. 6.33. The load is one of your 8-Ω speakers. But what will you do if someone shorts the output connection to ground? Large currents through Q_1 or Q_2 could lead to a toasted system.

Figure 6.34 shows one solution featuring two extra BJTs (Q_5 and Q_6) and a pair of R_x resistors in series with the Q_1 and Q_2 emitter terminals. Consider a positive output excursion. The output voltage does not change significantly if $R_x \ll R_L$. Nevertheless, Q_5 begins to draw collector current at the expense of base current for Q_1 if the load current is sufficiently large to make the voltage drop across R_x comparable to 0.7 V. Thus, Q_5 and Q_6 serve protective roles that limit the available output current.

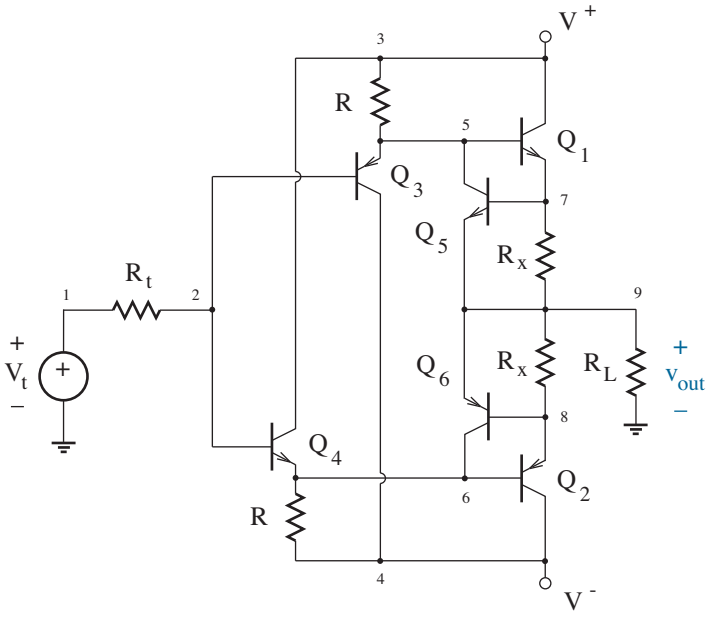


Figure 6.34: Push-pull power amplifier with output protection.

Think about the last few pages. We started with one BJT in Fig. 6.27, and now we have six BJTs in Fig. 6.34. The gradual increase in complexity does not lead to difficult analysis if one understands the role of each device.

Exercise 6.9 The power amplifier of Fig. 6.34 features $V^+ = +15$ V, $V^- = -15$ V, and $\beta_F = 100$ for all transistors. The load resistance is 8 Ω. Determine R so that the output voltage can swing between ±12 V.

Ans: $R = 170\ \Omega$

Switch Action

Like the MOSFET, the BJT switch has three major applications:

- Digital logic circuits
- Analog switching (primarily for current diversion)
- Power conditioning and control

The BJT switch has a well-defined 0.7-V turn-on voltage, and it can handle large currents with a modest device area and cost. Nevertheless, designers must tolerate non-zero base current and relatively slow transitions between saturation and cutoff (a liability that is explored in the chapter Appendix). Thus, the BJT switch has gradually given way to the MOSFET.

That said, we seek practice with saturation and cutoff analysis.

Example 6.5

Determine i_c for the circuit of Fig. 6.35.

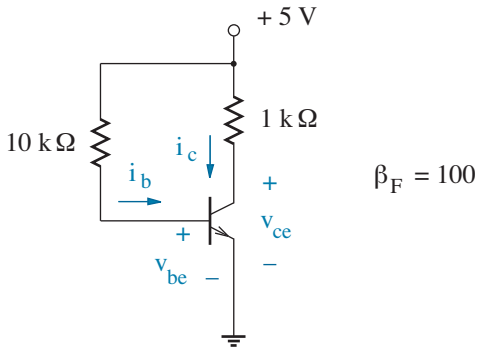


Figure 6.35: Circuit for Example 6.5.

Solution

The base-emitter junction is clearly forward biased. So with $v_{be} \approx 0.7$ V, $i_b = (5 - 0.7)/10 = 0.43$ mA. From experience, this seems like a rather large base current—large enough perhaps to support the saturation mode. Then with $v_{ce} = v_{ce,sat} \approx 0.2$ V, $i_c = (5 - 0.2)/1 = 4.8$ mA. The saturation condition is easily verified, since $\beta_{\text{forced}} = i_c/i_b \approx 11 < \beta_F$.

If we had initially assumed the forward active mode, we would require $i_c = \beta_F i_b = 43$ mA. This yields an unacceptable $v_{ce} = 5 - 1 \times 43 = -38$ V, and we are subsequently forced to consider the saturation mode.

Example 6.6

The transistors in Fig. 6.36 are either "off" (cutoff) or "on" (saturation). Determine the on-state collector currents, then find the maximum R value that preserves the observed conditions.

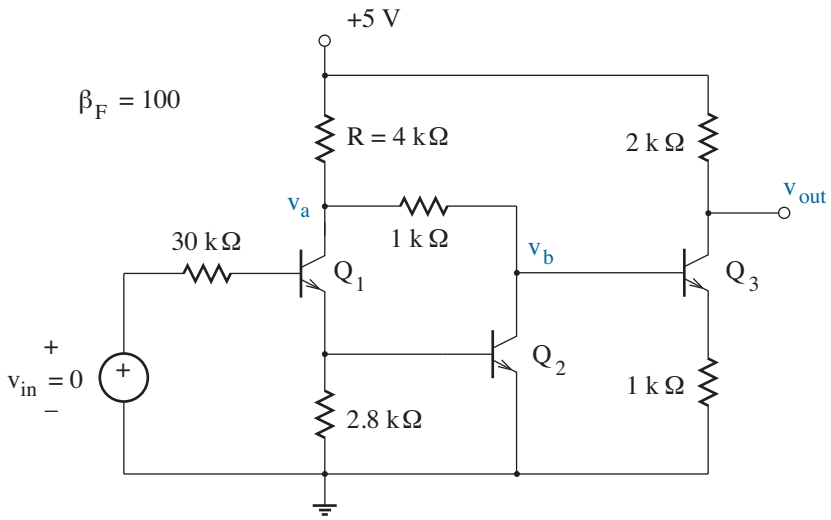


Figure 6.36: Circuit for Example 6.6.

Solution

Since $v_{in} = 0$, we expect $v_{be}(Q_1) = 0$ so that Q_1 is cutoff and no current passes through the 2.8-kΩ resistor. In turn, $v_{be}(Q_2) = 0$, and Q_2 is cutoff. Thus, Q_1 and Q_2 are effectively removed from the circuit.

It now seems reasonable to expect that Q_3 is "on" since its base node is tied to +5 V through an effective 5-kΩ resistance. We let $v_{be}(Q_3) = 0.7$ V and $v_{ce}(Q_3) = 0.2$ V (as in the first-order saturation model). Then for the Q_3 collector-emitter loop,

$$5 = 2\, i_c(Q_3) + 0.2 + 1\,[\, i_c(Q_3) + i_b(Q_3)\,]\,. \tag{6.68}$$

And for the Q_3 base-emitter loop,

$$5 = 5\, i_b(Q_3) + 0.7 + 1\,[\, i_c(Q_3) + i_b(Q_3)\,]\,. \tag{6.69}$$

Observe that we have expressed the Q_3 emitter current as $i_c(Q_3) + i_b(Q_3)$. If Q_3 is in saturation, $i_b(Q_3)$ can be significant in relation to $i_c(Q_3)$.

We subtract Eq. 6.69 from Eq. 6.68 to obtain $i_c(Q_3) = 2.5\, i_b(Q_3) + 0.25$. Then we substitute this relation back into Eq. 6.68 or Eq. 6.69 to find $i_c(Q_3) = 1.44$ mA, $i_b(Q_3) = 0.477$ mA, and $\beta_{\text{forced}} = i_c(Q_3)/i_b(Q_3) = 3.0$. The latter is less than $\beta_F = 100$, so Q_3 is indeed saturated.

Just as a check, we note that node voltage v_a is $5 - 4\, i_b(Q_3) = 3.1$ V, and node voltage v_b is $5 - 5\, i_b(Q_3) = 2.6$ V. These values are sufficient to ensure $v_{bc} < 0$ and cutoff for both Q_1 and Q_2.

If R becomes too large, Q_3 will operate in the forward active mode with $i_b(Q_3) = i_c(Q_3)/\beta_F$. To find $i_c(Q_3)$ at the *onset* of the forward active mode, we use Eq. 6.68, but we neglect $i_b(Q_3)$ so that $i_c(Q_3) \approx 1.6$ mA. In turn, $v_a = 1\,(1.6) + 0.7 + 1\,(0.016) \approx 2.3$ V, and $R = (5 - 2.3)/0.016 = 170$ kΩ.

Exercise 6.10 Determine β_{forced} for each of the circuits in Fig. 6.37.

Ans: (a) $\beta_{\text{forced}} = 32$ (b) $\beta_{\text{forced}} = 14$

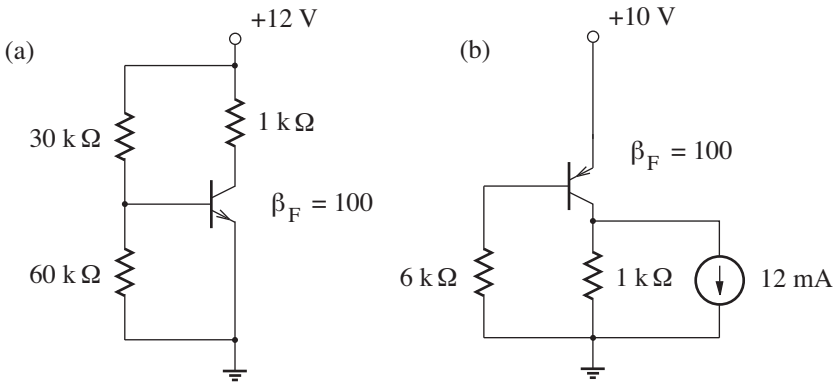


Figure 6.37: Circuits for Exercise 6.10.

Exercise 6.11 Consider the circuit of Fig. 6.36, but let $v_{in} = 5$ V. Determine β_{forced} for each BJT in saturation.

Ans: Q_1: $\beta_{\text{forced}} = 2.7$ Q_2: $\beta_{\text{forced}} = 3.6$ Q_3: cutoff

Concept Summary

For valve applications with a single BJT (in the forward active mode)

- An npn device
  - Has its emitter at the lowest circuit potential in relation to the base and collector (stands upright).
  - Establishes $v_{be} \approx 0.7$ V so that external circuit conditions can impose a particular $i_b > 0$ (current flows into the base).
  - Establishes $i_c = \beta_F\, i_b$ so that external circuit conditions can impose a particular v_{ce}.
  - Requires a consistent $v_{ce} \gtrapprox 0.2$ V.
- A pnp device
  - Has its emitter at the highest circuit potential in relation to the base and collector (hangs upside down).
  - Establishes $v_{be} \approx -0.7$ V so that external circuit conditions can impose a particular $i_b < 0$ (current flows out from the base).
  - Establishes $i_c = \beta_F\, i_b$ so that external circuit conditions can impose a particular v_{ce}.
  - Requires a consistent $v_{ce} \lessapprox -0.2$ V.

For power amplifiers (that sometimes combine BJTs),

- An emitter follower
  - Uses an npn device to push current into a ground-connected load so that the load and controlling base voltages differ by ~ 0.7 V.
  - Uses a pnp device to pull current from a ground-connected load so that the load and controlling base voltages differ by ~ 0.7 V.
- A push-pull amplifier
  - Uses npn and pnp devices for a full range of output excursions.
  - Suffers cross-over distortion without diode-like counteroffsets.
  - Can feature BJTs that protect the output from large currents.

Switch applications are deferred to the chapter Appendix.

6.4 SPICE Analysis

SPICE BJT circuit simulations require a .model statement of the form

.model < name > NPN [model parameters]

or

.model < name > PNP [model parameters]

for npn and pnp devices, respectively. Table 6.2 lists the most significant static BJT .model parameters. Those that apply to pending complications are highlighted in blue.

Although the large number of .model parameters may seem formidable, many of them can be ignored without peril. In particular, all of the reverse-mode parameters, the three parasitic resistance values, and the temperature coefficients are often unnecessary for successful SPICE simulations.

Of the remaining SPICE parameters, we invariably require a BF value (as determined from BJT curve-tracer measurements). IS is often required, and IKF, ISE and NE tend to assume default values. Since one frequently lacks a Gummel plot (Fig. 6.18), it is sometimes desirable to obtain a *crude* estimate for IS by measuring v_{be} at some moderate i_c value—say 1 mA—subject to the forward active mode. Then

$$\text{IS} \approx i_c \,/\, \exp\left(\frac{v_{be}}{kT/q}\right), \tag{6.70}$$

where $kT/q = 25.9$ mV at room temperature. Similarly, we can obtain a *crude* estimate for IKF by assuming

$$\beta_F \approx \frac{\text{BF}}{1 + \dfrac{i_c}{\text{IKF}}}, \tag{6.71}$$

so that IKF is the collector current for which β_F has a half-maximum value. (This is usually indicated on the manufacturer's data sheet.) Estimates for ISE and NE are unnecessary if the BJT collector current is always moderate or large (> 0.1 mA). Note that SPICE does not simulate the effects of base-collector junction breakdown at large v_{ce} values.

The temperature dependence of I_s is expressed terms of parameter XTI. Specifically,

$$\frac{I_s(T_2)}{I_s(T_1)} = \left(\frac{T_2}{T_1}\right)^{XTI} \exp\left[\frac{qE_g}{kT_2} - \frac{qE_g}{kT_1}\right], \tag{6.72}$$

where T is the absolute temperature in degrees Kelvin and E_g is the semiconductor bandgap energy. Typically, $X_{TI} = 3$.

Table 6.2: Static BJT Parameters

Symbol	SPICE keyword	Parameter Name	Default value	Unit
I_s	IS	Saturation current	10^{-16}	A
β_F	BF	Maximum forward current gain	100	—
V_A	VAF	Forward Early voltage	∞	V
I_{KF}	IKF	Corner for β_F high-current roll-off	∞	A
$C_2 I_s$	ISE	Base-emitter leakage saturation current	0	A
n_{EL}	NE	Base-emitter leakage emission coefficient	1.5	—
n_F	NF	Forward ideality factor	1	—
β_R	BR	Maximum reverse current gain	1	—
V_B	VAR	Reverse Early voltage	∞	V
I_{KR}	IKR	Corner for β_R high-current roll-off	∞	A
$C_4 I_s$	ISC	Base-collector leakage saturation current	0	A
n_{CL}	NC	Base-emitter leakage emission coefficient	2	—
n_R	NR	Reverse ideality factor	1	—
r_b	RB	Zero-bias base resistance	0	Ω
r_c	RC	Collector ohmic resistance	0	Ω
r_e	RE	Emitter ohmic resistance	0	Ω
X_{TB}	XTB	Forward and reverse β temperature coefficient	0	—
X_{TI}	XTI	Saturation current temperature coefficient	3	—
E_g	EG	Bandgap Energy	1.11	eV

Example 6.7

The push-pull amplifier of Fig. 6.34 features $V^+ = +15$ V, $V^- = -15$ V, $R_t = 100\ \Omega$, $R = 200\ \Omega$, $R' = 2\ \Omega$, and $R_L = 10\ \Omega$.

For the npn BJTs: $I_s = 5.0 \times 10^{-14}$ A, $\beta_F = 150$, $V_A = 100$ V
For the pnp BJTs: $I_s = 5.0 \times 10^{-14}$ A, $\beta_F = 150$, $V_A = 100$ V

Use SPICE to determine the Q_1 and Q_2 collector currents when $V_t = 0$.

Solution

The SPICE code takes the following form:

```
* Test Circuit - Example 6.7

Vt      1       0       0
Rt      1       2       100
V+      3       0       +15
V-      4       0       -15
R1      3       5       200
R2      4       6       200
Rp1     7       9       2
Rp2     8       9       2
RL      9       0       10
Q1      3       5       7       BJTN
Q2      4       6       8       BJTP
Q3      4       2       5       BJTP
Q4      3       2       6       BJTN
Q5      5       7       9       BJTN
Q6      6       8       9       BJTP

.model  BJTN    NPN(IS = 50f, BF = 150, VAF = 100)
.model  BJTP    PNP(IS = 50f, BF = 150, VAF = 100)

.op
.end
```

The BJT nodes are specified in the order of the collector, base, and emitter connections, and the .op statement provides the dc operating conditions. We look to the output file to find

IC (Q1) = +17.7 mA IC (Q2) = -17.7 mA

while the load receives no current. Thus, the power devices Q_1 and Q_2 are "idling" and ready to provide load current as needed in response to V_t.

BJT Capacitance

Dynamic SPICE simulations of analog and digital BJT circuits necessitate the specification of several device parameters that model capacitive effects. BJT capacitance between the base and emitter (C_{be} or C_π) and capacitance between the base and collector (C_{bc} or C_μ) have two components:

$$C_{be}\ (C_\pi) = C_{be}\ (\text{depletion}) + C_{be}\ (\text{diffusion}) \tag{6.73}$$

and

$$C_{bc}\ (C_\mu) = C_{bc}\ (\text{depletion}) + C_{bc}\ (\text{diffusion})\ . \tag{6.74}$$

The first component is the familiar voltage-dependent depletion capacitance linked to the carrier-free regions in the vicinity of a particular pn junction. As for the SPICE diode model considered in Chapter 3,

$$C_{be}\ (\text{depletion}) = CJE\left(1 - \frac{v_{be}}{VJE}\right)^{-MJE} \tag{6.75}$$

and

$$C_{bc}\ (\text{depletion}) = CJC\left(1 - \frac{v_{bc}}{VJC}\right)^{-MJC}\ . \tag{6.76}$$

To find the zero-bias CJE capacitance, one must perform a measurement (details in Chapter 8) or an extraction based on a particular BJT layout. Figure 6.38 shows a common geometry. The areal zero-bias base-emitter junction capacitance is known from fabrication characterization or Eq. 3.17. In turn, CJE is the product of the areal capacitance and the emitter area. Perimeter effects are often ignored. Similar considerations apply to CJC.

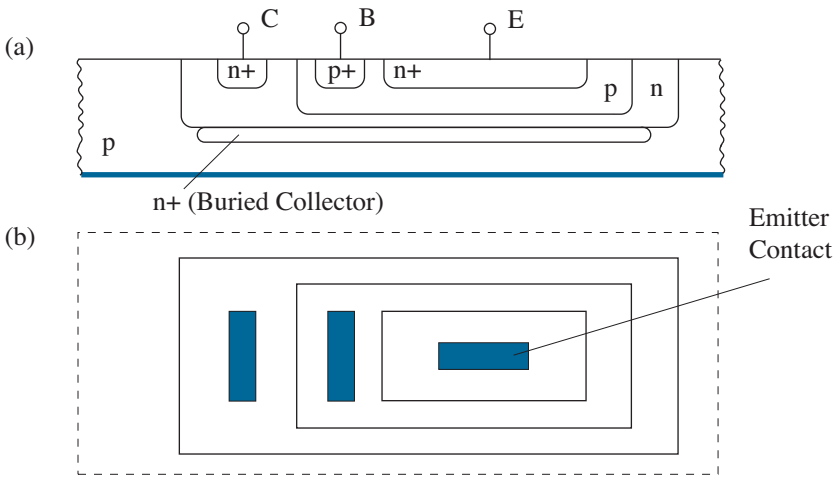


Figure 6.38: Integrated-circuit BJT layout: (a) side view; (b) top view.

Parameters VJE and VJC are found through experiment or by applying Eq. 2.19 for V_{bi}. The MJE and MJC parameters that relate to pn junction grading tend to assume values of 0.33 and 0.5, respectively.

Integrated BJTs need to account for a collector-to-substrate depletion capacitance (C_{cs}) that varies according to the relation

$$C_{cs} = CJS \left(1 - \frac{v_{bc}}{VJS}\right)^{-MJS}. \tag{6.77}$$

SPICE assumes that this capacitance connects between the collector and ground as shown in the large-signal BJT model of Fig. 6.39. Alternatively, when the p-substrate node is at the most negative circuit potential, the BJT net-list statements require an extra node variable. For example,

```
Q1      3       5       7       4       BJTN
```

indicates a BJT with its substrate connection at node 4. Parameters CJS, VJS, and MJS are found in the manner described for CJE, VJE, and MJE. Nevertheless, MJS ≈ 0.5. Simulations with discrete BJTs have CJS $= 0$.

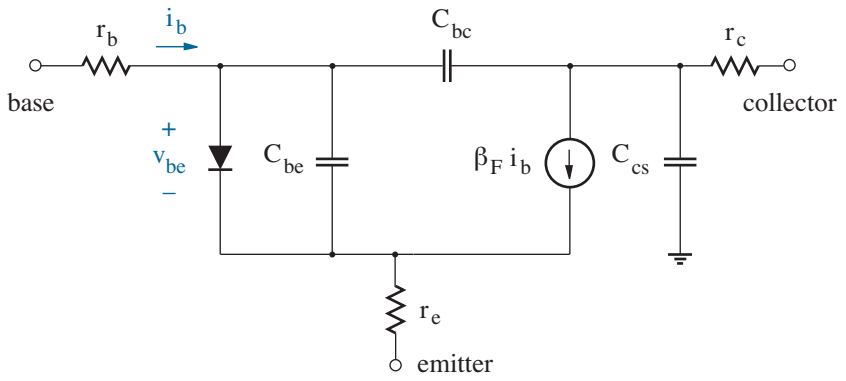


Figure 6.39: Large-signal forward-active BJT model with C_{be}, C_{bc}, and C_{cs}.

Integrated BJTs frequently exhibit the same areal base-emitter, base-collector, and collector-substrate capacitance values, and the corresponding junction areas are proportional to one another with the same scale factor. In this case, it is convenient to have a single .model statement that applies to a set of devices while providing individual AREA values. For example,

```
Q1      3       5       7       4       BJTN    AREA=5
Q4      3       2       6       4       BJTN    AREA=1
```

shows that Q_1 is five times as large as Q_4. The AREA factor directly scales BJT current parameters (IS, IKF, ISE, IKR, and ISC); however, it inversely scales BJT resistive parameters (RB, RC, and RE).

Apart from depletion capacitance effects, BJT dynamics imply changes in the forward- or reverse-injected electron (npn) or hole (pnp) charge that diffuses through the base. Electrical neutrality in the base is maintained by a similar distribution of opposite carriers that are supplied via base current. —In a crude npn scenario, hole "troops" are needed to watch nervously over the army of electrons that marches through p territory. The change of the neutralizing charge with respect to the change in the junction voltage that provokes it reflects a **diffusion capacitance**.

To quantify the second component of BJT junction-related capacitance, we assume that the stored base charge with forward injection (see Fig. 6.40) is proportional to the collector current. Specifically,

$$Q_f = \tau_f \, i_c \, . \tag{6.78}$$

In this expression, τ_f is the **forward transit time**. —Think of a freeway. If the transit time is very long over a short length of road, one can envision an electronic traffic jam as carriers struggle to get out of the congestion. Plenty of opposite carriers stand by to watch. Then subject to constant τ_f, we differentiate both sides of Eq, 6.78 to obtain

$$C_{be}\,(\text{diffusion}) = \frac{\partial Q_f}{\partial v_{be}} = \tau_f \frac{\partial i_c}{\partial v_{be}} \, . \tag{6.79}$$

Similarly, for reverse injection,

$$Q_r = \tau_r \, (-i_e) \, , \tag{6.80}$$

where τ_r is the **reverse transit time**. In turn,

$$C_{bc}\,(\text{diffusion}) = \frac{\partial Q_r}{\partial v_{bc}} = \tau_r \frac{\partial (-i_e)}{\partial v_{bc}} \, . \tag{6.81}$$

The diffusion capacitances of Eqs. 6.79 and 6.81 are inherently nonlinear.

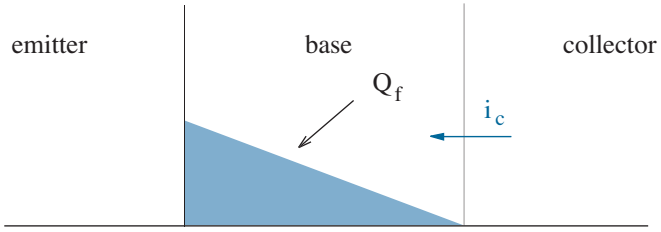


Figure 6.40: Excess charge-store Q_f for forward minority-carrier injection. As majority carriers, the Q_f charge constituents add to an already large concentration in the base, and they tend not to diffuse.

Now consider a somewhat different perspective.

In Eq. 6.78, the excess charge-store Q_f depends on collector current i_c. However, it is convenient to take the opposite point of view:

$$i_c = \frac{Q_f}{\tau_f} \,, \tag{6.82}$$

so that i_c is effectively controlled by Q_f. Similarly, the base current i_b can be viewed as a current governed by the rate of change for Q_f and the rate of hole/electron recombination in the base, which is proportional to Q_f. Specifically,

$$i_b = \frac{dQ_f}{dt} + \frac{Q_f}{\tau_{bf}} \,, \tag{6.83}$$

where τ_{bf} is the **effective lifetime** that applies to the excess charge-store. Note that the time derivatives vanish in the steady state, and

$$\frac{i_c}{i_b} = \frac{\tau_{bf}}{\tau_f} = \beta_F \,. \tag{6.84}$$

This useful result determines τ_f, given τ_{bf} and β_F.

Suppose a simple BJT circuit like that of Fig. 6.22 experiences a step increase in i_b so that the transistor progresses from cutoff to the forward active mode. Further assume that the eventual $i_c = \beta_F i_b$ is large so that the effects of diffusion capacitance dominate those of depletion capacitance. With $Q_f = 0$ at $t = 0$, the solution to Eq. 6.83 is

$$Q_f = i_b \, \tau_{bf} \left(1 - e^{-t/\tau_{bf}}\right) \,. \tag{6.85}$$

Then with the help of Eqs. 6.82 and 6.84,

$$i_c = \beta_F \, i_b \left(1 - e^{-t/\tau_{bf}}\right) \,. \tag{6.86}$$

The functional form of this expression suggests a means for measuring τ_{bf} and by extension τ_f. Similar theoretical developments apply to $\tau_{br} = \beta_R \tau_r$ when a BJT transitions from cutoff to the reverse active mode.

Equations 6.82 and 6.83 are **charge control relations** that support hand calculations of BJT transients. It would once have been appropriate to devote many pages of discussion to promote their proper application. We are now content to let SPICE do the work. Nevertheless, we can profit from the general observation that it takes time to build up or remove Q_f and that the time is limited by the base current. Things slow down even more during saturation, and this has led to the demise of BJT digital logic. The chapter Appendix explores some of the related issues.

Table 6.3 lists the dynamic BJT .model parameters.

Table 6.3: Dynamic BJT Parameters

Symbol	SPICE keyword	Parameter Name	Default value	Unit
C_{JE}	CJE	Zero-bias base-emitter depletion capacitance	0	F
ϕ_E	VJE	Base-emitter built-in potential	0.75	V
m_E	MJE	Base-emitter junction grading coefficient	0.33	V
C_{JC}	CJC	Zero-bias base-collector depletion capacitance	0	F
ϕ_C	VJC	Base-collector built-in potential	0.75	V
m_C	MJC	Base-collector junction grading coefficient	0.33	V
C_{JE}	CJE	Zero-bias collector-substrate depletion capacitance	0	F
ϕ_E	VJE	Collector-substrate built-in potential	0.75	V
m_E	MJE	Collector-substrate junction grading coefficient	0	V
τ_f	TF	Forward transit time	0	s
τ_r	TR	Reverse transit time	0	s

Appendix: BJT Switching Dynamics

This Appendix examines the dynamic behavior of a BJT switching circuit. While the physical discussion is somewhat involved, the design implications are relatively simple (and partly indicative of the demise of BJT logic).

The primitive **Resistor Transistor Logic** (RTL) circuit of Fig. 6.41a not only put a man on the moon, it is a practical test vehicle for BJT speed. The input-output transfer characteristic of Fig. 6.41b has three regimes as v_{in} increases: First, for $v_{in} < v_{be,on} \sim 0.7$ V, the BJT is effectively cutoff, $i_c \approx 0$, and $v_{out} \approx V^+ = 5$ V. Next, for v_{in} slightly greater than $v_{be,on}$, the BJT assumes the forward active operating mode, i_c rapidly increases, and v_{out} rapidly decreases. Finally, for v_{in} significantly greater than $v_{be,on}$, the BJT is in saturation, i_c has maximum value, and $v_{out} = v_{ce,sat} \sim 0.2$ V. Qualitatively, a LOW v_{in} produces a HIGH v_{out}, and a HIGH v_{in} produces a LOW v_{out}. This is the logical function of an inverter.

If the RTL inverter is to operate properly, a "HIGH" input voltage must be sufficient to place the BJT in saturation. For the base current, we have

$$i_b = \frac{V_{IH} - v_{be,on}}{R_b} , \tag{6.87}$$

where V_{IH} is the minimum HIGH-level input voltage that is to guarantee a LOW-level output voltage (see Fig. 1.18). And for the collector current,

$$i_c = \frac{V^+ - v_{ce,sat}}{R_c} . \tag{6.88}$$

The appropriate V_{IH} is at the edge of saturation with $i_c = \beta_F i_b$. Thus,

$$R_b = \beta_F \left(\frac{V_{IH} - v_{be,on}}{V^+ - v_{ce,sat}} \right) R_c . \tag{6.89}$$

For $V_{IH} = 1.2$ V with $\beta_F = 40$ and $R_C = 1$ kΩ, we find $R_b = 4.2$ kΩ.

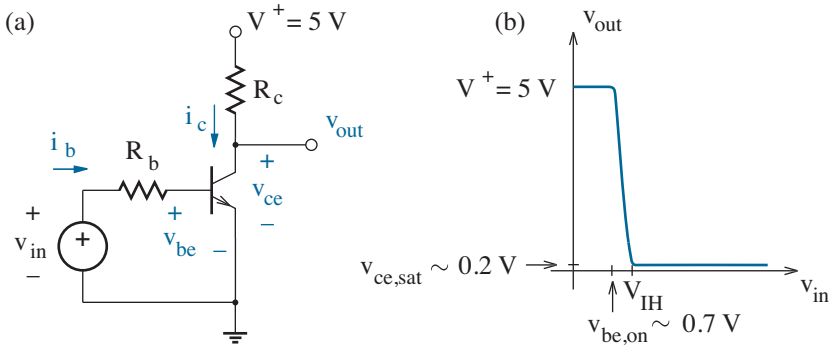


Figure 6.41: RTL inverter and associated transfer characteristic.

The requirement of a HIGH v_{in} lower than V^+ arises when the output of a particular inverter needs to "drive" the parallel-connected inputs of N similar gates. A HIGH state for the driver yields LOW states for the driven and a set of N identical base currents given by Eq. 6.87. So internal to the driver inverter (with a cutoff BJT and worst-case $v_{out} = V_{IH}$),

$$\frac{V^+ - V_{IH}}{R_c} = N\left(\frac{V_{IH} - v_{be,on}}{R_b}\right) . \tag{6.90}$$

Then with $V_{IH} = 1.2$ V and $R_b = 4.2$ kΩ, the inverter **fanout** is $N \approx 32$. The fanout increases with $(V^+ - V_{IH})$.

We now focus on the BJT collector current and five characteristic time intervals that relate its dynamic response to an input voltage that pulses from LOW to HIGH to LOW as shown in Fig. 6.42. The intervals are:

- The **delay time** ($t_d = t_1 - t_0$), during which the BJT base-emitter junction assumes forward bias with "turned-on" status;
- The **rise time** ($t_r = t_2 - t_1$), during which the BJT operates in the forward active mode with its collector current increasing toward the saturation limit;
- **A saturation period** ($t_2 < t < t_3$), during which the BJT operates in the saturation mode (while building up and maintaining forward-bias conditions for the base-collector junction);
- The **storage time** ($t_s = t_4 - t_3$), during which the BJT operates in the saturation mode (while breaking down forward-bias conditions for the base-collector junction);
- The **fall time** ($t_f = t_5 - t_4$), during which the BJT operates in the forward active mode with its collector current decreasing to zero.

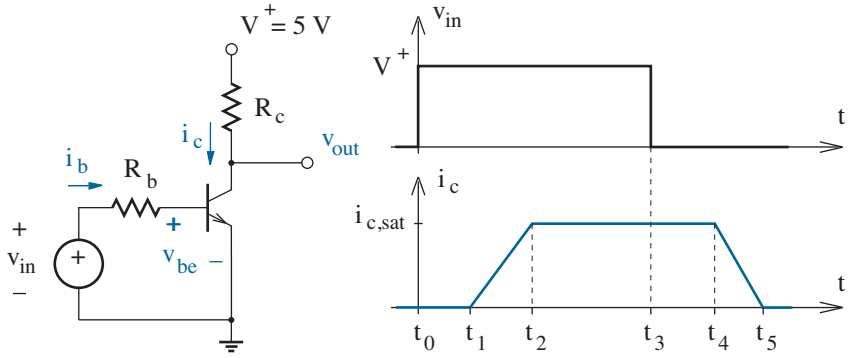


Figure 6.42: Qualitative dynamic response for an RTL inverter.

For the moment, we put aside the delay time, which is often short in comparison with the subsequent intervals.

The rise time typically reflects the need to establish diffusion charge within the BJT base in support of a particular collector current as described in Section 6.4. If the end-of–rise-time objective has $v_{ce} \approx v_{ce,sat} \sim 0.2$ V while just barely maintaining the forward active mode, the collector current increases according to the relation

$$i_c = \beta_F \, i_b \left(1 - e^{-t'/\tau_{bf}}\right) , \tag{6.91}$$

where $t' = t - t_1$, and τ_{bf} is an effective lifetime for the stored base charge. The latter relates to the forward transit time τ_f through the expression

$$\tau_{bf} = \beta_F \, \tau_f \, . \tag{6.92}$$

Section 6.4 derived these results in support of a simple measurement for τ_f, which is needed for SPICE. Figure 6.43 shows the i_c variation.

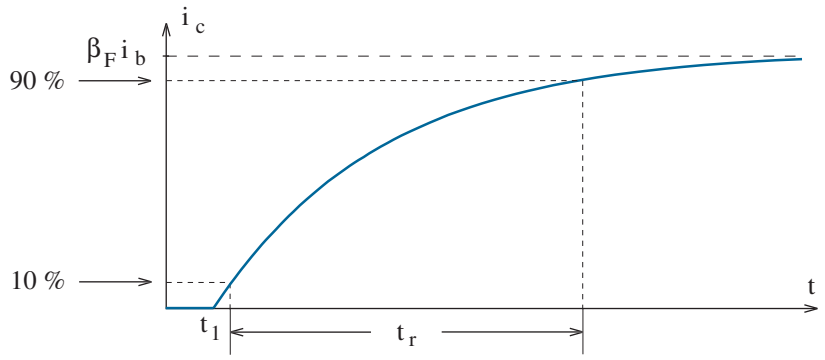


Figure 6.43: Transient BJT collector current, subject to $\beta_F i_b < i_{c,sat}$. The rise time is the interval between 10 % and 90 % of maximum value.

We define the rise interval as the time needed for the collector current to increase from 10 % to 90 % of maximum value. The 10-% time $t_{0.1}$ satisfies Eq. 6.91 with $i_c/\beta_F i_b = 0.1$. In turn,

$$t_{0.1} = \tau_{bf} \ln (10/9) \, . \tag{6.93}$$

Similarly, the 90-% time $t_{0.9}$ satisfies Eq. 6.91 with $i_c/\beta_F i_b = 0.9$, and

$$t_{0.9} = \tau_{bf} \ln (10) \, . \tag{6.94}$$

Thus,

$$t_r = t_{0.9} - t_{0.1} = \tau_{bf} \ln (9) = 2.2 \, \tau_{bf} \, . \tag{6.95}$$

For a BJT with $\beta_F = 40$ and $\tau_f = 0.5$ ns, $\tau_{bf} = 20$ ns and $t_r = 44$ ns.

Fortunately, rise time is significantly reduced by making the base current large with $\beta_F\, i_b \gg i_{c,sat}$. Equation 6.90 still applies for small values of i_c. However, i_c quickly reaches the saturation limit so that the steep portion of the curve is truncated. Figure 6.44 shows the rise-time effect.

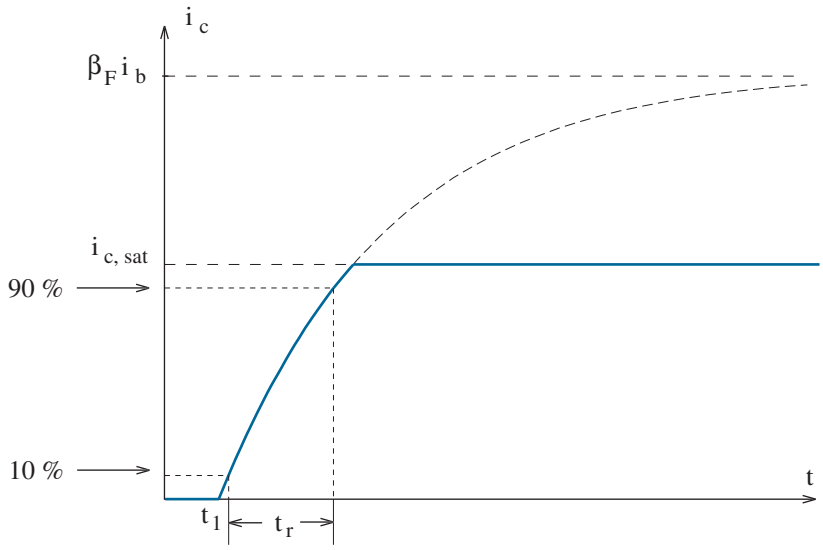


Figure 6.44: Transient BJT collector current, subject to $\beta_F i_b \gg i_{c,sat}$.

To compute the new rise time, apply Eq. 6.90 at the saturation limit. Specifically,

$$\frac{i_{c,sat}}{\beta_F\, i_b} = \frac{\beta_{\text{forced}}}{\beta_F} = 1 - e^{-t'/\tau_{bf}}\ . \tag{6.96}$$

Then solve for t' and grind through some algebra to obtain

$$t_r = t_{0.9} - t_{0.1} = \tau_{bf} \ln \left(\frac{1 - 0.1\, \dfrac{\beta_{\text{forced}}}{\beta_F}}{1 - 0.9\, \dfrac{\beta_{\text{forced}}}{\beta_F}} \right)\ . \tag{6.97}$$

If we assume $\beta_{\text{forced}}/\beta_F \ll 1$ (as for a BJT in strong saturation), a Taylor series expansion of Eq. 6.97 yields

$$t_r \approx 0.8\, \tau_{bf}\, \frac{\beta_{\text{forced}}}{\beta_F}\ . \tag{6.98}$$

With $R_c = 1$ kΩ and $R_b = 4.2$ kΩ, the inverter design of Fig. 6.41 has $i_c = 4.8$ mA and $i_b = 1.02$ mA when $v_{in} = 5$ V. Then with $\beta_{\text{forced}} = i_c/i_b = 4.71$, $\beta_F = 40$, and $\tau_{bf} = 20$ ns, $t_r \approx 1.9$ ns, a significant rise-time improvement. We can do better by decreasing R_b, but the fanout is also reduced.

Our desire for an accelerated rise time conditions the saturation period. The forward base-emitter junction voltage remains essentially constant at 0.7 V throughout (even though the electron injection from the emitter to the base may substantially increase). Meanwhile, the now *forward* base-collector junction voltage increases from a small value (zero) to a larger value (about 0.7 V), and electrons are reverse injected from the collector to the base. As shown in Fig. 6.45, the positive stored base charge that reflects the extent of forward electron injection (Q_f) and the positive stored base charge that reflects the extent of reverse electron injection (Q_r) combine to yield a "surplus" of stored saturation charge (Q_s) in addition to the stored forward-injection charge at the onset of saturation ($Q_{fo} = i_{c,sat}\,\tau_f$).

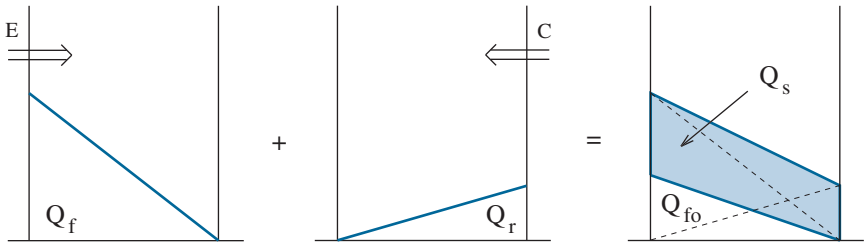


Figure 6.45: Injected base-charge stores during the saturation period. Forward (Q_f) and reverse (Q_r) components combine to yield a "surplus" saturation component (Q_s) that adds to the injected forward component at the onset of saturation— $Q_{fo} = i_{c,sat}\,\tau_f$.

In what follows, we recall the overdrive base current:

$$i_{bs} = i_b - \frac{i_c}{\beta_F}\,. \tag{6.99}$$

This excess current is effectively controlled by the rate of change in the stored saturation charge and the surplus hole/electron recombination rate, which is proportional to the stored saturation charge. Thus,

$$i_{bs} = \frac{dQ_s}{dt} + \frac{Q_s}{\tau_s}\,, \tag{6.100}$$

where τ_s is the effective lifetime of the stored saturation charge given by[3]

$$\tau_s = \frac{\alpha_F(\tau_f + \alpha_R\tau_r)}{1 - \alpha_F\alpha_R}\,. \tag{6.101}$$

Parameter τ_r is the **reverse transit time**. Subject to the steady state, when time derivatives vanish, Eq. 6.100 implies a maximum $Q_s = i_{bs}\tau_s$.

[3]P. E. Gray, D. DeWitt, A. R. Boothroyd, and J. F. Gibbons, *Physical Electronics and Circuit Models of Transistors*, Semiconductor Electronics Education Committee, Vol. 2, (Wiley, New York, 1964), pp. 226-229.

Non-zero Q_s saturation charge sets the stage for a non-zero storage time. As indicated in Fig. 6.42, the storage time interval commences at $t = t_3$ when V_b is returned to zero. Despite this change, the base-emitter junction voltage is *continuous*. If we use a prime ($\prime$) to denote circuit properties that apply during transitions to the BJT off-state, the effective base current is

$$i_b{}' = \frac{0 - 0.7}{R_b} < 0\,. \tag{6.102}$$

Meanwhile, the "surplus" of saturation charge (Q_s) is gradually removed, and the BJT approaches its status at the onset of saturation with $i_c = i_{c,sat}$. The storage time interval ends when $Q_s = 0$.

Like the preceding saturation period, the storage time interval features a charge control relation of the form

$$i_{bs} = \frac{dQ_s}{dt'} + \frac{Q_s}{\tau_s}\,, \tag{6.103}$$

where $t' = t - t_3$. In contrast, the overdrive base current is now

$$i_{bs}{}' = i_b{}' - \frac{i_{c,sat}}{\beta_F}\,. \tag{6.104}$$

The solution to Eq. 6.103 is

$$Q_s = K\, e^{-t'/\tau_s} + i_{bs}{}'\tau_s\,. \tag{6.105}$$

However, $Q_s = Q_{so}$ at $t' = 0$. Thus, $K = Q_{so} - i_{bs}{}'\tau_s$ and

$$Q_s = Q_{so}\, e^{-t'/\tau_s} + i_{bs}{}'\tau_s \left(1 - e^{-t'/\tau_s}\right)\,. \tag{6.106}$$

To determine the storage time, we set $Q_s = 0$ at $t' = t_s$ to obtain

$$t_s = \tau_s \ln\left(1 - \frac{Q_{so}}{i_{bs}{}'\tau_s}\right)\,. \tag{6.107}$$

A less general result features a saturation period that is sufficiently long to ensure "steady-state" $Q_{so} = i_{bs}\tau_s$. In this case,

$$t_s = \tau_s \ln\left(1 - \frac{i_{bs}}{i_{bs}{}'}\right)\,. \tag{6.108}$$

For the BJT at hand, let $\beta_R = 1$ and $\tau_R = 5$ ns —one typically has $\tau_R \gg \tau_f$. Then $\alpha_F = 40/41$, $\alpha_R = 1/2$, and $\tau_s = 5.71$ ns. The HIGH-input overdrive base current for the inverter design of Fig. 6.41 is $i_{bs} = 1.07$ mA, and the LOW-input overdrive base current is $i_{bs}{}' = -0.287$ mA. Thus, $t_s = 8.9$ ns under the worst-case conditions of Eq. 6.108. All this suggests a trade-off: *Improvements in rise time that reflect increasingly strong BJT saturation are made at the expense of increased storage time.*

At last, we arrive at the fall-time interval for which physical changes are the opposite of those observed during the rise-time interval. The (negative) change in the stored base charge is given by

$$\Delta Q_f = \tau_f \, \Delta i_c \, , \tag{6.109}$$

with $Q_f = \tau_f \, i_c$ subject to a charge control relation of the form

$$i_b{}' = \frac{dQ_f}{dt'} + \frac{Q_f}{\tau_{bf}} \, . \tag{6.110}$$

Here, τ_{bf} is the effective lifetime of the stored base charge, and $t' = t - t_4$. The solution is

$$Q_f = K \, e^{-t'/\tau_{bf}} + i_b{}' \, \tau_{bf} \, . \tag{6.111}$$

However, $Q_f = Q_{fo} = i_{c,sat} \, \tau_f$ at $t' = 0$. Thus, $K = Q_{fo} - i_b{}' \tau_{bf}$, and

$$Q_f = Q_{fo} \, e^{-t'/\tau_{bf}} + i_b{}' \tau_{bf} \left(1 - e^{-t'/\tau_{bf}} \right) \, . \tag{6.112}$$

In turn,

$$i_c = \frac{Q_{fo}}{\tau_f} \, e^{-t'/\tau_{bf}} + \beta_F i_b{}' \left(1 - e^{-t'/\tau_{bf}} \right) \, , \tag{6.113}$$

(since $\tau_{bf}/\tau_f = \beta_F$). Equations 6.112 and 6.113 are only valid for $Q_f > 0$ and $i_c > 0$, respectively.

The base-charge contribution to the fall time is defined as the time needed for Q_f to decrease from 90 % to 10 % of the maximum Q_{fo} value. The 90-% time $t_{0.9}$ satisfies Eq. 6.112 with $Q_f = 0.9 \, Q_{fo}$. Specifically,

$$t_{0.9} = \tau_{bf} \ln \left(\frac{Q_{fo} - i_b{}' \tau_{bf}}{0.9 \, Q_{fo} - i_b{}' \tau_{bf}} \right) \, . \tag{6.114}$$

Similarly, the 10-% time $t_{0.1}$ satisfies Eq. 6.112 with $Q_f = 0.1 \, Q_{fo}$, and

$$t_{0.1} = \tau_{bf} \ln \left(\frac{Q_{fo} - i_b{}' \tau_{bf}}{0.1 \, Q_{fo} - i_b{}' \tau_{bf}} \right) \, . \tag{6.115}$$

Thus,

$$t_f = t_{0.1} - t_{0.9} = \tau_{bf} \ln \left(\frac{0.9 \, Q_{fo} - i_b{}' \tau_{bf}}{0.1 \, Q_{fo} - i_b{}' \tau_{bf}} \right) \, . \tag{6.116}$$

In the limit of small $|i_b{}'|$ (relative to $Q_{fo}/\tau_{bf} = i_{c,sat}/\beta_F$),

$$t_f \approx \tau_{bf} \ln (9) = 2.2 \, \tau_{bf} \, . \tag{6.117}$$

This result is the same as the leisurely rise time of Eq. 6.95.

Our inverter has $Q_{fo} = 2.4 \times 10^{-12}$ C and ${i_b}'\tau_{bf} = -3.34 \times 10^{-12}$ C. Thus, Eq. 6.116 gives $t_f = 7.3$ ns. Any improvement requires a smaller R_b. Nevertheless, there is also the possibility of making v_{in} negative so that ${i_b}'$ is very large and negative. For example, if $|Q_{fo}/{i_b}'\tau_{bf}| \ll 1$, a Taylor series expansion of Eq. 6.116 yields

$$t_f \approx 0.8\,\tau_{bf}\,\left|\frac{Q_{fo}}{{i_b}'\tau_{bf}}\right| . \tag{6.118}$$

But with $Q_{fo} = i_{c,sat}\,\tau_f$,

$$t_f \approx 0.8\,\tau_{bf}\,\left|\frac{{\beta_{\text{forced}}}'}{\beta_F}\right| . \tag{6.119}$$

The form of this result is similar to Eq. 6.98.

The delay, rise, and fall times are all influenced by the need to supply or remove base charge as the widths of the depletion regions associated with the base-emitter or base-collector pn junctions change with applied voltage. As noted in Section 6.4, the base-emitter junction capacitance has the form

$$C_{je} = \frac{\Delta Q_{je}}{\Delta v_{be}} = C_{je}(0)\,[1 - v_{be}/\phi_e]^{-m_e} , \tag{6.120}$$

where Q_{je} is the positive-side junction depletion charge, $C_{je}(0)$ is the zero-bias junction capacitance, v_{be} and ϕ_e are the applied and built-in junction voltages, respectively, and m_e is the base-emitter junction grading factor. Thus for a particular change in v_{be} over the time interval $[t_a, t_b]$, the change in the requisite depletion charge is given by

$$\Delta Q_{je} = \int_{v_{be}(t_a)}^{v_{be}(t_b)} C_{je}\,dv_{be} , \tag{6.121}$$

or

$$\Delta Q_{je} = \frac{\phi_e C_{je}(0)}{1 - m_e}\left\{\left[1 - \frac{v_{be}(t_a)}{\phi_e}\right]^{1-m_e} - \left[1 - \frac{v_{be}(t_b)}{\phi_e}\right]^{1-m_e}\right\} . \tag{6.122}$$

Similarly, for the base-collector junction,

$$\Delta Q_{jc} = \frac{\phi_c C_{jc}(0)}{1 - m_c}\left\{\left[1 - \frac{v_{bc}(t_a)}{\phi_c}\right]^{1-m_c} - \left[1 - \frac{v_{bc}(t_b)}{\phi_c}\right]^{1-m_c}\right\} . \tag{6.123}$$

The length of the time interval is

$$t_b - t_a = \frac{\Delta Q_{je} + \Delta Q_{jc}}{|< i_b >|} , \tag{6.124}$$

where $< i_b >$ is an average base current throughout (see Problem 6.77). These calculations are best reserved for SPICE.

Digital BJT circuits that came after RTL offered small-scale integration while confronting speed-limiting saturation effects in three different ways.

- Use an additional BJT to facilitate base-charge removal. Figure 6.46 shows the replacement of the RTL base resistor with transistor Q_2 to form part of a **Transistor Transistor Logic** (TTL) inverter. When the input is HIGH, Q_2 supplies base current to put Q_1 in quick saturation, and the rare reverse-active configuration ensures minimal input current demand so as to promote large fanout. When the input changes to LOW, Q_2 acts like a vacuum cleaner to suck out the Q_1 base charge, thereby shortening the storage and fall times.

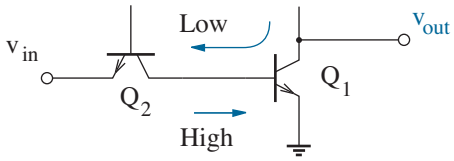


Figure 6.46: Portion of a TTL inverter.

- Use a parallel-connected metal-semiconductor Schottky diode to clamp the BJT base-collector voltage, thereby preventing the transistor from entering the saturation mode by absorbing the forward bias current. Figure 6.47 shows the ease of circuit integration.

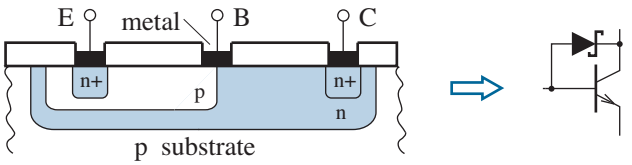


Figure 6.47: Integrated Scholttky/BJT pair.

- Avoid saturation completely. This is the basis for **Emitter Coupled Logic** (ECL), a circuit format in which one of two transistors is either "on" or "off" without saturation. Quick transitions are ensured by the inclusion of a current source that demands satisfaction (Fig. 6.48).

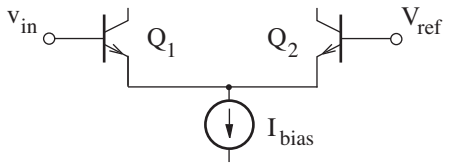


Figure 6.48: Emitter coupled pair.

All of these improvements have given way to low-power CMOS.

Problems

Section 6.1

6.1 The quadrants in the diagram of Fig. P6.1 have particular pairs of BJT pn-junction bias conditions. Identify the corresponding modes of operation.

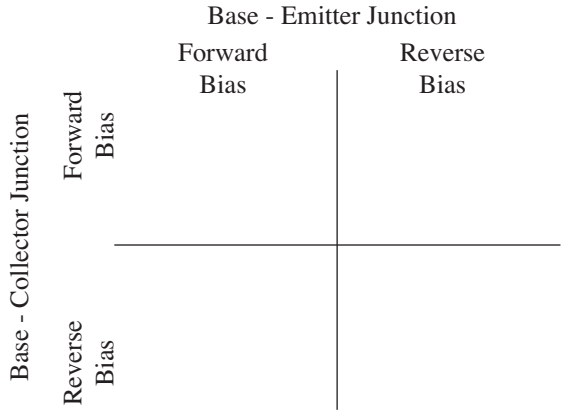


Figure P6.1

6.2 For each BJT configuration in Fig. P6.2, specify whether the device is in the forward active, reverse active, saturation, or cutoff mode of operation.

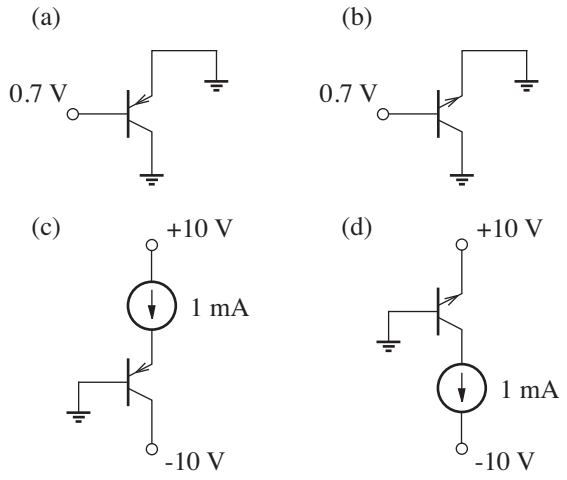


Figure P6.2

6.3 Consider an npn BJT with emitter doping $N_d = 2 \times 10^{18}$ cm^{-3}.

(a) Let the base/emitter form a **homojunction**—the same semiconductor material is on each side. Determine the base doping that is consistent with an emitter defect of 0.001.

(b) Let the base-emitter form a **heterojunction**—different semiconductor materials apply on the p and n sides. The emitter is Si with bandgap $E_{g1} = 1.12$ eV and ${n_{i1}}^2 = K_1 T^3 \exp(-E_{g1}/kT)$. The base is a Si-Ge alloy (20% Ge) with bandgap $E_{g2} = 0.92$ eV and ${n_{i2}}^2 = K_2 T^3 \exp(-E_{g2}/kT)$. Assume $K_1 \approx K_2$, and $kT = 0.0259$ eV at room temperature. Determine the base doping that is consistent with an emitter defect of 0.001.

Note: The relatively large part-b base doping allows a BJT design with low parasitic base resistance.

6.4 A particular npn BJT has neutral base width w. The electron current density crossing the base is

$$J_e = qD_e \nabla n \, ,$$

where D_e is the electron diffusion coefficient, and ∇n is the electron concentration gradient. The electron recombination current density in the base is

$$J_r = \frac{q}{\tau_e} \int n \, dx \, .$$

Here, the integration over the neutral base represents the total number of electrons per unit area therein, and τ_e is the average electron lifetime.

In the forward active mode, the base-side electron concentration at the edge of the base-emitter depletion region is $n(x_1)$ and the base-side electron concentration at the edge of the base-collector depletion region is $n(x_2) \approx 0$. The electron profile in the base is nearly linear if the level of recombination is small.

(a) Derive an expression for the base defect δ_b.

(b) Let $\sqrt{D_e \tau_e} = 1$ μm. Determine the base width that is consistent with a base defect of 0.01.

6.5 Let δ_e and δ_b denote emitter and base defects, respectively. Show that for $\delta_e \ll 1$ and $\delta_b \ll 1$,

$$\beta_F{}^{-1} \approx \delta_e + \delta_b \, .$$

(Thus, β_F is dominated by the larger of the defects —usually δ_b.)

6.6 Complete the Table of BJT data that follows. Enter a cross (X) for any undetermined information. Caution: There may be multiple solutions.

i_c (mA)	i_b (μA)	i_e (mA)	β_F	β_R	v_{ce} (V)
3	20				+5
		-5.05	100		-5
	-100		200	5	+5
-10				5	-5
5		5.05		2.5	+5
	-10		75	1	-5

6.7 Complete the Table of BJT data that follows. Enter a cross (X) for any undetermined information. Caution: There may be multiple solutions.

i_c (mA)	i_b (μA)	i_e (mA)	β_F	β_R	v_{ce} (V)
-2			100	4	-1.0
	-50		100		-2.4
4	200			5	+0.2
10 mA		10.1		1	+0.7
	- 150	-6.15			-0.2
	-20		80	2	+1.0

6.8 Both circuits in Fig. P6.8 feature a BJT with $\beta_F = 100$, $\beta_R = 4$, and $v_{ce} = -4$ V. Find each I_x.

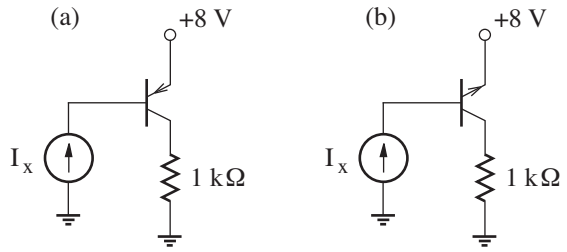


Figure P6.8

6.9 Both circuits in Fig. P6.9 feature a BJT with $\beta_F = 100$ and $\beta_R = 4$. Complete the designs so that v_{ce} has the value indicated.

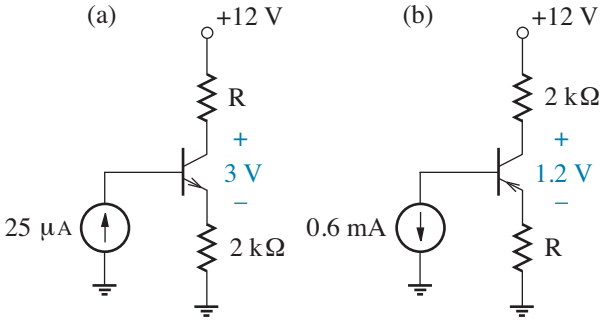


Figure P6.9

6.10 The BJT in Fig. P6.10 has $\beta_F = 100$, $\beta_R = 4$, and $|v_{ce,sat}| = 0.2$ V. Determine the ranges of I_x that correspond to BJT cutoff and saturation.

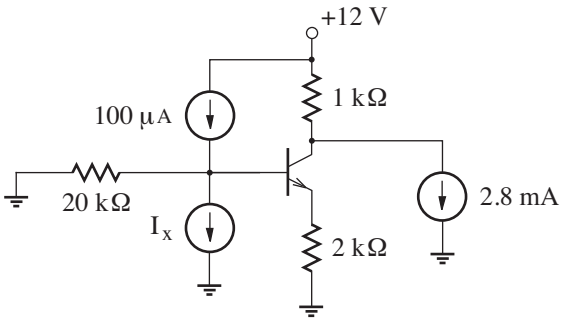


Figure P6.10

6.11 The BJTs in the circuit of Fig. P6.11 have $\beta_F = 100$, $\beta_R = 5$ and $\beta_F = 50$, $\beta_R = 2$ for transistors Q_1 and Q_2, respectively. The node voltages satisfy $v_4 > v_3 > v_2 > v_1 > 0$.

(a) Determine i_2/i_1.

(b) Repeat part a, but interchange the Q_2 emitter and collector terminals.

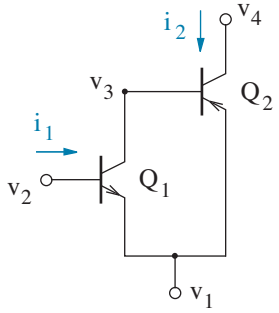


Figure P6.11

6.12 The BJTs In the circuit of Fig. P6.12 have $\beta_F = 120$, $\beta_R = 4$ and $\beta_F = 80$, $\beta_R = 4$ for transistors Q_1 and Q_2, respectively. The node voltages satisfy $v_5 > v_4 > v_3 > v_2 > v_1 > 0$.

(a) Determine i_2/i_1.

(b) Repeat part a, but interchange the Q_2 emitter and collector terminals.

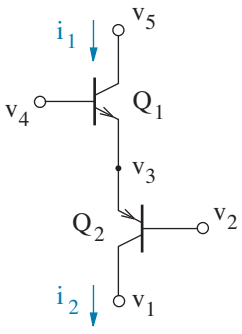


Figure P6.12

6.13 The circuit of Fig. P6.12 features $v_5 = 3$ V, and both BJTs are in saturation. Find consistent values for the other node voltages.

6.14 Consider the curve-tracer display of Fig. P6.14.

(a) Is the BJT npn or pnp?

(b) Estimate β_F.

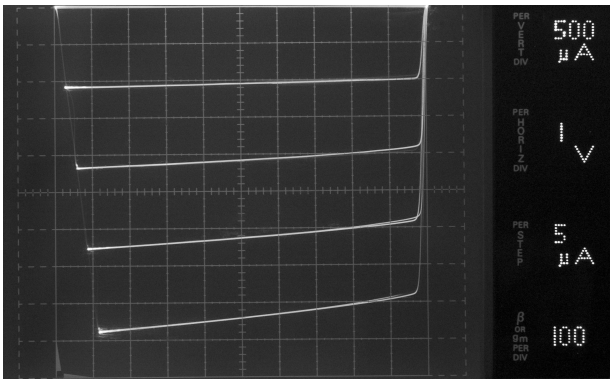


Figure P6.14

6.15 Consider the curve-tracer display of Fig. P6.15.

(a) Is the BJT npn or pnp?

(b) Estimate β_F.

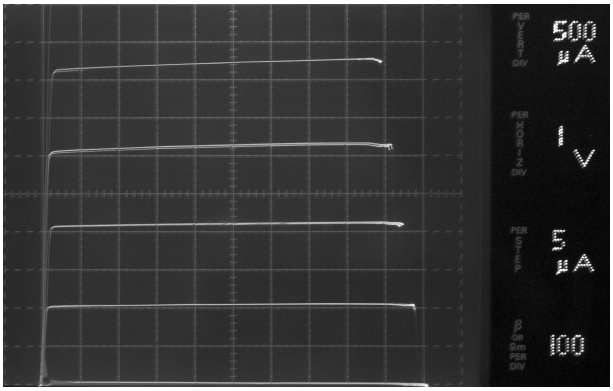


Figure P6.15

Section 6.2

6.16 Verify Eqs. 6.29 and 6.30.

6.17 Starting with the Ebers-Moll relations given in Eqs. 6.19a and 6.19b, show that the common-base operation implies the linear two-port equations

$$\begin{aligned} i_c &= y_{11}\, v_{cb} + y_{12}\, v_{eb} \\ -i_e &= y_{21}\, v_{cb} + y_{22}\, v_{eb} \end{aligned}$$

when v_{eb} and v_{cb} are very small (see Fig. P6.17). Specify values for each of the y coefficients, then show the consequences of $y_{12} = y_{21}$.

Note: A proof of the two-port reciprocity condition ($y_{12} = y_{21}$) and related restrictions is available in most texts on circuit theory.

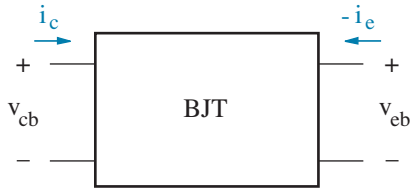


Figure P6.17

6.18 Use the Ebers-Moll relations to determine the current-voltage characteristic (i vs. v) for the npn "diode" configuration shown in Fig. P6.18.

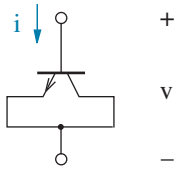


Figure P6.18

6.19 In terms of v_{be}, v_{bc}, β_F, and β_R, write expressions for the terminal currents of a pnp BJT that operates in the saturation mode.

6.20 A BJT has the following current-voltage data. Determine parameters I_s, I_{KF}, β_F, $C_2 I_s$, and n_{EL}.

$\frac{qv_{be}}{kT}$	i_c (A)	i_b (A)
16	4.44×10^{-8}	3.35×10^{-9}
18	3.28×10^{-7}	1.08×10^{-8}
20	2.43×10^{-6}	4.22×10^{-8}
22	1.79×10^{-5}	2.09×10^{-7}
24	1.32×10^{-4}	1.27×10^{-6}
26	9.74×10^{-4}	8.60×10^{-6}
28	7.14×10^{-3}	6.15×10^{-5}
30	5.17×10^{-2}	4.49×10^{-4}
32	3.61×10^{-1}	3.30×10^{-3}
34	2.33	2.43×10^{-2}

6.21 A BJT has the following β_F vs. temperature (°C) pairings: 102, -40; 114, -25, 126, -10; 140, 5; 153, 20; 168, 35; 183, 50; 198, 65. Determine a value for parameter X_{TB}.

6.22 A particular BJT features $I_s = 2 \times 10^{-16}$ A, $I_{KF} = 30$ mA, $\beta_F = 160$, $C_2 I_s = 0.2$ pA, $n_{EL} = 1.8$. Plot the actual β_F vs. i_c for $i_c > 1$ μA.

6.23 The circuit of Fig. P6.23 operates with $\beta_F = 200$ and $v_{out} = 4$ V at room temperature (300 K). Determine the temperature at which the BJT enters saturation with $v_{ce} = 0.2$ V. Assume $X_{TB} = 1.4$.

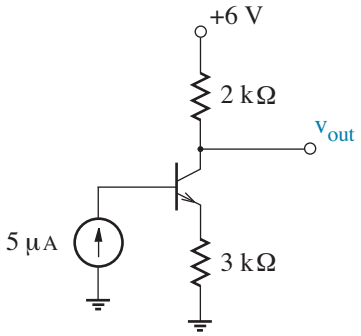


Figure P6.23

6.24 Show that the circuits of Fig. P6.24 operate in saturation by applying the tests of Eqs. 6.42, 6.43, and 6.44. Assume $\beta_F = 100$.

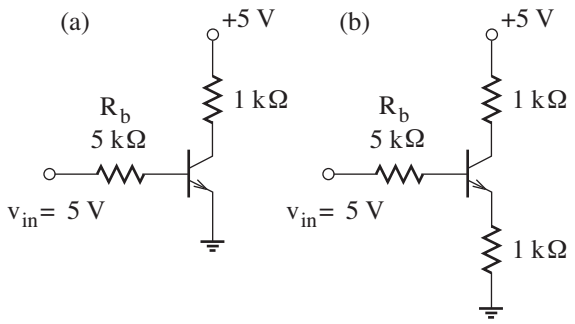


Figure P6.24

6.25 Determine the input voltage v_{in} that produces $\beta_{\text{forced}} = 20$ in the BJT circuits of Fig. P6.24. Assume $\beta_F = 100$.

6.26 Determine the base resistor R_b that produces a 50-μA overdrive base current in the BJT circuits of Fig. P6.24 when $v_{in} = 5$ V. Assume $\beta_F = 100$.

6.27 A BJT operates with $\beta_F = 180$, $\beta_R = 10$, and $\beta_{\text{forced}} = 20$.

(a) Determine $v_{ce,sat}$ for $r_c = 0$.

(b) Determine the value of r_c that makes $v_{ce,sat} = 0.2$ V if $i_b = 0.2$ mA.

6.28 A BJT exhibits the following data subject to v_{ce} at the edge of saturation. Find parameter r_c.

$i_c(mA)$	$v_{ce,sat}$ (V)
10	0.070
20	0.113
30	0.156
40	0.199
50	0.242

Section 6.3

dc BJT Circuits

6.29 Show that the use of "non-conventional" circuit variables in the analysis of the circuit of Fig. 6.23b leads to the results of Eqs. 6.58 and 6.59.

6.30 The BJT in Fig. P6.30 has $\beta_F = 100$.

(a) Determine v_{out} when $R = 470$ kΩ.

(b) How does the result of part a change when β_F increases to 125?

(c) Let v_{out} have the value of part a, when an added 1-kΩ resistor is in series with the BJT emitter. Determine R.

(d) How does the v_{out} of part c change when β_F increases to 125?

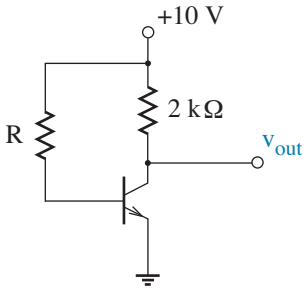


Figure P6.30

6.31 The BJT in Fig. P6.31 has $\beta_F = 100$.

(a) Determine v_{out} when $R = 400$ kΩ.

(b) How does the result of part a change when β_F decreases to 80?

(c) Let v_{out} have the value of part a, when an added 10-kΩ resistor connects from node x to ground. Determine R.

(d) How does the v_{out} of part c change when β_F decreases to 80?

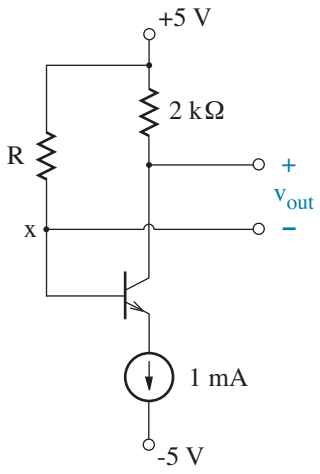


Figure P6.31

6.32 The circuit of Fig. P6.32 is to control a relay. Assume $\beta_F = 80$ and neglect relay-coil resistance.

(a) Complete the design so that $i = 20$ mA when the switch is closed.

(b) Determine i when the switch is open.

(c) Explain the function of the diode.

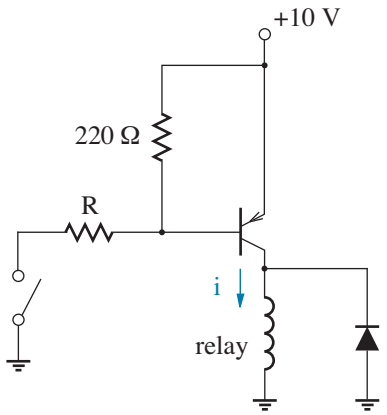


Figure P6.32

6.33 The BJT shown in Fig. P6.33 has $\beta_F = 50$.

(a) Complete the design so that $v_{out} = 0$ V.

(b) Find the minimum R consistent with circuit operation in the forward active mode.

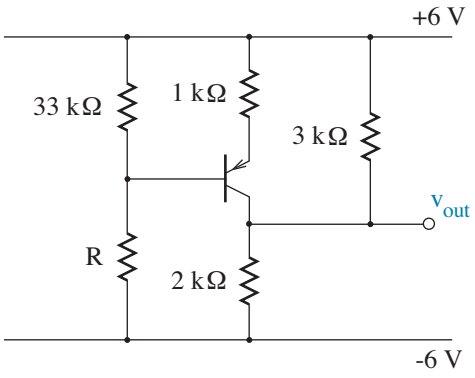


Figure P6.33

6.34 The circuit of Fig. P6.34 features a BJT with $\beta_F = 100$ and a Zener diode with 5-V breakdown.

(a) Determine v_{out}.

(b) Determine the total dissipated power.

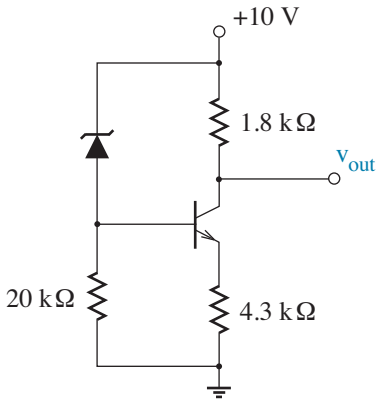


Figure P6.34

6.35 The circuit of Fig. P6.35 is intended to function as a voltage regulator. The unregulated input can assume any v_{in} between 7.5 and 8.5 V, and the load resistance can be any R_L between 75 Ω and 125 Ω. The BJT has $\beta_F = 100$. The Zener diode has 6.8-V breakdown if its reverse current exceeds 0.1 mA.

(a) Determine the regulated output voltage subject to proper design.

(b) Find R so that the circuit functions as desired for all possible input and load conditions.

(c) The Zener diode has 30-Ω series resistance. Given the R value of part b, find the expected worst-case variation of v_{out}.

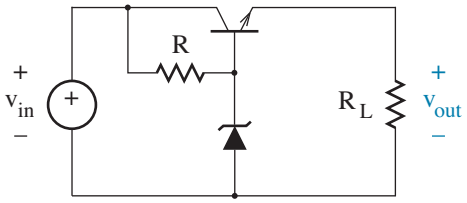


Figure P6.35

6.36 The BJTs shown in Fig. P6.36 have $\beta_F = 100$. Complete the design for $v_1 = 10$ V and $v_2 = 5$ V.

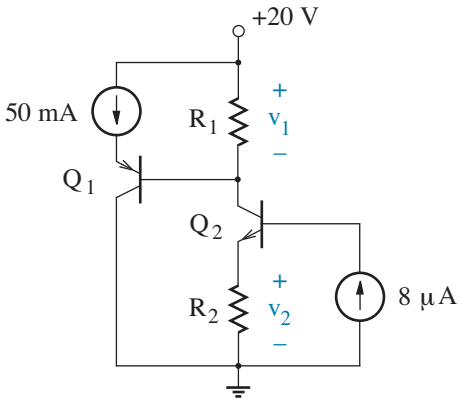


Figure P6.36

6.37 The BJTs shown in Fig. P6.37 have $\beta_F = 100$. Complete the design so that $i_{c2} = 2$ mA subject to $R_1 \parallel R_2 = 10$ kΩ.

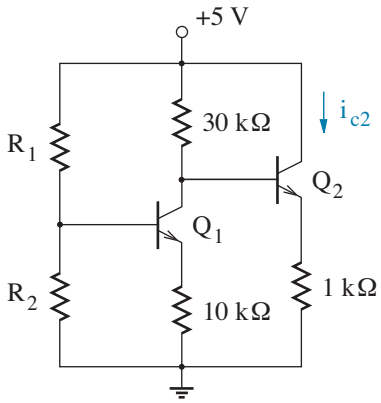


Figure P6.37

6.38 The BJTs shown in Fig. P6.38 have $\beta_F = 100$. Complete the design so that node voltages v_x and v_y are 6 V and 4 V, respectively.

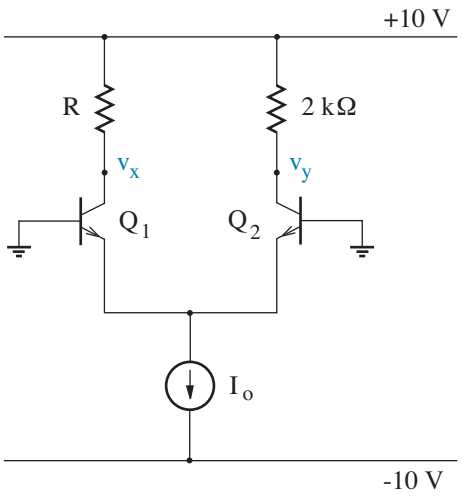


Figure P6.38

6.39 The BJTs shown in Fig. P6.39 have $\beta_F = 100$.

(a) Complete the design so that $i_x = 10$ mA and $v_2 = 5$ V.

(b) Determine the changes in i_x and v_2 when the power supply reduces to 8 V.

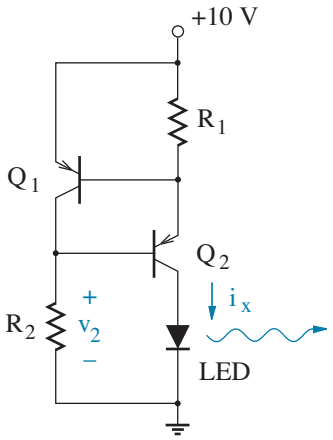


Figure P6.39

6.40 The circuit of Fig. P6.40 with identical BJTs is to produce 10 mA of LED current when $v_{in} = 3$ V despite the substantial 5-MΩ Thevenin resistance. Estimate the minimum acceptable β_F specification.

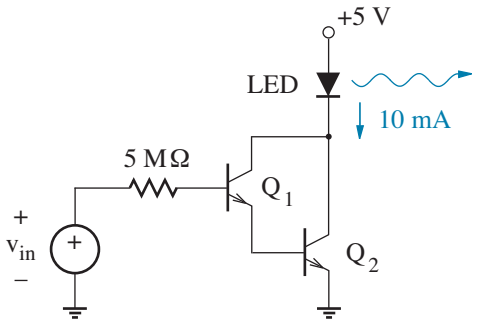


Figure P6.40

6.41 The circuit in Fig. P6.41 has a complementary Darlington pair for which $\beta_{F1} = 100$ and $\beta_{F2} = 50$. Complete the design so that $v_{out} = 2$ V.

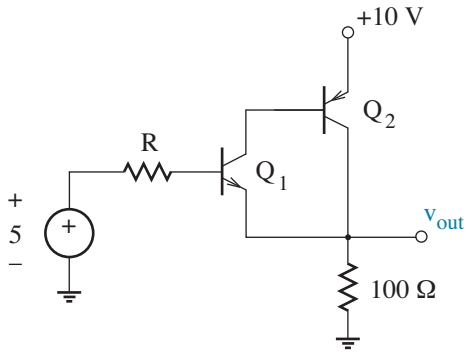


Figure P6.41

Power Amplification

6.42 The BJT shown in the emitter follower circuit of Fig. P6.42 has $\beta_F = 100$.

(a) Determine the current gain (i_2/i_1).

(b) Determine the power gain (power dissipated in R_2 / power from v_{in}).

(c) Determine the range of v_{in} for which the BJT operates in the forward active mode.

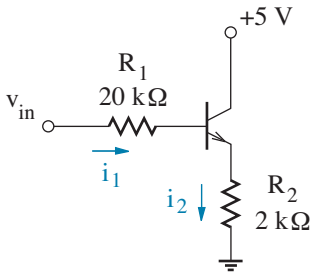


Figure P6.42

6.43 Consider the BJT push-pull power amplifier of Fig. 6.30 with $R_t = 1$ kΩ, $R_L = 100$ Ω, ±10-V power supplies, $\beta_{F1} = 120$ (npn), and $\beta_{F2} = 60$ (pnp). Sketch the collector currents vs. v_{in} over the range -12 V $\leq v_{in} \leq +12$ V.

6.44 Consider the BJT push-pull power amplifier of Fig. 6.32 with $R_t = 1$ kΩ, $R_L = 100$ Ω, ±10-V power supplies, $\beta_{F1} = 120$ (npn), and $\beta_{F2} = 60$ (pnp).

(a) Sketch v_{out} vs. v_{in} for -12 V $\leq v_{in} \leq +12$ V when $R = 80$ kΩ.

(b) Find the R value that promotes linear operation over the range -8 V $\leq v_{in} \leq +8$ V.

6.45 Consider the BJT push-pull power amplifier of Fig. 6.32 with $R_t = 1$ kΩ, $R_L = 100$ Ω, ±10-V power supplies, $\beta_{F1} = 80$ (npn), and $\beta_{F2} = 80$ (pnp). Let $R = 10$ kΩ.

(a) Suppose the diodes and the base-emitter pn junctions have the same saturation currents. Determine i_{c1} when $v_{out} = v_{in} = 0$.

(b) Repeat part a, but let the diodes have saturation currents that are ten times greater than those for the base-emitter pn junctions.

(c) Discuss the implications of the preceding results when applied to the design of the BJT push-pull power amplifier in Fig. 6.33.

6.46 Consider the BJT push-pull power amplifier of Fig. 6.32 with $R_t = 1$ kΩ, $R_L = 8$ Ω, ±12-V power supplies, $\beta_{F1} = \beta_{F2} = 180$ (pnp), and $R = 5$ kΩ. Determine R' so that the short-circuit load current is limited to 2 A.

6.47 Consider the BJT push-pull power amplifier of Fig. 6.30 with $R_t = 1$ kΩ, $R_L = 100$ Ω, ±10-V power supplies, $\beta_{F1} = 80$ (npn), and $\beta_{F2} = 80$ (pnp).

(a) The amplifier **power efficiency** is defined as the ratio of the time-averaged power dissipated in the load to the time-averaged power provided by the dc supplies. Find the power efficiency for a sinusoidal input signal with 5-V amplitude.

(b) Determine the maximum power efficiency and the input conditions that promote it.

Saturation Circuits

6.48 The BJT shown in Fig. P6.48 has $\beta_F = 80$. Determine the maximum value for i_x (a combined current from other circuits) such that v_{out} is LOW at 0.2 V when v_{in} is HIGH at +5 V.

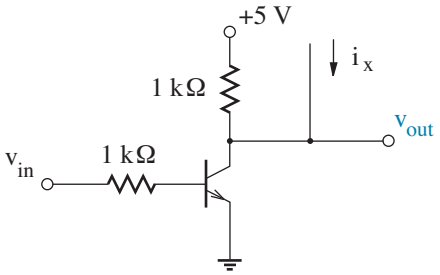


Figure P6.48

6.49 The BJT shown in Fig. P6.49 has $\beta_F = 50$. Determine the maximum value for i_x (a combined current from other circuits) such that v_{out} is HIGH at 4.8 V when v_{in} is LOW at 0 V.

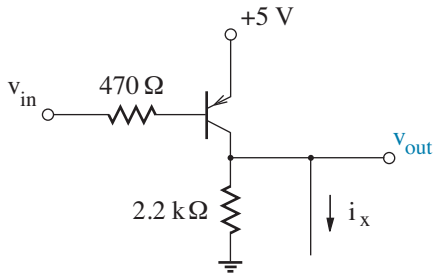


Figure P6.49

6.50 The circuit of Fig. P6.50 shows three of many subcircuits with identical BJTs ($\beta_F = 50$) and base resistors ($R_b = 82$ kΩ). For some strange reason, an electronically challenged engineer has designed the circuit so that v_{out} is LOW at 0.2 V when any two v_{in} inputs are HIGH at +5 V.

(a) Confirm the validity of the peculiar design.

(b) Suggest a better design criterion, and make suitable circuit changes, if necessary.

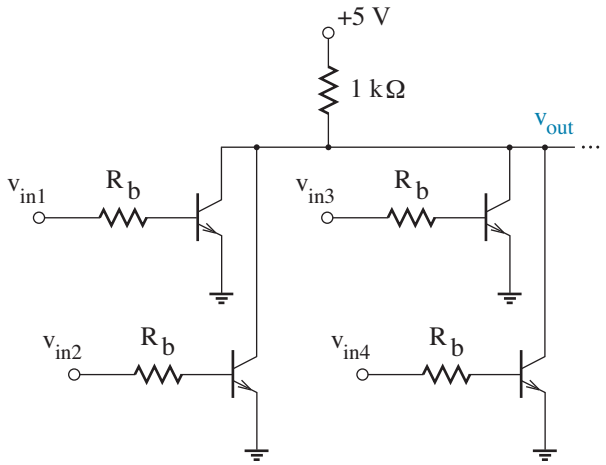


Figure P6.50

6.51 Now obsolete, the **Resistor Transistor Logic (RTL)** inverter of Fig. P6.51 features a BJT with $\beta_F = 80$. An acceptable LOW output is any voltage less than 0.7 V, and an acceptable HIGH output is any voltage greater than 1.2 V. The output connects to the inputs of N similar inverters. Determine the inverter **fanout**, the maximum N value that assures proper logical functionality throughout the circuit. Assume $v_{in} = 0$ or $v_{in} = +5$ V. (See Appendix.)

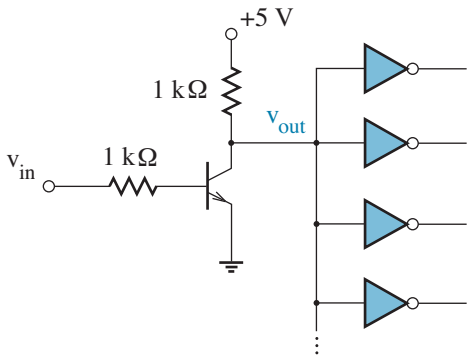


Figure P6.51

6.52 The input voltage source in the circuit of Fig. P6.52 has a large Thevenin resistance, so an emitter follower (Q_1) is used to drive switch (Q_2). Both BJTs have $\beta_F = 50$. Determine the minimum acceptable value for R such that $v_{out} = 0.2$ V (LOW) when $v_{in} = +5$ V (HIGH).

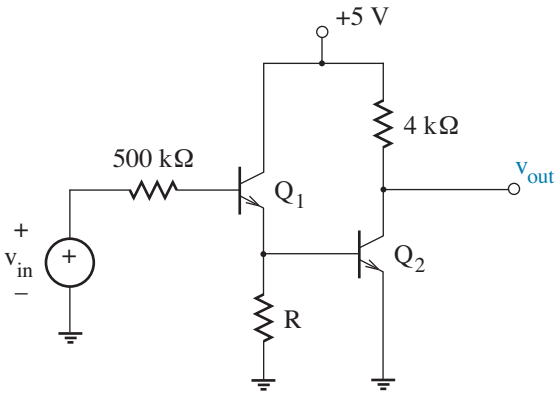


Figure P6.52

Section 6.4

SPICE parameter IS=10f unless otherwise indicated.

Elementary BJT Circuits

6.53 Use SPICE to verify the results of Exercise 6.3.

6.54 Use SPICE verify the cutoff and saturation conditions for Exercise 6.4 by sweeping current I over an appropriate range.

6.55 Use SPICE to verify the results of Exercise 6.7.

6.56 Use SPICE to verify the results of Exercise 6.8.

6.57 Use SPICE to verify the results of Example 6.4.

6.58 Use SPICE to verify the results of Example 6.6.

6.59 The circuit of Problem 6.30 has $R = 470$ kΩ.

(a) Use SPICE to show the behavior of v_{out} as β_F varies from 50 to 150.

(b) Repeat part a, but add a 1-kΩ resistor in series with the BJT emitter.

6.60 The circuit of Problem 6.31 has $R = 400$ kΩ.

(a) Use SPICE to show the behavior of v_{out} as β_F varies from 50 to 150.

(b) Repeat part a, but add a 10-kΩ resistor from the BJT base to ground.

6.61 The circuit of Problem 6.35 has $R = 1$ kΩ.

(a) Use SPICE to show the behavior of v_{out} as v_{in} varies from 7.5 V to 8.5 V.

(b) Repeat part a, but let the Zener diode have 30-Ω series R_z.

Complicating Circuit Effects

6.62 Use SPICE to demonstrate the behavior of v_{out} in the circuit of Problem 6.33 as temperature varies between -40 °C and +85 °C. Assume $X_{TB} = 1.5$.

6.63 Consider the β_F collector-current dependence shown in Fig. 6.17.

(a) Estimate SPICE parameters BF, IKF, and ISE.

(b) Use SPICE to replicate Fig. 6.17 subject to the parameters of part a.

6.64 The NE68833 BJT has IS=0.38f, BF=136, IKF=0.6, ISE=3.8f, and NE=1.49. Use SPICE to construct a Gummel plot with the form of Fig. 6.18.

Power Amplifiers

6.65 Use SPICE to validate the emitter "follower" results of Figs. 6.28 and 6.29.

6.66 Consider the push-pull amplifier of Fig. 6.30 with $R_t = R_L = 1$ kΩ, $V^+ = 15$ V, $V^- = -15$ V. Assume $\beta_F = 200$ for both transistors.

(a) Use SPICE to show cross-over distortion when the sinusoidal input signal has 10-V amplitude.

(b) Repeat part a, but let $R_L = 1$ MΩ. Explain any difference in the cross-over behavior.

6.67 Consider the push-pull amplifier of Fig. 6.33 with $R_t = 1$ kΩ, $R_L = 100$ Ω, $V^+ = 15$ V, and $V^- = -15$ V. Let $\beta_F = 200$ for all BJTs.

(a) Use SPICE to examine cross-over distortion for a sinusoidal input signal with 10-V amplitude. Consider three different R values: 400 Ω, 4 kΩ and 40 kΩ.

(b) Repeat part a, but investigate the load current as well as the collector currents for Q_1 and Q_2. Explain the results.

(c) Repeat part b with $R = 4$ kΩ, but let IS = 100f for Q_1 and Q_2. Explain the results.

6.68 Consider the push-pull amplifier of Fig. 6.34 with $R_t = 1$ kΩ, $R_L = 8$ Ω, $R = 200$ Ω, Ω, $V^+ = 15$ V, and $V^- = -15$ V. Let $\beta_F = 100$ for all BJTs. The sinusoidal input has 15-V amplitude.

(a) Use SPICE to examine v_{out} and the load current for three different values of R': 75 mΩ, 0.75 Ω and 7.5 Ω. Explain the results.

(b) Repeat part a with $R' = 0.75$ Ω, but let the BJT have IKF = 100m. Explain the results.

BJT Capacitance

6.69 An integrated npn BJT has the structure of Fig. 6.38 with the following doping concentrations and dimensional characteristics:

Layer	Doping Conc.	Dimensions
n^+ emitter	8×10^{18} cm^{-3}	8 μm x 4 μm
p base well	5×10^{17} cm^{-3}	14 μm x 6 μm
n collector well	2×10^{16} cm^{-3}	21 μm x 8 μm
n^+ buried collector	5×10^{18} cm^{-3}	21 μm x 8 μm
p substrate	2×10^{15} cm^{-3}	

Estimate SPICE parameters CJE, VJE, CJC, VJC, CJS, and VJS.

6.70 Figure P6.70 shows a lateral pnp BJT structure that allows for integration with an npn transistor on the same substrate. (Compare this with Fig. 6.38.) The device has the following doping concentrations and dimensional characteristics:

Layer	Doping Conc.	Dimensions
p^+ emitter	6×10^{18} cm^{-3}	4 μm x 6 μm
n base well	2×10^{17} cm^{-3}	21 μm x 8 μm
p collector (2)	4×10^{16} cm^{-3}	4 μm x 6 μm
n^+ buried layer	5×10^{18} cm^{-3}	21 μm x 8 μm
p substrate	2×10^{15} cm^{-3}	

Find SPICE parameters CJE, VJE, CJC, and VJC.

Note: In contrast with the npn BJT, the integrated pnp device has a relatively poor β_F due to surface minority-carrier recombination.

6.71 Consider the non-linear depletion capacitance

$$C = C_j \left[1 - \frac{v}{\phi}\right]^{-m}.$$

Let v vary from $v(t_a)$ at time t_a to $v(t_b)$ at time t_b. Show that the change in the stored depletion charge is given by

$$\Delta Q = \frac{\phi C_j}{1-m}\left\{\left[1 - \frac{v(t_a)}{\phi}\right]^{1-m} - \left[1 - \frac{v(t_b)}{\phi}\right]^{1-m}\right\}.$$

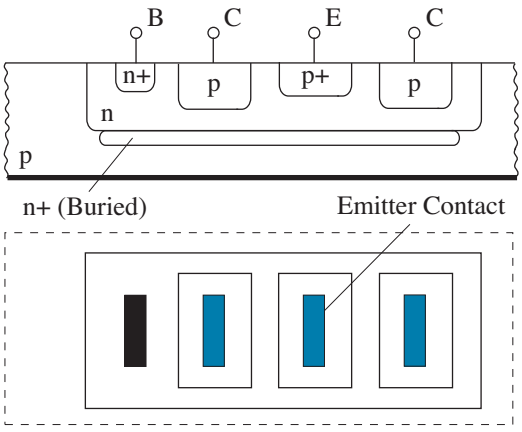


Figure P6.70

Appendix

6.72 The inverter circuit of Fig. 6.42 has $R_c = 2$ kΩ. The BJT has $\beta_F = 80$ and $\tau_f = 0.25$ ns. Let v_{in} step from zero to +5 V at $t = 0$.

(a) Determine the rise time subject to the maximum R_b that allows $v_{out} = 0.2$ V.

(b) Determine the rise time when $R_b = 500$ Ω.

Ignore BJT depletion capacitance.

6.73 The circuit of Fig. 6.42 has $R_c = R_b = 2$ kΩ. The BJT has $\beta_F = 80$, $\beta_R = 4$, $\tau_f = 0.25$ ns, and $\tau_r = 8$ ns. Let v_{in} step from zero to +5 V at $t = 0$.

(a) Determine the value of the "surplus" saturation charge Q_s achieved in the steady state.

(b) Find the time needed to establish Q_s at 80% of its steady-state value.

6.74 Consider the circuit of Fig. 6.42 subject to the specifications of Problem 6.73.

(a) Estimate the base-charge contribution to the storage time when v_{in} steps from +5 V to zero following a saturation period of 1 ms.

(b) Repeat the preceding calculation, but limit the saturation period to 100 ns.

6.75 Consider the circuit of Fig. 6.42 subject to the specifications of Problem 6.73. Let v_{in} step from +5 V to zero at $t = 0$ after a long saturation period.

(a) Estimate the base-charge contribution to the fall time.

(b) Repeat the preceding calculation, but let v_{in} transition to -2 V.

Ignore BJT depletion capacitance.

6.76 This problem concerns a so-called "speed-up capacitor" for the improvement of BJT rise time. The capacitor is positioned in parallel with the BJT base resistor as shown in Fig. P6.76. Let v_{in} step from zero to +5 V at $t = 0$.

If one neglects the base-emitter junction voltage,

$$i_b \approx C\,\frac{dv_{in}}{dt} + \frac{v_{in}}{R_b}\,.$$

(a) Prove that the rise time is essentially zero when $R_bC = \tau_{bf}$. Confirm with SPICE.

(b) Use SPICE to demonstrate the circuit behavior subject to $R_bC = 2\,\tau_{bf}$. Explain the results.

(c) Use SPICE to demonstrate the circuit behavior subject to $R_bC = 0.5\,\tau_{bf}$. Explain the results.

Ignore BJT depletion capacitance.

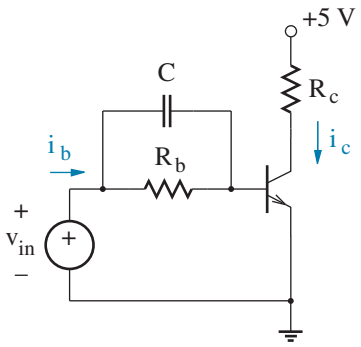


Figure P6.76

6.77 Consider the circuit of Fig. 6.42 subject to the specifications of Problem 6.73. Let v_{in} step from zero to +5 V at $t = 0$.

The BJT also has the following SPICE parameters:

CJE=4.5p, VJE=0.75, MJE=0.5
CJC=3.5p, VJD=0.75, MJC=0.33

(a) Estimate the change in the base-emitter depletion charge ΔQ_{je} during the delay period.

(b) Repeat the estimate for the collector-emitter depletion charge ΔQ_{jc}.

(c) Estimate the average base current during the delay time period, and use this and preceding results to estimate the delay time.

(d) Compare the estimated delay time with that obtained with SPICE.

6.78 Estimate the depletion-charge contributions to the BJT rise time as in Problem 6.78. Use SPICE to simulate the rise time with and without (TF=0) contributions from stored diffusion charge.

Perspective:

Circuit Strategies for Amplification

Back in Chapter 1, we found that one of the more useful processes that can be applied to an analog signal is amplification to make it significantly larger. The operational amplifier was our weapon of choice, and we treated it as a black box that adheres to a simple set of rules. Nevertheless, we noted that one could peel back its cover to see a collection of transistors that function like valves for through-current control, and we surmised that the limitations of these devices would help us to predict the breakdown of black-box rules. Chapters 5 and 6 have placed the transistors at hand. And while we are not quite ready to integrate them into an op-amp, we are prepared to examine some basic amplifier circuits containing one or perhaps a few transistors. Although discrete-component applications have given way to the op-amp at low or moderate signal frequencies, they still enjoy widespread use at very high frequencies that are required for numerous RF communication circuits. When treated as building blocks, single- and two-transistor amplifiers serve as starting points for internal op-amp design, and the factors that corrupt their performance have similar implications in the integrated realm.

At the risk of becoming rather abstract, the present Perspective seeks an overview of Chapter 7 by examining the behavior of a generic circuit that functions as an analog voltage amplifier. But first some definitions ...

In earlier discussion, we assigned a Q-point label to indicate a voltage and current that satisfy separate circuit constraints *at any particular time*. Thus, the Q-point has been free to adjust in response to time-dependent source variations. In what follows, however, it is advantageous to designate a restricted operating point that reflects a particular **biasing** condition in which sources have time-averaged (dc) values. Thus, in anticipation of time-dependent (ac) source variations, the Q-point serves as a truly quiescent *point of reference* for voltage and current excursions that emerge from it. The excursions also satisfy load-line and device constraints, but they receive no special instantaneous significance.

Apart from the Q-point variables, every other current or voltage in an analog amplifier can be expressed as the sum of a quiescent (dc) and a changing (ac) component. In this text, we use the following notation:

$$i = i|_Q + \Delta i\,, \tag{C1}$$

and

$$v = v|_Q + \Delta v\,. \tag{C2}$$

Subscripts are used to denote particular currents and voltages.[4]

Consider the prospective amplifier of Fig. C1. The amplifier input is represented by Thevenin source V_g, which consists of a constant dc voltage and a superimposed ac voltage. Specifically,

$$V_g(t) = V_{gg} + \Delta V_g(t)\,. \tag{C3}$$

The amplifier output is v_{ds}. Nevertheless, we also take interest in i_d as the companion load-line variable.

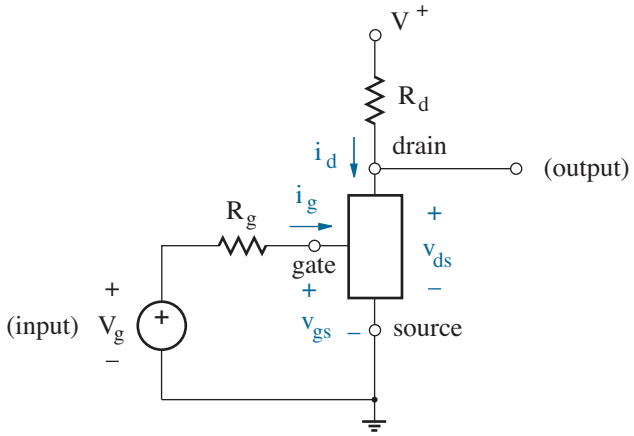


Figure C1: Prospective amplifier with a generic three-terminal transistor.

The distinction between dc and ac input components suggests that we approach the amplifier circuit with a two-pronged attack. First, we identify $i_d|_Q$ and $v_{ds}|_Q$, the Q-point of reference. To do this, we set ac components to zero so that $V_g = V_{gg}$. (V^+ is already constant.) In addition, we replace the three-terminal transistor with a **large-signal** model that describes its non-linear behavior. Chapters 5 and 6 have established some candidates.

[4] Ages ago, when the author was a student, he was advised of a "standard" notation in which dc components have capital letters with capital subscripts and ac components have small letters with small subscripts. The painful experience of trying to differentiate between V_C and v_c on a distant blackboard has prompted the alternative used here.

Based on our previous experience, we assume that we can determine a Q-point solution to the modified circuit of Fig. C2 if the characteristics of the non-linear model elements are known. We hope for a circuit design that ensures some "favorable" Q-point position on the load line. And we hope for a *stable* Q-point with weak dependence on uncertain device parameters and ambient conditions such as temperature.

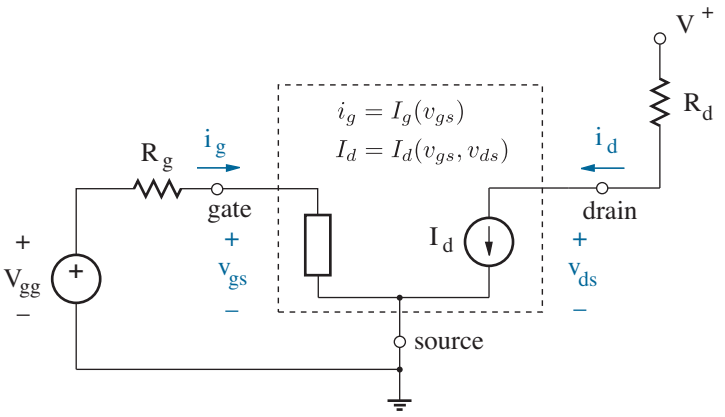


Figure C2: Modified circuit used to determine quiescent (dc) currents and voltages in the amplifier of Fig. C1. The circuit features a generic large-signal hybrid-π transistor model, and all ac sources are set to zero.

Having found the Q-point, we can qualitatively envision the effects of ac input changes with the help of Fig. C3. An amplifier with large voltage gain produces large changes in Δv_{ds} for small changes in ΔV_g.

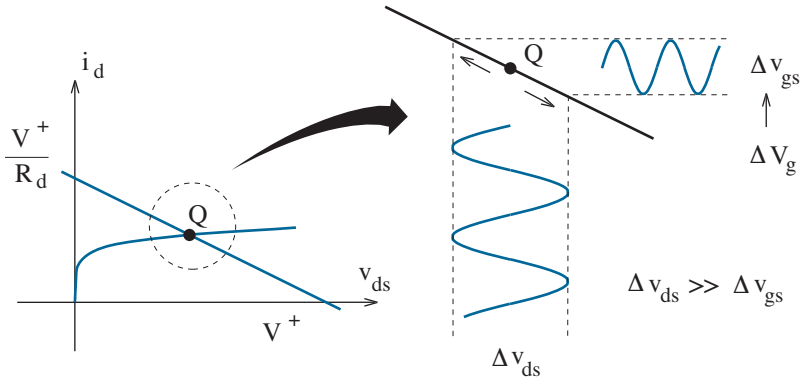


Figure C3: Qualitative view of the amplification process.

The inevitable need for a more quantitative relationship between Δv_{ds} (the ac output) and ΔV_g (the ac input) brings us to the second phase of our analysis, which would be elementary if the transistor were a linear device. We would use the same large-signal model, but with dc sources set to zero. Then we would superimpose the independent dc and ac solutions to obtain a total amplifier response. But the transistor is *not* a linear device!

No problem. *The transistor is effectively linear, provided ac current and voltage changes are small.*

To demonstrate the means for linear transistor behavior, we calculate the total differentials that apply to the non-linear characteristic relations for i_g and i_d in the generic large-signal hybrid-π model. Specifically,

$$\Delta i_g = \left(\frac{\partial i_g}{\partial v_{gs}}\right)\bigg|_Q \Delta v_{gs} = \frac{\Delta v_{gs}}{r_\pi} \,, \tag{C4}$$

and

$$\Delta i_d = \left(\frac{\partial i_d}{\partial v_{gs}}\right)\bigg|_Q \Delta v_{gs} + \left(\frac{\partial i_d}{\partial v_{ds}}\right)\bigg|_Q \Delta v_{ds} = g_m \Delta v_{gs} + \frac{\Delta v_{ds}}{r_o} \,. \tag{C5}$$

In these expressions, the coefficients for Δv_{gs} and Δv_{ds} are constant since they are evaluated under quiescent operating conditions. By interpreting the coefficients as parameters applicable to standard linear circuit elements, we can readily synthesize the equivalent generic **small-signal** hybrid-π transistor model shown in Fig. C4. The model is "static" in the sense that its development does not depend on the time derivative.

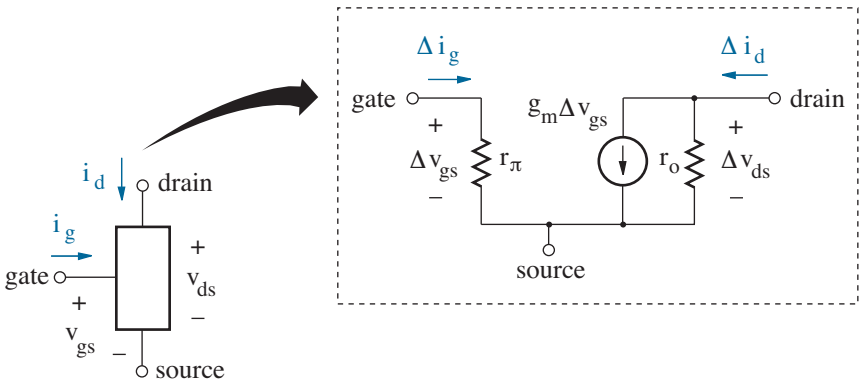


Figure C4: Generic small-signal hybrid-π transistor model.

Each model parameter in Fig. C4 has physical significance that depends on the circuit application. For the common-source transistor configuration at hand, the transistor source terminal is shared between input and output. Thus, r_π, g_m, and r_o are indicative of the small-signal **input resistance**, **transconductance**, and **output resistance**, respectively.

The linear small-signal transistor model allows us to obtain an ac circuit solution that can be superimposed on an independent dc solution. Following the rules of superposition, we construct the modified ac circuit of Fig. C5 *by setting dc sources to zero*. The new circuit follows directly from Fig. C1. However, with $V^+ = 0$, R_d is now connected to ground.

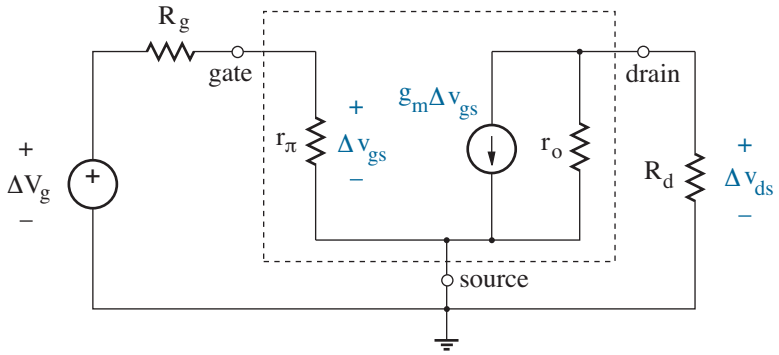


Figure C5: Modified circuit used to relate small-signal (ac) currents and voltages in the amplifier of Fig. C1. The circuit features a small-signal hybrid-π transistor model, and dc sources are set to zero.

From Fig. C5, we determine the voltage gain, which is defined as $\Delta v_{ds}/\Delta V_g$. By inspection,

$$A_v = \frac{\Delta v_{ds}}{\Delta V_g} = \frac{\Delta v_{ds}}{\Delta v_{gs}} \frac{\Delta v_{gs}}{\Delta V_g} = -g_m(R_d \parallel r_o)\left(\frac{r_\pi}{r_\pi + R_g}\right). \qquad \text{(C6)}$$

So it appears that an amplifier with large voltage gain requires a transistor with large g_m, r_π, and r_o.

What does this mean? Large g_m implies a drain-current characteristic that varies strongly with v_{gs} for sensitive valve control. Large r_π implies a low-level gate current, which otherwise reduces the drive capability of an input source with finite Thevenin resistance (R_g). Finally, large r_o implies a set of device characteristic curves that are flat (not tilted) so that current is efficiently delivered to the load (R_d). Chapters 7 and 8 let $r_0 \to \infty$.

The preceding analysis suggests that small-signal model parameters are important as guides for choosing a good transistor for amplifier applications. Thus, we need to develop a catalog of parametric expressions that apply to the MOSFET and BJT. (Perhaps another device will someday be superior.) We are open to amplifiers differing from Fig. C1. And we look for measures of performance other than voltage gain. So much for Chapter 7.

In practice, a complete single-stage ac amplifier circuit typically derives from a dc biasing circuit through the connection of an ac signal source and a resistive load. These connections must not be allowed to disrupt a carefully established Q-point, so they are often made through **coupling capacitors** that serve as ac short circuits at medium frequencies of interest and higher. No matter how large, the coupling capacitors become open circuits in the low-frequency limit, coupling is lost, and the voltage gain approaches zero. Meanwhile, we acknowledge that there is always a capacitance between any two nodes, a reflection of proximity, yet often ignored as an ac open circuit. No matter how small, these **parasitic capacitors** become short circuits in the high-frequency limit and thus tend to destroy the voltage gain through a variety of mechanisms. (Consider, for example, C_π in parallel with r_π.) Figure C6 shows a typical amplifier frequency response.

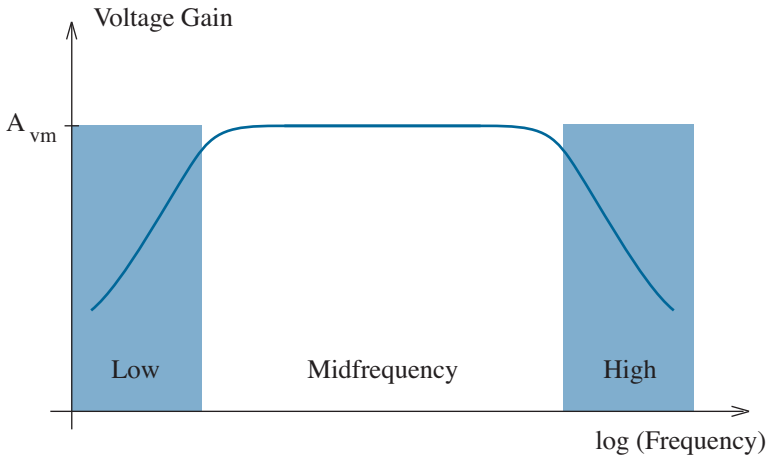
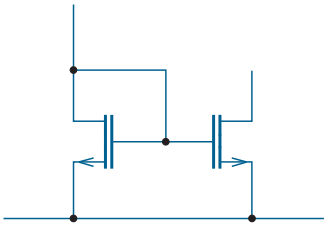


Figure C6: Typical amplifier frequency response and "midfrequency" range.

Whereas Chapter 7 considers capacitors as either useful or non-existent, the results therein apply to **midfrequency**, the unshaded region in Fig. C6. Chapter 8 examines the physical basis for high-frequency capacitive effects, a procedure for estimating the high-side demise of midfrequency operation, and design practices for extending or otherwise improving high-frequency circuit behavior. Chapter 8 also covers low-frequency analysis and design. We defer the prospect for zero-frequency (dc) operation to Chapter 9.



Chapter 7

Midfrequency Amplifiers

This chapter concerns the relationships between ac changes in three basic amplifier circuits used as building blocks for numerous analog applications. We begin by demonstrating biasing techniques that establish Q-points of reference for excursions along a load line. Then we determine small-signal transistor parameters that generally support four measures of performance for each of the basic single-stage amplifiers. We conclude with some simple design procedures for discrete-transistor (BJT) amplifiers.

Warning: The chapter is rather long, but we have much to derive.

Objectives

After studying this chapter and working through the problems, you should be able to perform each of the following:

- Analyze and design a "negotiated" or "imposed" biasing circuit for a MOSFET or BJT (Section 7.1).
- Calculate the g_m and r_π small-signal hybrid-π model parameters for an arbitrary transistor (Section 7.2).
- Determine the midfrequency voltage gain, input resistance, current gain, and output resistance for three basic amplifier configurations (Section 7.3).
- Use SPICE to specify an amplifier's total midfrequency voltage gain (Section 7.3)
- Analyze a multi-stage cascaded amplifier (Section 7.3).
- Design a single-stage BJT common-emitter amplifier (Section 7.4).

7.1 Elementary Biasing Principles

Most ac transistor amplifiers require one or more well-established Q-points that serve as references for small-signal current and voltage excursions along particular load lines. The associated **biasing** has one of two forms:

- In what could be described as a *negotiated* biasing circuit (Fig. 7.1a), drain (collector) current is determined as a compromise between the expectations of governing transistor principles and those of a set of external components. Good designs require parametric insensitivity, so the latter expectations are generally favored.

- In what could be described as an *imposed* biasing circuit (Fig. 7.1b), drain (collector) current is determined through forced influence of an external current source that has been designed to be insensitive to variations in supply voltage and/or temperature.

The Q-point voltage component reflects external constraints in both cases.

In most applications, the negotiated form of biasing tends to be used in circuits containing discrete transistors, whereas the imposed form of biasing has importance for analog integrated circuits (Chapter 9). For simplicity, our development of ac analytical procedures uses the negotiated form.

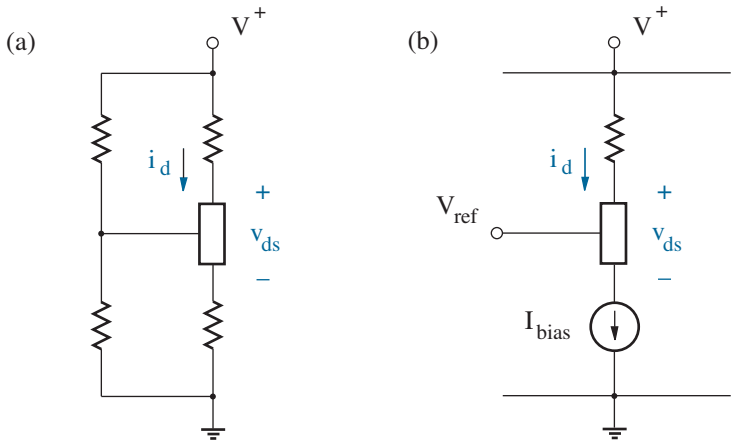


Figure 7.1: Two forms of transistor biasing: (a) the standard negotiated biasing circuit establishes a current compromise between a transistor, four external resistors, and a power supply; (b) the imposed biasing circuit has a specially designed current source that forces accommodation.

Negotiated Biasing: MOSFETs

Figure 7.2 shows the four-resistor negotiated MOSFET biasing circuit and its graphical solution for quiescent drain current. The MOSFET draws zero gate current, so the gate node voltage follows a divider relation:

$$v_g = V^+ \left(\frac{R_2}{R_1 + R_2} \right) . \tag{7.1}$$

And with equal drain and source currents, the source node voltage is $i_d R_s$. Thus, the gate-to-source voltage is

$$v_{gs} = V^+ \left(\frac{R_2}{R_1 + R_2} \right) - i_d R_s . \tag{7.2}$$

This is the drain-current-v_{gs} constraint required by the external circuit. Meanwhile, the MOSFET requires

$$i_d = \frac{1}{2} K' \frac{W}{L} (v_{gs} - V_T)^2 \tag{7.3}$$

provided that it operates in the saturation mode. Negotiations yield the Q-point at the intersection of both constraints. Then with $i_s = i_d = i_d|_Q$,

$$v_{ds}|_Q = V^+ - i_d|_Q (R_d + R_s) . \tag{7.4}$$

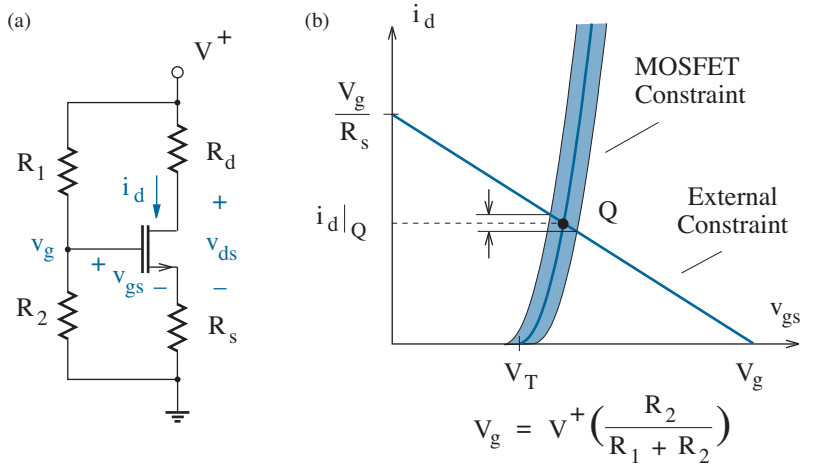


Figure 7.2: Negotiated MOSFET biasing: (a) Four-resistor biasing circuit; (b) Graphical solution showing the effect of uncertainty in the MOSFET drain-current constraint, a curve that lies within the shaded region.

Caution! The preceding solution for $i_d|_Q$ is only valid when

$$v_{ds}|_Q > v_{gs}|_Q - V_T > 0 \tag{7.5}$$

so that the MOSFET operates in saturation subject to Eq. 7.3. The form of Eq. 7.4 shows that a solution has the opportunity to fail if R_d is large. In this case, Eq. 7.3 must be replaced with the more complicated resistive-mode expression for the MOSFET drain current:

$$i_d = \frac{1}{2}K'\frac{W}{L}\left[2(v_{gs} - V_T)v_{ds} - {v_{ds}}^2\right] . \tag{7.6}$$

However, it is probably more productive to recognize a poor biasing design.

It will soon become apparent that only modest algebraic effort is needed to negotiate the differing expectations of Eqs. 7.2 and 7.3. Nevertheless, the graphical solution helps to illustrate the influence of source resistor R_s. A real MOSFET is expected to have uncertainty in both $K'W/L$ and V_T through process and temperature variations. Thus, the MOSFET's drain-current constraint is a half-parabolic curve that lies somewhere within the shaded region in Fig. 7.2b. The related drain-current uncertainty *decreases* as the external drain-current constraint becomes more horizontal through increased R_s. (A worst case eliminates R_s entirely so that v_{gs} is set by V_g). The R_s resistor provides a means of drain-current *feedback* for v_{gs} control. Were i_d to increase by some mysterious process, the source node voltage would also increase, thereby decreasing v_{gs} and offsetting the change in i_d.

Example 7.1

Determine $i_d|_Q$ and $v_{ds}|_Q$ in the MOSFET biasing circuit of Fig. 7.3.

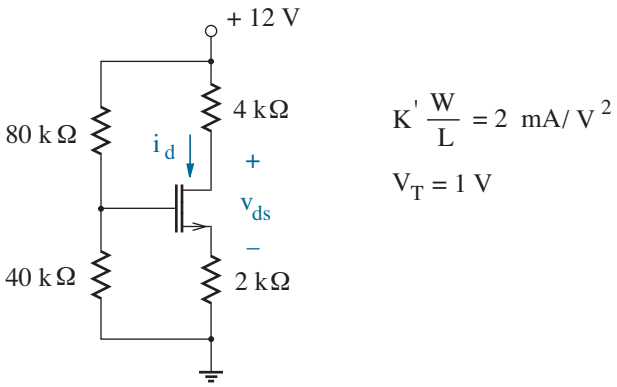


Figure 7.3: MOSFET biasing circuit for Example 7.1.

Solution

The node voltage at the MOSFET gate terminal is

$$v_g = 12\left(\frac{40}{40+80}\right) = 4\text{ V}, \tag{7.7}$$

and the node voltage at the MOSFET source terminal is

$$v_s = 2i_s = 2i_d\,, \tag{7.8}$$

(where i_d has units of mA). Thus,

$$v_{gs} = 4 - 2i_d\,. \tag{7.9}$$

But if the MOSFET is in saturation,

$$i_d = \frac{1}{2}(2)(v_{gs}-1)^2\,. \tag{7.10}$$

We substitute Eq. 7.10 into Eq. 7.9 and rearrange to obtain

$$2{v_{gs}}^2 - 3v_{gs} - 2 = 0\,. \tag{7.11}$$

The solutions to this quadratic equation are $v_{gs} = 2$ V and $v_{gs} = -0.5$ V. Our worry over the prospect of two solutions is brief when we note that the second is invalid—it is less than the threshold voltage V_T. So we accept the first solution and return to Eq. 7.10 to find $i_d = i_d|_Q = 1$ mA. In turn,

$$v_{ds}|_Q = 12 - 1\,(4+2) = 6\text{ V}. \tag{7.12}$$

The final (and most important) step is to check the solution for consistency with the assumed mode of MOSFET operation. Under saturation,

$$v_{ds}|_Q > v_{gs}|_Q - V_T = 1\text{ V}. \tag{7.13}$$

This is indeed the case, so the Q-point solution is acceptable.

We pause to note that the preceding analysis fails to account for two complications that influence the MOSFET drain current. Channel-length modulation (described via parameter λ) can be significant when v_{ds} is large, and the body effect (described via parameter γ) has importance when the source node voltage is at a different potential than the substrate (body). As noted in Chapter 5, these are issues for Chapter 9.

Although relatively rare, negotiated biasing circuits for depletion-mode MOSFETs take similar forms of analysis (with device-specific i_d constraints). Apart from its role in decreasing bias uncertainty, the R_s source resistor is needed to ensure negative (positive) v_{gs} for n-channel (p-channel) devices.

Exercise 7.1 Determine the Q-points in the circuits of Fig. 7.4. Assume $K'W/L = 0.5$ mA/V^2, $V_T = 1$ V.

Ans: (a) $i_d|_Q = 0.69$ mA, $v_{ds}|_Q = 5.0$ V

(b) $i_d|_Q = 2.1$ mA, $v_{ds}|_Q = 3.7$ V

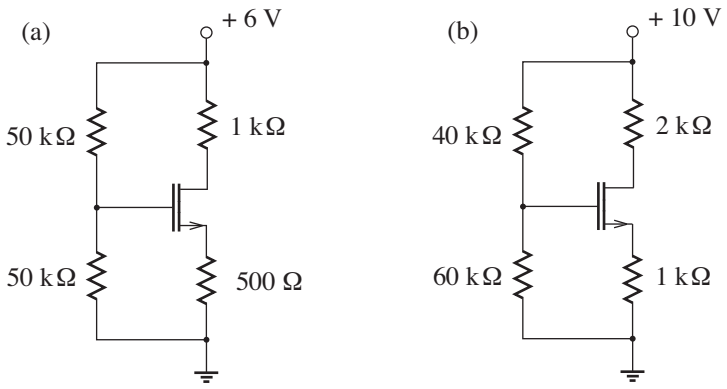


Figure 7.4: MOSFET biasing circuits for Exercise 7.1.

Exercise 7.2 Determine the Q-points in the circuits of Fig. 7.5. Assume $K'W/L = 2.0$ mA/V^2, $V_T = -1$ V.

Ans: (a) $i_d|_Q = -3.2$ mA, $v_{ds}|_Q = -7.2$ V

(b) $i_d|_Q = -0.69$ mA, $v_{ds}|_Q = -5.9$ V

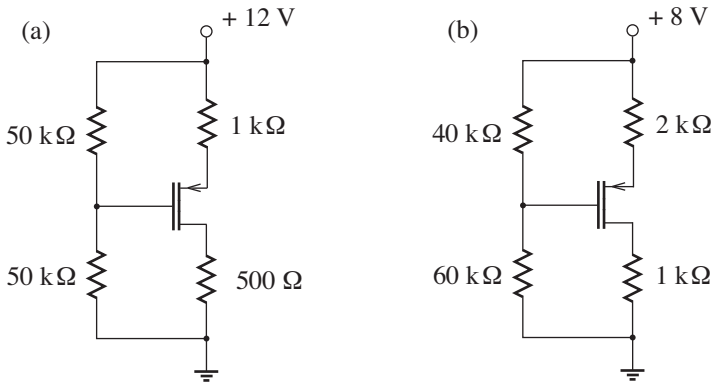


Figure 7.5: MOSFET biasing circuits for Exercise 7.2

Example 7.2

Determine R_1 and R_2 such that $i_d|_Q = 0.25$ mA and $R_1 \parallel R_2 = 60$ kΩ in the circuit of Fig. 7.6.

(This is part of the design process for negotiated biasing.)

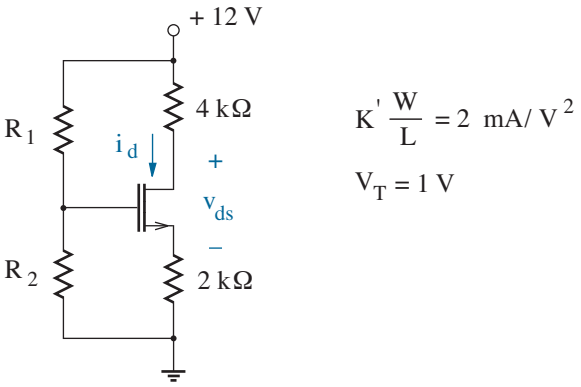


Figure 7.6: MOSFET biasing circuit for Example 7.2.

Solution

Our first objective is to find a numerical value for the gate node voltage. If $i_d|_Q = 0.25$ mA, the source node voltage is

$$v_s = i_d|_Q R_s = 0.5 \text{ V}. \tag{7.14}$$

And since

$$v_{gs}|_Q = V_T + \sqrt{\frac{2\, i_d|_Q}{K' W/L}} = 1.5 \text{ V}, \tag{7.15}$$

the gate node voltage is

$$v_g = 12 \left(\frac{R_2}{R_1 + R_2} \right) = v_s + v_{gs}|_Q = 2.0 \text{ V}. \tag{7.16}$$

The result is an equation with unknowns R_1 and R_2. Meanwhile,

$$R_1 \parallel R_2 = \frac{R_1 R_2}{R_1 + R_2} = 60 \text{ kΩ}. \tag{7.17}$$

So dividing Eq. 7.16 into Eq. 7.17, we obtain $R_1 = 360$ kΩ. This value is inserted back into Eq. 7.16 (or Eq. 7.17) to yield $R_2 = 72$ kΩ.

Negotiated Biasing: BJTs

Figure 7.7a shows the four-resistor negotiated biasing circuit for a bipolar junction transistor. The analysis is complicated by non-zero base current, which disallows a simple divider relation for the node voltage at the base. Nevertheless, the governing algebra avoids unpleasant quadratic equations that are generally associated with field-effect devices.

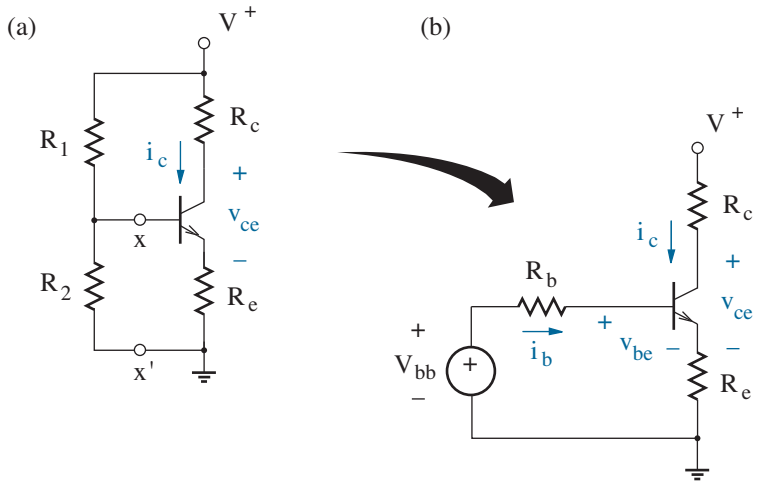


Figure 7.7: Negotiated BJT biasing: (a) Four-resistor biasing circuit; (b) reconfigured circuit with a Thevenin equivalent of the base component.

To account for the effect of base current, it is helpful to find the Thevenin equivalent of the portion of the biasing circuit that is connected to the base. The Thevenin voltage is the open-circuit voltage at terminals x - x' when the base is disconnected:

$$V_{bb} = V^+ \frac{R_2}{R_1 + R_2} \tag{7.18}$$

—this would be the base node voltage if the BJT drew zero base current. The Thevenin resistance is the equivalent resistance looking back into the terminals x - x' when the biasing network is "dead" (turned off). So split V^+ into two components: one connected at the top of R_c, and the other connected at the top of R_1. Turn off or ground the latter, and one finds

$$R_b = R_1 \parallel R_2 \,. \tag{7.19}$$

Thus, we have the revised biasing circuit of Fig. 7.7b.

Having transformed the negotiated biasing circuit, we seek a solution for the base current that reflects BJT operation in the forward active mode. The consistent emitter current is given by

$$i_e = (\beta_F + 1)\, i_b \,. \tag{7.20}$$

Thus, application of Kirchhoff's Voltage Law (KVL) to the base-emitter loop yields the expression

$$V_{bb} = i_b R_b + v_{be} + (\beta_F + 1)\, i_b R_e \,. \tag{7.21}$$

Strictly speaking, the base-emitter voltage v_{be} is logarithmically dependent upon the collector current, which, in turn, is proportional to base current. But the logarithmic functionality is weak, so

$$v_{be} \approx 0.7 \text{ V} \tag{7.22}$$

as for a typical diode in forward bias. We now have

$$i_b = \frac{V_{bb} - 0.7}{R_b + (\beta_F + 1)R_e} \,. \tag{7.23}$$

Then with $i_c = \beta_F i_b$ in the forward active mode,

$$i_c|_Q = \beta_F \, \frac{V_{bb} - 0.7}{R_b + (\beta_F + 1)R_e} \tag{7.24}$$

and

$$\begin{aligned} v_{ce}|_Q &\approx V^+ - i_c|_Q (R_c + R_e) \\ &= V^+ - \beta_F \, \frac{V_{bb} - 0.7}{R_b + (\beta_F + 1)R_e} \, (R_c + R_e) \,, \end{aligned} \tag{7.25}$$

since $i_e|_Q \approx i_c|_Q$.

Equation 7.24 reveals the role of the emitter resistor R_e in providing bias stability against β_F variations with device selection and temperature. If $R_b \gg (\beta_F + 1)R_e$, the first term in the denominator dominates, and

$$i_c|_Q \approx \beta_F \, \frac{V_{bb} - 0.7}{R_b} \tag{7.26}$$

with linear β_F dependence. Conversely, if $R_b \ll (\beta_F + 1)R_e$, the second term in the denominator dominates, and

$$i_c|_Q \approx \left(\frac{\beta_F}{\beta_F + 1} \right) \frac{V_{bb} - 0.7}{R_e} \approx \frac{V_{bb} - 0.7}{R_e} \tag{7.27}$$

(for $\beta_F \gg 1$). Once again, the control mechanism is a form of feedback. Were i_c to increase by some mysterious process, the emitter node voltage would also increase, thereby decreasing v_{be} and offsetting the change in i_c.

Example 7.3

Determine $i_c|_Q$ and $v_{ce}|_Q$ in the BJT biasing circuit of Fig. 7.8a.

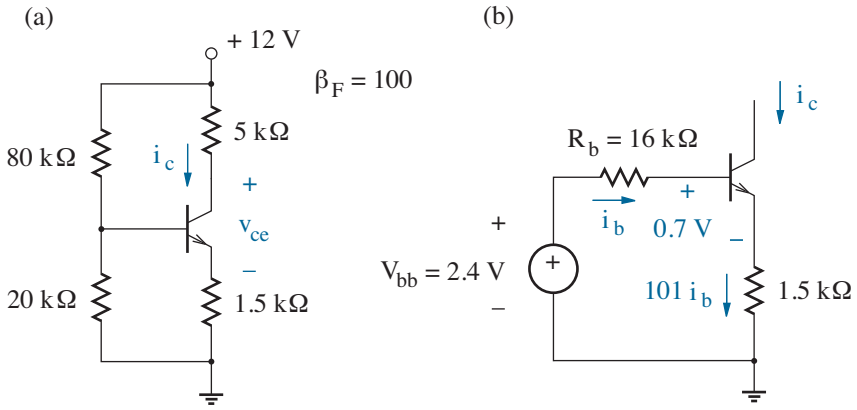


Figure 7.8: BJT biasing circuit and Thevenin equivalent for Example 7.3.

Solution

We draw the Thevenin equivalent of the base-connected portion of the biasing circuit with $V_{bb} = 12 \times 20\,/(80 + 20) = 2.4$ V and $R_b = 80 \parallel 20 = 16$ kΩ as shown in Fig. 7.8b. When drawing the equivalent circuit, we take care to label the base current i_b, the emitter current $i_e = (\beta_F + 1)i_b = 101 i_b$, and the base-to-emitter voltage $v_{be}|_Q \approx 0.7$ V. The latter two conditions assume the forward active mode of BJT operation.

Next, we write a KVL equation around the base-emitter loop:

$$2.4 = 16 i_b + 0.7 + 1.5 \times 101 i_b\,. \tag{7.28}$$

We solve to find $i_b \approx 0.01$ mA. And since $i_c|_Q = \beta_F\, i_b|_Q$ in the forward active mode, $i_c|_Q \approx 1$ mA. The companion Q-point co-ordinate is given by

$$v_{ce}|_Q \approx 12 - (5 + 1.5)\, i_c|_Q = 5.5 \text{ V}. \tag{7.29}$$

The final (and most important) step is to check the solution for consistency with the assumed mode of BJT operation. In the forward active mode,

$$v_{ce}|_Q > v_{ce,sat} \approx 0.2 \text{ V}. \tag{7.30}$$

No problem here, so the Q-point solution is acceptable.

Exercise 7.3 Determine the Q-point in the circuits of Fig. 7.9.

Ans: (a) $i_c|_Q = 4.5$ mA, $v_{ce}|_Q = 4.5$ V

(b) $i_c|_Q = 5.2$ mA, $v_{ce}|_Q = 2.5$ V

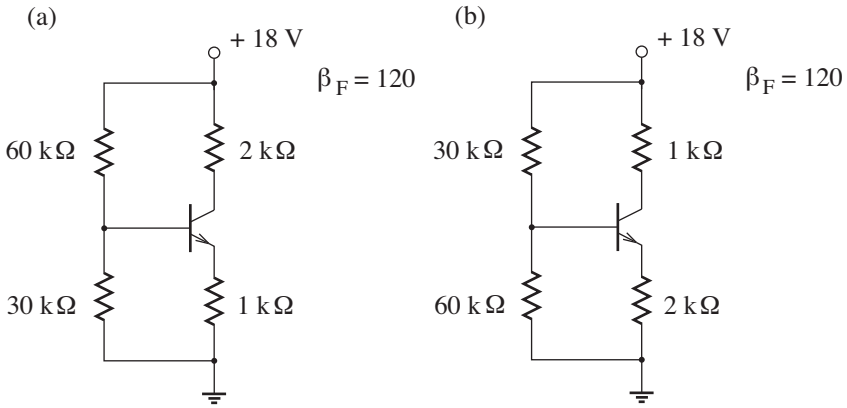


Figure 7.9: BJT biasing circuits for Exercise 7.3.

Exercise 7.4 Determine the Q-point in the circuits of Fig. 7.10.

Ans: (a) $i_c|_Q = -6.0$ mA, $v_{ce}|_Q = -3.6$ V

(b) $i_c|_Q = -6.7$ mA, $v_{ce}|_Q = -1.9$ V

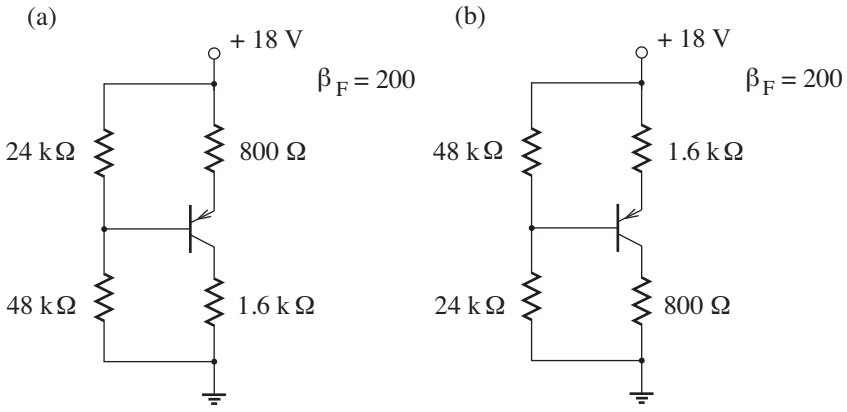


Figure 7.10: BJT biasing circuits for Exercise 7.4.

Biasing Considerations for the Discrete BJT

Discrete-transistor amplifiers are more likely to contain BJTs as opposed to MOSFETs or other transistors. So when considering this type of circuit, we need to understand not just *how* to set the Q-point, but *where* to set it in order to establish optimum performance.

A good BJT biasing design avoids the following:

- Saturation conditions (for which $v_{ce} < v_{ce,sat} \approx 0.2$ V)
- Excessive v_{ce} (that may lead to base-collector junction breakdown)
- Overheating (whereby $i_c v_{ce} > P_{max}$, the BJT power rating)

Meanwhile, it is appropriate to choose a Q-point that maximizes excursions along the *ac load line*, which is potentially distinct from the dc load line. For example, it is sometimes desirable to avoid wasteful quiescent power dissipation in a load by decoupling it from the BJT amplifier output with a capacitor that acts as an open circuit under dc biasing conditions but a short circuit under ac operating conditions.

Figure 7.11 shows distinct dc and ac load lines.

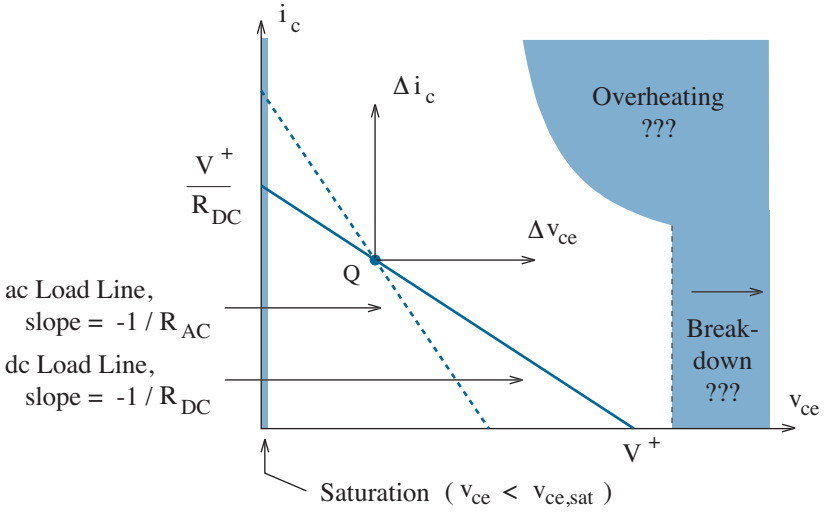


Figure 7.11: BJT dc and ac load lines, and proscribed operating regions. The shaded areas show saturation, breakdown, and overheating portions of the $i_c - v_{ce}$ plane. Seek them not.

With the goal of placing the Q-point in the middle of the ac load line, we derive equations for the dc and ac load lines with respect to i_c and v_{ce}. For the ac load line, we have

$$i_c = \frac{-v_{ce}}{R_{AC}} + b\,, \tag{7.31}$$

where $b = 2\,i_c|_Q$ is the desired i_c-axis intercept. So at the Q-point,

$$i_c|_Q = \frac{-\,v_{ce}|_Q}{R_{AC}} + 2\,i_c|_Q\,. \tag{7.32}$$

We solve this equation for $v_{ce}|_Q$ to obtain

$$v_{ce}|_Q = R_{AC}\,i_c|_Q\,. \tag{7.33}$$

Then for the dc load line,

$$i_c = \frac{V^+ - v_{ce}}{R_{DC}}\,. \tag{7.34}$$

So at the Q-point, and with the help of Eq. 7.33,

$$i_c|_Q = \frac{V^+ - R_{AC}\,i_c|_Q}{R_{DC}}\,. \tag{7.35}$$

We solve this equation for $i_c|_Q$ to obtain

$$i_c|_Q = \frac{V^+}{R_{DC} + R_{AC}}\,. \tag{7.36}$$

In general, the quantities R_{DC} and R_{AC} are determined by inspection according to the following rules:

- R_{DC} is the total resistance looking away from the collector and the emitter under dc conditions when capacitors are *open* circuits.
- R_{AC} is the total resistance looking away from the collector and the emitter under ac conditions when capacitors are *short* circuits and all dc sources are set to zero.

Equations 7.33 and 7.36 supporting Q-point *design* are DANGEROUS. When the Q-point is free to be adjusted, as when R_1 and R_2 are unspecified in the circuit of Fig. 7.7, the design equations are convenient for finding an optimum Q-point location. However, once the Q-point has been negotiated through specific component selection, the design equations are irrelevant. We have no way of knowing what design standards have been imposed, so the Q-point must be determined by the procedure of Example 7.3.

Example 7.4

Determine the Q-point that maximizes the voltage swing along the ac load line for the circuit of Fig. 7.12, then find consistent R_1 and R_2 values.

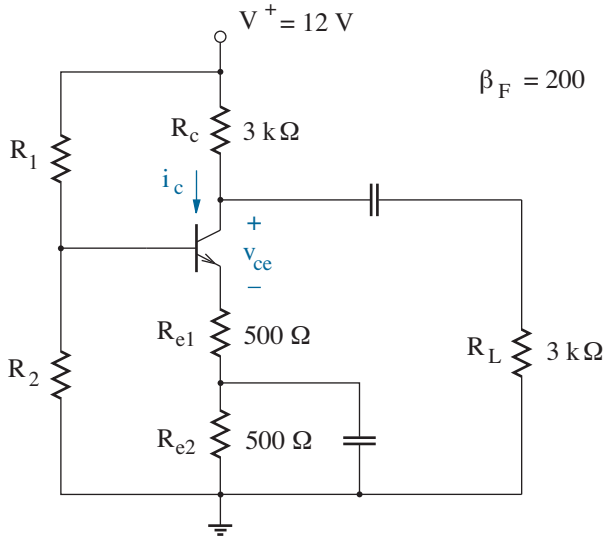


Figure 7.12: BJT biasing circuit for Example 7.4.

Solution

Under dc conditions, capacitors are open circuits . So looking away from the collector (towards V^+), one sees R_c, and looking away from the emitter (towards ground), one sees R_{e1} in series with R_{e2}. Thus,

$$R_{DC} = R_C + R_{e1} + R_{e2} = 4\ \text{k}\Omega. \tag{7.37}$$

Under ac conditions, capacitors are effectively short circuits and dc sources are set to zero. So looking away from the collector, one sees a path to ac ground through R_c and a parallel path to the actual ground through R_L. And looking away from the emitter, one sees a path through R_{e1} followed by a shortcut around R_{e2} as the ac route to ground. Thus,

$$R_{AC} = R_C \parallel R_L + R_{e1} = 2\ \text{k}\Omega. \tag{7.38}$$

Then for maximum voltage swing along the ac load line,

$$i_c|_Q = \frac{V^+}{R_{DC} + R_{AC}} = \frac{12}{4+2} = 2.0\ \text{mA}. \tag{7.39}$$

The companion Q-point component is

$$v_{ce}|_Q = R_{AC}\, i_c|_Q = 2 \times 2.0 = 4.0 \text{ V}. \tag{7.40}$$

In turn, and taking account for the 0.2-V-wide saturation region (see Fig. 7.11), the available swing along the ac load line is approximately ± 3.8 V.

Having identified a desirable Q-point, we determine consistent biasing resistors with the help of the Thevenin circuit of Fig. 7.13.

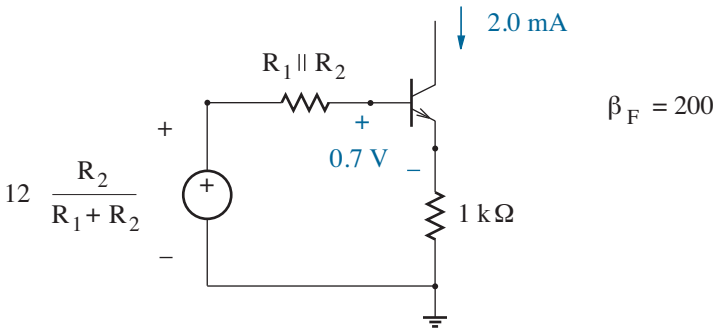


Figure 7.13: Thevenin biasing circuit for Example 7.4. Capacitors are open circuits for dc biasing calculations, so the two emitter resistors have been combined as a single 1-kΩ component.

The unspecified Thevenin resistance in Fig. 7.13 demands special care. If $R_1 \parallel R_2 \ll (\beta_F + 1)R_e$, bias stability is achieved with $i_c|_Q$ proportional to $\beta_F/(1 + \beta_F)$ — see Eq. 7.27. On the other hand, very small $R_1 \parallel R_2$ leads to large R_1 and R_2 currents (and difficulties we have yet to discover). So we make a design compromise:

$$R_1 \parallel R_2 = \frac{R_1 R_2}{R_1 + R_2} = \frac{(\beta_F + 1)}{10}\, R_e \approx 20 \text{ k}\Omega. \tag{7.41}$$

We now work backwards from beneath the emitter resistor to assign a numerical value to the Thevenin bias voltage. Subject to $i_c|_Q = 2.0$ mA, the node voltage at the emitter is approximately $2.0 \times 1 = 2.0$ V, the node voltage at the base is $2.0 + 0.7 = 2.7$ V, and the Thevenin bias voltage is $2.7 + (2/200) \times 20 = 2.9$ V. Thus,

$$12 \left(\frac{R_2}{R_1 + R_2} \right) = 2.9 \text{ V}. \tag{7.42}$$

Our last step is to divide Eq. 7.42 into Eq. 7.41 to obtain $R_1 = 83$ kΩ. Then we return to Eq. 7.41 (or Eq. 7.42) to find $R_2 = 26$ kΩ.

Imposed Biasing for Integrated Circuits

Analyzing either of the imposed biasing circuits of Fig. 7.14 is relatively straightforward. For the MOSFET, the gate-to-source voltage adjusts to a value consistent with $i_d = I_{bias}$. The gate node voltage is fixed at V_{ref}, so the source node voltage v_s is $V_{ref} - v_{gs}$. The BJT behavior is similar: $i_c = I_{bias}$, and v_e conforms to $V_{ref} - v_{be}$. However, the npn BJT also draws base current i_c/β_F from the reference supply.

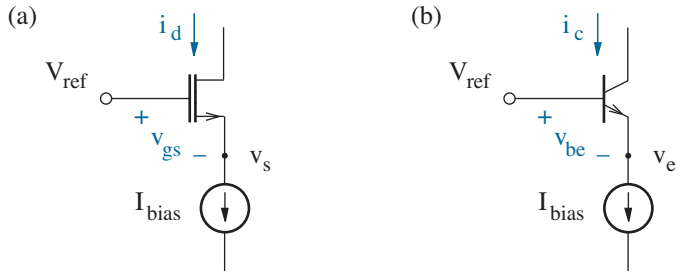


Figure 7.14: Imposed biasing: (a) n-channel MOSFET; (b) npn BJT.

CMOS analog integrated circuits often establish I_{bias} with the simple **current mirror** shown in Fig. 7.15 for n and p-channel implementations. A separate I_{bias} current source—of which more later—forces a particular v_{gs1} for M_1, and v_{gs2}, v_{gs3} ... are necessarily equivalent. If the MOSFETs are designed to have the same W/L ratios, $I_{bias} = i_{d1} = i_{d2} = i_{d3} = \ldots$ The integrated circuit designer also enjoys the freedom to assign different W/L ratios to achieve a set of scaled drain currents. All this is subject to the complication of channel-length modulation considered in Chapter 9.

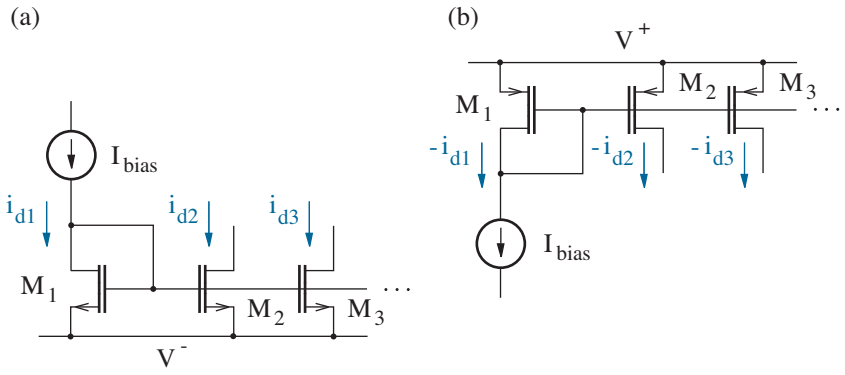


Figure 7.15: MOSFET current mirror: (a) n-channel; (b) p-channel.

The concepts that apply to the npn and pnp BJT current mirrors of Fig. 7.16 are much the same as those for the preceding MOSFET circuits. I_{bias} imposes a particular v_{be} that is the same for each device. If the emitter areas are equal, $i_{c1} = i_{c2} = i_{c3} = \ldots$. And once more, the designer is free to assign different emitter areas to achieve a set of scaled collector currents. Complications similar to those for the MOSFET will arise in Chapter 9.

Unfortunately, $i_{c1} \neq I_{bias}$ as a consequence of non-zero base currents. For example, an npn current mirror with N base-connected transistors features individual base currents that follow the relation

$$\frac{i_{c1}}{\beta_F} = i_{b1} = i_{b2} = i_{b3} = \ldots = i_{bN} \, . \tag{7.43}$$

So in consideration of Kirchhoff's current law at the Q_1 collector node,

$$i_{c1} = \frac{I_{bias}}{1 + N/\beta_F} \, , \tag{7.44}$$

which reduces to $i_{c1} \approx I_{bias}$ when $N \ll \beta_F$.

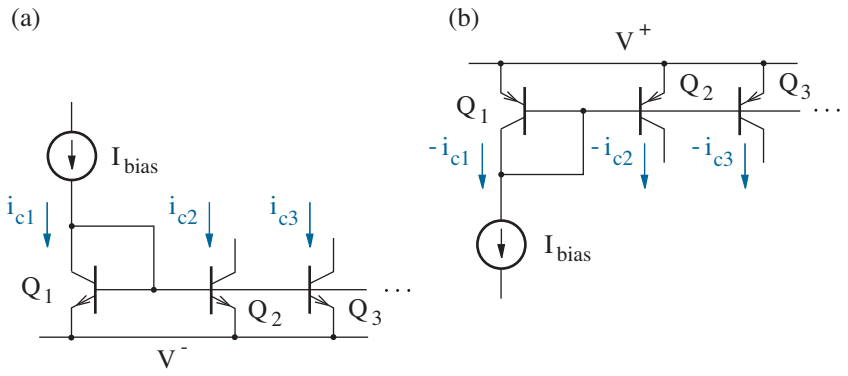


Figure 7.16: BJT current mirror: (a) npn; (b) pnp.

Now the astute reader would be right to complain that the MOSFET and BJT current-mirror constituents are simply slaves to an imposing I_{bias} current source that has yet to be implemented. Why bother to add extra transistors before even starting the I_{bias} design?

Part of the answer, which will be more fully appreciated in Chapter 9, is the desire to have a *single* special current source that mirrors into a set of proportional currents flowing downward toward the V^- rail (Fig. 7.15a) or emanating from the V^+ rail (Fig. 7.15b, after a second mirror action). The special current source can take a simple form that uses a single resistor. However, it is common to design a current source that is independent of supply voltage and resistant to temperature variations.

Example 7.5

Design a circuit so that two LEDs operate with 1- and 2-mA bias currents. Use the MPQ2222 npn transistor array shown in Fig. 7.17.

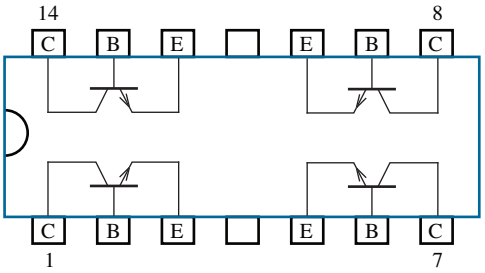


Figure 7.17: MPQ2222 npn transistor array.

Solution

We construct the BJT current mirror of Fig. 7.18. Current i mirrors to the collector current for Q_2, so $i = 1$ mA establishes $i_{c2} = 1$ mA for LED #1. The same mirror action applies to Q_3 and Q_4, so we connect the collectors in parallel to obtain 2 mA for LED #2. To find R, we assume $v_{be1} \approx 0.7$ V. Then $R = (5.0 - 0.7)\text{ V}/\,1\text{ mA} = 4.3\ \text{k}\Omega$ (a standard value).

How good is our design? SPICE data for the MPQ2222 indicate IS=14f. In turn, $v_{be1} = kT/q\,\ln(10^{-3}/14\times10^{-15}) \approx 0.65$ V. The LEDs are brighter. But who would care?

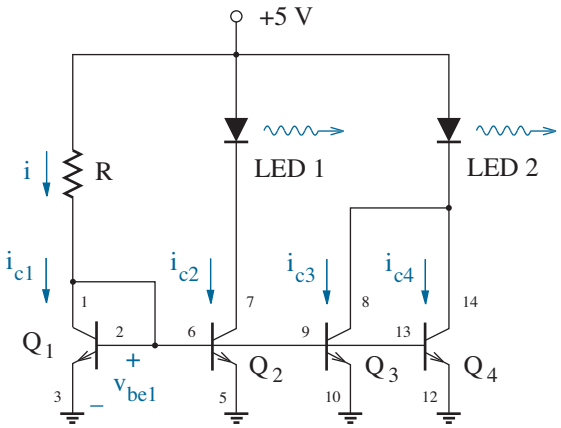


Figure 7.18: Circuit for Example 7.5 with pin connections for the MPQ2222.

7.2 Small-Signal Transistor Models

The small-signal transistor model is the foundation for the relation between input and output changes in an amplifier circuit. This section derives the first-order g_m and r_π model parameters used for the MOSFET and BJT. As shown in Fig. 7.19, the generic three-terminal hybrid-π transistor model also has a second-order resistance r_o in parallel with the dependent source. Chapter 9 explores the second-order complications that arise from finite r_o, and it derives a four-terminal model for MOSFETs with $v_{bs} \neq 0$.

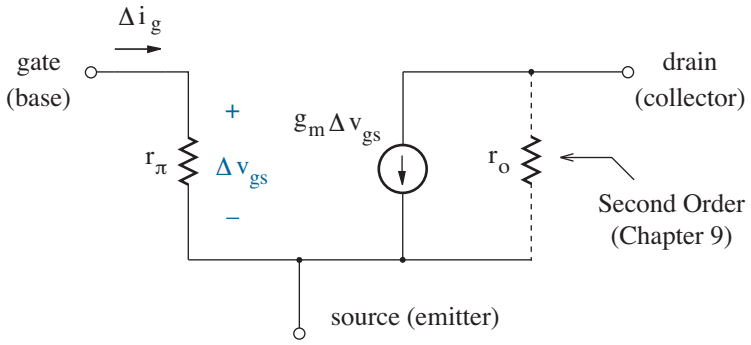


Figure 7.19: Small-signal hybrid-π model for a generic transistor.

MOSFET Small-Signal Model

The first-order small-signal model for the n-channel MOSFET in *saturation* features a Δv_{gs}-dependent current source with a transconductance that derives from Eq. 5.7:

$$g_m = \left.\frac{\partial i_d}{\partial v_{gs}}\right|_Q = K' \frac{W}{L}(v_{gs}|_Q - V_T) = \sqrt{2\,K' \frac{W}{L}\, i_d|_Q}\,. \tag{7.45}$$

In addition, we have

$$\frac{1}{r_\pi} = \left.\frac{\partial i_g}{\partial v_{gs}}\right|_Q = 0\,, \tag{7.46}$$

so r_π is infinite.

The p-channel MOSFET functions much like the n-channel MOSFET, but all currents and voltages are altered by a change in sign. Despite these modifications, the small-signal models are unchanged (see Problem 7.31).

BJT Small-Signal Model

The first-order small-signal model for the npn BJT in the *forward active mode* features a Δv_{be}-dependent current source with a transconductance that derives from Eq. 6.34:

$$g_m = \left.\frac{\partial i_c}{\partial v_{be}}\right|_Q = I_s \left(\frac{q}{kT}\right) \exp\left(\frac{q\, v_{be}|_Q}{kT}\right) \approx \frac{i_c|_Q}{kT/q} \,. \tag{7.47}$$

In addition, we have

$$\frac{1}{r_\pi} = \left.\frac{\partial i_b}{\partial v_{be}}\right|_Q = \left.\frac{\partial i_b}{\partial i_c}\right|_Q \left.\frac{\partial i_c}{\partial v_{be}}\right|_Q = \frac{1}{\beta_o} \left.\frac{\partial i_c}{\partial v_{be}}\right|_Q . \tag{7.48}$$

So with the help of Eq. 7.47,

$$r_\pi = \frac{\beta_o}{g_m} \,. \tag{7.49}$$

In deriving the preceding expressions, we have assumed that the reverse-biased base-collector junction does not contribute to the collector current. We have also introduced β_o as a small-signal (ac) parameter that indicates the ratio between Δi_c and Δi_b at the Q-point. *It is not equivalent to* β_F, a large-signal (dc) parameter that gives the ratio between $i_c|_Q$ and $i_b|_Q$. However, it is common to take $\beta_o \approx \beta_F$ in the absence of contrary data, and we will tend to do so unless otherwise indicated. BJTs that are subject to SPICE parameters such as IKF that alter β_F exhibit different results for g_m and β_o (see Problem 7.35). We reserve these complications for SPICE.

The pnp BJT functions like the npn BJT, but all currents and voltages are altered by a change in sign. As for the MOSFET, these modifications do not change the small-signal models.

Summary

Table 7.1 summarizes the first-order small-signal transistor parameters. Most readers will find it helpful to memorize the entries.

Table 7.1: First-order small-signal transistor parameters.

Device	g_m	r_π
MOSFET	$\sqrt{2\,K'\frac{W}{L}\,i_d\|_Q}$	∞
BJT	$\frac{i_c\|_Q}{kT/q}$	$\frac{\beta_o}{g_m}$

Model Comparisons

The last Perspective showed the benefit of large g_m, r_π, and r_o when seeking large voltage gain in a simple amplifier circuit. The MOSFET is obviously far superior to a BJT in terms of r_π (as for field-effect devices, generally). Less apparent are the relative device merits in terms of g_m.

Or so it would seem. For the BJT, we have

$$g_m = \frac{i_c|_Q}{kT/q} \,. \tag{7.50}$$

And for the MOSFET (after some algebraic manipulation),

$$g_m = \frac{2\, i_d|_Q}{v_{gs}|_Q - V_T} \,. \tag{7.51}$$

Thus,

$$\frac{g_m(\text{BJT})}{g_m(\text{MOSFET})} = \frac{i_c|_Q}{i_d|_Q} \frac{(v_{gs}|_Q - V_T)}{2kT/q} \,. \tag{7.52}$$

For a meaningful comparison, we examine the transconductance ratio when $i_c|_Q = i_d|_Q$. The MOSFET condition is unspecified. Nevertheless, it is reasonable to take $v_{gs}|_Q = 2$ V and $V_T = 1$ V. In turn, the transconductance ratio is 1 V / 52 mV ≈ 20. We are lead to conclude that the BJT is superior to the MOSFET in terms of g_m or "drive-current" capability.

Exercise 7.5 A MOSFET with $K'W/L = 2$ mA/V^2 and $V_T = 1$ V is biased so that $i_d|_Q = 10$ mA. Determine g_m and r_π.

Ans: $g_m = 6.32 \times 10^{-3}$ ℧, $r_\pi = \infty$

Exercise 7.6 A MOSFET with $K' = 50\ \mu$A/V^2 and $V_T = 0.5$ V is biased so that $i_d|_Q = 2$ mA. Find W/L so that $g_m = 8 \times 10^{-3}$ ℧.

Ans: $W/L = 320$

Exercise 7.7 A BJT with $\beta_F = \beta_o = 200$ is biased so that $i_c|_Q = 4$ mA. Determine g_m and r_π.

Ans: $g_m = 0.154$ ℧, $r_\pi = 1.30$ kΩ

Exercise 7.8 A BJT with $\beta_F = \beta_o = 100$ requires $g_m = 8 \times 10^{-3}$ ℧. Determine the necessary collector bias current.

Ans: $i_c|_Q = 0.21$ mA

A common convention for discrete BJTs warrants brief digression.

BJT h Parameters

In the study of circuit theory, we learn that the linear two-port shown in Fig. 7.20 can be described by expressing two terminal variables as a linear combination of the other two. There are four ways to do this. For example, in the h-parameter representation, we have

$$v_1 = h_i\, i_1 + h_r\, v_2\,, \tag{7.53a}$$
$$i_2 = h_f\, i_1 + h_o\, v_2\,. \tag{7.53b}$$

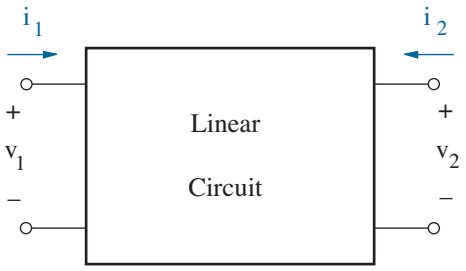


Figure 7.20: Linear two-port circuit.

Each of the h parameters has physical significance. Specifically,

$$h_f = \left.\frac{i_2}{i_1}\right|_{v_2=0} \tag{7.54}$$

is the **forward current gain** when the output is shorted;

$$h_o = \left.\frac{i_2}{v_2}\right|_{i_1=0} \tag{7.55}$$

is the **output admittance** when the input is open;

$$h_i = \left.\frac{v_1}{i_1}\right|_{v_2=0} \tag{7.56}$$

is the **input impedance** when the output is shorted; and,

$$h_r = \left.\frac{v_1}{v_2}\right|_{i_1=0} \tag{7.57}$$

is the **voltage feedback ratio** when the input is open. The h parameters are generally complex numbers, but they are real for a resistive two-port.

When applied to the linearized BJT, the h-parameter representation uses an additional subscript to indicate the device connection that is shared between the input and output for the purpose of parameter measurement. The small-signal hybrid-π model of Fig. 7.19 suggests the common emitter configuration, so we restrict our attention to h_{fe}, h_{oe}, h_{ie}, and h_{re}.

Now consider the general two-port equivalent circuit shown in Fig. 7.21. The circuit obeys the h-parameter relations of Eq. 7.53.

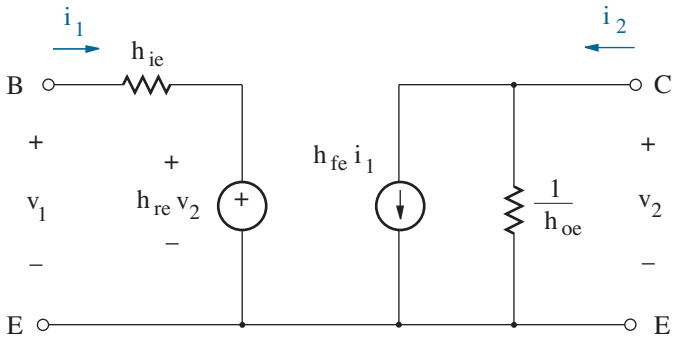


Figure 7.21: General equivalent circuit for a two-port with h-parameter characterization. Note the similarity to the hybrid-π small-signal model.

Our task at hand is to relate the individual h parameters to the small-signal parameters that were derived from the governing device equations. So we compare Fig. 7.21 with the first-order small-signal model of Fig. 7.19 (with subscripts appropriate to the BJT). Whereas $i_1 = \Delta i_b$,

$$h_{fe}\,\Delta i_b = g_m\,\Delta v_{be}\,. \tag{7.58}$$

But with $\Delta i_b = \Delta v_{be}/r_\pi$,

$$h_{fe} = g_m r_\pi = \beta_o\,. \tag{7.59}$$

Similar comparisons yield

$$h_{ie} = r_\pi\,. \tag{7.60}$$

We also have

$$h_{oe} = \frac{1}{r_o} = 0 \tag{7.61}$$

(barring the second-order considerations of Chapter 9), and

$$h_{re} = 0\,. \tag{7.62}$$

Equation 7.62 implies no feedback from the collector to the base.

Manufacturer's data sheets often provide the results of h-parameter measurements as a means of describing small-signal BJT characteristics. Figure 7.22 shows typical data for a randomly selected 2N3904 BJT. Of the four plots, we are mainly interested in $h_{fe} = \beta_o$. The other data are guides for BJT comparison, *not* substitutes for small-signal calculations.

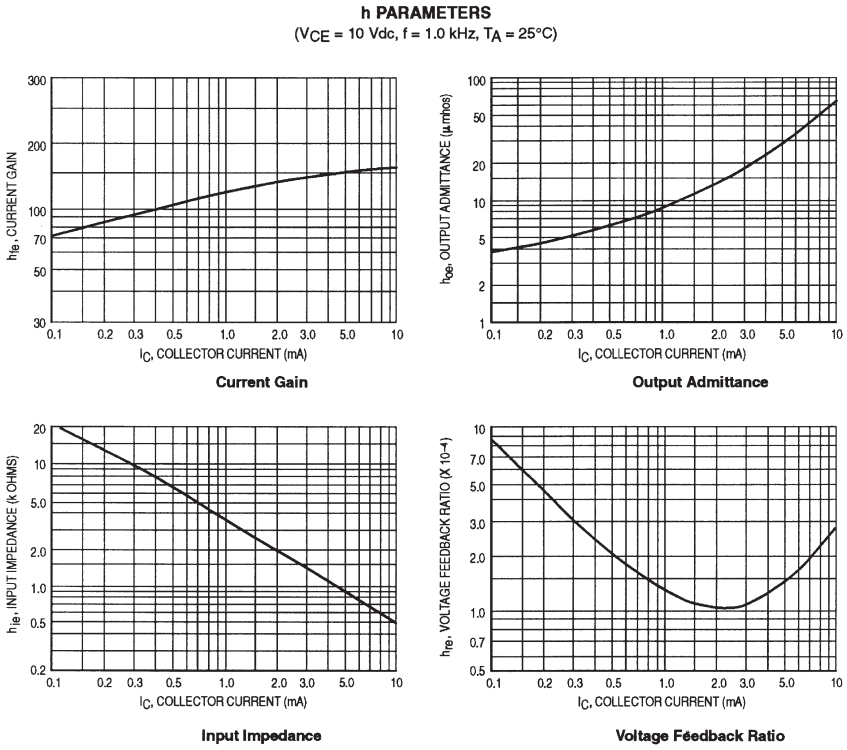


Figure 7.22: Manufacturer's h-parameter data for a typical 2N3904 BJT. Parameters vary with current i_c. Used with permission from SCILLC dba ON Semiconductor.

Note: Some data sheets specify β_F as h_{FE}. Upper-case subscripts generally denote large-signal characteristics, whereas lower-case subscripts denote small-signal characteristics.

Apart from the h parameters, one can apply g-, y-, and z-parameter relations to describe an arbitrary two-port. Of these, the y parameters are sometimes useful to describe discrete-MOSFET small-signal characteristics. The parameter sets are also convenient for high-frequency device analysis (Chapter 8) and feedback circuits (Chapters 11 and 12).

7.3 Small-Signal Amplifier Analysis

Despite their crucial function, the inherently dormant biasing circuits that establish Q-points of reference for the MOSFET, BJT, (and other devices) are not particularly exciting. Yet they quickly come alive with noteworthy dynamic behavior when we add an ac input excitation and an output load to form a one-transistor or **single-stage** amplifier as shown in Fig. 7.23. Here, the input and output are capacitively coupled so that the Q-point remains unaffected—the capacitors block dc current.

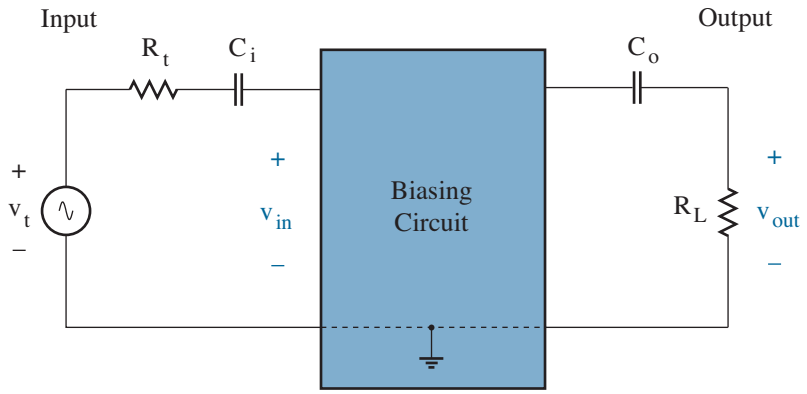


Figure 7.23: Circuit format for a single-stage amplifier with capacitively coupled input and output. The capacitors leave the Q-point unchanged.

Individual connections to a transistor biasing circuit are made according to the the following rules:

- One transistor node is the amplifier input.
- A second transistor node is the amplifier output.
- A third transistor node is shared between input and output, given its connection to ac ground, sometimes through a **bypass capacitor**.

The drain (collector) never serves as the input node, while the gate (base) never serves as the output node (see Problem 7.39). Thus, there are only three possible amplifier configurations:

- **common-source** (common-emitter),
- **common-gate** (common-base),
- **common-drain** (common-collector).

Figure 7.24 indicates the requisite node connections.

Amplifier Name	Terminal Connection: Source (Emitter)	Gate (Base)	Drain (Collector)
Common source (Common emitter)	ac ground	Input	Output
Common gate (Common base)	Input	ac ground	Output
Common drain (Common collector)	Output	Input	ac ground

Figure 7.24: Transistor node connections for basic single-stage amplifiers. Each amplifier is named for its ac ground connection.

In what follows, we examine the three basic single-stage amplifiers with the objective of tabulating expressions for ac voltage gain, input resistance, current gain, and output resistance. These expressions help to minimize future analytical effort, and they are useful as guides for amplifier design. We limit our discussion to **mid-frequency** conditions in which coupling and bypass capacitors function as effective short circuits for ac signals. Chapter 8 considers low-frequency conditions (in which coupling and bypass capacitors assume less conductive roles) and high-frequency conditions (in which small-signal transistor models display new dynamic behavior).

Section 7.2 showed that the first-order small-signal characteristics of any transistor could be described with a three-terminal circuit model featuring r_π and a voltage-dependent current source with transconductance value g_m —only the parametric expressions for r_π and g_m differed between devices. Thus, to avoid needless repetition, we derive each midfrequency amplifier characteristic using the generic model, and we determine results for specific transistors by taking appropriate limits. Of course we will need refinements: Second-order small-signal models feature resistance r_o, and the MOSFET is subject to the action of a fourth terminal under conditions when $v_{bs} \neq 0$. We mostly postpone these considerations to Chapter 9.

Before we proceed, it will prove helpful to adopt a special notation for the ac resistance to ac ground looking *away* from each transistor terminal. Specifically,

$R_d{}'$ is the ac resistance to ground looking away from the drain.

$R_g{}'$ is the ac resistance to ground looking away from the gate.

$R_s{}'$ is the ac resistance to ground looking away from the source.

Figure 7.25 shows the orientation of the looking-away primed resistances. The BJT counterparts are $R_c{}'$, $R_b{}'$, and $R_e{}'$.

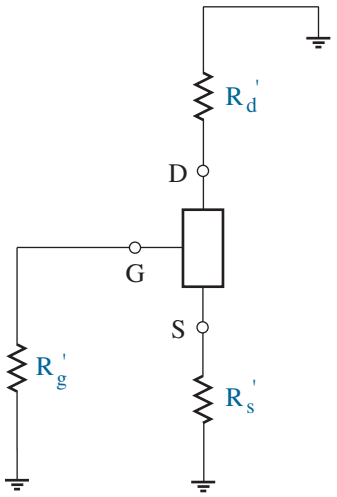


Figure 7.25: Primed ac resistances looking away from the drain, gate, and source transistor terminals.

Exercise 7.9 Determine $R_c{}'$, $R_b{}'$, and $R_e{}'$ for the circuit of Fig. 7.12.

Ans: $R_c{}' = R_c \parallel R_L = 1.5\ \text{k}\Omega \quad R_b{}' = R_1 \parallel R_2$

$R_e{}' = R_{e1} = 500\ \Omega$

Our motivation for a special notation is to simplify pending results and to provide insight for amplifier behavior.

We are fully prepared. Tables 7.2 and 7.3 summarize the results of the remainder of this Section. Do the exercises, but feel free to read quickly.

Common-Source (Common-Emitter) Amplifier

In relation to the "standard" transistor biasing circuit, the common-source (common-emitter) amplifier accepts an ac input signal at the gate (base), and it imparts an ac output signal at the drain (collector) —see Fig. 7.26. Bypass capacitor C_s (C_e) ensures a source (emitter) that is at ac ground.

Figure 7.27 shows the applicable generic small-signal equivalent circuit. It is sometimes desirable to eliminate the bypass capacitor, so we leave R_s.

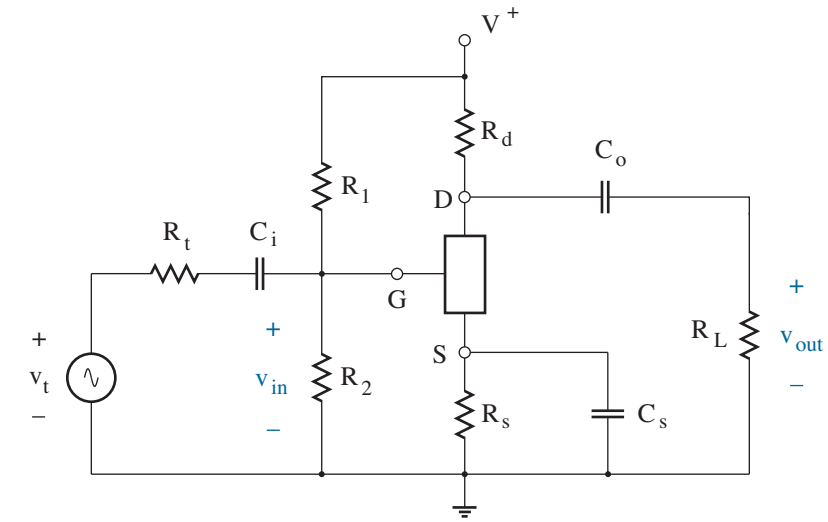


Figure 7.26: Common-source (common-emitter) amplifier configuration.

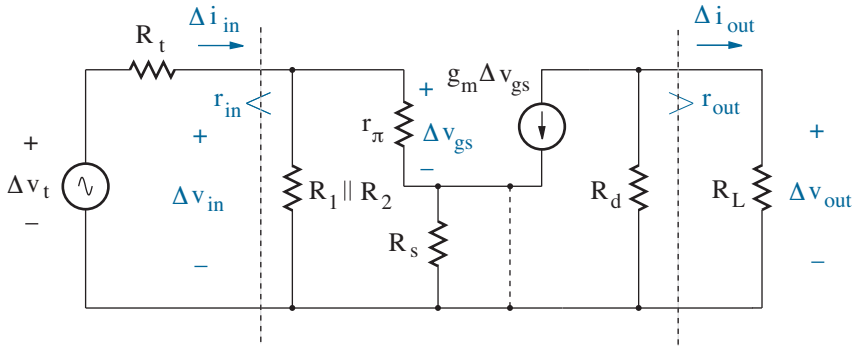


Figure 7.27: Small-signal ac circuit that represents the common-source (common-emitter) amplifier. R_s is shorted when C_s is present.

Voltage Gain

Currents through r_π and the dependent source are proportional to Δv_{gs}. So from KCL and Ohm's law, the voltage across R_s is given by

$$\Delta v_s = (g_m + g_\pi)\Delta v_{gs} R_s \,, \tag{7.63}$$

where $g_\pi = 1/r_\pi$. And from KVL,

$$\Delta v_{in} = \Delta v_{gs} + \Delta v_s \,. \tag{7.64}$$

We eliminate Δv_s from Eqs. 7.63 and 7.64 to obtain

$$\Delta v_{gs} = \frac{\Delta v_{in}}{1 + (g_m + g_\pi)R_s} \,. \tag{7.65}$$

In turn, Δv_{gs} controls the amplifier output. Specifically,

$$\Delta v_{out} = -g_m \Delta v_{gs}(R_d \parallel R_L) \,. \tag{7.66}$$

By definition, the midfrequency voltage gain is $A_{vm} = \Delta v_{out}/\Delta v_{in}$. Thus, with $R_d{}' = R_d \parallel R_L$ and $R_s{}' = R_s$, we have

$$A_{vm} = \frac{-g_m R_d{}'}{1 + (g_m + g_\pi)R_s{}'} = \frac{-g_m R_d{}'}{1 + g_m R_s{}' \left(1 + \dfrac{1}{g_m r_\pi}\right)} \,. \tag{7.67}$$

We use Eq. 7.67 to determine the midfrequency voltage gain for the MOSFET and BJT.

Case 1A - MOSFET (with source bypass capacitor)

This reflects a true common-source configuration. The MOSFET small-signal model features $r_\pi \to \infty$. And thanks to the bypass capacitor, $R_s{}' = 0$ (R_s is shorted). Then,

$$A_{vm} = -g_m R_d{}' \,. \tag{7.68}$$

Case 1B - MOSFET (no source bypass capacitor)

Again, we let $r_\pi \to \infty$, but $R_s{}' \neq 0$. Thus, we obtain

$$A_{vm} = \frac{-g_m R_d{}'}{1 + g_m R_s{}'} \,. \tag{7.69}$$

The absence of the bypass capacitor reduces the midfrequency voltage gain due to the $1/(1 + g_m R_s{}')$ factor, which is always less than unity.

Case 2A - BJT (with emitter bypass capacitor)

This reflects the true common-emitter configuration. A conventional circuit assumes $R_d \to R_c$ and $R_s \to R_e$. So with $g_m r_\pi = \beta_o$ and $R_e{}' = 0$ (R_e is shorted),

$$A_{vm} = -g_m R_c{}' \,. \tag{7.70}$$

Case 2B - BJT (no emitter bypass capacitor)

Again, we let $g_m r_\pi = \beta_o$, but $R_e{}' \neq 0$. Thus, we obtain

$$A_{vm} = \frac{-g_m R_c{}'}{1 + g_m R_e{}'(1 + 1/\beta_o)} \,. \tag{7.71}$$

The absence of the bypass capacitor reduces the midfrequency voltage gain. We also note that if $g_m R_e{}' \gg 1$,

$$A_{vm} \approx \frac{-R_c{}'}{R_e{}'} \,. \tag{7.72}$$

This relatively simple expression is useful for the design process.

If the absence of a source (emitter) bypass capacitor leads to reduced amplifier voltage gain, then why bother to remove it? The answer lies in the decreased voltage-gain sensitivity to g_m, which varies with temperature and the transistor fabrication process. Here we have yet another example of feedback stabilization that is provided by the source (emitter) resistor.

Finally, we observe that the common-source (common-emitter) circuit yields an **inverting amplifier** due to the negative sign in each expression for midfrequency voltage gain. Consequently, the amplifier output is 180^o out of phase with respect to the input. This is seldom a problem that warrants correction.

Input Resistance

The current Δi_{in} is given by

$$\Delta i_{in} = \frac{\Delta v_{in}}{R_1 \parallel R_2} + g_\pi \Delta v_{gs} \,. \tag{7.73}$$

And with the help of Eq. 7.65,

$$\Delta i_{in} = \frac{\Delta v_{in}}{R_1 \parallel R_2} + \frac{\Delta v_{in}}{r_\pi + (g_m r_\pi + 1) R_s} \,. \tag{7.74}$$

By definition, the midfrequency input resistance is $r_{in} = \Delta v_{in}/\Delta i_{in}$. Thus, with $R_s{}' = R_s$, we have

$$\frac{1}{r_{in}} = \frac{1}{R_1 \parallel R_2} + \frac{1}{r_\pi + (g_m r_\pi + 1){R_s}'} , \tag{7.75}$$

or

$$r_{in} = R_1 \parallel R_2 \parallel [r_\pi + (g_m r_\pi + 1){R_s}'] . \tag{7.76}$$

We use Eq. 7.76 to determine the midfrequency input resistance for the MOSFET and BJT.

Case 1 -MOSFET

Either with (Case 1A) or without (Case 1B) the source bypass capacitor, the input resistance reduces to

$$r_{in} = R_1 \parallel R_2 \tag{7.77}$$

in the limit as $r_\pi \to \infty$. This r_{in} can be quite large.

Case 2A - BJT (with emitter bypass capacitor)

We let $R_s \to R_e$. Nevertheless, ${R_e}' = 0$, and

$$r_{in} = R_1 \parallel R_2 \parallel r_\pi . \tag{7.78}$$

Typically, $R_1 \parallel R_2 \gg r_\pi$. Thus,

$$r_{in} \approx r_\pi . \tag{7.79}$$

Case 2B -BJT (no emitter bypass capacitor)

We let $g_m r_\pi = \beta_o$, but ${R_e}' \neq 0$. Then,

$$r_{in} = R_1 \parallel R_2 \parallel [r_\pi + (\beta_o + 1){R_e}'] . \tag{7.80}$$

In this case, we usually find $R_1 \parallel R_2 \ll [r_\pi + (\beta_o + 1){R_e}']$ (as in the design condition for bias stability). So now,

$$r_{in} \approx R_1 \parallel R_2 . \tag{7.81}$$

Comparing Eqs. 7.79 and 7.81, we find that the emitter bypass capacitor promotes large voltage gain, but at the expense of reduced input resistance. This trade-off does not occur for the MOSFET amplifier.

The input resistance is important since the full amplitude of the Thevenin signal v_t is not transferred to the amplifier input if $R_t \neq 0$. For the circuit of Fig. 7.26, we usually want to find the total midfrequency voltage gain, which is defined as

$$A_{vm}^{\text{Total}} = \frac{\Delta v_{out}}{\Delta v_t} = \frac{\Delta v_{out}}{\Delta v_{in}} \frac{\Delta v_{in}}{\Delta v_t} . \tag{7.82}$$

In this expression, $\Delta v_{out}/\Delta v_{in}$ is the midfrequency voltage gain that is *inherent* to the amplifier and thus independent of specific input conditions (such as finite R_t). The term $\Delta v_{in}/\Delta v_t$ is the **loading factor**, given by

$$\text{LF} = \frac{\Delta v_{in}}{\Delta v_t} = \frac{r_{in}}{r_{in} + R_t} . \tag{7.83}$$

So in general,

$$A_{vm}^{\text{Total}} = A_{vm} \times \text{LF} . \tag{7.84}$$

We aim to achieve large total voltage gain, in part, through a loading factor that is close to unity. This requires an amplifier for which $r_{in} \gg R_t$.

The specification of r_{in} is also important when we consider the analysis and design of multistage amplifiers.

Current Gain

The midfrequency current gain is defined as $A_{im} = \Delta i_{out}/\Delta i_{in}$, and it is easily determined once we have A_{vm} and r_{in}. At the amplifier output,

$$\Delta v_{out} = \Delta i_{out} R_L . \tag{7.85}$$

And at the amplifier input,

$$\Delta v_{in} = \Delta i_{in} r_{in} . \tag{7.86}$$

We divide Eq. 7.86 into Eq. 7.87:

$$\frac{\Delta v_{out}}{\Delta v_{in}} = \frac{\Delta i_{out}}{\Delta i_{in}} \frac{R_L}{r_{in}} . \tag{7.87}$$

Then we rearrange this expression to obtain the general relation

$$A_{im} = A_{vm} \frac{r_{in}}{R_L} . \tag{7.88}$$

The midfrequency current gain is independent of R_t (see Problem 7.47).

We use Eq. 7.88 and previously derived results to determine the midfrequency current gain for the MOSFET and BJT.

Case 1A - MOSFET (with source bypass capacitor)

$$A_{im} = -g_m(R_1 \parallel R_2)\,\frac{R_d}{R_d + R_L}\,. \tag{7.89}$$

Case 1B - MOSFET (no source bypass capacitor)

$$A_{im} = \frac{-g_m}{1 + g_m R_s{}'}\,(R_1 \parallel R_2)\,\frac{R_d}{R_d + R_L}\,. \tag{7.90}$$

Case 2A - BJT (with emitter bypass capacitor)

$$A_{im} = -g_m(R_1 \parallel R_2 \parallel r_\pi)\,\frac{R_c}{R_c + R_L}\,. \tag{7.91}$$

However, for $R_1 \parallel R_2 \gg r_\pi$,

$$A_{im} \approx -g_m r_\pi\,\frac{R_c}{R_c + R_L} = -\beta_o\,\frac{R_c}{R_c + R_L}\,. \tag{7.92}$$

Thus, in the $R_L \ll R_c$ limit, the maximum amplifier current gain is $-\beta_o$ (as expected from our discussion of common-emitter BJT characteristics in Chapter 6).

Case 2B - BJT (no emitter bypass capacitor)

$$\begin{aligned} A_{im} &= \frac{-g_m\{R_1 \parallel R_2 \parallel [r_\pi + (\beta_o + 1)R_e{}']\}}{1 + g_m R_e{}'(1 + 1/\beta_o)}\,\frac{R_c}{R_c + R_L} \\ &\approx -\beta_o\,\frac{R_1 \parallel R_2}{r_\pi + (\beta_o + 1)R_e{}'}\,\frac{R_c}{R_c + R_L}\,. \end{aligned} \tag{7.93}$$

If we apply the special bias-stability design condition whereby $R_1 \parallel R_2 \sim (\beta_F + 1)R_e/10$ (and we assume comparable β_F and β_o with $R_e{}' = R_e$),

$$A_{im} \approx \frac{-\beta_o}{10}\,\frac{R_c}{R_c + R_L}\,. \tag{7.94}$$

This is an order of magnitude less than the midfrequency current gain when the emitter bypass capacitor is present.

Output Resistance

To determine the midfrequency output resistance, we remove the load R_L, and we turn off (zero) the independent voltage source at the amplifier input. Then we apply a test voltage or current source across R_d. Figure 7.28 shows the measurement with Δv_{test}. We find Δi_{test}, and $r_{out} = \Delta v_{\text{test}}/\Delta i_{\text{test}}$.

It is important to note that the dependent current source with value $g_m \Delta v_{gs}$ cannot be turned off arbitrarily. However, from Eq. 7.65, we note that $\Delta v_{gs} = 0$ if $\Delta v_{in} = 0$ (as when $\Delta v_t = 0$). Thus, the dependent current source is an open circuit under the conditions at hand, and we are left with a single resistance R_d in parallel with the test source. The midfrequency output resistance is simply

$$r_{out} = R_d \tag{7.95}$$

for the case of a MOSFET (Cases 1A and 1B), or

$$r_{out} = R_c \tag{7.96}$$

for the case of a BJT (Cases 2A and 2B).

The midfrequency output resistance is important in the design process, where we seek to optimize power transfer to a given load. This problem is considered in Section 7.4.

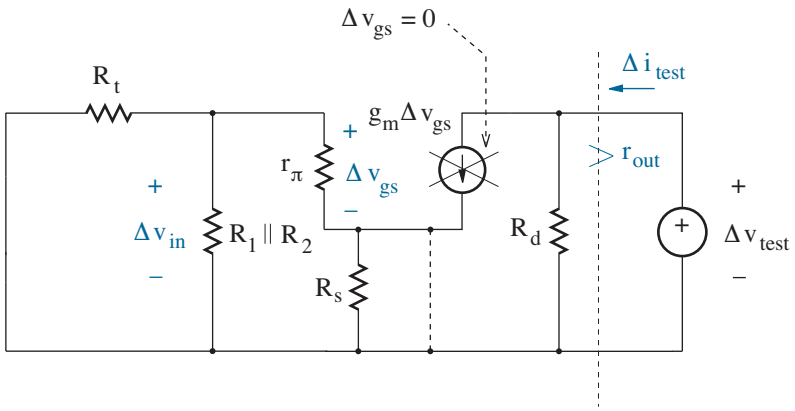


Figure 7.28: Measurement procedure for determining output resistance.

All of the preceding characteristics of the common-source or common-emitter amplifier are listed in Tables 7.2 and 7.3 at the end of this section. Take a moment to examine each entry.

Second-order Effects

Second-order small-signal transistor models include the shunt resistance r_o in parallel with the $g_m \Delta v_{gs}$ dependent current source. When the source (emitter) resistor is bypassed, the first-order amplifier characteristics are easily modified—we simply replace $R_d{}'$ ($R_c{}'$) with $R_d{}' \parallel r_o$ ($R_c{}' \parallel r_o$). Finite r_o is thus a limiting factor for midfrequency voltage gain. The output resistance is similarly limited: $R_d \parallel r_o$ ($R_c \parallel r_o$). When the source (emitter) resistor is not bypassed, the amplifier characteristics are more complicated. We defer further discussion to Chapter 9.

Example 7.6

Determine the total midfrequency voltage gain for the MOSFET common-source amplifier of Fig. 7.29 with and without the source bypass capacitor.

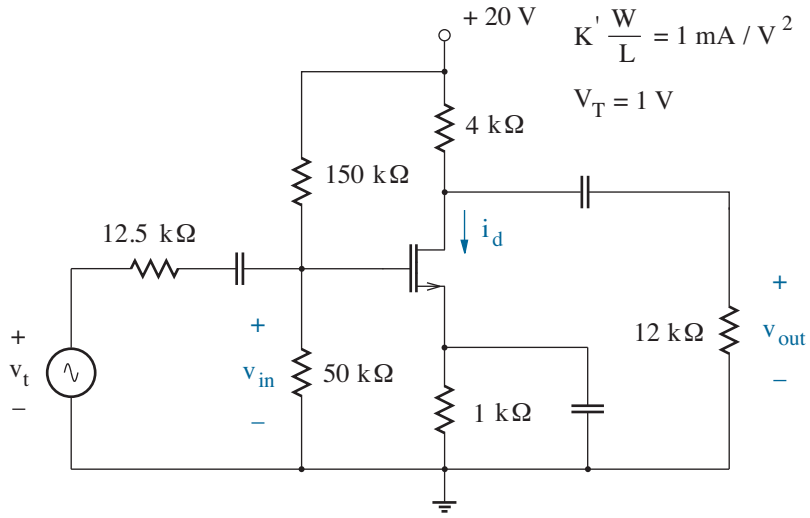


Figure 7.29: Circuit for Example 7.6.

Solution

We first determine the Q-point. Briefly, we assume that the MOSFET has a saturated drain current given by

$$i_d = \frac{1}{2}\,(1)\,(v_{gs} - 1)^2. \tag{7.97}$$

Meanwhile, the gate-to-source voltage is

$$v_{gs} = 20\,\left(\frac{50}{150+50}\right) - 1\,i_d = 5 - i_d\,. \tag{7.98}$$

We eliminate i_d to obtain a quadratic equation,

$$0.5\,{v_{gs}}^2 - 4.5 = 0\,, \tag{7.99}$$

which we easily solve to obtain $v_{gs}|_Q = +3$ V or $v_{gs}|_Q = -3$ V. The latter solution is less than the MOSFET threshold voltage. So with $v_{gs}|_Q = 3$ V, we use Eq. 7.97 to find $i_d|_Q = 2$ mA. This is consistent with the assumed MOSFET saturation condition since $v_{ds}|_Q = 20 - 2 \times (4 + 1) = 10$ V $> v_{ds,sat} = v_{gs} - V_T = 2$ V.

Next, we use the Q-point data to determine the MOSFET small-signal transconductance. Specifically,

$$g_m = \sqrt{2\,K'\,\frac{W}{L}\,i_d|_Q} = \sqrt{2 \times 1\ \text{mA/V}^2 \times 2\ \text{mA}} = 2 \times 10^{-3}\ \mho\,. \tag{7.100}$$

We need not consider r_π, since it is infinite.

When the source bypass capacitor is present, the amplifier features an inherent midfrequency voltage gain given by Eq. 7.68. The ac resistance looking away from the MOSFET drain is ${R_d}' = 4\ \text{k}\Omega \parallel 12\ \text{k}\Omega = 3\ \text{k}\Omega$. Thus,

$$A_{vm} = \frac{\Delta v_{out}}{\Delta v_{in}} = -2 \times 10^{-3}\ \mho \times 3 \times 10^3\ \Omega = -6\,. \tag{7.101}$$

However, this is not the total midfrequency voltage gain. We must also consider the midfrequency input resistance, which partially determines the amplifier loading factor. For this example, $r_{in} = 150\,\text{k}\Omega \parallel 50\,\text{k}\Omega = 37.5\,\text{k}\Omega$ (Eq. 7.77) and LF $= 37.5/(37.5 + 12.5) = 0.75$ (Eq. 7.83). So now

$$A_{vm}^{\text{Total}} = \frac{\Delta v_{out}}{\Delta v_t} = -6 \times 0.75 = -4.5\,. \tag{7.102}$$

When the source bypass capacitor is absent, the amplifier features an inherent midfrequency voltage gain given by Eq. 7.69. The ac resistance looking away from the MOSFET source is ${R_s}' = 1\ \text{k}\Omega$. In turn,

$$A_{vm} = \frac{\Delta v_{out}}{\Delta v_{in}} = \left(\frac{-6}{1 + 2 \times 10^{-3}\ \mho \times 1 \times 10^3\ \Omega}\right) = -2\,. \tag{7.103}$$

The midfrequency input resistance and loading factor remain unchanged. Thus, the total midfrequency voltage gain is

$$A_{vm}^{\text{Total}} = \frac{\Delta v_{out}}{\Delta v_t} = -2 \times 0.75 = -1.5\,. \tag{7.104}$$

Example 7.7

Determine the total midfrequency voltage gain for the BJT common-emitter amplifier of Fig. 7.30 with and without the emitter bypass capacitor.

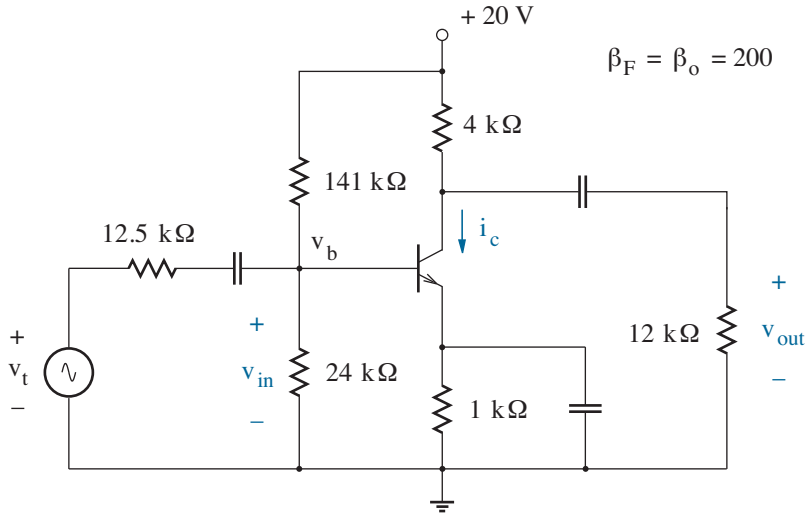


Figure 7.30: Circuit for Example 7.7.

Solution

We first determine the Q-point. If the BJT did not absorb any base current, the node voltage v_b could be obtained by using a simple divider relation. But this condition does not apply to the expected forward active mode—the effect of finite base current is to reduce v_b from its open-circuit value. Thus, we have

$$v_b = 20\left(\frac{24}{141+24}\right) - (141 \parallel 24)\, i_b = 2.9 - 20.5\, i_b\,. \tag{7.105}$$

On the other hand, with $i_e = (\beta_F + 1)\, i_b$ and $v_{be} \approx 0.7$ V,

$$v_b = 1 \times 201\, i_b + 0.7\,. \tag{7.106}$$

So we eliminate v_b from Eqs. 7.105 and 7.106 to obtain $i_b|_Q = 0.01$ mA. Then $i_c|_Q = \beta_F\, i_b|_Q = 2.0$ mA. This is consistent with the forward-active mode since $v_{ce}|_Q \approx 20 - 2.0 \times (4+1) = 10$ V $> v_{ce,sat} \sim 0.2$ V.

Take a moment to compare this Q-point solution technique with that of Example 7.3, which called for an explicit Thevenin biasing circuit.

Our next step is to use the Q-point data to determine the BJT small-signal transconductance and base-emitter resistance. Specifically,

$$g_m = \frac{i_c|_Q}{kT/q} = \frac{2.0\,\text{mA}}{25.9\,\text{mV}} = 77.2 \times 10^{-3}\,\text{℧}\,, \tag{7.107}$$

and

$$r_\pi = \frac{\beta_o}{g_m} = \frac{200}{77.2 \times 10^{-3}} = 2.59\,\text{k}\Omega\,. \tag{7.108}$$

When the emitter bypass capacitor is present, the amplifier features an inherent midfrequency voltage gain given by Eq. 7.70. The ac resistance looking away from the BJT emitter is $R_c{}' = 4\ \text{k}\Omega \parallel 12\ \text{k}\Omega = 3\ \text{k}\Omega$. Thus,

$$A_{vm} = \frac{\Delta v_{out}}{\Delta v_{in}} = -77.2 \times 10^{-3}\,\text{℧} \times 3 \times 10^3\,\Omega = -232\,. \tag{7.109}$$

The midfrequency input resistance partially contributes to the amplifier loading factor. For this example, $r_{in} = 141\,\text{k}\Omega \parallel 24\,\text{k}\Omega \parallel 2.59\,\text{k}\Omega = 2.30\,\text{k}\Omega$ (Eq. 7.78) and $\text{LF} = 2.30\,/(2.30 + 12.5) = 0.155$ (Eq. 7.83). So now,

$$A_{vm}^{\text{Total}} = \frac{\Delta v_{out}}{\Delta v_t} = -232 \times 0.155 = -36\,. \tag{7.110}$$

The severe degree of input loading suggests an inefficient amplifier.

When the emitter bypass capacitor is absent, the amplifier features an inherent midfrequency voltage gain given by Eq. 7.71. The ac resistance looking away from the BJT emitter is $R_e{}' = 1\ \text{k}\Omega$. In turn,

$$A_{vm} = \frac{\Delta v_{out}}{\Delta v_{in}} = \left[\frac{-232}{1 + 77.2 \times 10^{-3}\,\text{℧} \times 1 \times 10^3\,\Omega \times (1 + 1/200)}\right] = -2.95\,. \tag{7.111}$$

The midfrequency input resistance and loading factor are also changed. These are: $r_{in} = 141\ \text{k}\Omega \parallel 24\ \text{k}\Omega \parallel [2.59 + (200 + 1) \times 1]\,\text{k}\Omega = 18.6\,\text{k}\Omega$ (Eq. 7.80) and $\text{LF} = 18.6\,/(18.6 + 12.5) = 0.598$ (Eq. 7.83), respectively. Thus, the total midfrequency voltage gain is

$$A_{vm}^{\text{Total}} = \frac{\Delta v_{out}}{\Delta v_t} = -2.95 \times 0.598 = -1.8\,. \tag{7.112}$$

The input loading is less severe, but it is still significant.

It is interesting to compare the result of Eq. 7.112 with that obtained using the approximate expressions for A_{vm} and r_{in} (Eqs. 7.72 and 7.81). Here one finds $A_{vm}^{\text{Total}} \approx -3.00 \times 0.621 = -1.8$ —not a bad estimate at all when two significant figures of accuracy suffice.

Example 7.8

Use SPICE to examine the total midfrequency voltage gain as a function of temperature for the "common-source" circuits of Examples 7.6 and 7.7. Consider the range $-55\,^{o}\text{C} < \text{T} < 125\,^{o}\text{C}$.

Solution

The simulation code for the MOSFET common-source amplifier (with source bypass capacitor) takes the following form:

```
* Common-Source MOSFET Amplifier (with bypass capacitor)
* Temperature Dependence of Midfrequency Voltage Gain

V+      4       0       20
Vt      1       0       ac      1m
Rt      1       2       12.5K
R1      4       3       150K
R2      3       0       50K
Rd      4       5       4K
Rs      6       0       1K
RL      7       0       12K
Ci      2       3       1
Co      5       7       1
Cs      6       0       1
* Remove preceding line for unbypassed case.
M1      5       3       6       0       MOSN

.model  MOSN    NMOS    (KP=1m, VTO=1)

.ac     LIN     1       1K      1K
.temp   -55  -35  -15  5  25  45  65  85  105  125
.print  AC      v(7)    v(1)

.end
```

Here, input v_t is specified as a 1-kHz ac signal source with 1-mV amplitude. The frequency condition is imposed by the .ac statement, which forces a one-point sweep that begins and ends at 1 kHz. (We assume that the very large coupling and bypass capacitors effectively function as ac short circuits at this frequency.) The .temp command initiates a *series* of simulation runs at various temperatures within the range of interest. After each simulation, the .print command passes ac node amplitudes (7) and (1) to the output file. The ratio of the amplitude values is the total midfrequency voltage gain. Finally, the MOSFET $K'W/L$ and V_T temperature dependence follows the SPICE Level-1 behavior outlined in Chapter 5.

The SPICE simulation code for the BJT common-emitter amplifier (with emitter bypass capacitor) has a similar form apart from the BJT .model statement:

```
.model    BJT    NPN    (IS=5f, BF=200, XTB=1.15)
```

The particular XTB value provides a realistic β_o temperature variation (as discussed in Chapter 6).

Here are the results of the simulations:

Total Midfrequency Voltage Gain

T ^{o}C	MOSFET with C_s	MOSFET no C_s	BJT with C_e	BJT no C_e
-55	-6.064	-1.641	-28.05	-1.736
-35	-5.592	-1.604	-30.20	-1.745
-15	-5.188	-1.569	-32.22	-1.752
5	-4.836	-1.536	-34.10	-1.758
25	-4.529	-1.503	-35.85	-1.763
45	-4.257	-1.472	-37.47	-1.768
65	-4.015	-1.442	-38.97	-1.771
85	-3.799	-1.413	-40.34	-1.775
105	-3.603	-1.385	-41.60	-1.778
125	-3.427	-1.358	-42.76	-1.780

The improved stability in the absence of a bypass capacitor is obvious, particularly at low temperatures. However, the stability is achieved at the expense of midfrequency voltage gain.

Observe that the R_s-bypassed MOSFET amplifier gain decreases with increasing temperature as a consequence of $K'W/L$, which varies as $T^{-3/2}$. The temperature dependence on the threshold voltage is less significant.

The temperature dependence of the R_e-bypassed BJT amplifier gain is more difficult to explain. Transconductance g_m decreases with increasing temperature, which decreases A_{vm}. On the other hand, r_π increases with increasing temperature— β_o increases due to the positive XTB parameter in the .model statement, and $r_\pi = \beta_o/g_m$. In turn, this increases the loading factor, which is nearly proportional to r_π (provided that $R_t \gg r_\pi$). The r_π and loading-factor temperature changes are clearly dominant.

Exercise 7.10 Find A_{vm}, r_{in}, and A_{vm}^{Total} for the amplifier of Fig. 7.31.

Ans: $A_{vm} = -6.00$, $r_{in} = 37.5$ kΩ, $A_{vm}^{\text{Total}} = -5.3$

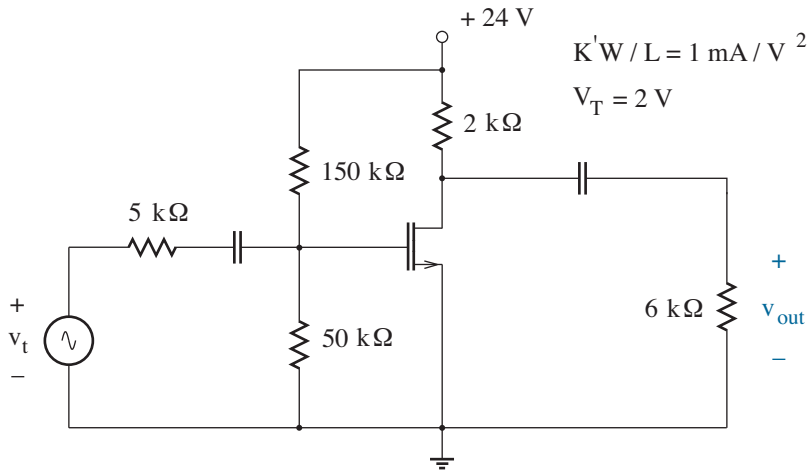


Figure 7.31: Circuit for Exercise 7.10.

Exercise 7.11 Find A_{vm}, r_{in}, and A_{vm}^{Total} for the amplifier of Fig. 7.32.

Ans: $A_{vm} = -5.06$, $r_{in} = 16.5$ kΩ, $A_{vm}^{\text{Total}} = -3.9$

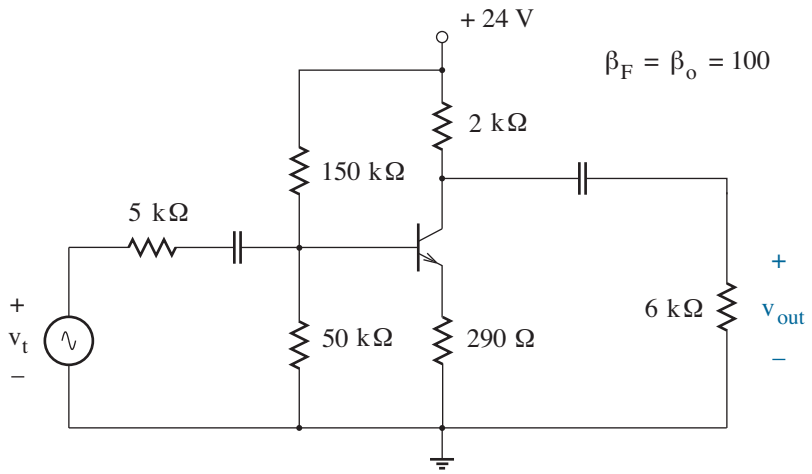


Figure 7.32: Circuit for Exercise 7.11.

Common-Drain (Common-Collector) Amplifier

In relation to the "standard" transistor biasing circuit, the common-drain (common-collector) amplifier accepts an ac input signal at the gate (base), and it imparts an ac output signal at the source (emitter) —see Fig. 7.33. Note that the drain (collector) node is tied directly to the V^+ power supply, which serves as an ac ground. A drain (collector) resistor is not necessary.

Figure 7.34 shows the applicable generic small-signal equivalent circuit.

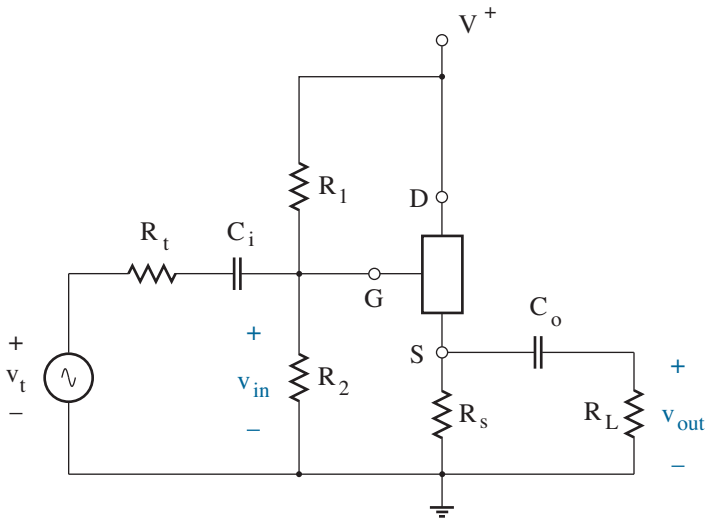


Figure 7.33: Common-drain (common-collector) amplifier configuration.

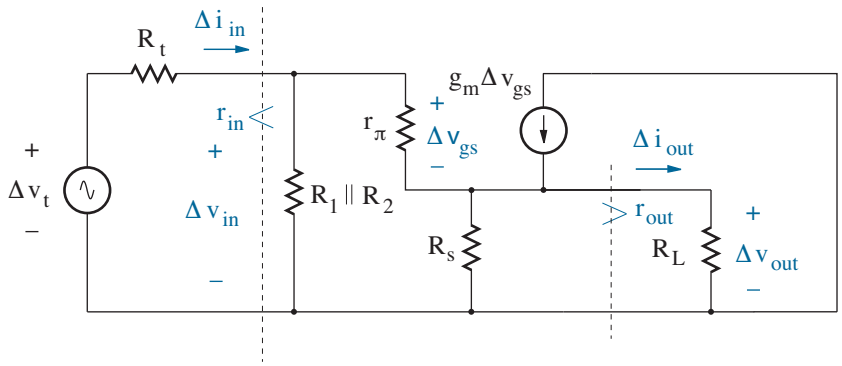


Figure 7.34: Small-signal ac circuit that represents the common-drain (common-collector) amplifier.

Voltage Gain

The voltage across $R_s{}' = R_s \parallel R_L$ is given by

$$\Delta v_{out} = (g_m + g_\pi)\Delta v_{gs} R_s{}' \,, \tag{7.113}$$

where $g_\pi = 1/r_\pi$. And from KVL,

$$\Delta v_{in} = \Delta v_{gs} + \Delta v_{out} \,. \tag{7.114}$$

We eliminate Δv_{out} from Eqs. 7.113 and 7.114 to find

$$\Delta v_{gs} = \frac{\Delta v_{in}}{1 + (g_m + g_\pi)R_s{}'} \,. \tag{7.115}$$

Then returning to Eq. 7.113, and with $A_{vm} = \Delta v_{out}/\Delta v_{in}$,

$$A_{vm} = \frac{(g_m + g_\pi)R_s{}'}{1 + (g_m + g_\pi)R_s{}'} = \frac{g_m R_s{}'(1 + 1/g_m r_\pi)}{1 + g_m R_s{}'(1 + 1/g_m r_\pi)} \,. \tag{7.116}$$

Equation 7.116 applies to the MOSFET and BJT.

Case 1 - MOSFET

We let $r_\pi \to \infty$. Then,

$$A_{vm} = \frac{g_m R_s{}'}{1 + g_m R_s{}'} \,, \tag{7.117}$$

which is close to unity if $g_m R_s{}' \gg 1$. For this reason, the common-drain amplifier is sometimes called a **source follower**. The output at the source follows the input at the gate.

Case 2 - BJT

We let $R_s \to R_e$ and $g_m r_\pi = \beta_o$. Then,

$$A_{vm} = \frac{g_m R_e{}'(1 + 1/\beta_o)}{1 + g_m R_e{}'(1 + 1/\beta_o)} \,. \tag{7.118}$$

Nevertheless, if $g_m R_e{}' \gg 1$,

$$A_{vm} \approx 1 \,. \tag{7.119}$$

The common-collector amplifier is sometimes called an **emitter follower** —the output at the emitter follows the input at the base.

Small voltage gains are not encouraging. But read on.

Input Resistance

The current Δi_{in} is given by

$$\Delta i_{in} = \frac{\Delta v_{in}}{R_1 \parallel R_2} + g_\pi \, \Delta v_{gs} \, . \tag{7.120}$$

And with the help of Eq. 7.115,

$$\Delta i_{in} = \frac{\Delta v_{in}}{R_1 \parallel R_2} + \frac{\Delta v_{in}}{r_\pi + (g_m r_\pi + 1) R_s{}'} \, . \tag{7.121}$$

It follows that

$$\frac{\Delta i_{in}}{\Delta v_{in}} = \frac{1}{r_{in}} = \frac{1}{R_1 \parallel R_2} + \frac{1}{r_\pi + (g_m r_\pi + 1) R_s{}'} \, , \tag{7.122}$$

or

$$r_{in} = R_1 \parallel R_2 \parallel [\, r_\pi + (g_m r_\pi + 1) R_s{}' \,] \, . \tag{7.123}$$

Equation 7.123 applies to the MOSFET and BJT.

Case 1 - MOSFET

In the limit as $r_\pi \to \infty$,

$$r_{in} = R_1 \parallel R_2 \, . \tag{7.124}$$

This is the same input resistance derived for the common-source amplifier. There is no dependence on the load R_L.

Case 2 - BJT

We let $R_s \to R_e$ and $g_m r_\pi = \beta_o$. Then,

$$r_{in} = R_1 \parallel R_2 \parallel [\, r_\pi + (\beta_o + 1) R_e{}' \,] \, . \tag{7.125}$$

Often, $r_\pi \ll (1 + \beta_o) R_e{}' \ll R_1 \parallel R_2$, and $R_L \ll R_e$. Thus, the approximate midfrequency input resistance is given by

$$r_{in} \approx (\beta_o + 1) R_L \, . \tag{7.126}$$

In turn, we have a potentially useful amplifier characteristic: The common-collector amplifier "buffers" the Thevenin signal source from the load R_L, effectively multiplying the latter by a factor of $(\beta_o + 1)$.

Output Resistance

To calculate the midfrequency output resistance, we disconnect load R_L, and we turn off the independent voltage source at the input. Then we apply a test source across R_s (R_e) as shown in Fig. 7.35. In the case of a test source with voltage Δv_{test}, we find Δi_{test}, and $r_{out} = \Delta v_{\text{test}}/\Delta i_{\text{test}}$.

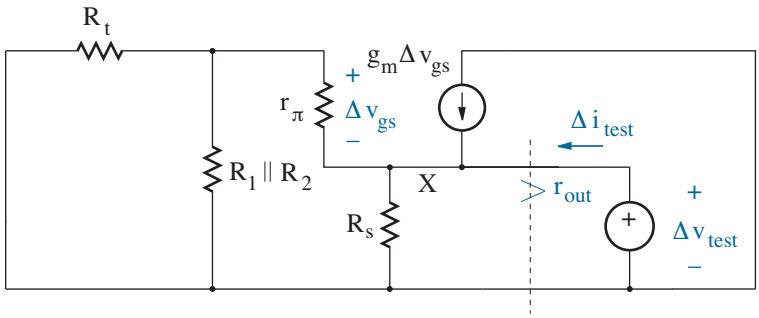


Figure 7.35: Measurement procedure for determining output resistance.

Once again, we note that the dependent current source with value $g_m \Delta v_{gs}$ cannot be turned off arbitrarily. Subject to $R_g{}' = R_1 \parallel R_2 \parallel R_t$, we write a KVL equation around the left loop:

$$\Delta v_{gs} + \Delta v_{\text{test}} = -g_\pi \Delta v_{gs} R_g{}' \,. \tag{7.127}$$

In turn, we solve for Δv_{gs} to find

$$\Delta v_{gs} = \frac{-\Delta v_{\text{test}}}{1 + g_\pi R_g{}'} \,. \tag{7.128}$$

Next, we sum the currents into node X. Specifically,

$$\Delta i_{\text{test}} + g_m \Delta v_{gs} + g_\pi \Delta v_{gs} - \frac{\Delta v_{\text{test}}}{R_s} = 0 \,. \tag{7.129}$$

Then we go back to Eq. 7.128 to obtain

$$\Delta i_{\text{test}} = \frac{\Delta v_{\text{test}}}{R_s} + \frac{-\Delta v_{\text{test}}(g_m + g_\pi)}{1 + g_\pi R_g{}'} \,. \tag{7.130}$$

Thus,

$$\frac{\Delta i_{\text{test}}}{\Delta v_{\text{test}}} = \frac{1}{r_{out}} = \frac{1}{R_s} + \frac{g_m r_\pi + 1}{r_\pi + R_g{}'} \,, \tag{7.131}$$

or

$$r_{out} = R_s \parallel \left(\frac{r_\pi + R_g{}'}{g_m r_\pi + 1} \right) = R_s \parallel \left[\frac{1 + R_g{}'/r_\pi}{g_m (1 + 1/g_m r_\pi)} \right] \,. \tag{7.132}$$

Equation 7.132 applies to the MOSFET and BJT.

Case 1 - MOSFET

We let $r_\pi \to \infty$. Then,

$$r_{out} = R_s \parallel \frac{1}{g_m} \,. \tag{7.133}$$

Case 2 BJT

We let $g_m r_\pi = \beta_o$. Then with $R_g{}' \to R_b{}' = R_1 \parallel R_2 \parallel R_t$,

$$r_{out} = R_e \parallel \left[\frac{1 + R_b{}'/r_\pi}{g_m(1 + 1/\beta_o)} \right] . \tag{7.134}$$

In this case, r_{out} is relatively small.

Current Gain

The midfrequency current gain for the common-drain (common-collector) amplifier derives from the general result of Eq. 7.88.

Case 1 - MOSFET

$$A_{im} = \frac{-g_m}{1 + g_m R_s{}'} \, (R_1 \parallel R_2) \, \frac{R_s}{R_s + R_L} \,. \tag{7.135}$$

Case 2 - BJT

After some tedious algebra,

$$A_{im} = \frac{-g_m(1 + 1/\beta_o)}{1 + g_m R_e{}'(1 + 1/\beta_o) + \dfrac{R_1 \parallel R_2}{r_\pi}} \, (R_1 \parallel R_2) \, \frac{R_e}{R_e + R_L} \,. \tag{7.136}$$

Second-order Effects

The second-order resistance r_o that appears in parallel with the $g_m \Delta v_{gs}$ dependent current source also appears in parallel with R_s (R_e) in the small-signal circuit of Fig. 7.34. Usually, $R_s \ll r_o$ ($R_e \ll r_o$), so the influence of r_o is negligible.

All of the preceding characteristics of the common-drain or common-collector amplifier are listed in Tables 7.2 and 7.3 at the end of this section. Take a moment to examine each entry.

Exercise 7.12 Find A_{vm}, r_{in}, and A_{vm}^{Total} for the amplifier of Fig. 7.36.

Ans: $A_{vm} = 0.667$, $r_{in} = 7.5$ kΩ, $A_{vm}^{\text{Total}} = 0.53$

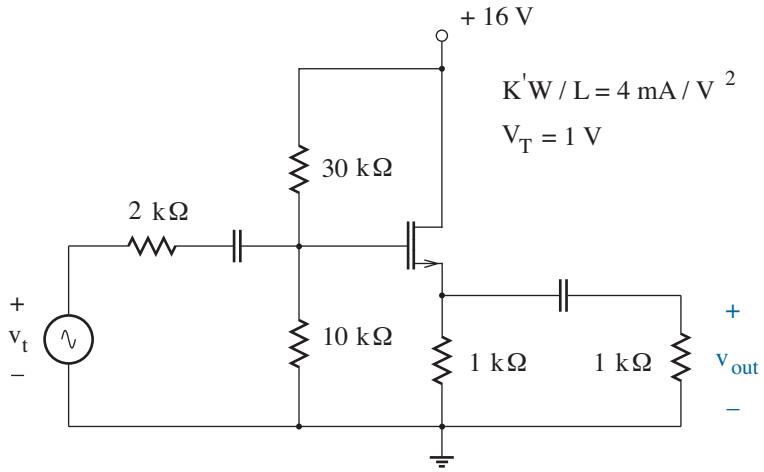


Figure 7.36: Circuit for Exercise 7.12.

Exercise 7.13 Find A_{vm}, r_{in}, and A_{vm}^{Total} for the amplifier of Fig. 7.37.

Ans: $A_{vm} = 0.990$, $r_{in} = 832\ \Omega$, $A_{vm}^{\text{Total}} = 0.88$

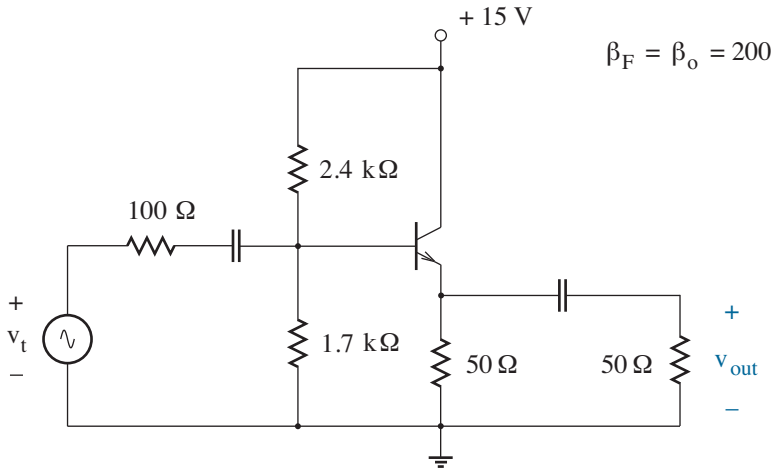


Figure 7.37: Circuit for Exercise 7.13.

Common-Gate (Common-Base) Amplifier

In relation to the "standard" transistor biasing circuit, the common-gate (common-base) amplifier accepts an ac input signal at the source (emitter), and it imparts an ac output signal at the drain (collector) —see Fig. 7.38. The gate (base) is connected to ac ground through capacitor C_g, either to actual ground, as shown, or to the supply voltage V^+.

Figure 7.39 shows the applicable generic small-signal equivalent circuit.

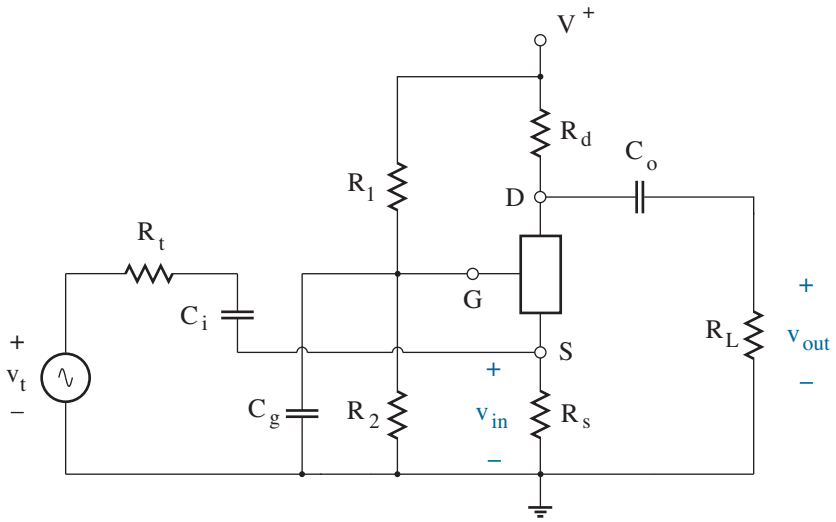


Figure 7.38: Common-gate (common-base) amplifier configuration.

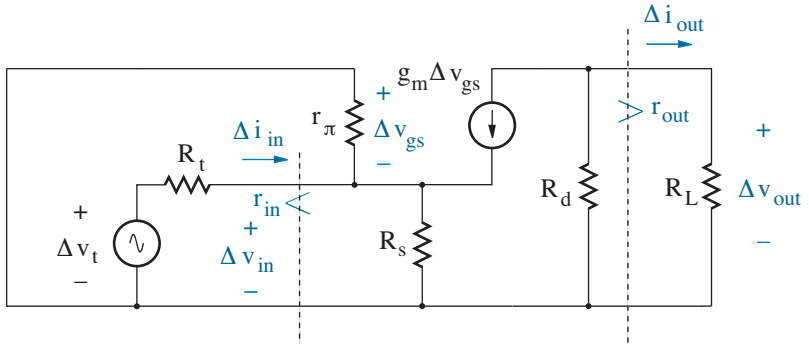


Figure 7.39: Small-signal ac circuit that represents the common-gate (common-base) amplifier.

Voltage Gain

The uppermost r_π terminal is connected to ground, so

$$\Delta v_{gs} = -\Delta v_{in} \, . \tag{7.137}$$

Thus with $R_d{}' = R_d \parallel R_L$,

$$\Delta v_{out} = -g_m \Delta v_{gs} R_d{}' = g_m \Delta v_{in} R_d{}' \, . \tag{7.138}$$

And with $A_{vm} = \Delta v_{out}/\Delta v_{in}$,

$$A_{vm} = g_m R_d{}' \, . \tag{7.139}$$

Equation 7.139 applies to the MOSFET and BJT.

Case 1 - MOSFET

$$A_{vm} = g_m R_d{}' \, . \tag{7.140}$$

Case 2 - BJT

$$A_{vm} = g_m R_c{}' \, . \tag{7.141}$$

In either case, we have the same midfrequency voltage gain as for the common-source (common-emitter) circuit that features a bypass capacitor. Note that the amplifier is **non-inverting** with positive A_{vm}.

Input Resistance

The current Δi_{in} is given by

$$\Delta i_{in} = \frac{\Delta v_{in}}{R_s \parallel r_\pi} - g_m \Delta v_{gs} \, . \tag{7.142}$$

So with the help of Eq. 7.137,

$$\Delta i_{in} = \frac{\Delta v_{in}}{R_s \parallel r_\pi} + g_m \Delta v_{in} \, . \tag{7.143}$$

It follows that

$$\frac{\Delta i_{in}}{\Delta v_{in}} = \frac{1}{r_{in}} = \frac{1}{R_s \parallel r_\pi} + g_m \, , \tag{7.144}$$

or

$$r_{in} = R_s \parallel r_\pi \parallel \frac{1}{g_m} = R_s \parallel \frac{1}{g_m(1 + 1/g_m r_\pi)} \, . \tag{7.145}$$

Equation 7.145 applies to the MOSFET and BJT.

Case 1 - MOSFET

In the limit as $r_\pi \to \infty$,

$$r_{in} = R_s \parallel \frac{1}{g_m} \,. \tag{7.146}$$

Case 2 - BJT

With $R_s \to R_e$ and $g_m r_\pi = \beta_o$, we have

$$r_{in} = R_e \parallel \frac{1}{g_m(1 + 1/\beta_o)} \,. \tag{7.147}$$

Generally, $\beta_o \gg 1$, and

$$r_{in} \approx R_e \parallel \frac{1}{g_m} \,. \tag{7.148}$$

In either of the preceding cases, $1/g_m$ is usually small, so the midfrequency input resistance is also small. This contributes to a small loading factor. Thus, the common-gate (common-base) amplifier is not a popular circuit. However, it enjoys some special high-frequency advantages (among other redeeming features), which are considered in Chapter 8 and beyond.

Output Resistance

Since the output of the common-gate (common-base) amplifier is at the drain (collector) terminal, the midfrequency output resistance is the same as that for the common-source (common-emitter) amplifier. Specifically,

Case 1 - MOSFET

$$r_{out} = R_d \,. \tag{7.149}$$

Case 2 - BJT

$$r_{out} = R_c \,. \tag{7.150}$$

Current Gain

The midfrequency current gain for the common-drain (common-collector) amplifier derives from the general result of Eq. 7.88.

All of the preceding characteristics of the common-gate or common-base amplifier are listed in Tables 7.2 and 7.3 at the end of this section. Take a moment to examine each entry.

Exercise 7.14 Find A_{vm}, r_{in}, and A_{vm}^{Total} for the amplifier of Fig. 7.40.

Ans: $A_{vm} = 12.0$, $r_{in} = 250\ \Omega$, $A_{vm}^{\text{Total}} = 10$

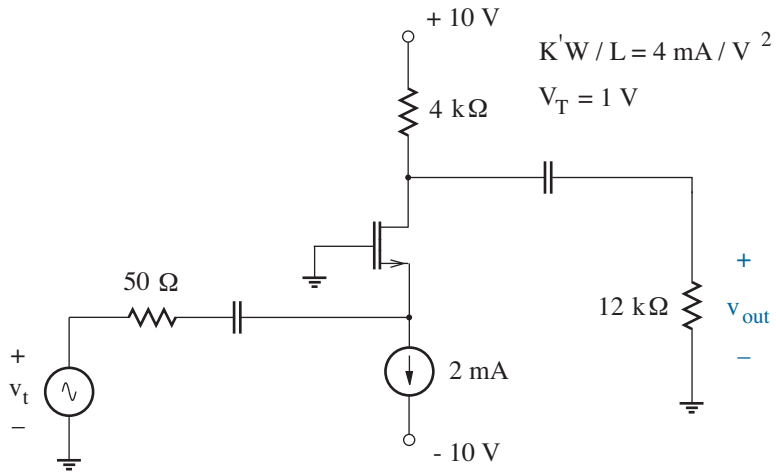


Figure 7.40: Circuit for Exercise 7.14.

Exercise 7.15 Find A_{vm}, r_{in}, and A_{vm}^{Total} for the amplifier of Fig. 7.41.

Ans: $A_{vm} = 232$, $r_{in} = 12.9\ \Omega$, $A_{vm}^{\text{Total}} = 48$

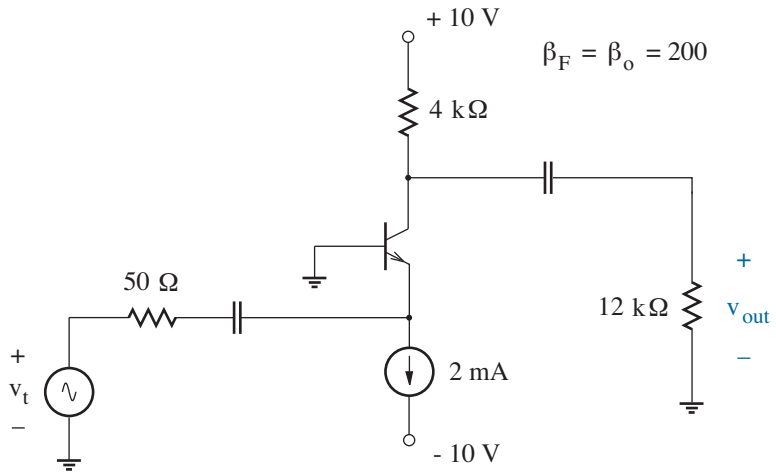


Figure 7.41: Circuit for Exercise 7.15.

Multistage Analysis

Common experience shows that teamwork often leads to improved results when individual efforts fail. The electronic parallel to this concept exploits the best features of two or more single-stage amplifiers using a multistage **cascade** configuration in which the output of one amplifier serves as the input to another. Judicious application of the "looking-away" resistances (such as $R_d{}'$) and just a little care yields straightforward analysis.

Consider the generic three-stage cascade amplifier of Fig. 7.42.

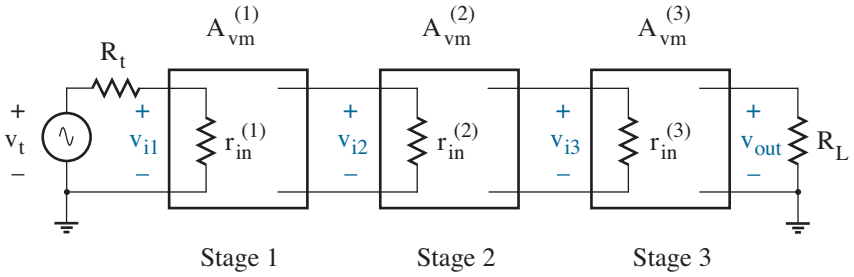


Figure 7.42: Three-stage cascade amplifier.

The total midfrequency voltage gain is the ratio of Δv_{out} to Δv_t, which can be expressed as

$$\frac{\Delta v_{out}}{\Delta v_t} = \left(\frac{\Delta v_{i1}}{\Delta v_t}\right)\left(\frac{\Delta v_{i2}}{\Delta v_{i1}}\right)\left(\frac{\Delta v_{i3}}{\Delta v_{i2}}\right)\left(\frac{\Delta v_{out}}{\Delta v_{i3}}\right) . \tag{7.151}$$

This is simply the product of a loading factor and three separate inherent midfrequency voltage gains:

$$\frac{\Delta v_{out}}{\Delta v_t} = LF \times {A_{vm}}^{(1)} \times {A_{vm}}^{(2)} \times {A_{vm}}^{(3)} . \tag{7.152}$$

The loading factor is always calculated in terms of the ac input resistance of the first stage that adjoins the Thevenin signal source:

$$LF = \frac{{r_{in}}^{(1)}}{{r_{in}}^{(1)} + R_t} . \tag{7.153}$$

When calculating the other terms, the ac input resistance for a particular stage is treated as the effective ac load resistance for the preceding stage. The stage closest to the output encounters the actual load. Thus, it is wise to analyze individual stages in reverse order, moving from output to input.

Example 7.9

Determine the total midfrequency voltage gain that applies to the 3-stage cascade amplifier of Fig. 7.43. The MOSFETs are identical.

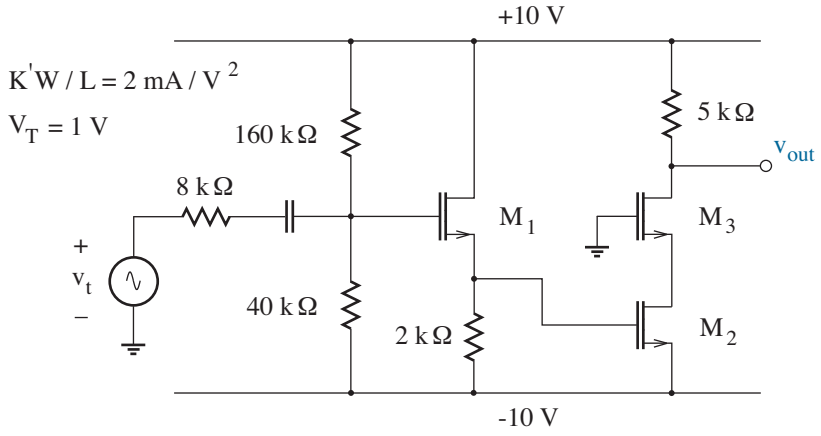


Figure 7.43: Circuit for Example 7.9.

Solution

The two bias equations for M_1 are $v_{gs1} = 4 - 2\,i_{d1}$ and $i_{d1} = (v_{gs1} - 1)^2$ —assuming saturation conditions—with v_{gs1} solutions at -0.5 V and 2 V. Only the second solution exceeds the M_1 threshold voltage, so $i_{d1} = 1$ mA. This establishes $v_{gs2} = 2$ V, and $i_{d2} = i_{d3} = 1$ mA. The forced biasing for M_3 yields a consistent source node voltage at -2 V. Having determined the MOSFET drain currents, we obtain $g_{m1} = g_{m2} = g_{m3} = g_m = 2 \times 10^{-3}$ ℧ as the small-signal transconductance values. The r_π values are all infinite.

The 3-stage analysis now proceeds from the output back to the input.

Stage 3 featuring M_3 is a common-gate amplifier—input at the source, output at the drain, and the gate at ac ground. The inherent midfrequency voltage gain for this stage is

$$A_{vm}{}^{(3)} = g_m R_d{}' = 2 \times 10^{-3}\ \text{℧} \times 5 \times 10^3\ \Omega = 10\,, \tag{7.154}$$

and the ac input resistance is

$$r_{in}{}^{(3)} = 1/g_m = 500\ \Omega\,. \tag{7.155}$$

(Compare with Fig. 7.38 to note that $R_s = \infty$.) Figure 7.44 shows the new stage-2 perspective with stage 3 replaced by a 500-Ω resistor ($r_{in}{}^{(3)}$).

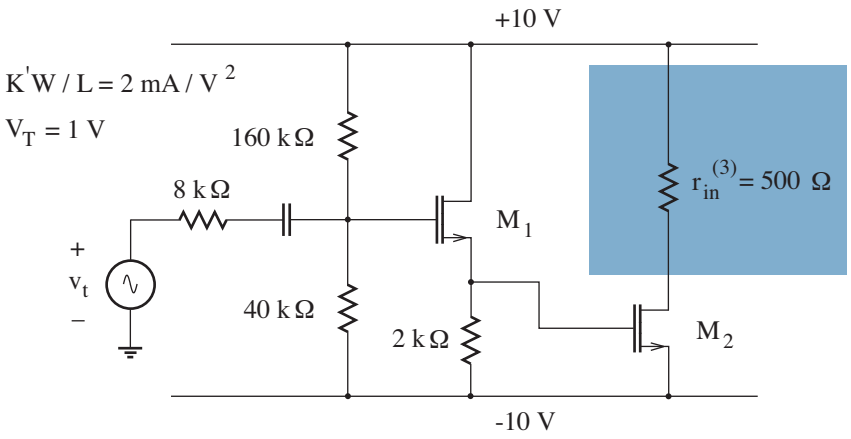


Figure 7.44: Revised circuit for Example 7.9: The common-gate stage-3 amplifier has been replaced by an equivalent ac resistance.

Stage 2 featuring M_2 is a common-source amplifier—input at the gate, output at the drain, and source at ac ground. The inherent midfrequency voltage gain for this stage is

$$A_{vm}{}^{(2)} = -g_m R_d{}' = -2 \times 10^{-3}\ \mho \times 0.5 \times 10^3\ \Omega = -1\,, \tag{7.156}$$

and the ac input resistance is infinite. Thus, the new stage-1 perspective at midfrequency has stages 2 and 3 replaced with an open circuit.

Stage 1 featuring M_1 is a common-drain amplifier—input at the gate, output at the source, and drain at ac ground. The inherent midfrequency voltage gain for this stage is

$$A_{vm}{}^{(1)} = \frac{g_m R_s{}'}{1 + g_m R_s{}'} = 0.8 \tag{7.157}$$

(subject to $g_m R_s{}' = 2 \times 10^{-3}\ \mho \times 2 \times 10^3\ \Omega = 4$). The ac input resistance is 160 kΩ ∥ 40 kΩ = 32 kΩ.

At last, we are positioned to determine the loading factor, which is

$$LF = \frac{r_{in}{}^{(1)}}{r_{in}{}^{(1)} + R_t} = \frac{32\ \text{k}\Omega}{32\ \text{k}\Omega + 8\ \text{k}\Omega} = 0.8\,. \tag{7.158}$$

So putting everything together, we have

$$A_{vm}^{\text{Total}} = LF \times A_{vm}{}^{(1)} \times A_{vm}{}^{(2)} \times A_{vm}{}^{(3)} = 0.8 \times 0.8 \times (-1) \times 10 = -6.4\,. \tag{7.159}$$

Not all that large, but not every electronic team is a winner.

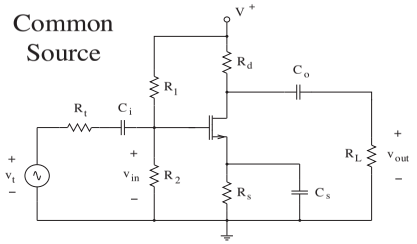
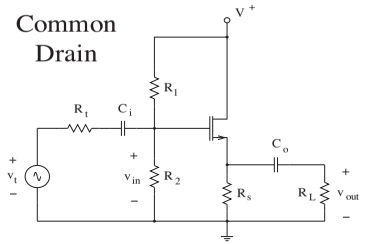
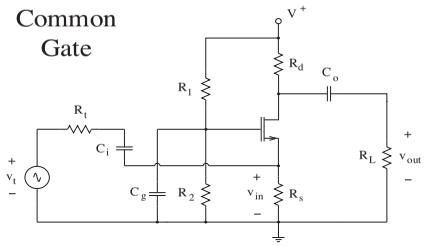
Configuration	Inherent Voltage Gain	Input Resistance	Output Resistance
Common Source 	$\dfrac{-g_m R_d'}{1+g_m R_s'}$ $R_d' = (R_d \parallel R_L)$ with C_s : $R_s' = 0$ w/o C_s : $R_s' = R_s$	$R_1 \parallel R_2$	R_d
Common Drain 	$\dfrac{g_m R_s'}{1+g_m R_s'}$ $R_s' = (R_s \parallel R_L)$	$R_1 \parallel R_2$	$R_s \parallel \dfrac{1}{g_m}$
Common Gate 	$g_m R_d'$ $R_d' = (R_d \parallel R_L)$	$R_s \parallel \dfrac{1}{g_m}$	R_d

Table 7.2 MOSFET Amplifier Characteristics:

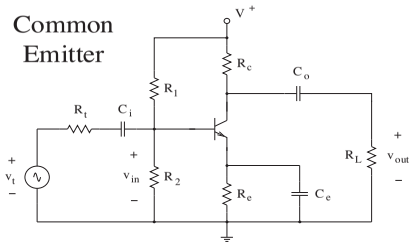
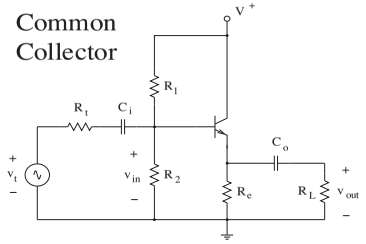
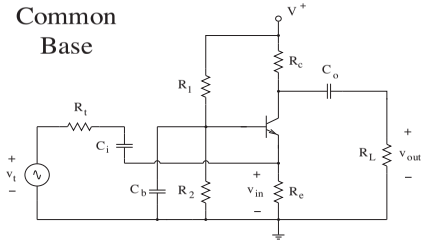
Configuration	Inherent Voltage Gain	Input Resistance	Output Resistance
Common Emitter 	$\dfrac{-g_m R_c'}{1 + g_m R_e'(1 + 1/\beta_o)}$ $R_c' = (R_c \parallel R_L)$ with C_e: $R_e' = 0$ w/o C_e: $R_e' = R_e$	$R_1 \parallel R_2 \parallel [r_\pi + (\beta_o + 1)R_e']$	R_c
Common Collector 	$\dfrac{g_m R_e'(1 + 1/\beta_o)}{1 + g_m R_e'(1 + 1/\beta_o)}$ $R_e' = (R_e \parallel R_L)$	$R_1 \parallel R_2 \parallel [r_\pi + (\beta_o + 1)R_e']$	$R_e \parallel \dfrac{1 + R_b'/r_\pi}{g_m(1 + 1/\beta_o)}$ $R_b' = R_1 \parallel R_2 \parallel R_t$
Common Base 	$g_m R_c'$ $R_c' = (R_c \parallel R_L)$	$R_e \parallel \dfrac{1}{g_m(1 + 1/\beta_o)}$	R_c

Table 7.3 BJT Amplifier Characteristics:

7.4 Single-Stage Amplifier Design

Not that many years ago, a textbook such as this would have devoted much to the topic of amplifier design using discrete transistors, primarily BJTs. The widespread use of high-performance integrated operational amplifiers makes discrete-transistor amplifier designs increasingly rare. Nevertheless, some consideration is warranted as a prelude to modern practice.

The Common-Emitter Amplifier

The common-emitter amplifier of Fig. 7.45 is a popular design candidate when moderate-to-large total midfrequency voltage gains are necessary. In a typical situation, input (v_t, R_t) and output (R_L) circuit parameters are known, and V^+ is fixed by system constraints. The design problem is to determine R_c and R_e. Then the biasing guidelines of Section 7.1 are used to find R_1 and R_2 values that support an optimum and stable Q-point. Designs that meet intended performance standards are not guaranteed!

(Problem 7.75 examines the common-collector BJT configuration.)

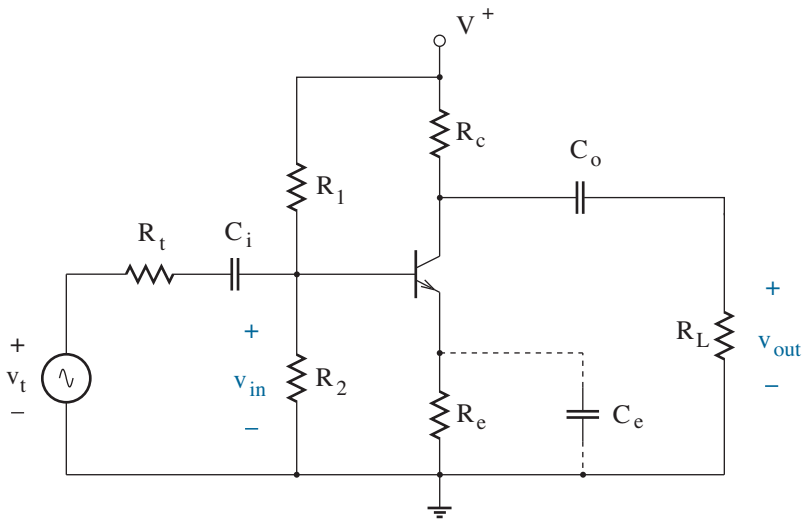


Figure 7.45: A popular design candidate — the common-emitter amplifier.

At the risk of implying a cookbook procedure that is carved in stone, we consider a 12-step design process. Most steps are straightforward and require little justification. Only the first introduces a new design guideline. And two steps are opportunities to trash designs with hopeless prospects.

Step 1 - Choose R_c.

From the load perspective, the Thevenin equivalent of the common-emitter amplifier has the form shown in Fig. 7.46. The Thevenin resistance is r_{out}, and the Thevenin signal voltage is $v_t{}'$.

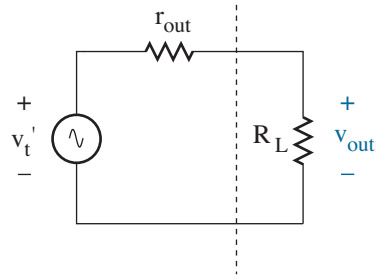


Figure 7.46: Thevenin equivalent of the common-emitter amplifier.

Consider the power delivered to R_L. Specifically,

$$P_L = \frac{v_t{}'^2 \, R_L}{(r_{out} + R_L)^2} \, . \tag{7.160}$$

A classic circuit problem is to determine the load resistance that maximizes P_L when r_{out} is fixed. For maximum P_L, we require

$$\left. \frac{dP_L}{dR_L} \right|_{r_{out} \text{ fixed}} = 0 \, , \tag{7.161}$$

with the simple result that $R_L = r_{out}$. Although seldom explicitly stated, the reverse condition is desirable for most designs—we want to maximize P_L when R_L is fixed. In this case,

$$\left. \frac{dP_L}{dr_{out}} \right|_{R_L \text{ fixed}} = 0 \, , \tag{7.162}$$

with the dubious result that $r_{out} = 0$. For the common-emitter amplifier, $r_{out} = R_c$, and the midfrequency voltage gain is proportional to $R_c \parallel R_L$. (See Table 7.4.) Thus, $r_{out} = 0$ implies $A_{vm} = 0$.

The trade-off between midfrequency voltage gain and power transfer to the load suggests a design compromise of the form

$$r_{out} \approx R_L. \tag{7.163}$$

For the design at hand, we let $r_{out} = R_c = R_L$.

Step 2 -Satisfy the inherent voltage-gain requirement.

The total midfrequency voltage gain A_{vm}^{Total} is the product of an inherent midfrequency voltage gain A_{vm} and a loading factor LF. The loading factor is unknown, so we assume a trial value of unity (pending readjustment). In turn, our trial design seeks $A_{vm} = A_{vm}^{\text{Total}}$. Two cases warrant attention:

Case A - We need to satisfy a large voltage-gain requirement, and we are not especially worried about temperature sensitivity. Thus, we bypass R_e with a capacitor, and

$$g_m = \frac{-A_{vm}}{R_c \parallel R_L} = \frac{i_c|_Q}{kT/q} \,. \tag{7.164}$$

A particular A_{vm} determines $i_c|_Q$.

Case B - We need to satisfy a moderate voltage-gain requirement subject to minimal temperature sensitivity. Thus, we choose not to bypass R_e. Parameter g_m is unknown but $g_m R_e \gg 1$ is likely. So in view of Eq. 7.72,

$$R_e \approx \frac{R_c \parallel R_L}{-A_{vm}} \,. \tag{7.165}$$

A particular A_{vm} determines R_e.

Step 3 - Choose an optimum Q-point.

In Section 7.1, we developed a pair of design equations that optimized the Q-point by placing it in the center of the ac load line (with the expectation of large output excursions). Specifically,

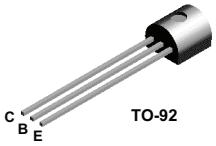
$$i_c|_Q = \frac{V^+}{R_{AC} + R_{DC}} \tag{7.166}$$

and

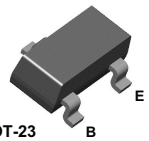
$$v_{ce}|_Q = R_{AC}\, i_c|_Q \,. \tag{7.167}$$

For Design A, $R_{AC} = R_c \parallel R_L$ and $R_{DC} = R_c + R_e$. Step 2 determined the quiescent collector current, so we find consistent values for $v_{ce}|_Q$ and R_e. For Design B, $R_{AC} = R_c \parallel R_L + R_e$ and $R_{DC} = R_c + R_e$, with R_e known. Values for $i_c|_Q$ and $v_{ce}|_Q$ are directly obtained from Eqs. 7.166 and 7.167.

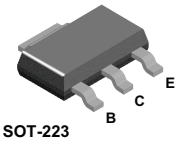
The chosen Q point must be consistent with i_c, v_{ce}, and quiescent power dissipation values that are not too close to the maximum transistor ratings. Figure 7.47 shows the related specifications for the popular 2N3904 BJT. A Q point with modest power dissipation is often desirable if the anticipated output excursions are small.



TO-92



SOT-23
Mark: 1A



SOT-223

NPN General Purpose Amplifier

This device is designed as a general purpose amplifier and switch. The useful dynamic range extends to 100 mA as a switch and to 100 MHz as an amplifier.

Absolute Maximum Ratings*

T_A = 25°C unless otherwise noted

Symbol	Parameter	Value	Units
V_{CEO}	Collector-Emitter Voltage	40	V
V_{CBO}	Collector-Base Voltage	60	V
V_{EBO}	Emitter-Base Voltage	6.0	V
I_C	Collector Current - Continuous	200	mA
T_J, T_{stg}	Operating and Storage Junction Temperature Range	-55 to +150	°C

*These ratings are limiting values above which the serviceability of any semiconductor device may be impaired.

NOTES:
1) These ratings are based on a maximum junction temperature of 150 degrees C.
2) These are steady state limits. The factory should be consulted on applications involving pulsed or low duty cycle operations.

Thermal Characteristics

T_A = 25°C unless otherwise noted

Symbol	Characteristic	Max			Units
		2N3904	*MMBT3904	**PZT3904	
P_D	Total Device Dissipation Derate above 25°C	625 5.0	350 2.8	1,000 8.0	mW mW/°C
$R_{\theta JC}$	Thermal Resistance, Junction to Case	83.3			°C/W
$R_{\theta JA}$	Thermal Resistance, Junction to Ambient	200	357	125	°C/W

*Device mounted on FR-4 PCB 1.6" X 1.6" X 0.06."

**Device mounted on FR-4 PCB 36 mm X 18 mm X 1.5 mm; mounting pad for the collector lead min. 6 cm^2.

Figure 7.47: 2N3904 data sheet showing the maximum bias conditions. Courtesy Fairchild Semiconductor.

Step 4 - Inspect the ac and dc load lines.

We evaluate the Q-point and the ac and dc load lines in relation to BJT operation in the forward active mode. We need to avoid saturation, cutoff, and excessive quiescent power dissipation.

Apart from its presumed indication of a contented BJT, the ac load line determines the maximum permissible output voltage swing. Some general comments are appropriate for each of the design cases under consideration. In Design A, we have

$$v_{ce}|_Q = R_{AC}\, i_c|_Q = (R_c \parallel R_L)\, i_c|_Q \,, \tag{7.168}$$

and from Eq. 7.164,

$$i_c|_Q = \frac{A_{vm}}{R_c \parallel R_L} \left(\frac{kT}{q}\right). \tag{7.169}$$

Thus, the maximum output voltage swing is approximately

$$2\, v_{ce}|_Q = 2 \left(\frac{kT}{q}\right) A_{vm} \approx (50\,\text{mV}) \times A_{vm} \,. \tag{7.170}$$

Equation 7.170 is a surprising result since it does not depend on the value of the power supply voltage. Notwithstanding, it shows that large maximum output voltage swing requires large A_{vm}. For this reason (among others), we usually do not want to bypass R_e if A_{vm} is moderate or small.

Now consider Design B. Here, we have

$$\begin{aligned} v_{ce}|_Q &= R_{AC}\, i_c|_Q = \frac{R_{AC}\, V^+}{R_{AC} + R_{DC}} \\ &= \frac{(R_c \parallel R_L + R_e)\, V^+}{R_c \parallel R_L + R_c + 2R_e} \,. \end{aligned} \tag{7.171}$$

Then with $R_c = R_L$ and $R_e \ll (R_c \parallel R_L)$ (for $A_{vm} \gg 1$), the maximum output voltage swing is approximately

$$\Delta v_{out}\,(\text{max}) \approx v_{ce}|_Q \pm V^+/3 \,. \tag{7.172}$$

Equation 7.172 is a useful design aid. In practice, we still observe increased output voltage swing with increasing A_{vm}, but the effect is relatively small.

Step 5 - Determine g_m (if still unknown) and r_π.

We use the relation

$$g_m = \frac{\beta_o}{r_\pi} = \frac{i_c|_Q}{kT/q} \,. \tag{7.173}$$

Step 6 - Determine the actual A_{vm} (if necessary).

This step only applies to Design B, which used an approximate expression for A_{vm} to find R_e. Since we have now obtained a value for g_m, the actual inherent midfrequency voltage gain is easily determined using the complete expression that appears in Table 7.3. The revised A_{vm} is typically about 5% less than its initial value.

Step 7 - Choose $R_1 \parallel R_2$.

For this calculation, we recall the design condition for BJT bias stability:

$$R_1 \parallel R_2 \sim \frac{1}{10} \left(\beta_F + 1\right) R_e \, . \tag{7.174}$$

We postpone the choice of individual R_1 and R_2 values, pending the results of the next few design steps.

Step 8 - Determine r_{in}.

We know the value for r_π, so the midfrequency input resistance is easily determined using the complete expression that appears in Table 7.3.

Step 9 -Determine the loading factor and A_{vm}^{Total}.

Both calculations are trivial when we use Eqs. 7.83 and 7.84, respectively.

Step 10 - Test the calculated A_{vm}^{Total}.

We now compare A_{vm}^{Total} with the desired total midfrequency voltage gain. If the calculated value is unacceptably low (as is often the case), we must consider several corrective design options that depend on the loading factor determined in Step 9.

- The loading factor is close to unity. — In this case, the low value for A_{vm}^{Total} results from the use of an approximate expression for the inherent midfrequency voltage gain (as in Design B). The error can be corrected by increasing R_c or by decreasing R_e. Often, we prefer the latter method, since it has less effect on $v_{ce}|_Q$.

- The loading factor is slightly less than unity. — In this case, we need to return to Step 2 with a modified estimate for the loading factor. For example, if the calculated loading factor is 0.8, we might try for a new design with A_{vm} increased by a factor of $1/0.8 = 1.25$. We can expect that this will lead to a new r_{in}, a new loading factor, and a more favorable value for $A_{vm} \times \text{LF} = A_{vm}^{\text{Total}}$.

- The loading factor is significantly less than unity. — In this case, $r_{in} \ll R_t$, and the BJT single-stage common-emitter amplifier is not a good design prospect. If the required total midfrequency voltage gain is small, a MOSFET common-source amplifier may be acceptable. If the required total midfrequency voltage gain is large, we need to consider an amplifier with more than one stage.

With some experience, we can sometimes anticipate one of the preceding options when we are at Step 2.

Step 11 - Determine R_1 and R_2.

If the amplifier design survives Step 10, it is worth the effort to determine R_1 and R_2 individually. This is a standard biasing problem that is quickly resolved once R_e, $(R_1 \parallel R_2)$, and $i_c|_Q$ have been specified.

Step 12 - Simulate the amplifier design with SPICE.

This step tests for errors at the *end* of the design process. One should never design a circuit by means of trial-and-error simulations that use random component values.

Example 7.10

Design a common-emitter amplifier with $A_{vm}^{\text{Total}} = -8$ when $R_t = 100\,\Omega$, $R_L = 4\text{ k}\Omega$, and $V^+ = 15$ V. For the BJT, $\beta_F = \beta_o = 100$.

Solution

We follow the procedure outlined previously.

Step 1 - For efficient power transfer, we let $R_c = R_L = 4\text{ k}\Omega$.

Step 2 - We assume LF = 1, and we design for $A_{vm} = -8$. This voltage gain is rather small, so we choose not to bypass R_e. Then in consideration of the approximate expression for the midfrequency voltage gain,

$$R_e = \frac{R_c \parallel R_L}{-A_{vm}} = \frac{2000\,\Omega}{8} = 250\,\Omega\,. \tag{7.175}$$

Step 3 - We place the Q-point in the center of the ac load line. For this calculation, we have $R_{AC} = R_c \parallel R_L + R_e = 2.25\text{ k}\Omega$ and $R_{DC} = R_c + R_e = 4.25\text{ k}\Omega$. In turn,

$$i_c|_Q = \frac{V^+}{R_{AC} + R_{DC}} = \frac{15\,V}{6.5\text{ k}\Omega} = 2.3\text{ mA} \tag{7.176}$$

and

$$v_{ce}|_Q = R_{AC}\, i_c|_Q = 2.25\,\text{k}\Omega \times 2.3\,\text{mA} = 5.2\,\text{V}. \tag{7.177}$$

Step 4 - The ac and dc load lines are shown in Fig. 7.48. The maximum permissible voltage swing along the ac load line is roughly ±5 V.

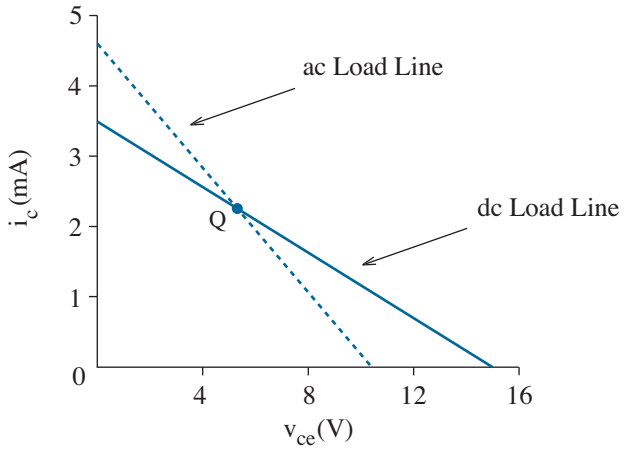


Figure 7.48: ac and dc load lines for the amplifier of Example 7.10.

Step 5 - Given $i_c|_Q$, we find

$$g_m = \frac{i_c|_Q}{kT/q} = \frac{2.3\,\text{mA}}{25.9\,\text{mV}} = 88.8 \times 10^{-3}\,\text{℧}\,, \tag{7.178}$$

and

$$r_\pi = \frac{\beta_o}{g_m} = \frac{100}{88.8 \times 10^{-3}\,\text{℧}} = 1.13\,\text{k}\Omega\,. \tag{7.179}$$

Step 6 - We use Table 7.3 to find the actual midfrequency voltage gain. Specifically,

$$A_{vm} = \frac{-g_m(R_c \parallel R_L)}{1 + g_m R_e(1 + 1/\beta_o)} = \frac{-88.8 \times 2}{1 + 88.8 \times 0.25 \times 1.01} = -7.58\,. \tag{7.180}$$

Step 7 - To ensure bias stability, we require

$$R_1 \parallel R_2 \sim \frac{(\beta_F + 1)R_e}{10} \approx 2.5\,\text{k}\Omega\,. \tag{7.181}$$

Step 8 - We calculate r_{in} using the complete expression from Table 7.3 with the result

$$r_{in} = R_1 \parallel R_2 \parallel [r_\pi + (1 + \beta_o)R_e] = 2.5\,\text{k}\Omega \parallel 26.4\,\text{k}\Omega = 2.28\,\text{k}\Omega\,. \tag{7.182}$$

Step 9 - The loading factor is

$$\text{LF} = \frac{r_{in}}{r_{in} + R_t} = \frac{2.28}{2.28 + 0.1} = 0.958\,, \tag{7.183}$$

and the total midfrequency voltage gain is given by

$$A_{vm}^{\text{Total}} = A_{vm} \times \text{LF} = -7.58 \times 0.958 = -7.3\,. \tag{7.184}$$

Step 10 - So far, our design yields an A_{vm}^{Total} that is slightly less than our desired goal (-8). This error is less than 10 %, and perhaps we could live with it. However, we choose to improve our design. Since the loading factor is not too far from unity, we adjust the total midfrequency voltage gain by decreasing R_e. Specifically, we let

$$R_e \to R_e \times \frac{7.3}{8} = 230\,\Omega\,. \tag{7.185}$$

We do not bother to recalculate the optimum Q-point, $(R_1 \parallel R_2)$ or r_{in}. But we return to Eqs. 7.180 and 7.184 to find that A_{vm}^{Total} is now -7.9.

Not bad! Any further revisions are pointless, since we can expect some degree of component uncertainty when we attempt to build our design.

Step 11 - We have already encountered the biasing design problem in which we seek individual values for R_1 and R_2 given $i_c|_Q$ and $(R_1 \parallel R_2)$. Thus, we skip over the details of this calculation. The result is $R_1 = 29$ kΩ and $R_2 = 2.7$ kΩ.

Step 12 - To check our design, we follow the SPICE analysis outlined in Example 7.8. But we delete the .temp statement (since we are only interested in room-temperature amplifier performance), and we include the .op statement (in order to obtain Q-point information). With $R_1 = 29$ kΩ, $R_2 = 2.7$ kΩ, $R_c = 4$ kΩ, and $R_e = 230$ kΩ, the simulation yields $A_{vm}^{\text{Total}} = -7.86$, $i_c|_Q = 2.27$ mA, and $v_{ce}|_Q = 5.39$ V. Once again, not bad!

Example 7.11

Use SPICE to examine the expected **yield statistics** for the design of Example 7.10 if $|A_{vm}|$ must be 8.0 ± 0.5. Consider 10% and 5% tolerance values for resistors R_1, R_2, R_c, and R_e.

Solution

The format of the SPICE code is similar to that for Example 7.4. However, we describe the uncertain resistors by including a model reference as in the following statements:

```
R1      4       3       Rmod    29K
R2      3       0       Rmod    2.7K
Rc      4       5       Rmod    4K
Re      6       0       Rmod    230
```

The corresponding .model statement takes the form

```
.model      Rmod        RES(R = 1, DEV = 10%)
```

This individually multiplies the Rmod resistors by $1 + 0.1m$, where m is a random number between -1 and 1.

So much for the resistors. The key to the SPICE code is the command sequence

```
.mc         100         AC          v(7)        YMAX
+LIST       OUTPUT      ALL
.ac         LIN         1           1000        1000
```

that initiates a 100-trial **Monte Carlo** ac analysis at 1000 Hz. Specifically, the first SPICE trial determines the ac amplitude at node 7 (the amplifier output) when the Rmod resistors have their "nominal" values ($m = 0$). The next 99 SPICE trials determine the ac amplitude at node 7; however, the Rmod resistor values are allowed to fluctuate randomly between runs as if chosen in a (Monte Carlo) casino. When the 100 trials are completed, SPICE computes the average and the standard deviation of the *difference* between each node-7 amplitude and the nominal value. Parameter YMAX summarizes the trial results in order of the decreasing absolute difference. (In contrast, the optional MAX parameter summarizes the trial results in order of the decreasing actual difference.)

One can show that the statistical error of the Monte Carlo analysis varies inversely with $\sqrt{N}$, where N is the number trials. Thus, the error for this example is ± 10 %. A significant improvement in accuracy would require a very large output file.

Here are the SPICE results:

- The nominal midfrequency voltage gain is -7.86. (This is determined by dividing the nominal ac amplitude at node 7 by the ac amplitude of the input signal generator.)
- For the case of 10-% resistor tolerance and 100 trials, the difference between A_{vm} and the nominal A_{vm} value has an average of -0.113 and a standard deviation of 0.507. The number of trials that produce A_{vm} values between -7.5 and -8.5 is 59, so the design yield is 59 %.
- The worst-case trials feature: a) $R_1 = 29.9$ kΩ, $R_2 = 2.78$ kΩ, $R_c = 4.32$ kΩ, and $R_e = 207$ Ω (with $A_{vm} = -9.05$); and, b) $R_1 = 29.3$ kΩ, $R_2 = 2.47$ kΩ, $R_c = 3.66$ kΩ, and $R_e = 251$ Ω (with $A_{vm} = -6.80$). Of course, the random nature of the Monte Carlo analysis will change the worst-case conditions if the simulation is repeated.
- When the number of trials is increased to 400 (in order to obtain only 5 % statistical error), the A_{vm} difference from the nominal A_{vm} value is reduced (-0.047), but the standard deviation is about the same (0.515). The design yield is unchanged (59 %). These data demonstrate the adequacy of 100-trial simulations.
- For the case of 5-% resistor tolerance and 100 trials, the difference between A_{vm} and the nominal A_{vm} value has an average of -0.059 and a standard deviation of 0.253. The design yield increases to 89 %, an obvious improvement.

In a more realistic simulation, it would be appropriate to account for an uncertain β_F by including DEV = 50 after BF = 100 in the BJT model statement. (In this case, $50 < \beta_F < 150$.) However, it is easy to show that the effect of random β_F variation becomes small when R_e is not bypassed (as in this design). Recall that R_e is intended to reduce bias sensitivity to uncertain β_F.

> Note: You should verify some of the preceding SPICE results by performing your own simulation. In particular, you should understand the contents of the SPICE output file, which is too long to present here.

Monte Carlo analysis is important since it simulates the uncertainty of the manufacturing process. Often, several different circuit designs can be used to achieve a given set of performance objectives. If a large number of identical circuits must be produced, the best design is that which maximizes yield and profit.

Concept Summary

Transistors are commonly used as midfrequency amplifiers.

- The transistor biasing process establishes a quiescent operating point that provides a frame of reference for current and voltage changes.
  - Negotiated biasing mutually satisfies current-voltage constraints for the transistor and an external resistive circuit.
    - A source (emitter) resistor promotes bias-current stability.
  - Imposed biasing forces a transistor to accommodate the value of an external current source, often in the form of a current mirror.
- Small-signal transistor behavior reflects a particular Q point.
  - A transconductance parameter g_m indicates the strength of the dependent source for drain or collector current.
    - MOSFETs have g_m proportional to $\sqrt{i_d|_Q}$.
    - BJTs have g_m proportional to $i_d|_Q$.
  - A resistance parameter r_π governs gate or base current.
    - MOSFETs have infinite r_π.
    - BJTs have r_π proportional to ${i_d|_Q}^{-1}$.
- Only three basic amplifier configurations are available.
  - The common-source (common-emitter) amplifier
    - Has large or moderate negative voltage gain, depending on the presence of a source (emitter) bypass capacitor;
    - Has pronounced temperature sensitivity under bypass.
  - The common-drain (common-collector) amplifier
    - Has voltage gain typically just under unity;
    - Has large or moderate current gain that varies in proportion to the amplifier input resistance.
  - The common-gate (common-base) amplifier
    - Has large positive voltage gain;
    - Has small input resistance for a reduced loading factor.
- The total voltage gain produced by a cascade of amplifiers is equal to the product of the individual inherent voltage gains and the loading factor that applies to the first amplifying stage.

Problems

Section 7.1

Negotiated Biasing

7.1 The MOSFET shown in Fig. P7.1 has $K'W/L = 8$ mA/V^2 and $V_T = 1$ V. Determine the Q point.

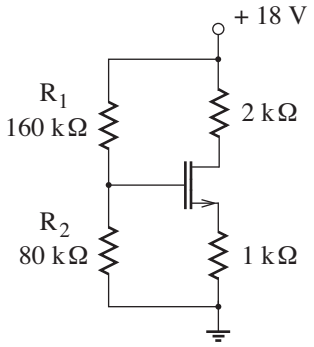


Figure P7.1

7.2 Repeat Problem P7.2, but let $R_2 = 40$ kΩ.

7.3 Redesign the circuit of Fig. P7.1 subject to $v_{ds}|_Q = 9$ V and $R_1 \parallel R_2 = 80$ kΩ.

7.4 Redesign the circuit of Fig. P7.1 subject to $i_d|_Q = 2$ mA and $R_1 \parallel R_2 = 100$ kΩ.

7.5 The MOSFET shown in Fig. P7.5 has $K'W/L = 3$ mA/V^2 and $V_T = -0.5$ V. Determine the Q point.

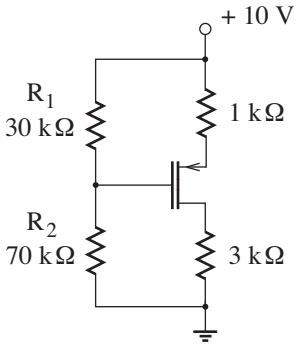


Figure P7.5

7.6 Repeat Problem P7.5, but let $R_2 = 170$ kΩ.

7.7 Redesign the circuit of Fig. P7.5 subject to $v_{ds}|_Q = -5$ V and $R_1 \parallel R_2 = 50$ kΩ.

7.8 Redesign the circuit of Fig. P7.5 subject to $i_d|_Q = -1$ mA and $R_1 \parallel R_2 = 8$ kΩ.

7.9 The BJT shown in Fig. P7.9 has $\beta_F = 100$. Determine the Q point.

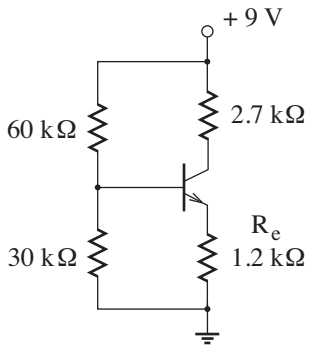


Figure P7.9

7.10 Repeat Problem P7.9, but let $R_e = 1.8$ kΩ.

7.11 The BJT shown in Fig. P7.11 has $\beta_F = 120$. Determine the Q point.

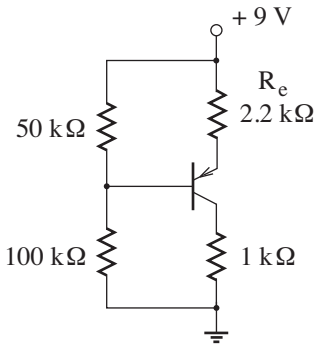


Figure P7.11

7.12 Repeat Problem P7.11, but let $R_e = 2.7$ kΩ.

7.13 The BJT shown in Fig. P7.13 has $\beta_F = 150$.

(a) Determine the Q point.

(b) Specify a design condition for R_b that promotes bias stability.

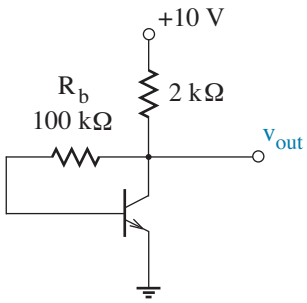


Figure P7.13

7.14 The circuit of Fig. P7.14 uses a diode to compensate for the effect of the base-emitter junction voltage in setting the BJT Q point. Let $\beta_F = 180$.

(a) Assume that the diode and the base-emitter junction both have an approximate 0.7-V drop. Complete the design to ensure $i_c|_Q = 1$ mA. Use Eq. 7.40 as a guide for bias stability.

(b) Specify how the forward diode voltage should relate to v_{be} if the two voltages are to cancel in the expression for quiescent base current.

(c) Use SPICE to show the variation of $i_c|_Q$ vs. temperature ($-40\,^\circ\text{C} \le T \le +85\,^\circ\text{C}$) for the biasing design of part a. Assume $X_{TB} = 1.5$. Let IS=10f for the BJT, and let the diode have the relative IS value consistent with part b.

(d) Redesign the circuit without the diode using Eq. 7.40 as a guide for bias stability.

(e) Use SPICE to show the variation of $i_c|_Q$ vs. temperature ($-40\,^\circ\text{C} \le T \le +85\,^\circ\text{C}$) for the biasing design of part d. Assume $X_{TB} = 1.5$, and let IS=10f for the BJT.

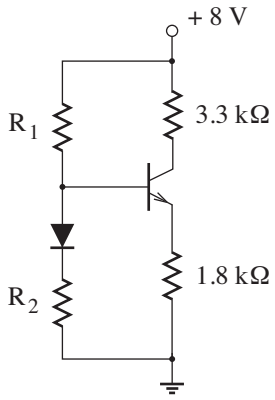


Figure P7.14

7.15 The BJT shown in Fig. P7.15 has $\beta_F = 100$.

(a) Find the Q point that maximizes excursions along the ac load line.

(b) Draw the dc and ac load lines. Be sure to label your sketch appropriately.

(c) Complete the design for the Q point of part a. Ensure bias stability.

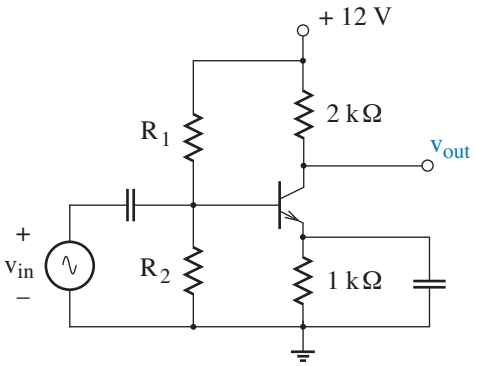


Figure P7.15

7.16 Consider the circuit of Problem P7.15 with $R_1 = 90$ kΩ and $R_2 = 30$ kΩ. Draw the dc and ac load lines. Label your sketch appropriately.

7.17 The BJT shown in Fig. P7.17 has $\beta_F = 130$.

(a) Find the Q point that maximizes excursions along the ac load line.

(b) Draw the dc and ac load lines. Be sure to label your sketch appropriately.

(c) Complete the design for the Q point of part a. Ensure bias stability.

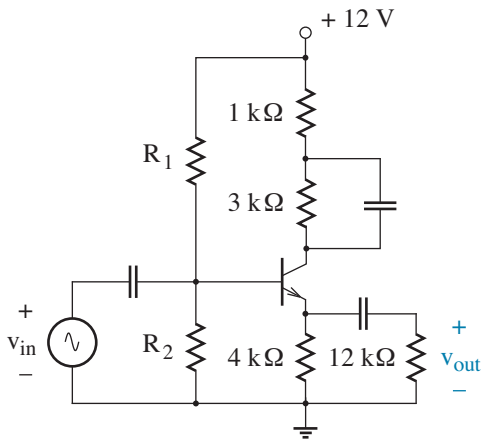


Figure P7.17

7.18 Consider the circuit of Problem P7.17 with $R_1 = 47$ kΩ and $R_2 = 75$ kΩ. Draw the dc and ac load lines. Label your sketch appropriately.

7.19 Use SPICE and the .op command to verify the results of Exercise 7.1.

7.20 Consider the biasing circuit of Example 7.1.

(a) Use SPICE to examine bias stability subject to $K'W/L$ varying from 1.75 to 2.25 mA/V^2 (parametric uncertainty). Hint: Perform a dc model-parameter sweep with the statement

.dc NMOS NAME(KP) 1.75 2.25 0.05

where NAME is the MOSFET model name.

(b) Redesign the circuit so that $i_d|_Q = 1$ mA, but change the 2-kΩ source resistor to 4 kΩ.

(c) Use SPICE to examine the bias stability of the revised circuit.

7.21 Consider the biasing circuit of Example 7.3.

(a) Use SPICE to examine bias stability with β_F varying from 50 to 150 (parametric uncertainty). Assume IS=10f. Hint: Perform a dc model-parameter sweep with the statement

.dc NPN NAME(BF) 50 150 1

where NAME is the BJT model name.

(b) Redesign the circuit so that $i_c|_Q = 1$ mA with $R_1 \parallel R_2 = 160$ kΩ for the base resistors.

(c) Use SPICE to examine the bias stability of the revised circuit.

7.22 Use SPICE to examine the bias stability of the circuit in Exercise 7.4a over the temperature range $-40\,°\text{C} \le T \le +85\,°\text{C}$. Let IS=10f and $X_{TB} = 1.5$.

Imposed Biasing

7.23 The MOSFET in Fig. P7.23 features $K_n' = 50\ \mu$A/V^2 and $V_T = 0.5$ V. Complete the design so that $i_d|_Q = 2$ mA and $v_{ds}|_Q = 2.5$ V.

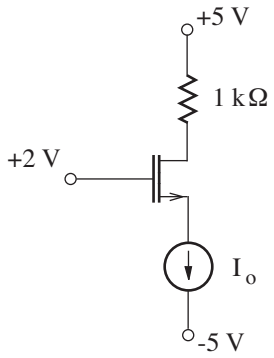


Figure P7.23

7.24 The MOSFETs in Fig. P7.24 feature $K_p' = 20\ \mu\text{A/V}^2$ and $V_T = -0.5$ V. Find the M_1 Q point.

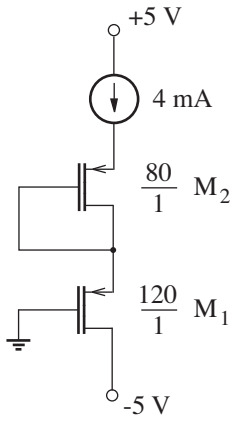


Figure P7.24

7.25 The MOSFETs in Fig. P7.25 feature $K_n' = 50\ \mu\text{A/V}^2$ and $V_T = 0.5$ V.

(a) Complete the design so that M_1, M_2, and M_3 have the indicated drain currents.

(b) Use SPICE to verify the design of part a.

(c) Find the maximum R_1, R_2, and R_3 values that are consistent with saturation for the associated MOSFETs.

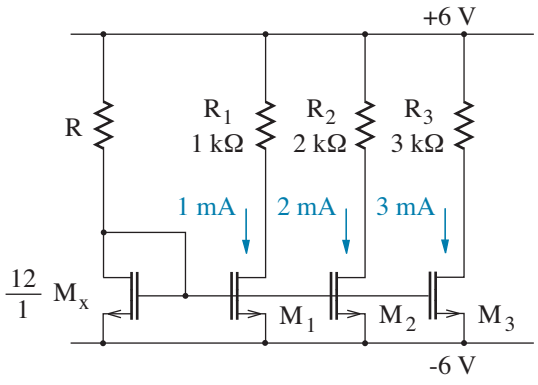


Figure P7.25

7.26 The BJTs shown in Fig. P7.26 have $\beta_F = 120$. The integrated transistors are sized in relation to Q_2 and Q_3 for which AREA=1. Assume IS = 10f.

(a) Complete the design to achieve the indicated collector currents.

(b) Use SPICE to verify the design of part a.

(c) Find the maximum R_4 and R_5 values that are consistent with the forward active mode for the associated BJTs.

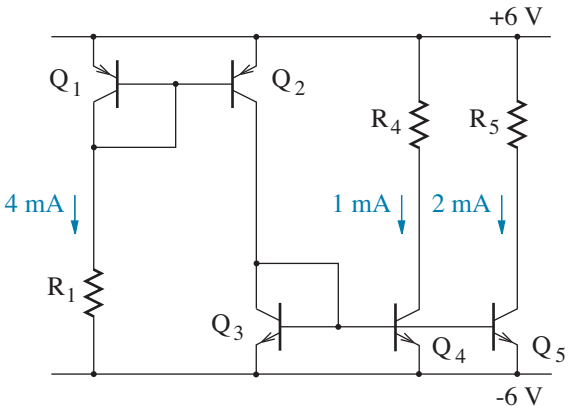


Figure P7.26

Section 7.2

7.27 A MOSFET with $K'W/L = 4\ \text{mA/V}^2$ and $V_T = 0.6$ V is biased such that $i_d|_Q = 2.5$ mA and $v_{ds}|_Q = 7.2$ V. Find g_m and r_π.

7.28 A MOSFET with $K'W/L = 6\ \text{mA/V}^2$ and $V_T = 0.8$ V has $g_m = 5.88 \times 10^{-3}$ ℧. Find the consistent Q-point component.

7.29 Prove the equivalence of n- and p-channel small-signal MOSFET models.

7.30 An npn BJT featuring $\beta_F = \beta_o = 220$ is biased such that $i_c|_Q = 3.2$ mA and $v_{ce}|_Q = 8$ V at room temperature. Find g_m and r_π.

7.31 A pnp BJT with $\beta_F = \beta_o = 85$ has $g_m = 0.1$ ℧ at $T = 5\,^\circ$C. Find the consistent Q-point component.

7.32 Consider a BJT with SPICE parameters IS, IKF, and BF.

(a) Show that

$$g_m \approx \frac{i_c|_Q}{kT/q}\left[\frac{\sqrt{1+\dfrac{4\,IKF}{i_c|_Q}}}{1+\sqrt{1+\dfrac{4\,IKF}{i_c|_Q}}}\right] .$$

(b) Use the SPICE .op statement to demonstrate the preceding result for a BJT with IS=10f, IKF=20m, BF=120, and $i_c|_Q = 15$ mA.

7.33 Consider a BJT with SPICE parameters IS, IKF, and BF.

(a) Show that

$$\beta_o \approx \frac{\beta_F}{1+\dfrac{i_c|_Q}{IKF}\left(1+\dfrac{i_c|_Q}{IKF}\right)^{-1}} .$$

(b) Use the SPICE .op statement to demonstrate the preceding result for a BJT with IS=10f, IKF=20m, BF=120, and $i_c|_Q = 28$ mA.

7.34 Chapter 6 used the Ebers-Moll model to show that for an npn BJT in the forward active mode,

$$i_c = I_{\text{common}} - \frac{I_s}{\beta_R}\left(e^{\,qv_{bc}/kT} - 1\right) ,$$

where

$$I_{\text{common}} = I_s\left(e^{\,qv_{be}/kT} - e^{\,qv_{bc}/kT}\right) .$$

(a) Derive expressions for small-signal parameters $g_{mr} = \partial i_c/\partial v_{bc}$ (the reverse transconductance) and $1/r_\mu = \partial i_b/\partial v_{bc}$. Make approximations.

Hint:

$$\frac{\partial i_b}{\partial v_{bc}} = \left[\frac{\partial i_b}{\partial(-i_e)}\right]\left[\frac{\partial(-i_e)}{\partial i_c}\right]\left[\frac{\partial i_c}{\partial v_{bc}}\right] .$$

(b) Show where the g_{mr}- and r_μ-related elements appear in the small-signal hybird-π model.

(c) A BJT with IS=10f and $\beta_R = 5$ is biased such that $v_{bc}|_Q = -5$ V. Find values for g_{mr} and r_μ at room temperature. Are they significant?

Section 7.3

Assume HUGE 1-Farad capacitors in all SPICE simulations (pending later developments in Chapter 8). The BJTs have IS=10f.

7.35 Use the generic hybrid-π model to show that an amplifier has no practical function when its input is at the drain (collector), its output is at the gate (base), and its source (emitter) is at ac ground.

7.36 Determine R_c', R_b', and R_e' in the circuit of Fig. P7.17.

7.37 The MOSFET in the amplifier of Fig. P7.37 features $K'W/L = 1.6$ mA/V^2 and $V_T = 0.6$ V. Determine the total midfrequency voltage gain.

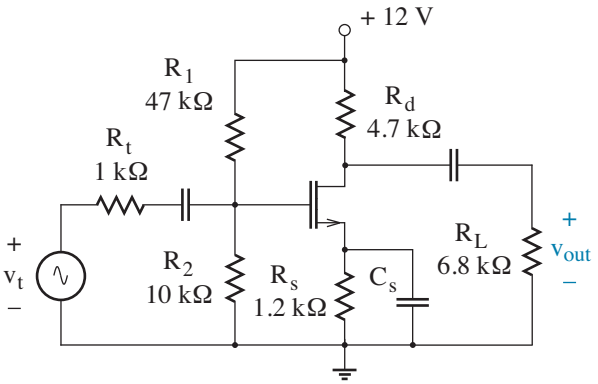


Figure P7.37

7.38 Repeat Problem P7.37, but use a 10-V supply.

7.39 Repeat Problem P7.37, but let $R_d = 3.3$ kΩ and remove C_s.

7.40 Consider the amplifier of Problem P7.37.

(a) Change R_1 and R_2 so that $A_{vm}^{Total} = -5$.

(b) Use SPICE to verify your design.

7.41 The BJT in the circuit of Fig. P7.41 has $\beta_F = \beta_o = 200$. Find the total midfrequency voltage gain.

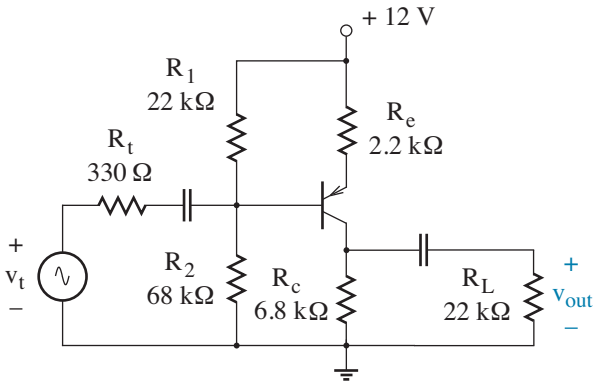


Figure P7.41

7.42 Repeat Problem P7.41, but use a 10-V supply.

7.43 Repeat Problem P7.41, but let $R_e = 3.3$ kΩ and add a parallel capacitor.

7.44 Consider the amplifier of Problem P7.41.

(a) Change R_e so that $A_{vm}^{Total} = -5$.

(b) Use SPICE to verify your design.

7.45 Prove that the midfrequency current gain (A_{im}) for an amplifier does not depend on the Thevenin resistance (R_t) of the input signal source.

7.46 The MOSFET in the amplifier of Fig. P7.46 has $K'W/L = 3.6$ mA/V^2 and $V_T = -0.5$ V. Find the total midfrequency voltage gain.

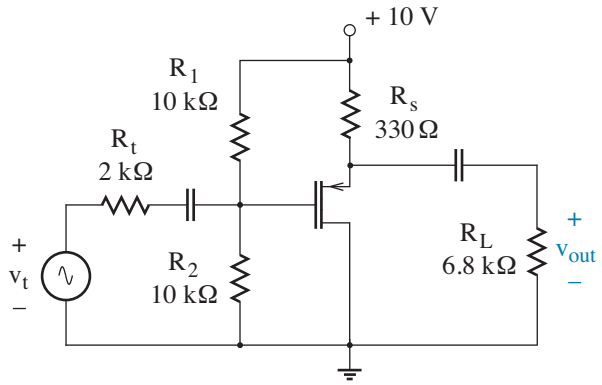


Figure P7.46

7.47 Repeat Problem P7.46, but let $R_2 = 18$ kΩ and $R_L = 4.7$ kΩ.

7.48 Consider the amplifier of Problem P7.46.

(a) Change R_s so that $A_{im} = 2.5$.

(b) Use SPICE to verify your design.

7.49 The BJT in the circuit of Fig. P7.49 has $\beta_F = \beta_o = 150$. Find the total midfrequency current gain.

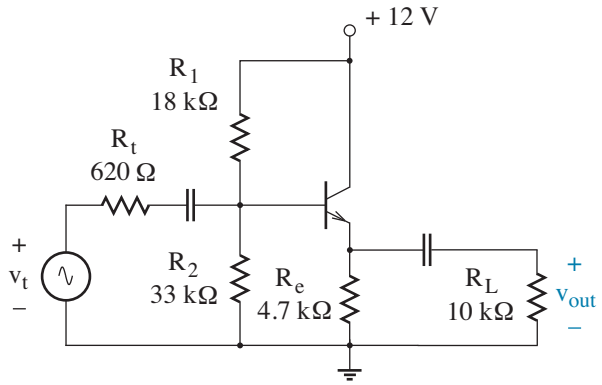


Figure P7.49

7.50 Repeat Problem P7.49, but let $R_e = 1.2$ kΩ and let $\beta_F = \beta_o = 60$.

7.51 Consider the amplifier of Problem P7.49.

(a) Change R_1 and R_2 so that $i_c|_Q = 1$ mA and the loading factor is 0.9.

(b) Use SPICE to verify your design.

7.52 This problem concerns a **bootstapping** bias technique that improves the input resistance of a common-collector amplifier. Consider the circuit of Fig. P7.52 in which $\beta_F = \beta_o = 80$.

(a) Complete a bias design such that $v_{ce}|_Q = 5$ V. To promote bias stability, make the dc Thevenin resistance in the circuit that connects to the BJT base equal to about $\beta_F/10$ times the dc resistance looking away from the emitter, and let $R_x = (R_1 \parallel R_2)/4$.

(b) Draw an appropriate small-signal circuit model, then find an expression and numerical value for the input resistance r_{in}.

(c) Repeat the biasing design with $v_{ce}|_Q = 5$ V, but let $R_x \to 0$ and eliminate C_x. Apply the design conditions of part a to promote bias stability.

(d) Find r_{in} for the design of part c and compare with that of part b.

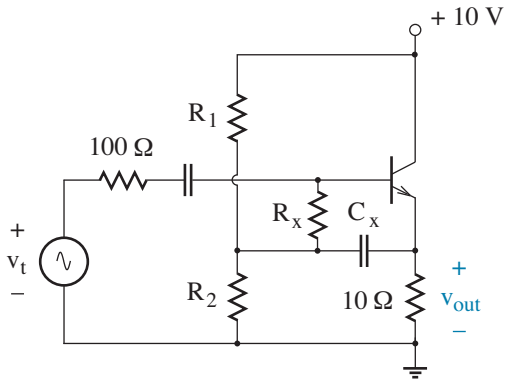


Figure P7.52

7.53 The MOSFET in the amplifier of Fig. P7.53 has $K'W/L = 2.8$ mA/V^2 and $V_T = 0.5$ V. Find the total midfrequency voltage gain.

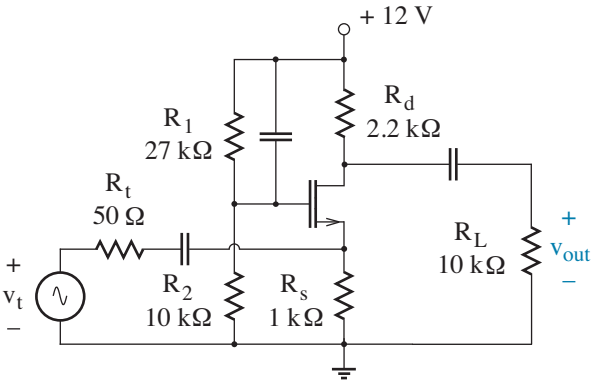


Figure P7.53

7.54 Repeat Problem P7.53, but use a 10-V supply.

7.55 Repeat Problem P7.53, but let $R_s = 1.8$ kΩ.

7.56 Consider the amplifier of Problem P7.53.

(a) Change R_1 and R_d so that $i_d|_Q = 2$ mA and $A_{vm}^{Total} = 8$.

(b) Use SPICE to verify your design.

7.57 The BJT in the circuit of Fig. P7.57 has $\beta_F = \beta_o = 120$. Find the total midfrequency voltage gain.

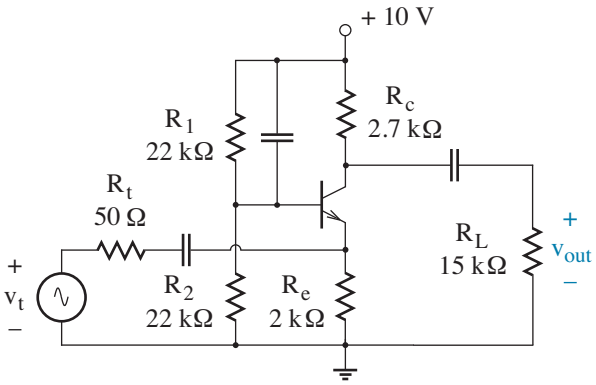


Figure P7.57

7.58 Repeat Problem P7.57, but let $R_e = 3.9$ kΩ.

7.59 Consider the amplifier of Problem P7.57.

(a) Change R_1 and R_2 so that the Q point is in the middle of the ac load line, and specify the total midfrequency gain that results.

(b) Use SPICE to verify your design.

7.60 Determine the total midfrquency voltage gain for the circuit of Fig. P7.60. The BJTs have $\beta_F = \beta_o = 100$, and quiescent collector currents are shown.

7.61 Determine the total midfrquency voltage gain for the circuit of Fig. P7.61. The BJT has $\beta_F = \beta_o = 100$, the MOSFET has $K'W/L = 4.2$ mA/V^2 and $V_T = 0.6$ V, and quiescent currents are shown.

7.62 Determine the total midfrquency voltage gain for the circuit of Fig. P7.62. The BJT has $\beta_F = \beta_o = 120$, the MOSFET has $K'W/L = 2.2$ mA/V^2 and $V_T = -0.6$ V, and quiescent currents are shown.

7.63 Determine the total midfrquency voltage gain for the circuit of Fig. P7.63. The BJT has $\beta_F = \beta_o = 140$, the MOSFET has $K'W/L = 1.8$ mA/V^2 and $V_T = 0.7$ V, and quiescent currents are shown.

Section 7.4

Assume HUGE 1-Farad capacitors in all SPICE simulations (pending later developments in Chapter 8). The BJTs have IS=10f.

7.64 Design a common-emitter amplifier with $A_{vm}^{Total} = -5$, $R_t = 220\ \Omega$, $R_L = 8.2$ kΩ, $\beta_F = \beta_o = 150$, and $V^+ = 15$ V. Verify with SPICE.

7.65 Design a common-emitter amplifier with $A_{vm}^{Total} = -120$, $R_t = 100\ \Omega$, $R_L = 4.7$ kΩ, $\beta_F = \beta_o = 120$, and $V^+ = 15$ V. Verify with SPICE.

7.66 Design a common-emitter amplifier with $A_{vm}^{Total} = -10$, $R_t = 180\ \Omega$, $R_L = 6.2$ kΩ, $\beta_F = \beta_o = 200$, and $V^+ = 15$ V. The total dissipated dc power must not exceed 50 mW. Verify with SPICE.

7.67 Design a common-emitter amplifier with $A_{vm}^{Total} = -150$, $R_t = 50\ \Omega$, $R_L = 2.2$ kΩ, $\beta_F = \beta_o = 65$, and $V^+ = 15$ V. Use a pnp BJT, and verify with SPICE.

7.68 Design a common-collector amplifier with $A_{im} = 10$, $R_t = 50\ \Omega$, $R_L = 8\ \Omega$, $\beta_F = \beta_o = 160$, and $V^+ = 15$ V. The quiescent load current is zero. Verify with SPICE.

7.69 Design a two-stage BJT amplifier for which:

- $A_{vm} = -4$.
- r_{in} (minimum) $= 120\ \Omega$.
- $R_L = 20\ \Omega$.
- The minimum acceptable output swing is ± 5 V.
- $V^+ = 36$ V.
- $\beta_F = \beta_o = 100$.

Specify all of the resistor values to no more than two significant figures. Verify the design with SPICE.

7.70 Design a two-stage BJT amplifier for which:

- $A_{vm} = +80$.
- r_{in} (minimum) $= 1$ kΩ.
- $R_L = 10$ kΩ.
- The minimum acceptable output swing is ± 5 V.
- $V^+ = 24$ V.
- $\beta_F = \beta_o = 200$.

Specify all of the resistor values to no more than two significant figures. Verify the design with SPICE.

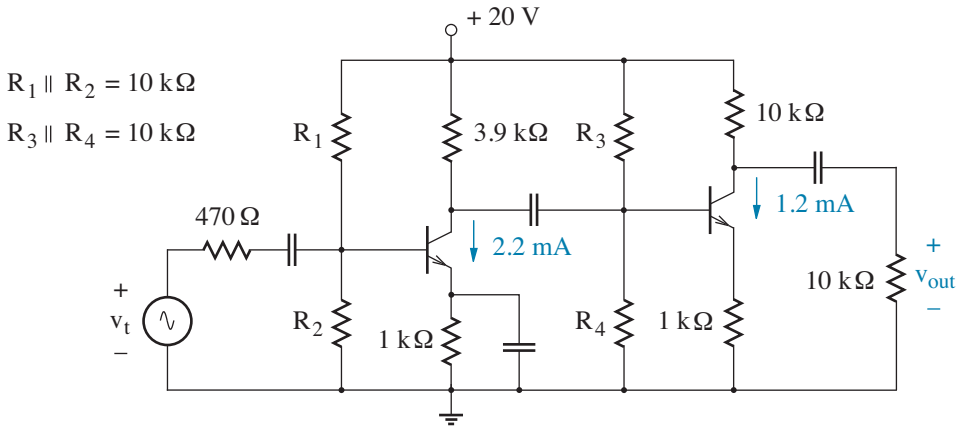


Figure P7.60

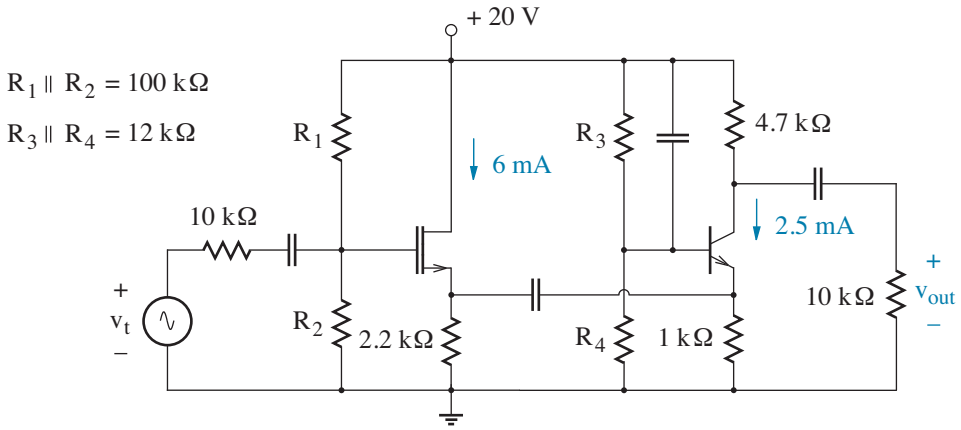


Figure P7.61

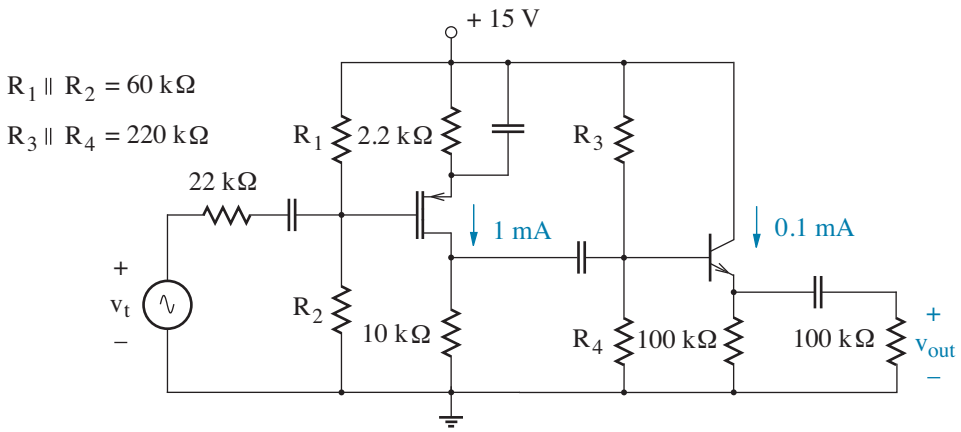


Figure P7.62

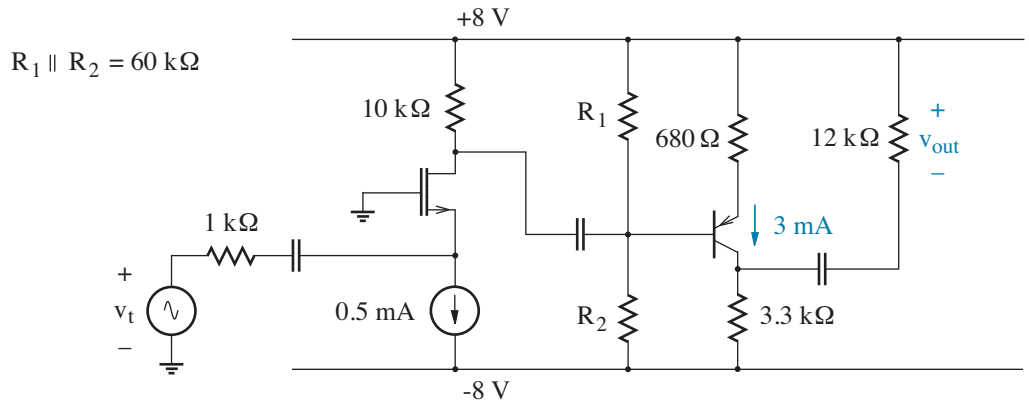
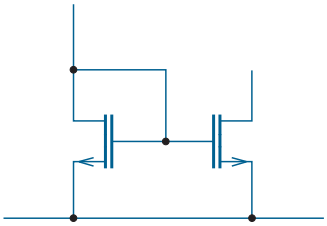


Figure P7.63



Chapter 8

Amplifier Frequency Response

Thus far, we have ignored the possibility of capacitance in the small-signal transistor models, effectively treating any capacitors as ac open circuits. And we have considered coupling and bypass capacitors as ac short circuits. This chapter abandons the restricted viewpoint. Our goal is to determine the onset of both high- and low-frequency deviations from midfrequency amplifier performance and to take corrective action, when possible.

Following appropriate modification of the small-signal transistor models, we develop a generalized approach to high-frequency analysis using simple time constants that are easily calculated. This motivates design techniques that put off the demise of midfrequency behavior. Time constants also play a central role in our consideration of low-frequency analysis and design.

Objectives

After studying this chapter and working through the problems, you should be able to perform each of the following:

- Specify small-signal models for the MOSFET and BJT that apply at high frequencies, and evaluate the capacitors therein (Section 8.1).
- Use the method of open-circuit time constants to determine or impose a particular high-frequency amplifier cutoff (Section 8.2).
- Analyze and design a cascode or tuned amplifier (Section 8.3).
- Use the method of short-circuit time constants to determine or impose a particular low-frequency amplifier cutoff (Section 8.4).

8.1 High-Frequency Transistor Models

Before we plunge into a discussion of high-frequency amplifier behavior, we need to consider little capacitances that were previously neglected in the small-signal transistor models used for midfrequency analysis. Our goal is to characterize charge-storage effects on the basis of simple measurements.

MOSFET Small-Signal Capacitance

At high-frequency, the first-order small-signal model for the three-terminal MOSFET includes two capacitors C_{gs} and C_{gd} (Fig. 8.1). Physical origins have been examined in Section 5.4. Both capacitors are Q-point dependent. We can often neglect the tiny capacitance between the far-removed source and drain regions as well as the interlead capacitances related to packaging.

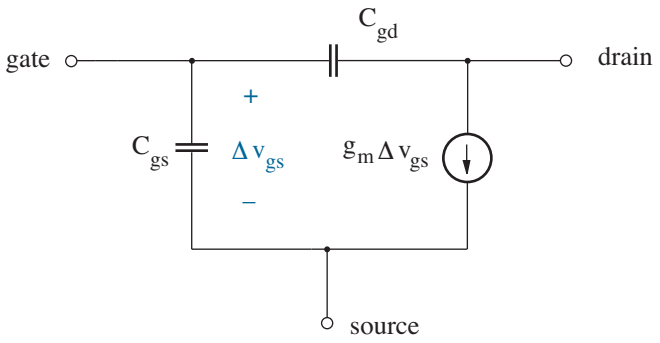


Figure 8.1: Three-terminal small-signal MOSFET model.

A first measurement of interest determines the admittance between the gate and drain terminals when the source is an open circuit for ac signals. In this case,

$$\jmath\omega C_{gs}\Delta v_{gs} + g_m \Delta v_{gs} = 0 \tag{8.1}$$

so that $\Delta v_{gs} = 0$. In turn, the dependent current source is an open circuit, and C_{gd} can be measured directly as

$$C_{gd} = C_{rss}\ . \tag{8.2}$$

A second measurement of interest determines the admittance between the gate and source terminals when the drain is an ac short to the source. The measured admittance is $\jmath\omega(C_{gs} + C_{gd})$, which is specified as $\jmath\omega C_{iss}$. Then in consideration of Eq. 8.2,

$$C_{gs} = C_{iss} - C_{rss}\ . \tag{8.3}$$

Example 8.1

A 2N7000 MOSFET has $C_{iss} = 20$ pF and $C_{rss} = 4$ pF. Find C_{gs} and C_{gd}.

Solution

The solution is trivial. By definition, $C_{gd} = C_{rss} = 4$ pF. Similarly, $C_{gs} = C_{iss} - C_{rss} = 16$ pF.

Integrated MOSFET Complications

Figure 8.2 shows the complete four-termal MOSFET model that includes C_{gb}, C_{bs}, and C_{bd} capacitive coupling in relation to the body terminal. These capacitances are also Q-point dependent, and they are evaluated as described in Section 5.4. Look to Chapter 9 for second-order complications.

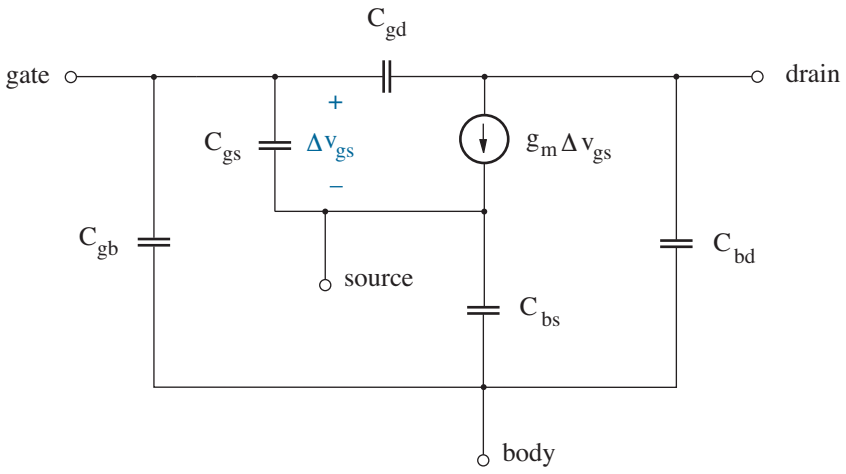


Figure 8.2: Four-terminal first-order small-signal MOSFET model.

The four-termal MOSFET model is mandatory for integrated circuits. Discrete MOSFET circuits tend to have the source and body terminals connected so that C_{bs} is eliminated. In turn, C_{gb} contributes to C_{gs}, and C_{bd} appears as an effective drain-to-source capacitance.

SPICE Parameters

SPICE parameters that describe small-signal MOSFET capacitance have been presented in Section 5.4 (see Tables 5.2 and 5.3).

BJT Small-Signal Capacitance

At high-frequency, the first-order small-signal BJT model has two oddly named capacitors C_π and C_μ (large-signal C_{be} and C_{bc}) as shown in Fig. 8.3. These capacitors are associated with the base-emitter and base-collector pn junctions, respectively. Unless we are concerned with very high frequencies, we neglect the relatively tiny capacitance between the far-removed emitter and collector regions and the interlead capacitances related to packaging.

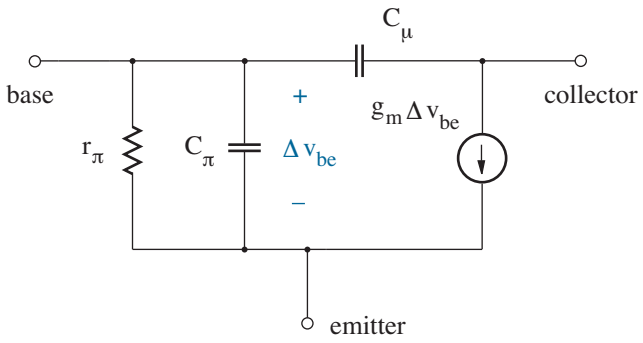


Figure 8.3: High-frequency small-signal model for the BJT.

In Chapter 6, we found that the BJT base-emitter and base-collector capacitances consisted of the sum of a depletion component and a diffusion component (Eqs. 6.73 and 6.74). Thus, for the BJT at hand,

$$C_\pi\,(C_{be}) = \underbrace{C_{je}(0)\,[1 - v_{be}/\phi_e]^{-m_e}}_{\text{depletion } C_{\pi d}} + \underbrace{\tau_f\,\frac{I_s}{kT/q}\,e^{\,qv_{be}/kT}}_{\text{diffusion } C_{\pi s}} \tag{8.4}$$

and

$$C_\mu\,(C_{bc}) = \underbrace{C_{jc}(0)\,[1 - v_{bc}/\phi_c]^{-m_c}}_{\text{depletion } C_{\mu d}} + \underbrace{\tau_r\,\frac{I_s}{kT/q}\,e^{\,qv_{bc}/kT}}_{\text{diffusion } C_{\mu s}}, \tag{8.5}$$

where ϕ_e and ϕ_c are built-in junction potentials, and $m_e = 0.5$, $m_c = 0.33$ (junction grades). Under small-signal conditions, C_π and C_μ are constant, since $v_{be} = v_{be}|_Q$ and $v_{bc} = v_{bc}|_Q$.

It is important to note the relative magnitudes of each of the diffusion capacitances when the BJT operates in the common forward active mode. With $v_{be}|_Q \gg kT/q$, the second-term exponent factor in Eq. 8.4 is large and positive. In turn, C_π features a *substantial* diffusion capacitance that tends to dominate the depletion capacitance at large collector currents. However, with $v_{bc}|_Q \ll -kT/q$, the similar exponent factor in Eq. 8.5 is large and negative. In turn, C_μ features a *negligible* diffusion capacitance, and the depletion capacitance dominates.

To better appreciate the physical basis for the C_π diffusion capacitance, it is constructive to recall the electron distribution in the base of an npn transistor that is biased in the forward active mode. Figure 8.4 shows a typical situation in which the injected electron concentration is very large at the base-emitter junction and almost zero at the base-collector junction. Electrical neutrality in the base is maintained through a similar distribution of holes that are supplied via positive base current. As noted in Chapter 6, this is analogous to a military situation: hole "troops" are needed to watch nervously over the army of electrons that is marching through p territory. The area between the charge concentration profiles that correspond to base-emitter voltages $v_{be}|_Q$ and $v_{be}|_Q + \Delta v_{be}$ is ΔQ_s, where $\Delta Q_s = C_{\pi s}\,\Delta v_{be}$. Parameter $C_{\pi s}$ relates differential voltage and charge, so it is a small-signal diffusion capacitance.

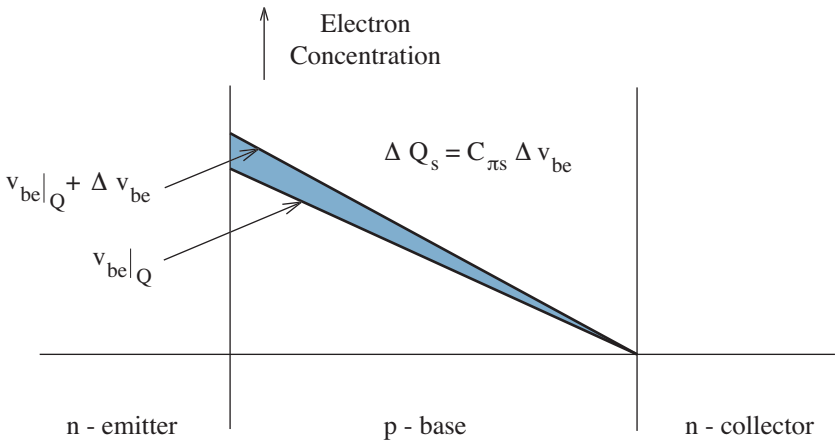


Figure 8.4: Charge storage in the base of an npn BJT that is biased in the forward active mode. Parameter $C_{\pi s}$ is the diffusion component of the small-signal base-emitter capacitance.

As indicated in Eq. 8.4, the C_π diffusion capacitance increases in proportion to parameter τ_f. In a BJT, this is the **forward transit time**—the time for an emitter-injected carrier to cross the base to the collector. It is straightforward to show that τ_f is proportional to the *square* of the base width (see Problem 8.5). Thus, narrow-base BJTs are advantageous for high-speed applications.

When the BJT is operated in the reverse active mode, the C_π diffusion capacitance is negligible and the C_μ diffusion capacitance is substantial. Parameter τ_r in Eq. 8.5 is the **reverse transit time**.

We now consider two measurements that determine C_π and C_μ for a BJT operating in the forward active mode.

Figure 8.5 shows a test circuit used to find the admittance between the BJT base and collector terminals when the emitter is an open circuit for small signals. At the emitter node,

$$\frac{\Delta v_{be}}{r_\pi} + j\omega C_\pi \Delta v_{be} + g_m \Delta v_{be} = 0\,, \tag{8.6}$$

which is consistent with $\Delta v_{be} = 0$. It follows that the dependent current source is an open circuit, and the measured admittance is simply $j\omega C_\mu$.

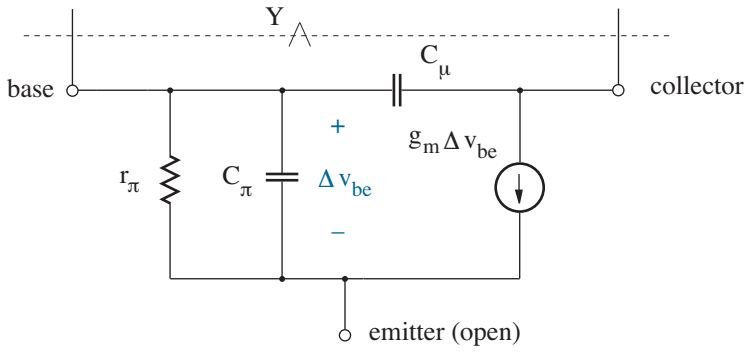


Figure 8.5: Test circuit for base/collector admittance measurement.

On the basis of the preceding measurement, most manufacturers specify C_μ as C_{obo}. Typical C_{obo} dependence on v_{bc} is generally provided on the BJT data sheet as shown in Fig. 8.6. Note the accompanying C_{ibo} data, which indicates C_π for a BJT operating in the *reverse active mode*.

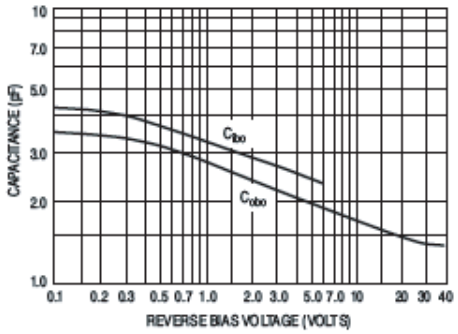


Figure 8.6: C_{obo} data for a 2N3904 BJT. Used with permission from SCILLC dba ON Semiconductor.

Figure 8.7 shows another test circuit used to measure the frequency dependence of $\beta = \Delta i_c / \Delta i_b$ when the collector and emitter experience a short circuit for small signals. In this case, $\Delta v_{ce} = 0$, and $\Delta v_{be} = \Delta v_{bc}$.

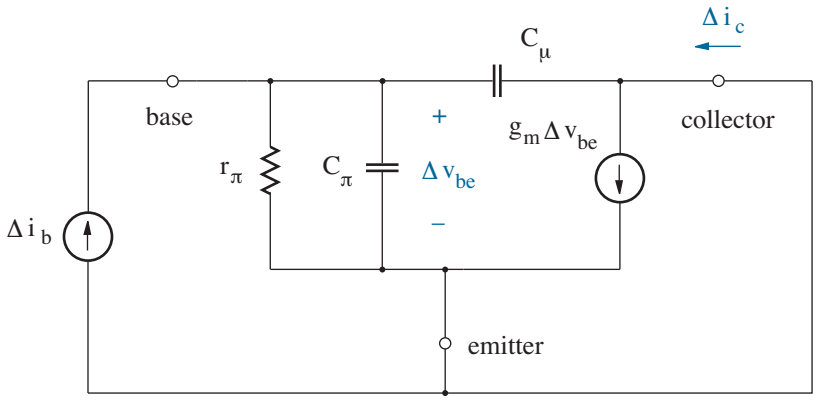


Figure 8.7: Test circuit for $\beta(\omega)$ measurement.

At the collector node,

$$\Delta i_c = g_m \Delta v_{be} - j\omega C_\mu \Delta v_{be} \, . \tag{8.7}$$

And at the base node,

$$\Delta i_b = g_\pi \Delta v_{be} + j\omega C_\pi \Delta v_{be} + j\omega C_\mu \Delta v_{be} \, , \tag{8.8}$$

where $g_\pi = 1/r_\pi$. So by dividing Eq. 8.7 by Eq. 8.8, we have

$$\beta(\omega) = \frac{\Delta i_c}{\Delta i_b} = \frac{g_m - j\omega C_\mu}{g_\pi + j\omega(C_\pi + C_\mu)} \, . \tag{8.9}$$

Subject to later verification, we assume $g_m \gg \omega C_\mu$. Then

$$\beta(\omega) \approx \frac{\beta_o}{1 + j\omega r_\pi (C_\pi + C_\mu)} \, , \tag{8.10}$$

since $g_m r_\pi = \beta_o$. For convenience, we rewrite Eq. 8.10 in the form

$$\beta(\omega) \approx \frac{\beta_o}{1 + j\omega/\omega_\beta} \, , \tag{8.11}$$

where

$$\omega_\beta = \frac{1}{r_\pi (C_\pi + C_\mu)} \, . \tag{8.12}$$

Figure 8.8 shows the Bode plot for $\beta(\omega)$. At low frequency, $\beta = \beta_o$. Then as frequency increases, we encounter a breakpoint at $\omega = \omega_\beta$, and β begins an asymptotic decrease at a rate of -20 dB per decade. Eventually, $\beta = 1$ (or $\log \beta = 0$) at $\omega = \omega_T$. The BJT loses its capability for current gain ($\beta < 1$) when $\omega > \omega_T$.

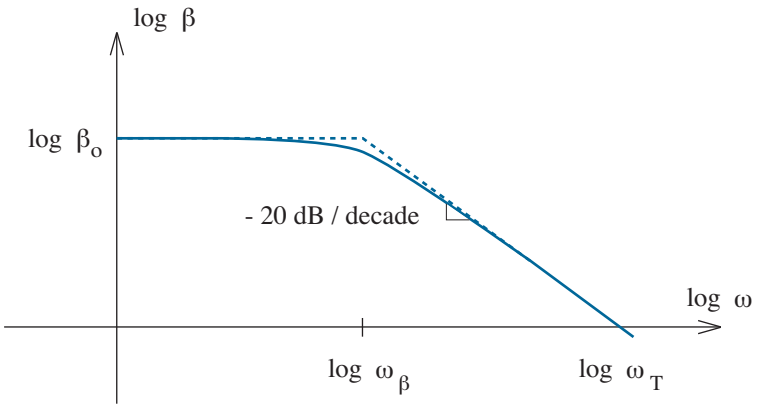


Figure 8.8: Bode plot for $\beta(\omega)$.

When applied to Eq. 8.11, the preceding definition of ω_T implies

$$\frac{\omega_T}{\omega_\beta} \approx \beta_o \, . \tag{8.13}$$

So in consideration of Eq. 8.12,

$$\omega_T = \frac{\beta_o}{r_\pi (C_\pi + C_\mu)} \, . \tag{8.14}$$

But $\beta_o / r_\pi = g_m$. Thus,

$$\omega_T = \frac{g_m}{C_\pi + C_\mu} \, . \tag{8.15}$$

In practice, ω_T is measured by extrapolating from the ω_β breakpoint in a partial Bode plot.

Before continuing, we must verify the assumption whereby $g_m \gg \omega C_\mu$. If we had retained the imaginary component in the numerator of Eq. 8.9, the Bode plot for $\beta(\omega)$ would have exhibited the effect of a zero. However,

$$\omega_{zero} = \frac{g_m}{C_\mu} \gg \frac{g_m}{C_\pi + C_\mu} = \omega_T \, , \tag{8.16}$$

so the zero bears little interest.

Further digression— With $v_{be} = v_{be}|_Q$,

$$\frac{I_s}{kT/q}\, e^{\,qv_{be}/kT} = g_m\,. \tag{8.17}$$

So Eq. 8.4 takes the form

$$C_\pi = C_{je} + g_m\tau_f\,, \tag{8.18}$$

where C_{je} is the depletion capacitance for the base-emitter junction. Similarly, with $v_{bc} = v_{bc}|_Q$,

$$\frac{I_s}{kT/q}\, e^{\,qv_{bc}/kT} = g_m{}'\,. \tag{8.19}$$

And Eq. 8.5 takes the form

$$C_\mu = C_{jc} + g_m{}'\tau_r\,, \tag{8.20}$$

where C_{jc} is the depletion capacitance for the base-collector junction. In the forward-active mode, $g_m{}' \approx 0$ and $C_\mu \approx C_{jc}$.

Equation 8.15 gives a relation for ω_T that is always valid. But in light of the preceding digression, the implied linear dependence on g_m is misleading. This is apparent when we substitute Eq. 8.18 for C_π and let $C_\mu \approx C_{jc}$ to obtain

$$\omega_T = \left(\tau_f + \frac{C_{je} + C_{jc}}{g_m}\right)^{-1}\,. \tag{8.21}$$

Thus, ω_T increases towards a steady maximum value as g_m becomes large (due to large $i_c|_Q$). New mechanisms cause ω_T to decrease again at very large collector currents. However, these need not concern us.

Most manufacturers specify $f_T = \omega_T/2\pi$ as the BJT **cutoff frequency**. The f_T measurement generally has large collector current—often indicated, otherwise implied—so that

$$f_T = \frac{\omega_{T,max}}{2\pi} \approx \frac{1}{2\pi\tau_f}\,. \tag{8.22}$$

In turn,

$$C_\pi \approx \frac{g_m}{2\pi f_T} + C_{je}\,. \tag{8.23}$$

After all this work, we find that the diffusion component of C_π can be determined from g_m and f_T, whereas we can only *estimate* the depletion component (C_{je}). Even a crude estimate is not necessary if $g_m/2\pi f_T \gg C_{je}$ (as is often the case). Otherwise, it is reasonable to take $C_{je} \approx C_{ibo}$ at zero bias voltage (see data of Fig. 8.6). Example 8.3 offers another estimate. Improved precision is best left to simulation tools such as SPICE.

Example 8.2

A 2N3904 BJT is biased such that $i_c|_Q = 5$ mA and $v_{bc}|_Q = -5$ V. Determine C_π and C_μ if $f_T = 300$ MHz.

Solution

We refer to Fig. 8.6 (from a 2N3904 data sheet) to find $C_\mu = C_{obo} = 2$ pF at the specified reverse voltage (5 V). With $g_m = i_c|_Q/(kT/q) = 0.193$ ℧, Eq. 8.20 yields $C_\pi = 102$ pF when C_{je} is ignored. The data of Fig. 8.6 specifies $C_{je} \approx C_{ibo} \approx 4.4$ pF at zero bias voltage. In turn, $C_\pi \approx 106$ pF. The correction is insignificant, particularly since f_T is an averaged value.

Example 8.3

A BJT features $C_\pi = 10$ pF when $v_{be} = 0$. Use SPICE to examine the variation of the C_π depletion capacitance with BJT collector current.

Solution

We simulate the circuit of Fig. 8.9. The BJT .model statement features typical IS=10f, CJE=10p, TF=0 (no diffusion capacitance), and CJC=0. We include the .op statement to find conditions at the BJT operating point, and we observe the value of CBE in the output file.

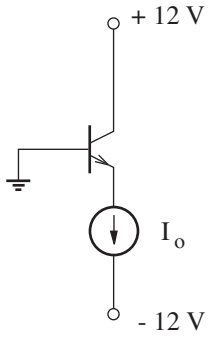


Figure 8.9: Test circuit for Example 8.3.

Here are the results:

10 μA (14.4 pF), 100 μA (15.0 pF), 1 mA (15.7 pF), 10 mA (16.3 pF)

So the C_π depletion component is about 1.5 times the depletion capacitance over a broad range of quiescent collector currents.

8.2 High-Frequency Response

As operating frequencies become large, the capacitors in the small-signal transistor models evolve into ac short circuits, thus reducing the prospects for voltage or current gain. This is quite apparent in the case of C_{gs} (C_π), which shorts out Δv_{gs} (Δv_{be}) in the circuit model of Fig. 8.1 (Fig. 8.3). In turn, the effect of the small-signal dependent current source diminishes. The influence of C_{gd} (C_μ) is less obvious, but it is often more significant.

High-Frequency Cutoff

Why not try this—

The small-signal amplifier is a linear circuit, so perform a comprehensive analysis in the frequency domain: Replace the capacitors with equivalent impedances, specifically $1/sC_x$, where s is the Laplace transform variable. Use standard techniques of resistive circuit analysis to determine $A_v^{\text{Total}}(s)$, the s-dependent total voltage gain. Make a pole-zero plot, which shows the values of s for which $A_v^{\text{Total}} = 0$ (a zero) or $A_v^{\text{Total}} \to \infty$ (a pole). Finally, construct a Bode plot, which shows the frequency response when $s = \jmath\omega$. All very straightforward. But what an algebraic nightmare when more than two or three capacitors are present.

Fortunately, we do not require all of the comprehensive analytical data. As noted previously, small-signal capacitors tend to limit the prospects for voltage gain, perhaps with the frequency dependence shown in Fig. 8.10. We arbitrarily define the **high-frequency cutoff** (f_h) as the frequency at which the midfrequency voltage (power) gain is reduced by $1/\sqrt{2}$ (1/2). At frequencies above f_h, the amplifier is assumed to have lost its intended function, and further details of frequency behavior often bear little interest. We are content to know only f_h (and an *approximate* procedure to find it).

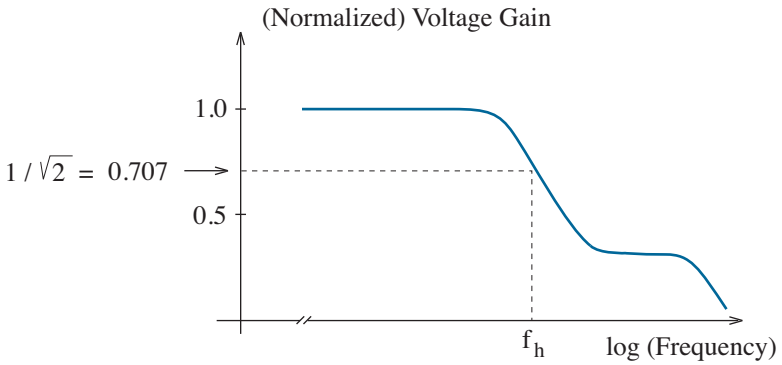


Figure 8.10: High-frequency voltage gain for a particular amplifier.

Looking Around

As in Chapter 7, it will prove helpful to adopt a special notation for an ac resistance to ground, now looking *into* each transistor terminal. Specifically,

r_d' is the ac resistance to ground looking into the drain.

r_g' is the ac resistance to ground looking into the gate.

r_s' is the ac resistance to ground looking into the source.

Figure 8.11 shows the orientation of the three looking-in primed resistances. The BJT counterparts are r_c', r_b', and r_e'.

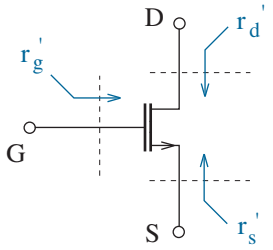


Figure 8.11: Primed ac resistances looking into the drain, gate, and source.

You will do well to memorize Table 8.1. Each of the entries is valid in the absence of separate active feedback paths between the transistor terminals (see Problem 8.8). Chapter 9 modifies r_d' and r_c' for integrated circuits.

Table 8.1: Small-signal "looking in" ac resistances.

Terminal	MOSFET	BJT
drain/collector (r_d'/r_c')	∞	∞
gate/base (r_g'/r_b')	∞	$r_\pi + (1+\beta_o)R_e'$
source/emitter (r_s'/r_e')	$\dfrac{1}{g_m}$	$\dfrac{r_\pi + R_b'}{1+\beta_o}$

For the BJT, upper-case R_e' and R_b' are the Thevenin resistances looking away from the emitter and base terminals, respectively.

Exercise 8.1 Determine the small-signal resistances to ground specified in the circuits of Fig. 8.12. Assume $g_m = 4 \times 10^{-3}$ ℧ for the MOSFETs, $r_\pi = 3.2$ kΩ and $\beta_o = 99$ for the BJTs.

Ans: (a) $r = 1$ kΩ (b) $r = 200$ Ω (c) $r = 20$ kΩ (d) $r = 188$ Ω
(e) $r = 2$ kΩ (f) $r = 16.8$ kΩ

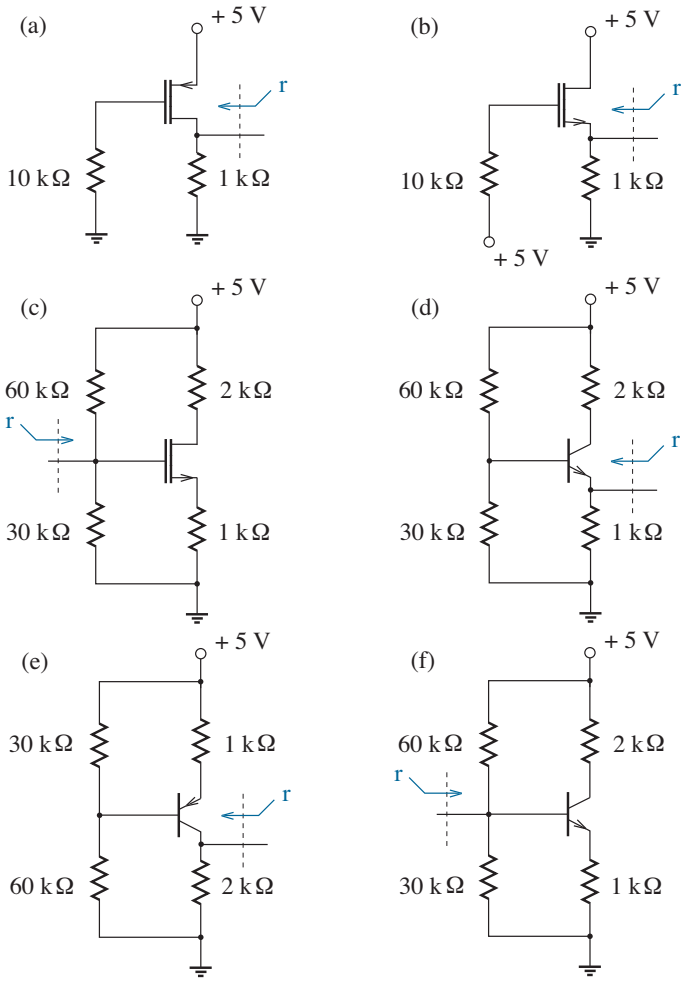


Figure 8.12: Circuits for Exercise 8.1.

Common-Gate (Common-Base) Amplifier

The common-gate (common-base) amplifier features small input resistance, so it is not particularly fashionable as a stand-alone circuit. Nevertheless, high-frequency analysis is relatively easy by virtue of the grounded gate, which allows C_{gs} and C_{gd} to hang like tree ornaments as shown in Fig. 8.13.

(There is no need to draw a complete small-signal circuit—our experience at midfrequency is sufficient to account for other small-signal components.)

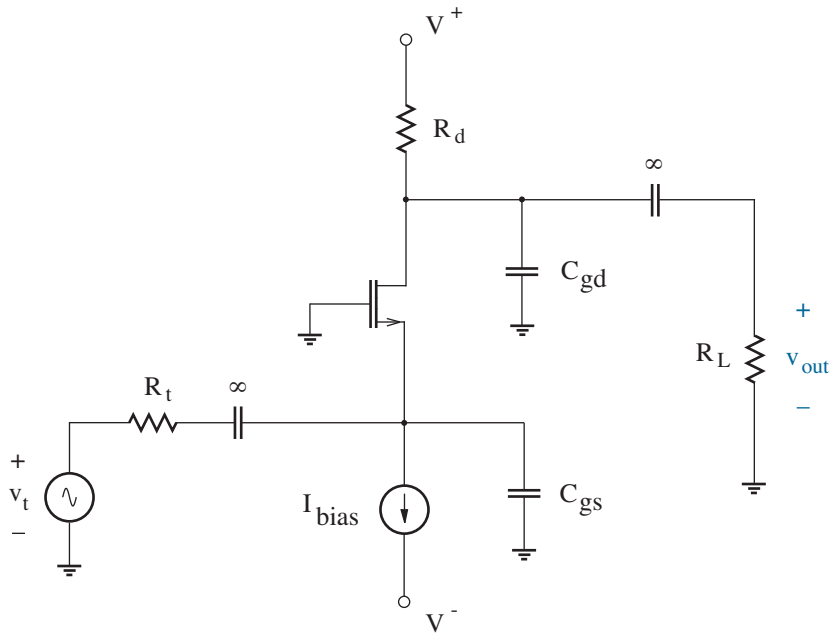


Figure 8.13: Common-gate amplifier at high-frequency. The C_{gs} and C_{gd} couplings to ground reflect the same connection at the MOSFET gate.

It is straightforward to determine the frequency-dependent total voltage gain for the circuit of Fig. 8.13, since two parts of the circuit act separately: C_{gs} decreases the loading factor; C_{gd} reduces the effective load resistance. The normalized total voltage gain has the form

$$\frac{A_v^{\text{Total}}}{A_{vm}^{\text{Total}}} = \left(\frac{1}{1+j\omega\tau_1}\right)\left(\frac{1}{1+j\omega\tau_2}\right), \tag{8.24}$$

where

$$A_{vm}^{\text{Total}} = \left[\frac{1/g_m}{R_t + 1/g_m}\right] g_m(R_d \parallel R_L). \tag{8.25}$$

The time constants in Eq. 8.24 are given by (see Problem 8.9)

$$\tau_1 = (R_t \parallel 1/g_m)\, C_{gs}\,, \tag{8.26}$$

and

$$\tau_2 = (R_d \parallel R_L)\, C_{gd}\,. \tag{8.27}$$

Note that each time constant has physical significance as the product of a capacitance and the equivalent Thevenin resistance "seen" by that capacitor under midfrequency conditions (see Table 8.1). Typically, $\tau_1 \ll \tau_2$.

We restrict attention to the cutoff frequency (f_h) where the normalized total voltage gain decreases by $1/\sqrt{2}$. Then with

$$\frac{A_v^{\text{Total}}}{A_{vm}^{\text{Total}}} = \frac{1}{1 + j\omega(\tau_1 + \tau_2) - \omega^2\tau_1\tau_2}\,, \tag{8.28}$$

the angular cutoff frequency (ω_h) satisfies

$$\sqrt{(1 - \omega^2\tau_1\tau_2)^2 + \omega^2(\tau_1 + \tau_2)^2} = \sqrt{2}\,. \tag{8.29}$$

The solution to Eq. 8.29 is not particularly difficult. Nevertheless, we can obtain an adequate estimate for f_h by neglecting the $\omega^2\tau_1\tau_2$ term. In turn,

$$f_h = \frac{\omega_h}{2\pi} \approx \frac{1}{2\pi(\tau_1 + \tau_2)}\,. \tag{8.30}$$

Figure 8.14 shows the percent error of the f_h estimate as a function of α when $\tau_1 = \tau$ and $\tau_2 = \alpha\tau$. The error is maximum when the τ values are equal and about -10% when the τ values differ by an order of magnitude. The negative percent error indicates a conservative underestimate.

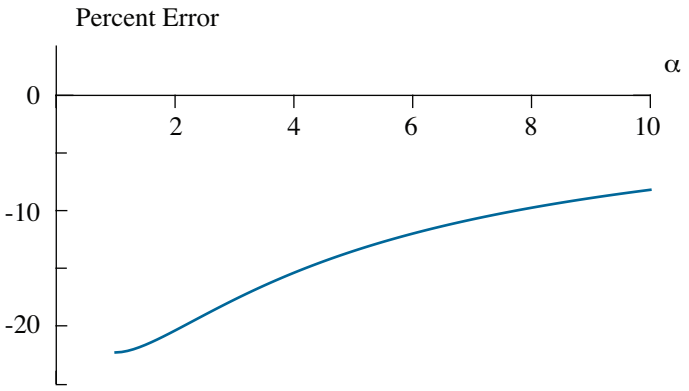


Figure 8.14: Estimated high-frequency cutoff error for $\tau_1 = \tau$, $\tau_2 = \alpha\tau$.

The Method of Open-Circuit Time Constants

Equation 8.30 provided a rough estimate of the high-frequency cutoff of a particular amplifier in terms of two time constants that are easily calculated. In general, we can estimate the high-frequency cutoff of capacitively loaded amplifiers using the expression

$$f_{h1} \approx \frac{1}{2\pi}\left(\sum \tau_{oc}\right)^{-1}, \tag{8.31}$$

where the τ_{oc} are **open-circuit time constants** given by the product of a capacitance and the equivalent resistance "seen" by that capacitor when *all other capacitors are treated as ac open circuits (as at midfrequency).*

Caution: The estimates exclude amplifier coupling and bypass capacitors. At frequencies in the vicinity of f_h, these capacitors are ac short circuits.

Proof [1]

Consider a large amplifier circuit with n capacitors that influence the total voltage gain at high frequencies. The system function that describes the total voltage gain must satisfy three conditions: (1) There are n poles—one for each capacitor. (2) The total voltage gain is constant at arbitrarily low frequencies (midfrequency) when the capacitors act as ac open circuits. (3) There is at least one less zero than the number of poles so that the total voltage gain vanishes at infinite frequency. Subject to evaluation at $s = \jmath\omega$, the required system function takes the form

$$A(\jmath\omega) = K\,\frac{(1+\jmath\omega\tau_{z1})(1+\jmath\omega\tau_{z2})\ldots(1+\jmath\omega\tau_{z[n-1]})}{(1+\jmath\omega\tau_1)(1+\jmath\omega\tau_2)(1+\jmath\tau_3)\ldots(1+\jmath\omega\tau_n)}, \tag{8.32}$$

where K is a constant. Pending future verification, we cautiously assume that the system zeros occur at relatively large $|1/\tau_z|$ values in relation to the poles. Then we expand the denominator of Eq. 8.32 to obtain

$$A(\jmath\omega) \approx \frac{K}{1+\jmath\omega\sum\tau-\omega^2\left[\left(\sum\tau\right)^2-\sum\tau^2\right]/2\;+\;\ldots}. \tag{8.33}$$

The high-frequency cutoff applies when the complex denominator in Eq. 8.33 has $\sqrt{2}$ magnitude. A rough estimate ignores powers of ω^2 or higher, and

$$f_h \approx \frac{1}{2\pi}\left(\sum\tau\right)^{-1}. \tag{8.34}$$

What remains is to determine the sum of the τ factors.

[1] Portions of this argument are adapted from R. D. Thornton, C. L Searle, D. O Pederson, R. B. Adler, and E. J. Angelo, Jr., *Multistage Transistor Circuits*, Semiconductor Electronics Education Committee, Vol. 5, (Wiley, New York, 1965), pp. 13 - 16.

Consider an n-port linear resistive network that is connected to external capacitors as shown in Fig. 8.15 (for a particular $n = 3$). If all independent sources within the network are turned off and the capacitors are removed, then in general

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} R_{11} & R_{12} & \dots & R_{1n} \\ R_{21} & R_{22} & \dots & R_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ R_{n1} & R_{n2} & \dots & R_{nn} \end{pmatrix} \begin{pmatrix} i_1 \\ i_2 \\ \vdots \\ i_n \end{pmatrix} . \tag{8.35}$$

Note that the diagonal terms in the **R** matrix indicate the resistance looking into a terminal pair if all of the other terminal pairs exhibit open circuits. For example, $v_1 = R_{11} i_1$ at terminal pair 1 in Fig. 8.15 when $i_2 = i_3 = 0$.

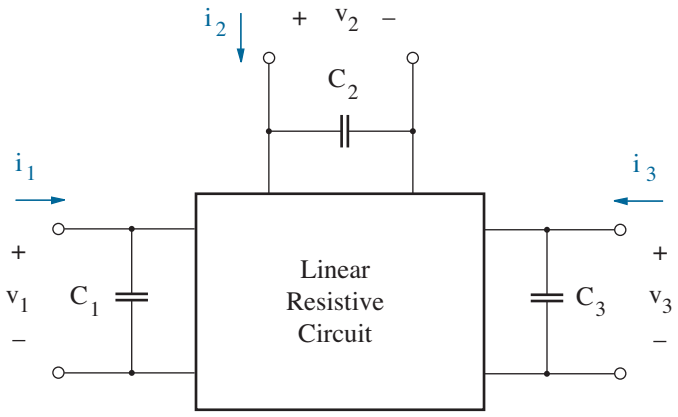


Figure 8.15: General network model for high-frequency calculations.

Alternatively, we can express terminal currents in terms of voltages:

$$\begin{pmatrix} i_1 \\ i_2 \\ \vdots \\ i_n \end{pmatrix} = \begin{pmatrix} G_{11} & G_{12} & \dots & G_{1n} \\ G_{21} & G_{22} & \dots & G_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ G_{n1} & G_{n2} & \dots & G_{nn} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} . \tag{8.36}$$

Here, the diagonal terms in the **G** matrix indicate the conductance looking into a terminal pair if all of the other terminal pairs exhibit short circuits. For example, $i_1 = G_{11} v_1$ at terminal pair 1 in Fig. 8.15 when $v_2 = v_3 = 0$. The matrix product **RG** is the identity matrix

$$\mathbf{I} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} . \tag{8.37}$$

When the external capacitors are restored, a current at a particular port (as defined in Fig. 8.15) must include the capacitor current at that port. Thus, Eq. 8.36 is modified to take the form

$$\mathbf{i} = [\mathbf{G} + s\mathbf{C}]\,\mathbf{v}\,. \tag{8.38}$$

In this expression, $\mathbf{i}$ and $\mathbf{v}$ are the current and voltage vectors, respectively, and $\mathbf{C}$ is a capacitance matrix given by

$$\mathbf{C} = \begin{pmatrix} C_1 & 0 & \dots & 0 \\ 0 & C_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & C_n \end{pmatrix}. \tag{8.39}$$

The system poles are the values of s that allow $\mathbf{v} \neq 0$ when $\mathbf{i} = 0$, and they are found by determining the roots to the equation

$$det\,[\mathbf{G} + s\mathbf{C}] = 0\,. \tag{8.40}$$

The determinant in Eq. 8.40 is greatly simplified when we observe that for arbitrary square matrices $\mathbf{A}$ and $\mathbf{B}$,

$$det\,\mathbf{AB} = det\,\mathbf{A}\;det\,\mathbf{B}\,. \tag{8.41}$$

So we multiply Eq. 8.40 by matrix $\mathbf{R}$ to obtain

$$det\,[\mathbf{I} + s\mathbf{RC}] = 0 \tag{8.42}$$

or

$$det \begin{pmatrix} 1 + sR_{11}C_1 & sR_{12}C_2 & \dots & sR_{1n}C_n \\ sR_{21}C_1 & 1 + sR_{22}C_2 & \dots & sR_{2n}C_n \\ \vdots & \vdots & \ddots & \vdots \\ sR_{n1}C_1 & sR_{n2}C_2 & \dots & 1 + sR_{nn}C_n \end{pmatrix} = 0\,. \tag{8.43}$$

The s^0 (constant) and s^1 terms in the expression for the determinant derive exclusively from a portion of the product of the diagonal matrix elements. Thus, the condition for the poles implies that the denominator of the system function has the form

$$D = 1 + \jmath\omega\,(R_{11}C_1 + R_{22}C_2 + \ldots + R_{nn}C_n) + \ldots \tag{8.44}$$

where $s = \jmath\omega$. When compared with Eq. 8.33, it follows that

$$R_{11}C_1 + R_{22}C_2 + \ldots + R_{nn}C_n = \sum \tau\,. \tag{8.45}$$

Thus, the sum of the τ_{oc} open-circuit time constants defined previously through the interpretation of R_{xx} is the same as the sum of the τ factors. Notwithstanding, the individual τ_{oc} and τ factors are usually different.

Common-Source (Common-Emitter) Amplifier

Figure 8.16 shows the genuine common-source amplifier at high frequency. As for the common-gate amplifier, C_{gs} hangs like a tree ornament to ground since one of the transistor terminals (in this case, the source) is at ac ground. No such luck with C_{gd}, which couples the output to the input.

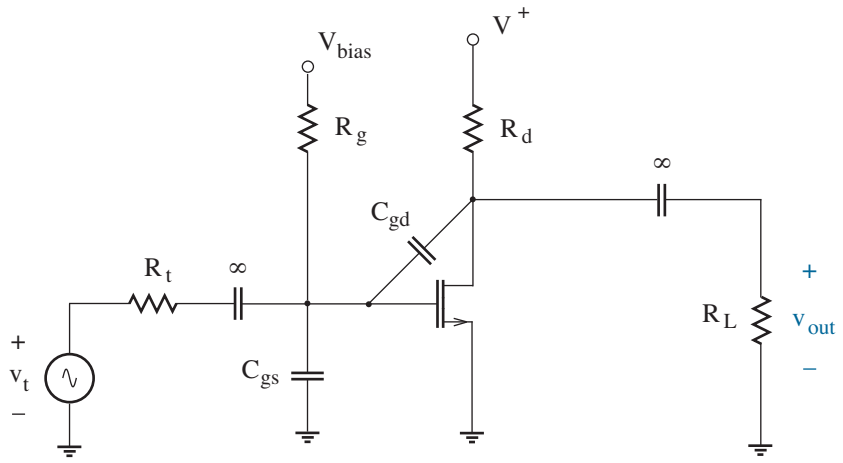


Figure 8.16: Common-source amplifier at high-frequency.

The ac resistance "seen" by C_{gs} is simply $R_g{}' \parallel r_g{}' = R_t \parallel R_g \parallel \infty$ when C_{gd} is an open circuit. Thus,

$$\tau_1 = (R_T \parallel R_g)\, C_{gs}\,. \tag{8.46}$$

As for C_{gd}, it is tempting to assume the same ac resistance to ground when looking to the left and $R_d{}' \parallel r_d{}' = R_d \parallel R_L \parallel \infty$ when looking to the right. In turn, one might propose

$$\tau_2 = (R_T \parallel R_g + R_d \parallel R_L)\, C_{gd} \tag{8.47}$$

as a time constant consistent with Fig. 8.17. But this analysis is *invalid* since it ignores the active influence of the parallel-connected MOSFET.

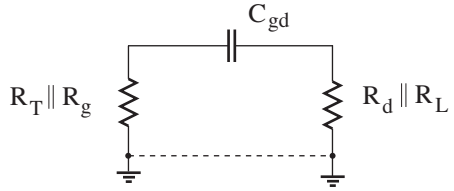


Figure 8.17: Invalid ac circuit for C_{gd}.

A valid analysis considers the circuit transformation shown in Fig. 8.18. The starting circuit (a) features impedance Z in a feedback path from the output to the input of an amplifier for which $v_2 = {A_{vm}}^F v_1$, where ${A_{vm}}^F$ is the **forward midfrequency voltage gain.** The current i_1 that is drawn from the v_1 input is given by

$$i_1 = \frac{v_1 - v_2}{Z} = \frac{v_1(1 - {A_{vm}}^F)}{Z} \,. \tag{8.48}$$

However, the same current flows away from the v_1 input through impedance $Z/(1-{A_{vm}}^F)$ in the transformed circuit (b). So the circuits are equivalent. Similarly, the current i_2 that is drawn from the v_2 output is given by

$$i_2 = \frac{v_2 - v_1}{Z} = \frac{v_2(1 - {A_{vm}}^R)}{Z} \,, \tag{8.49}$$

where ${A_{vm}}^R$ is the **reverse midfrequency voltage gain.** Again, the same current flows away from the v_2 output through impedance $Z/(1-{A_{vm}}^R)$ in the transformed circuit (b). This splitting of a feedback impedance into two effective input and output impedances reflects a **Miller transformation**. When the impedance is that of a capacitor ($Z = 1/j\omega C$), the transformed input and output capacitances feature the factors $1-{A_{vm}}^F$ and $1-{A_{vm}}^R$, respectively.[2] The former is often large; the latter is often at or near unity.

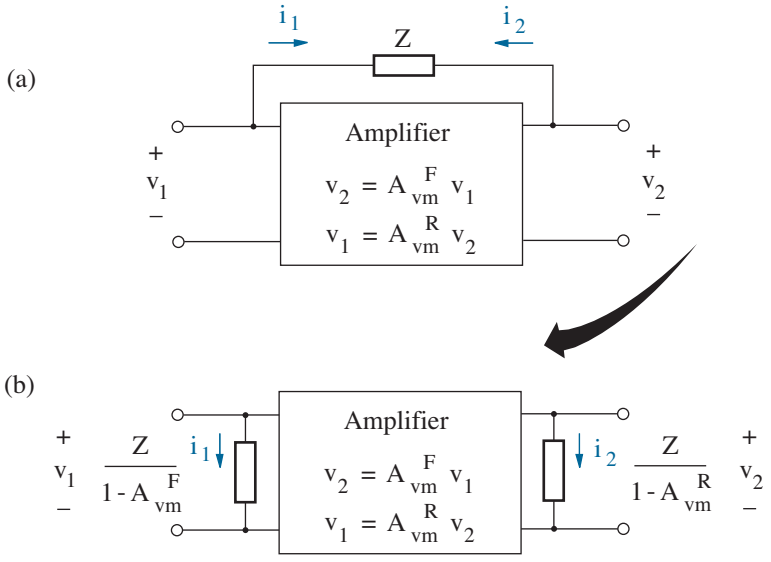


Figure 8.18: Miller transformation for an amplifier feedback impedance.

[2]Warning! Many texts incorrectly assume ${A_{vm}}^R = 1/{A_{vm}}^F$.

In the case of the common-source amplifier, the forward midfrequency voltage gain is ${A_{vm}}^F = -g_m(R_d \parallel R_L)$ (the ordinary A_{vm} of Chapter 7), and the reverse midfrequency voltage gain is ${A_{vm}}^R = 0$ (see Problem 8.15). Thus, C_{gd} transforms to $C_{gd}\,[1 + g_m(R_d \parallel R_L)]$ at the amplifier input and C_{gd} at the amplifier output. Figure 8.19 shows the new circuit.

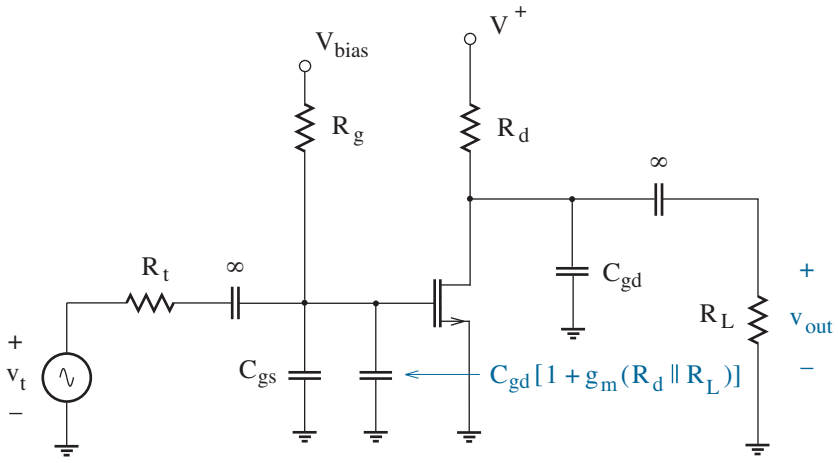


Figure 8.19: The circuit of Fig. 8.16 after the Miller transformation.

The remaining analysis uses the method of open-circuit time constants. At the input, we find

$$\tau_1 = (R_t \parallel R_g)\{C_{gs} + C_{gd}\,[1 + g_m(R_d \parallel R_L)]\}\,. \tag{8.50}$$

And at the output,

$$\tau_2 = (R_d \parallel R_L)\,C_{gd}\,. \tag{8.51}$$

Finally,

$$f_h \approx \frac{1}{2\pi(\tau_1 + \tau_2)}\,. \tag{8.52}$$

We would have found the same time-constant sum, but with greater effort, had we derived the equivalent resistance "seen" by C_{gd} (see Problem 8.16).

The C_{gd} feedback capacitor plays a particularly pernicious role in the limitation of the high-frequency cutoff of a common-source amplifier since its seemingly innocuous value becomes multiplied by a factor of $(1 - {A_{vm}}^F)$. In a so-called **broadband** amplifier, the **bandwidth** is approximately f_h when the low-frequency cutoff (f_l) is negligible. One achieves large gain at the expense of bandwidth, and the reverse tradeoff applies as well.

Example 8.4

Estimate the high-frequency cutoff for the common-emitter amplifier of Fig. 8.20, then compare with the SPICE result. Assume $f_T = 800$ MHz, $C_{ibo} = 4$ pF, $\phi_e = 0.95$ V, $C_{obo} = 3$ pF, and $\phi_c = 0.75$ V.

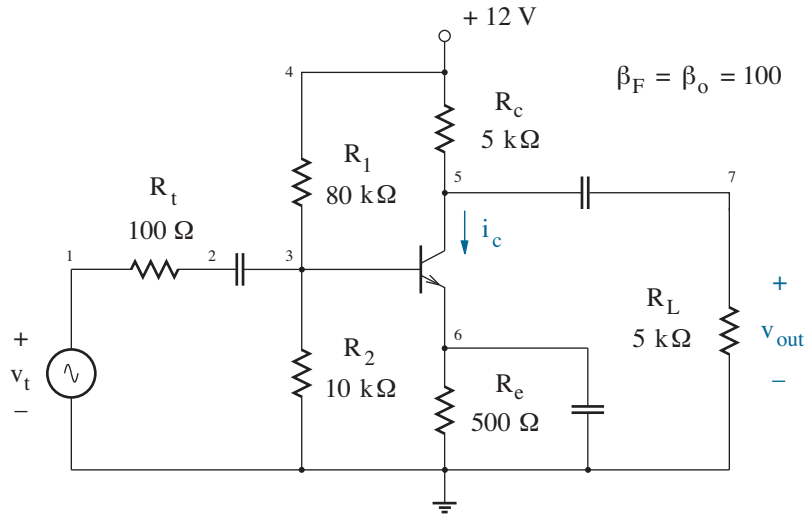


Figure 8.20: Circuit for Example 8.4.

Solution

A quick dc analysis—no need for details—reveals that $i_c|_Q = 1.06$ mA. In turn, $g_m = i_c|_Q/(kT/q) = 40.9 \times 10^{-3}$ ℧ and $r_\pi = \beta_o/g_m = 2.44$ kΩ. The forward midfrequency voltage gain is ${A_{vm}}^F = -g_m(R_c \parallel R_L) = -102$.

For C_π, we use Eq. 8.23 and the result of Example 8.3 to find

$$C_\pi \approx \frac{g_m}{2\pi f_T} + 1.5\,C_{ibo} = 14.1 \text{ pF}\,.$$

For C_μ, we first obtain $v_{bc}|_Q = (i_c|_Q R_e + 0.7) - (12 - i_c|_Q R_c) = -5.47$ V. Then with the help of Eq. 8.5,

$$C_\mu = C_{obo}\left(1 - \frac{v_{bc}|_Q}{\phi_c}\right)^{-0.33} = 1.49 \text{ pF}\,.$$

The latter capacitance is small, but its role is augmented through feedback. So we subject it to a Miller transformation and combine with C_π to obtain $C_1 = C_\pi + C_\mu(1+102) = 168$ pF as the total capacitance between the BJT base and ground. The reverse common-emitter amplifier gain is zero, so we obtain $C_2 = C_\mu(1-0) = 1.49$ pF as the total capacitance between the BJT collector and ground.

Now find the open-circuit time constants. At the input (with $r_b' = r_\pi$),

$$\tau_1 = (R_t \parallel R_1 \parallel R_2 \parallel r_\pi)\, C_1 = 16.0 \text{ ns}\,.$$

And at the output (with $r_c' = \infty$),

$$\tau_2 = (R_c \parallel R_L)\, C_2 = 3.7 \text{ ns}\,.$$

So finally, $f_h \approx 1/2\pi(\tau_1 + \tau_2) = 8.1$ MHz.

The requisite SPICE analysis is frequently used for analog circuit design, so we list an appropriate code.

```
* Test Circuit - Example 8.4

V+      4       0       +12
Vt      1       0       ac      1m
Rt      1       2       100
Cin     2       3       1
R1      4       3       80k
R2      3       0       10k
Q1      5       3       6       BJTN
Rc      4       5       5k
Cout    5       7       1
RL      7       0       5k
Re      6       0       500
Ce      6       0       1

.model  BJTN    NPN     (IS = 10f, BF = 100, TF = 199p
+ CJE=4p, VJE=0.95, CJC=3p, VJC=0.75)

.op
.ac     DEC     100     10k     100MEG
.probe

.end
```

The V_t descriptor reflects an ac source with *arbitrary* 1-mV amplitude—the actual value is irrelevant when voltage gains are measured as the ratio of voltage amplitudes at two different nodes within the .probe environment. The .ac control statement sweeps the signal-source frequency from 10 kHz to 100 MHz with 100 points per decade. For the BJTN .model statement, we determine TF=199p ($= 1/2\pi f_T$), and we let IS=10f (a reasonable value). Since we have no interest in low-frequency behavior, we assign 1-F values to the coupling and bypass capacitors. These HUGE specifications ensure that the capacitors function as ac short circuits at midfrequency.

The preceding code supports the .probe plot shown in Fig. 8.21.

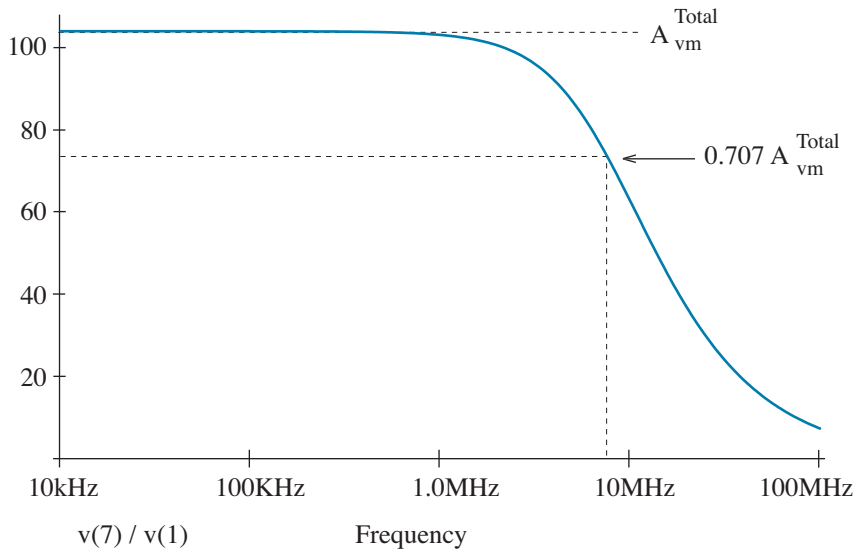


Figure 8.21: SPICE .probe plot for Example 8.4. The voltage gain is the ratio of two ac node amplitudes.

To find the high-frequency cutoff, we note that the magnitude of the total midfrequency voltage gain is 104. Then at the cutoff, the voltage-gain magnitude reduces to $0.707 \times 104 = 73.5$, and $f_h = 7.6$ MHz.

It is instructive to examine some of the small-signal BJT characteristics that appear in the output file (by virtue of the .op statement).

IC	1.14E-03
VBE	6.58E-01
VBC	-5.09E+00
GM	4.39E-02
RPI	2.28E+03
CBE	1.44E-11
CBC	1.52E-12

With $v_{be}|_Q = 0.658$ V (as opposed to 0.7 V), the quiescent collector current and transconductance are somewhat higher than hand-calculated values, and the midfrequency voltage gain increases. And with less negative $v_{bc}|_Q$, the C_μ (CBC) value also increases, but only slightly. Due to these changes, the simulated high-frequency cutoff is lower than anticipated.

Exercise 8.2 Estimate f_h for the amplifier of Fig. 8.22.

Ans: $f_h \approx 40$ MHz

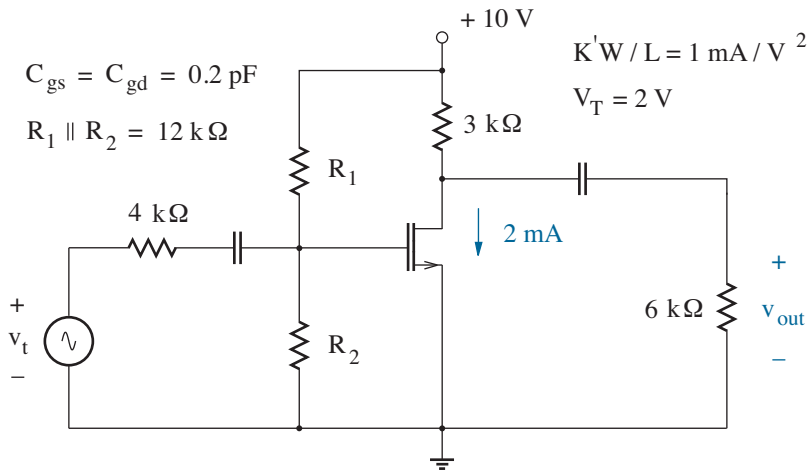


Figure 8.22: Circuit for Exercise 8.2.

Exercise 8.3 Estimate f_h for the amplifier of Fig. 8.23.

Ans: $f_h \approx 130$ MHz

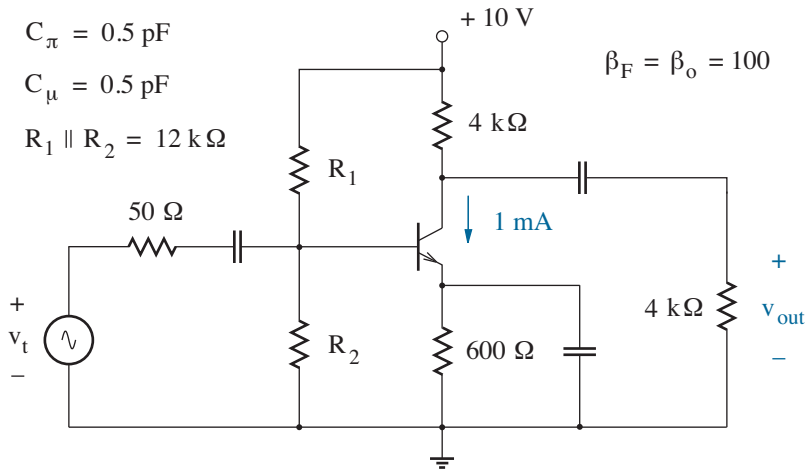


Figure 8.23: Circuit for Exercise 8.3.

Common-Drain (Common-Collector) Amplifier

Figure 8.24 shows the high-frequency common-drain amplifier. Here, C_{gd} hangs like a tree ornament to ac ground while C_{gs} couples output to input.

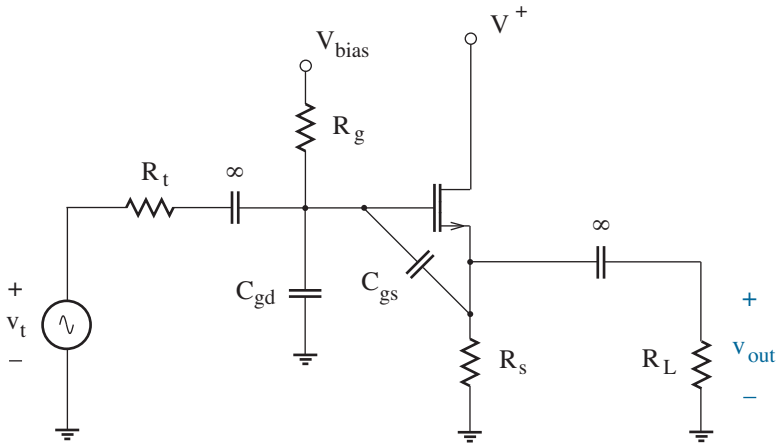


Figure 8.24: Common-drain amplifier at high-frequency.

To determine f_h, one applies the Miller transformation. At the input,

$$C_1 = C_{gd} + C_{gs}\,[\,1 - {A_{vm}}^{F}\,] = C_{gd} + \frac{C_{gs}}{1 + g_m {R_s}'}\,, \tag{8.53}$$

where ${R_s}' = R_s \parallel R_L$. In turn,

$$\tau_1 = ({R_g}' \parallel {r_g}')\,C_1 = (R_t \parallel R_g \parallel \infty)\,C_1\,. \tag{8.54}$$

The reverse midfrequency voltage gain is given by (see Problem 8.21)

$${A_{vm}}^{R} = \frac{{R_g}'}{r_\pi + {R_g}'}\,. \tag{8.55}$$

So with $r_\pi \to \infty$, ${A_{vm}}^{R} = 0$. Thus,

$$C_2 = C_{gs}\,[\,1 - {A_{vm}}^{R}\,] = C_{gs}\,, \tag{8.56}$$

and

$$\tau_2 = ({R_s}' \parallel {r_s}')\,C_2 = (R_s \parallel R_L \parallel 1/g_m)\,C_2\,. \tag{8.57}$$

Finally,

$$f_h \approx \frac{1}{2\pi(\tau_1 + \tau_2)}\,. \tag{8.58}$$

We would have found the same time-constant sum, but with greater effort, had we derived the equivalent resistance "seen" by C_{gs} (see Problem 8.22).

Exercise 8.4 Estimate f_h for the amplifier of Fig. 8.25.

Ans: $f_h \approx 140$ MHz

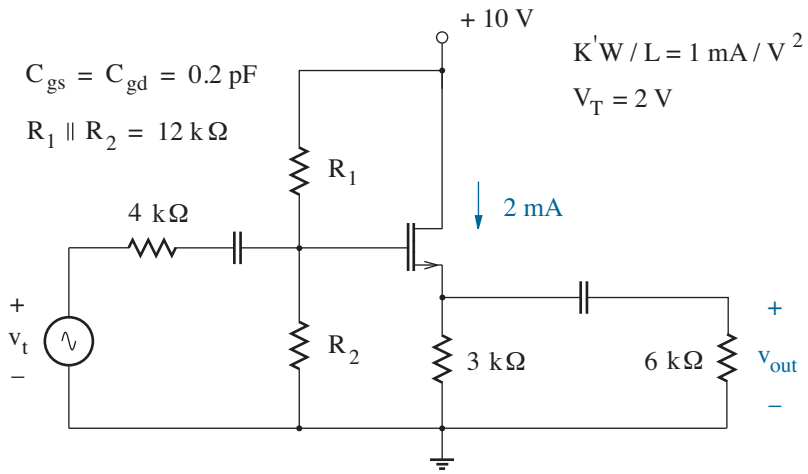


Figure 8.25: Circuit for Exercise 8.4.

Exercise 8.5 Estimate f_h for the amplifier of Fig. 8.26.

Ans: $f_h \approx 4.3$ GHz

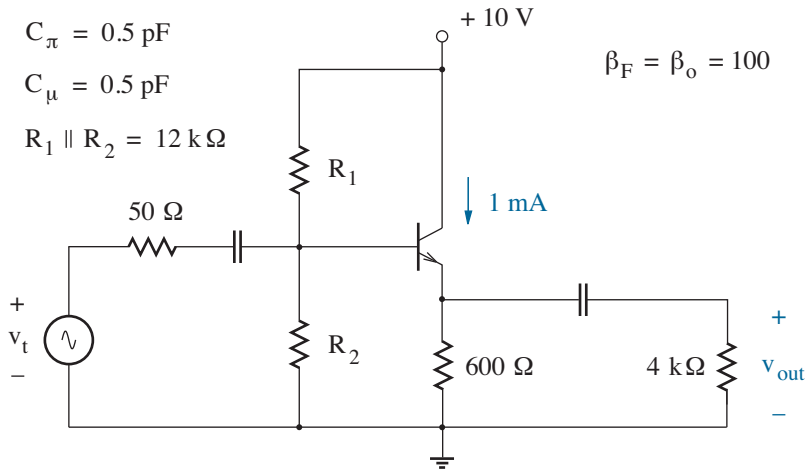


Figure 8.26: Circuit for Exercise 8.5.

Example 8.5

Estimate the high-frequency cutoff for the amplifier of Fig. 8.27.

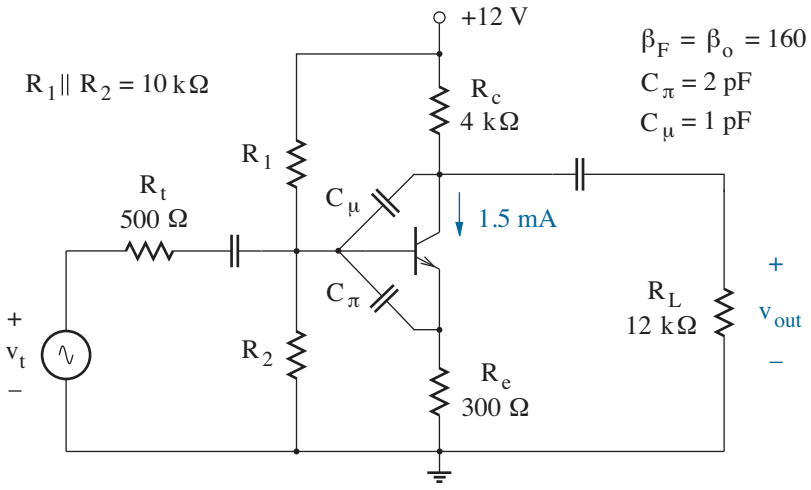


Figure 8.27: Circuit for Example 8.5.

Solution

The circuit functions as a common-emitter amplifier with no emitter bypass. As a base-to-collector capacitance, C_μ couples the input to the output and is thus subject to a Miller transformation. Meanwhile, the circuit effectively functions as a common-collector amplifier with an unused emitter output. As a base-to-emitter capacitance, C_π renders similar input/ouput coupling, and it is also subject to Miller transformation.

The small-signal parameters are $g_m = i_c|_Q/(kT/q) = 57.9\times10^{-3}$ ℧ and $r_\pi = \beta_o/g_m = 2.76$ kΩ. The upper-case (looking away) primed resistance values are ${R_c}' = R_c \parallel R_L = 3$ kΩ, ${R_b}' = R_1 \parallel R_2 \parallel R_t = 476$ Ω, and ${R_e}' = R_e = 300$ Ω. In turn, the lower-case (looking in) primed resistance values are ${r_c}' = \infty$ (pending Chapter 9), ${r_b}' = r_\pi + (1+\beta_o){R_e}' = 51.1$ kΩ, and ${r_e}' = (r_\pi + {R_b}')/(1+\beta_o) = 20.1$ Ω.

For the common-emitter forward voltage gain, we find

$$ {A_{vm}}^F = -\frac{g_m {R_c}'}{1 + g_m {R_e}'(1+1/\beta_o)} = -9.40\,. $$

And for the reverse common-emitter voltage gain,

$$ {A_{vm}}^R = 0\,. $$

Thus, the Miller-transformed C_μ component between the base and ground is $C_\mu(1-{A_{vm}}^F) = 10.4$ pF. The Miller-transformed C_μ component between the collector and ground is $C_\mu(1-{A_{vm}}^R) = 1$ pF.

For the common-collector forward voltage gain, we find

$$A_{vm}{}^{F} = \frac{g_m R_e{}'(1 + 1/\beta_o)}{1 + g_m R_e{}'(1 + 1/\beta_o)} = 0.946 \, .$$

And for the reverse common-collector voltage gain,

$$A_{vm}{}^{R} = \frac{R_b{}'}{r_\pi + R_b{}'} = 0.147 \, .$$

Thus, the Miller-transformed C_π component between the base and ground is $C_\pi(1 - A_{vm}{}^{F}) = 0.108$ pF. The Miller-transformed C_π component between the emitter and ground is $C_\pi(1 - A_{vm}{}^{R}) = 1.71$ pF.

Figure 8.28 shows the amplifier circuit with the Miller-transformed high-frequency capacitances. The total base capacitance is 10.5 pF.

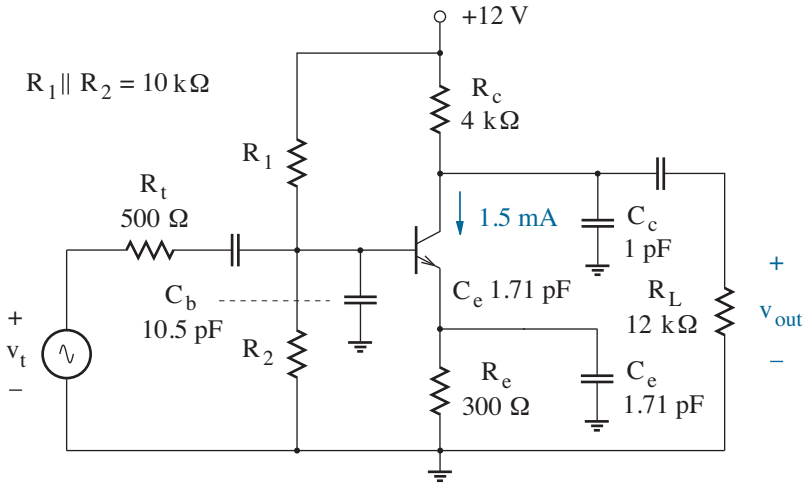


Figure 8.28: Amplifier circuit following C_μ and C_π Miller transformations.

The open-circuit time constants are:

- τ_c (collector) = 1 pF × $(R_c{}' \parallel r_c{}')$ = 3.00 ns
- τ_b (base) = 10.5 pF × $(R_b{}' \parallel r_b{}')$ = 4.96 ns
- τ_e (emitter) = 1.71 pF × $(R_e{}' \parallel r_e{}')$ = 0.032 ns

Finally, the high-frequency cutoff is

$$f_h \approx \frac{1}{2\pi(\tau_c + \tau_b + \tau_e)} = 20 \text{ MHz} \, .$$

A SPICE simulation (using $R_1 = 96$ kΩ and $R_2 = 11.2$ kΩ for the specified quiescent collector current) yields good agreement with $f_h = 20.1$ MHz.

8.3 Bandwidth Considerations

One tends not to design for a pre-specified high-frequency cutoff. Instead, one designs for a particular amplifier gain that is achieved despite capacitive device constraints, and any f_h will do as long as it exceeds some minimum. When this becomes difficult, there are some clever circuit options.

Look at the common-source amplifier of Fig. 8.29. With $i_{d1}|_Q = 1$ mA, $g_{m1} = 8 \times 10^{-3}$ ℧. The forward and reverse midfrequency voltage gains are ${A_{vm1}}^F = -g_{m1}R_{d1} = -20$ and ${A_{vm1}}^R = 0$, respectively. Capacitance C_{gd1} couples the input and output, so we apply a Miller transformation to find

$$C_1 = C_{gs1} + C_{gd1}\,(1 - {A_{vm1}}^F) = 11 \text{ pF} \tag{8.59}$$

as the capacitance at the M_1 gate and

$$C_2 = C_{gd1}\,(1 - {A_{vm1}}^R) = 0.5 \text{ pF} \tag{8.60}$$

as the capacitance at the M_1 drain. The open-circuit time constants are:

- $\tau_1 = C_1 \times (R_{g1}' \parallel r_{g1}') = C_1 \times 200\ \Omega = 2.2$ ns
- $\tau_2 = C_2 \times (R_{d1}' \parallel r_{d1}') = C_2 \times 2.5\ \text{k}\Omega = 1.25$ ns

Then with $\tau_1 + \tau_2 = 3.45$ ns, $f_h \approx 46$ MHz —all very straightforward by now. The Miller-transformed C_{gd1} is the obvious spoiler. If there were a way to reduce the extent of this transformation without sacrificing voltage gain, τ_1 would decrease and f_h would increase.

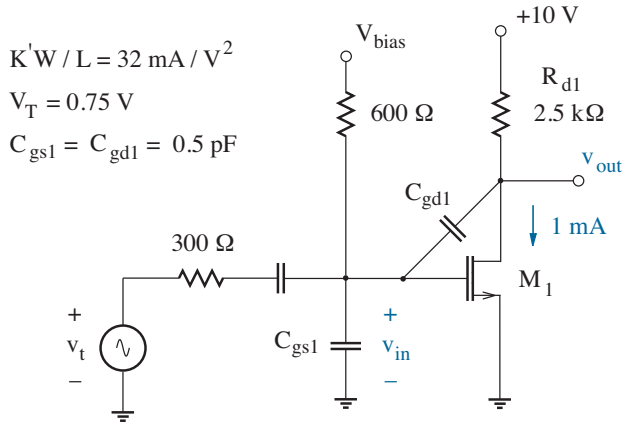


Figure 8.29: Common-source amplifier with $A_{vm} = -20$.

Cascode Amplifier

Recall that the midfrequency voltage-gain expressions for a common-source and a common-gate amplifier are given by $-g_m R_d{}'$ and $g_m R_d{}'$, respectively. Thus, the latter circuit configuration can achieve the same gain without a Miller-transformed capacitance. The downside is a small input resistance. However, this effect is mitigated if the input to the common-gate amplifier derives from the output of a common-source amplifier as shown in Fig. 8.30. The midfrequency voltage gain for the common-source stage is

$$A_{vm1} = -g_{m1} R_{d1}{}' = -g_{m1} \left(\frac{1}{g_{m2}} \right) , \tag{8.61}$$

and $A_{vm1} = -1$ if $g_{m1} = g_{m2}$. Of course any reliance on a common-source amplifier brings us back to the evils of the Miller transformation for C_{gd1}, but the damage is far less significant than before.

The two-transistor circuit of Fig. 8.30 is called a **cascode amplifier**. It is one of only two redeeming venues for the common-gate configuration. (See Chapter 9 for the other.)

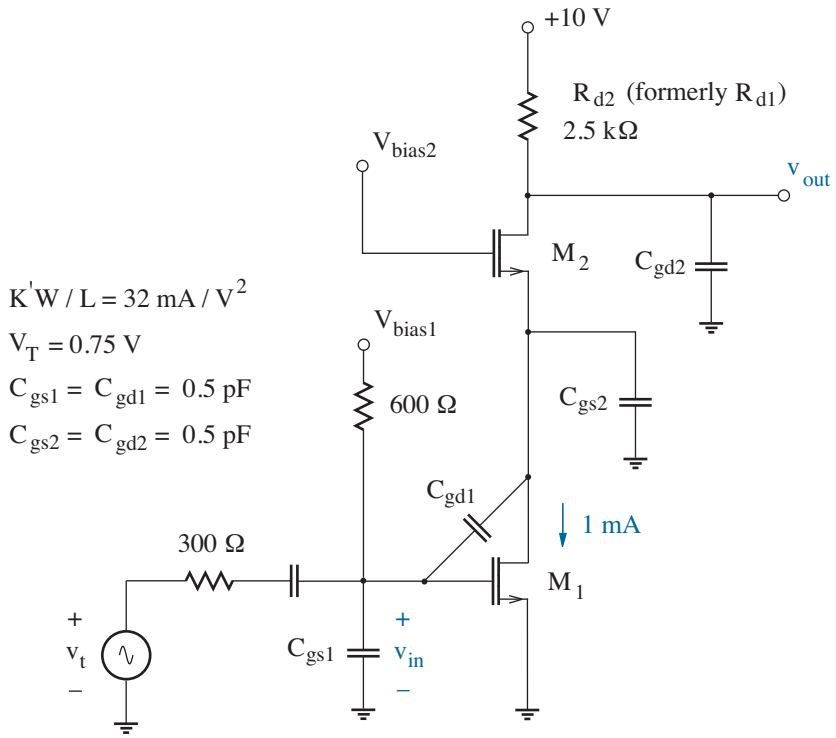


Figure 8.30: Two-stage Cascode amplifier.

Treating the single-stage common-source amplifier, we define C_1 as the capacitance at the M_1 gate. Then with $g_{m1} = g_{m2}$ and ${A_{vm1}}^F = -1$,

$$C_1 = C_{gs1} + C_{gd1}\,(1 - {A_{vm1}}^F) = 1.5 \text{ pF}\,. \tag{8.62}$$

We also define C_2 as the capacitance at the M_2 drain (to keep it near the 2.5-kΩ load resistance, now designated as R_{d2}). Thus,

$$C_2 = C_{gd2} = 0.5 \text{ pF}\,. \tag{8.63}$$

Finally, we take C_3 as the capacitance at the M_1 drain or the M_2 source. With ${A_{vm1}}^R = 0$,

$$C_3 = C_{gd1}\,(1 - {A_{vm1}}^R) + C_{gs2} = 1.0 \text{ pF}\,. \tag{8.64}$$

The open-circuit time constants are:

- $\tau_1 = C_1 \times ({R_{g1}}' \parallel {r_{g1}}') = C_1 \times 200\ \Omega = 0.3$ ns (slashed !)
- $\tau_2 = C_2 \times ({R_{d2}}' \parallel {r_{d2}}') = C_2 \times 2.5\ \text{k}\Omega = 1.25$ ns (unchanged)
- $\tau_3 = C_3 \times ({r_{d1}}' \parallel {r_{s2}}') = C_3 \times 125\ \Omega = 0.125$ ns (new)

In turn, we have $\tau_1 + \tau_2 + \tau_3 = 1.425$ ns and $f_h \approx 110$ MHz —much better. Further increases in f_h would seem to point to τ_2. Nevertheless, this time constant is difficult to change when the load resistance is large and fixed. Unlike brute-force SPICE simulations, open-circuit-time-constant analysis typically tells us where or where not to go for high-frequency improvements.

Before ending our discussion, some practical design issues are in order. If we put aside the loading factor, the product of the midfrequency voltage gains for the two-stage cascode amplifier is

$$A_{vm1}A_{vm2} = -g_{m1}\left(\frac{1}{g_{m2}}\right) g_{m2}{R_{d2}}' = -g_{m1}{R_{d2}}'\,. \tag{8.65}$$

So the result is the same as that for the single-stage amplifier regardless of the g_{m2} value. One is tempted to allow $g_{m2} \gg g_{m1}$ to ensure $|{A_{vm1}}^F| \ll 1$ and a more favorable Miller transformation of C_{gd1}. However, this requires $(W_2/L_2) \gg (W_1/L_1)$ since M_2 and M_1 have equal quiescent bias currents. The size imbalance will likely imply larger C_{gs2} and C_{gd2}.

A second design issue concerns the choice of the MOSFET bias voltages V_{bias1} and V_{bias2}. The former voltage sets $i_{d1}|_Q = 1$ mA for M_1 when $V_{bias1} = v_{gs1}|_Q = 1.0$ V. Whereas M_1 bias current is imposed upon M_2, $v_{gs2}|_Q$ is forced to adjust to 1.0 V when the devices are similarly sized. Thus, the node voltage at the M_2 source (and by connection, the M_1 drain) is $V_{bias2} - 1.0$ V. Meanwhile, M_1 is required to operate in saturation with $v_{ds1} > v_{gs1} - V_T = 0.25$ V. A conservative minimum for $v_{ds1}|_Q$ is 0.75 V, which is consistent with $V_{bias2} = 1.75$ V.

Shunt Peaking

The preceding discussion of the cascode amplifier left us with a stubborn open-circuit time constant (τ_2) at the M_2 drain. Another way to understand its influence is to examine the effective load impedance presented to M_2. As shown in Fig. 8.31, this is

$$Z = R_{d2} \parallel \frac{1}{\jmath\omega C_2} = \frac{R_{d2}}{1 + \jmath\omega\tau_2}\,, \tag{8.66}$$

where $\tau_2 = R_{d2}C_2$. The voltage gain is $A_v = -g_{m1}Z$ (no loading factor). Time constant τ_2 provides for a reduction in $|Z|$, while τ_1 and τ_3 provide for a deterioration of g_{m1}.

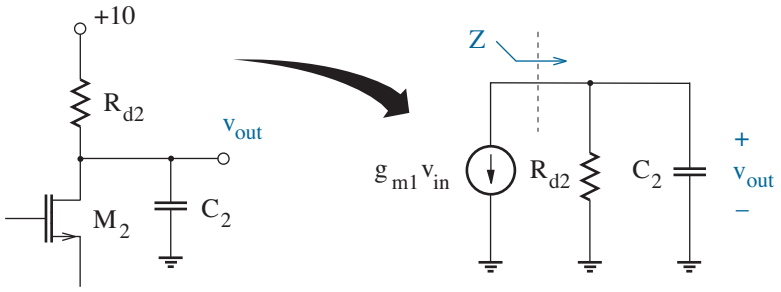


Figure 8.31: M_2 load impedance for the cascode amplifier of Fig. 8.30.

One means of improving the frequency dependence of the load presented to M_2 is to add an inductor L in series with R_{d2} as shown in Fig. 8.32. As frequency increases, the impedance $Z_1 = 1/\jmath\omega C_2$ decreases while the parallel or shunt impedance $Z_2 = R_{d2} + \jmath\omega L$ increases, and the two effects tend to offset each other to maintain $|Z|$ over an extended frequency range.

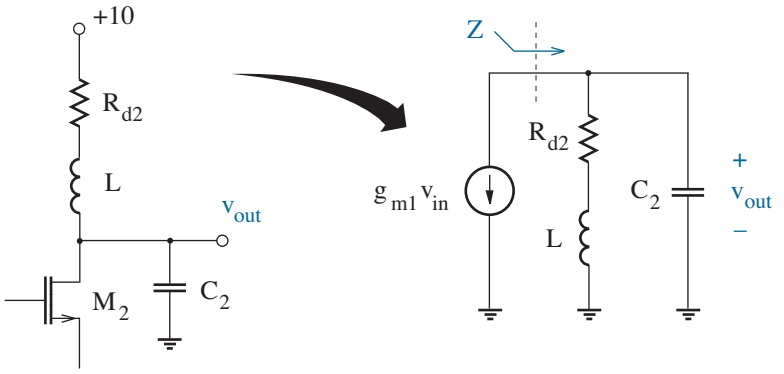


Figure 8.32: Revised "shunt-peaking" load impedance for M_2.

What follows is a remarkable amount of math for a three-element circuit. The impedance of interest is

$$Z = (R_{d2} + \jmath\omega L) \parallel \frac{1}{\jmath\omega C_2}\,, \tag{8.67}$$

which "simplifies" to

$$Z = \frac{R_{d2}}{1 + \jmath\omega\tau_2}\left[1 + \frac{\jmath\omega\tau_2/(\omega_o\tau_2)^2}{1 - (\omega/\omega_o)^2 + \jmath\omega\tau_2}\right], \tag{8.68}$$

where $\omega_o = 1/\sqrt{LC_2}$ is the resonant angular frequency that links L and C_2. At $\omega = \omega_o$, we find

$$Z = \frac{R_{d2}}{1 + \jmath\omega_o\tau_2}\left[1 + \frac{1}{(\omega_o\tau_2)^2}\right] \tag{8.69}$$

so that $|Z| = \sqrt{2}\,R_{d2}$ if $\omega_o = 1/\tau_2$. But increasing $|Z|$ is not our objective. The problem is to choose ω_o (and by extension L) so that $|Z|$ remains flat at R_{d2} over the broadest possible range of frequency. The so-called quality factor associated with the L-C_2 resonance is $Q = 1/(\omega_o\tau_2)$. If $\omega_o \ll 1/\tau_2$, $Q \gg 1$, and the second term within the brackets in Eq. 8.68 contributes sharp undesirable peaking to $|Z|$. At the other extreme, $\omega_o \gg 1/\tau_2$, $Q \ll 1$, and there is little benefit from the added inductance.

An approach that might best be classified as mathematics by experiment is to see what happens if ω_o is chosen so that $|Z| = R_{d2}$ at $\omega = \omega_2 = 1/\tau_2$. To this end, we let $\omega_o = m\,\omega_2$. Then we substitute into Eq. 8.68 to obtain

$$Z(\omega_2) = \frac{R_{d2}}{1 + \jmath}\left[1 + \frac{\jmath/m^2}{1 - 1/m^2 + \jmath}\right]. \tag{8.70}$$

Whereas $|1 + \jmath| = \sqrt{2}$, we require

$$\left[1 + \frac{\jmath/m^2}{1 - 1/m^2 + \jmath}\right] = \sqrt{2}\,, \tag{8.71}$$

which implies (after adding the terms within the brackets)

$$\frac{(1 - 1/m^2)^2 + (1 + 1/m^2)^2}{(1 - 1/m^2)^2 + 1} = 2\,. \tag{8.72}$$

The solution to Eq. 8.72 is $m = \sqrt{2}$. Accordingly, with $\omega_2/2\pi = 127$ MHz, $\omega_o/2\pi = 180$ MHz. As for L, we have

$$L = \frac{{R_{d2}}^2 C_2}{m^2} \tag{8.73}$$

or $L = 1.56\ \mu$H for the case of the R_{d2} and C_2 values considered previously.

Figure 8.33 displays SPICE simulation results for the shunt-peaking design at hand with $m = \sqrt{2}$. As expected, the load impedance is 2.5 kΩ at dc and $\omega_2/2\pi = 127$ MHz. The impedance at intermediate frequencies has a broad peak at about 2.57 kΩ, roughly 3 % higher than the intended value. The frequency at which the impedance declines to 2.5 kΩ/$\sqrt{2}$ is 229 MHz.

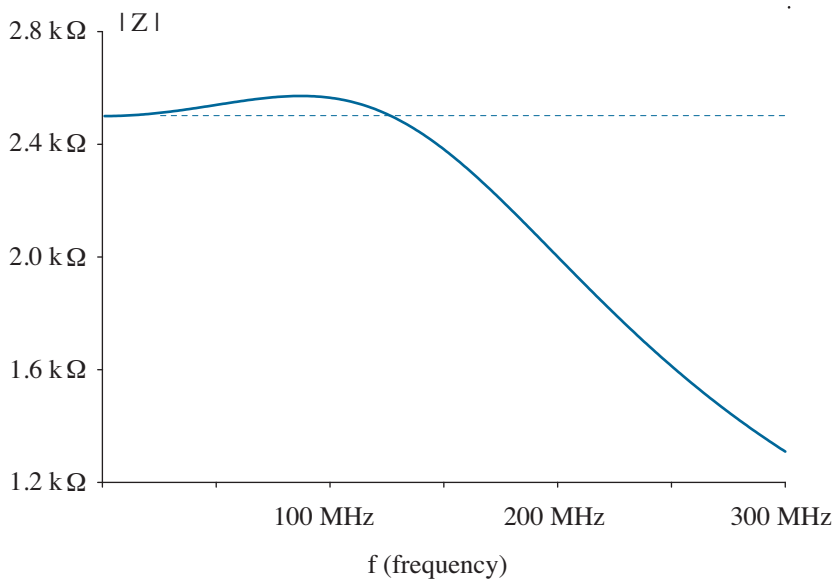


Figure 8.33: "Shunt-peaking" SPICE simulation for $m = \sqrt{2}$.

When the shunt peaking is applied to the cascode amplifier of Fig. 8.30, it is tempting to use the SPICE results of Fig. 8.33 to define an effective $\tau_2' = 0.69$ ns. This leads to $\tau_1 + \tau_2' + \tau_3 = 1.115$ ns and $f_h \approx 117$ MHz. Nevertheless, the concept of open-circuit time constants is meaningless for circuits with capacitors and inductors. So we turn to another simulation, which produces $f_h = 229$ MHz —far better, indeed.

Had we opted for a shunt-peaking design with a maximum degree of "flatness" over the broadband range of frequencies, we would have required $m = 1.55$ and $L = 1.29$ μH. The corresponding f_h is smaller (193 MHz). However, designers tend to take the former route, partly because $m = \sqrt{2}$ is easy to remember, but more likely in anticipation of inductive parasitics, uncertainty in the C_{gd2} capacitance, and other factors. A certain degree of "tweaking" is usually necessary to complete a successful design.

Another broadbanding strategy presents one side of an RLC two-port network to the M_2 drain while taking the other side as the amplifier output. Numerous networks have been developed.

Tuned Amplifier

When the design tradeoff between bandwidth and gain becomes formidable, it is sometimes advantageous to settle for gain in a narrow frequency band. For example, some cell phones receive signals in the vicinity of 1900 MHz, and anything far removed is uninteresting or noise. The **tuned amplifier** of Fig. 8.34 uses a parallel LCR resonant filter to define a frequency band. The load admittance presented to the common-emitter circuit is

$$Y = \frac{1}{\jmath\omega L} + \jmath\omega C + \frac{1}{R}\,. \tag{8.74}$$

Thus, at the resonant frequency

$$f_o = \frac{1}{2\pi\sqrt{LC}}\,, \tag{8.75}$$

the imaginary contributions to the load admittance cancel so that

$$A_v = -g_m R\,. \tag{8.76}$$

For $f \ll f_o$, the inductor looks more like a short circuit, while for $f \gg f_o$, the capacitor assumes a similar character. In either case, the magnitude of the admittance increases and gain decreases. The amplifier bandwidth is

$$\text{BW} = \frac{f_o}{Q}\,, \tag{8.77}$$

where

$$Q = 2\pi f_o RC = \frac{R}{2\pi f_o L} \tag{8.78}$$

is the **quality factor** associated with the parallel resonator.

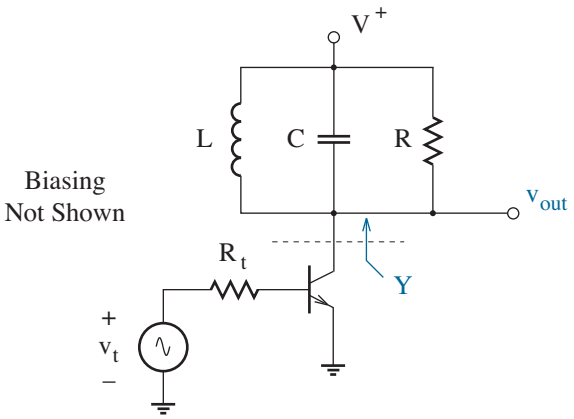


Figure 8.34: Tuned amplifier circuit with parallel LCR resonant load.

The practical tuned amplifier suffers from two complications—

Whereas the BJT exhibits base-to-collector pn-junction capacitance C_μ, the common-emitter circuit has capacitance $C_\mu(1 - A_{vm}{}^R) = C_\mu$ between collector and ground. Thus, the tuned-amplifier center frequency derives, in part, through the parallel combination of C_μ and C. A more rigorous analysis reveals voltage-gain degradation for $R_t \neq 0$ (see Problem 8.42). The two-transistor cascode tuned amplifier avoids this gain problem.

A second complication concerns the parasitic series resistance R_s that inevitably accompanies L. To show how R_s affects the parallel resonator, we consider the series-to-parallel transformation of Fig. 8.35. At $\omega_o = 2\pi f_o$,

$$Z = \jmath\omega_o L_s + R_s = \frac{\jmath\omega_o L_p R_p}{\jmath\omega L_p + R_p} \,, \tag{8.79}$$

where subscripts s and p denote series and parallel, respectively. We equate the real and imaginary parts of Eq. 8.79 to find

$$R_p = R_s\,(1 + Q^2) \tag{8.80}$$

and

$$L_p = L_s\,(1 + \frac{1}{Q^2})\,, \tag{8.81}$$

where

$$Q = \frac{R_p}{\omega_o L_p} = \frac{\omega_o L_s}{R_s}\,. \tag{8.82}$$

Equation 8.80 shows that even a small R_s value can have great significance for the parallel resonator when Q is large. Resistor R_p is in parallel with R in the load of Fig. 8.34, so the revised resonant gain is $A_v = -g_m(R_p \parallel R)$. Manufacturers often cite their inductors as having a particular Q_L when participating in a series resonant circuit with no resistance other than R_s. The specified Q_L value determines R_s or R_p when used in Eq. 8.82.

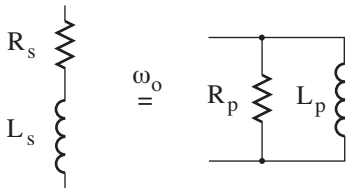


Figure 8.35: Series to parallel RL transformation at angular frequency ω_o

Finally, we observe that inductor L acts as a short circuit at dc. Thus, the tuned amplifier of Fig. 8.34 can be biased with large $i_c|_Q$ for large g_m without the worry of BJT saturation (as for a resistive load).

Example 8.6

Design the tuned amplifier of Fig. 8.36 so that the peak voltage gain is -40, the center frequency is 500 MHz, and the available bandwidth is 50 MHz. Assume that inductors are available with $Q_L = 15$, and let $C = 10$ pF.

The BJTs feature IS=0.3f, BF=125, TF=20p, CJE=120f, and CJC=80f.

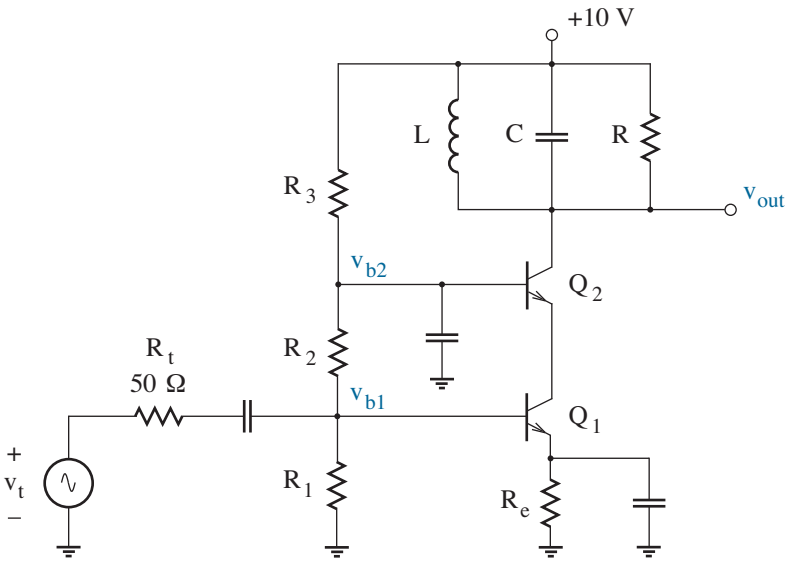


Figure 8.36: Circuit for Example 8.6.

Solution

With $f_o = 500$ MHz and $C = 10$ pF, we use Eq. 8.75 to obtain

$$L = \frac{1}{(2\pi \times 500 \times 10^6 \text{ s}^{-1})^2 \times 10^{-11} \text{ F}} = 10.1 \text{ nH}\,.$$

However, we choose $L = 10$ nH as a standard value. The inductor Q_L is 15. Thus, Eq. 8.82 establishes $R_s = 2.09\ \Omega$ as the parasitic series resistance. We apply Eq. 8.80 (or Eq. 8.82) with the same Q_L to find $R_p = 471\ \Omega$ as an effective parallel resistance in the resonator load. If resistor R were absent, the resonator would have $Q = Q_L = 15$. No larger Q can be achieved.

The combination of $R_p \parallel R$ determines the actual Q for the resonator. From the bandwidth specification, we require

$$Q = \frac{500 \text{ MHz}}{50 \text{ MHz}} = 10\,.$$

Then with the support of Eq. 8.78, $R_p \parallel R = 314\ \Omega$. In turn, $R = 942\ \Omega$. The closest standard value is $R = 910\ \Omega$.

To obtain the desired voltage gain, we use Eq. 8.76 with $R_p \parallel R = 310\ \Omega$. Thus, we seek

$$g_m = \frac{40}{310\ \Omega} = 0.129\ \mho\ .$$

The consistent bias current is

$$i_c|_Q = 0.129\ \mho \times 25.9\ \text{mV} = 3.34\ \text{mA}\ .$$

What remains is a biasing problem. We let $R_e = 1\ \text{k}\Omega$ and $R_1 \parallel (R_2 + R_3) = (\beta_F + 1)R_e \approx 12.5\ \text{k}\Omega$ to promote bias stability. The required node voltage at the base of Q_1 is about 4.0 V, and a reasonable design value for the node voltage at the base of Q_2 is 6.0 V. We leave it as an exercise to show that $R_1 = 24\ \text{k}\Omega$, $R_2 = 9.1\ \text{k}\Omega$, and $R_3 = 18\ \text{k}\Omega$ establish proper conditions.

For SPICE simulations, we take care to include $R_s = 2.09\ \Omega$ in series with L, we assign huge 1-F values to the coupling and bypass capacitors, and we perform a *linear* frequency sweep to force detail at and around f_o. A first simulation shows a peak voltage gain of -34.5, a result that reflects the influence of relatively large $C_\pi = 2.85$ pF (as obtained from the SPICE output file). A painless fix sets R_3 to 12 kΩ so that $i_c|_Q$ and g_m increase. Figure 8.37 shows the final results. The peak gain is -39.5 at 502.5 MHz. The bandwidth between the "0.707" frequencies is 50.8 MHz. Not bad.

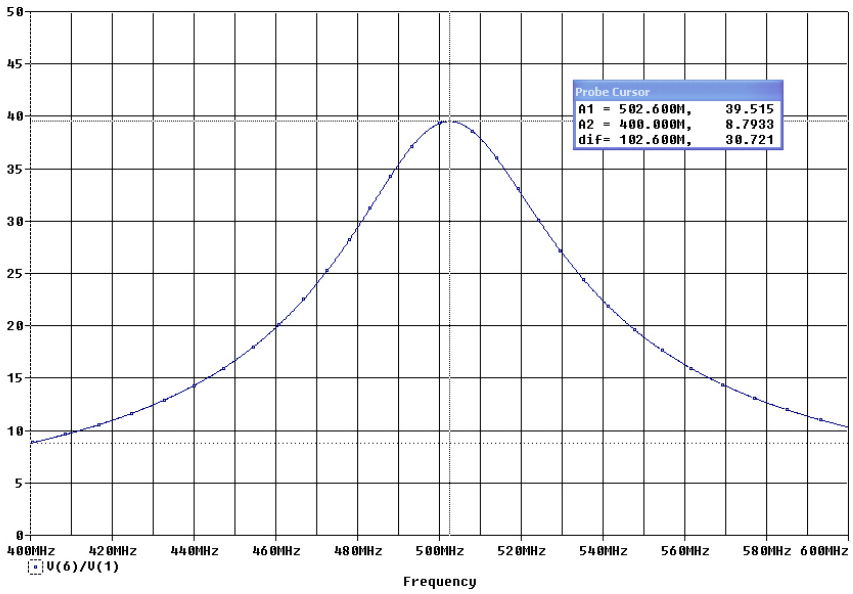


Figure 8.37: SPICE results for Example 8.6.

Unpleasant Parasitics

When choosing transistors for broadband amplifier circuits, one typically endeavors for devices with minimal internal capacitance at minimal cost. But there is more to the story.

As in preceding discussion, we examine a benchmark circuit, this time with a BJT as shown in Fig. 8.38. With $i_c|_Q = 2.59$ mA, $g_m = 0.1$ ℧, and $r_\pi = \beta_o/g_m = 2$ kΩ. The forward and reverse midfrequency voltage gains are ${A_{vm}}^F = -g_m R_c = -200$ and ${A_{vm}}^R = 0$, respectively. Capacitance C_μ couples the input and output, so we apply a Miller transformation to find

$$C_1 = C_\pi + C_\mu\,(1 - {A_{vm}}^F) = 102 \text{ pF} \tag{8.83}$$

as the capacitance at the Q_1 base and

$$C_2 = C_\mu\,(1 - {A_{vm}}^R) = 0.5 \text{ pF} \tag{8.84}$$

as the capacitance at the Q_1 collector. Then subject to ${R_b}' = 100\ \Omega, {r_b}' = r_\pi = 2$ kΩ, ${R_c}' = 2$ kΩ, and ${r_c}' \to \infty$, the open-circuit time constants are:

- $\tau_1 = C_1 \times ({R_b}' \parallel {r_b}') = C_1 \times 95.2\ \Omega = 9.71$ ns
- $\tau_2 = C_2 \times ({R_c}' \parallel {r_c}') = C_2 \times 2$ kΩ $= 1.00$ ns

Thus $\tau_1 + \tau_2 = 10.71$ ns and $f_h \approx 15$ MHz —straightforward as ever.

With the hope of improving the high-frequency cutoff, we modify the signal source so that the Thevenin series resistance $R_t \to 0$ and $\tau_1 \to 0$. This leaves a single 1-ns open-circuit time constant (τ_2), and $f_h \approx 160$ MHz. Or does it?

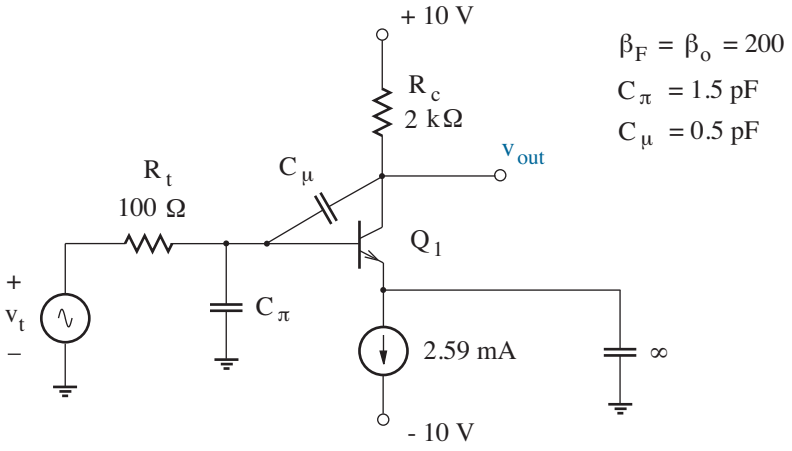


Figure 8.38: Common-emitter amplifier with $A_{vm} = -200$.

Consider Fig. 8.39, which shows a typical npn BJT in cross section. The emitter-to-collector electron current is mostly vertical across the base, and separate capacitive contributions to C_π and C_μ are mostly confined to the indicated horizontal planes. The layered BJT structure requires that a base contact lie to the side of the emitter. In turn, there is a significant parasitic base resistance r_b that limits the base current needed to placate C_π and C_μ. Narrow base width (for large f_T cufoff) makes matters worse.

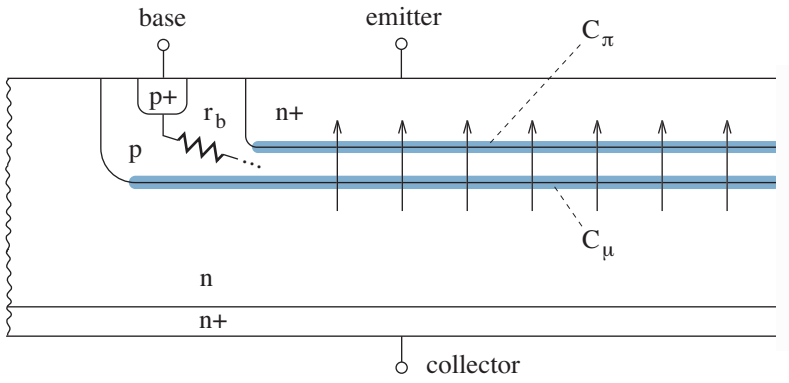


Figure 8.39: BJT cross section showing C_π and C_μ device capacitance in relation to parasitic base resistance r_b. The horizontal scale is compressed.

Strictly speaking, the small-signal resistance looking into the BJT base is r_b plus the former r_b' (Fig. 8.40a). However, it is convenient to adopt an inward and outward perspective at the node common to C_π and C_μ so that the procedures of Section 8.2 apply. This has r_b external to the BJT, where it adds to the former R_b' (Fig. 8.40b).

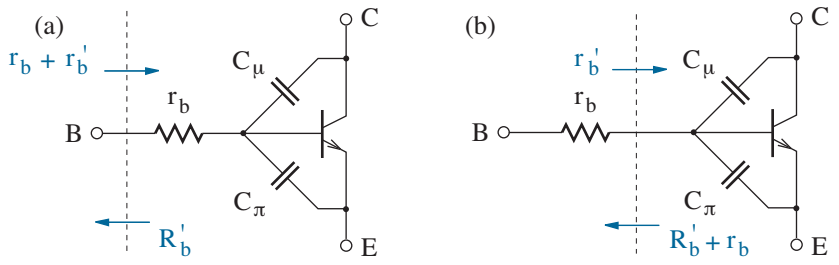


Figure 8.40: Two perspectives in relation to parasitic base resistance.

Back to the benchmark circuit. Time constant $\tau_1 = C_1 \times (R_b' + r_b) \parallel r_b'$. Let $r_b = 10\ \Omega$. Then with $R_t \to 0$, $\tau_1 + \tau_2 = 2.02$ ns, and $f_h \approx 79$ MHz. Choosing to ignore r_b invites unrealistic expectations.

In our haste to play out a dramatic and treacherous role for resistor r_b, we neglected some issues that threaten its use in a small-signal BJT model.

- The portion of r_b near (and in series with) the base contact reflects an *inactive* base region where transistor action does not occur.
- The portion of r_b along the gap between the collector and emitter reflects an *active* base region with minority-carrier cross-currrent.
- The boundary between the preceding regions is poorly defined.
- Resistivity in the active base region tends to reduce the base-emitter voltage as one moves into the collector/emitter gap. Whereas carrier injection is exponential with a local v_{be}, the maximum emitter current density occurs at the emitter periphery. In poor transistor designs, the emitter current density steadily falls to zero (pinch out).
- Nonuniform emitter current density implies nonuniform C_π.
- Resistivity in either base region is *modulated* (or specifically reduced) in the presence of large currents.

The preceding statements suggest that r_b represents a *distributed* circuit with localized contributions that experience varying capacitive interactions throughout the base. In turn, it is impossible to assign a single and constant lumped resistance in series with the base terminal to account for everything. However, we can approximate the effects of base resistivity through a single series resistance by making it *frequency dependent*. The r_b value will have greatest importance at high frequencies in the vicinity of cutoff.

Since r_b is typically much smaller than r_π, it has only modest influence at low frequencies. Thus, BJT simulations often get by with a constant r_b value that is consistent with experimental observations at high frequencies and harmless otherwise.

To measure r_b we look to a set of y-parameters that relate the currents and voltages of a linear two-port (see Fig. 7.20):

$$i_1 = y_i\, v_1 + y_r\, v_2 \,, \tag{8.85a}$$

$$i_2 = y_f\, v_1 + y_o\, v_2 \,. \tag{8.85b}$$

The individual parameters are evaluated subject to $v_1 = 0$ at the input or $v_2 = 0$ at the output. For example,

$$y_i = \left.\frac{i_1}{v_1}\right|_{v_2=0} \tag{8.86}$$

is the **input admittance** when the output is a short circuit. When applied to the small-signal hybrid-π BJT model, the y-parameter representation uses an additional subscript to indicate the device connection that is shared between the input and output. So under the common-emitter configuration, our focus is y_{ie}, y_{re}, y_{fe}, and y_{oe}.

Consider y-parameter data for a high-performance "RF" npn BJT.

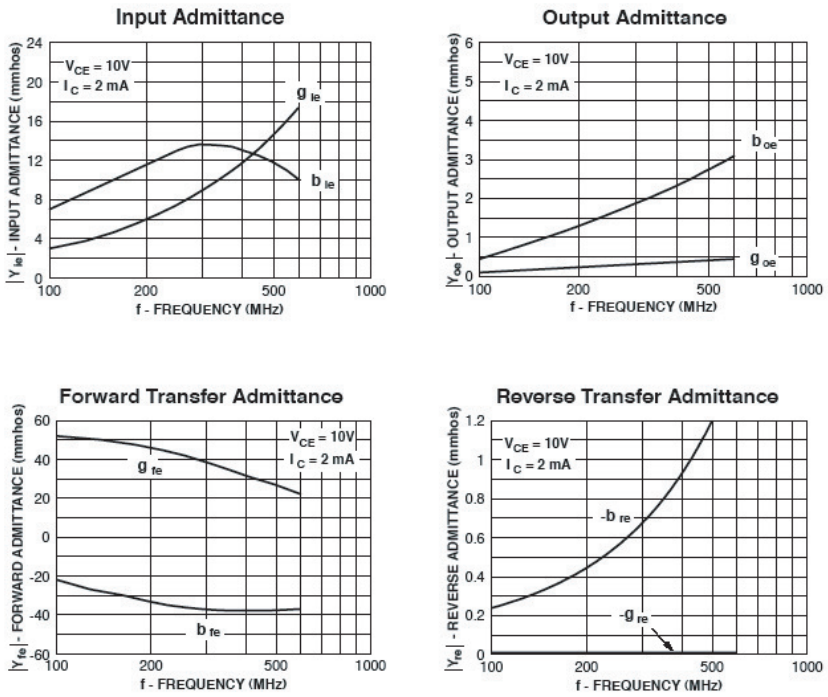


Figure 8.41: Manufacturer's y-parameter data for the MMBTH10 BJT. Courtesy Fairchild Semiconductor.

For the case of the small-signal model of Fig. 8.3 with resistance r_b in series with the base terminal and a short circuit between collector and emitter, we have

$$y_{ie} = \frac{g_b\,[\,g_\pi + \jmath\omega(C_\pi + C_\mu)\,]}{g_b + g_\pi + \jmath\omega(C_\pi + C_\mu)}\,, \tag{8.87}$$

where $g_b = 1/r_b$ and $g_\pi = 1/r_\pi$. The imaginary y_{ie} part (susceptance b_{ie}) exhibits a maximum at $\omega = (g_b + g_\pi)/(C_\pi + C_\mu)$ with corresponding value

$$b_{ie}\,|_{\max} = \frac{g_b}{2}\left(\frac{g_b}{g_b + g_\pi}\right) \tag{8.88}$$

(see Problem 8.50). Then with $g_b \gg g_\pi$,

$$b_{ie}\,|_{\max} \approx \frac{g_b}{2}\,. \tag{8.89}$$

The data of Fig. 8.41 has $g_b/2 \approx 14 \times 10^{-3}$ $\mho$, so $r_b \approx 35\ \Omega$.

We are getting dangerously beyond practical hand calculations for f_h. Nevertheless, SPICE can partially account for the effects of distributed base resistance by partitioning the C_μ capacitance. With reference to Fig. 8.39, the poorly defined boundary between the inactive and active base regions is near the left edge of the emitter. If we assign a node there in support of an extended lumped small-signal circuit model, the base-collector capacitance splits as shown in Fig. 8.42. In turn, we have

$$C_\mu \approx X_{CJC}\, C_{jc}\, [1 - v_{bc}/\phi_c]^{-m_c} \tag{8.90}$$

and

$$C_{jx} \approx (1 - X_{CJC})\, C_{jc}\, [1 - v_{bc}/\phi_c]^{-m_c}\,. \tag{8.91}$$

SPICE parameter XCJC has a value between zero and unity, and it is can be found in consideration of the relative areas of the emitter and base regions. The C_{jx} portion of the base-collector capacitance tends to be more benign than troublesome C_μ, so XCJC can be significant at very high frequencies.

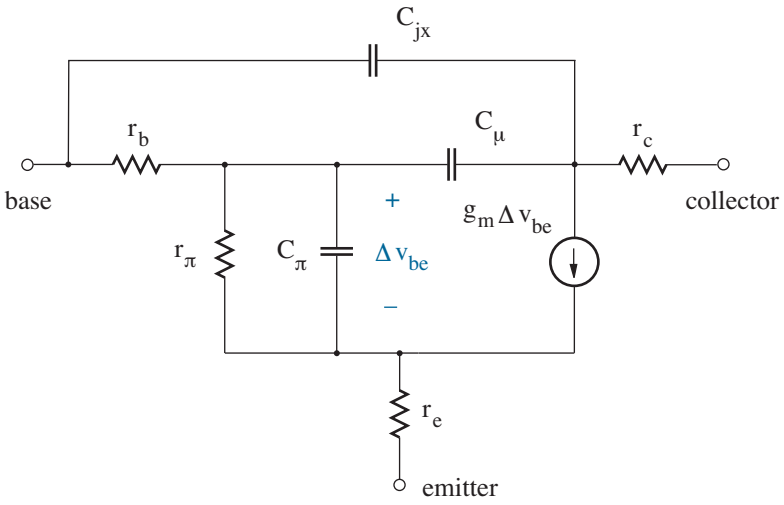


Figure 8.42: High-frequency small-signal BJT model with the base-collector capacitance partitioned between internal and external base nodes.

The previously noted decline of base resistance with increases in BJT terminal currents is difficult to model. SPICE assumes a maximum r_b value (RB) at low base current, a minimum value (RBM) at high base current, and a complex transition between the two extremes that is characterized with parameter IRB, the base current at the halfway point. It is advisable to avoid this complexity by measuring r_b for anticipated quiescent conditions so that parameter RB is sufficient.

A modern technology that alleviates some of the problems associated with base resistance utilizes a **heterojunction bipolar transistor (HBT)** in place of a conventional BJT. While available in several material systems, the HBT counterpart to the silicon BJT at hand features a $Si_{1-x}Ge_x$ base. To understand this device, we need to recall the discussion of energy-band models in Chapter 2 as we examine the base-emitter junction of Fig. 8.43. With $x = 0.13$, the reduced bandgap for SiGe is primarily manifest as a valence-band discontinuity of about 0.1 eV. Thus, a base that is p-type and heavily doped so as to reduce r_b has hole confinement—the valence-band discontinuity prevents back injection into the emitter that would otherwise promote a poor β_F. Consider this a triumph of bandgap engineering.

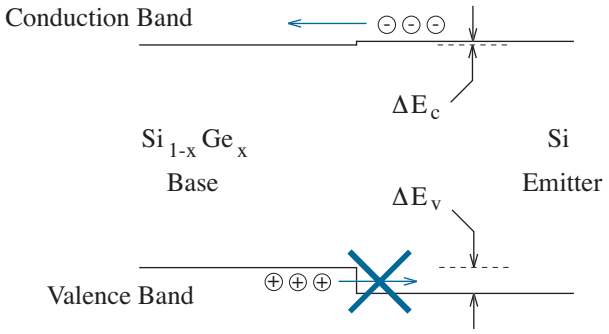


Figure 8.43: Base-emitter junction for an npn SiGe HBT.

SiGe HBTs have become important for very-high-speed communication circuits operating at tens of GHz. Figure 8.44 shows a typical structure. Apart from a SiGe base, devices in this technology have polysilicon emitters that allow for small feature sizes and reduced capacitance.

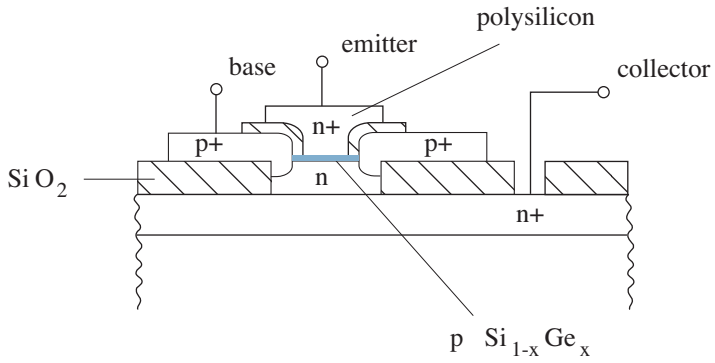


Figure 8.44: Typical npn SiGe HBT structure.

8.4 Low-Frequency Response

In the preceding chapter, we treated all capacitors as ac short circuits that allowed input coupling, output coupling, or terminal bypass connections in a single-stage amplifier while preserving appropriate dc transistor biasing. This section explores the "gray" region of frequency wherein the ac-to-dc transition occurs, and it develops design procedures for capacitor selection.

Capacitors tend to block low-frequency ac signals, so we expect reduced voltage gain, perhaps with the frequency dependence shown in Fig. 8.45. We arbitrarily define the **low-frequency cutoff** (f_l) as the frequency at which the midfrequency voltage (power) gain is reduced by $1/\sqrt{2}$ (1/2).

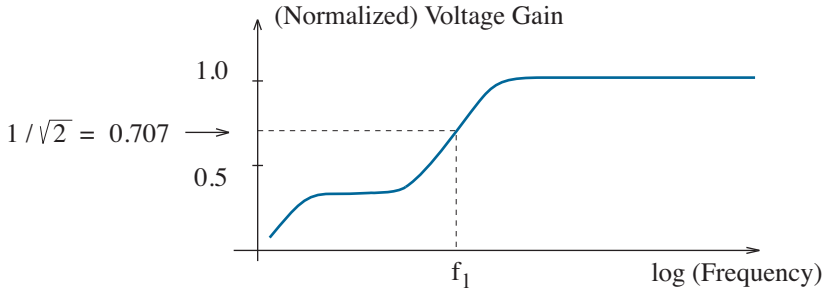


Figure 8.45: Low-frequency voltage gain for a particular amplifier.

Before determining a process that estimates f_l for an arbitrary amplifier, it will be helpful to examine the exact low-frequency cutoff that results when one capacitor is the first to experience an ac-to-dc transition and all other capacitors persist as ac short circuits. Three cases concern us—

Case I: Input Coupling Capacitor

We examine a "dominant" input coupling capacitor in the general amplifier circuit of Fig. 8.46.

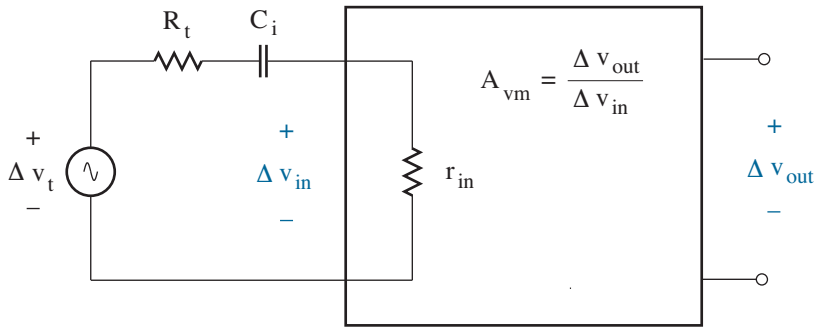


Figure 8.46: Amplifier circuit with input coupling capacitor.

If C_i assumes an impedance given by $1/\jmath\omega C_i$, the total voltage gain includes a frequency-dependent loading factor. Specifically,

$$A_v^{\text{Total}} = A_{vm} \frac{r_{in}}{r_{in} + R_t + 1/\jmath\omega C_i} , \tag{8.92}$$

In the limit as $C_i \to \infty$, this reduces to

$$A_{vm}^{\text{Total}} = A_{vm} \frac{r_{in}}{r_{in} + R_t} , \tag{8.93}$$

the total midfrequency voltage gain. Thus, we obtain a normalized total voltage gain by dividing Eq. 8.92 by Eq. 8.93. After some algebra,

$$\frac{A_v^{\text{Total}}}{A_{vm}^{\text{Total}}} = \frac{\jmath\omega\tau_i}{\jmath\omega\tau_i + 1} , \tag{8.94}$$

where

$$\tau_i = (r_{in} + R_t)\, C_i . \tag{8.95}$$

Note that τ_i has physical significance as an RC time constant containing the equivalent resistance "seen" by the capacitor when Δv_t is a short circuit.

We are now prepared to construct the Bode plot shown in Fig. 8.47. For large ω, the normalized total voltage gain is unity. As ω decreases, we encounter a breakpoint at $1/\tau_i$, and the normalized total voltage gain asymptotically decreases at a rate of -20 dB per decade.

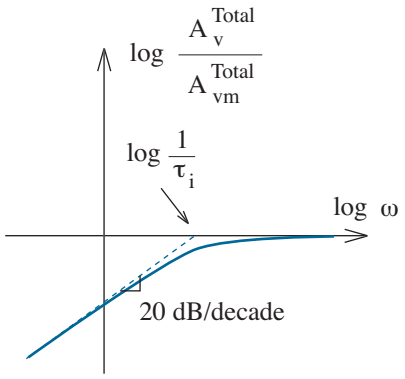


Figure 8.47: Bode (magnitude) plot for the amplifier of Fig. 8.46.

At breakpoint, the normalized total voltage gain is -3 dB ($1/\sqrt{2} = 0.707$), so we have

$$f_l = \frac{\omega_l}{2\pi} = \frac{1}{2\pi(r_{in} + R_t)C_i} . \tag{8.96}$$

Case II: Output Coupling Capacitor

We examine a dominant output coupling capacitor in the general amplifier circuit of Fig. 8.48.

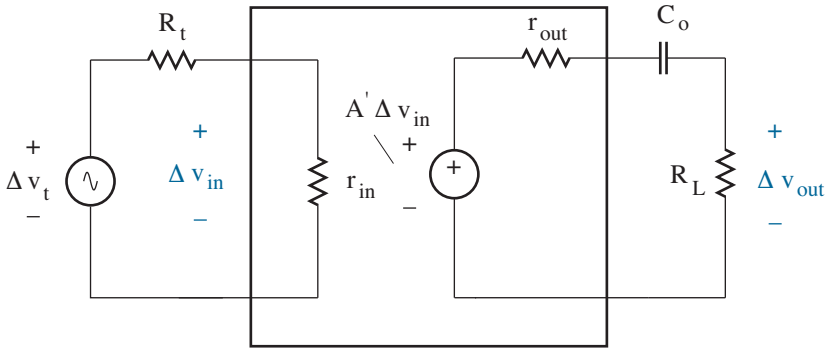


Figure 8.48: Amplifier circuit with output coupling capacitor.

If C_o assumes an impedance given by $1/\jmath\omega C_o$, the total voltage gain reflects a frequency-dependent voltage division at the output. Specifically,

$$A_v^{\text{Total}} = \left(\frac{r_{in}}{r_{in}+R_t}\right) A' \left(\frac{R_L}{R_L + r_{out} + 1/\jmath\omega C_o}\right) . \tag{8.97}$$

This result reduces to the total midfrequency voltage gain in the limit as $C_o \to \infty$. With a little algebra, we obtain a normalized total voltage gain:

$$\frac{A_v^{\text{Total}}}{A_{vm}^{\text{Total}}} = \frac{\jmath\omega\tau_o}{\jmath\omega\tau_o + 1} , \tag{8.98}$$

where

$$\tau_o = (R_L + r_{out})\, C_o . \tag{8.99}$$

Again, we note that τ_o has physical significance as an RC time constant containing the equivalent resistance "seen" by the capacitor when Δv_t is a short circuit (so that the dependent source is also a short circuit).

Whereas Eq. 8.98 has the same form as Eq. 8.94, we expect a Bode plot that is similar to that of Fig. 8.47. The low-frequency cutoff is

$$f_l = \frac{\omega_l}{2\pi} = \frac{1}{2\pi(R_L + r_{out})C_o} . \tag{8.100}$$

Case III: Emitter (Source) Bypass Capacitor

We examine a dominant emitter bypass capacitor in the BJT amplifier of Fig. 8.49. All other capacitors are assumed to behave as short circuits. Similar results are expected for the source-bypassed MOSFET amplifier when we let $r_\pi \to \infty$, $R_e \to R_s$, and $C_e \to C_s$.

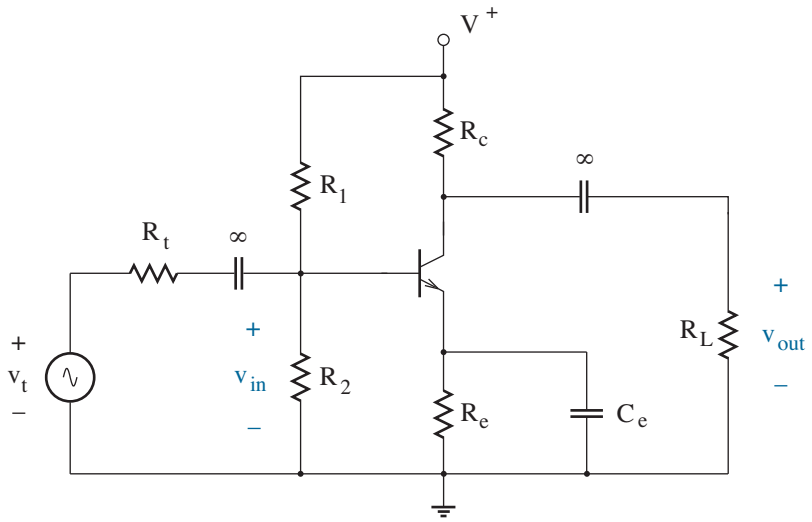


Figure 8.49: Amplifier circuit with a dominant emitter bypass capacitor. The large input and output coupling capacitors persist as ac short circuits.

If C_e assumes an impedance given by $1/\jmath\omega C_e$, the total voltage gain has a frequency-dependent inherent component and loading factor. Specifically,

$$A_v^{\text{Total}} = \frac{-g_m(R_c \parallel R_L)}{1 + g_m(R_e \parallel 1/\jmath\omega C_e)(1 + 1/\beta_o)} \left(\frac{r_{in}}{r_{in} + R_t}\right), \tag{8.101}$$

where

$$r_{in} = R_1 \parallel R_2 \parallel [r_\pi + (\beta_o + 1)(R_e \parallel 1/\jmath\omega C_e)]\,. \tag{8.102}$$

In the limit as $C_e \to \infty$, the preceding expressions reduce to their values at midfrequency:

$$A_{vm}^{\text{Total}} = -g_m(R_c \parallel R_L) \left(\frac{r_{in}}{r_{in} + R_t}\right), \tag{8.103}$$

and

$$r_{in} = R_1 \parallel R_2 \parallel r_\pi\,. \tag{8.104}$$

After some tedious algebra, we find

$$\frac{A_v^{\text{Total}}}{A_{vm}^{\text{Total}}} = \frac{\jmath\omega + 1/\tau_{ez}}{\jmath\omega + 1/\tau_{ep}} , \tag{8.105}$$

where

$$\tau_{ez} = R_e C_e \tag{8.106}$$

and

$$\tau_{ep} = \left[R_e \parallel \frac{r_\pi + (R_1 \parallel R_2 \parallel R_t)}{1 + \beta_o} \right] C_e . \tag{8.107}$$

Yet again, we note that τ_{ep} has physical significance as an RC time constant containing the equivalent resistance "seen" by the capacitor $(R_e \parallel r_e')$ when Δv_t is a short circuit. We also note that $\tau_{ez} \gg \tau_{ep}$, since the latter contains R_e in parallel with a small resistance.

Figure 8.50a shows the applicable pole-zero plot, which features a pole at $\jmath\omega = -1/\tau_{ep}$ and a zero at $\jmath\omega = -1/\tau_{ez}$. The corresponding Bode plot of Fig. 8.50b shows unity normalized total voltage gain for large values of ω. As ω decreases, we encounter a breakpoint at $1/\tau_{ep}$, and the normalized total voltage gain asymptotically decreases at a rate of -20 dB per decade. We subsequently encounter another breakpoint at $1/\tau_{ez}$, which eventually restores the normalized total voltage gain to a constant level. The -3 dB condition is related to the first breakpoint, so

$$f_l = \frac{\omega_l}{2\pi} = \left\{ 2\pi \left[R_e \parallel \frac{r_\pi + (R_1 \parallel R_2 \parallel R_t)}{1 + \beta_o} \right] C_e \right\}^{-1} . \tag{8.108}$$

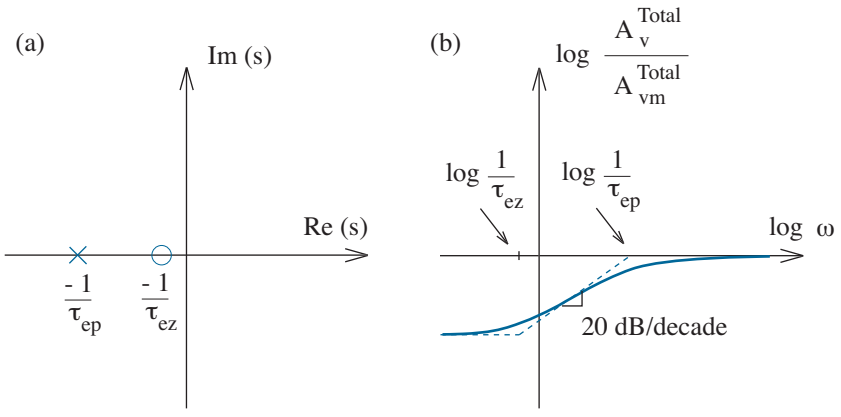


Figure 8.50: Plots used to determine the low-frequency cutoff for the BJT emitter-bypassed amplifier of Fig. 8.49: (a) pole-zero; (b) Bode.

The Method of Short-Circuit Time Constants

By now, it should be apparent that the low-frequency cutoff is determined if a particular capacitor "dominates" through its unchallenged experience of an ac-to-dc transition. But how do we proceed when several capacitors compete for this distinction?

In general, we can estimate the low-frequency cutoff of an amplifier using the following approximation:

$$f_l \approx \frac{1}{2\pi} \sum \frac{1}{\tau_{sc}} , \tag{8.109}$$

where τ_{sc} is a **short-circuit time constant** given by the product of a capacitance and the equivalent resistance "seen" by that capacitor when *all other capacitors are treated as ac short circuits (as at midfrequency).*

Proof

Consider an amplifier with n coupling/bypass capacitors. The system function that describes the total voltage gain must satisfy three conditions: (1) There are n poles—one for each capacitor. (2) The total voltage gain is constant at arbitrarily high frequencies when all capacitors function as ac short circuits. (3) There is at least one zero at zero frequency when, among other things, the input coupling capacitor becomes an open circuit. We conclude that the required system function takes the form

$$A(s) = K \frac{s(s+z_1)(s+z_2)\dots(s+z_{n-1})}{(s+p_1)(s+p_2)\dots(s+p_n)} , \tag{8.110}$$

where K is a constant. If we assume that the zeros occur at relatively small s values in relation to the poles—this applied in the preceding examples—the system function is approximately

$$A(s) \approx K \frac{s^n}{(s+p_1)(s+p_2)\dots(s+p_n)} . \tag{8.111}$$

We divide numerator and denominator by s^{n-1} and expand to obtain

$$A(s) \approx K \frac{s}{s + (p_1 + p_2 + \dots + p_n) + c_1/s + c_2/s^2 + \dots + c_{n-1}/s^{n-1}} , \tag{8.112}$$

where $c_1, c_2 \dots c_{n-1}$ are constants.

Since we are interested in the first signs of non-constant behavior in $A(s)$ as $s = j\omega$ is reduced from a large imaginary value, it is reasonable to ignore the third and all successive terms in the denominator of Eq. 8.112. In turn, we estimate the low-frequency cutoff as

$$f_l \approx \frac{1}{2\pi} (p_1 + p_2 + \dots + p_n) . \tag{8.113}$$

Note that the p_i terms have the dimensions of reciprocal time constants.

Recall (from Section 8.2) that the terminal variables of an n-port linear resistive network with external capacitors are governed by the relation

$$\mathbf{i} = [\mathbf{G} + s\mathbf{C}]\,\mathbf{v}\,. \tag{8.114}$$

Here, $\mathbf{i}$ and $\mathbf{v}$ are network current and voltage vectors, respectively, $\mathbf{G}$ is a conductance matrix (Eq. 8.36), and $\mathbf{C}$ is a capacitance matrix (Eq. 8.39). The diagonal terms in the $\mathbf{G}$ matrix represent the conductance looking into a terminal pair if all of the other terminal pairs exhibit short circuits. (Figure 8.15 shows the n-port circuit.)

The poles of the system are the values of s that allow $\mathbf{v} \neq 0$ when $\mathbf{i} = 0$, and they are found by determining the roots to the equation

$$det\,[\mathbf{G} + s\mathbf{C}] = 0 \tag{8.115}$$

or

$$det \begin{pmatrix} G_{11} + sC_1 & G_{12} & \dots & G_{1n} \\ G_{21} & G_{22} + sC_2 & \dots & G_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ G_{n1} & G_{n2} & \dots & G_{nn} + sC_n \end{pmatrix} = 0\,. \tag{8.116}$$

If we expand the determinant and divide both sides of the resulting equation by the product of the n capacitor values, we find

$$s^n + s^{n-1}\left(\frac{G_{11}}{C_1} + \frac{G_{22}}{C_2} + \ldots + \frac{G_{nn}}{C_n}\right) + \ldots = 0\,. \tag{8.117}$$

But in terms of the actual poles of the system,

$$(s+p_1)(s+p_2)\ldots(s+p_n) = s^n + s^{n-1}(p_1+p_2+\ldots+p_n) + \ldots = 0. \tag{8.118}$$

So a comparison between Eq. 8.117 and Eq. 8.118 reveals the relation

$$\frac{G_{11}}{C_1} + \frac{G_{22}}{C_2} + \ldots + \frac{G_{nn}}{C_n} = p_1 + p_2 + \ldots + p_n\,. \tag{8.119}$$

The left-hand side of this expression is the sum of n short-circuit time constants, since $1/G_{ii}$ is the resistance "seen" by capacitor C_i when all other capacitors are replaced by short circuits (so that only $v_i \neq 0$). Thus, in consideration of Eq. 8.113,

$$f_l \approx \frac{1}{2\pi} \sum \frac{1}{\tau_{sc}}\,. \tag{8.120}$$

Caution: The estimate excludes capacitors that are internal to transistors. At frequencies in the vicinity of f_l, these capacitors are ac open circuits.

Example 8.7

Estimate the low-frequency cutoff for the amplifier of Fig. 8.51.

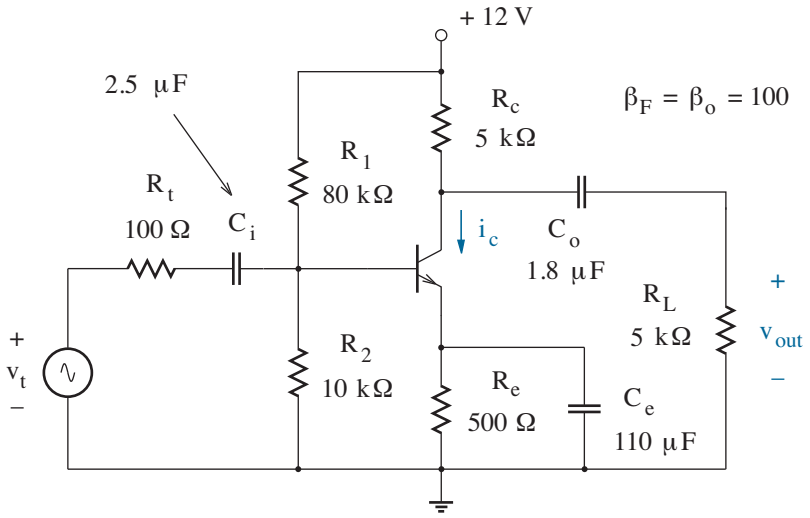


Figure 8.51: Circuit for Example 8.7.

Solution

A quick dc analysis—no need for details—reveals that $i_c|_Q = 1.06$ mA. Thus, $r_\pi = \beta_o(kT/q)/i_c|_Q = 2.44$ kΩ. Then with the help of Table 8.1 (or Table 7.3, if preferred), we calculate the short-circuit time constants associated with each capacitor.

Capacitor C_i: $\tau_{sc}{}^{(1)} = r_{eq}{}^{(1)}C_i$
$r_{eq}{}^{(1)} = r_{in} + R_t = R_1 \parallel R_2 \parallel r_\pi + R_t = 2.02$ kΩ.
$\tau_{sc}{}^{(1)} = 5.05 \times 10^{-3}$ s, $1/2\pi\tau_{sc}{}^{(1)} = 31.5$ Hz.

Capacitor C_o: $\tau_{sc}{}^{(2)} = r_{eq}{}^{(2)}C_o$
$r_{eq}{}^{(2)} = R_L + r_{out} = R_L + R_c = 10$ kΩ.
$\tau_{sc}{}^{(2)} = 18 \times 10^{-3}$ s, $1/2\pi\tau_{sc}{}^{(2)} = 8.8$ Hz.

Capacitor C_e: $\tau_{sc}{}^{(3)} = r_{eq}{}^{(3)}C_e$
$r_{eq}{}^{(3)} = R_e \parallel [(r_\pi + R_1 \parallel R_2 \parallel R_t)/(1 + \beta_o)] = 24$ Ω.
$\tau_{sc}{}^{(3)} = 2.64 \times 10^{-3}$ s, $1/2\pi\tau_{sc}{}^{(3)} = 60.3$ Hz.

Finally, $f_l = (1/2\pi)\sum(1/\tau_{sc}) = 100.6 \approx 100$ Hz.

Of course, a more exact procedure that determines low-frequency cutoff uses a circuit analysis program such as SPICE. Following Example 8.4, we describe the input signal source with a statement of the form

```
Vt      1       0       ac      10m
```

where the 10-mV ac amplitude is *arbitrary* —SPICE does not worry about swing limitations. Second, we include a control statement of the form

```
.ac     DEC     10      1       10k
```

which sweeps the signal-source frequency from 1 Hz to 10kHz with 10 points per decade. Together, these statements support the .probe plot of Fig. 8.52.

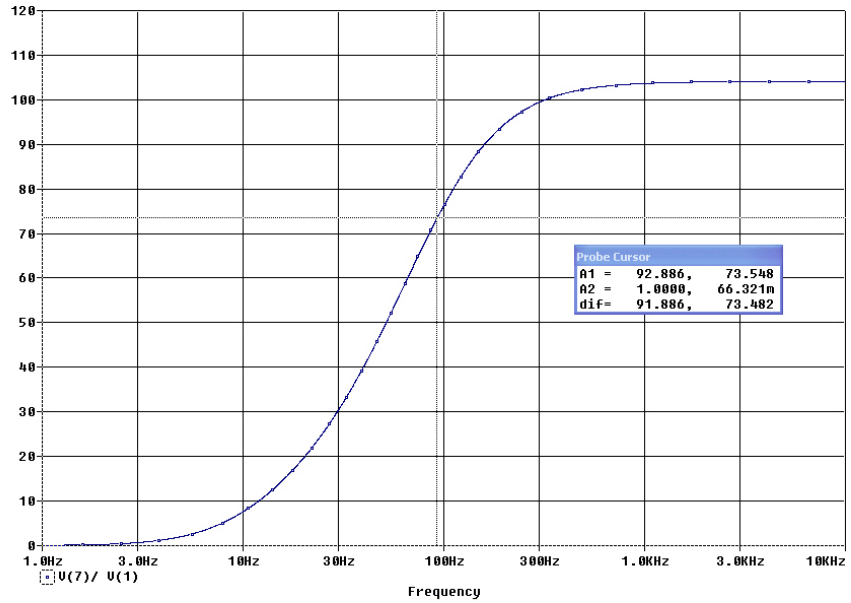


Figure 8.52: SPICE .probe plot for Example 8.7.

To find the low-frequency cutoff, we note that the magnitude of the total midfrequency voltage gain is 104. At cutoff, the magnitude of the voltage gain is reduced to $0.707 \times 104 = 73.5$, so $f_l = 93$ Hz.

You might wonder why we have bothered to discuss a procedure that estimates the low-frequency cutoff with roughly 10-% error when a much more accurate computer simulation is available. (And in view of this error, you might justifiably criticize the degree of precision used in Example 8.7.) Notwithstanding, the method of short-circuit time constants is often useful, since it allows us to determine which of several capacitors provides the greatest influence at low frequencies. In Example 8.7, the emitter bypass capacitor clearly dominates.

Exercise 8.6 Estimate f_l for the amplifier of Fig. 8.53.

Ans: $f_l = 7.5$ Hz

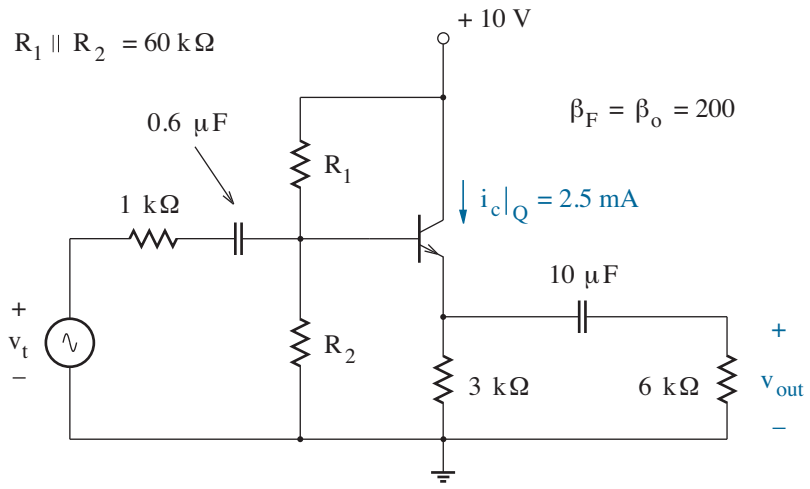


Figure 8.53: Circuit for Exercise 8.6.

Exercise 8.7 Estimate f_l for the amplifier of Fig. 8.54.

Ans: $f_l = 6.6$ Hz

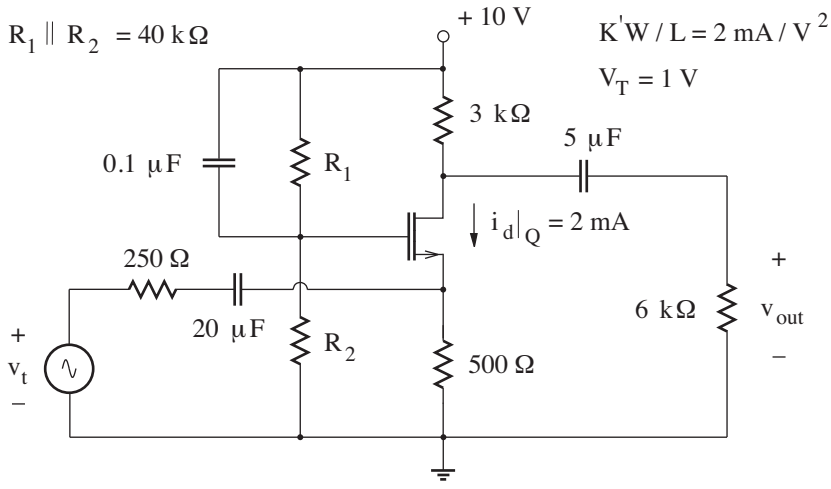


Figure 8.54: Circuit for Exercise 8.7.

Low-Frequency Amplifier Design

Low-frequency amplifier design selects a set of reasonable capacitor values, given a desired cutoff frequency. Several design procedures that are based on the inverse application of Eq. 8.109 are available (see Problem 8.62). However, the following method is particularly simple:

1. Match capacitors to f_l with appropriate short-circuit time constants.
2. Leave one capacitor value alone, and multiply the others by 10.

The first step establishes the same f_l for each of several *independent* design problems in which one capacitor "dominates" while the other capacitors function as ac short circuits. In practice, the capacitors act in concert. So the second step shifts all except one breakpoint to $f_l/10$ where the amplifier is not operated, and we are left with a single breakpoint at f_l. Put another way, the second step forces the approximate validity of one of the independent design problems.

An alternate procedure scales each f_l-matched capacitor value by the same factor (see Problem 8.62). However, the process typically results in greater total circuit capacitance than the preceding 10X method.

Example 8.8

The two-stage amplifier of Fig. 8.55 contains a MOSFET for which $K'W/L = 6$ mA, $V_T = 1$ V, and a BJT for which $\beta_F = \beta_o = 100$. Determine capacitor values that ensure a low-frequency cutoff at 25 Hz.

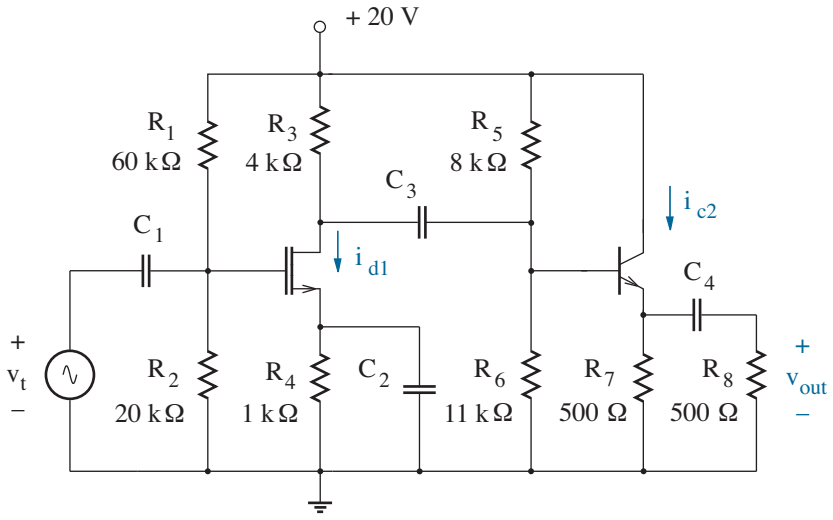


Figure 8.55: Circuit for Example 8.8.

Solution

As in Example 8.7, we skip over the details of a dc analysis that determines $i_{d1}|_Q = 3$ mA, $i_{c2}|_Q = 20$ mA, $g_{m1} = 6 \times 10^{-3}$ ℧ for the stage-1 MOSFET, and $r_{\pi 2} = 130\ \Omega$ for the stage-2 BJT.

Our next step is to match the capacitor values to their respective short-circuit time constants with $f_l = 25$ Hz. Expressions for the Thevenin small-signal resistances are obtained with the help of Table 8.1.

Capacitor C_1: $C_1 = [2\pi f_l {r_{eq}}^{(1)}]^{-1}$
${r_{eq}}^{(1)} = R_1 \parallel R_2 = 15$ kΩ.
$C_1 = 0.42\ \mu$F.

Capacitor C_2: $C_2 = [2\pi f_l {r_{eq}}^{(2)}]^{-1}$
${r_{eq}}^{(2)} = R_4 \parallel 1/g_{m1} = 143\ \Omega$.
$C_2 = 44\ \mu$F.

Capacitor C_3: $C_3 = [2\pi f_l {r_{eq}}^{(3)}]^{-1}$
${r_{eq}}^{(3)} = R_3 + R_5 \parallel R_6 \parallel [r_{\pi 2} + (1 + \beta_o)(R_7 \parallel R_8)] = 7.89$ kΩ.
$C_3 = 0.81\ \mu$F.

Capacitor C_4: $C_4 = [2\pi f_l {r_{eq}}^{(4)}]^{-1}$
${r_{eq}}^{(4)} = R_8 + R_7 \parallel [(r_{\pi 2} + R_5 \parallel R_6 \parallel R_3)/(1 + \beta_o)] = 522\ \Omega$.
(Note that R_3 is the Thevenin small-signal output resistance of stage 1.)
$C_4 = 12\ \mu$F.

Having determined the capacitor values on the basis of separate design calculations, we leave one capacitor value alone, and we multiply the others by 10. It is generally desirable to minimize the total capacitance in a circuit (to minimize cost, among other things), so we preserve the largest capacitor value (C_2). Thus,

$$\begin{aligned} C_1 &\rightarrow 4.2\ \mu\text{F}\,, \\ C_2 &= 44\ \mu\text{F}\,, \\ C_3 &\rightarrow 8.1\ \mu\text{F}\,, \\ C_4 &\rightarrow 120\ \mu\text{F}\,. \end{aligned}$$

How good is our design? Given the preceding capacitor values, a quick SPICE simulation yields $f_l = 25.7$ Hz. Not bad.

Low-frequency design has become less common over the years due to the rise of op-amps, which do not contain coupling and bypass capacitors.

Concept Summary

Transistor amplifiers suffer reduced voltage gains at high frequencies.

- The deterioration reflects internal small-signal parasitic capacitance.
  - C_{gs} (MOSFET) or C_π (C_{be}, BJT) weaken the drain or collector voltage-dependent current sources at high frequencies.
  - C_{gd} (MOSFET) or C_μ (C_{bc}, BJT) are problematic as feedback capacitors in common-source or common-emitter amplifiers.
  - MOSFETs can have significant C_{bd} and C_{bs} capacitance.
  - Small-signal capacitors are Q-point dependent.
- An inductor-less amplifier has an approximate high-frequency cutoff (voltage gain reduced by a factor of $1/\sqrt{2}$) that is inversely proportional to the sum of its open-circuit time constants:

$$f_h \approx \frac{1}{2\pi}\left(\sum \tau_{oc}\right)^{-1},$$

  where τ_{oc} is the product of a capacitance and the resistance "seen" by the relevant capacitor when all other capacitors are open circuits.
- Given feedback capacitance C_f, the Miller effect
  - Produces an effective $C_f(1 - {A_{vm}}^F)$ capacitance at the input,
  - Produces an effective $C_f(1 - {A_{vm}}^R)$ capacitance at the output,
  - Often promotes a tradeoff between voltage gain and bandwidth.
- A cascode amplifier partially mitigates the Miller effect with the help of common-gate/source or common-base/emitter amplifier pairings.

Transistor amplifiers can suffer reduced voltage gains at low frequencies.

- The deterioration reflects external coupling or bypass capacitance.
- An inductor-less amplifier has an approximate low-frequency cutoff (voltage gain reduced by a factor of $1/\sqrt{2}$) that is proportional to the sum of its reciprocal short-circuit time constants:

$$f_h \approx \frac{1}{2\pi}\sum \frac{1}{\tau_{sc}},$$

  where τ_{sc} is the product of a capacitance and the resistance "seen" by the relevant capacitor when all other capacitors are short circuits.

Problems

Section 8.1

8.1 Download the 2SJ148 MOSFET data sheet www.toshiba.com/taec), then use it to determine values for C_{gs} and C_{gd}.

8.2 Download the NTZD5110N MOSFET data sheet (www.onsemi.com), then use it to determine values for C_{gs} and C_{gd}.

8.3 The 2N2222 BJT has the SPICE parameters: CJE=22p, VJE=0.75, MJE=0.38, CJC=7.3p, VJC=0.75, MJC=0.34, TF=410p, and TR=47p. The device operates in the forward active mode with $i_c|_Q = 4$ mA and $v_{bc}|_Q = -3.5$ V. Find C_π and C_μ.

8.4 The NE68833 BJT has the SPICE parameters: CJE=0.8p, VJE=0.71, MJE=0.38, CJC=0.55p, VJC=0.65, MJC=0.48, TF=11p, and TR=32p. The device operates in the forward active mode with $i_c|_Q = 2$ mA and $v_{bc}|_Q = -7.2$ V. Find C_π and C_μ.

8.5 An npn BJT has emitter area A, base width w. The injected electron concentration at the emitter edge is n', and the electron concentration at the collector edge is approximately zero. Assume negligible electron/hole recombination in the base.

(a) Derive a relation for the electron current that diffuses through the base in terms of n', w, and D_e (the electron diffusion coefficient).

(b) Derive a relation for the charge store of holes within the base in terms of n' and w.

(c) Apply the results of parts a and b to show that

$$\tau_f = \frac{w^2}{2D_e}.$$

8.6 The 2N2222 BJT in the circuit of Fig. P8.6 has the SPICE parameters of Problem P8.3 in addition to IS=14f and BF=256. Devise SPICE experiments to show that f_t approaches $1/2\pi\tau_f$ as $i_c|_Q$ increases.

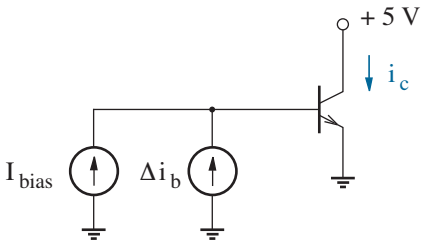


Figure P8.6

Section 8.2

8.7 Derive the expression for r_e' in Table 8.1.

8.8 Consider the BJT of Fig. P8.8. When finding the small-signal resistance r', careless application of Table 8.1 suggests $r' = r_b' \parallel r_c' = r_\pi \parallel \infty = r_\pi$. Apply the midfrequency small-signal BJT model to show a different result.

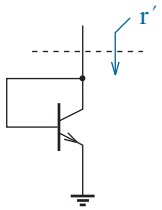


Figure P8.8

8.9 Verify Eqs. 8.26 and 8.27 for the common-gate amplifier.

8.10 Show that an amplifier with normalized gain of the form

$$\frac{A_v}{A_{vm}} = \left(\frac{1}{1+j\omega\tau_1}\right)\left(\frac{1}{1+j\omega\tau_2}\right)\left(\frac{1}{1+j\omega\tau_3}\right)\cdots$$

has the following approximate cutoff frequency:

$$f_h \approx \frac{1}{2\pi\sum \tau_i{}^2}.$$

What conditions apply?

8.11 The common-gate amplifier of Exercise 7.14 has $C_{gs} = 2.5$ pF and $C_{gd} = 1.5$ pF. Estimate the high-frequency cutoff, and compare with SPICE.

8.12 The common-base amplifier of Exercise 7.15 has $C_\pi = 2.5$ pF and $C_\mu = 1.5$ pF. Estimate the high-frequency cutoff, and compare with SPICE. (IS=10f)

8.13 The common-gate amplifier of Problem 7.53 has $C_{gs} = 2.8$ pF and $C_{gd} = 1.8$ pF. Estimate the high-frequency cutoff, and compare with SPICE.

8.14 The common-base amplifier of Problem 7.57 has CJE=2.5p, VJE=0.78, MJE=0.45, CJC = 1.5p, VJC=0.70, MJC=0.38. Estimate the high-frequency cutoff, and compare with SPICE. (IS=10f)

8.15 Use a small-signal transistor model to prove that the reverse midfrequency voltage gain is zero for a common-source or common-emitter amplifier. Assume $R_s{}'$ or $R_e{}'$ "looking-away" resistances.

8.16 Derive an exact result for the resistance "seen" by capacitor C_{gd} in the common-source amplifier of Fig. 8.16, and show that the corresponding time constant is equivalent to the sum of the time constants that are obtained using the Miller transformation.

8.17 The common-source amplifier of Exercise 7.10 has $C_{gs} = 2.5$ pF and $C_{gd} = 1.5$ pF. Estimate the high-frequency cutoff, and compare with SPICE.

8.18 The common-emitter amplifier of Exercise 7.11 has $C_\pi = 2.5$ pF and $C_\mu = 1.5$ pF. Estimate the high-frequency cutoff subject to an emitter bypass capacitor, and compare with SPICE. (IS=10f)

8.19 The common-source amplifier of Problem 7.37 has $C_{gs} = 2.8$ pF and $C_{gd} = 1.8$ pF. Estimate the high-frequency cutoff, and compare with SPICE.

8.20 The common-emitter amplifier of Problem 7.41 has CJE = 2.5p, VJE=0.78, MJE=0.45, CJC = 1.5p, VJC=0.70, MJC=0.38. Estimate the high-frequency cutoff with an emitter bypass capacitor, and compare with SPICE. (IS=10f)

8.21 Use a small-signal transistor model to verify Eq. 8.55 as the reverse midfrequency voltage gain for the common-drain or common-collector amplifier. Assume $R_s{}'$ ($R_e{}'$) and $R_g{}'$ ($R_b{}'$) looking away from the source (emitter) and gate (base), respectively.

8.22 Derive an exact result for the resistance "seen" by capacitor C_{gs} in the common-drain amplifier of Fig. 8.24, and show that the corresponding time constant is equivalent to the sum of the time constants that are obtained using the Miller transformation.

8.23 The common-drain amplifier of Exercise 7.12 has $C_{gs} = 2.5$ pF and $C_{gd} = 1.5$ pF. Estimate the high-frequency cutoff, and compare with SPICE.

8.24 The common-collector amplifier of Exercise 7.13 has $C_\pi = 2.5$ pF and $C_\mu = 1.5$ pF. Estimate the high-frequency cutoff, and compare with SPICE.

8.25 The common-drain amplifier of Problem 7.46 has $C_{gs} = 2.8$ pF and $C_{gd} = 1.8$ pF. Estimate the high-frequency cutoff, and compare with SPICE.

8.26 The common-collector amplifier of Problem 7.49 has CJE=2.5p, VJE=0.78, MJE=0.45, CJC= 1.5p, VJC=0.70, MJC=0.38. Estimate the high-frequency cutoff, and compare with SPICE. (IS=10f)

8.27 Repeat Problem P8.18, but leave out the bypass capacitor.

8.28 Repeat Problem P8.20, but leave out the bypass capacitor.

8.29 The multistage amplifier of Problem 7.60 has $C_\pi = 2.5$ pF and $C_\mu = 1.5$ pF for both BJTs. Estimate the high-frequency cutoff, and compare with SPICE. (IS=10f)

8.30 The multistage amplifier of Problem 7.61 has $C_{gs} = 1.8$ pF and $C_{gd} = 1.2$ pF for the MOSFET and $C_\pi = 2.5$ pF, and $C_\mu = 1.5$ pF for the BJT. Estimate the high-frequency cutoff, and compare with SPICE. (IS=10f)

8.31 The multistage amplifier of Problem 7.62 has $C_{gs} = 1.8$ pF and $C_{gd} = 1.2$ pF for the MOSFET and $C_\pi = 2.5$ pF, and $C_\mu = 1.5$ pF for the BJT. Estimate the high-frequency cutoff, and compare with SPICE. Assume IS=10f for the BJT.

8.32 The multistage amplifier of Problem 7.63 has $C_{gs} = 1.8$ pF and $C_{gd} = 1.2$ pF for the MOSFET and $C_\pi = 2.5$ pF, and $C_\mu = 1.5$ pF for the BJT. Estimate the high-frequency cutoff, and compare with SPICE. Assume IS=10f for the BJT.

Section 8.3

8.33 The MOSFETs in the circuit of Fig. P8.33 have $K'W/L = 2$ mA/V^2, $V_T = 0.7$ V, $C_{gs} = 0.2$ pF, and $C_{gd} = 0.1$ pF. Both devices also have 0.2-pF capacitance from drain or source to the most negative node potential. Estimate the high-frequency cutoff, and compare with SPICE.

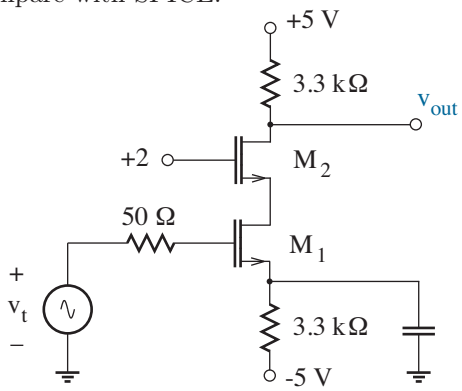


Figure P8.33

8.34 Redesign the cascode amplifier of Problem P8.33 so that $A_{vm} = -5$ and $v_{ds1}|_Q$ is 1 V higher than that for M_1 saturation. Do this in a manner that maximizes f_h, and verify with SPICE.

8.35 Redesign the cascode amplifier of Problem P8.33 to include M_3 as part of a common-drain source follower that "drives" the capacitive load. Assume M_3 has the same electrical characteristics as M_1 and M_2, and use a current source for biasing. Find the new f_h, and compare with SPICE.

8.36 Compare the midfrequency voltage gains that apply to a simple common-emitter amplifier and the cascode amplifier that derives from it.

8.37 The 2N2222 BJTs in the cascode circuit of Fig. P8.37 have IS=14f, BF=256, CJE=22p, VJE = 0.75, MJE=0.38, CJC=7.3p, VJC=0.75, MJC=0.34, TF=410p, and TR=47p.

(a) Find f_h, and compare with SPICE.

(b) Repeat part a, but eliminate Q_2 in favor of a short circuit from Q_1 to the 2-kΩ load resistor.

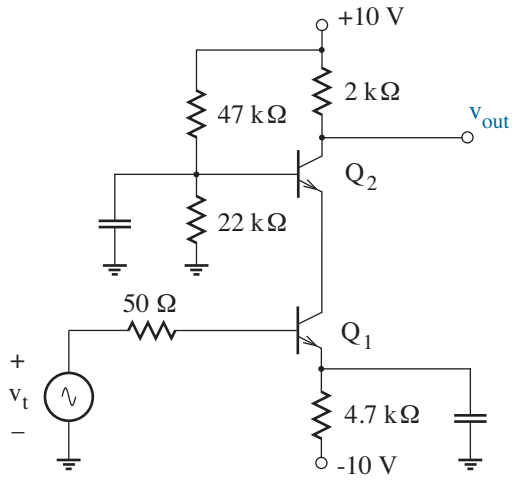


Figure P8.37

8.38 Redesign the circuit of Problem P8.37 with the same bias currents and 2-kΩ collector resistor, but change $v_{ce}|_Q$ for Q_1 and Q_2 so as to maximize f_h. Assume 50-mV peak-to-peak input amplitude.

8.39 Consider the circuit of Problem P8.33.

(a) Apply shunt peaking to maximize the amplifier bandwidth.

(b) Apply shunt peaking to increase the amplifier bandwidth subject to a maximally "flat" magnitude response.

Use SPICE to demonstrate the design results.

8.40 Repeat Problem P8.39, but consider the circuit of Problem P8.37.

8.41 Apart from promoting maximum bandwidth or a maximally flat magnitude response, an alternative choice of the m factor defined in Eq. 8.73 promotes a maximally flat delay so that amplified signals have minimal phase distortion.

(a) Use SPICE and a trial-by-error process to determine m (and thus L) for the cascode amplifier of Fig. P8.33 so that the delay is maximally flat.

Hint: To find the group delay with .probe, plot VG(x) vs. frequency. Node x is the output.

(b) Show the change in group delay changes when m is chosen to achieve a maximally flat magnitude response at the output.

(c) Demonstrate a compromise choice for m that balances the objectives of parts a and b.

8.42 Consider the tuned amplifier circuit of Fig. 8.34.

(a) Draw the ac circuit that includes a small-signal BJT model with C_π and C_μ. Assume $r_o \to \infty$.

(b) Determine the real and imaginary components of the admittance looking into the BJT collector for non-zero R_t. What do the results imply?

(c) Determine the real and imaginary components of the admittance looking into the BJT base. What unusual conditions apply?

8.43 Consider the single-stage tuned amplifier circuit of Fig. P8.43. The BJT features IS=4f, BF=180, TF=65p, CJE=1.4p, and CJC=1.0p.

(a) Design the circuit so that the peak voltage gain is -25 at 400 MHz when $C = 12$ pF. The inductor has $Q_L = 18$. Specify the amplifier bandwidth. Assume $R_t = 0$. Ignore C_μ (for now).

(b) Use SPICE to demonstrate the design of part a subject to $R_t = 0$. Show appropriate adjustments that establish the intended voltage gain and center frequency.

(c) Repeat part b for $R_t = 75\ \Omega$. Adjustments?

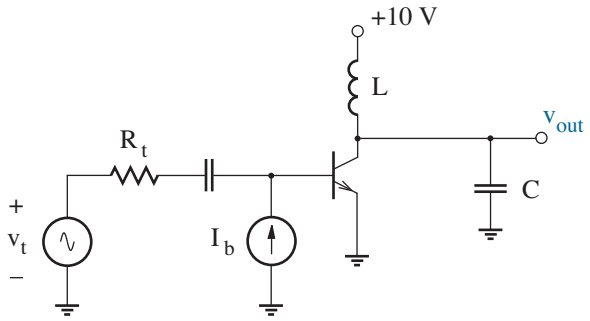


Figure P8.43

8.44 Design the tuned amplifier of Fig. 8.36 so that the peak voltage gain is -32, the center frequency is 650 MHz, and the available bandwidth is 40 MHz. Assume that inductors are available with $Q_L = 22$, and assume $C = 15$ pF. Use SPICE to verify the design with the BJT parameters of Example 8.6.

8.45 Design the tuned amplifier of Fig. 8.36 so that the peak voltage gain is -25, the center frequency is 420 MHz, and the available bandwidth is 65 MHz. Assume that inductors are available with $Q_L = 18$, and assume $C = 18$ pF. Use SPICE to verify the design with the BJT parameters of Example 8.6, but let CJE=1.2p and CJC=0.8p.

8.46 Find f_h for Example 8.4 with $r_b = 12\ \Omega$.

8.47 Find f_h for Exercise 8.3 with $r_b = 20\ \Omega$.

8.48 Repeat Problem 8.14, but let $r_b = 10\ \Omega$.

8.49 Repeat Problem 8.26, but let $r_b = 10\ \Omega$.

8.50 Consider the high-frequency BJT model of Fig. 8.42 with $C_{jx} = 0$ and $r_e = r_c = 0$.

(a) Use the model to demonstrate the b_{ie} maximum of Eq. 8.89.

(b) Use the model to explain the g_{ie} behavior in Fig. 8.41.

8.51 Estimate r_b given the y_{ie} data of Fig. P8.51. Data used with permission from SCILLC dba ON Semiconductor.

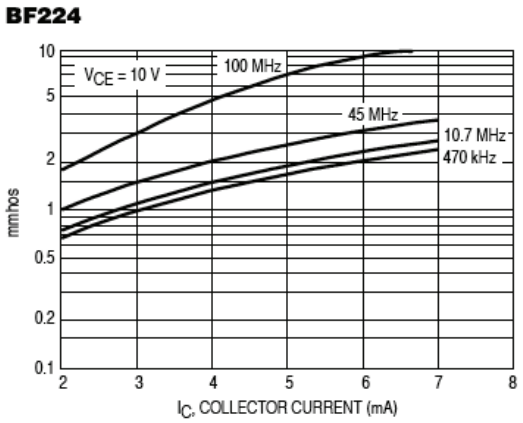


Figure P8.51

8.52 Consider the high-frequency BJT model of Fig. 8.42 with $C_{jx} = 0$ and $r_e = r_c = 0$, and consider the y-parameter data of Fig. 8.41.

(a) Use the model to explain the y_{fe} behavior.

(b) Use the model to explain the y_{re} behavior.

(c) Use the model to explain the y_{oe} behavior.

8.53 Determine f_h for the common-source amplifier of Exercise 8.2 if the MOSFET has parasitic gate resistance $r_g = 10\ \Omega$.

Section 8.4

8.54 Consider the circuit of Problem P7.60. Let all coupling and bypass capacitors have 10-μF value. Estimate f_l and compare with SPICE.

8.55 Design the circuit of Problem P7.61 so that $f_l = 20$ Hz. Use SPICE to verify your design.

8.56 Consider the circuit of Problem P7.61. Let all coupling and bypass capacitors have 10-μF value. Estimate f_l and compare with SPICE.

8.57 Design the circuit of Problem P7.61 so that $f_l = 20$ Hz. Use SPICE to verify your design.

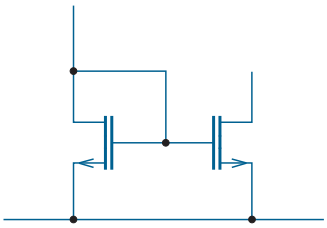
8.58 Consider the circuit of Problem P7.62. Let all coupling and bypass capacitors have 10-μF value. Estimate f_l and compare with SPICE.

8.59 Design the circuit of Problem P7.62 so that $f_l = 20$ Hz. Use SPICE to verify your design.

8.60 Consider the circuit of Problem P7.63. Let all coupling and bypass capacitors have 10-μF value. Estimate f_l and compare with SPICE.

8.61 Design the circuit of Problem P7.63 so that $f_l = 20$ Hz. Use SPICE to verify your design.

8.62 A circuit has n coupling or bypass capacitors. Determine a design procedure so that each capacitor makes an equal contribution to the low-frequency cutoff as determined by the method of short-circuit time constants. Apply this procedure to the circuit of Problem P7.63 to achieve $f_l = 50$ Hz, and use SPICE to verify your design.



Interlude

Whereas we are about to examine integrated circuits in Chapters 9 and 10, we need an appreciation of modern integrated-circuit fabrication processes and the circuit constraints that are associated with a particular "layout." We concentrate on the MOSFET and a CMOS process, leaving the BJT considerations for other texts.

Fundamental Fabrication Processes

Don't try this at home.

Start with quartzite (SiO_2) from the sands of the earth. Add coal, coke, and wood chips to provide sources of carbon. Mix well. Place in a special furnace under high heat to support the reaction:

$$SiO_2 + 2C \rightarrow Si + 2CO\,(gas) \uparrow .$$

Draw off the liquid byproduct, and cool. Pulverize the metallurgical-grade (98-% pure) silicon, and react with gaseous hydrogen chloride:

$$Si + 3HCl \rightarrow SiHCl_3\,(gas) \uparrow + H_2\,(gas) \uparrow + \text{heat} .$$

Refine the trichlorosilane ($SiHCl_3$), a room-temperature liquid, by passing it through a distillation column. Return to a second special furnace—

$$SiHCl_3(gas) + H_2(gas) \rightarrow Si + 3HCl\,(gas) \uparrow$$

$$SiHCl_3(gas) + HCl(gas) \rightarrow SiCl_4 + H_2\,(gas) \uparrow$$

—with electronic-grade silicon as the result of the first competing reaction. This material is amazingly pure (one part per billion), but not single crystal, the form of silicon needed for electronic devices.

To obtain single-crystal semiconductor, melt the electronic-grade silicon in yet another special furnace featuring an inert (typically Ar) environment. Then lower a single-crystal "seed" that is carefully oriented at the end of a metal shaft until it touches the liquid surface. While rotating the seed, raise it at a very slow rate, typically about 2 mm/minute. As the molten silicon in contact with the seed cools, it assumes the same crystalline orientation. Thus, and over the course of many hours, "pull" a cylindrical silicon boule with over 200-mm diameter and 2-m length. Shape the boule to a uniform 200-mm diameter with appropriate cutting tools, then saw into thin slices. Polish each silicon "wafer" on one side until it is uniformly flat.

Now you are ready to make an integrated circuit.

Semiconductor device fabrication, like most other manufacturing efforts, involves three basic process categories:

- Additive processes (building up)
- Subtractive processes (tearing down)
- Modification of material properties

We briefly examine each topic for a silicon technology prior to describing typical CMOS and BJT processes.

Additive Processes

Arguably the most important additive silicon process reflects the ability of Si to oxidize or "rust" with the build-up of a silicon dioxide (SiO_2) insulator. Thermal oxidation yields an SiO_2 layer with a thickness dependent on time, temperature (typically 900 - 1100 °C), and ambient (dry oxygen or steam), and the SiO_2/Si interface is of high quality. Two other growth processes, expitaxy (Si onto Si) and heteroepitaxy ($Si_{1-x}Ge_x$ onto Si, for example), establish single-crystal material over a crystalline seed. Growth is effected through interactions with appropriate chemical vapors or molecular beams.

Deposition provides an overcoat layer. Chemical vapor deposition (CVD) is a process in which gases react at moderate temperatures to leave behind a thin residual coating of SiO_2, albeit with an interface of meager quality, or thermally resilient conductive materials such as polysilicon. Evaporation is a thermal or electron-beam-assisted non-reactive process that deposits thin metal films such as Al. Electrochemical reactions produce Cu films. Sputtering is a snow-storm-like process that moves moderately conductive compounds such as $TaSi_2$ from a purified source to a receptive substrate. Exposure to liquid suspensions followed by solvent evaporation is a means for depositing insulating polymer films including polyimide and paralene.

The uniformity of an additive process depends on process conditions and the topography of the recipient material system. In Fig. I1a, for example, a silicon step has a conformal coating of SiO_2 with the same thickness on every surface. Figure I1b shows non-conformal step coverage.

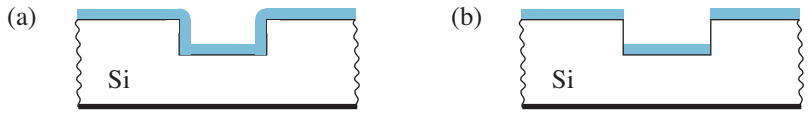


Figure I1: Step coverage: (a) conformal; (b) non-conformal.

Additive processes are broadly applied—there is no selective paint brush. However, a notable exception is the localized oxidation process of Fig. I2. Silicon areas covered by Si_3N_4 are not subject to oxidation apart from slight incursions at the edges of the protective layer. The local oxidation process is used to define active integrated-circuit regions that are electrically isolated.

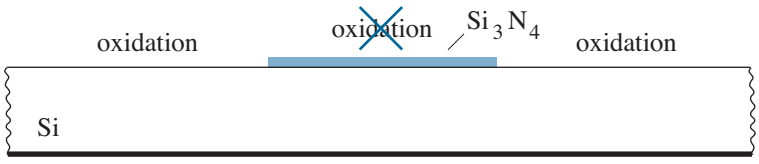


Figure I2: Selective oxidation process. A Si_3N_4 film prevents SiO_2 growth.

Subtractive Processes

Subtractive integrated-circuit processes typically involve wet or dry etching. Wet etching is a chemical process that removes material A while preserving material B. For example, hydroflouric acid (HF) removes SiO_2 and not Si, while a solution of potassium hydroxide (KOH) eliminates Si and not SiO_2. Dry etching achieves similar selectivity when materials A and B are exposed to a low-pressure gaseous discharge or plasma with energetic reactants.

Etching can impart non-uniformity to a materials system with the help of a companion process, lithography. Figure I3 shows the four basic steps: Step 1 deposits material called resist over a broad area by wetting a rapidly spinning silicon wafer with a polymer suspended in a volatile liquid solvent. A short low-temperature bake evaporates the solvent after spinning to leave behind a film with a thickness of about 1-2 μm. Step 2 exposes portions of the resist to radiation. In the case of ultraviolet exposure, a mask defines a desired pattern in relation to a previously established alignment mark. In the case of electron-beam exposure, as for very-high-resolution patterns, direct-write capability avoids a mask entirely. Step 3 develops the resist. Positive resists become more soluble in exposed areas, which are dissolved away during development. A low-temperature bake toughens what remains. Step 4 subjects unprotected material to wet or dry etching. When the etch is complete, the resist is stripped away in its entirety.

Although there are some exceptions, wet etches are generally isotropic—the rate of material removal is independent of direction. Thus, wet etching of thick films produces undercutting with respect to the resist window, and there is loss of resolution. In contrast, dry etches are generally anisotropic—the vertical etch rate is highly favored. Modern integrated-circut processes tend to use dry etching to achieve small feature sizes.

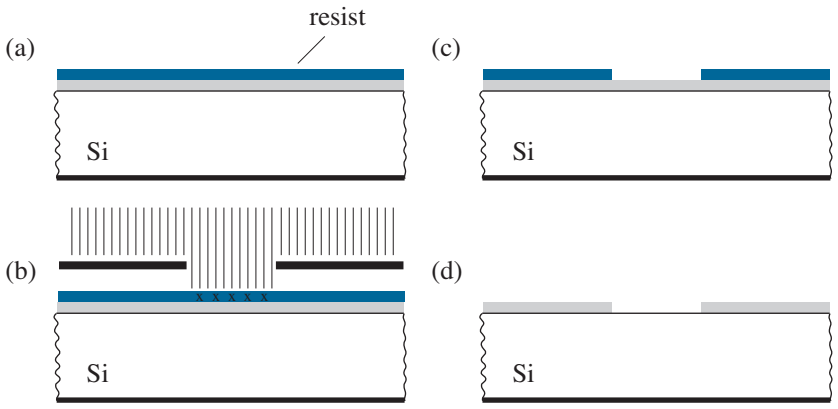


Figure I3: Basic four-step lithography process: (a) Resist deposition; (b) exposure; (c) development; (d) etching and resist removal.

Material Modification

Processes that modify electronic materials usually involve the introduction and redistribution of dopant impurities (such as B, P, or As for silicon). Two fundamental processes are of importance.

An "ancient" doping process deposits an impurity-rich SiO_2 (glass) layer as an impurity source over a silicon substrate. In turn, dopants are allowed to diffuse into the substrate. The rate of solid-state diffusion is very slow. Nevertheless, high temperatures of the order of 1000 °C are sufficient for μm-scale redistribution in the course of a few hours. Boron diffuses quickly, phosphorus, and particularly arsenic, are relatively sluggish.

Modern fabrication processes insert impurities using ion implantation. The procedure begins with what amounts to an ion "soup" produced by an electrical discharge in a rarified gas such as BF_3. The ionic constituents—B^+_{10}, B^+_{11}, F^+_{19}, $BF^+_{2\,(49)}$, etc. —are extracted by means of an electric field and separated with a magnetic field (as in the action of a mass spectrometer). Then one constituent, say B^+_{11}, is accelerated to high energy (50 - 300 keV) and made to collide with the material surface under process. To first order, the ions come to rest beneath the surface with the Gaussian distribution shown in Fig. I4. Range R and standard deviation Δ are energy dependent, and the area under the curve, the number of implanted ions per unit area, relates to the time integral of the impinging ion-beam current density.

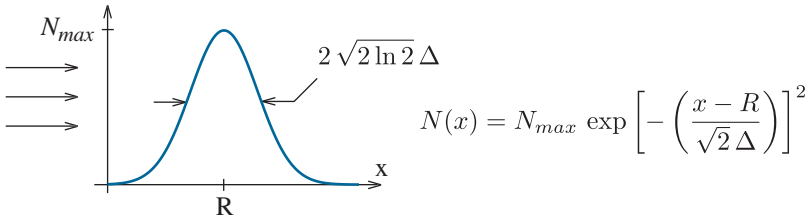


Figure I4: Impurity concentration distribution following ion implantation.

Unfortunately, ion implantation induces significant damage to a semiconductor crystal as ions undergo collisions during the stopping process. A subsequent thermal annealing step is required so that this damage can be repaired. In the early days of ion implantation, annealing meant high-temperature treatment in a furnace with an inert (non-oxidizing) ambient, and the process invariably led to some degree of impurity diffusion and loss of localization. Later, it was discovered that annealing could be effected with minimal diffusion using a laser as the recrystallization energy source. Still later, a non-coherent arc lamp was found to be sufficient, so this is the method used today. When ion implantation is intended to provide a fixed amount of impurities prior to drive-in diffusion, the issue of bombardment damage is not a problem. The high temperatures needed for diffusion allow for coincident lattice repair.

A CMOS Fabrication Process

Now apply the basic fabrication principles to a typical CMOS process. Apart from understanding the numerous fabrication steps, we need to be familiar with the associated design rules that determine integrated-circuit geometries with minimum dimensions. The process design rules will be used to realize a compact CMOS inverter with moderate parasitic capacitance. Our more immediate concern is a pair of equally-sized and adjacent n- and p-channel MOSFETs shown schematically in Fig. I5.

Figure I5: End-products of the typical CMOS process under discussion. Metal interconnections yield an inverter or part of another CMOS circuit.

> The design rules that follow reflect the "SCMOS" Scalable CMOS process administered by MOSIS, an affiliate of the Information Sciences Institute at the Unversity of Southern California.

The first few process steps localize the n- and p-channel MOSFETs.

1. Start with a p-type Si wafer. Inspect it for visual defects and flatness, and measure the wafer resistivity to determine the background boron concentration.

2. Clean the Si wafer, first in a hot solution of H_2O_2 and NH_4OH, and second in a solution of H_2O_2 and HCl to remove organic and metallic contaminants, respectively. This is called an "RCA" clean after the company that developed it in the 1970s.

3. Grow a uniform and moderately thick (200-nm) SiO_2 "barrier" layer. Then apply lithography with **Mask 1** and etch the exposed SiO_2 to define **n-wells**, the body regions designated for p-channel MOSFETs. n-channel devices will reside in the protected substrate. A similar p-well process applies when the starting material is n-type.

4. Grow a thin (20-nm) SiO_2 "pre-implant" layer in the n-well regions. Among other things, this helps to reduce the directional dependence of ion stopping during implantation, and it helps to establish favorable surface conditions during drive-in diffusion. In theory, one could have removed just enough oxide in step 3 to leave exactly 20 nm behind. In practice, the inherent difficulty of etch-rate control makes it more reliable to remove everything and then put the right thickness back.

5. Perform the n-well implant (phosphorus). Then diffuse the implanted dopants until a desired pn junction depth has been achieved on the basis of a carefully tested process simulation model and experiment. The diffusion typically requires considerable time.

Figure I6 shows the integrated circuit at the completion of Step 5.

Figure I7 exemplifies circuit-layout design rules that apply to Mask 1. Minimum dimensional requirements are specified in units of λ ("lambda"), equal to half the channel length used to characterize the CMOS process. The minimum n-well width leaves enough room for a functional MOSFET. The minimum spacings help to ensure adequate electrical isolation.

The next process steps promote electrical isolation between devices.

6. Remove all SiO_2 that originates from Steps 3 and 4.

7. Grow a thin (20-nm) SiO_2 pad layer, then deposit a somewhat thicker (50-nm) Si_3N_4 film. The pad layer helps to alleviate stress caused by the difference in the coefficients of thermal expansion for Si and Si_3N_4.

8. Apply lithography with **Mask 2** and etch the exposed Si_3N_4 to define **field regions**. These are the inverse of the **active regions** that are receptive to MOSFET gates, drains, sources, and body contacts.

9. Perform the channel-stop implant (boron). The boron concentration that is introduced should be large enough to help raise the threshold voltage in the field yet small enough to avoid reverse breakdown at n^+p source or drain junctions that have yet to be formed. The boron implant energy should be sufficiently low to allow the uneteched Si_3N_4 and SiO_2 films to serve as a protective mask.

10. Grow a thick (1000-nm) SiO_2 layer in the Mask-2-defined field regions. This local oxidation (LOCOS) step further increases the threshold voltage in the field. Care should be taken to minimize the so-called "bird's beak", the profile that results when oxidation is allowed to make substantial undercutting at the Si_3N_4 periphery.

Figure I8 shows the integrated circuit at the completion of Step 10.

Figure I9 exemplifies circuit-layout design rules that apply to Mask 2. The n^+ select and p^+ select masks define the corresponding source/drain implant regions and are considered shortly.

Apart from providing isolation between devices, the field regions support the meanderings of interconnections that give CMOS circuit functionality. The parasitic areal capacitance between an overlying interconnect and the Si substrate is reduced as the field SiO_2 is made thicker.

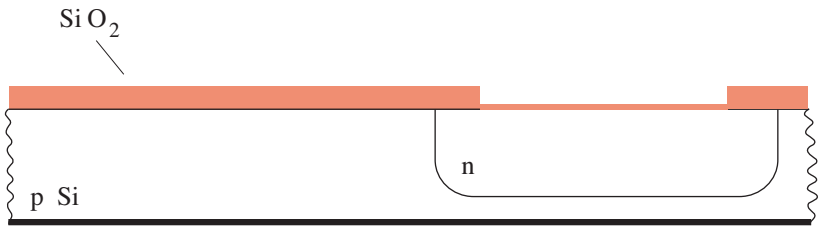


Figure I6: Integrated circuit profile after Step 5 in the CMOS process.

Layout Rules: n-well

Rule	Description	Lambda
1.1	Minimum Width	10
1.2	Minimum Spacing Between Wells at Different Potential	9
1.3	Minimum Spacing Between Wells at Same Potential	6

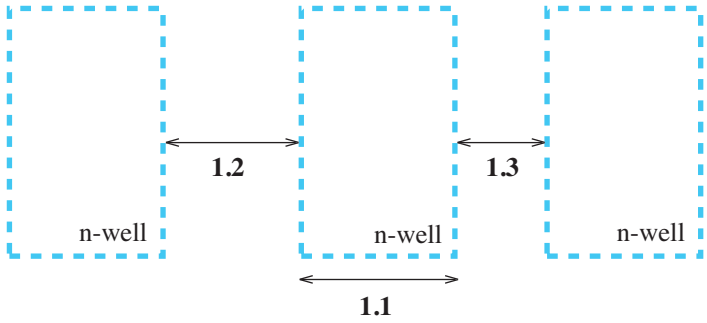


Figure I7: Top view and layout design rules relating to Mask 1 (n-well). p-channel MOSFETs reside within the n wells; n-channel MOSFETs reside in the p-substrate region outside.

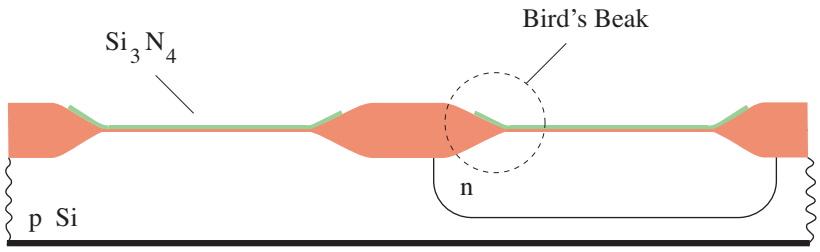


Figure I8: Integrated circuit profile after Step 10 in the CMOS process.

Layout Rules: Active

Rule	Description	Lambda
2.1	Minimum Width	3
2.2	Minimum Spacing	3
2.3	Spacing to Edge of Well (Active Region Contains MOSFET)	5
2.4	Spacing to Edge of Well (Active Region Contains Well Contact)	3
2.5	Minimum Spacing Between Active Regions of Different Implant	4

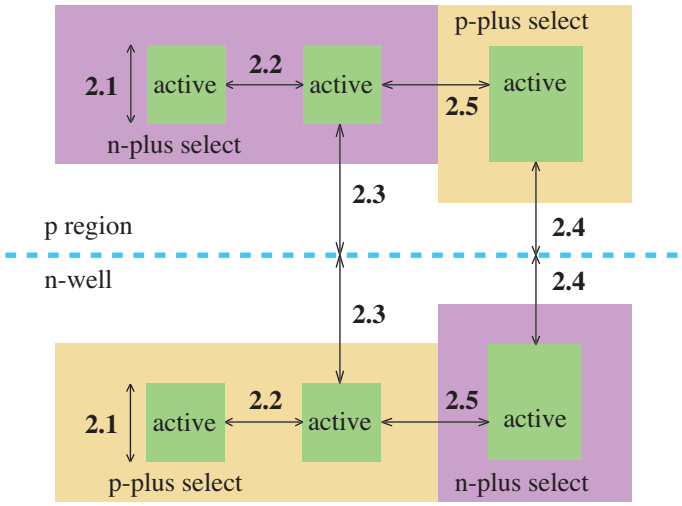


Figure I9: Top view and layout design rules relating to Mask 2 (active). Active regions contain a MOSFET source, drain, and channel.

Steps 11 through 15 establish the MOSFET channel regions.

11. Remove the Si_3N_4 and pad oxide from Step 7. The latter film is thin, so the SiO_2 etch has minimal impact in the thick field regions.
12. Grow a high-quality SiO_2 gate oxide layer with 20-80 nm thickness. The growth conditions increase the field-oxide thickness slightly.
13. Deposit a medium thick (400-nm) polysilicon film. In some processes, an additional layer of silicide material (such as $TaSi_2$) is subsequently deposited to improve the overall conductivity.
14. Apply lithography with **Mask 3** and etch the exposed polysilicon to define **poly regions**. These serve as the MOSFET gates and short-distance interconnections to same.
15. Grow a thin (50-mn) oxide layer over the remaining poly.

Figure I10 shows the integrated circuit at the completion of Step 15.

Figure I11 exemplifies circuit-layout design rules that apply to Mask 3. Rule 3.3 is particularly critical. If violated, a source and drain will become connected when these regions are formed by ion implantation.

Steps 16 through 20 create n^+ and p^+ transistor and contact regions.

16. Apply lithography with **Mask 4A** to define **n^+ select regions**. These allow n^+ sources and drains for the n-channel MOSFETs and n^+ contacts to the n-well body for p-channel devices. Use the resist remaining after development to mask against arsenic implantation. The poly gate masks the channel region to implement self alignment between the edges of a gate and the adjoining n^+ source and drain. The thick SiO_2 field regions provide peripheral masking.
17. Strip the resist used in Step 16.
18. Apply lithography with **Mask 4B** to define **p^+ select regions**. These allow p^+ sources and drains for the p-channel MOSFETs and p^+ contacts to the substrate for n-channel devices. Use the resist remaining after development to mask against boron implantation. The poly gate and SiO_2 field have the function noted in Step 16.
19. Strip the resist used in Step 18.
20. Perform a rapid thermal anneal to eliminate ion implantation damage.

Figure I12 shows the integrated circuit at the completion of Step 20.

Figure I13 exemplifies circuit-layout design rules that apply to Mask 4 (both n^+ select and p^+ select).

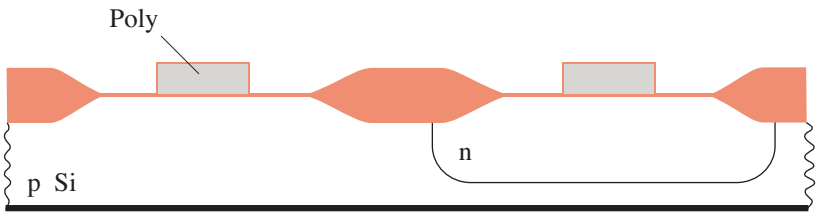


Figure I10: Integrated circuit profile after Step 15 in the CMOS process.

Layout Rules: Poly

Rule	Description	Lambda
3.1	Minimum Width	2
3.2	Minimum Spacing	2
3.3	Minimum Gate Extension over Active	2
3.4	Minimum Active Extension from Poly	3
3.5	Minimum Field Poly to Active	1

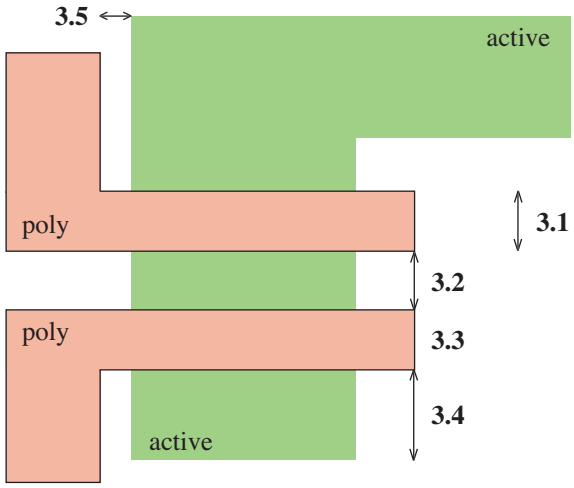


Figure I11: Top view and layout design rules relating to Mask 3 (poly). MOSFET channels will exist where the poly and active areas intersect.

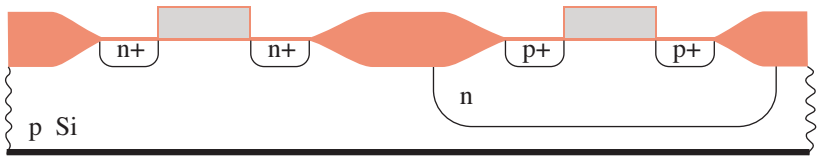


Figure I12: Integrated circuit profile after Step 20 in the CMOS process.

Layout Rules: Select

Rule	Description	Lambda
4.1	Minimum Spacing to Poly Edge over Channel	3
4.2	Minimum Overlap of Active	2
4.3	Minimum Overlap of Contact	1
4.4	Minimum Select Width and Spacing (not illustrated)	2

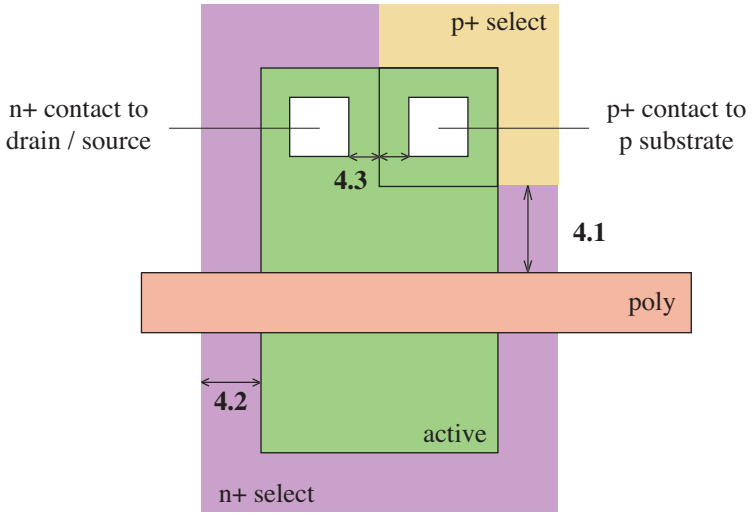


Figure I13: Design rules relating to Mask 4 (n+ select and p+ select). Similar rules apply when the select regions are reversed.

The remaining steps add features needed for electrical interconnections.

21. Deposit a thick (1-μm) layer of SiO_2 on the entire wafer.

22. Apply lithography with **Mask 5** and etch the exposed SiO_2 to define **contact cuts**. These are made to a source, drain, or body region.

Figure I14 shows the integrated circuit at the completion of Step 22.

Figure I15 exemplifies circuit-layout design rules that apply to Mask 5. Rule set 5.x applies to poly contacts. Rule set 6.x is the same and applicable to body contacts. Specifying an exact contact size ($2\lambda \times 2\lambda$) necessitates many contact structures as opposed to a huge contact in large-area regions. Redundant contacts promote reliability and favorable current distributions.

> We deliberately avoid the designation of a Mask 6 so that the standard numbering used in conjunction with the various design-rule sets generally corresponds to the mask that is being used.

23. Deposit a layer of Al. Then apply lithography with **Mask 7** and etch the exposed metal to define **metal 1**. This serves as the lowest level of interconnect. Perform a low-temperature "sintering" treatment to form an ohmic contact between the metal and semiconductor.

Figure I16 shows the integrated circuit at the completion of Step 23.

Figure I17 exemplifies circuit-layout design rules that apply to Mask 7.

Further processing is needed to produce additional metal layers when Metal 1 does not suffice.

24. Deposit a thick (1-μm) layer of polymer insulator. Apply lithography with **Mask 8** and etch the exposed polymer to define holes or **vias**. The vias allow connections from metal 1 to metal 2. In the event that metal 2 needs to connect to a poly or active region, a metal-1 pad should have been previously formed over the lower-level feature.

25. Deposit a layer of Al. Then apply lithography with **Mask 9** and etch the exposed metal to define **metal 2**. This serves as the second level of interconnect. Advanced integrated-circuit processes that support complex systems sometimes feature additional metallization layers. In this case, Steps 24 and 25 are repeated with minor variations.

Steps 24 and 25 are not depicted—they go beyond the needs of this text. We have also avoided a set of final steps that pertain to corrosion protection and packaging. These steps are important, nonetheless.

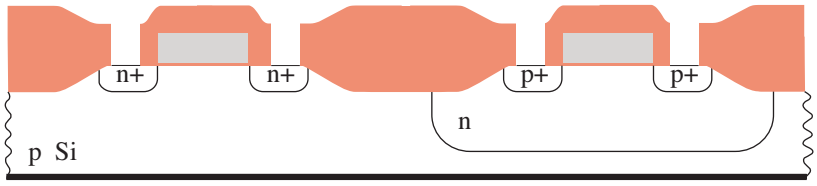


Figure I14: Integrated circuit profile after Step 22 in the CMOS process.

Layout Rules: Set 5 -Contact to Poly (Set 6 - Contact to Active)

Rule	Description	Lambda
5.1, 6.1	Exact Contact Size	2×2
5.2, 6.2	Minimum Poly Overlap	$2^{\dagger}$
5.3, 6.3	Minimum Contact Spacing	2
5.4, 6.4	Minimum Spacing to Transistor Gate (Poly)	2

†Less conservative processes sometimes feature 1.5λ for rules 5.2 and 6.2.

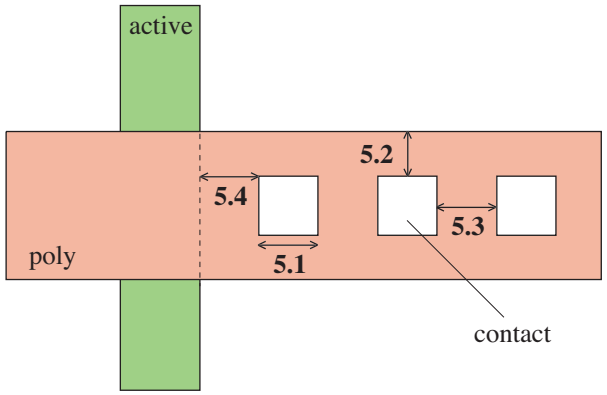


Figure I15: Top view and layout design rules relating to Mask 5 (contact). Contact cuts are made to poly (gate) or active (source/drain) regions.

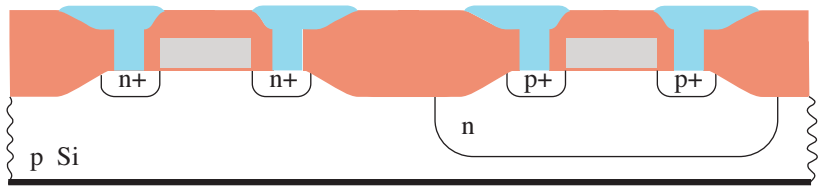


Figure I16: Integrated circuit profile after Step 23 in the CMOS process.

Layout Rules: Metal 1

Rule	Description	Lambda
7.1	Minimum Width	3
7.2	Minimum Spacing	2
7.3	Minimum Overlap of any Contact	1
7.4	Minimum Spacing When Either Line Wider than 10 λ	4

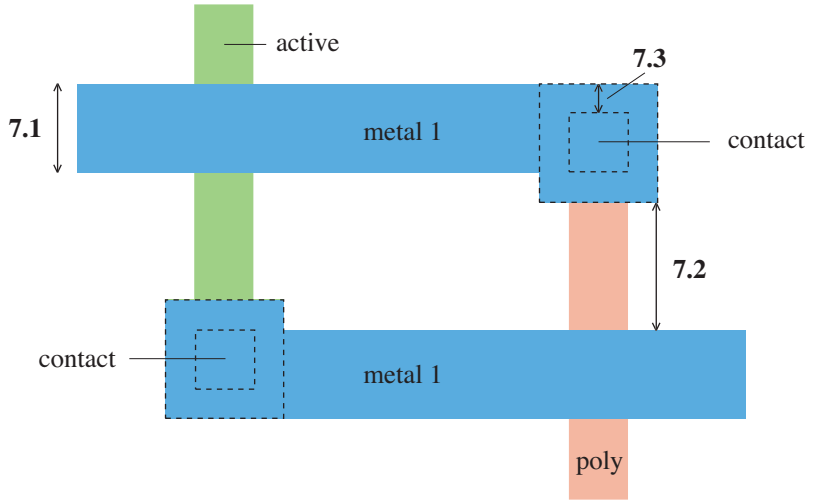


Figure I17: Top view and layout design rules relating to Mask 7 (metal 1). Metal is used for most circuit interconnections.

CMOS Layout and Capacitance

At last we are ready to specify a set of mask patterns that determine the layout of a CMOS circuit. The practice is an art that is rooted in science, and it demands both cleverness and experience. We will be content to view a few examples, especially in support of the CMOS inverter of Chapter 10. We also consider the *extraction* of circuit parasitics such as capacitance.

Consider a single isolated n-channel MOSFET. The active source and drain regions on either side of the poly gate must measure at least $6\lambda \times 6\lambda$ — $2\lambda \times 2\lambda$ for a contact cut (rule 6.1), 2λ spacings on the sides (rule 6.2). The poly gate width is 2λ, by definition. Thus, the minimum practical active area is $6\lambda \times 14\lambda$ with $W/L = 3/1$ as shown in Fig. I18. Note that there is plenty of room for metal source/drain connections per rule 7.3.

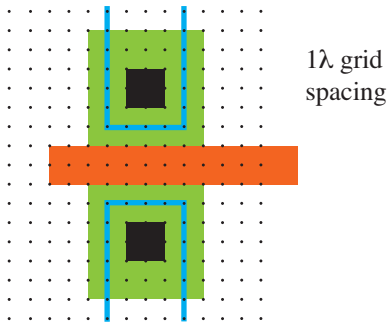


Figure I18: Layout for an n-channel MOSFET with minimum-size active area and $W/L = 3/1$. The poly contact is in the field region to the right.

Figure I19 shows a somewhat larger n-channel MOSFET ($W/L = 9/1$). The active area is $18\lambda \times 14\lambda$, and there is an abundance of source/drain contact cuts to ensure even current distributions.

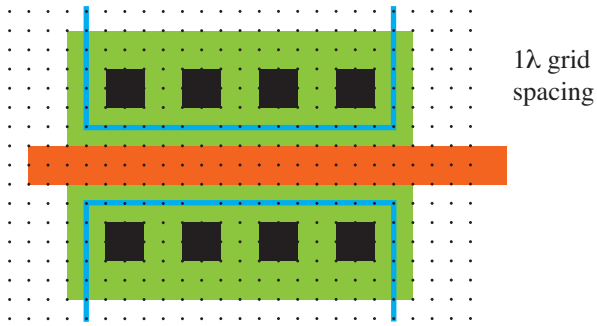


Figure I19: Layout for an n-channel MOSFET with $W/L = 9/1$.

When very large W/L values are required (as in certain analog circuits), it is usually desirable to "finger" the poly gate with two or more branches so that smaller MOSFETs are connected in parallel. Figure I20 shows a three-finger layout with $W/L = 27/1$. The active area is $18\lambda \times 30\lambda$.

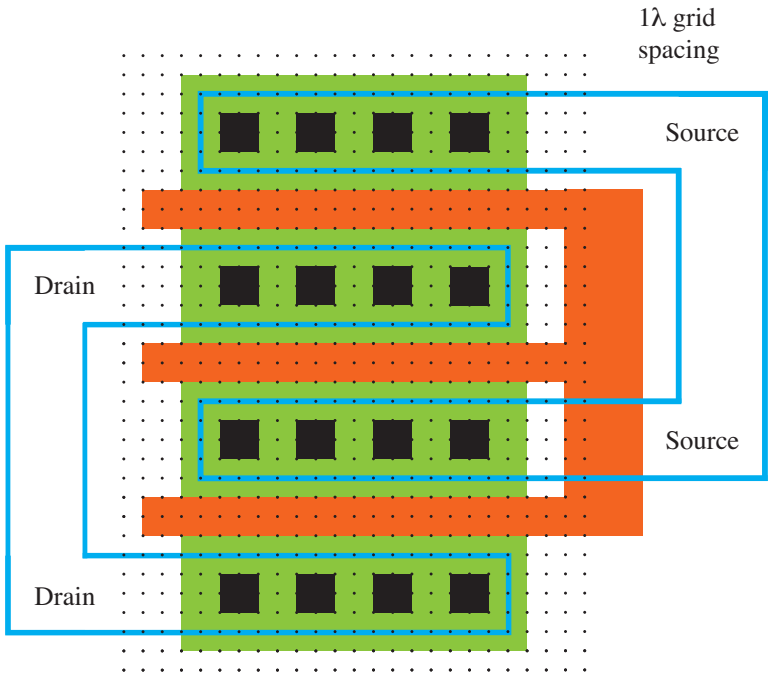


Figure I20: Fingered n-channel MOSFET with $W/L = 36/1$.

Channels are sometimes bent. Figure I21 shows an n-channel MOSFET with $W/L = 6/1$ and a relatively small drain area for reduced capacitance.

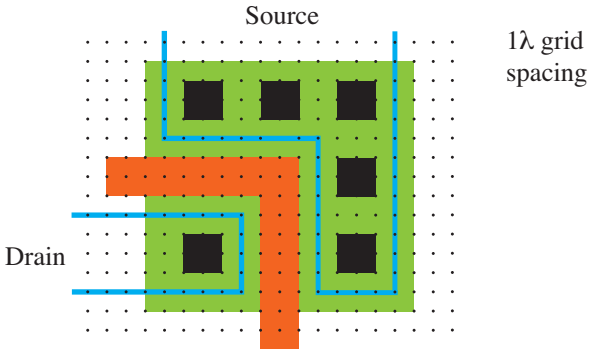


Figure I21: L-shaped n-channel MOSFET with $W/L = 6/1$.

The preceding layout considerations also apply to p-channel MOSFETs, but with the constraint of a surrounding n-well with appropriate device-to-boundary spacings. An n-well can contain more than one transistor.

Further note that the examples so far have individual mask patterns that feature rectangular polygons, which, in turn, are merged so that mask-layer paths go either up/down or left/right in a so-called "Manhattan" geometry. Irregular shapes are generally avoided in order to simplify the data needed for computer-aided design. Paths with 45° orientation are used sparingly.

Now put everything together to form two adjacent CMOS inverters as shown in Fig. I22. With the output of one inverter connected to the input of the other, this will form part of a circuit to be examined in Chapter 10. The most significant areal capacitance contributions are those associated with the source and drain portions of each MOSFET. Thus, the layout achieves compactness with minimal capacitance as follows:

- The minimum active width and length for the n-channel devices are 6λ and 14λ, respectively (see Fig. I18).
- The active regions for the p-channel devices have the same length as those for the n-channel devices, but width scales by a factor of 2.5 —the reason awaits discussion in Chapter 10. Widths of 15λ result.
- The n and p active regions in separate inverters have 3-λ spacings (rule 2.2). The n and p active regions within a single inverter have 5-λ spacings to the n-well boundary (rule 2.3).

Further considerations apply to the integrated-circuit wiring:

- The MOSFET gates connect with a single poly film, an acceptable practice over short distances. Note the 2-λ gate extensions (rule 3.3) and a $6\lambda \times 6\lambda$ poly abutment to allow for an input contact.
- The MOSFET drains connect to the gate input of another inverter with Metal 1. Note the 1-λ overlaps for the various contacts.
- The horizontal V_{DD} and GND rails make appropriate connections to the MOSFET sources. The supply lines also connect to underlying n^+-to-n-well (V_{DD}) or p^+-to-p-substrate (GND) active regions with 6-λ width and 4-λ spacings to opposite-type active regions (rule 2.5). Note that the select rule 4.1 does not apply to the poly that extends from a device-containing active region.
- Metal 2 (not shown) may be needed in support of circuit connections that provide specific functionality. The extent of the overall circuit determines the perimeter of the n-well boundary.

The inverters in Fig. I22 measure 17λ by 46λ. Heroic efforts to achieve less area often waste design time with meager capacitive savings.

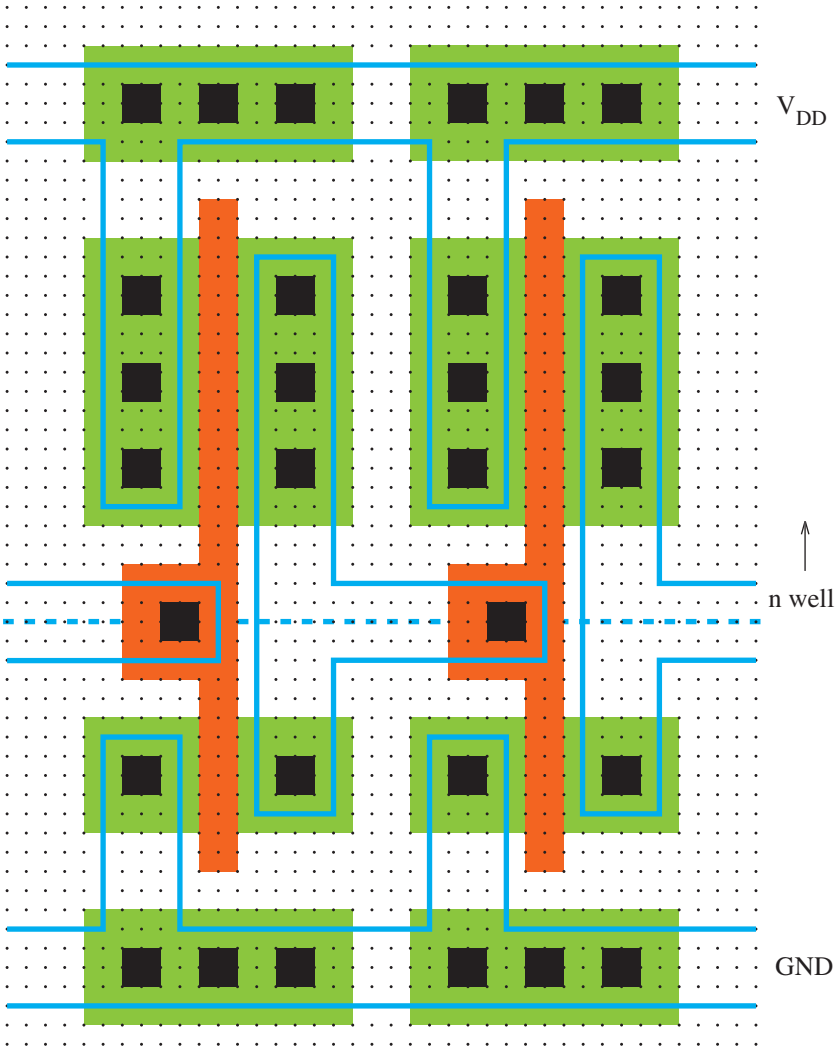


Figure I22: Coupled CMOS inverter layout. The grid spacing is 1 λ.

Capacitance Extraction

The MOSFET capacitances that typically dominate a particular layout are those that couple the source or drain regions to the body (C_{bs} and C_{bd}). Both capacitances feature bias-dependent area and perimeter components. Figure I23 provides a portion of the CMOS inverter layout shown in Fig. I22. For the particular MOSFET, the source and drain each measure $15\lambda \times 6\lambda$, the areas are $90\lambda^2$, and the perimeters are 42λ. With $W = 15\lambda$ and $L = 2\lambda$, and letting $\lambda = 0.5\ \mu$m, the MOSFET SPICE description takes the form:

```
M1    3       2       1       0       MOSFET   W=7.5u   L=1u
+     AS=22.5p        PS=21u          AD=22.5p PD=21u
```

Area and perimeter values are specified in units of m^2 and m, respectively. The corresponding body capacitances are

$$C_{bs} = CJ \cdot AS \left(1 - \frac{v_{bs}}{PB}\right)^{-MJ} + CJSW \cdot PS \left(1 - \frac{v_{bs}}{PB}\right)^{-MJSW} \qquad \text{(I1)}$$

and

$$C_{bd} = CJ \cdot AD \left(1 - \frac{v_{bd}}{PB}\right)^{-MJ} + CJSW \cdot PD \left(1 - \frac{v_{bd}}{PB}\right)^{-MJSW} . \qquad \text{(I2)}$$

SPICE parameters such as CJ and CJSW appear in the .model statement.

MOSFET channel dimensions W and L are required together with the gate-oxide thickness t_{ox} and capacitance parameters C_{gso}, C_{gdo}, and C_{gbo} to determine capacitances C_{gs}, C_{gd}, and C_{gb} as specified in Chapter 5.

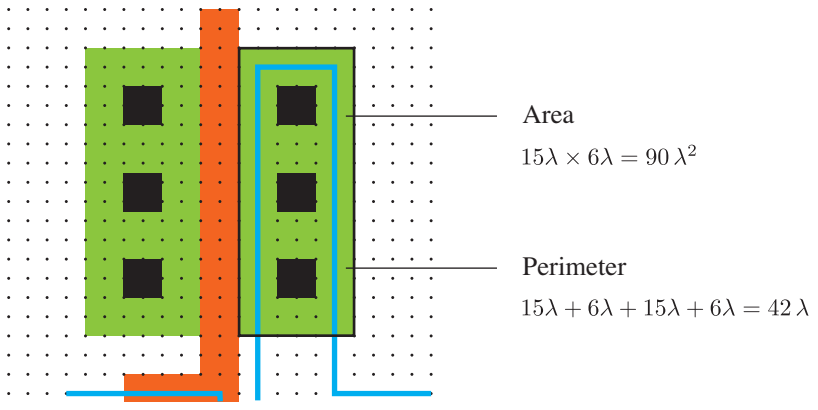


Figure I23: Area and perimeter taken from part of a CMOS inverter layout. The grid spacing is 1 λ.

Apart from the capacitances localized to MOSFETs (or other devices), integrated circuits feature parasitic capacitance associated with polysilicon or metal interconnections. Figure I24 shows a cross section of a conductor with width w and thickness t that traverses a field region with thickness h. Capacitance between the conductor and the ground plane has two parts: An "area" contribution is the parallel-plate capacitance given by

$$C_{area} = \frac{\epsilon_{ox}}{h}\, w\, l\,, \tag{I3}$$

where ϵ_{ox} is the SiO_2 dielectric permittivity and l is the interconnect length. An additive "perimeter" contribution reflects fringing fields that emanate from the sides and top of the conductor. Specifically,

$$C_{perimeter} = C(h, t, w)\, l\,, \tag{I4}$$

where $C(h, t, w)$ is some complicated function of the interconnect geometry. This contribution is significant when $w \sim h$. Integrated circuit processes typically have data for the calculation of both capacitive contributions for polysilicon, metal 1, and metal 2.

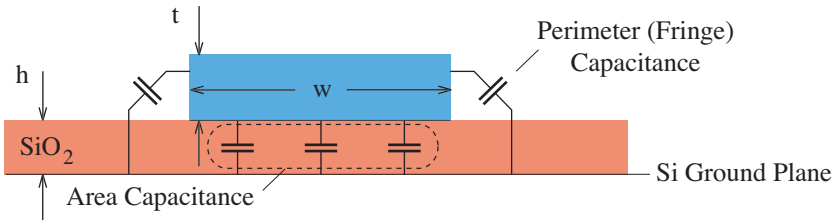


Figure I24: Parallel-plate and fringing capacitances to a ground plane.

A second form of parasitic capacitance, which is often difficult to model, is that between separate interconnects (polysilicon, metal 1, or metal 2). Figure I25 shows an example. The capacitance depends on the relative orientation of the interconnects, and it is concentrated at crossing points.

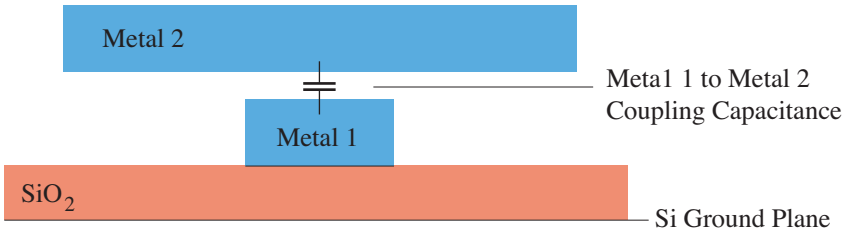
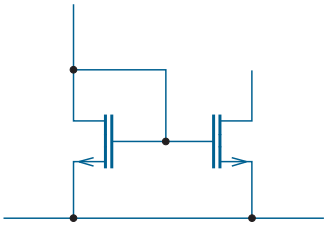


Figure I25: Capacitances between different types of interconnect.

Interconnect capacitance is important for the calculation of signal delays between far-removed portions of an integrated circuit. Such calculations and the design implications that follow go beyond the scope of this text.



Chapter 9

Gray Boxes: Inside the Op-Amp

Having witnessed the extensive scope of op-amp applications in Chapter 1, we are eager to apply subsequently acquired electronic principles so that black-box rules become gray in relation to specific internal op-amp circuity. The chapter begins with three forms of analysis for the differential amplifier, a circuit that serves at the input of every op-amp. Then after a discussion of complicating factors that are common for integrated MOSFETs and BJTs, we gradually develop circuit modifications for improved differential gain and output drive capability. Throughout, we emphasize small-signal transistor behavior in modern integrated circuits that contain few if any resistors.

Objectives

After studying this chapter and working through the problems, you should be able to perform each of the following:

- Calculate the differential and common-mode gains for an elementary differential amplifier (Section 9.1).
- Apply half-circuit analysis to a symmetric circuit (Section 9.1).
- Describe channel-length modulation and body effect for a MOSFET, and describe base-width modulation for a BJT (Section 9.2).
- Determine small-signal relationships subject to finite r_o (Section 9.2).
- Design a differential amplifier with a current-mirror load (Section 9.3).
- Design a current source with supply independence (Section 9.3).
- Design a simple two-stage CMOS op-amp (Section 9.4).

9.1 The Differential Amplifier

Our quest for an appropriate op-amp "front end" begins with the somewhat abstract circuit of Fig. 9.1 in which two inputs are amplified separately to produce outputs $v_{o1} = A_{v1}\, v_{i1}$ and $v_{o2} = A_{v2}\, v_{i2}$. The composite output v_{od} derives from the difference between between the v_{o1} and v_{o2} node voltages. Thus, we have

$$v_{od} = A_{v1}\, v_{i1} - A_{v2}\, v_{i2}\ . \tag{9.1}$$

In what follows, we will find it convenient to express the composite output in terms of a differential signal $v_{id} = v_{i1} - v_{i2}$ and a common-mode (average) signal $v_{ic} = (v_{i1} + v_{i2})/2$. Specifically,

$$v_{od} = A_{dd}\,(v_{i1} - v_{i2}) + A_{cd}\left(\frac{v_{i1} + v_{i2}}{2}\right)\ , \tag{9.2}$$

where A_{dd} is the **differential voltage gain** and A_{cd} is the **common-to-differential-mode voltage gain** . The factors that multiply v_{i1} and v_{i2} in Eqs. 9.1 and 9.2 are equivalent. In turn,

$$A_{dd} = \frac{A_{v1} + A_{v2}}{2} \tag{9.3}$$

and

$$A_{cd} = A_{v1} - A_{v2}\ . \tag{9.4}$$

A good **differential amplifier** produces the same output voltage regardless of the common-mode input component. For example, $v_{i1} = 1.001$ V and $v_{i2} = 0.999$ V have the same influence as $v_{i1} = 5.001$ V and $v_{i2} = 4.999$ V even though the averages of the inputs differ by 4 V. This requires $A_{cd} = 0$ through the action of a *symmetric* circuit with $A_{v1} = A_{v2}$.

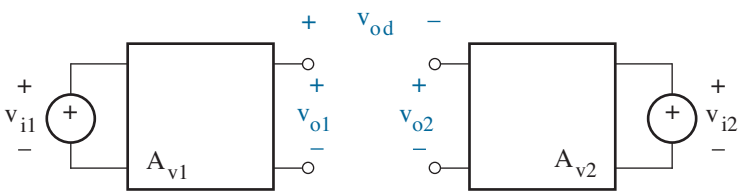


Figure 9.1: Two-input amplifier circuit with composite output.

With the basic amplifier circuits of Chapter 7 in hand, it is tempting to construct a differential amplifier from common-source (common-emitter) amplifiers that are symmetrically arranged as shown in Fig. 9.2.

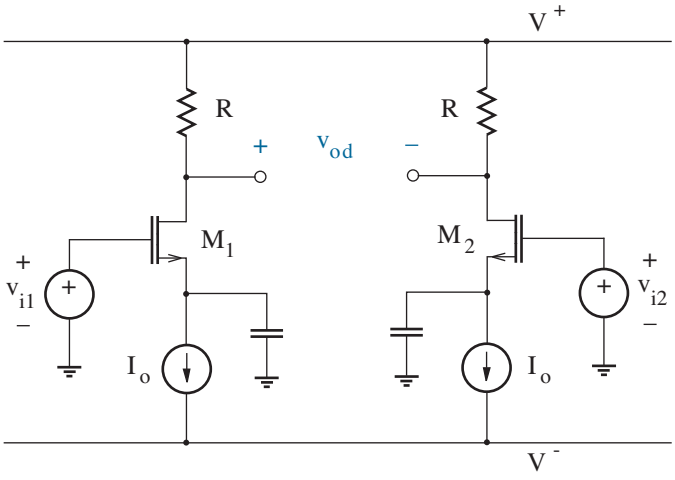


Figure 9.2: Differential amplifier with symmetric common-source circuits.

But the circuit of Fig. 9.2 is unacceptable—the capacitors become open circuits in the dc limit to ruin the differential gain for op-amp functionality. The easy fix is to form a cross-connected **source-coupled pair** as shown in Fig. 9.3 so that each source "sees" a low-resistance ac path to ground when looking into the opposite source. Capacitors are no longer needed.

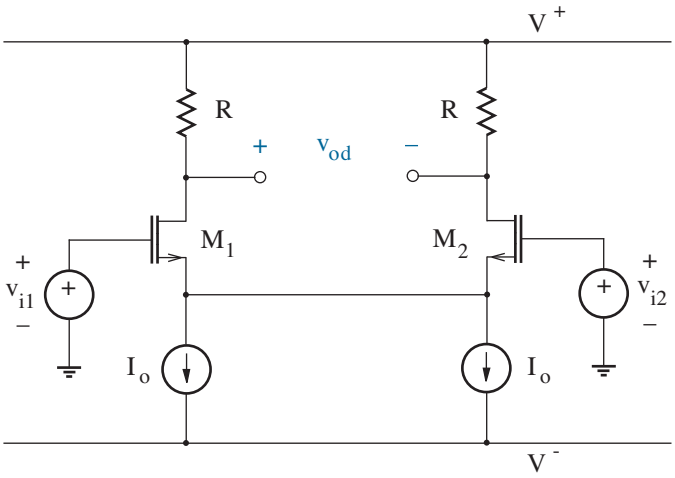


Figure 9.3: Revised differential amplifier with a source-coupled pair.

Differential-Mode Analysis

An immediate task is to find an expression for the differential gain that is applicable to the general circuit of Fig. 9.4, and we do so in three ways. As in Chapter 7, the results for the BJT circuit extend to MOSFETs and other field-effect devices through appropriate limits ($\beta_o \to \infty$, $r_\pi \to \infty$). The current source at the bottom of the circuit is assumed to be non-ideal with finite small-signal Norton (shunt) resistance r_x. Section 9.2 explores just how this comes about.

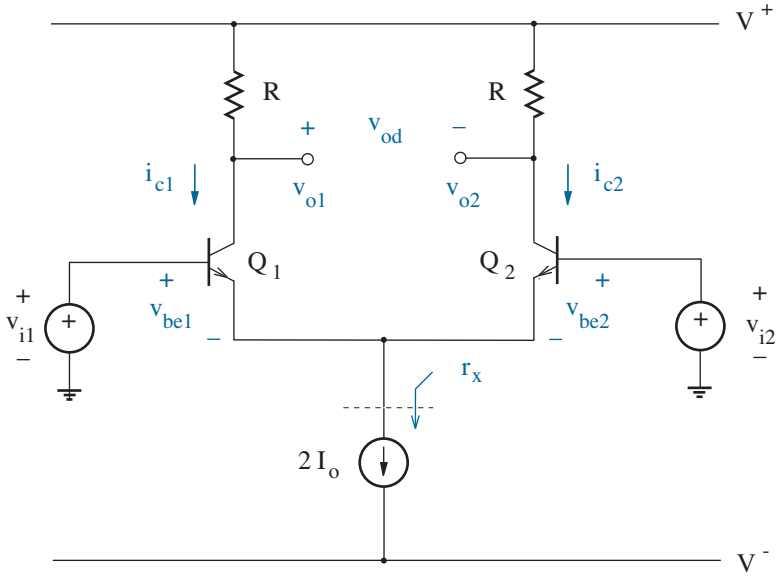


Figure 9.4: Differential amplifier with an emitter-coupled pair for analysis. The ac resistance looking into the current source is r_x.

Before we proceed, we note that the BJTs in the differential amplifier are matched in every respect—a straightforward outcome in integrated circuits, a reflection of tedious measurements and sorting in older discrete circuits. Thus, by symmetry

$$i_{c1}|_Q = i_{c2}|_Q = I_o \, . \tag{9.5}$$

In turn,

$$g_{m1} = g_{m2} = g_m \, . \tag{9.6}$$

And with $\beta_{o1} = \beta_{o2} = \beta_o$,

$$r_{\pi 1} = r_{\pi 2} = r_\pi \, . \tag{9.7}$$

Subscript distinctions for small-signal parameters are no longer needed.

Method 1

A first analytical method tests our understanding of the three basic amplifiers of Chapter 7. The change in v_{o1} due to small-signal changes in v_{i1} alone ($v_{i2} = 0$) reflects the voltage gain of a common-emitter amplifier:

$$\Delta v_{o1} = \frac{-g_m R}{1 + g_m {R_{e1}}'(1 + 1/\beta_o)} \Delta v_{i1} \,, \tag{9.8}$$

where ${R_{e1}}'$ is the small-signal resistance looking away from the Q_1 emitter. Meanwhile, the change in v_{o1} due to changes in v_{i2} alone ($v_{i1} = 0$) reflects the voltage gain of a two-stage amplifier in which a common-base circuit follows an effective common-collector circuit:

$$\Delta v_{o1} = \frac{g_m {R_{e2}}'(1 + 1/\beta_o)}{1 + g_m {R_{e2}}'(1 + 1/\beta_o)} (g_m R) \Delta v_{i2} \,, \tag{9.9}$$

where ${R_{e2}}'$ is the small-signal resistance looking away from the Q_2 emitter. By symmetry, and in consideration of the small-signal resistance looking into the Q_1 or Q_2 emitter,

$${R_{e1}}' = {R_{e2}}' = {R_e}' = \frac{1}{g_m(1 + 1/\beta_o)} \parallel r_x \,. \tag{9.10}$$

Thus with superposition,

$$\Delta v_{o1} = -g_m R \left(\frac{1}{1 + \zeta} \right) \Delta v_{i1} + g_m R \left(\frac{\zeta}{1 + \zeta} \right) \Delta v_{i2} \,, \tag{9.11}$$

where

$$\zeta = g_m {R_e}'(1 + 1/\beta_o) \,. \tag{9.12}$$

In the limit as $r_x \to \infty$ (an ideal current source), $\zeta \to 1$. Generally, $\zeta < 1$. Symmetry conditions dictate the response on the other side of the circuit:

$$\Delta v_{o2} = +g_m R \left(\frac{\zeta}{1 + \zeta} \right) \Delta v_{i1} - g_m R \left(\frac{1}{1 + \zeta} \right) \Delta v_{i2} \,. \tag{9.13}$$

So for the differential gain,

$$A_{dd} = \frac{\Delta v_{o1} - \Delta v_{o2}}{\Delta v_{i1} - \Delta v_{i2}} = -g_m R \left(\frac{1 + \zeta}{1 + \zeta} \right) = -g_m R \,. \tag{9.14}$$

Whereas the ζ factor drops out of Eq. 9.14, we conclude that a non-ideal current source does not adversely affect the ideal differential gain. Indeed, the current source could be replaced with a resistor.

MOSFET circuits yield the same result but with different g_m.

Method 2

The second analytical method traces large-signal voltages and currents. The node voltage at the Q_1 emitter is $v_{i1} - v_{be1}$ and the node voltage at the Q_2 emitter is $v_{i2} - v_{be2}$. These reflect a shared node, so

$$v_{i1} - v_{be1} = v_{i2} - v_{be2} \, . \tag{9.15}$$

And after rearrangement,

$$v_{be1} - v_{be2} = v_{i1} - v_{i2} = v_{id} \, . \tag{9.16}$$

Now with $v_{be1} \approx kT/q \, \ln(i_{c1}/I_{s1})$, $v_{be2} \approx kT/q \, \ln(i_{c2}/I_{s2})$, and $I_{s1} = I_{s2}$, the difference of the base-to-emitter voltages is

$$v_{be1} - v_{be2} = \frac{kT}{q} \ln\left(\frac{i_{c1}}{i_{c2}}\right) \, . \tag{9.17}$$

It follows that

$$\frac{i_{c1}}{i_{c2}} = x \, , \tag{9.18}$$

where $x = \exp(qv_{id}/kT)$. Meanwhile, from KCL at the emitter node,

$$i_{c1} + i_{c2} \approx 2I_o \tag{9.19}$$

(subject to large $\beta_{F1} = \beta_{F2}$ for nearly equal collector and emitter currents). The simultaneous collector-current solutions to Eqs. 9.18 and 9.19 are

$$i_{c1} = \frac{2I_o}{1 + 1/x} \tag{9.20}$$

and

$$i_{c2} = \frac{2I_o}{1 + x} \, . \tag{9.21}$$

Thus with

$$v_{od} = (V^+ - i_{c1}R) - (V^+ - i_{c2}R) = -(i_{c1} - i_{c2})R \tag{9.22}$$

(the difference of two node voltages), a little algebra reveals

$$v_{od} = -2I_oR\left(\frac{x-1}{x+1}\right) \, . \tag{9.23}$$

Finally, we substitute for x and simplify to obtain

$$v_{od} = -2I_oR \tanh\left(\frac{qv_{id}}{2kT}\right) \, . \tag{9.24}$$

In this expression, tanh( ) represents the hyperbolic tangent function.

Figure 9.5 shows the transfer characteristic that relates v_{od} and v_{id}. For $v_{id} \gg 2kT/q$ and $v_{id} \ll -2kT/q$, the hyperbolic tangent function approaches +1 and -1, respectively. Corresponding v_{od} values are $\mp 2I_oR$. This assumes, of course, that neither Q_1 or Q_2 are allowed to operate in the saturation mode. In turn, $|v_{od}|_{\max} \approx V^+ + 0.5$ V (see Problem 9.1).

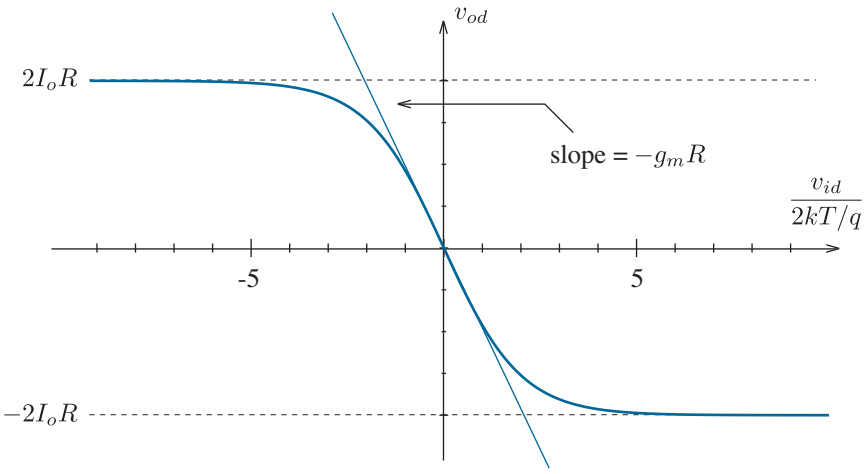


Figure 9.5: Transfer characteristic for the differential amplifier of Fig. 9.4.

No doubt most readers are keenly aware that the derivative of $\tanh(x)$ is $\text{sech}^2(x)$ (the square of the hyperbolic secant function) and $\text{sech}(0) = 1$. Thus, the small-signal differential gain in the vicinity of $v_{id} = 0$ is

$$A_{dd} = \left.\frac{dv_{od}}{dv_{id}}\right|_0 = -2I_oR\left(\frac{q}{2kT}\right)\text{sech}^2\left(\frac{qv_{id}}{2kT}\right)\Big|_0 = \frac{-I_oR}{kT/q}\,. \tag{9.25}$$

The quiescent collector current I_o divided by kT/q is tranconductance g_m. So we have

$$A_{dd} = -g_mR \tag{9.26}$$

in agreement with the result of Method 1.

The preceding analysis does not readily apply to a differential amplifier containing MOSFETs with a square-law relationship between i_d and v_{gs}. We leave it to Problem 9.2 to show that over a certain range of R

$$v_{od} = -g_mR\,v_{id}\sqrt{1 - \frac{{v_{id}}^2}{4\left(v_{gs}|_Q - V_T\right)^2}} \tag{9.27}$$

is the comparable transfer characteristic with small-signal $A_{dd} = -g_mR$. The limiting v_{od} values are $\mp I_oR$, but they are more gradually achieved.

Method 3

The third analytical method will prove to be easiest with a new concept. Given the definitions of both differential and common-mode input signals, we can express the left- and right-side inputs as follows:

$$v_{i1} = \frac{+v_{id}}{2} + v_{ic}\,, \tag{9.28}$$

and

$$v_{i2} = \frac{-v_{id}}{2} + v_{ic}\,. \tag{9.29}$$

Similarly, for the left- and right-side output node voltages (see Fig. 9.1),

$$v_{o1} = \frac{+v_{od}}{2} + v_{oc}\,, \tag{9.30}$$

and

$$v_{o2} = \frac{-v_{od}}{2} + v_{oc}\,. \tag{9.31}$$

Here, v_{od} and v_{oc} are the differential and common-mode output signals. The former has been considered previously. The latter averages v_{o1} and v_{o2}. If we restrict our attention to purely differential signals, $v_{ic} = 0$, $v_{oc} = 0$, and the circuit of Fig. 9.4 takes the form of Fig. 9.6.

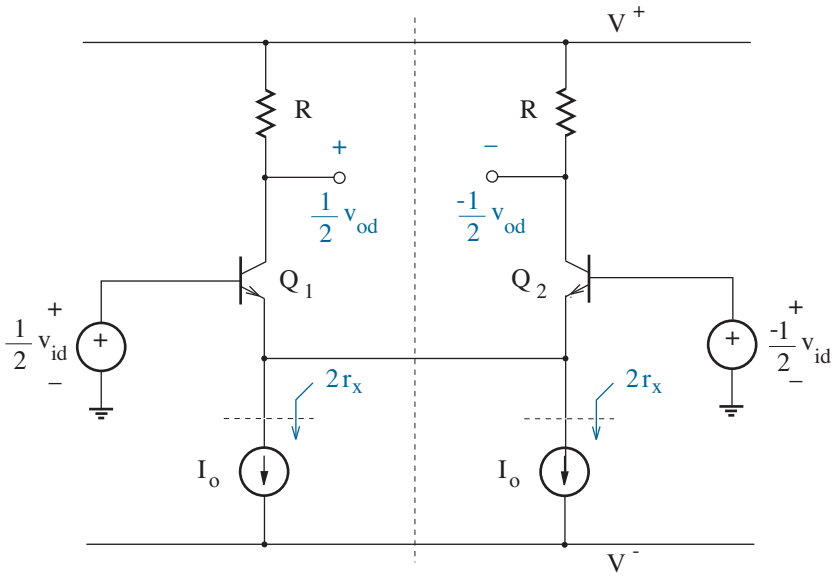


Figure 9.6: Differential amplifier with input and output assignments for differential signal analysis. Note how the $2I_o$ current source has been split.

The symmetry of Fig. 9.6 suggests that any voltage or current that increases in the left half will decrease by the same extent in the right half. Thus, we can get by with half of the analytical effort—actually even less—by constructing an effective **differential half circuit** as shown in Fig. 9.7. In doing so, we apply the following rule:

> *Nodes that are common to symmetric differential half circuits function as effective ac grounds.*

The validity of this rule should be obvious. In a handwaving proof, voltage increases and decreases induced by the opposing half circuits tend to cancel. (See Problem 9.3 if you remain unconvinced.)

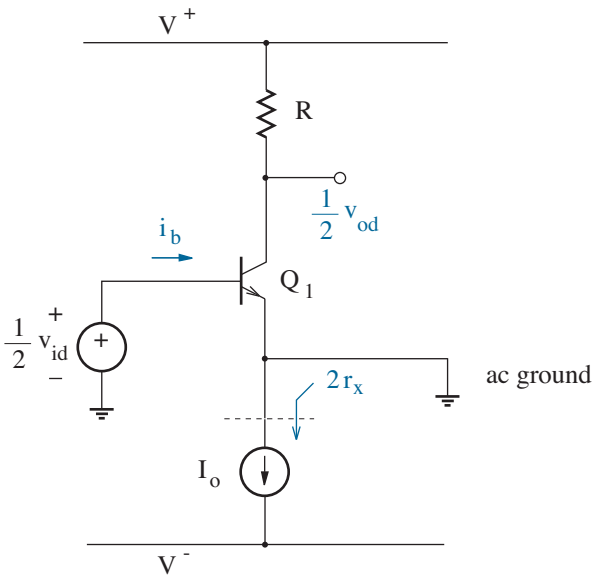


Figure 9.7: Differential half circuit for analysis. Note the ac ground.

As a consequence of the ac ground at the Q_1 emitter, we have a simple common-emitter amplifier for which

$$A_{dd} = \frac{v_{od}/2}{v_{id}/2} = -g_m R\,. \tag{9.32}$$

After catching our breath from the rigors of this analysis, we can calculate the applicable **differential input resistance**. With $(v_{id}/2)/i_b = r_b{}' = r_\pi$,

$$\frac{v_{id}}{i_b} = r_{id} = 2\,r_\pi\,. \tag{9.33}$$

The r_{id} value is infinite for source-coupled MOSFETs.

Exercise 9.1 Use a half circuit to find A_{dd} for the amplifier of Fig. 9.8.

Ans: $A_{dd} = -18$

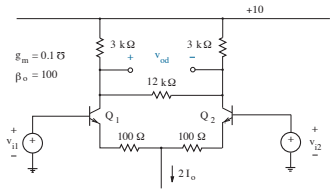


Figure 9.8: Circuit for Exercise 9.1.

Exercise 9.2 Use a half circuit to find A_{dd} for the amplifier of Fig. 9.9.

Ans: $A_{dd} = 240$

Common-Mode Analysis

If we restrict our attention to common-mode signals, Eqs. 9.28 and 9.29 tell us to apply a v_{ic} input source at both inputs to a differential amplifier, and Eqs. 9.30 and 9.31 tell us to observe a v_{oc} response at both outputs. Circuit symmetry suggests that any voltage or current changes by the same extent in each **common-mode half circuit**. So we have the following rule:

> *Passive elements that connect symmetrically positioned nodes in common-mode half circuits function as effective open circuits.*

We apply this rule to reduce the differential amplifier of Fig. 9.4 to the half-circuit shown in Fig. 9.10. The other half circuit is redundant.

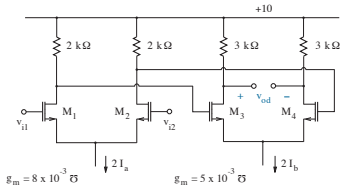


Figure 9.9: Circuit for Exercise 9.2.

As for differential analysis, we have a simple common-emitter amplifier. But with $R_e{}' = 2\,r_x$, the **common-mode gain** is

$$A_{cc} = \frac{v_{oc}}{v_{ic}} = \frac{-g_m R}{1 + 2\,g_m r_x (1 + 1/\beta_o)}\,. \tag{9.34}$$

Next, we can calculate the applicable **common-mode input resistance**. Specifically,

$$\frac{v_{ic}}{i_b} = r_{ic} = r_\pi + (1 + \beta_o)\,2\,r_x\,. \tag{9.35}$$

The r_{ic} value is infinite for source-coupled MOSFETs.

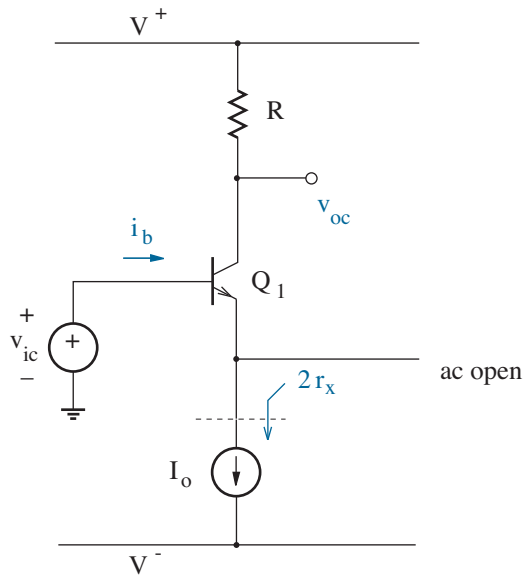


Figure 9.10: Common-mode half circuit for analysis. Note the ac open.

Exercise 9.3 Use a half circuit to find A_{cc} for the amplifier of Fig. 9.8. Assume $r_x = 50$ kΩ.

Ans: $A_{cc} = -2.97 \times 10^{-2}$

Exercise 9.4 Use a half circuit to find A_{cc} for the amplifier of Fig. 9.9. Assume $r_x = 25$ kΩ.

Ans: $A_{cc} = 2.38 \times 10^{-3}$

When differential and common-mode signals act in concert, they follow the matrix relation

$$\begin{pmatrix} v_{od} \\ v_{oc} \end{pmatrix} = \begin{pmatrix} A_{dd} & A_{cd} \\ A_{dc} & A_{cc} \end{pmatrix} \begin{pmatrix} v_{id} \\ v_{ic} \end{pmatrix} . \tag{9.36}$$

Like A_{cd}, the A_{dc} **differential-to-common-mode voltage gain** is zero in circuits with left-right symmetry. The A_{dd} differential gain is often large, and the A_{cc} common-mode gain tends to be very small.

Circuit designers use the **common-mode rejection ratio** or **CMRR** as a figure of merit that indicates the degree to which common-mode input

signals are suppressed at a differential amplifier output. Specifically,

$$\text{CMRR} = \left| \frac{A_{dd}}{A_{cd}} \right| . \tag{9.37}$$

CMRR values are typically expressed in dB ($20 \log |A_{dd}/A_{cd}|$). The CMRR is infinite for a symmetric amplifier with a differential output as in Fig. 9.4. Nevertheless, small transistor mismatches generally lead to finite CMRR.

Reassigning the differential amplifier output voltage changes everything. For example, a circuit that is to function as an operational amplifier has a pair of differential inputs and a single-ended output (one side at ground). A cavalier approach takes the output at node v_{o1} in Fig. 9.4 to yield

$$A_{dd} = \frac{-g_m R}{2} \tag{9.38}$$

—the half circuit producing half of the former differential gain is lost—and

$$A_{cd} = \frac{-g_m R}{1 + 2\, g_m r_x (1 + 1/\beta_o)} \tag{9.39}$$

—the half circuit producing v_{oc} also accounts for v_{o1}. In turn,

$$\text{CMRR} = 20 \log \left[\frac{1}{2} + g_m r_x (1 + 1/\beta_o) \right] . \tag{9.40}$$

This may be unacceptable if r_x is not very large.

Asymmetry Effects

All of the preceding analysis assumed a differential amplifier with perfect left-right symmetry. When symmetry fails, we have unfortunate results.

Consider the BJT differential amplifier of Fig. 9.4 with $R_1 = R + \Delta R/2$ on the left and $R_2 = R - \Delta R/2$ on the right. Left-right symmetry is lost, so half circuits no longer apply. We look to the complete circuit to find

$$v_{o1} = V^+ - i_{c1}\left(R + \frac{\Delta R}{2}\right) \tag{9.41}$$

and

$$v_{o2} = V^+ - i_{c2}\left(R - \frac{\Delta R}{2}\right) . \tag{9.42}$$

Then with $v_{i1} = v_{i2} = 0$, the BJT collector currents remain equally divided ($i_{c1} = i_{c2} = I_o$) so that

$$v_{od} = v_{o1} - v_{o2} = -I_o\,\Delta R\,. \tag{9.43}$$

We define the **input offset voltage** as the differential input that produces the differential output of Eq. 9.43 under *symmetric* circuit conditions. In turn,

$$v_{os} = \frac{v_{od}}{A_{dd}} = \frac{-I_o\,\Delta R}{-g_m R} . \tag{9.44}$$

And subject to $g_m = I_o/(kT/q)$,

$$v_{os} = \frac{kT}{q}\left(\frac{\Delta R}{R}\right) . \tag{9.45}$$

Equation 9.45 shows that a 1-% change in R is consistent with an input offset voltage of 0.259 mV at room temperature.

Now consider a separate form of asymmetry in which $I_{s1} = I_s + \Delta I_s/2$ on the left and $I_{s2} = I_s - \Delta I_s/2$ on the right. In this case, $v_{i1} = v_{i2} = 0$ requires $v_{be1} = v_{be2}$ or

$$\frac{kT}{q}\ln\left(\frac{i_{c1}}{I_s + \Delta I_s/2}\right) = \frac{kT}{q}\ln\left(\frac{i_{c2}}{I_s - \Delta I_s/2}\right) . \tag{9.46}$$

The logarithmic arguments are necessarily equal and $i_{c1} + i_{c2} = 2I_o$. Thus,

$$i_{c1} = I_o\left(1 - \frac{\Delta I_s}{2I_s}\right) \tag{9.47}$$

and

$$i_{c2} = I_o\left(1 + \frac{\Delta I_s}{2I_s}\right) . \tag{9.48}$$

Subject to equal values for R, we apply Eq. 9.22 to find

$$v_{od} = -I_o R \frac{\Delta I_s}{I_s} \,. \tag{9.49}$$

So with $A_{dd} = -g_m R$ and $g_m = I_o/(kT/q)$, the input offset voltage is

$$v_{os} = \frac{v_{od}}{A_{dd}} = \frac{kT}{q}\left(\frac{\Delta I_s}{I_s}\right) \,. \tag{9.50}$$

Independent effects arising from asymmetry in R and I_s combine separately. When circuit parameters are governed by statistics, we use the relation

$$\Delta v_{os} \approx \frac{kT}{q}\sqrt{\left(\frac{\Delta R}{R}\right)^2 + \left(\frac{\Delta I_s}{I_s}\right)^2} \,, \tag{9.51}$$

where Δv_{os}, ΔR, and ΔI_s represent standard deviations.

The input offset voltage for the asymmetric differential amplifier with MOSFETs derives in a similar fashion, but the algebra is somewhat tedious. We leave it to the end-of-chapter problems to demonstrate three separate contributions that combine to yield the statistical estimate

$$\Delta v_{os} \approx \frac{v_{gs} - V_T}{2}\sqrt{\left(\frac{\Delta R}{R}\right)^2 + \left(\frac{\Delta(W/L)}{W/L}\right)^2 + \left(\frac{2\Delta V_T}{v_{gs} - V_T}\right)^2} \,. \tag{9.52}$$

Whereas $v_{gs} - V_T$ is typically much larger than kT/q, the input offset voltage for a MOSFET amplifier is far more significant that that for a BJT amplifier given comparable ΔR, for example. Nevertheless, the BJT input offset voltage tends to be pernicious as it varies *linearly* with temperature. The MOSFET input offset voltage has weaker temperature dependence.

Problem 9.14 explores **input offset current** observed in asymmetric BJT differential amplifiers. There is no MOSFET-related counterpart.

Exercise 9.5 A BJT differential amplifier has $I_{s1} = 10\text{f}$ and $I_{s2} = 10.5\text{f}$ for the transistors and $R_1 = 4.9\ \text{k}\Omega$ and $R_2 = 5.2\ \text{k}\Omega$ for the resistive loads. Determine the input offset voltage at room temperature.

Ans: $v_{os} = -1.26\ \text{mV} + 1.54\ \text{mV} = 0.28\ \text{mV}$

Exercise 9.6 A MOSFET differential amplifier subject to $2I_o = 2\ \text{mA}$ has $K_n{}' = 50\ \mu\text{A/V}^2$ and $V_T = 0.6\ \text{V}$ for the transistors. The aspect ratios are $W/L = 60/1.2 \pm 0.5\ \%$ and the resistive loads are $2.6\ \text{k}\Omega \pm 0.4\ \%$. Estimate the standard deviation of the input offset voltage.

Ans: $\Delta v_{os} = 1.43\ \text{mV}$

9.2 Analog Integrated-Circuit Complications

The preceding section puts us well along the path toward a modern op-amp. Nevertheless, we need to step back to consider some MOSFET and BJT complications that arise in integrated circuits having compact dimensions. Our reassessment of large- and small-signal terminal characteristics extends to new expressions for amplifier gain and ac "looking-around" resistances. New design opportunities await.

MOSFET Channel-Length Modulation

A major limitation of the elementary MOSFET analysis and the associated current-voltage relations is the assumption of a constant channel length L. In practice, the *effective* L decreases slightly to $L - \Delta L$ as the cutoff point X moves toward the source with increasing v_{ds} (Fig. 9.11). Whereas drain current is proportional to W/L, small $\Delta L/L$, implies

$$i_d \sim \frac{W}{L - \Delta L} \approx \frac{W}{L}\left(1 + \frac{\Delta L}{L}\right). \tag{9.53}$$

And if $\Delta L/L$ is proportional to v_{ds},

$$i_d \sim \frac{W}{L}\left(1 + \lambda\, v_{ds}\right). \tag{9.54}$$

Strictly speaking, this form of **channel-length modulation** is limited to the saturation mode. Nevertheless, drain-current continuity requires

$$i_d = \begin{cases} 0 & \text{for } v_{gs} - V_T \le 0\ ; \\ \frac{1}{2}K'\dfrac{W}{L}\left[2(v_{gs} - V_T)v_{ds} - {v_{ds}}^2\right](1 + \lambda\, v_{ds}) & \text{for } 0 < v_{ds} \le v_{gs} - V_T\ ; \\ \frac{1}{2}K'\dfrac{W}{L}(v_{gs} - V_T)^2\,(1 + \lambda\, v_{ds}) & \text{for } 0 < v_{gs} - V_T \le v_{ds}\ . \end{cases} \tag{9.55}$$

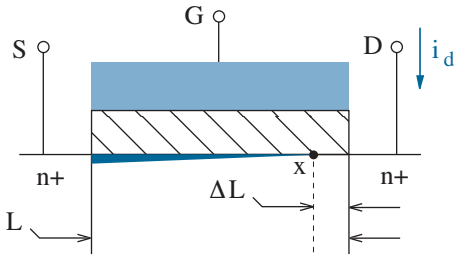


Figure 9.11: Channel length modulation. Cutoff point X varies with v_{ds}.

The most pronounced effect of non-zero λ is a tilting of the formerly flat characteristic curves for $v_{ds} > v_{ds,sat}$. A straightforward geometric analysis shows that each of the sloped curves can be extrapolated to a point on the v_{ds} axis with value $-1/\lambda$ as shown in Fig. 9.12.

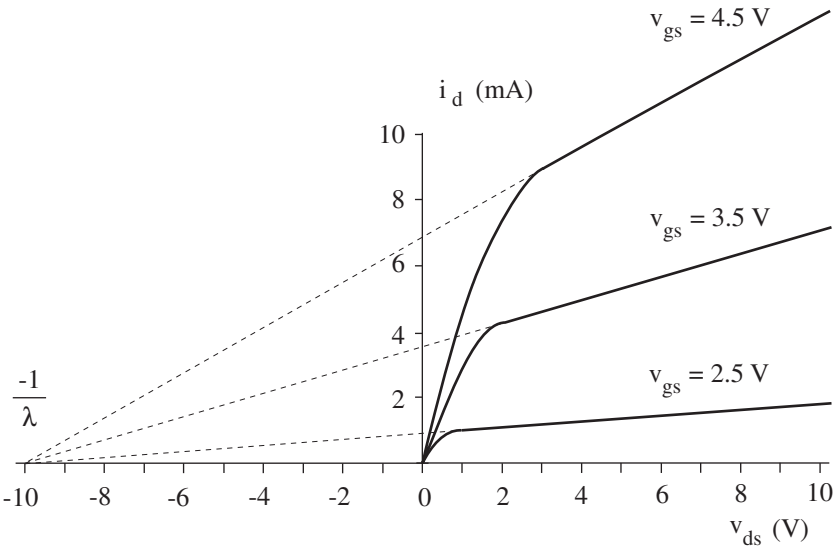


Figure 9.12: Set of three non-ideal characteristic curves for an n-channel enhancement-mode MOSFET with $K'\,W/L = 2$ mA/V^2, $V_T = 1.5$ V, and $\lambda = 0.1$ V^{-1}. The extrapolated curves intersect at $-1/\lambda = -10$ V.

To determine λ, we measure the slope of a drain-current characteristic curve at some point (i_d, v_{ds}) that lies well within the region of transistor saturation (see Fig. 9.13). In turn, with $\eta = (\Delta i_d/i_d)\,/\,(\Delta v_{ds}/v_{ds})$,

$$\lambda = \left(\frac{1}{v_{ds}}\right)\frac{\eta}{1-\eta}\,. \tag{9.56}$$

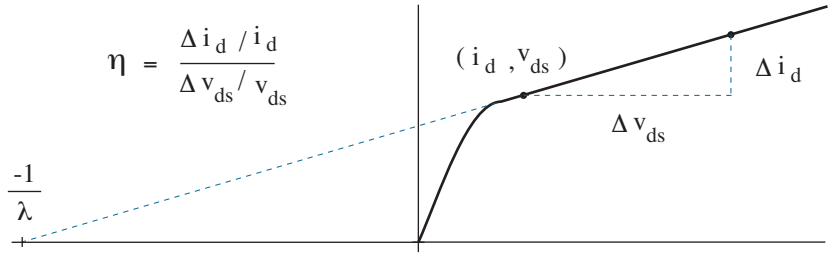


Figure 9.13: Measurement procedure for λ.

MOSFET Substrate Bias (Body) Effect

MOSFETs in integrated circuits are sometimes operated with non-zero v_{bs}. The body (fourth) MOSFET terminal is necessarily held at a potential that ensures reverse bias for both the source and drain pn junctions, so non-zero v_{bs} does not produce body-terminal current. Nevertheless, $v_{bs} \neq 0$ makes it more difficult to form an inversion layer—carriers hang around the source. In turn, the magnitude of the source-referenced threshold voltage increases. Specifically,

$$V_T \text{ (source)} = V_{T0} + \gamma(\sqrt{2\phi_f - v_{bs}} - \sqrt{2\phi_f}) \tag{9.57}$$

for n-channel devices ($2\phi_f > 0$), and

$$V_T \text{ (source)} = V_{T0} - \gamma(\sqrt{|2\phi_f| + v_{bs}} - \sqrt{|2\phi_f|}) \tag{9.58}$$

for p-channel devices ($2\phi_f < 0$). In these expressions, V_{T0} is the threshold voltage that is in effect when $v_{bs} = 0$, γ is the **body effect parameter**, and $2\phi_f$ is the **surface inversion potential**, the change in the MOSFET surface potential that leads to inversion when the flat-band voltage is zero. All are process dependent. (The Chapter-5 Appendix has physical details.)

Parameter γ is relatively difficult to determine, since the body effect is also conditioned by $2\phi_f$. If we are positioned to know the substrate doping, we can obtain $2\phi_f$ from the relation

$$2\phi_f = \left(\frac{kT}{q}\right) \ln \frac{N_a}{n_i} \,. \tag{9.59}$$

More likely, we will have to settle for $2\phi_F \approx 0.75$ V as a reasonable estimate. We then perform a series of threshold-voltage measurements using different values for v_{bs} In turn, we plot V_T vs. $(\sqrt{2\phi_f - v_{bs}} - \sqrt{2\phi_f})$ with the hope of obtaining a straight line through the data points. If the line is straight, the slope is γ. If the line is not straight, we must revise our $2\phi_f$ estimate. Figure 9.14 shows typical data.

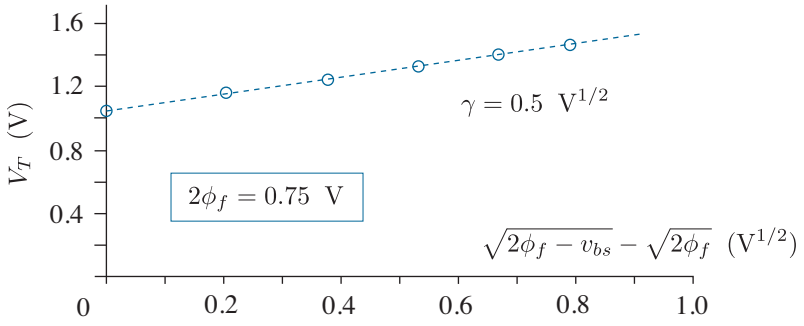


Figure 9.14: Measurement procedure for γ.

MOSFET Subthreshold Current

The modified elementary current-voltage relations of Eq. 9.55 demand that the MOSFET drain current steadily reduce to zero as v_{gs} decreases to V_T. In practice, the drain current is reduced to a very small but non-zero value. Further reduction proceeds according to the relation

$$i_d = I_s \exp\left[\frac{q(\phi_o - 2\phi_f)}{kT}\right]\left[1 - \exp\left(\frac{-qv_{ds}}{kT}\right)\right], \tag{9.60}$$

where

$$I_s = \frac{1}{2}K'\frac{W}{L}\gamma\left(\frac{kT}{q}\right)^{3/2}\left(\frac{2\phi_f}{kT/q}\right)^{-1/2}. \tag{9.61}$$

Here, ϕ_o is the surface potential at the semiconductor/insulator interface, $2\phi_f$ is the surface potential for inversion, and γ is the body-effect parameter. Equation 9.61 reflects the **subthreshold** mode of MOSFET operation.

We thus find that the subthreshold drain current is nearly independent of drain voltage for v_{ds} greater than a few kT/q. The mode of operation is like that of a "valve" with ϕ_o control over a tiny current, even at maximum. For example, with $K'W/L = 0.5$ mA/V^2, $\gamma = 0.6$ V$^{1/2}$, and $2\phi_f = 0.75$ V, the maximum subthreshold drain current is only 0.12 μA at $\phi_o = 2\phi_f$.

The actual control mechanism is by means of v_{gs}, which relates to ϕ_o. As demonstrated in the Chapter-5 Appendix, the relationship is almost linear for large arguments. So we perform a Taylor expansion to find

$$\phi_o - 2\phi_f \approx (\phi_o - 2\phi_f)|_{v_{gs}=V_T} + \left.\frac{\partial(\phi_o - 2\phi_f)}{\partial(v_{gs} - V_T)}\right|_{v_{gs}=V_T}(v_{gs} - V_T). \tag{9.62}$$

The first term in Eq. 9.62 is zero since $\phi_o = 2\phi_f$ at threshold. The second term is proportional to the factor $1/n$ given by

$$\frac{1}{n} = 1 - \left[1 + \frac{4(V_T - V_{fb})}{\gamma^2}\right]^{-1/2}, \tag{9.63}$$

where V_{fb} is the flat-band voltage. In turn, for large v_{ds},

$$i_d = I_s \exp\left[\frac{q(v_{gs} - V_T)}{nkT}\right]. \tag{9.64}$$

Parameter n is complicated by the characteristics of the Si/SiO$_2$ system. Thus, n (and for that matter, I_s) are measured quantities.

Some circuit designers (who are familiar with material in Chapter 7) become excited over the ramifications of exponential subthreshold control. Good designers outgrow this phase. Apart from the small currents involved, it is difficult to cope with V_T process variations of the order of kT/q.

Example 9.1

The MOSFET in the circuit of Fig. 9.15 features $K'W/L = 2.0$ mA/V^2 and $V_T = 1$ V. Determine voltage v_x subject to: (a) the information given; (b) $\lambda = 0.1$ V^{-1}; (c) $\lambda = 0.1$ V^{-1}, $\gamma = 0.5$ V$^{1/2}$, and $2\phi_f = 0.6$ V.

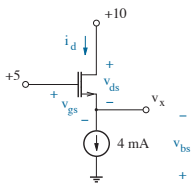


Figure 9.15: Circuit for Example 9.1.

Solution

(a) We assume saturation with $i_d = 4$ mA. Then

$$4 \text{ mA} = 1/2 \times 2 \text{ mA} \times (v_{gs} - 1)^2 .$$

There are two solutions: $v_{gs} = 3$ V and $v_{gs} = -1$ V. But the latter is invalid because it is less than V_T. So with $v_{gs} = 5 - v_x$, $v_x = 2$ V. The assumed saturation condition holds since $v_{ds} = 10 - v_x = 8 \text{ V} > v_{gs} - V_T = 2$ V.

(b) Whereas $v_{gs} = 5 - v_x$ and $v_{ds} = 10 - v_x$, we have

$$4 \text{ mA} = 1/2 \times 2 \text{ mA} \times (5 - v_x - 1)^2 \left[1 + 0.1(10 - v_x)\right] .$$

A trial-and-error solution yields $v_x = 2.49$ V. Saturation still applies.

(c) Here, we have

$$4 \text{ mA} = 1/2 \times 2 \text{ mA} \times (5 - v_x - V_T)^2 \left[1 + 0.1(10 - v_x)\right] .$$

And with $v_{bs} = -v_x$,

$$V_T = 1 + 0.5 \left(\sqrt{0.6 + v_x} - \sqrt{0.6}\right) .$$

Another trial-and-error solution yields $v_x = 2.08$ V. Saturation once again.

The calculations of parts (b) and (c) are typically reserved for SPICE.

BJT Base-Width Modulation

The BJT characteristic curves shown in Fig. 6.6 are ideal since they neglect the influence of **base-width modulation** when v_{ce} $(-v_{bc})$ is made large. Consider the BJT base profile in Fig. 9.16. The metallurgical base width between the base-emitter junction and the base-collector junction is w_b, and the neutral base width, which excludes the junction depletion regions, is $w = x_2 - x_1$. When the base-emitter junction is subject to forward bias, v_{be} is never much more than about 0.7 V, and x_1 is nearly fixed. In contrast, x_2 tends to decrease (move to the left) as v_{bc} is made arbitrarily negative. Thus, the neutral base width decreases as v_{ce} $(-v_{bc})$ increases.

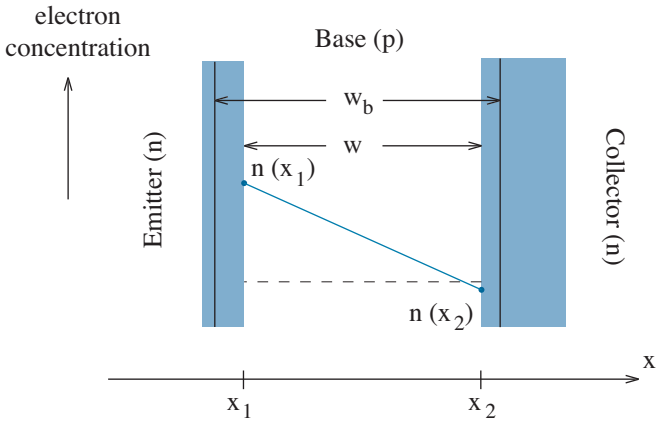


Figure 9.16: Close-up of the BJT (npn) base region showing metallurgical and neutral base widths under forward-active bias conditions. The junction depletion regions are shaded.

Electrons that are injected into the base cross by means of a diffusion process (since the electric field is negligible in the neutral base regions). The rate of diffusion is proportional to the electron concentration gradient, so for an approximate linear carrier profile we have

$$i_c \sim \frac{n(x_1) - n(x_2)}{w} \,. \tag{9.65}$$

A decrease in the neutral base width w leads to increased collector current. In turn, it is reasonable to expect a collector current of the form

$$i_c = i_{co}\,\frac{w_o}{w_o - \Delta w} \approx i_{co}\left(1 + \frac{\Delta w}{w_o}\right) \,, \tag{9.66}$$

where i_{co} and w_o correspond to forward-active conditions in the limit as $v_{ce} \to 0$. If we assume that $\Delta w / w_o$ is *roughly* proportional to v_{ce},

$$i_c \approx i_{co} \left(1 + \frac{v_{ce}}{V_A}\right) . \tag{9.67}$$

Thus, we have

$$i_c = \beta_F \, i_b \left(1 + \frac{v_{ce}}{V_A}\right) \tag{9.68}$$

so that the effective forward "Beta" is

$$\beta_F{}^{\text{effective}} = \beta_F \left(1 + \frac{v_{ce}}{V_A}\right) . \tag{9.69}$$

In the preceding equations, V_A is the **forward Early voltage** (named after James Early, who explained the phenomenon of base-width modulation). The V_A parameter typically exceeds 100 V for discrete BJTs.

The effect of finite V_A is a tilting of otherwise flat characteristic curves, and the curves intersect the v_{ce} axis at $-V_A$. Compare with the channel-length-modulated MOSFET curves of Fig. 9.12, which intersect at $-1/\lambda$. The V_A measurement procedure is similar to that shown in Fig. 9.13.

We also observe the effects of base-width modulation when the BJT is operated in the reverse active mode. However,

$$i_e = \beta_R \, i_b \left(1 + \frac{v_{ec}}{V_B}\right) , \tag{9.70}$$

and

$$\beta_R{}^{\text{effective}} = \beta_R \left(1 + \frac{v_{ec}}{V_B}\right) , \tag{9.71}$$

where V_B is the **reverse Early voltage**. This parameter is significantly less than V_A, since the large doping concentration in the emitter relative to that in the base results in greater near-junction depletion within the base. The base-side depletion region broadens substantially with increasing $-v_{be}$ —yet another strike against the reverse active mode.

Exercise 9.7 A BJT with $\beta_F = 120$ and $V_A = 40$ V operates in the forward active mode. Determine i_b and i_e if $i_c = 6$ mA and $v_{ce} = 25$ V.

Ans: $i_b = 30.8$ μA, $i_e = 6.03$ mA

Exercise 9.8 A particular BJT operates in the forward active mode. Determine V_A if the collector current increases by 5 % as the collector-emitter voltage increases from 10 V to 20 V.

Ans: $V_A = 190$ V

Revised Small-Signal Models

Second-order small-signal models for the three-terminal MOSFET ($v_{bs} = 0$) and the BJT feature a finite output resistance r_o as shown in Fig. 9.17. For the n-channel MOSFET,

$$r_o = \left(\left. \frac{\partial i_d}{\partial v_{ds}} \right|_Q \right)^{-1} = \frac{1/\lambda + v_{ds}|_Q}{i_d|_Q} \,. \tag{9.72}$$

And for the npn BJT,

$$r_o = \left(\left. \frac{\partial i_c}{\partial v_{ce}} \right|_Q \right)^{-1} = \frac{V_A + v_{ce}|_Q}{i_c|_Q} \,. \tag{9.73}$$

The n-channel MOSFET also has a revised transconductance:

$$g_m = \left. \frac{\partial i_d}{\partial v_{gs}} \right|_Q = \sqrt{2\,K' \frac{W}{L}\, i_d|_Q \,(1 + \lambda\, v_{ds}|_Q)} \,. \tag{9.74}$$

Similar parameter relations apply to p-channel MOSFETs or pnp BJTs, but the bias voltages and currents assume absolute values.

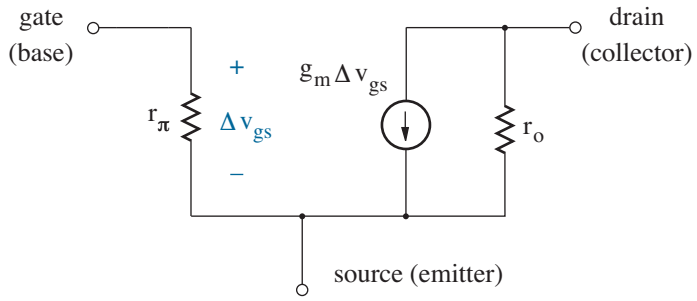


Figure 9.17: Three-terminal small-signal MOSFET and BJT model.

Readers who will be working with MOSFETs as discrete components (and even those relative few who will design MOSFET integrated circuits) can look forward to happy hours sifting through the practical ramifications of the three-terminal small-signal models. Nevertheless, the MOSFET has a fourth body or substrate terminal, which does not necessarily have the same potential as the source terminal. The body-to-source voltage shifts the n-channel threshold voltage through the relation

$$V_T = V_{T0} + \gamma \left(\sqrt{2\phi_f - v_{bs}} - \sqrt{2\phi_f} \right) , \tag{9.75}$$

where V_{T0} is the zero-bias threshold voltage, γ is the body-effect parameter, and $2\phi_F$ is the inversion potential. Thus, a particular Δv_{bs} induces Δi_d, thereby altering the small-signal circuit behavior.

To account for the body effect, the four-terminal small-signal MOSFET model includes a Δv_{bs}-dependent current source as shown in Fig. 9.18. Note the source orientation.

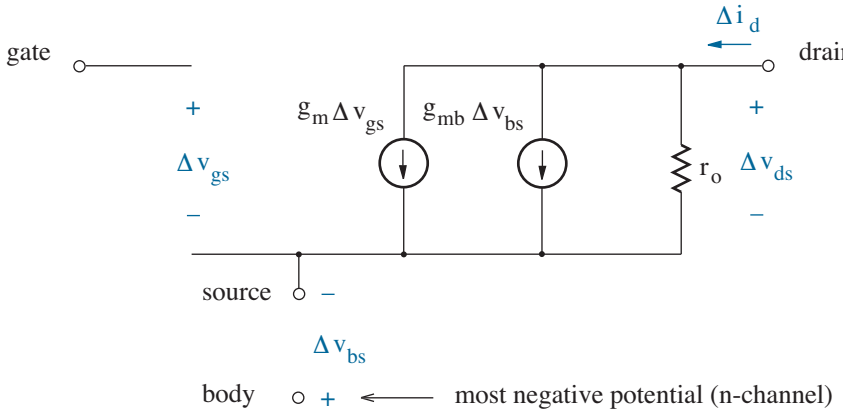


Figure 9.18: Four-terminal small-signal MOSFET model.

The differential drain current consistent with Fig. 9.18 has the form

$$\Delta i_d\,(v_{gs}, v_{ds}, v_{bs}) = g_m\,\Delta v_{gs} + \frac{1}{r_o}\,\Delta v_{ds} + g_{mb}\,\Delta v_{bs}\,. \tag{9.76}$$

Thus, the body transconductance parameter g_{mb} is given by

$$g_{mb} = \left.\frac{\partial i_d}{\partial v_{bs}}\right|_Q = \left.\frac{\partial i_d}{\partial V_T}\right|_Q \left.\frac{\partial V_T}{\partial v_{bs}}\right|_Q . \tag{9.77}$$

For the derivative with respect to V_T, we find

$$\left.\frac{\partial i_d}{\partial V_T}\right|_Q = -K'\,\frac{W}{L}\,(\,v_{gs}|_Q - V_T)\,(1 + \lambda\, v_{ds}|_Q) = -g_m\,. \tag{9.78}$$

And for the derivative with respect to v_{bs},

$$\left.\frac{\partial V_T}{\partial v_{bs}}\right|_Q = \frac{-\gamma\,/2}{\sqrt{2\phi_f - v_{bs}|_Q}}\,. \tag{9.79}$$

In turn,

$$g_{mb} = g_m \frac{\gamma\,/2}{\sqrt{2\phi_f - v_{bs}|_Q}} = g_m\,\chi\,. \tag{9.80}$$

The χ parameter is typically 0.15 or less.

Revised Small-Signal Amplifier Behavior

Table 9.1 presents revised expressions for the inherent voltage gain in each of the three basic amplifier configurations (see Problems 9.44 and 9.45). The expressions are *approximate* results that generally assume $r_o \gg 1/g_m$. The MOSFET body effect reduces gain in the common-source and common-drain configurations when the source terminal is not at ac ground.

Caution: We apply Table 9.1 judiciously—there are perils in precision.

Configuration	MOSFET	BJT
CS / CE	$\dfrac{-g_m(R_d' \parallel r_o)}{1+(g_m+g_{mb})R_s'/(1+R_d'/r_o)}$	$\dfrac{-g_m(R_c' \parallel r_o)}{1+g_mR_e'(1+1/\beta_o)/(1+R_c'/r_o)}$ $(R_e' \ll r_o)$
CD / CC	$\dfrac{g_mR_s'}{1+(g_m+g_{mb})R_s'}$	$\dfrac{g_mR_e'(1+1/\beta_o)}{1+g_mR_e'(1+1/\beta_o)}$
CG / CB	$g_m(R_d' \parallel r_o)$	$g_m(R_c' \parallel r_o)$

Table 9.1: Revised inherent voltage gains for various forms of amplifier.

Table 9.2 provides revised, *approximate* expressions for the small-signal resistances looking into various transistor terminals given finite r_o.

Resistance	MOSFET	BJT
r_d' / r_c'	$r_o\left[1+(g_m+g_{mb})R_s'\right]$	$r_o\left[1+\dfrac{g_mR_e'}{1+(R_b'+R_e')/r_\pi}\right]$
r_g' / r_b'	∞	$r_\pi+\dfrac{(1+\beta_o)R_e'}{1+(R_c'+R_e')/r_o}$
r_s' / r_e'	$\dfrac{1+R_d'/r_o}{g_m+g_{mb}}$	$\dfrac{r_\pi+R_b'}{1+\beta_o/(1+R_c'/r_o)}$

Table 9.2: Revised inward-looking small-signal transistor resistance values.

Table 9.2 comes to particular use when we consider the $2I_o$ current source that biases the differential amplifier source- or emitter-coupled pair. Figure 9.19 shows a simple current mirror for which $r_x = r_d{}' = r_{o2}$.

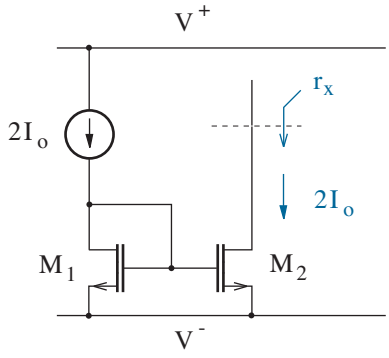


Figure 9.19: Current mirror with small-signal shunt resistance.

If we find the preceding r_x too small, we can add $R_s{}'$ resistance in series with each MOSFET source. Integrated resistors often consume large areas, so a better solution is to have the M_2 source look into a separate M_4 drain as in the **cascode current source** of Fig. 9.20. Here, we have

$$r_x = r_{o2}\left[\,1 + (g_{m2} + g_{mb2})\,r_{o4}\,\right] . \tag{9.81}$$

The factor in brackets typically offers two-order-of-magnitude improvement.

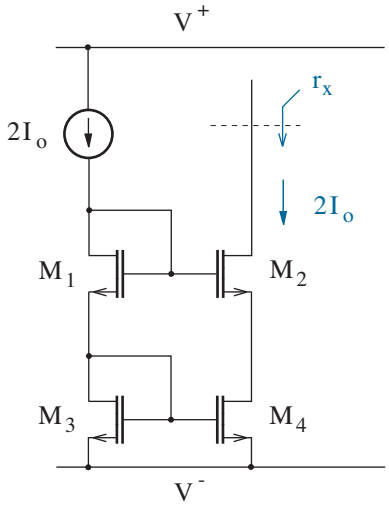


Figure 9.20: Cascode current source.

Housekeeping concluded. We are ready for op-amp design examples.

9.3 Front-End Designs

Figure 9.21 shows a simple candidate for an op-amp front-end design with two source-coupled MOSFETs and resistor loads. Input currents are zero, as desired, and the differential gain is $-g_m R$. To first order,

$$g_m = \sqrt{2K_n{}'(W/L)\, i_d|_Q}\,. \tag{9.82}$$

Thus, we seek large, but reasonable (W/L) and $i_d|_Q$ —excessive values lead to impractical geometric constraints and power consumption, respectively. We also seek large R.

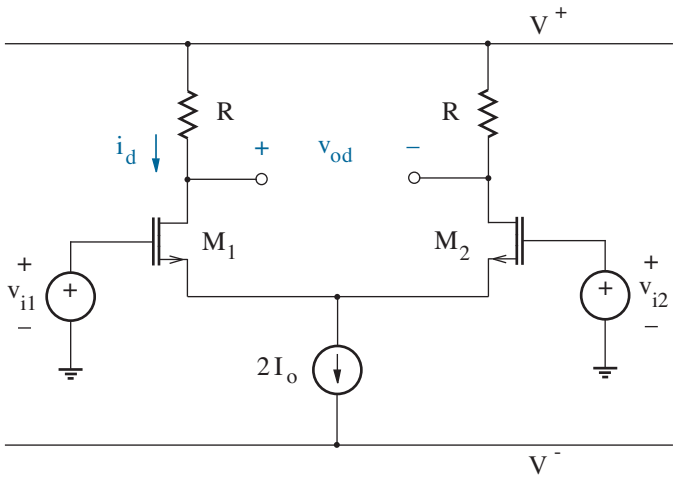


Figure 9.21: Differential amplifier with resistor loads.

But there is an unhappy large-R penalty. The MOSFETs have quiescent node voltages: $V^+ - i_d|_Q R$ at either drain and $0 - v_{gs}|_Q$ at either source. We subtract the drain and source values to find

$$v_{ds}|_Q = V^+ - i_d|_Q R + v_{gs}|_Q\,. \tag{9.83}$$

Meanwhile, for MOSFET operation in the saturation mode, we require

$$v_{ds}|_Q \;\geq\; v_{gs}|_Q - V_T\,. \tag{9.84}$$

So a little algebra reveals an upper bound for R. Specifically,

$$R \;\leq\; \frac{V^+ + V_T}{i_d|_Q}\,. \tag{9.85}$$

Not liking resistors is a recurring theme in modern integrated electronics. We need a better load, one that presents a large resistance to *small signals* while limiting the quiescent voltage drop induced by large-signal current. The requisite behavior is that of a non-linear device.

Example 9.2

Figure 9.22 shows a differential amplifier with small-signal MOSFET loads. Complete the design so that $A_{dd} = -20$ and $i_{d5}|_Q = 200\ \mu$A.

n-channel MOSFETs:

$K_n{}' = 50\ \mu\text{A/V}^2$, $V_{Tn} = +0.5$ V, $\lambda_n = 0.05\ \text{V}^{-1}$, $\gamma_n = 0.2\ \text{V}^{1/2}$.

p-channel MOSFETs:

$K_p{}' = 20\ \mu\text{A/V}^2$, $V_{Tp} = -0.5$ V, $\lambda_p = 0.08\ \text{V}^{-1}$, $\gamma_p = 0.2\ \text{V}^{1/2}$.

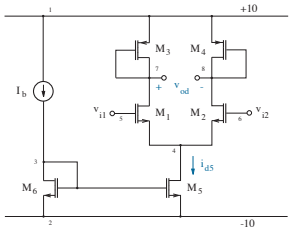


Figure 9.22: Circuit for Example 9.2.

Solution

We start with zero values for λ and γ. The differential voltage gain is

$$A_{dd} = -g_{m1} R_{d1}{}' \,, \tag{9.86}$$

where

$$R_{d1}{}' = \frac{1}{g_{m3}} \,. \tag{9.87}$$

Thus,

$$A_{dd} = -\frac{g_{m1}}{g_{m3}} = -\sqrt{\frac{K_n{}'(W/L)_1}{K_p{}'(W/L)_3}}\,, \tag{9.88}$$

subject to $i_{d1}|_Q = |\,i_{d3}|_Q|$. The ratio $K_n{}'/K_p{}' = 2.5$ is favorable.

Now let $(W/L)_3 = 4$ (a reasonable sizing that promotes modest $v_{gs3}|_Q$). Then $(W/L)_1 = 4 \times (20)^2 / 2.5 = 640$. In turn, by symmetry, $(W/L)_4 = 4$ and $(W/L)_2 = 640$. With $(W/L)_5 = (W/L)_6 = 4$, the current mirror has $i_{d5}|_Q = 200$ μA when $I_b = 200$ μA. We put off further I_b considerations.

The SPICE analysis that checks our design at this point is unremarkable apart from two ac signal sources, one at v_{i1} (node 5 in Fig. 9.15)

```
Vi1        5          0          ac          1m
```

and the other with reverse polarity at v_{i2} (node 6 in Fig. 9.15)

```
Vi2        0          6          ac          1m
```

The .model statements used for the MOSFETs do not yet include λ or γ. We add a .op statement to recover biasing data, and we use .ac analysis (as in Example 8.4) with no expectation of frequency-dependent behavior. All of this yields $|A_{dd}| = 20.000$ when we plot [v(7) - v(8)] / [v(5) - v(6)] in the "probe" environment. The A_{dd} result is encouraging.

But $\lambda \neq 0$. Thanks to the .op statement, we have the following:

	M_1	M_2	M_3	M_4	M_5	M_6
ID	1.00E–04	1.00E–04	-1.00E–04	-1.00E–04	2.00E–04	2.00E–04
VGS	5.79E–01	5.79E–01	-2.08E+00	-2.08E+00	1.91E+00	1.91E+00
VDS	8.50E+00	8.50E+00	-2.08E+00	-2.08E+00	9.42E+00	1.91E+00
VBS	-9.42E+00	-9.42E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
GM	2.53E–03	2.53E–03	1.26E–04	1.26E–04	2.83E–04	2.83E–04

So with

$$r_o = \frac{1/\lambda + |\, v_{ds}|_Q|}{|\, i_d|_Q|}\,, \tag{9.89}$$

we apply the M_1 and M_3 data to estimate $r_{01} = 285$ kΩ and $r_{03} = 146$ kΩ. The M_1 and M_3 transconductances are also transformed. Specifically,

$$g_{m1} \to g_{m1}\sqrt{1 + \lambda_n\, v_{ds1}|_Q} = 3.02 \times 10^{-3}\ \mho\,, \tag{9.90}$$

$$g_{m3} \to g_{m3}\sqrt{1 + \lambda_p |\, v_{ds3}|_Q|} = 1.36 \times 10^{-4}\ \mho\,. \tag{9.91}$$

The latter is somewhat small since M_3 is being used as a load. Accordingly, it is appropriate to determine $R_{d1}{}' = 1/g_{m3} \parallel r_{o3} = 7.00$ kΩ even though we have tended to assume $r_o \gg 1/g_m$. Whereas A_{dd} varies with $R_{d1}{}' \parallel r_{o1}$ (Table 9.1, CS half circuit, $R_{s1}{}' = 0$), the estimated A_{dd} value is 20.6. Thus, we let $(W/L)_1 = (W/L)_2 = 640 \times (20/20.6)^2 = 603$.

The preceding estimates assume $i_{d5}|_Q = 200\ \mu$A. Nevertheless, $\lambda \neq 0$ alters the mirror bias currents, which vary in proportion to $(1 + \lambda v_{ds}|_Q)$. For the intended mirror action we require

$$i_{d5} = \frac{1}{2} K_n'(W/L)_5 \left(v_{gs5} - V_{Tn}\right)^2 \left(1 + \lambda_n v_{ds5}|_Q\right) = 200\ \mu\text{A}\,, \qquad (9.92)$$

and

$$i_{d6} = \frac{1}{2} K_n'(W/L)_6 \left(v_{gs6} - V_{Tn}\right)^2 \left(1 + \lambda_n v_{ds6}|_Q\right) = 200\ \mu\text{A}\,. \qquad (9.93)$$

Then with $v_{gs5}|_Q = v_{gs6}|_Q$, and with the help of the VDS data, we divide Eq. 9.93 by Eq. 9.92 to obtain $(W/L)_6 = 4 \times (1.462/1.096) = 5.3$.

Back to SPICE. The revised MOSFET sizings yield $|A_{dd}| = 20.011$ and the following ...

	M_1	M_2	M_3	M_4	M_5	M_6
ID	1.02E–04	1.02E–04	-1.02E–04	-1.02E–04	2.05E–04	2.00E–04
VGS	5.69E–01	5.69E–01	-1.99E+00	-1.99E+00	1.68E+00	1.68E+00
VDS	8.58E+00	8.58E+00	-1.99E+00	-1.99E+00	9.43E+00	1.68E+00
VBS	-9.43E+00	-9.43E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
GM	2.97E–03	2.97E–03	1.38E–04	1.38E–04	3.47E–04	3.39E–04
GDS	3.58E–06	3.58E–06	7.07E–06	7.07E–06	6.96E–06	9.23E–06

Here, GDS $= 1/r_o$. The bias conditions differ slightly from our estimates.

A final design iteration must account for $\gamma \neq 0$. MOSFETs M_1 and M_2 are the only devices that experience a body effect through non-zero v_{bs}. This alters the threshold voltage, but it does not alter the design for A_{dd} — the M_1 and M_2 sources are at ac ground by virtue of half-circuit symmetry. So we run a final SPICE simulation without changing the MOSFET sizings. The new $|A_{dd}|$ is 20.202, a reflection of modified bias conditions.

We can determine common-mode performance with SPICE by changing the polarity of the Vi2 signal and plotting [v(7) - v(8)] / v(5) with "probe". For the symmetric design at hand, we find a perfect $A_{cd} = 0$. If we allow for some asymmetry, say $(W/L)_1 = 600$ and $(W/L)_2 = 606$, we obtain $|A_{cd}| = 9.56 \times 10^{-5}$. So with $|A_{dd}|$ barely changed, the CMRR is 107 dB.

How good is our design? A differential voltage gain of only -20 is not what we would expect for an op-amp. And an already large $(W/L)_1$ sizing that increases with the square of A_{dd} spells trouble for further improvement.

Current-Mirror Loading

Having witnessed the lackluster effort of two-terminal MOSFET loads, we look to another solution such as current sources with ac resistance r_o. Nevertheless, an exceptionally clever technique applies a current mirror as shown in Fig. 9.23 to achieve the added benefit of a single-ended output.

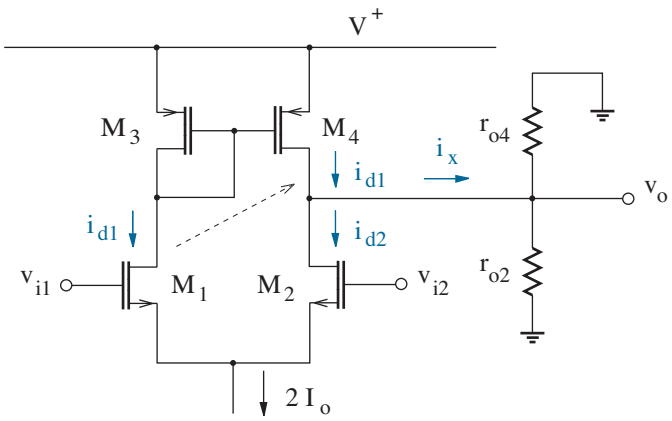


Figure 9.23: Current mirror applied as a load for a differential amplifier. To clarify analysis, the small-signal resistances looking into the M_2 and M_4 drains are presented as external to the transistors.

Consider the half circuit that would apply to M_1 or M_2 with symmetric drain connections at the V^+ rail. The shared source node is at ac ground. So to first order, $i_{d1} \approx g_m v_{i1}$ and $i_{d2} \approx g_m v_{i2}$. Inserting the M_3 - M_4 current mirror does little to affect i_{d1} and i_{d2}. However, the mirror action establishes i_{d1} as the drain current for M_4. In turn, we find

$$i_x = g_m v_{i1} - g_m v_{i2} = g_m v_{id} \tag{9.94}$$

as a differential current that flows from the node common to M_2 and M_4. The ac resistances looking into the M_2 and M_4 drains are r_{o2} and r_{o4}, respectively, and the differential i_x current effectively applies to the parallel combination of these resistances. Thus,

$$A_{dd} = \frac{v_o}{v_{id}} = g_m(r_{o2} \parallel r_{o4}) \, . \tag{9.95}$$

The r_{o2} and r_{o4} values can be very large.

A downside for the current-mirror load is the slight asymmetry that is introduced when the M_3 drain is connected to the M_3 gate (unlike M_4). This establishes $r_{o1} \neq r_{o2}$ and $r_{o3} \neq r_{04}$ for $A_{cd} \neq 0$. On the other hand, common-mode components in i_{d1} and i_{d2} tend to be suppressed through the subtractive process that determines i_x.

Example 9.3

Figure 9.24 shows a differential amplifier with current-mirror loading. Complete the design so that $A_{dd} = 200$ and $i_{d5}|_Q = 200$ μA.

Use the MOSFET parameters of Example 9.2.

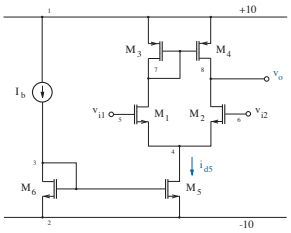


Figure 9.24: Circuit for Example 9.3.

Solution

We require non-zero λ to account for r_o values. But $\gamma = 0$, for now.

Rather than agonize over precise values for the various $|v_{gs}|_Q|$, we start with *crude* estimates at 1 V so that the quiescent M_2 source is set at -1 V. Thus, $v_{ds2}|_Q = |v_{ds4}|_Q| = 5.5$ V if the quiescent value for v_o partitions equal $|v_{ds}|$ voltages. Then with $i_{d2}|_Q = |i_{d4}|_Q| = 100$ μA, we apply Eq. 9.89 to find $r_{02} \approx 255$ kΩ, $r_{04} \approx 180$ kΩ, and $r_{02} \parallel r_{o4} \approx 106$ kΩ.

Whereas $A_{dd} = g_m(r_{02} \parallel r_{o4})$, we require $g_m = 1.89 \times 10^{-3}$ $\mho$. In turn,

$$(W/L)_2 = \frac{{g_m}^2}{2\,{K_n}'\,i_{d2}|_Q(1 + \lambda_n\, v_{ds2}|_Q)} = 279\,. \tag{9.96}$$

Symmetry conditions require the same value for $(W/L)_1$. The $(W/L)_3$ and

$(W/L)_4$ values are arbitrary, so we set them both to 4 as in Example 9.2. We also arbitrarily let $(W/L)_5 = 4$. Then with $(1 + \lambda_n v_{ds5}|_Q) \approx 1.45$ and $(1 + \lambda_n v_{ds6}|_Q) \approx 1.05$, $(W/L)_6 = 4 \times (1.45/1.05) = 5.5$.

Onward to SPICE. The simulation is similar to that of Example 9.2, but the "probe" output of interest is v(8) / [v(5) - v(6)]. In not much time, we have $A_{dd} = 188.65$ and the following:

	M_1	M_2	M_3	M_4	M_5	M_6
ID	9.87E–05	9.87E–05	-9.87E–05	-9.87E–05	1.97E–04	2.00E–04
VGS	5.99E–01	5.99E–01	-1.96E+00	-1.96E+00	1.66E+00	1.66E+00
VDS	8.64E+00	8.64E+00	-1.96E+00	-1.96E+00	9.40E+00	1.66E+00
VBS	-9.40E+00	-9.40E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
GM	1.99E–03	1.99E–03	1.35E–04	1.35E–04	3.41E–04	3.45E–04
GDS	3.45E–06	3.45E–06	6.83E–06	6.83E–06	6.72E–06	9.23E–06

As expected, the $|v_{gs}|_Q|$ values are off, in some cases by as much as 1 V. Nevertheless, the r_o (reciprocal GDS) values are not particularly erroneous, and A_{dd} is close to the intended objective. The $i_{d5}|_Q$ value is acceptable. An easy fix makes $(W/L)_1 = (W/L)_2 = 279 \times (200/189)^2 = 312$.

Back to SPICE to obtain $A_{dd} = 199.44$. Then back again with $\gamma \neq 0$ to obtain $A_{dd} = 203.72$ and a new set of data:

	M_1	M_2	M_3	M_4	M_5	M_6
ID	9.72E–05	9.72E–05	-9.72E–05	-9.72E–05	1.94E–04	2.00E–04
VGS	1.06E+00	1.06E+00	-1.95E+00	-1.95E+00	1.66E+00	1.66E+00
VDS	9.11E+00	9.11E+00	-1.95E+00	-1.95E+00	8.49E+00	1.66E+00
VBS	-8.94E+00	-8.94E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
GM	2.10E–03	2.10E–03	1.34E–04	1.34E–04	3.35E–04	3.45E–04
GDS	3.34E–06	3.34E–06	6.73E–06	6.73E–06	6.72E–06	9.23E–06
VTH	9.63E–01	9.63E–01	-5.00E–01	-5.00E–01	5.00E–01	5.00E–01
GMB	6.80E–05	6.80E–05	1.73E–05	1.73E–05	4.33E–05	4.46E–05

Note that the body effect enhances g_{m2}, which varies with $(1 + \lambda v_{ds2}|_Q)$. However, the overall improvement is small, since r_{o2} and r_{o4} are diminished. One last "tweak" makes $(W/L)_1 = (W/L)_2 = 301$ so that $A_{dd} = 200.11$.

Once again, we determine common-mode performance with SPICE by changing the polarity of the Vi2 signal, but we plot v(8)/v(5) with "probe". For this example, we obtain $A_{cd} = 2.30 \times 10^{-2}$. The CMRR is 79 dB.

How good is our design? A differential gain of 200 is far better than 20. But it is still below par for an op-amp, and further improvement remains dependent on the square of (W/L). We need a revised circuit configuration with either a larger load resistance or another stage for gain.

A current-mirror load with substantially higher ac resistance takes the form of Fig. 9.20. However, with reference to Fig. 9.23, improvements in an effective r_{o4} that remains in parallel with an unmodified r_{o2} bear little advantage unless the latter is also increased. This suggests the combination of a cascode differential amplifier and a cascode current-mirror load shown in Fig. 9.25. The total differential gain is

$$A_{dd} = g_m \left\{ r_{o2a} \left[1 + (g_m + g_{mb})\, r_{o2} \right] \parallel r_{o4} \left[1 + (g_m + g_{mb})\, r_{o4a} \right] \right\} . \quad (9.97)$$

So if $r_{o2} \approx r_{o2a} \approx r_{o4} \approx r_{o4a} = r_o$ and $g_{mb} \approx 0$,

$$A_{dd} \approx (g_m r_o)^2 . \quad (9.98)$$

Example 9.3 suggests that $A_{dd} \sim 10,000$ should be easily achieved.

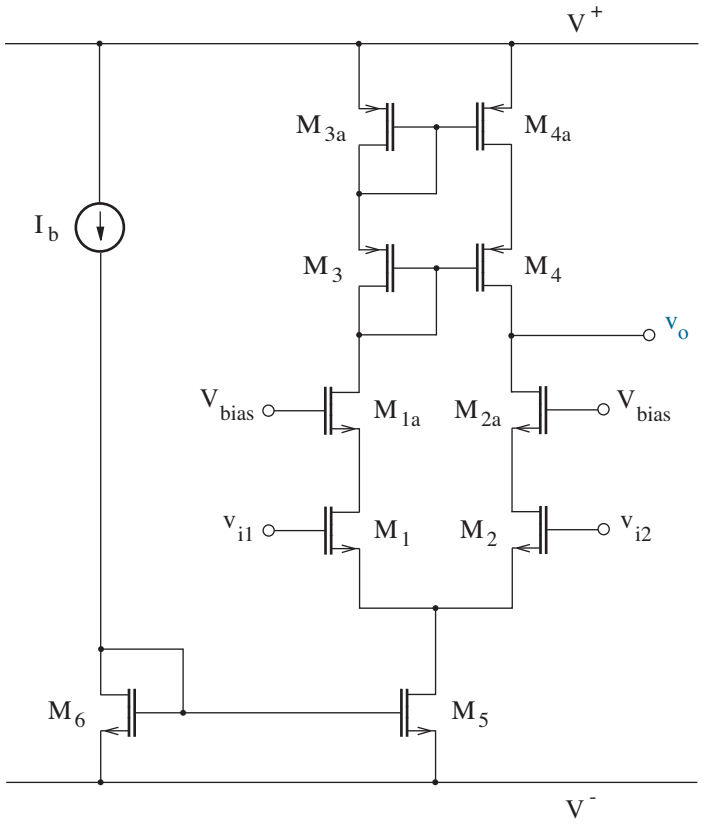


Figure 9.25: Cascode differential amplifier with cascode current-mirror.

The circuit of Fig. 9.25 has one major disadvantage: Five MOSFET layers between the power-supply rails limits available output voltage swing (see Problem 9.62).

An improved cascode differential amplifier is "folded" as in Fig. 9.26. Here, M_1 and M_2 assume p-channel roles but with unaltered ac connections. The current mirror featuring M_5 and M_6 is also reconfigured to hang from the V^+ supply rail, and two additional current sources establish $I_b/2$ bias currents for M_{1a}, M_{2a}, M_3, M_4, M_{3a}, and M_{4a}. The sources have no other influence on A_{dd} since they are effectively open circuits under ac conditions.

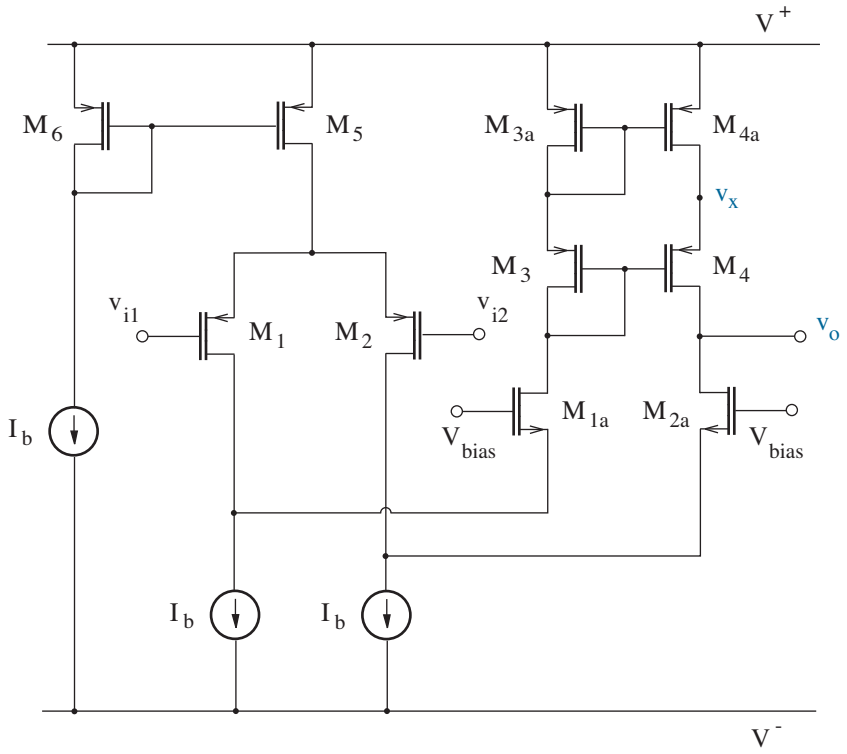


Figure 9.26: Folded cascode differential amplifier.

Folding does not improve the available high-side output swing when the current-mirror load has p-channel MOSFETs. With reference to Fig. 9.26, we apply KVL to the symmetric circuit to find the node voltage

$$v_x = V^+ - |v_{gs3a}| - |v_{gs3}| + |v_{gs4}| = V^+ - |v_{gs3a}| \,. \tag{9.99}$$

If M_4 is to remain in saturation, $|v_{ds4}| = v_x - v_o \geq |v_{gs4}| - |V_{Tp}|$. In turn,

$$V^+ - v_o \geq |V_{Tp}| + \sqrt{\frac{I_b}{K_p{}'(W/L)_3}} + \sqrt{\frac{I_b}{K_p{}'(W/L)_{3a}}} \tag{9.100}$$

subject to $(W/L)_3 = (W/L)_4$.

Front-End Bias Current

In our discussion thus far, we have used current mirrors to propagate the influence of an ideal current source for the purpose of transistor biasing. How does this current originate? An easy method negotiates a MOSFET and a resistor as shown in Fig. 9.27. For MOSFET M_1, which operates in saturation by virtue of its gate-to-drain connection,

$$I_b = \frac{1}{2} K_n{}' \left(\frac{W}{L} \right) (v_{gs1} - V_{Tn})^2 \, . \tag{9.101}$$

And for the resistor,

$$I_b = \frac{(V^+ - V^-) - v_{gs1}}{R} \, . \tag{9.102}$$

A straightforward solution determines v_{gs} and I_b. Nevertheless, the solution is far from ideal since it depends on the values of the power-supply voltages. Good op-amps "reject" the power supplies when establishing A_{dd}.

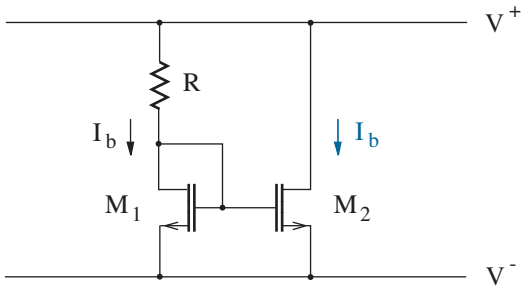


Figure 9.27: Negotiated bias circuit featuring a MOSFET and resistor.

An intriguing alternative replaces R with a p-channel current mirror as shown in Fig. 9.28. The circular mirror arrangement makes I_b independent of the power supplies. But I_b is undefined—any value appears to suffice.

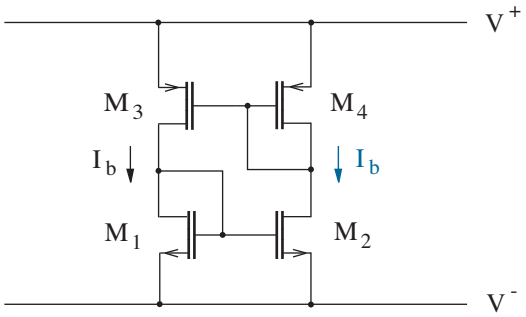


Figure 9.28: Two-current-mirror bias circuit.

The two-current-mirror bias circuit produces a well-defined I_b when we insert two-terminal elements with different current-voltage characteristics below the M_1 and M_2 sources. Figure 9.29 shows a typical arrangement. With $v_{gs1} = v_{gs2}$, the voltages across M_5 and R are also equal. Specifically,

$$v_{gs5} = \sqrt{\frac{2\,I_b}{K_n'(W/L)_5}} + V_{Tn} = I_b R\,. \tag{9.103}$$

So for a particular I_b, we require

$$(W/L)_5 = \frac{I_b}{(1/2)\,K_n'(I_b R - V_{Tn})^2}\,. \tag{9.104}$$

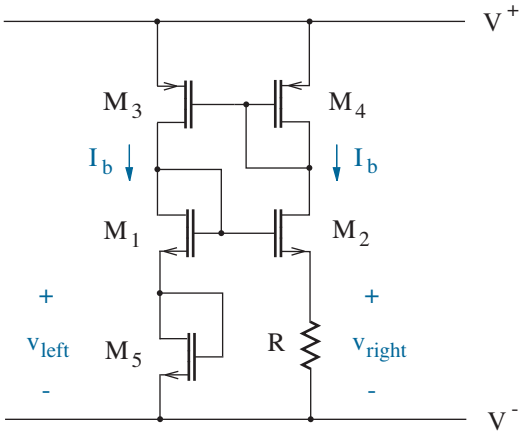


Figure 9.29: Supply-independent biasing circuit.

Another arrangement features a pnp transistor with emitter area A in place of M_5 and a pnp transistor with emitter area XA in series with R. Both transistors act as as diodes with $v_{bc} = 0$. At the left of the bias circuit,

$$v_{\,\text{left}} = \frac{kT}{q} \ln \frac{I_b}{AJ_s}\,, \tag{9.105}$$

where J_s is the emitter saturation current density. Similarly,

$$v_{\,\text{right}} = \frac{kT}{q} \ln \frac{I_b}{XAJ_s} + I_b R\,. \tag{9.106}$$

So with equal left- and right-side voltages,

$$I_b = \frac{(kT/q)\ln X}{R}\,. \tag{9.107}$$

Unfortunately, the supply-independent biasing circuit of Fig. 9.29 also supports a stable mode of operation with $I_b = 0$. Thus, a start-up circuit must be included to force $I_b \neq 0$.

9.4 Intermediate and Output Stages

Intermediate Stage Design

In our quest for more gain, we add a second stage to the differential amplifier of Fig. 9.24 as shown in Fig. 9.30. The new stage sports a common-source topology with a current-source load, and it functions as a **level shifter**—the quiescent output voltage can be made zero (as expected for an op-amp).

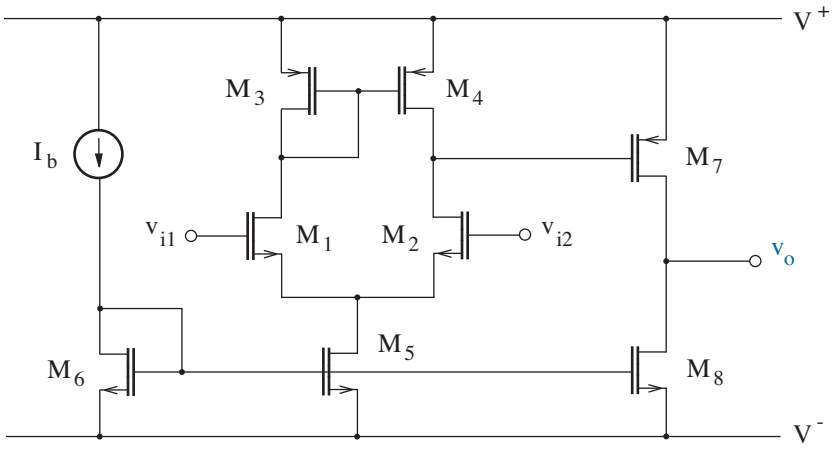


Figure 9.30: Two-stage CMOS operational amplifier: n-channel inputs.

If we interchange the roles of the n-channel and p-channel MOSFETs, we obtain the two-stage op-amp of Fig. 9.31. The alternative configuration has improved noise characteristics (see Chapter 13).

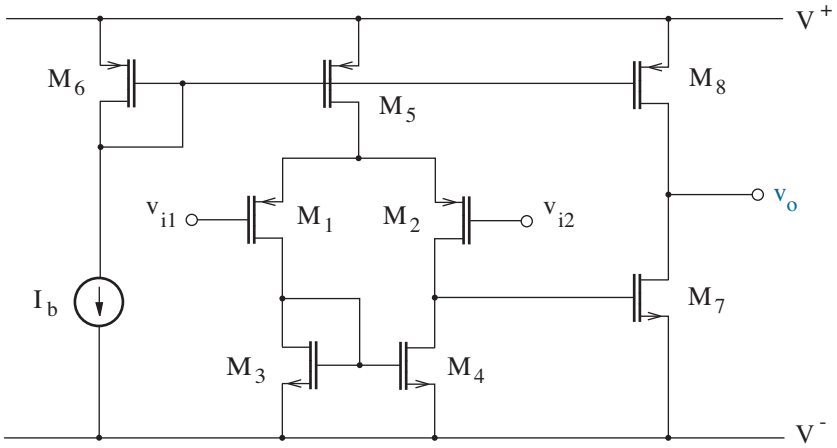


Figure 9.31: Two-stage CMOS operational amplifier: p-channel inputs.

In what follows, we focus upon the "alternative" op-amp of Fig. 9.31. The results are easily extended to the circuit of Fig. 9.30, notwithstanding.

The multistage amplifiers in Chapter 7 were coupled with the aid of capacitors so that the biasing in individual stages could be independent. Not so for an op-amp, whose differential gain must extend to the dc limit. For the circuit at hand, coupling and symmetry conditions require

$$v_{gs7}|_Q = v_{ds4}|_Q = v_{ds3}|_Q = v_{gs3}|_Q \,. \tag{9.108}$$

We express $v_{gs7}|_Q$ and $v_{gs3}|_Q$ in terms of $i_{d7}|_Q$ and $i_{d3}|_Q$, respectively, then equate the results to find

$$\sqrt{\frac{2\,i_{d7}|_Q}{K_n{}'(W/L)_7(1+\lambda_n\,v_{ds7}|_Q)}}+V_{Tn}=\sqrt{\frac{2\,i_{d3}|_Q}{K_n{}'(W/L)_3(1+\lambda_n\,v_{ds3}|_Q)}}+V_{Tn}\,. \tag{9.109}$$

In turn,

$$\frac{i_{d7}|_Q}{i_{d3}|_Q}=\frac{(W/L)_7(1+\lambda_n\,v_{ds7}|_Q)}{(W/L)_3(1+\lambda_n\,v_{ds3}|_Q)}\,. \tag{9.110}$$

Meanwhile,

$$i_{d7}|_Q = |\,i_{d8}|_Q| \tag{9.111}$$

and

$$i_{d3}|_Q=\frac{|\,i_{d5}|_Q|}{2}\,. \tag{9.112}$$

So we divide Eq. 9.111 by Eq. 9.112 to obtain

$$\frac{i_{d7}|_Q}{i_{d3}|_Q}=\frac{2\,|\,i_{d8}|_Q|}{|\,i_{d5}|_Q|}=\frac{2\,(W/L)_8(1+\lambda_p|\,v_{ds8}|_Q|)}{(W/L)_5(1+\lambda_p|\,v_{ds5}|_Q|)}\,. \tag{9.113}$$

(in consideration of appropriate scaling for the M_5 and M_8 mirror currents). Finally, with the help of Eq. 9.110,

$$\frac{(W/L)_7(1+\lambda_n\,v_{ds7}|_Q)}{(W/L)_3(1+\lambda_n\,v_{ds3}|_Q)}=\frac{2\,(W/L)_8(1+\lambda_p|\,v_{ds8}|_Q|)}{(W/L)_5(1+\lambda_p|\,v_{ds5}|_Q|)}\,. \tag{9.114}$$

Equation 9.114 is a relation that avoids **systemic offset** ($v_o|_Q \neq 0$.) When properly used for design, (W/L) factors are minimized. For example, $(W/L)_7$ is necessarily large to support large g_{m7}, so $(W/L)_3$ is also large. This is unfortunate, since it forces a large value for $(W/L)_4$. Nevertheless, Eq. 9.114 allows relatively small $(W/L)_5$ and $2\,(W/L)_8$. The latter sizing does not need to be large to establish a substantial r_{o8} load resistance.

Example 9.4

Figure 9.32 shows a two-stage differential amplifier with p-channel inputs. Complete the design so that $A_{dd} = 20,000$ and $|i_{d5}|_Q| = 200$ μA.

Use the MOSFET parameters of Example 9.2.

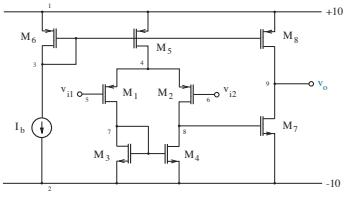


Figure 9.32: Circuit for Example 9.4.

Solution

As in Example 9.3, we start with $\lambda \neq 0$, $\gamma = 0$, and crude 1-V estimates for the various $|v_{gs}|_Q|$. The differential voltage gain for the first stage is

$$A_{dd}{}^{(1)} = g_{m2} \left(r_{o2} \parallel r_{04} \right). \tag{9.115}$$

Whereas $|v_{ds2}|_Q| \approx 10$ V, $v_{ds4}|_Q \approx 1$ V, and $|i_{d2}|_Q| = |i_{d5}|_Q| / 2 = 100$ μA, we use Eq. 9.89 to find $r_{o2} = 225$ kΩ, $r_{o4} = 210$ kΩ, and $r_{o2} \parallel r_{o4} = 101$ kΩ. Now let $(W/L)_1 = (W/L)_2 = (W/L)_3 = (W/L)_4 = X$, an unknown factor. In turn, $g_{m2} = 8.49 \times 10^{-5} \sqrt{X}$ $\mho$ and $A_{dd}{}^{(1)} = 9.08 \sqrt{X}$.

The differential voltage gain for the second stage is

$$A_{dd}{}^{(2)} = g_{m7} \left(r_{o7} \parallel r_{08} \right). \tag{9.116}$$

However, we need to know $i_{d7}|_Q$ and $|i_{d8}|_Q|$, which are equal currents that are held hostage in relation to $|i_{d5}|_Q|$ by virtue of Eqs. 9.113 and 9.114. Subject to $(W/L)_7 = X$, $v_{ds7}|_Q = 10$ V, and $v_{ds3}|_Q \approx 1$ V, we have

$$\frac{|i_{d8}|_Q|}{|i_{d5}|_Q|} = \frac{(W/L)_7(1 + \lambda_n v_{ds7}|_Q)}{2\,(W/L)_3(1 + \lambda_n v_{ds3}|_Q)} = 0.714\,. \tag{9.117}$$

So with $|i_{d5}|_Q| = 200$ μA, $i_{d7}|_Q = |i_{d8}|_Q| = 143$ μA.

Now that $i_{d7}|_Q$ and $|i_{d8}|_Q|$ are known, $r_{o7} = 210$ kΩ, $r_{o8} = 157$ kΩ, and $r_{o7} \parallel r_{o8} = 89.8$ kΩ. Then $g_{m7} = 1.47 \times 10^{-4}\sqrt{X}$ ℧ and $A_{dd}{}^{(2)} = 13.2\sqrt{X}$. The total gain is $A_{dd} = A_{dd}{}^{(1)}A_{dd}{}^{(2)} = 120\,X$. Thus, $A_{dd} = 20,000$ requires $X = (W/L) = 167$ for M_1, M_2, M_3, M_4, and M_7.

What remains are the sizings for the p-channel MOSFETs that provide bias currents for the gain stages. We assign $(W/L)_8 = 4$ and $I_b = 200$ μA. Then we apply Eqs. 9.113 and 9.117 to find

$$(W/L)_5 = \frac{(W/L)_8}{0.714}\,\frac{1+\lambda_p|v_{ds8}|_Q|}{1+\lambda_p|v_{ds5}|_Q|} = 5.9 \tag{9.118}$$

and

$$(W/L)_6 = (W/L)_5\,\frac{1+\lambda_p|v_{ds5}|_Q|}{1+\lambda_p|v_{ds6}|_Q|} = 9.4\,. \tag{9.119}$$

Here, we assume $|v_{ds5}|_Q| \approx 9$ V, $|v_{ds6}|_Q| \approx 1$ V, and $|v_{ds8}|_Q| = 10$ V.

A quick SPICE simulation reveals that $|i_{d5}|_Q| = 191$ μA is too small. So we let $(W/L)_6 \to 9.4 \times 200/191 = 9.0$. Then a revised simulation yields $|i_{d5}|_Q| = 199$ μA, $A_{dd} = 20,148$, and the following:

	M_2	M_4	M_5	M_6	M_7	M_8
ID	-9.94E–05	9.94E–05	-1.99E–04	-2.00E–04	1.42E–04	-1.42E–04
VGS	-6.82E-01	6.52E-01	-1.89E+00	-1.89E+00	6.52E-01	-1.89E+00
VDS	-1.00E+01	6.52E-01	-9.32E+00	-1.89E+00	9.51E+00	-1.05E+01
VBS	9.32E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
GM	1.09E–03	1.31E–03	2.86E–04	2.88E–04	1.87E–03	2.04E–04
GDS	4.41E–06	4.81E–06	9.11E–06	1.39E–05	4.81E–06	6.18E–06

MOSFETs M_1 and M_3 track device data for M_2 and M_4, respectively.

We are clearly in the ballpark for gain, so we forge ahead with $\gamma \neq 0$. The related SPICE simulation yields $A_{dd} = 20,928$, which is too large. This allows $X \to 167 \times 20,000/20,928 = 160$. In turn, $A_{dd} = 20,056$.

The match with our gain objective is close enough. Thus, we turn our attention to the quiescent output voltage, specifically, $v_o|_Q = -0.215$ V. Whereas smaller $(W/L)_7$ shifts $v_o|_Q$ upward, some trial and error yields $v_o|_Q = 28$ mV when $(W/L)_7 = 157$. The final A_{dd} is 20,003.

How good is our design? Another SPICE simulation yields $A_{cd} = 0.576$ for a CMRR of 90.8 dB.

Output Stage Design

Although in many ways similar, a differential amplifier followed by a single-ended gain stage with zero offset is not an operational amplifier. The latter circuit includes an output stage that is designed to accommodate resistive or capacitive loads requiring large currents over a rail-to-rail range of swing. Output stages can be complex, and space is limited, so we will be content to examine a simple MOSFET implementation that highlights key issues. The end-of-chapter problems do the same for the LM741 with BJT devices.

Consider the CMOS output stage of Fig. 9.33 featuring complementary n-channel (M_1) and p-channel (M_2) MOSFETs and two V_b biasing sources. To keep matters simple, the circuit is symmetric with $V^+ = |V^-| = V_{rail}$, $V_{Tn} = |V_{Tp}| = V_T$, and

$$K_n'(W/L)_1 = K_p'(W/L)_2 = K \, . \tag{9.120}$$

Our initial objective will require the output stage to operate as a **push-pull** power amplifier: Subject to $v_i < 0$, M_1 is "off", M_2 operates in saturation, and current is "pushed" into the load. Conversely, when $v_i > 0$, M_2 is "off", M_1 operates in saturation, and current is "pulled" from the load. Finally, $v_i = 0$ is consistent with $v_o = 0$ (no offset). All this demands

$$V_b = V_{rail} - V_T \, . \tag{9.121}$$

The means for implementing the V_b voltage sources will be considered later. What remains is a design value for K given a minimum for R.

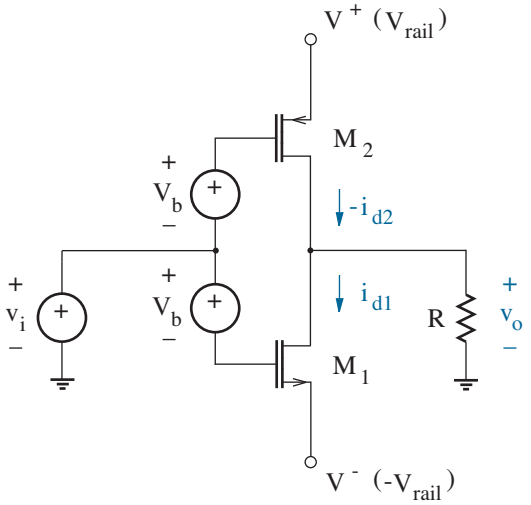


Figure 9.33: CMOS "push-pull" output stage for an operational amplifier.

As noted previously, $v_i > 0$ puts M_1 in saturation with v_o falling rapidly. However, a sufficiently low v_o leads to the M_1 resistive mode and a slower rate of decrease. Thus, we design for K so that voltage v_o at the edge of saturation for M_1 has the value v_{om} with suitable proximity to the V^- rail. For the M_1 terminal voltages of interest, $v_{gs1} = v_i - V_b + V_{rail} = v_i + V_T$ and $v_{ds1} = v_o + V_{rail}$. Then at the edge of saturation, $v_{ds1} = v_{gs1} - V_T$ or

$$v_i = v_{om} + V_{rail} \; . \tag{9.122}$$

Meanwhile, with M_1 in saturation and M_2 off,

$$\frac{1}{2}K(v_{gs1} - V_T)^2 = \frac{1}{2}K{v_i}^2 = \frac{-v_{om}}{R} \; . \tag{9.123}$$

We eliminate v_i from Eqs. 9.122 and 9.123 to find

$$K = \frac{2\,|v_{om}|}{R\,(V_{rail} - |v_{om}|)^2} \; . \tag{9.124}$$

Similar considerations for M_2 yield the same result.

Let $V_{rail} = 10$ V. Then with $|v_{om}| = 9$ V and $R = 1$ kΩ, Eq. 9.123 yields $K = 18$ mA/V^2. Subject to ${K_n}' = 50$ μA/V^2 and ${K_p}' = 20$ μA/V^2, the required MOSFET aspect ratios are $(W/L)_1 = 360$ and $(W/L)_2 = 900$, respectively. Increased load currents will necessitate even larger MOSFETs. Figure 9.34 shows the applicable SPICE results for $V_T = 0.5$ V.

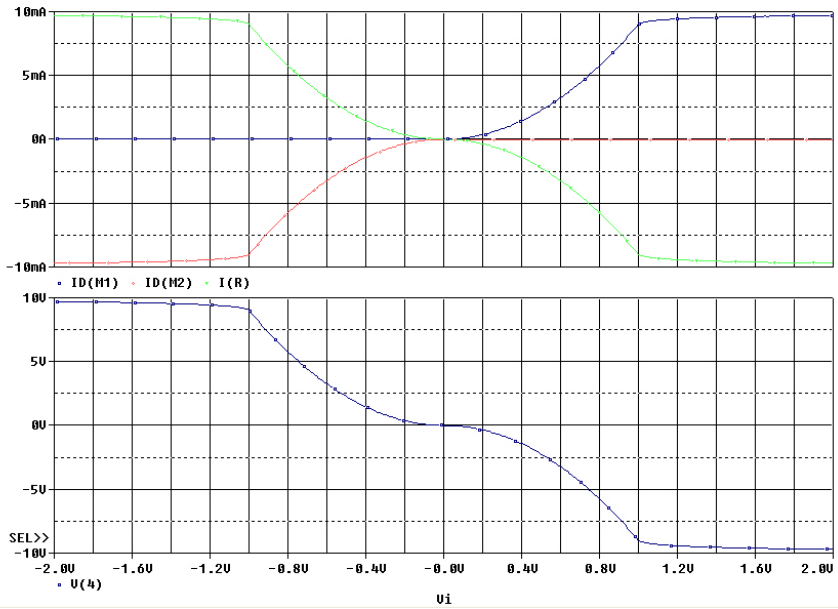


Figure 9.34: SPICE performance for the push-pull amplifier of Fig. 9.33. The MOSFETs are biased so that they are at threshold when $v_i = 0$.

When biased according to Eq. 9.121, the push-pull output stage clearly suffers from a non-linear input-output characteristic. And worse, the slope of the characteristic is zero at $v_i = 0$, thereby ruining the total op-amp gain. The flatness reflects **cross-over distortion**, a problem first encountered in Chapter 6 with the BJT push-pull power amplifier.

To correct for cross-over distortion, we alter the bias supplies so that

$$V_b = V_{rail} - V_T - v_x \,, \tag{9.125}$$

where v_x is a small decrease. This makes M_1 and M_2 slightly on at $v_i = 0$. Nevertheless, the drain-current magnitudes are equal, no current flows to the load, and $v_o = 0$. For M_1, $v_{gs1} = v_i - V_b + V_{rail} = v_i + V_T + v_x$, and

$$i_{d1} = \frac{1}{2}K(v_{gs1} - V_T)^2 = \frac{1}{2}K(v_i + v_x)^2 \,. \tag{9.126}$$

Whereas for M_2, $v_{gs2} = v_i + V_b - V_{rail} = v_i - V_T - v_x$, and

$$-i_{d2} = \frac{1}{2}K(v_{gs2} + V_T)^2 = \frac{1}{2}K(v_i - v_x)^2 \,. \tag{9.127}$$

Finally,

$$\frac{v_o}{R} = -i_{d2} - i_{d1} \,. \tag{9.128}$$

When we combine Eqs. 9.126, 9.127, and 9.128, the squared voltages cancel out to yield the *linear* result

$$v_o = -2KR\, v_x v_i \,, \tag{9.129}$$

which, as expected, vanishes in the limit as $v_x \to 0$. The penalty for the improved behavior is non-zero quiescent drain current. Specifically,

$$i_{d1}|_Q = -\,i_{d2}|_Q = \frac{1}{2}K{v_x}^2 \,. \tag{9.130}$$

Finite quiescent drain current implies a transconductance for M_1 and M_2. Thus, the gain of the output stage is

$$A_v = \frac{v_o}{v_i} = -2\, g_m R \,. \tag{9.131}$$

subject to $g_{m1} = g_{m2} = g_m$.

It is not too difficult to show that $v_x \neq 0$ is consistent with the design implications of Eq. 9.124 if M_2 is "off" at v_{om}. With $v_{gs2} = v_i - V_T - v_x$, this requires $v_i \geq v_x$. We apply Eq. 9.129 to find

$$\frac{v_{om}}{-2KR\, v_x} \geq v_x \,. \tag{9.132}$$

Then with the help of Eq. 9.130,

$$i_{d1}|_Q = -\,i_{d2}|_Q \geq \frac{|v_{om}|}{4R} \tag{9.133}$$

is the condition for linear operation over a range including v_{om}.

Back to the example at hand— With $|v_{om}| = 9$ V and $R = 1$ kΩ, Eq. 9.133 has $i_{d1}|_Q = -i_{d2}|_Q \geq 2.25$ mA. Nevertheless, we choose 2.25 mA as the minimum quiescent current to avoid excessive power consumption. Then with $K = 18$ mA/V^2, Eqs. 9.130 and 9.125 yield $v_x = 0.5$ V and $V_b = 9$ V, respectively. Figure 9.35 shows the revised SPICE results. The output has linear variation over the intended range, and the voltage gain of -18 is consistent with Eq. 9.131.

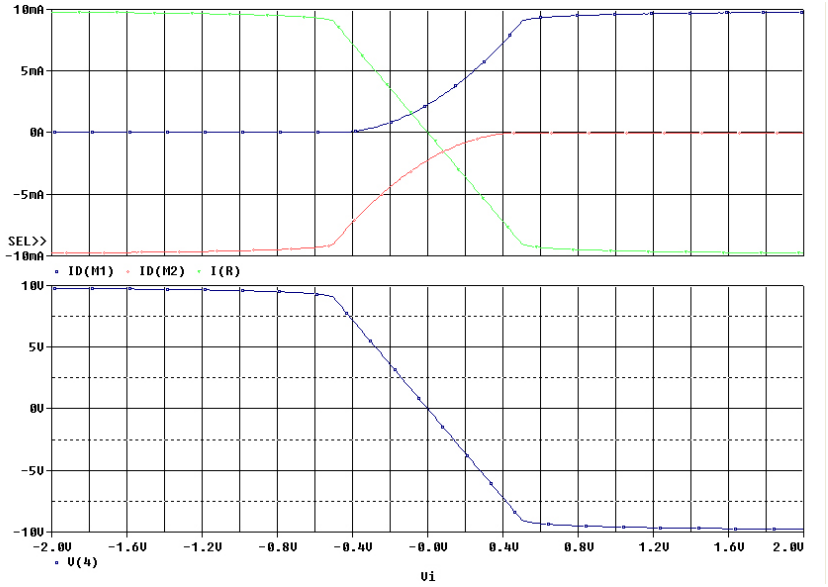


Figure 9.35: SPICE performance for the push-pull amplifier of Fig. 9.33. The MOSFETs are biased so that they are slightly on when $v_i = 0$.

Exercise 9.9 An output stage with the form of Fig. 9.33 features ±5-V supply voltages. The output is linear over ±4.75 V for $i_{load} \leq 0.25$ mA. Determine $(W/L)_1$, $(W/L)_2$, $i_{d1}|_Q$, and v_x.

$K_n' = 50\ \mu$A/V^2, $V_{Tn} = +0.6$ V, $K_p' = 20\ \mu$A/V^2, $V_{Tp} = -0.6$ V.

Ans: $(W/L)_1 = 160$ $(W/L)_2 = 400$ $i_{d1}|_Q = 62.5\ \mu$A $v_x = 0.125$ V

What remains is the implementation of the V_b biasing sources of Eq. 9.125. Better yet, we need to establish V_{bias} potentials in relation to the power-supply rails such that

$$V_{rail} - |V_{bias}| = V_T + v_x\ . \tag{9.134}$$

We also need an effective input, a corequisite that is surprisingly difficult.

Consider the circuit of Fig. 9.36. Here, M_{1c} and M_{2c} are symmetric ($K_n = K_p = K$, $V_{Tn} = |V_{Tp}| = V_T$) so that bias current $2I_b$ splits evenly: $i_{d1c}|_Q = -i_{d2c}|_Q = I_b$. Let $(W/L)_{1b} = (W/L)_{1c}$ and $(W/L)_{2b} = (W/L)_{2c}$. Then we have $V_{bias,1} = V_{b1}$ and $V_{bias,2} = V_{b2}$. In turn,

$$i_{d1}|_Q = I_b \frac{(W/L)_1}{(W/L)_{1a}} \tag{9.135}$$

and

$$-i_{d2}|_Q = I_b \frac{(W/L)_2}{(W/L)_{2a}} \, . \tag{9.136}$$

Equations 9.135 and 9.136 support a design that is blind to specific v_x. The input is positioned to allow desired output swing, but the quiescent input voltage is necessarily non-zero.

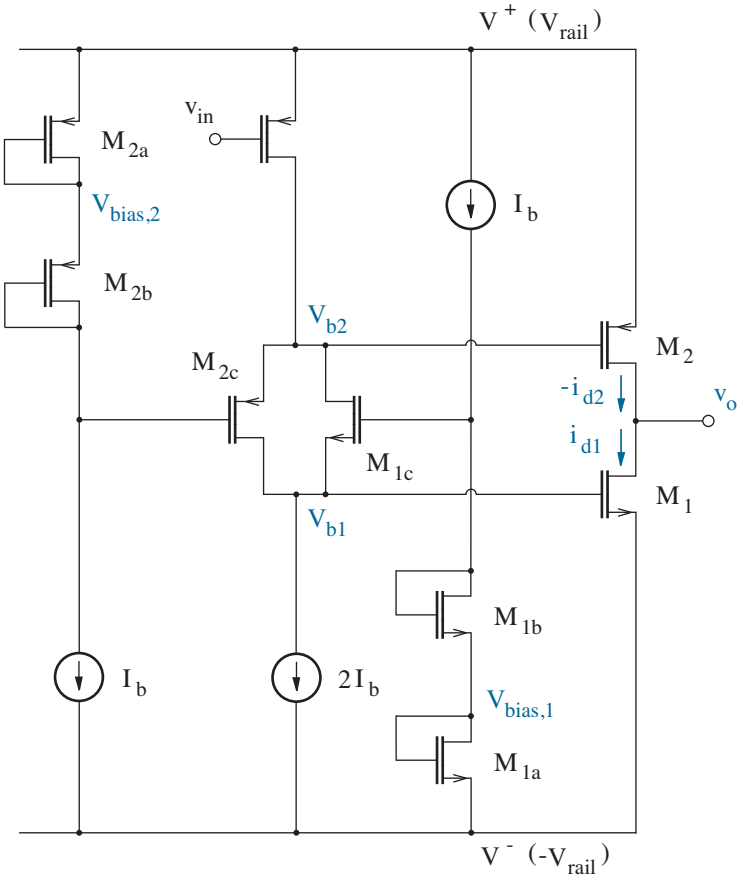


Figure 9.36: CMOS output stage with quiescent biasing for M_1 and M_2.

Concept Summary

Signal pairs have differential- and common-mode (average) components.

- The ideal differential voltage amplifier
  - Multiplies differential components by a differential gain (A_{vd}).
  - Rejects the influence of common-mode components ($A_{vc} = 0$).
  - Has infinite common-mode rejection ratio (CMRR $= A_{vd}/A_{vc}$).

Modern op-amps are designed as analog integrated circuits.

- Front end designs
  - Typically feature source- or emitter-coupled transistor pairs with left-right symmetry and current-source biasing (Fig. 9.3).
  - Use current-mirror loading to improve gain and provide a single-ended output with respect to ground (Fig. 9.24).
  - Are subject to non-zero input offset voltage when asymmetric.
  - Are often facilitated using half-circuits to determine differential or common-mode performance.
- Intermediate stages provide additional gain and output level shifting.
- Push-pull output stages allow for large output currents and a linear transfer characteristic over a broad output voltage range (Fig. 9.23).
- Op-amp designs are complicated by second-order transistor effects.
  - Channel-length modulation in a MOSFET or base-width modulation in a BJT yields a small-signal output resistance r_o.
    - Finite r_o alters expressions for amplifier voltage gain.
    - Finite r_o alters small-signal transistor-terminal resistances, particularly those looking into a drain or collector.
  - Substrate bias effects in a MOSFET provide additional small-signal modifications that tend to degrade performance.
- As noted, current sources are critical design components.
  - Current sources are used for transistor biasing, and they provide small-signal resistive loading.
  - A current mirror that is independent of supply voltage is often used as the basis for current-mirror biasing throughout a circuit.

Problems

Section 9.1

9.1 Show that the BJTs in the differential amplifier of Fig. 9.4 operate in the forward active mode when $|v_{od}| < V^+ + 0.5$ V.

9.2 Show that Eq. 9.27 is applicable to a MOSFET differential amplifier with $|v_{od}| < I_o R$ (see Fig. 9.4 with M_1 and M_2 in place of Q_1 and Q_2).

9.3 Consider the pair of half circuits shown in Fig. 9.6.

(a) Determine the small-signal variation of the Q_1 emitter with $+v_{id}/2$.

(b) Determine the small-signal variation of the Q_2 emitter with $-v_{id}/2$.

(c) Find the total small-signal variation of the shared half-circuit node.

9.4 Use half circuits to determine A_{dd} and A_{cc} for the amplifier of Fig. P9.4. Assume $\beta_o = 100$.

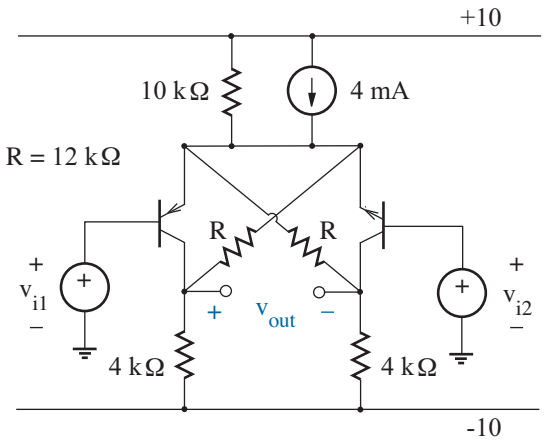


Figure P9.4

9.5 Use half circuits to determine A_{dd} and A_{cc} for the amplifier of Fig. P9.5. Assume $g_{m1} = g_{m2} = 5 \times 10^{-3}$ ℧.

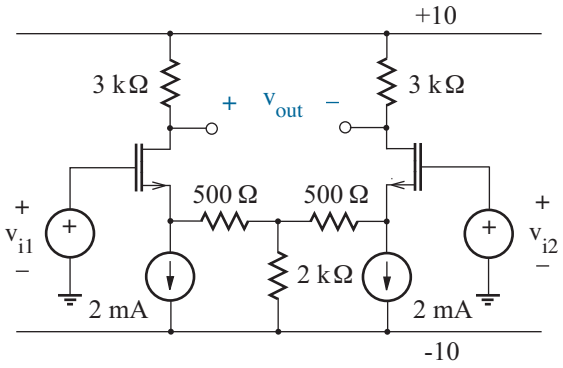


Figure P9.5

9.6 Use half circuits to determine A_{dd} and A_{cc} for the amplifier of Fig. P9.6. Let $\beta_o = 100$, and let $K'W/L = 1$ mA/V^2, $V_T = 1$ V for the MOSFETs.

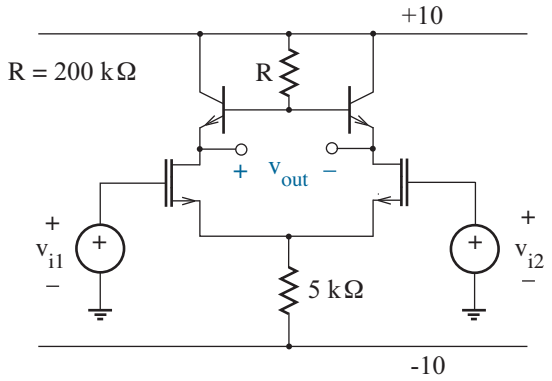


Figure P9.6

9.7 Find the CMRR (in dB) for the amplifier of Fig. P9.7 with $K'W/L = 2$ mA/V^2, $V_T = -0.5$ V.

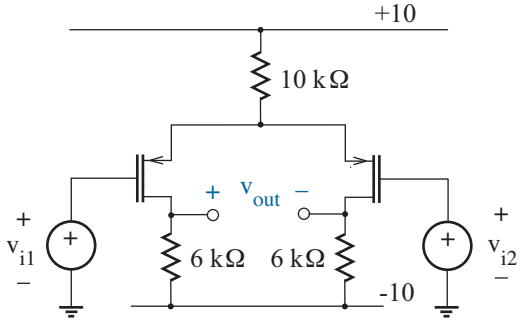


Figure P9.7

9.8 Find the CMRR (in dB) for the amplifier of Exericse 9.1 when the output is the node voltage at the Q_2 collector. Compare with the result when the output is voltage v_{od} (see also Exercise 9.3).

9.9 Show that an op-amp with a particular CMRR can be modeled as an ideal op-amp with an effective input offset voltage.

9.10 Use small-signal analysis to derive Eq. 9.39.

9.11 Show that a MOSFET differential amplifier with ΔR asymmetry has input offset voltage

$$v_{os} = \left(\frac{v_{gs} - V_T}{2}\right) \frac{\Delta R}{R} .$$

9.12 Show that a MOSFET differential amplifier with $\Delta(W/L)$ asymmetry has input offset voltage

$$v_{os} = \left(\frac{v_{gs} - V_T}{2}\right) \frac{\Delta(W/L)}{W/L} .$$

9.13 Show that a MOSFET differential amplifier with ΔV_T asymmetry has input offset voltage

$$v_{os} = \Delta V_T .$$

9.14 The input offset current (i_{os}) for a differential amplifier is defined as a difference of input currents. Let Q_1 and Q_2 in the differential amplifier of Fig. 9.4 reflect asymmetries with ΔI_s and $\Delta\beta_F$. Derive an expression for the standard deviation of i_{os}.

9.15 The differential amplifier of Fig. 9.4 has $R = 5$ kΩ, $I_o = 1$ mA, $V^+ = 5$ V, and $V^- = -5$ V. The BJTs have IS=10f and BF=100.

(a) Use SPICE to find v_{os} when $\Delta R = 50$ Ω, and compare with the theoretical result.

(b) Use SPICE to find v_{os} when $\Delta IS = 0.1$f, and compare with the theoretical result.

(c) Use SPICE to find v_{os} with both asymmetries.

9.16 The differential amplifier of Fig. 9.4 has the specifications in Problem 9.15, and standard deviations for R and IS are 50 Ω and 0.1f, respectively. Apply a 100-trial Monte Carlo SPICE analysis to find standard deviation Δv_{os}. (See Example 7.11.)

9.17 Consider a BJT differential amplifier with the form of Fig. 9.4 and $R = 2$ kΩ. Complete the design so that $A_{dd} = -100$ and indicate the r_x requirement so that $A_{cc} = 0.001$. Assume BJTs with $\beta_F = 100$.

9.18 Repeat Problem 9.17, but let the design have n-channel MOSFETs for which $K_n' = 50\ \mu\text{A/V}^2$ and $V_{Tn} = 0.5$ V.

9.19 The MOSFETs in the differential amplifier of Fig. 9.19 have $K_n' = 50\ \mu\text{A/V}^2$, $V_{Tn} = 0.5$ V, $C_{gs} = C_{gd} = 0.1$ pF, and $C_{db} = C_{sb} = 0.8$ pF. Estimate f_h for A_{dd} and compare with SPICE.

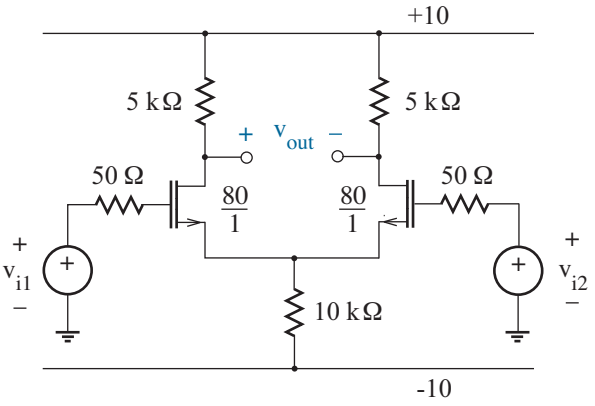


Figure P9.19

9.20 Repeat Problem 9.19, replacing MOSFETs with BJTs for which $I_s = 2 \times 10^{-15}$ A, $\beta_F = 200$, $C_\pi = 1.2$ pF, and $C_\mu = 0.5$ pF.

9.21 Use SPICE to determine the CMRR frequency response for the circuit of Problem 9.19.

9.22 Show that the circuit of Fig. P9.22 produces an output of the form

$$v_{out} = a v_x (v_y + b) \,,$$

and provide expressions for a and b.

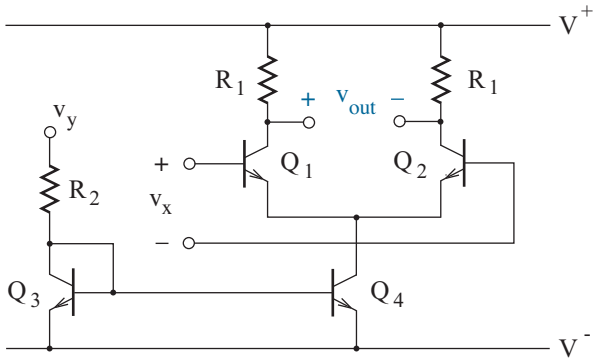


Figure P9.22

9.23 Figure. P9.23 shows a **Gilbert multiplier**.

(a) Prove that in general,

$$i_L - i_R = I_o \tanh\left(\frac{q v_x}{2kT}\right) \tanh\left(\frac{q v_y}{2kT}\right) \,.$$

(b) What conditions yield $i_L - i_R \approx a\, v_x v_y$?

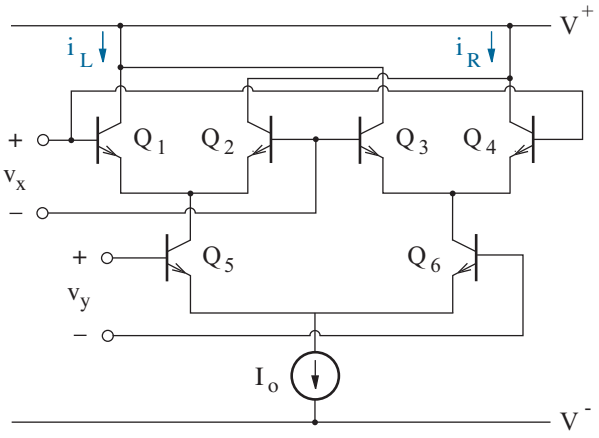


Figure P9.23

9.24 Figure. P9.24 shows the salient features of an **emitter-coupled logic (ECL)** inverter, a member of a very fast logic technology that has nevertheless succumbed to CMOS and is in near obsolescence. The BJTs have IS=10f and BF=100.

(a) Estimate V_{OL} and V_{OH} in terms of the current I_o and other factors. Consider two conditions: $v_{in} \ll V_{ref}$ and $v_{in} \gg V_{ref}$.

(b) Find R_3 so that V_{OL} and V_{OH} are symmetrically positioned in relation to V_{ref}.

(c) Use SPICE to plot the input-output transfer characteristic when the input spans the range between V_{OL} and V_{OH} as determined in part b.

(d) Determine V_{IL}, V_{IH}, and the HIGH and LOW noise margins from the results of part c.

(e) Suggest a way to realize an ECL NOR/OR gate. Consult some history books as a last resort.

(f) Chapter 6 argued that BJT logic speed suffers when constituent devices transition from saturation to cutoff—time is required to remove stored base charge. Explain why ECL is relatively fast.

(g) Suggest some reasons why CMOS has won out.

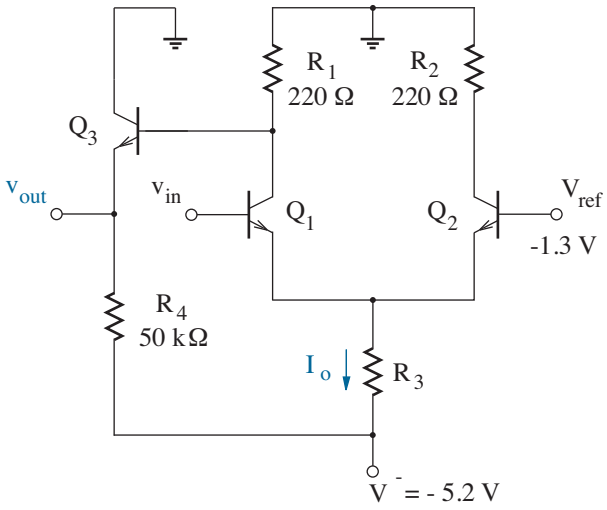


Figure P9.24

Section 9.2

9.25 An n-channel MOSFET in saturation exhibits $i_d = 10$ mA with $v_{ds} = 4$ V and $\lambda = 0.08$ V^{-1}. Determine i_d when $v_{ds} = 8$ V and $v_{ds} = 12$ V.

9.26 Show that the i_d vs. v_{ds} characteristic curves for an n-channel enhancement-mode MOSFET in saturation can be extrapolated to a common point of intersection at $v_{ds} = -1/\lambda$.

9.27 The MOSFETs in Fig. P9.27 have the same $K_n'W/L$ and V_{Tn}. However, $\lambda_n(M_1) = 0.08$ V^{-1} and $\lambda_n(M_2) = 0.12$ V^{-1}. The measured v_{ds1} is 5 V. Determine i_1 and i_2. Neglect body effects.

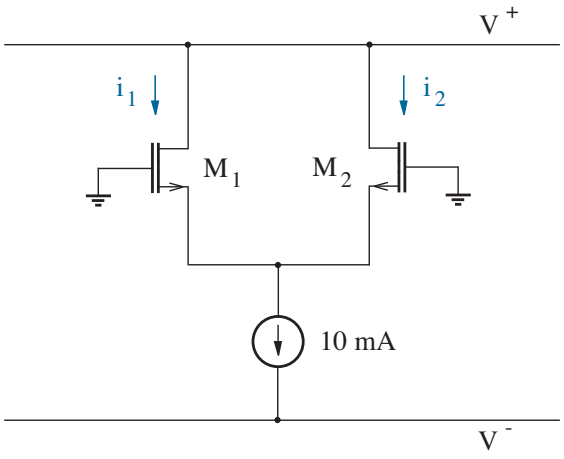


Figure P9.27

9.28 A MOSFET exhibits the following data with $v_{gs} = 2.0$ V and $v_{bs} = 0$. Determine λ.

i_d (mA)	v_{ds} (V)
0.000	0.0
2.240	1.0
2.620	2.0
2.873	3.0
3.123	4.0
3.380	5.0
3.634	6.0

9.29 A MOSFET exhibits the following data with $v_{gs} = -2.7$ V and $v_{bs} = 0$. Determine λ.

i_d (mA)	v_{ds} (V)
-9.05	-0.0
-11.56	-1.0
-13.07	-2.0
-14.57	-3.0
-16.08	-4.0
-17.59	-5.0
-19.10	-6.0

9.30 An n-channel MOSFET ($N_a = 5 \times 10^{16}$ cm^{-3}) features $\gamma = 0.5$ V$^{1/2}$. When $v_{bs} = 0$, $V_T = 0.8$ V. Find V_T when $v_{bs} = -2$ V and $v_{bs} = -5$ V.

9.31 An n-channel MOSFET features $K'W/L = 2$ mA/V^2, $V_{T0} = 0.5$ V, $\lambda = 0$, $\gamma = 0.4$ V$^{1/2}$, and $2\phi_f = 0.75$ V. Find v_{bs} such that i_d is reduced to half of its value when $v_{bs} = 0$.

9.32 A p-channel MOSFET ($N_d = 5 \times 10^{16}$ cm^{-3}) features $t_{ox} = 35$ nm. When $v_{bs} = 0$, $V_T = -0.6$ V. Find V_T when $v_{bs} = 4$ V and $v_{bs} = 8$ V.

9.33 The MOSFET in the circuit of Fig. P9.33 has $K'W/L = 1$ mA/V^2, $V_{T0} = 0.6$ V, and $\lambda = 0$.

(a) Find v_{out} when $\gamma = 0$.

(b) Find v_{out} when $\gamma = 0.5$ V$^{1/2}$ and $2\phi_f = 0.75$ V.

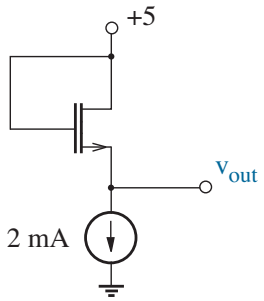


Figure P9.33

9.34 A MOSFET exhibits the following V_T data. Determine γ and $2\phi_f$.

V_T (V)	v_{bs} (V)
0.720	0.0
1.421	-2.0
1.871	-4.0
2.230	-6.0
2.538	-8.0
2.813	-9.0

9.35 A MOSFET exhibits the following V_T data. Determine γ and $2\phi_f$.

V_T (V)	v_{bs} (V)
-0.460	0.0
-1.071	2.0
-1.456	4.0
-1.762	6.0
-2.024	8.0
-2.257	9.0

9.36 An n-channel MOSFET has $K_n' = 50\ \mu\text{A/V}^2$, $V_{Tn} = 0.24$ V, $\gamma = 0.78\ \text{V}^{1/2}$, $2\phi_f = 0.64$ V, and $V_{fb} = -1.02$ V.

(a) Determine W/L such that $I_s = 1\ \mu$A for the subthreshold drain current.

(b) Find v_{ds} such that $i_d = I_s/2$ when $v_{gs} = V_T$.

(c) Find v_{gs} such that $i_d = I_s/2$ when $v_{ds} = 1$ V.

9.37 The **subthreshold gate swing** (S) is defined as the value of $|v_{gs} - V_T|$ needed to reduce the subthreshold drain current by one order of magnitude. Derive an expression for S, then determine the value of S that applies to the MOSFET of Problem 9.36.

9.38 An n-channel MOSFET has Level-7 SPICE parameters listed in Example 10.5. Perform SPICE simulations, as needed, to determine values for K_n', V_{Tn}, λ, and γ that are useful for hand calculations.

9.39 A p-channel MOSFET has Level-7 SPICE parameters listed in Example 10.5. Perform SPICE simulations, as needed, to determine values for K_n', V_{Tn}, λ, and γ that are useful for hand calculations.

9.40 Consider an npn BJT with emitter, base, and collector doping concentrations of 2×10^{19} cm^{-3}, 5×10^{17} cm^{-3}, and 5×10^{15} cm^{-3}, respectively. The metallurgical base width is 1 μm.

(a) Let $v_{be} = 0.7$ V. Find the depletion-region width on the base side of the base-emitter junction. Repeat for $v_{be} = 0.72$ V.

(b) Find the depletion-region width on the base side of the base-collector junction if $v_{ce} = 10$ V. Repeat for $v_{ce} = 20$ V.

(c) Estimate the forward Early voltage.

(d) Estimate the reverse Early voltage.

9.41 Estimate V_A for a BJT with the curve-tracer display of Problem 6.14.

9.42 Estimate V_A for a BJT with the curve-tracer display of Problem 6.15.

9.43 The BJTs in Fig. P9.43 feature $V_A = 25$ V. Determine i_1, i_2, and i_3.

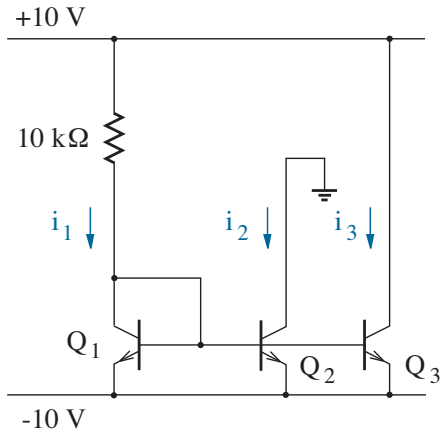


Figure P9.43

9.44 Derive the Table-9.1 voltage-gain expressions for the three MOSFET amplifier configurations.

9.45 Derive the Table-9.1 voltage-gain expressions for the three BJT amplifier configurations.

9.46 Derive the Table-9.2 small-signal resistances looking into three different MOSFET terminals.

9.47 Derive the Table-9.2 small-signal resistances looking into three different BJT terminals.

9.48 Figure P9.48 shows a **Widlar** current source used to achieve small integrated-circuit bias currents without large area-consuming resistors. The BJTs have $I_s = 2 \times 10^{-15}$ A, $\beta_F = 200$, and $V_A = 50$ V.

(a) Estimate I_o.

(b) Complete a design for R so that $i_x = 10$ μA. Hint: Write KVL around the bottom loop, and express each base-to-emitter voltage as

$$v_{be} \approx \frac{kT}{q} \ln \frac{i_c}{I_s}$$

—setting $v_{be} \approx 0.7$ V will not suffice.

(c) Determine the small-signal resistance r_x.

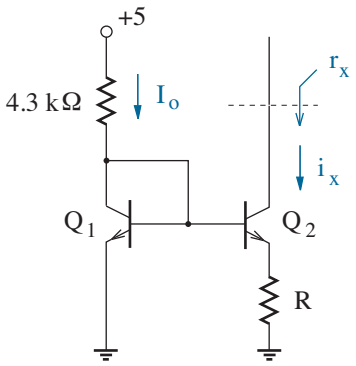


Figure P9.48

9.49 Figure P9.49 shows a **Wilson** current source. The MOSFETs feature the same g_m and r_0 values. Determine r_x.

Note: The "looking around" rules that determine small-signal resistances with the help of Table 9.2 do not apply as a consequence of feedback around the four-MOSFET loop. You will need to apply the small-signal MOSFET model to the entire circuit.

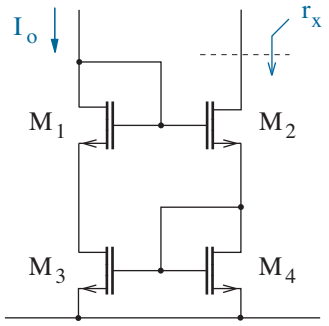


Figure P9.49

9.50 Determine A_{vm} for the amplifier of Problem P7.41 when the MOSFET has $\lambda = 0.12$ V^{-1} .

9.51 Determine A_{vm} for the amplifier of Problem P7.50 when the MOSFET has $\lambda = 0.08$ V^{-1} .

9.52 The BJTs in the circuit of Fig. P9.52 have $\beta_F = \beta_o = 100$ and $V_A = 40$ V. Determine A_{vm}. Assume appropriate base current for Q_1.

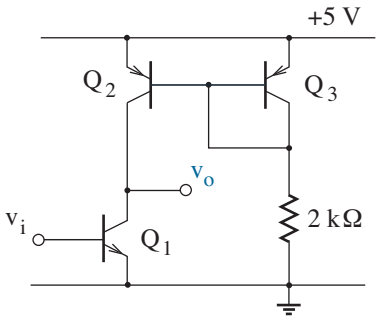


Figure P9.52

Section 9.3

9.53 Repeat Example 9.2, but design for $A_{dd} = 25$ and $i_{d5}|_Q = 400$ μA.

9.54 Repeat Example 9.2, but design for $A_{dd} = 15$ and $i_{d5}|_Q = 120$ μA.

9.55 Repeat Example 9.2, but use n-channel loads in place of M_3 and M_4.

9.56 A bewildered colleague has suggested the use of forward-biased diodes in place of the resistive loads in the front-end differential amplifier of Fig. 9.21. Explain why this design option is unattractive.

9.57 Repeat Example 9.3, but design for $A_{dd} = 280$ and $i_{d5}|_Q = 150$ μA.

9.58 Repeat Example 9.3, but design for $A_{dd} = 150$ and $i_{d5}|_Q = 300$ μA.

9.59 Repeat Example 9.3, but replace all of the MOSFETs with BJTs subject to the following:

npn - $I_s = 5 \times 10^{-15}$ A, $\beta_F = 150$, $V_A = 80$ V
pnp - $I_s = 2 \times 10^{-15}$ A, $\beta_F = 100$, $V_A = 50$ V

9.60 Repeat Example 9.3, but design for $A_{dd} = 400$ and $i_{d5}|_Q = 400$ μA, and replace the MOSFETs with the BJTs of Problem 9.59.

9.61 Design the cascode differential amplifier of Fig. 9.25 so that $A_{dd} = 10,000$ and $i_{d5}|_Q = 200$ μA. Use the MOSFET parameters of Example 9.2, and assume $\pm$10-V power supplies.

9.62 Determine the available output voltage swing for the cascode differential amplifier of Fig. 9.25. Assume $i_{d5}|_Q = I_b$.

9.63 Design the folded cascode differential amplifier of Fig. 9.26 for $A_{dd} = 10,000$ and $|i_{d5}|_Q| = 200$ μA. Use the MOSFET parameters of Example 9.2, and assume $\pm$10-V power supplies.

9.64 Design the supply-independent biasing circuit of Fig. 9.29 so that $I_b = 100$ μA. Use the MOSFET parameters of Example 9.2. Verify with SPICE.

Note: To implement the SPICE simulation, you will need to "kick" the circuit with an impulse of current.

9.65 Repeat Problem 9.64, but replace M_5 and R with BJTs subject to

BJT (left) - emitter area A
BJT (right) - emitter area XA

Use the MOSFET parameters of Example 9.2, and let IS=10f for the left BJT. Verify with SPICE.

Section 9.4

9.66 Repeat Example 9.4, but design for $A_{dd} = 15,000$ and $i_{d5}|_Q = 120$ μA.

9.67 Repeat Example 9.4, but design for $A_{dd} = 25,000$ and $i_{d5}|_Q = 150$ μA.

9.68 Repeat Example 9.4, but design for $A_{dd} = 12,000$ and $i_{d5}|_Q = 300$ μA.

9.69 Repeat Example 9.4, but use a circuit with n-channel MOSFETs connected to the input signals.

9.70 Design a CMOS output stage with the form of Fig. 9.33 so that the voltage gain is -10 over a range within 0.5 V of the $\pm$5-V power-supply rails. Assume 2-mA load current. The MOSFETs have $K_n' = 50$ μA/V^2, $V_{Tn} = 0.5$ V, $K_p' = 20$ μA/V^2, $V_{Tp} = -0.5$ V. Specify quiescent power dissipation, and verify the design with SPICE.

9.71 Design a CMOS output stage with the form of Fig. 9.33 so that the voltage gain is -15 over a range within 0.8 V of the $\pm$10-V power-supply rails. Assume 5-mA load current. The MOSFETs have $K_n' = 50$ μA/V^2, $V_{Tn} = 0.5$ V, $K_p' = 20$ μA/V^2, $V_{Tp} = -0.5$ V. Specify quiescent power dissipation, and verify the design with SPICE.

Supplementary Problems

The following problems all pertain to the LM741 op-amp shown in Fig. P9.72. For the BJTs,

npn - $I_s = 5 \times 10^{-15}$ A, $\beta_F = 120$, $V_A = 35$ V
pnp - $I_s = 2 \times 10^{-15}$ A, $\beta_F = 50$, $V_A = 35$ V

Assume $\beta_F = \beta_o$, and let $V^+ = |V^-| = 15$ V.

9.72 Find $i_{c11}|_Q$, then apply the Widlar current-source analysis used in Problem 9.36 to find $i_{c10}|_Q$.

Ans: $i_{c11}|_Q = 730\ \mu$A, $i_{c10}|_Q = 19\ \mu$A

9.73 Explain the role of Q_7.

9.74 Use half-circuit analysis to find the small-signal current leaving the Q_3 collector in terms of the differential input voltage v_{id}.

Ans: $-ic_3/v_{id} = 1.83 \times 10^{-4}$ ℧

9.75 Find $i_{c16}|_Q$ and $i_{c17}|_Q$ subject to $i_{c16}|_Q + i_{c17}|_Q = -\,i_{c13}|_Q = 730\ \mu$A.

Ans: $i_{c16}|_Q = 14\ \mu$A, $i_{c17}|_Q = 716\ \mu$A

9.76 Find $r_{b16}{}'$. Ignore Q_{21} and C_1.

Ans: $r_{b16}{}' = 1.26$ MΩ

9.77 Determine $A_{dd}{}^{(1)}$ for the front-end amplifier. Ignore Q_{22} and C_1.

Ans: $A_{dd}{}^{(1)} = -461$

9.78 Use half-circuit analysis to determine the small-signal differential input resistance.

Ans: $r_{id} = 1.3$ MΩ

9.79 Find r, the small-signal resistance of the R_7, R_8, and Q_{18} subcircuit. Assume $i|_Q = 730\ \mu$A.

Ans: $r = 101\ \Omega$

9.80 Find $r_{c13}{}'$. Neglect the bias current for Q_{14} (consistent with $i|_Q = 730\ \mu$A).

Ans: $r_{c13}{}' \approx 68.5$ kΩ.

9.81 Determine $A_{dd}{}^{(2)}$ for the gain stages involving Q_{16} and Q_{17}. Ignore Q_{20}, Q_{22}, and C_1.

Ans: $A_{dd}{}^{(2)} = -657$

(Thus, $A_{dd}{}^{(1)} A_{dd}{}^{(2)} \approx 300,000$.)

9.82 Describe the subcircuit involving Q_{14} and Q_{20}. Ignore Q_{15}. What function does the subcircuit of Problem 9.60 provide?

9.83 Transistor Q_{15} is intended to protect Q_{14} when the latter is subjected to large emitter current. Explain how this protection is effected, and estimate the maximum output source current.

Ans: i_{out} (max) ≈ 25 mA

9.84 Explain the function of Q_{22}.

Note: The role for C_1 is established in Chapter 12.

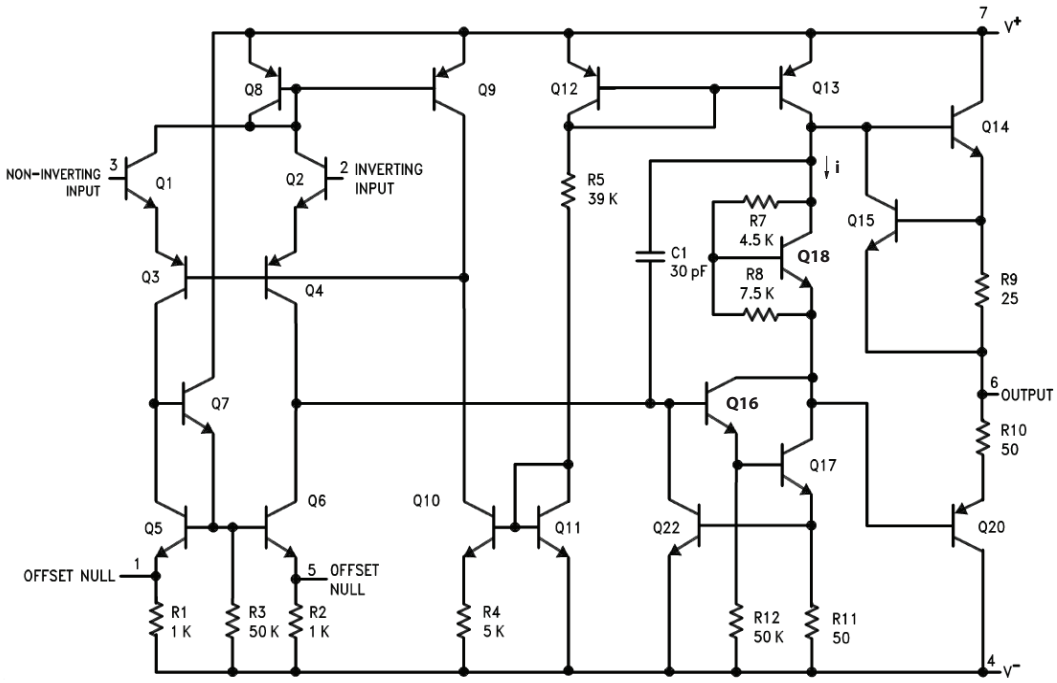
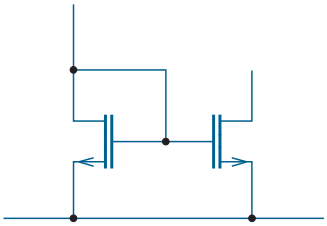


Fig. P9.72: LM741 Operational Amplifier. Courtesy Texas Instruments.



Chapter 10

Gray Boxes: Digital Integrated Circuits

At last, we have had sufficient exposure to electronic devices and fabrication processes to understand the internal workings of digital integrated circuits, thereby allowing the transformation of "black" boxes with ideal behavior to more practical "gray boxes" with peculiar limitations.

The chapter begins with performance issues that primarily relate to the periphery of a CMOS digital integrated circuit (as well as many analog ICs). Then we discuss representative CMOS digital circuits found in the interior. Of these, the CMOS inverter is the most basic constituent. Thus, we devote the remainder of the chapter to static and dynamic inverter design.

Objectives

After studying this chapter and working through the problems, you should be able to perform each of the following:

- Describe the transistor behavior that applies throughout the various regions of the CMOS inverter transfer characteristic (Section 10.1).
- Describe the faults that arise from ESD and latch-up, and specify the implementation and function of protective circuity (Section 10.1).
- Design a CMOS combinational logic circuit with appropriate sizings for the n- and p-channel MOSFET constituents (Section 10.2).
- Design a CMOS inverter for a particular switching threshold voltage, possibly subject to the effects of velocity saturation (Section 10.3).
- Design a CMOS inverter with balanced propagation delays, given an unbalanced transistor and capacitance configuration (Section 10.4).

10.1 The CMOS Digital Perimeter

Chapter 5 introduced the CMOS inverter as a pair of MOSFET switches that connect an output to LOW or HIGH levels when one switch is closed and the other is open. Our present task investigates the input and output characteristics that this representative circuit presents to the outside world. We defer a discussion of specific inverter design to Sections 10.3 and 10.4.

Figure 10.1 shows a CMOS inverter circuit with supply level V_{DD} and ground level GND (in deference to convention). Note that both transistors have zero body-to-source voltage, so there is no meddlesome body effect.

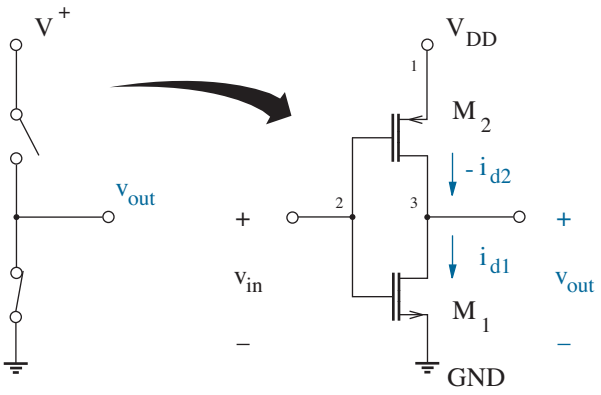


Figure 10.1: CMOS inverter circuit. The supply voltage is V_{DD}.

First examine M_1, the n-channel device, for which $v_{gs} = v_{in}$ and $v_{ds} = v_{out}$. We identify two modes of operation in terms of the input/output variables. Specifically,

$$v_{in} < V_{Tn} \quad \rightarrow \quad M_1 \text{ off}\,, \tag{10.1}$$

$$v_{out} > v_{in} - V_{Tn} > 0 \quad \rightarrow \quad M_1 \text{ saturation}\,, \tag{10.2}$$

where V_{Tn} is the n-channel threshold voltage. Then we do the same for M_2, the p-channel device, for which $v_{gs} = v_{in} - V_{DD}$ and $v_{ds} = v_{out} - V_{DD}$. After a bit of algebra, we find

$$v_{in} > V_{DD} - |V_{Tp}| \quad \rightarrow \quad M_2 \text{ off}\,, \tag{10.3}$$

$$v_{out} < v_{in} + |V_{Tp}| < V_{DD} \quad \rightarrow \quad M_2 \text{ saturation}\,, \tag{10.4}$$

where V_{Tp} is the p-channel threshold voltage (here taken at absolute value). These considerations allow us to construct the set of operational boundaries shown in Fig. 10.2. Resistive operational regions are identified by default.

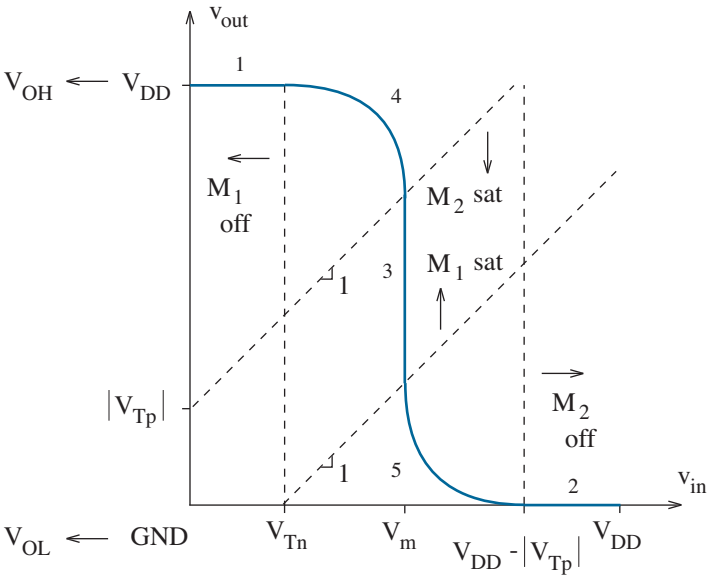


Figure 10.2: Transistor operational regions and transfer characteristic for a CMOS inverter. The input voltage V_m that corresponds to the point of maximum negative slope is called the **switching threshold voltage**.

Now that we have an operational roadmap, we can deduce portions of the inverter transfer characteristic within the confines of Fig. 10.2.

- In Region 1, M_1 is off, thereby presenting a near-infinite load to M_2. The latter operates in the resistive mode and pulls the output HIGH to V_{DD}. This is V_{OH}, the **high-level output voltage**.
- In Region 2, M_2 is off, thereby presenting a near-infinite load to M_1. The latter operates in the resistive mode and pulls the output LOW to GND. This is V_{OL}, the **low-level output voltage**.
- In Region 3, both M_1 and M_2 operate in saturation. To first order, the equal-magnitude MOSFET drain currents are v_{out} *independent*—channel-length modulation effects are needed to arbitrate a mutually acceptable output level that appropriately partitions v_{ds1} and $|v_{ds2}|$. And if v_{out} is undefined, the transfer curve has infinite negative slope. Figure 10.2 shows a desirable condition in which the corresponding input level V_m lies halfway between V_{DD} and GND.
- In Regions 4 and 5, M_1 and M_2 operate with saturation/resistive and resistive/saturation conditions, respectively. The slopes of the curve segments gradually transition between zero and large negative values. Except for intermediate points where the slope is -1, functional details seldom merit tedious analytical effort. (Leave the work to SPICE.)

The CMOS inverter transfer characteristic has two other special points, and two important performance measures derive from them:

- If $v_{in} < V_{IL}$, the **low-level input voltage**, the inverter output is *guaranteed* to be effectively HIGH.
- If $v_{in} > V_{IH}$, the **high-level input voltage**, the inverter output is *guaranteed* to be effectively LOW.

As shown in Fig. 10.3, V_{IL} and V_{IH} correspond to points on the transfer characteristic where the slope is -1. The output voltage is relatively immune to input-voltage changes in regions where the slope is less negative than -1.

Given V_{IL}, we can define a conventional, albeit fuzzy meaning of LOW. Output levels within the **LOW noise margin**

$$NM_L = V_{IL} - V_{OL} \tag{10.5}$$

present a "0" character to the input of another gate despite the corrupting influence of spurious signals. And given V_{IH}, we can do the same for HIGH. Output levels within the **HIGH noise margin**

$$NM_H = V_{OH} - V_{IH} \tag{10.6}$$

present a "1" character to the input of another gate despite the corrupting influence of spurious signals. Outputs beyond the noise margins are invalid. With $V_m = V_{DD}/2$, both CMOS noise margins are slightly lower than V_m in good fulfillment of the noise-immunity concept introduced in Chapter 1.

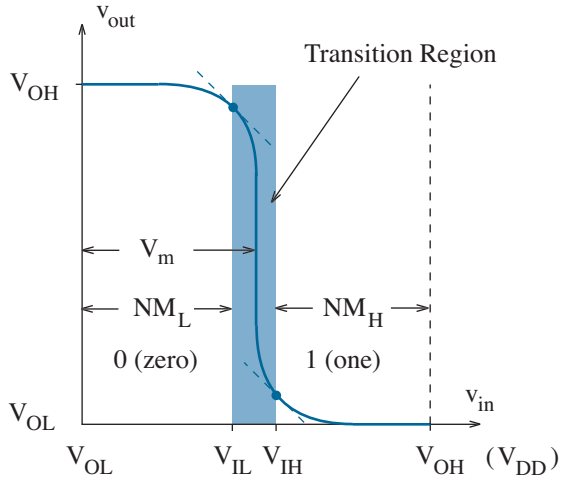


Figure 10.3: Low- and high-level inputs for which the transfer slope is -1. Inputs within the unshaded noise margins reflect valid logic levels.

As noted in Chapter 5, the CMOS inverter dissipates zero **static power** when its output rests in the HIGH or LOW state. Nevertheless, the inverter briefly experiences **crowbar current** and power dissipation during logical transitions when M_1 and M_2 are both conducting (Regions 3, 4, and 5). Intentionally slow-moving inputs exacerbate this effect—avoid them.

Dynamic power loss applies to CMOS inverters that are capacitively loaded as shown in Fig. 10.4. During LOW-to-HIGH transitions (Fig. 10.4a), the capacitor receives charge from V_{DD} through M_2, and the transfer rate is limited by the M_2 "on" resistance. The energy gained is ${CV_{DD}}^2/2$. During subsequent HIGH-to-LOW transitions (Fig. 10.4b), the capacitor gives up charge to ground through M_1, and the transfer rate is limited by the M_1 "on" resistance. The energy lost is ${CV_{DD}}^2/2$. Thus, the inverter participates in a two-step process that transfers charge from V_{DD} to ground, thereby releasing energy to the environment. The dynamic power is

$$P_d = \frac{1}{2}{CV_{DD}}^2 f \,, \tag{10.7}$$

where f is an average switching frequency that is less than or equal to the system clock rate. *It is a worthwhile objective to minimize load capacitance, which slows the output transition rates and increases the dynamic power.* The dependence on f explains why multi-GHz CMOS systems such as those found in a laptop computer require special packaging to avoid overheating.

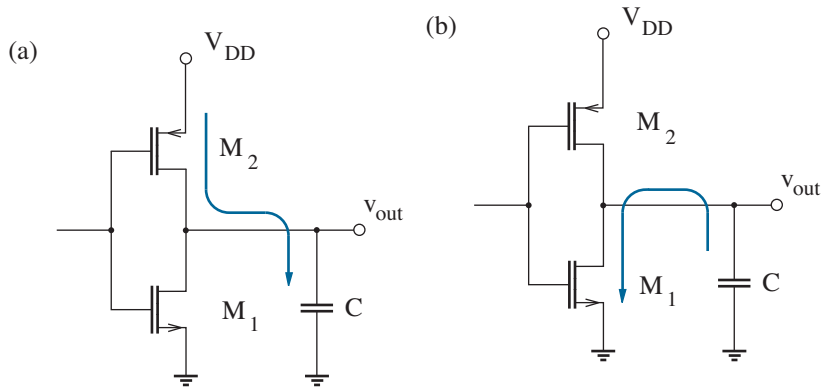


Figure 10.4: Capacitor-assisted charge transfer processes during CMOS output transitions: (a) LOW-to-HIGH; (b) HIGH-to-LOW.

Section 10.3 offers a more quantitative discussion regarding MOSFET sizing requirements for a specific switching threshold V_m, and Section 10.4 does the same for HIGH-to-LOW and LOW-to-HIGH timing attributes. The remainder of this section keeps us close to the CMOS digital perimeter.

Input Problem: Electrostatic Discharge (ESD)

The input to the CMOS inverter of Fig. 10.1 is purely capacitive, a reflection of MOSFET gate capacitance. M_1 and M_2 have very thin gate insulators, so they are susceptible to damage when exposed to electrostatic discharge.

Look at some numbers. Suppose a particular CMOS inverter has total gate area $A = 3\ \mu\text{m}^2$ with $t_{ox} = 10$ nm. Each MOSFET SiO_2 insulator has a dielectric strength of about 10^7 V/cm, so the maximum sustainable gate voltage is $V_{max} = 10^7\ \text{V/cm} \times 10^{-6}\ \text{cm} = 10$ V. For the gate capacitance, we find

$$C_{in} = \frac{\epsilon_{ox} A}{t_{ox}} = \frac{3.9 \times 8.854 \times 10^{-14}\ \text{C/V-cm} \times 3 \times 10^{-8}\ \text{cm}^2}{10^{-6}\ \text{cm}} = 10\ \text{fF}$$

—we neglect any mode dependence for "back-of-the-envelope" calculations. Thus, the maximum allowable gate charge is $Q_{max} = C_{in} V_{max} = 10^{-13}$ C.

Meanwhile, the so-called **human body model** finds you electrically equivalent to a 100-pF capacitor in series with a 1.5-kΩ resistor as shown in Fig. 10.5. When you walk across a vinyl floor in a low-humidity room, you become "charged" to about 10,000 V. And even working at a bench in a moderately humid room tends to charge you to about 100 V. The charge Q_{hbm} that relates to these conditions is 10^{-6} C and 10^{-8} C, respectively. Discharge events are clearly hazardous to CMOS inputs nearby.

People who regularly work with CMOS parts often wear grounded wrist straps and special antistatic gowns to minimize ESD failures. CMOS-laden circuit boards are shipped in special plastic wrappers. All these precautions are no substitute for protection at the circuit level, a topic examined shortly. But first another problem.

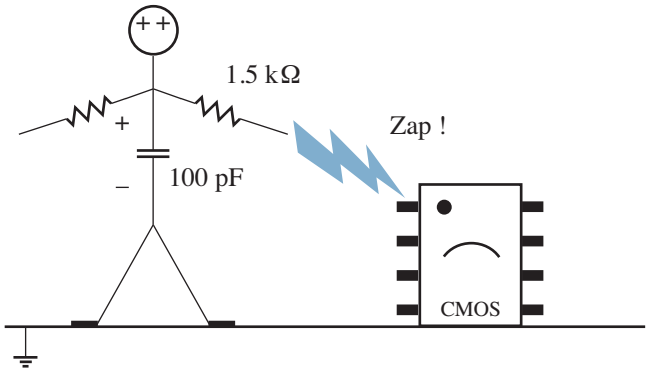


Figure 10.5: Human body model for electrostatic discharge events.

Output Problem: Latch-Up

Our discussion relating the static input-output characteristics of the CMOS inverter would be adequate if the circuit contained only discrete MOSFETs. The integrated implementation offers some unpleasant surprises.

Consider the CMOS inverter profile of Fig. 10.6 for an n-well process. As with any sequence of npn or pnp semiconductor layers, there exists the potential for bipolar transistor action, intended or not. In the case at hand, a vertically oriented pnp device (Q_1) has an emitter at the p^+ M_2 source, a base that is distributed in the n well, and a collector at the p substrate. The β_{F1} factor for Q_1 can be moderately large, perhaps of the order of 100, if the n well has shallow depth. A laterally oriented npn device (Q_2) has an emitter at the n^+ M_1 source, a base that is distributed in the p substrate, and a collector at the n well. Figure 10.6 has compressed horizontal scale. Thus, the Q_2 collector offers a low-profile target for electrons injected into a base region with a large volume (as well as a surface) for recombination. The β_{F2} factor for Q_2 is accordingly very small. The Q_1 base and the Q_2 collector are "connected" in the sense that they share the n-well region. There is no common circuit node that cries out for identification.

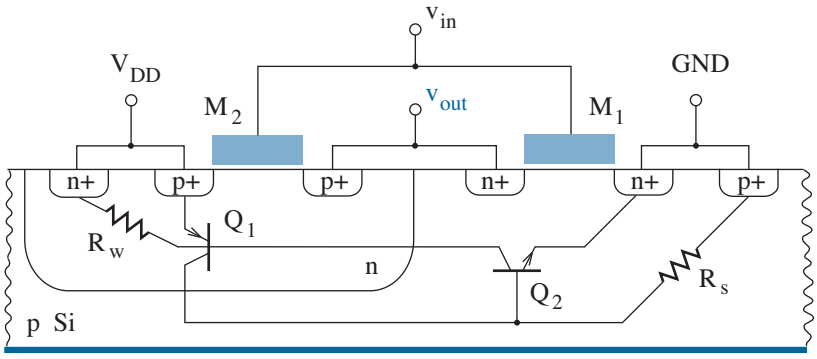


Figure 10.6: CMOS inverter profile with parasitic components.

Apart from the BJT devices, a parasitic resistor R_w "connects" from the Q_1 base to the n^+ substrate (body) terminal for M_2. The effective R_w value depends on the n-well doping concentration and a complex geometry, and it is typically a few kΩ. A similar parasitic resistor R_s "connects" from the Q_2 base to the p^+ well (body) terminal for M_1. This component has a value of tens or hundreds of ohms that depends upon the CMOS process. For simplicity, we ignore parasitic pnp and npn bipolar devices that engage the M_1 and M_2 drains as n^+ and p^+ emitters over differing ranges for v_{in}.

Figure 10.7 shows the applicable parasitic circuit model. With $V_{DD} > 0$, the Q_1 and Q_2 base-to-emitter junctions are both subject to forward bias. Nevertheless, the circuit can function with zero-value junction voltages and zero currents throughout. The circuit operation becomes unstable if one of the junction voltages is sufficient to support significant collector current. Suppose $R_s \to \infty$ and $R_w \to \infty$, and let i_{b1} suddenly increase from zero as a result of some transient process. Thus, $i_{b2} = i_{c1} = \beta_{F1} i_{b1}$. In turn, $i_{b1} = i_{c2} \to \beta_{F1}\beta_{F2} i_{b1}$. The result is a runaway condition if $\beta_{F1}\beta_{F2} > 1$. The eventual high-current circuit state is called **latch-up**.

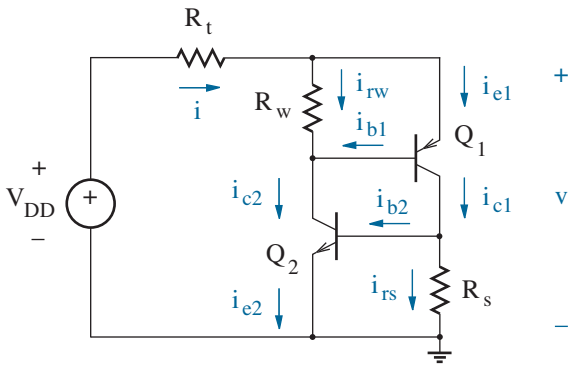


Figure 10.7: Circuit model featuring parasitic elements shown in Fig. 10.6. Resistor R_t is the Thevenin resistance associated with V_{DD}.

In a more quantitative general analysis, we have $i_{e1} = i - i_{rw}$ so that

$$i_{c1} = \alpha_{F1} i - \alpha_{F1} i_{rw} + I_{co1} \, . \tag{10.8}$$

Here, I_{co1} is a small "leakage" current for Q_1. Similarly, with $i_{e2} = i - i_{rs}$,

$$i_{c2} = \alpha_{F2} i - \alpha_{F2} i_{rs} + I_{co2} \, . \tag{10.9}$$

Finally,

$$i = i_{c1} + i_{c2} \, . \tag{10.10}$$

We eliminate i_{c1} and i_{c2} in Eqs. 10.8 to 10.10 to find

$$i = \frac{I_{co1} + I_{co2} - \alpha_{F1} i_{rw} - \alpha_{F2} i_{rs}}{1 - (\alpha_{F1} + \alpha_{F2})} \, . \tag{10.11}$$

The circuit enters latch-up with a demand for infinite current when

$$\alpha_{F1} + \alpha_{F2} = 1 \, . \tag{10.12}$$

This is equivalent to $\beta_{F1}\beta_{F2} = 1$.

We now direct attention to Fig. 10.8, which shows the current-voltage characteristic for the composite parasitic circuit of Fig. 10.7. The latch-up condition of Eq. 10.12 may apply at an intermediate voltage between zero and V_{DD}, or the latch-up may reflect a transient disturbance at $v = V_{DD}$. Either way, a current increase snaps v back to a **holding voltage**

$$V_H = |v_{be1}| + v_{ce2,sat} = v_{be2} + |v_{ce1,sat}| \tag{10.13}$$

with a value of about 0.9 V. The minimum **holding current** at V_H derives from Eq. 10.11 with $\alpha_{F1} + \alpha_{F2} \geq 1$. Specifically, with $i_{rw} = v_{be1}/R_w$ and $i_{rs} = v_{be2}/R_s$,

$$I_H = \frac{\alpha_{F1}\dfrac{v_{be1}}{R_w} + \alpha_{F2}\dfrac{v_{be2}}{R_s}}{(\alpha_{F1} + \alpha_{F2}) - 1} \; . \tag{10.14}$$

Equation 10.14 neglects small I_{co1} and I_{co2}, which were included in Eq. 10.11 to allow for latch-up behavior when open circuits substitute for R_w and R_s. The current dependence of α_{F1} and α_{F2} determines the composite-circuit behavior for $i > I_h$ since v_{be1} and v_{be2} are limited. A load-line analysis establishes the circuit operating point for a particular V_{DD} and R_t.

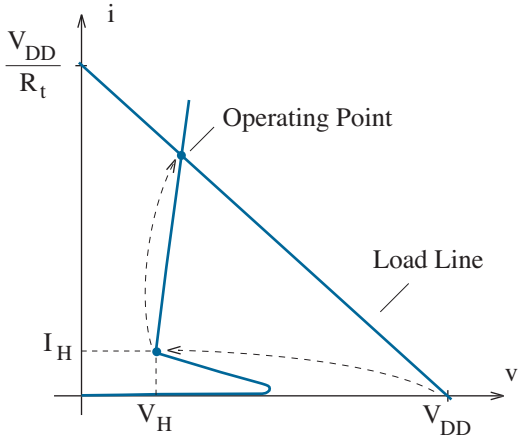


Figure 10.8: Graphical analysis for a latch-up circuit

.

If we assume a small difference between the operating voltage and V_H, latch-up requires $i \approx (V_{DD} - V_H)/R_t > I_H$. In turn, with Eq. 10.14,

$$\alpha_{F1} + \alpha_{F2} = \beta_{F1}\,\beta_{F2} > 1 + \frac{\left(\alpha_{F1}\dfrac{v_{be1}}{R_w} + \alpha_{F2}\dfrac{v_{be2}}{R_s}\right) R_t}{V_{DD} - V_H} \; . \tag{10.15}$$

Thus, latch-up is less easily achieved when R_w and R_s are small.

A **silicon controlled rectifier** exhibits the electrical model of Fig. 10.7, and it can be deliberately forced into latch-up to switch large currents in power systems. Latch-up is a problem for digital or analog CMOS circuits. Apart from loss of circuit functionality, power dissipation during latch-up can lead to overheating and permanent failure.

Several design techniques help to avoid latch-up in CMOS circuits:

- Minimize R_w and/or R_s. The parasitic well resistance R_w is reduced with the use of a retrograde well implant that increases the doping concentration near the bottom of the well. The substrate resistance R_s is also reduced in the presence of a deep heavily doped thin film. Both resistances are reduced through layout practices in which body contacts are made as close as possible to their respective MOSFET source regions. Numerous redundant contacts are advantageous.

- Degrade α_{F1} and/or α_{F2}. The use of deep wells limits the former, while large well-to-source spacings limits the latter.

- Avoid minority-carrier injection that can trigger latch-up. The n-well potential is V_{DD}, so p^+ sources should not connect to larger voltages. Similarly, the p-substrate potential is GND, so n^+ sources should not connect to negative voltages. While unlikely in the interior of a large digital system, these violations are often a problem at the periphery, particularly at output pin connections that can be subject to careless wiring or inductive transients. Protective circuitry at the output pins helps to limit unintended bias conditions (as examined shortly).

- Surround substrate transistors with heavily doped guard rings that absorb injected carriers before they can reach a nearby well collector. Guard rings limit circuit density, so they are generally reserved for peripheral regions with greatest vulnerability.

The ultimate solution to the the latch-up problem is the application of a **silicon-on-insulator (SOI)** technology in which n- and p-channel devices are completely isolated as shown in Fig. 10.9. Although generally expensive, SOI technology is also relatively immune to ionizing radiation.

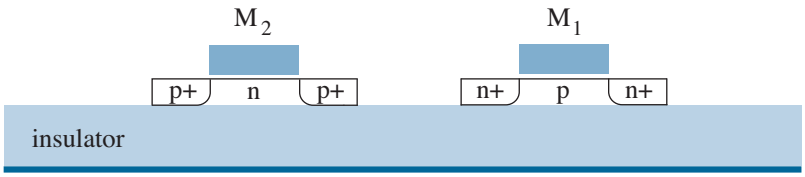


Figure 10.9: Silicon-on-insulator CMOS devices with electrical isolation.

Input/Output Protection

Two diode limiters are typically used to obtain on-chip protection from ESD and latch-up transients as shown in Fig. 10.10. The L_1 and L_2 inductors are associated with the bond wires that connect the silicon integrated circuit to its package, and R_t is the Thevenin resistance for the V_{DD} power supply. In the case of positive i_{esd}, the input (output) voltage is

$$v_x = V_{DD} + V_f + L_1 \frac{di_{esd}}{dt} + i_{esd}\, R_t \,, \tag{10.16}$$

where V_f is the characteristic diode turn-on voltage. Careful package design keeps L_1 small so that rapid changes in i_{esd} make little contribution to v_x. (Similar considerations apply to L_2 for negative i_{esd}.) The inclusion of an external bypass capacitor C_b near the V_{DD} pin suppresses the influence of R_t when i_{esd} is very large—see Chapter 1 for details.

Particularly vulnerable CMOS inputs (such as those connected to data communication lines) often have supplemental protection with an external IC containing a similar diode arrangement in relation to V_{DD} and ground. The package is specially designed to withstand ESD events of a particular magnitude, and the diodes are sometimes of the Shottky variety (small V_f). The MAX13202E (featuring pn diodes) is one example.

Note that a protected input becomes a low-impedance path to ground when $V_{DD} = 0$. Powered-down CMOS parts should not receive signals from other ICs. Unused inputs should be tied to V_{DD} or ground (as appropriate) through moderate-value resistors to avoid "banging" between the rails.

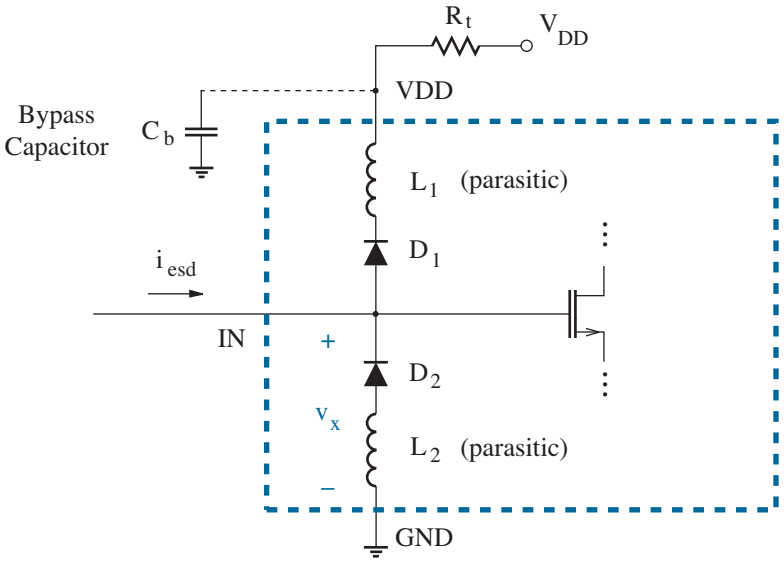


Figure 10.10: Protective input (output) circuitry for a CMOS IC package.

Output Options

CMOS digital circuits are available with three types of output:

True CMOS outputs have the form of the standard inverter (Fig. 10.1) and are primarily intended to connect to the inputs of other CMOS parts. Low-power operation and noise immunity are typical benefits. The sizes of the constituent n- and p-channel MOSFETs are usually much larger than CMOS circuitry elsewhere so that large output currents can be supported. This form of **output buffer** also accommodates multiple-load **fan-out**.

A true CMOS output in the LOW state that absorbs or *sinks* current as shown in Fig. 10.11 experiences a change in v_{out} according to the relation

$$I_{sink} = \frac{1}{2}K_n{}'\left(\frac{W}{L}\right)_n [\,2(V_{DD} - V_{Tn})v_{out} - {v_{out}}^2\,] \tag{10.17}$$

for resistive-mode operation. MOSFET M_1 is otherwise in saturation with v_{out} undefined to first order. The maximum allowable sink current is I_{OL}, the **low-level output current**. The corresponding v_{out} reflects V_{IL}.

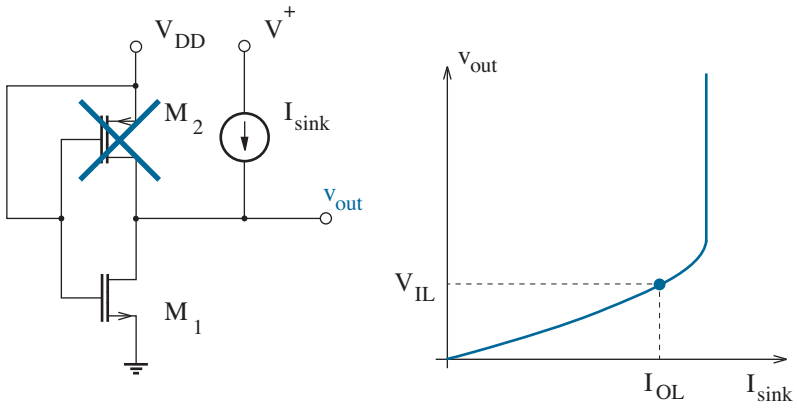


Figure 10.11: Output sink-current dependence for a true CMOS inverter.

A true CMOS output in the HIGH state that produces or *sources* current as shown in Fig. 10.12 experiences a change in v_{out} according to the relation

$$I_{source} = \frac{1}{2}K_p{}'\left(\frac{W}{L}\right)_p [\,2(V_{DD} - |V_{Tp}|)(V_{DD} - v_{out}) - (V_{DD} - v_{out})^2\,] \tag{10.18}$$

for resistive-mode operation. MOSFET M_2 is otherwise in saturation with v_{out} undefined to first order. The maximum allowable source current is I_{OH}, the **high-level output current**. The corresponding v_{out} reflects V_{IH}.

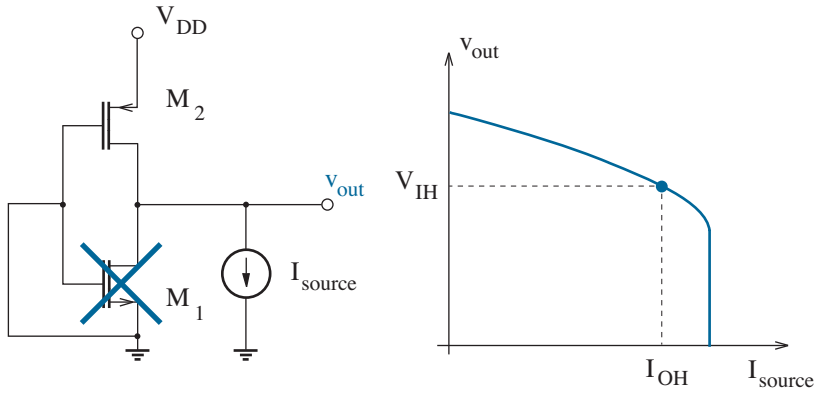


Figure 10.12: Output source-current dependence for a true CMOS inverter.

Open drain outputs are similar to the true CMOS configuration, but the p-channel MOSFET is removed. To achieve the proper output function, a **pull-up resistor** must be externally connected. This has two advantages: First, the supply voltage at the top of the pull-down resistor need not have the same value as V_{DD}. Much higher voltages are allowed (for special cases) if M_1 can accommodate large drain-to-source voltages. Second, the outputs of several open-drain gates can be tied to a common pull-down resistor as shown in Fig. 10.13. In this **wired OR** circuit, a mutual output falls LOW on demand of any individual gate—perhaps as a reflection of one of several logical conditions that warrant an interrupt signal for some digital process. Other gates that promote a HIGH output have their M_1 in the off state, so there is no adverse influence.

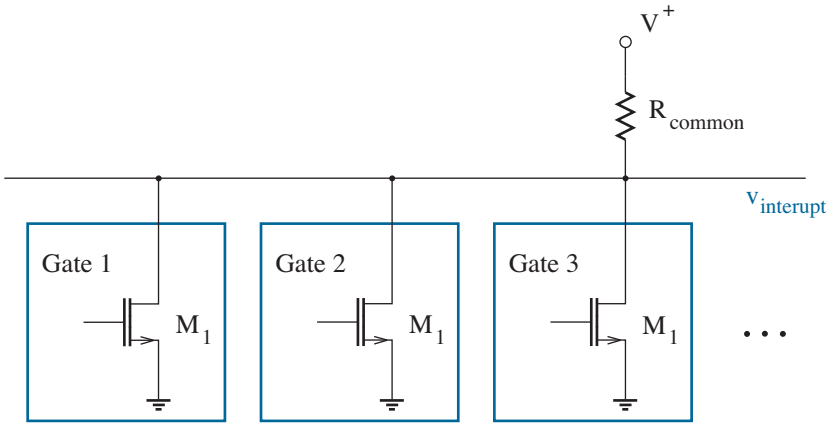


Figure 10.13: Open-drain CMOS gates in a wired OR configuration.

Tri-state outputs are similar to the true CMOS configuration, but two extra MOSFETs are included as shown in Fig. 10.14a. When $\overline{CS}$ is LOW, M_{1x} and M_{2x} are on so that the circuit functions as an ordinary inverter. However, when $\overline{CS}$ is HIGH, M_{1x} and M_{2x} are off, and the output is an effective open circuit. This is called a **high impedance (high Z)** state. The following truth table summarizes the tri-state action:

IN	$\overline{CS}$	OUT
0	0	1
1	0	0
—	1	High Z

Tri-state outputs are useful when CMOS parts need to communicate along a single "bus" line without interfering with one another. For example, having $\overline{CS1}$ LOW with the other chip-select inputs HIGH allows Part 1 to send serial data along the bus line to a distant receiver. Then setting $\overline{CS2}$ LOW with other controls HIGH transfers this "write" capability to Part 2.

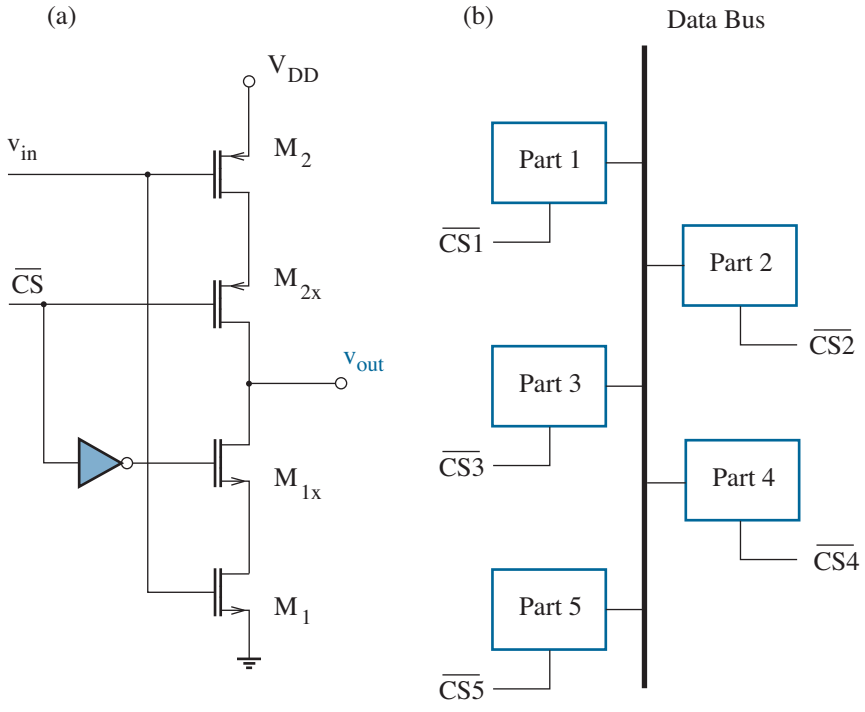


Figure 10.14: Tri-state outputs: (a) circuit implementation; (b) bus control.

10.2 Probing the CMOS Digital Interior

Having examined the salient features of the CMOS inverter, we are poised to implement a relentless hierarchy of digital logic as outlined in Chapter 1. Wired NOR and NAND gates easily extend to flip-flops, registers, counters, microprocessors, and beyond. Nevertheless, logic design at the circuit level requires optimization of performance factors such as speed and layout area. This section surveys some design alternatives and opportunities (memory).

> Reminder: Chapter 1 declared the "system" formalisms of VLSI design divergent to the objectives of this text. Look elsewhere to find realization methods and architectures for your next PC.

Combinational CMOS Logic

Now that we know the device implications of p-switch and n-switch as used for the NOR and NAND switching circuits of Fig. 1.29, it is straightforward to derive the CMOS versions of Fig. 10.15. The p-channel MOSFETs have their sources at or near V_{DD}, while the n-channel MOSFETs have their sources at or near ground. The output partitions the two types of switches. Finally, the requirement that the p- and n-channel devices have opposing "on" and "off" states indicates that their gates connect as paired inputs. Both circuits allow two arbitrary pairings by virtue of symmetry.

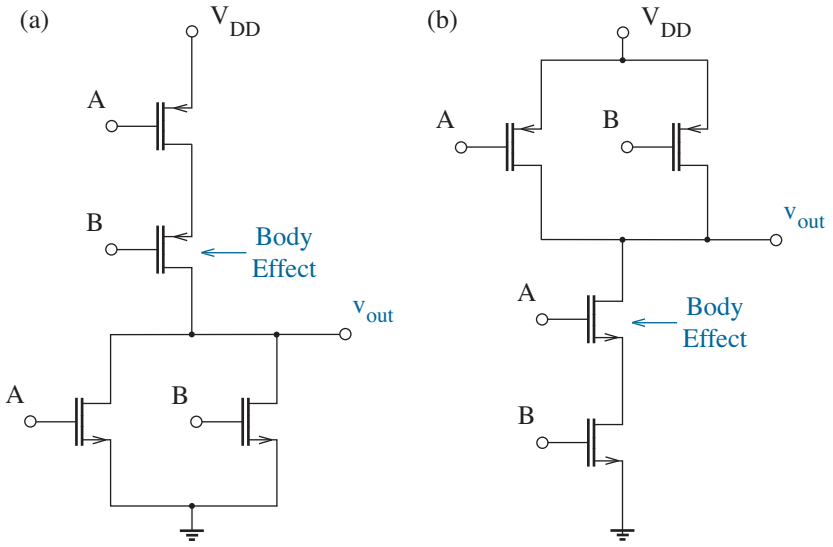


Figure 10.15: Elementary two-input CMOS gates: (a) NOR; (b) NAND. Wires connecting the A and B input pairs are not shown to preserve clarity. Marked transistors exhibit the body effect (see Problems 10.25 and 10.26).

Figure 10.16 shows the general form of CMOS combinational logic in which the output is either connected to V_{DD} through a **pull-up network** (PUN) containing p-channel MOSFETs or connected to ground through a **pull-down network** (PDN) containing n-channel MOSFETs. The latter satisfies the relation

$$OUT = \overline{F}\,, \tag{10.19}$$

where F is a logical expression involving one or more inputs. When $F = 1$, a combination of on-state n-channel switches has one or more ties to ground. For example, the NOR operation requires

$$OUT = \overline{A + B}\,. \tag{10.20}$$

Thus, the A and B n-channel MOSFET switches are connected in parallel—either one of them can be "on" to pull the output to ground. Conversely, the NAND operation requires

$$OUT = \overline{AB}\,. \tag{10.21}$$

Thus, the A and B n-channel MOSFET switches are connected in series—both of them must be "on" to pull the output to ground. To avoid conflicts, the PUN is a **dual circuit** in relation to the PDN. Series combinations of switches in one network correspond to parallel combinations in the other. Inputs apply to paired series and parallel devices in the two networks.

It is often helpful to use De Morgan's theorem so that logical expressions can be manipulated to take the form of Eq. 10.19. Specifically,

$$A + B = \overline{\overline{A}\,\overline{B}}\,, \tag{10.22}$$

$$AB = \overline{\overline{A} + \overline{B}}\,. \tag{10.23}$$

The circuitry needed to obtain an inverted input is seldom indicated.

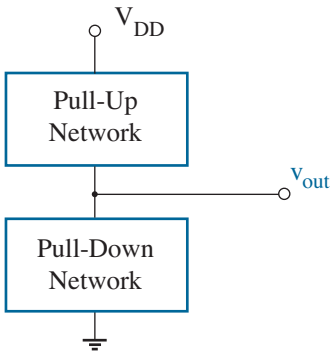


Figure 10.16: General circuit format for CMOS combinational logic.

Example 10.1

Sketch the CMOS circuit that implements $OUT = (A + B\overline{C})\overline{D}E$.

Solution

We apply two inversions and De Morgan's theorem to find

$$\begin{aligned} OUT &= \overline{\overline{(A + B\overline{C})\overline{D}E}} \\ &= \overline{\overline{A}\,\overline{B\overline{C}} + D + \overline{E}} \\ &= \overline{\overline{A}(\overline{B} + C) + D + \overline{E}}\,. \end{aligned}$$

Thus, with $F = \overline{A}(\overline{B} + C) + D + \overline{E}$, the PDN has two parallel n-channel switches controlled by D and $\overline{E}$ in parallel with a realization for $\overline{A}(\overline{B} + C)$. Figure 10.17 shows this PDN and its dual PUN. Note the input pairings.

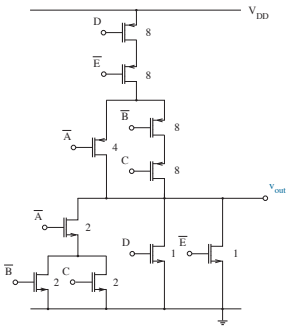


Figure 10.17: CMOS circuit solution for Example 10.1. Numbers next to the MOSFETs indicate optimum relative device widths (see Example 10.2).

Exercise 10.1 Sketch CMOS circuits that implement the following logic:

(a) $OUT = A\overline{B} + C$ (b) $OUT = A(B + \overline{C})$

Ans:

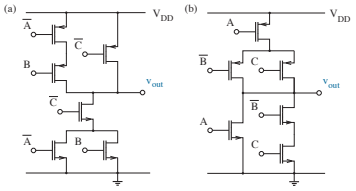


Figure 10.18: Circuits for Exercise 10.1.

Exercise 10.2 Determine logical functions for the circuits of Fig. 10.19.

Ans: (a) $OUT = \overline{\overline{A}(B + C)}$ (b) $OUT = A\overline{B} + \overline{A}B$ (exclusive OR)

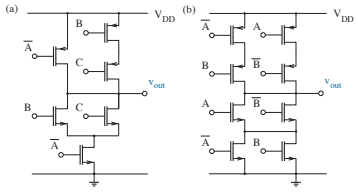


Figure 10.19: Circuits for Exercise 10.2.

Combinational CMOS logic design is complicated by the need to scale individual MOSFETs so that HIGH-to-LOW and LOW-to-HIGH timing transitions are optimized. We adopt the following rule:

> *Worst-case* "resistance" to ground through the PDN or to V_{DD} through the PUN is constant in relation to the CMOS inverter.

Sections 10.3 and 10.4 discuss the rationale for the CMOS inverter standard. For the sake of argument (and round numbers), we shall assume a design for which the p-channel MOSFET width is twice that of the n-channel device. This partially reflects the lower surface mobility for holes.

Consider the CMOS NOR gate. The n-channel MOSFETs in the PDN are in parallel, so they maintain unity width for worst-case "on" resistance. The p-channel MOSFETs in the PUN are in series, so they are each scaled to a width of four. The doubling of width cuts each "on" resistance in half, thereby preserving the total resistance.

Now consider the CMOS NAND gate. The reasoning used for the NOR gate scales each series-connected n-channel MOSFET to a width of two. The parallel-connected p-channel MOSFETs maintain widths of two.

The total integrated-circuit area consumed by a particular gate is roughly proportional to the sum of the relative n- and p-channel MOSFET widths—

ten and eight for the NOR- and NAND-gate configurations, respectively. Thus, the NAND logical function is slightly preferred for two inputs.

Example 10.2

Determine relative MOSFET widths for the CMOS circuit of Example 10.1. Assume a p/n width ratio of 2/1 for the inverter standard.

Solution

The n-channel MOSFETs controlled by D and $\overline{E}$ make direct connections to ground, so they maintain unity width. The worst-case paths to ground through the remaining n-channel devices have control signals $\overline{A}/\overline{B}$ or $\overline{A}/C$. Thus, the "on" resistance for each of these paths is regulated by doubling the transistor widths. The "on" resistance improves if the $\overline{B}$ and C devices are turned on simultaneously, but this is a *best-case* condition.

The worst-case path to V_{DD} has series-connected p-channel MOSFETs controlled by D, $\overline{E}$, $\overline{B}$, and C. Scaling each transistor to a width of eight cuts the individual contributions to the total "on" resistance by a quarter. The width of the p-channel MOSFET controlled by $\overline{A}$ only doubles to four.

Caution: Scale factors can change depending on the nature of an input. For example, input C could be a very infrequent "reset" that does not demand a quick response. Then C-MOSFET scaling may not be warranted.

No PUN Intended

Conventional CMOS gates tend to become unwieldly when the **fan-in** or number of inputs is large. For example, the four-input NOR gate requires four series-connnected p-channel MOSFETs with 4X-scaled relative widths as shown in Fig.10.20a. The integrated real estate and capacitance is large. Two clever design alternatives eliminate most of the PUN circuit baggage to improve speed and dynamic power.

A **pseudo nMOS** design replaces the PUN with a single p-channel load whose gate is connected to ground as shown in Fig. 10.20b.[1] When $i_2 = 0$ (as for open-circuit PDN conditions that warrant a HIGH output), M_2 has $v_{ds2} = 0$ so that $V_{OH} = V_{DD}$. When $i_2 \neq 0$ (conductive PDN conditions), v_{out} is LOW (presumably $\leq |V_{Tp}|$) and M_2 operates in saturation. In turn,

$$i_2 = \frac{1}{2} K_p{}' \left(\frac{W}{L}\right)_p (V_{DD} - |V_{Tp}|)^2 \,. \tag{10.24}$$

Meanwhile, an n-channel constituent in the PDN that is sufficiently large when "on" to make v_{out} LOW has

$$r_{on}{}^{-1} \approx K_n{}' \left(\frac{W}{L}\right)_n (V_{DD} - V_{Tn}) \tag{10.25}$$

since it operates in the resistive mode.

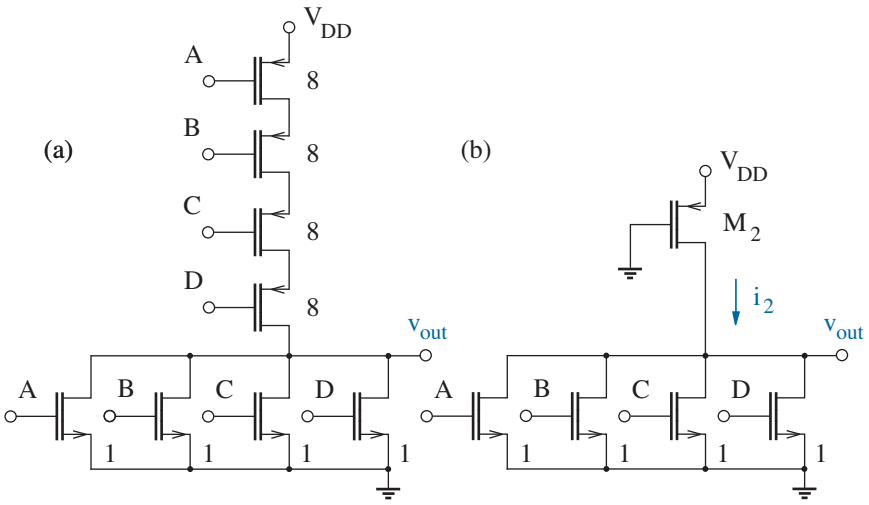


Figure 10.20: Four-input NOR gate: (a) CMOS; (b) pseudo-nMOS.

[1] Historically, "nMOS" refers to a similar circuit in which the load is a depletion-mode n-channel MOSFET with $v_{gs} = 0$. This technology has become obsolete.

With $V_{OL} \approx i_2\, r_{on}$, we apply Eqs. 10.24 and 10.25 to find

$$V_{OL} \approx \frac{K_p{}'(W/L)_p}{K_n{}'(W/L)_n}\left(\frac{V_{DD}}{2}\right) \tag{10.26}$$

subject to $V_{Tn} \sim |V_{Tp}| \ll V_{DD}$. The two transconductance parameters in this expression tend to support a small V_{OL} since $K_p{}'/K_n{}'$ is typically 0.4. Nevertheless, any V_{OL} other than zero exceeds that for conventional CMOS and reduces the low noise margin (N_{ML}). The LOW state is also achieved at the expense of static power, and LOW-to-HIGH transitions can be slow (see Problem 10.29).

A somewhat more complicated **dynamic logic** design eliminates static power dissipation that is characteristic of pseudo nMOS by turning off the p-channel load when the PDN is active. Figure 10.21 shows an example in which the load (M_2) and an additional n-channel "foot" (M_3) are controlled by clock signal ϕ. During the **pre-charge phase**, the clock input is LOW, M_2 is "on" and M_3 is "off" so that the inherent output capacitance C charges to V_{DD} regardless of the inputs to the PDN. During the subsequent **evaluation phase**, M_2 and M_3 change states so that the output becomes LOW if the PDN constituents promote a discharge path from C to ground. The capacitive output otherwise remains HIGH, but only for a limited time —parasitic discharge paths will eventually cause C to lose its charge. Thus, the clock period must be short compared to the parasitic discharge time. Periodic pre-charging *refreshes* a HIGH output state with long duration.

Unfortunately, dynamic-gate interconnections can offer invalid output states for certain input sequences unless each gate output is followed by a standard CMOS inverter—Problems 10.33 and 10.34 explore timing issues. This hybrid form of circuit is often called **domino logic** in recognition of a characteristic domino-like propagation of output states.

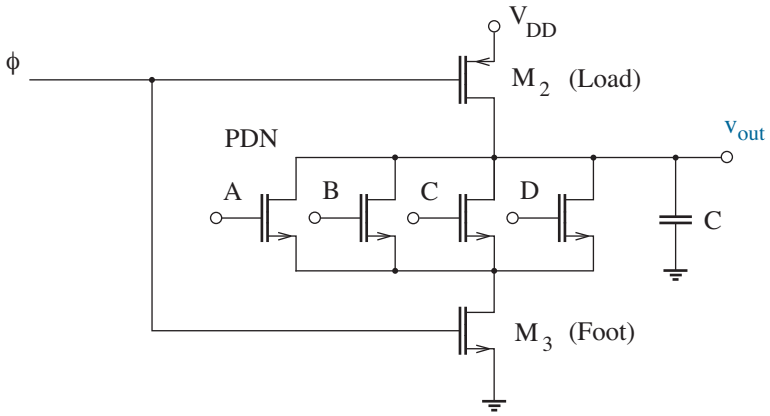


Figure 10.21: Four-input "dynamic" NOR gate.

Transmission-Gate Logic

Chapter 5 introduced a switch for analog signals in the form of two parallel-connected n- and p-channel MOSFETs with complementary HIGH or LOW gate voltages for on-off control. When used as a switch for digital signals, the circuit is called a **transmission gate** or **pass gate**. Figure 10.22 shows a common symbolic representation. The overlapping triangles are intended to convey the bidirectional character of the switching action.

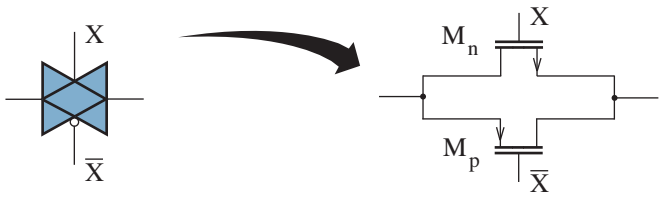


Figure 10.22: Transmission-gate switch with complementary MOSFETs. The switch is "on" when X is HIGH.

Transmission gates are noteworthy for the area-efficient implementation of certain logical functions such as XOR, the exclusive OR, which is used in adders and bit comparators. Figure 10.23 shows two XOR circuit examples. In the first, a LOW level for B forces the upper and lower transmission gates on and off, respectively, so that $y = A\overline{B}$. Conversely, a HIGH level for B reverses the switch states to produce $y = \overline{A}B$. Thus, $y = A \oplus B = A\overline{B} + \overline{A}B$ as a result of the wired-OR output connection. The transistor count is eight: four for the two transmission gates and four for the two inverters needed to derive $\overline{A}$ and $\overline{B}$. Compare with twelve transistors in support of Fig. 10.19b. In the second example, a LOW level for B turns on the transmission gate, but the inverter is dysfunctional with both transistors off since the power-supply voltages are reversed from their normal connections. Thus, $y = A\overline{B}$. However, a HIGH level for B turns off the transmission gate while properly configuring the inverter for $y = \overline{A}B$. The wired OR establishes $y = A \oplus B$. We now have a transistor count of six—no inverter for $\overline{A}$ is necessary.

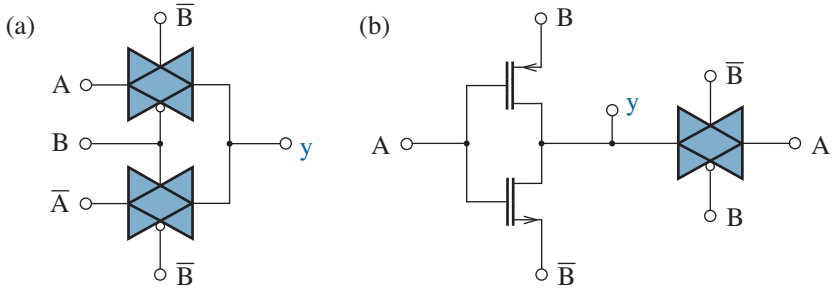


Figure 10.23: Two implementations of exclusive OR: $y = A \oplus B$.

Figure 10.24 shows a simple shift register featuring transmission gates and inverters. Data moves forward through one of the odd-numbered transmission gates to an inverter when the clock signal ϕ is HIGH, and inverted data moves forward though one of the even-numbered transmission gates for restorative inversion when the clock signal is LOW. The input capacitance of any inverter must be sufficient to hold a HIGH level when disconnected, an implication of a short clock period in comparison to the rate of discharge. The transistor count for one shift-register segment is eight. An ordinary D flip-flop made with CMOS combinational logic has eighteen (see Fig. 1.25).

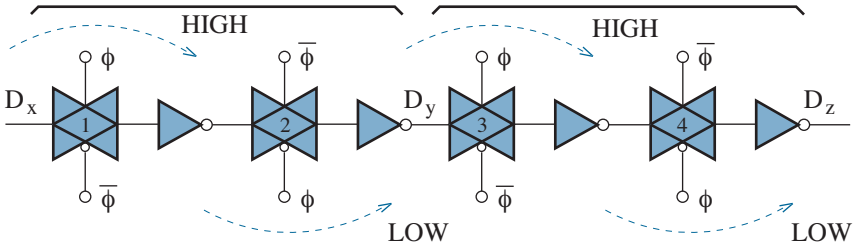


Figure 10.24: Transmission-gate shift register.

When designing a transmission gate, it is desirable to have comparable $K'W/L$ values for the n- and p-channel constituents so that each drain-to-source resistance r_{ds} is roughly the same when the switch is "on" and the potentials at either end are in the vicinity of $V_{DD}/2$. This typically implies a p-channel device that is about 2.5 times wider than the n-channel device on account of the difference in channel mobility. Upscaling the widths of both devices helps to reduce r_{ds}, but it adds capacitance that limits speed. Thus, minimally-sized devices are appropriate.

As noted in Chapter 5, n-channel r_{ds} increases as the potentials at either end of a switch increase from GND to V_{DD}, while p-channel r_{ds} increases as these potentials undergo the reverse transition. Nevertheless, the overall r_{ds} remains fairly constant through the parallel actions of the two devices. Individual body effects are not helpful (see Problem 10.38).

Having two transistors per switch requires more integrated-circuit area, not only from a second device but also from the inverter needed to obtain a complementary control signal. When the associated circuit requirements are too much of a burden, it is sometimes desirable to realize a switch that excludes the relatively large p-channel MOSFET. A single n-channel switch has no difficulty pulling a node voltage to a low potential, but it needs help for the other transition. Problem 10.40 explores one method, which avoids a complementary control signal for a p-channel assistant.

The preceding discussion shows that there are often numerous ways to implement logical functions, each with varying area requirement and speed. The successful integrated-circuit designer considers all options.

Memory

Wanderings through the CMOS interior are likely to engage special circuits that provide memory action, especially in microprocessors and computers. In what follows, we examine three major memory classes and, in two cases, the electronic devices that support them.

Static Random-Access Memory (sRAM)

Figure 10.25 shows a CMOS six-transistor static random-access memory (sRAM) cell. The MOSFET pairs M_{1L}, M_{2L} and M_{1R}, M_{2R} form separate inverters with cross-coupled inputs and outputs. Positive feedback assures that the node-voltage pair $[v_L,v_R]$ rest in [HIGH,LOW] or [LOW,HIGH] stable (static) states that indicate the memory-cell content. Nevertheless, the sRAM memory is *volatile*—the content is lost if V_{DD} falls to zero.

M_{3L} and M_{3R} serve as access transistors that connect to **bit lines** BL and $\overline{BL}$ so that particular memory content can be conveyed to or imposed from the outside environment. Together with an orthogonal **word line** WL that controls the access transistors, the bit lines specify a unique memory address in a typically gargantuan array of similar cells.

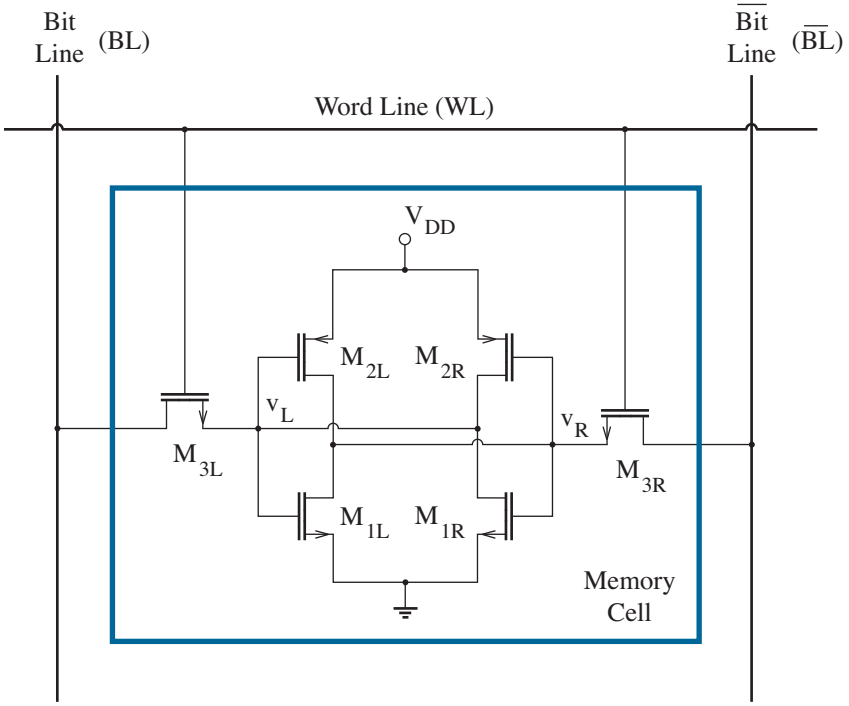


Figure 10.25: Six-transistor CMOS sRAM cell.

The sRAM **write** operation requires two steps: The BL and $\overline{BL}$ lines are initially set to (HIGH,LOW) or (LOW,HIGH) depending on the desired memory state. Then setting WL HIGH ties in the latch via M_{3L} and M_{3R}. If the latch state is consistent with the intended write, there is no change. Otherwise, the latch is forced into the opposite stable state.

The sRAM **read** operation is relatively complicated: The distributed capacitances associated with the BL and $\overline{BL}$ lines are charged to V_{DD} by setting pre-charge signal PC LOW so that transistors M_L and M_R are "on" at the top of a particular memory column (Fig. 10.26). A so-called balance transistor M_B is also "on" (through LOW PC) to ensure equal voltages on the two lines. Once the V_{DD} pre-charge condition has been established (as for the dynamic logic considered previously), PC is set HIGH, M_L, M_R, and M_B are turned "off", and the lines "float" at the same V_{DD} potential. In turn, WL is set HIGH so that M_{3L} and M_{3R} are "on". Whereas one side of the latch is already HIGH at V_{DD}, the associated line connection produces no voltage change. However, the other side of the latch is LOW, so the associated line connection produces a downward voltage transition. The large bit-line capacitance implies significant decay time. Nevertheless, there is no need to wait for the line to evolve to a continuous LOW state. A comparator-like **sense amplifier** (considered in Chapter 9) can detect a small differential voltage change between the changing and unchanging lines and interpret the sRAM memory content accordingly.

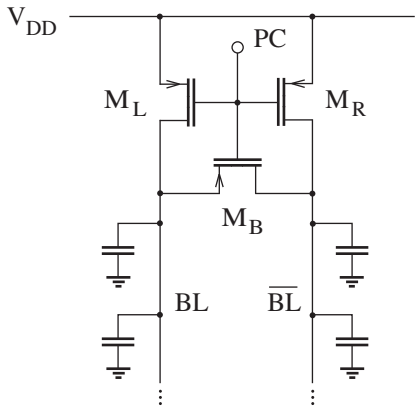


Figure 10.26: Precharge circuitry for sRAM BL and $\overline{BL}$ lines.

When designing M_{3L} and M_{3R}, one must ensure that the transistors are not so small as to suppress a necessary state transition during a write operation or so large as to disrupt the latch state during a read operation (see Problems 10.44 and 10.45).

sRAM memory is easily implemented, but each cell consumes precious integrated-circuit area with six transistors.

Dynamic Random-Access Memory (dRAM)

Compared with sRAM, a dynamic random-access memory (dRAM) cell is far more compact with only a single MOSFET and a capacitor localized to the intersection of a bit line and an orthogonal word line (Fig. 10.27). The penalty for area savings is a memory content that evolves with time unless it is periodically refreshed with the help of added peripheral circuitry.

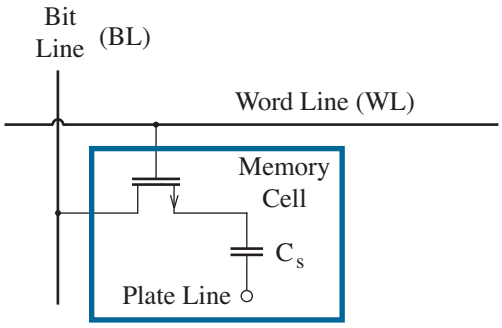


Figure 10.27: One-transistor dRAM cell.

As a starting point for understanding dRAM operation, we consider a primitive (but easily fabricated) version of the physical cell structure shown in Fig. 10.28. The n-channel MOSFET source and gate are connected to the bit and word lines, respectively, but there is no heavily doped drain region. Instead, the would-be drain merges with an MOS capacitor whose gate ties to a **plate line** at constant V_{DD}. The MOS capacitor is electrically isolated from the bit line when the word line (WL) is LOW.

Under equilibrium conditions, the MOS capacitor supports an inversion layer with areal charge density

$$Q_s \approx C_{ox}\left(V_{DD} - V_T\right), \tag{10.27}$$

where C_{ox} is the areal insulator capacitance and V_T is the threshold voltage. Moreover, the semiconductor surface potential near the inversion layer is $2\phi_f$ or roughly 0.6 V in a typical process (see the Chapter-5 Appendix). For reasons soon to be revealed, this capacitor state reflects a "zero".

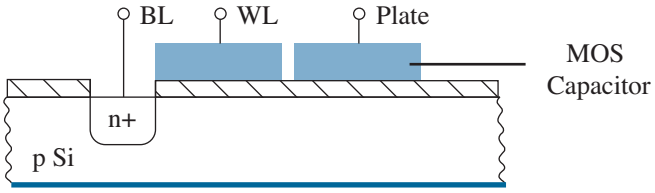


Figure 10.28: Primitive dRAM structure.

Under extreme *non-equilibrium* conditions, the MOS capacitor supports **deep depletion** without inversion ($Q_s = 0$), the semiconductor surface potential is slightly below V_{DD}, and the capacitor state reflects a "one". Eventually, a one evolves to a zero as the MOS capacitor seeks equilibrium through the creation of an inversion layer with appropriate areal density. Nevertheless, the characteristic time needed to equilibrate through thermal electron generation can be of the order of a second when isolation assures the absence of competing processes. Thus, there is ample opportunity to test the cell for a zero or a one and refresh the latter condition when necessary.

Figure 10.29 shows conditions that apply for dRAM **write** operations with $V_{DD} = 5$ V. The bit line is first set HIGH or LOW depending on the desired memory state. Then HIGH WL gives access to the MOS capacitor. For the case of a "one" (Fig. 10.29a), the n^+ source has a higher potential than the semiconductor surface beneath the plate, so electrons are drawn to the source, and none are left behind. For the case of a "zero" (Fig. 10.29b), the potential conditions are reversed, so electrons are removed from the source to contribute to the inversion layer in the MOS capacitor.

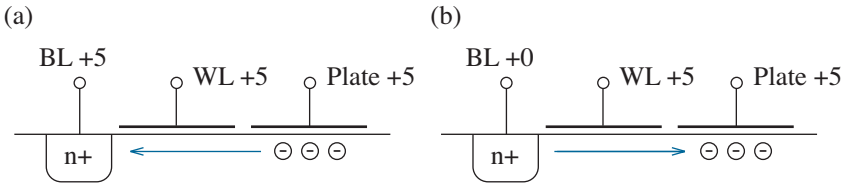


Figure 10.29: dRAM write operations: (a) one; (b) zero.

Figure 10.30 shows conditions that apply for dRAM **read** operations with $V_{DD} = 5$ V. The bit line is first precharged and allowed to "float" at $V_{DD}/2$ (2.5 V), and WL is set HIGH. For the case of a "one" (Fig. 10.30a), the semiconductor surface beneath the plate has a higher potential than the n^+ source, so electrons are removed from the source. The loss of negative charge effectively adds a positive charge, so the bit-line voltage *increases*. A comparator-like sense amplifier (Chapter 9) detects the voltage change. For the case of a "zero" (Fig. 10.30b), the potential conditions are reversed, electrons are added to the source, and the bit-line potential *decreases*.

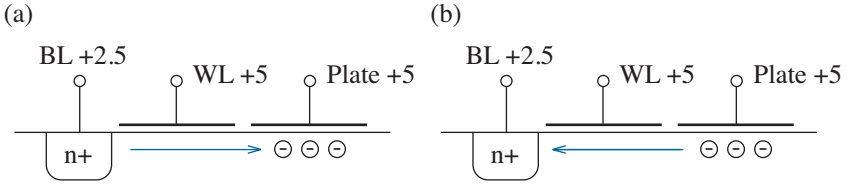


Figure 10.30: dRAM read operations: (a) one; (b) zero.

We leave it to the reader to show that the charge redistribution during a read operation leads to the bit-line voltage change (Problem 10.39)

$$\Delta v_{BL} = \frac{V_{DD}}{2}\,(2b-1)\,\frac{C_s}{C_s + C_{BL}}\,, \tag{10.28}$$

where b is the stored bit state (0 or 1), C_s is the MOS storage capacitance, and C_{BL} is the total bit-line capacitance. The latter tends to be large since the bit line spans a long column of identical dRAM cells. Thus, an analog Δv_{BL} signal is difficult to distinguish from additive noise when C_s is small. The problem is exacerbated as dRAM cells become smaller to accommodate demands for ever increasing memory density.

Numerous clever process technologies have been implemented to modify the primitive dRAM structure of Fig. 10.28 so that C_s is somehow increased. Perhaps the most successful applies a **trench capacitor** with large effective area as shown in Fig. 10.31. The storage element is formed by etching a deep recess with vertical sidewalls adjacent to the WL-contolled access transistor. The trench surface is then oxidized and filled with polysilicon to complete the capacitor. Such processes are invariably difficult and expensive.

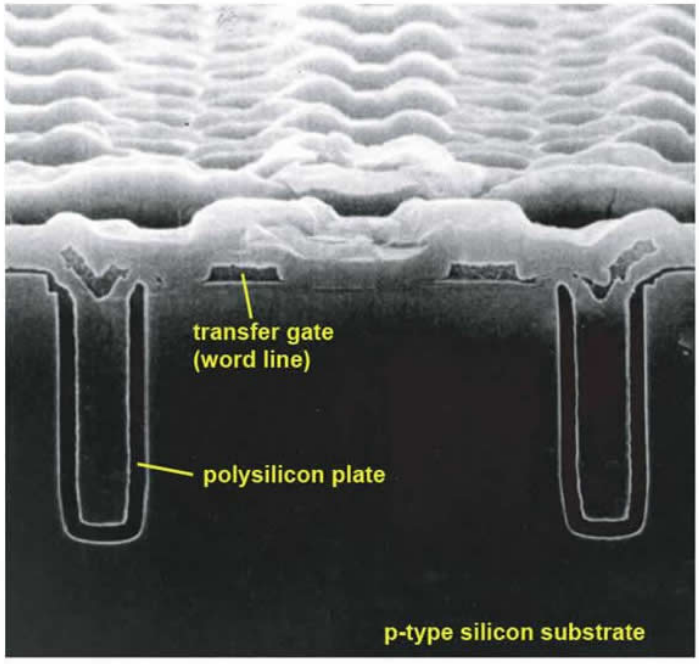


Figure 10.31: "Trench" capacitor profile for a compact dRAM technology. Courtesy of the Institute of Electrical and Electronics Engineers (IEEE).

Non-Volatile Memory

As noted previously, sRAM and dRAM lose their information content when V_{DD} is deliberately or otherwise set to zero. While there are several forms of **non-volatile memory** that are oblivious to power-supply interruptions, we accept **Erasable Programmable Read-Only Memory (EPROM)** as a representative in common use.

EPROM operation requires the use of a special type of MOSFET with two polysilicon gates as shown in Fig. 10.32. An upper **select gate** controls the MOSFET to produce the current-voltage characteristics of Chapter 5. The lower **floating gate** is encased in SiO_2 and positioned close to the silicon interface. There is no direct electrical connection.

In the **unprogrammed state**, the floating gate is nothing more than an internal conductive region in support of a moderate threshold voltage that depends on the series combination of the areal capacitance between the select gate and the floating gate and that between the floating gate and the silicon substrate. The **programmed state** is achieved by applying large voltages (typically around 20 V) for v_{gs} (of the select gate) and v_{ds}. All of the electrons in the channel region thus formed are attracted to the SiO_2 interface above it. However, some electrons are sufficiently energetic or "hot" to blast their way across the SiO_2 insulator into the floating gate. These electrons are subsequently trapped when v_{ds} and v_{gs} are reduced to lower levels, and the effect is to increase the threshold voltage significantly. An EPROM cell that is localized to a particular bit- and word-line address will either turn on or not depending on the status of its threshold voltage.

To erase an EPROM array, shine ultraviolet light through a window on the chip so that trapped electrons have sufficient energy to escape into the silicon substrate. This is a cumbersome and usually infrequent process (even though it can be repeated many times). So-called EEPROM systems are electrically erasable, but the details go beyond the scope of this text.

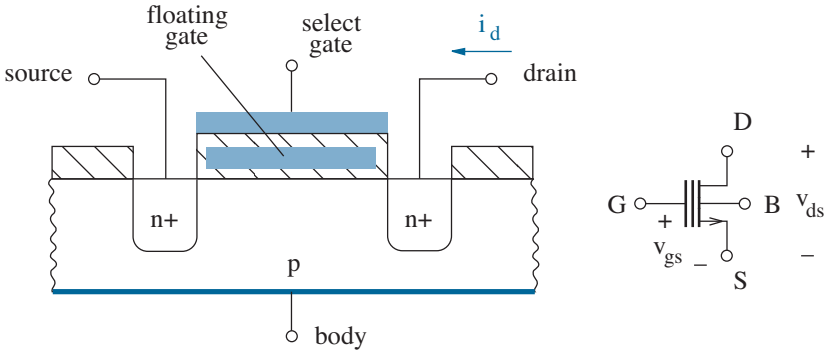


Figure 10.32: Floating-gate MOSFET for EPROM non-volatile memory.

10.3 Static CMOS Inverter Design

When designing to obtain a certain CMOS inverter transfer characteristic, perhaps the most obvious objective is the switching threshold voltage V_m that corresponds to the maximum negative slope in Region 3 of Fig. 10.2 — input changes above or below V_m yield significant (logical) output changes. Some texts designate V_m as the point on the transfer characteristic where $v_{out} = v_{in} = V_m$ to substantiate an easy measurement.

To estimate V_m, we equate the magnitudes of the M_1 and M_2 drain currents under saturation, no channel-length modulation, and $v_{in} = V_m$:

$$\underbrace{\frac{1}{2}\,K_n{}'\,(W/L)_n\,(V_m - V_{Tn})^2}_{i_{d1}\ \text{(saturation)}} = \underbrace{\frac{1}{2}\,K_p{}'\,(W/L)_p\,(V_{DD} - V_m - |V_{Tp}|)^2}_{-i_{d2}\ \text{(saturation)}} \,. \tag{10.29}$$

Then we equate the positive square root of the left-hand side with the positive square root of the right-hand side (the judicious root combination). In turn,

$$V_m = \frac{V_{DD} + \sqrt{K_R}\,V_{Tn} - |V_{Tp}|}{1 + \sqrt{K_R}} \,, \tag{10.30}$$

where

$$K_R = \frac{K_n{}'\,(W/L)_n}{K_p{}'\,(W/L)_p} \,. \tag{10.31}$$

A symmetric transfer characteristic features $V_{Tn} = |V_{Tp}|$ and $V_m = V_{DD}/2$. This requires $K_R = 1$, or

$$\frac{(W/L)_p}{(W/L)_n} = \frac{K_n{}'}{K_p{}'} = \frac{\mu_e\,C_{ox,n}}{\mu_h\,C_{ox,p}} \,. \tag{10.32}$$

Here, μ_e and μ_h are the low-field electron and hole surface mobilities and $C_{ox,n}$ and $C_{ox,p}$ are the areal n- and p-channel gate oxide capacitances. Areal capacitance parameters are generally equal, and μ_e/μ_h is about 2.5. Thus, the symmetry condition reflects $(W/L)_p\,/(W/L)_n \approx 2.5$.

Exercise 10.3 A particular CMOS technology features $V_{DD} = 3.3$ V, $V_{Tn} = +0.6$ V, and $V_{Tp} = -0.6$ V. Estimate the K_R factor and the p/n ratio of (W/L) values that establishes: (a) $V_m = 1.4$ V; (b) $V_m = 1.9$ V. Assume $\mu_e/\mu_h = 2.5$.

Ans: (a) $K_R = 2.64$, $(W/L)_p\,/(W/L)_n = 0.95$
(b) $K_R = 0.38$, $(W/L)_p\,/(W/L)_n = 6.58$

Determining V_{IL} or V_{IH} is conceptually straightforward yet analytically horrendous except for special-case conditions. Consider the voltage V_{IL}. Once again, we equate the magnitudes of the M_1 and M_2 drain currents, but we place M_1 in saturation and M_2 in the resistive mode. Specifically,

$$\underbrace{\frac{1}{2}\,K_n{}'\,(W/L)_n\,(v_{in}-V_{Tn})^2}_{i_{d1}\ \text{(saturation)}} = \tag{10.33}$$

$$\underbrace{\frac{1}{2}\,K_p{}'\,(W/L)_p\,[\,2\,(V_{DD}-v_{in}-|V_{Tp}|)(V_{DD}-v_{out})-(V_{DD}-v_{out})^2\,]}_{-i_{d2}\ \text{(resistive)}}\,.$$

(We are content to *estimate* V_{IL}, so we ignore channel-length modulation.) Careful calculus leads to dv_{out}/dv_{in}, which we set to -1. In turn,

$$v_{out}=\frac{1}{2}\,[\,(1+K_R)\,v_{in}-K_RV_{Tn}+|V_{Tp}|+V_{DD}\,]\,, \tag{10.34}$$

where K_R is given by Eq. 10.31. Finally, we substitute Eq. 10.34 back into Eq. 10.33 and solve to find $v_{in}=V_{IL}$. This is an obvious algebraic mess. Nevertheless, under symmetric conditions with $V_{Tn}=|V_{Tp}|$ and $K_R=1$,

$$V_{IL}=\frac{3V_{DD}+2V_{Tn}}{8}\,. \tag{10.35}$$

A similar procedure reveals V_{IH} subject to M_1 in the resistive mode and M_2 in saturation. With equal-magnitude drain currents,

$$\underbrace{\frac{1}{2}\,K_n{}'\,(W/L)_n\,[\,2\,(v_{in}-V_{Tn})\,v_{out}-{v_{out}}^2\,]}_{i_{d1}\ \text{(resistive)}} = \tag{10.36}$$

$$\underbrace{\frac{1}{2}\,K_p{}'\,(W/L)_p\,(V_{DD}-v_{in}-|V_{Tp}|)^2}_{-i_{d2}\ \text{(saturation)}}\,.$$

Then with $dv_{out}/dv_{in}=-1$,

$$v_{out}=\frac{1}{2K_R}\,[\,(1+K_R)\,v_{in}-K_RV_{Tn}+|V_{Tp}|-V_{DD}\,]\,. \tag{10.37}$$

Simultaneously solving Eqs. 10.36 and 10.37 for $v_{in}=V_{IH}$ seems formidable. Yet with $V_{Tn}=|V_{Tp}|$ and $K_R=1$,

$$V_{IH}=\frac{5V_{DD}-2V_{Tn}}{8}\,. \tag{10.38}$$

The symmetric noise margins are thus $NM_L=NM_H=3V_{DD}/8+V_{Tn}/4$.

The effects of channel-length modulation on V_m, V_{IL} and V_{IH} are minor (see Problem 10.53).

It should be obvious by now that the voltage levels V_m, V_{IL}, and V_{IH}, (and the noise margins, by extension), are functions of the n- and p-channel threshold voltages, V_{DD}, and GND. Process or system constraints tend to place all of these parameters beyond the realm of a lowly CMOS designer. The special voltage levels are also functions of a *single* K_R factor derived from an apparent plethora of geometric variables (two W's and two L's), which are mostly free for adjustment, and two transconductance parameters (K_n' and K_p'), which are not. Thus, in the typical case of a design that is directed toward V_m, one is forced to accept the other level outcomes.

Figure 10.33 shows the variation of voltages V_m, V_{IL}, and V_{IH} with K_R subject to $V_{Tn} = |V_{Tp}| = 0.6$ V and $V_{DD} = 3.3$ V. The variations are weak —note the logarithmic scale—as a consequence of the explicit square-root term in Eq. 10.30 and the square-root term in the quadratic formula needed to resolve either V_{IL} (Eqs. 10.33 and 10.34) or V_{IH} (Eqs. 10.36 and 10.37). This weakness helps to ensure *robust* designs with comparable electrical behavior despite statistical variations in actual vs. intended K_R values. Designs with $V_m < V_{DD}/2$ ($K_R > 1$) usually imply area efficiency with comparable n- and p-channel transistor sizes (equal sizes for $K_R \approx 2.5$). Designs with $V_m > V_{DD}/2$ ($K_R < 1$) usually imply area penalties with expensive p-channel transistor size requirements.

The noise-margin data in Fig. 10.33 show favorable high- and low-side conditions for $K_R > 1$ and $K_R < 1$, respectfully. The former is sometimes advantageous for CMOS interfaces to older logic families such as TTL.

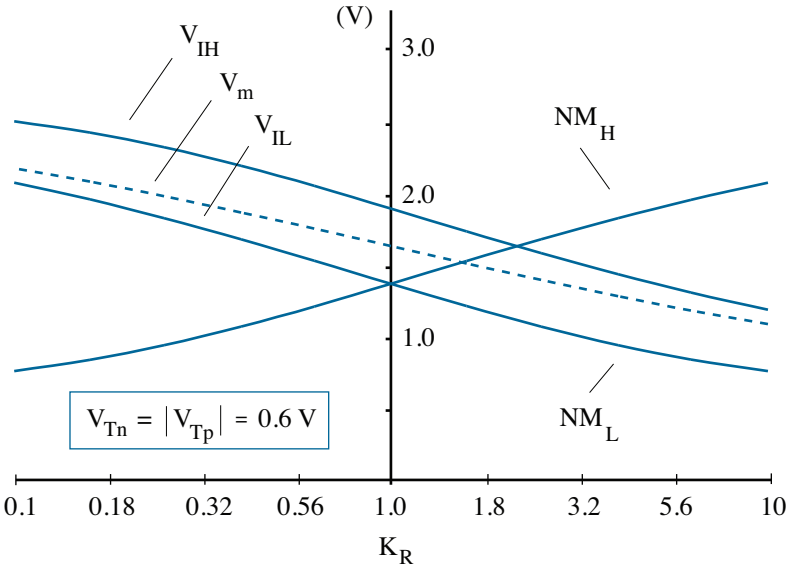


Figure 10.33: Special voltages V_m, V_{IL}, and V_{IH}, and noise margin data vs. parameter K_R (no channel length modulation, $V_{DD} = 3.3$ V).

Example 10.3

A CMOS inverter has the following transistor specifications:

n-channel — $K_n{}' = 50\ \mu\text{A/V}^2$, $V_{Tn} = +0.4$ V, $(W/L)_n = 2/1$
p-channel — $K_p{}' = 20\ \mu\text{A/V}^2$, $V_{Tp} = -0.8$ V, $(W/L)_p = 4/1$

Estimate the low and high noise margins by hand, then check with SPICE. Assume $V_{DD} = 3.3$ V.

Solution

Noise-margin parameters depend on K_R, so we apply Eq. 10.31 to find

$$K_R = \frac{50 \times 2}{20 \times 4} = 1.25\,.$$

That was easy—the remaining work is tedious and prone to errors.

To find V_{IL} we rewrite Eq. 10.33,

$$1.25\,(v_{in} - 0.4)^2 = 2\,(2.5 - v_{in})(3.3 - v_{out}) - (3.3 - v_{out})^2\,,$$

and Eq. 10.34,

$$\begin{aligned} v_{out} &= \frac{1}{2}\,[\,2.25\,v_{in} - 1.25(0.4) + 0.8 + 3.3\,] \\ &= 1.125\,v_{in} + 1.800\,. \end{aligned}$$

We substitute the second relation into the first to obtain

$$0.266\,{v_{in}}^2 + 4.250\,v_{in} - 5.050 = 0\,.$$

Of the two roots (1.111 and -17.09), we assign the former to V_{IL}. Then with $V_{OL} = 0$ V, $NM_L = V_{IL} - V_{OL} = 1.11$ V.

To find V_{IH} we rewrite Eq. 10.36,

$$1.25\,[\,2\,(v_{in} - 0.4)v_{out} - {v_{out}}^2\,] = (2.5 - v_{in})^2\,,$$

and Eq. 10.37,

$$\begin{aligned} v_{out} &= \frac{1}{2.5}\,[\,2.25\,v_{in} - 1.25(0.4) + 0.8 - 3.3\,] \\ &= 0.900\,v_{in} - 1.200\,. \end{aligned}$$

We substitute the second relation into the first to obtain

$$0.190\,{v_{in}}^2 + 3.040\,v_{in} - 5.480 = 0\,.$$

Of the two roots (1.635 and -17.64), we assign the former to V_{IH}. Then with $V_{OH} = 3.3$ V, $NM_H = V_{OH} - V_{IH} = 1.66$ V.

We can check our results with SPICE by running the following code, which sweeps the inverter input.

```
* VIL and VIH Analysis Code
* See Fig. 10.1 for circuit node numbers

Vdd      1      0      3.3
Vin      2      0      0
M1       3      2      0      0      MOSN   W=2u   L=1u
M2       3      2      1      1      MOSP   W=4u   L=1u

.model   MOSN   NMOS   ( KP=50u, VTO= +0.4 )
.model   MOSP   PMOS   ( KP=20u, VTO=  -0.8 )

.dc      Vin    0      3.3      0.01
.probe

.end
```

Plot v(3) to obtain the transfer characteristic, then superimpose constraint traces 1.125 * v(2) + 1.800 and 0.900 * v(2) - 1.200 as shown in Fig. 10.34. The two straight-line intercepts correspond to V_{IL} and V_{IH}, respectively. Noise margins follow from the low and high definitions.

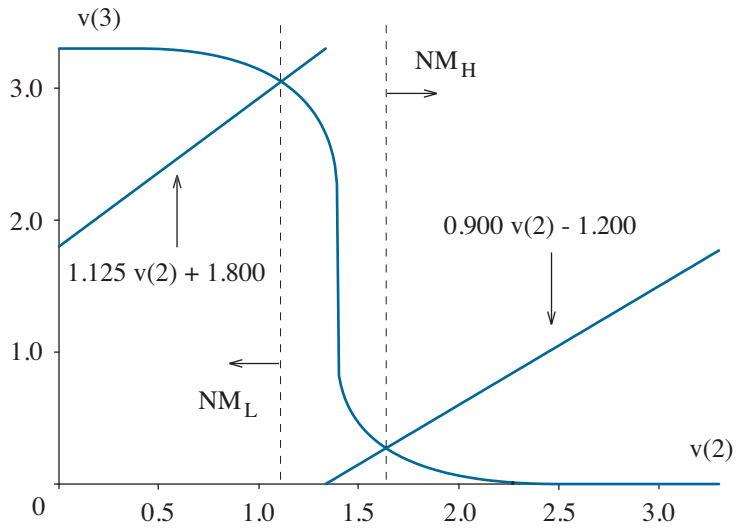


Figure 10.34: SPICE .probe results for Example 10.3.

Example 10.4

A CMOS inverter has the following transistor specifications:

n-channel — $K_n{}' = 50\ \mu\text{A/V}^2$, $V_{Tn} = +0.4$ V
p-channel — $K_p{}' = 20\ \mu\text{A/V}^2$, $V_{Tp} = -0.8$ V

Design to ensure $NM_H = 2.0$ V. Assume $V_{OH} = V_{DD} = 3.3$ V.

Solution

Our goal is to find K_R such that $V_{IH} = V_{OH} - NM_H = 1.3$ V. An algebraic solution that uses Eqs. 10.36 and 10.37 is just hopeless, so we look to SPICE. The following code essentially performs trial and error:

```
* Inverter Design for a Particular VIH
* See Fig. 10.1 for circuit node numbers

.param     KR=1

Vdd        1        0        3.3
Ein        2        0
+          VALUE = {(2*KR*v(3) + KR*0.4 - 0.8 + 3.3)/(1 + KR)}
M1         3        2        0        0        MOSN   W=1u    L=1u
M2         3        2        1        1        MOSP   W=1u    L=1u

.model     MOSN     NMOS     ( KP={KR*20u}, VTO= +0.4)
.model     MOSN     PMOS     ( KP=20u,      VTO= -0.8 )

.dc        PARAM    KR       0.2      3.0      0.01
.probe

.end
```

The code begins by declaring parameter KR, which is arbitrarily set to 1. Despite the actual $K_n{}'$ specification for the n-channel MOSFET, the SPICE value is set to KR*20u in the .model statement—note the curly brackets. Thus, the K_R definition (Eq. 10.31) is satisfied when the transistors have the same arbitrary 1-μm W and L dimensions. The Ein statement corresponds to a *dependent* voltage source at the inverter input whose value is a function of K_R and the output voltage through a rearranged Eq. 10.37. Specifically,

$$v_{in} = \frac{2K_R\, v_{out} + K_R V_{Tn} - |V_{Tp}| + V_{DD}}{1 + K_R}\,. \tag{10.39}$$

This forces $v(2) = v_{in} = V_{IH}$. The .dc statement sweeps KR.

Figure 10.35 show the resulting .probe display of v(2) vs. parameter KR. A quick manipulation of the cursor determines $V_{IH} = 1.3$ V for $K_R = 3.431$. In turn, $(W/L)_p \,/\, (W/L)_n = (K_n'/K_p') \,/\, K_R = 0.73$.

A simulation like that of Example 10.3 verifies the V_{IH} result.

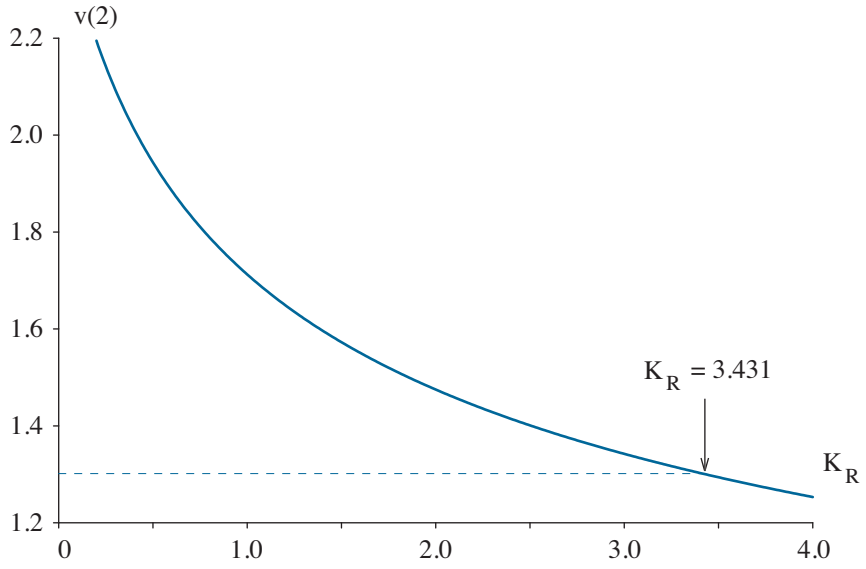


Figure 10.35: SPICE .probe results for Example 10.4.

The method of Example 10.4 easily extends to determining the K_R value that applies to a desired NM_L (through V_{IL}). The SPICE code is similar, but the value expression for the dependent voltage source derives through a rearranged Eq. 10.34. Specifically,

$$v_{in} = \frac{2\,v_{out} + K_R V_{Tn} - |V_{Tp}| - V_{DD}}{1 + K_R} \,. \tag{10.40}$$

The method can also determine K_R for a specific V_m subject to

$$v_{in} = v_{out} \,. \tag{10.41}$$

(This V_m procedure applies even for higher-level transistor models.)

If a K_R value is known, the constraint equations 10.36 and 10.37 reduce to the form $m\,v(3) + b$, where m and b are constants. The .param, .dc, and .probe statements are then replaced with the .op command., which directs SPICE to determine a particular operating point. The "bias solution" that corresponds to V_{IH} or V_{IL} appears as v(2) in the output file.

Velocity Saturation

The elementary long-channel drain-current expressions used to estimate V_m are hardly adequate for the deep sub-micron MOSFETs of modern times. In what follows, we examine the physical basis and circuit implications of **velocity saturation**, a mobility-limiting process in very short devices.

Chapter 2 showed that electrons and holes undergo numerous scattering events when subjected to an accelerating electric field. In turn, the collision-averaged drift velocity has magnitude

$$v = \mu_o \mathcal{E} \,, \tag{10.42}$$

where μ_o is a constant mobility. But drift velocity cannot grow indefinitely. As indicated in Fig. 10.36, it tends to "saturate" at v_{max}, which is roughly the average velocity of carriers in random thermal motion (about 10^7 cm/s). The actual limiting mechanisms do not concern us. We merely note that the large tangential electric fields needed to produce velocity saturation effects come into play as MOSFET channel lengths are substantially diminished.

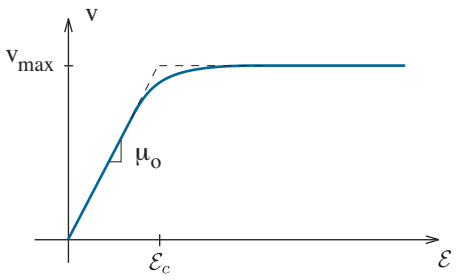


Figure 10.36: MOSFET carrier channel velocity as a function of tangential electric field. The linear portion of the curve defines low-field mobility.

An empirical expression for drift velocity that favors the smaller of two contributing factors takes the form

$$v = \mu_o \mathcal{E} \parallel v_{max} = \frac{\mu_o \mathcal{E}\, v_{max}}{\mu_o \mathcal{E} + v_{max}} \,. \tag{10.43}$$

So in consideration of Eq. 10.42, the **effective mobility** is

$$\mu_{eff} = \frac{\mu_o}{1 + \mathcal{E}/\mathcal{E}_c} \,, \tag{10.44}$$

where

$$\mathcal{E}_c = \frac{v_{max}}{\mu_o} \tag{10.45}$$

is the "critical" electric field (see Fig. 10.36). When $\mathcal{E} \ll \mathcal{E}_c$, Eq. 10.44 reduces to the low-field mobility (μ_o). And when $\mathcal{E} \gg \mathcal{E}_c$, $\mu_{eff}\,\mathcal{E} \approx v_{max}$.

Having obtained an expression for effective mobility, we can apply it to Eq. 5.2 to rederive the MOSFET drain current. Specifically,

$$i_d = \frac{q\mu_e W N_I(y)\,\dfrac{d\psi}{dy}}{1+\dfrac{1}{\mathcal{E}_c}\dfrac{d\psi}{dy}}\,, \tag{10.46}$$

subject to $|\mathcal{E}| = d\psi/dy$, and $\mu_o \to \mu_e$ for electron current. We rearrange and integrate over the interval [0,L] to obtain

$$\int_0^L i_d\left(1+\frac{1}{\mathcal{E}_c}\frac{d\psi}{dy}\right)dy = q\int_{\psi(0)}^{\psi(L)} \mu_e W N_I(\psi)\,d\psi\,. \tag{10.47}$$

The left-side integration yields $i_d L\,(1+v_{ds}/L\,\mathcal{E}_c)$. The right-side integration yields a previously encountered result with the help of Eq. 5.4 for $qN_I(\psi)$. Thus,

$$i_d = \frac{1}{2}K'\frac{W}{L}\frac{[2(v_{gs}-V_T)v_{ds}-{v_{ds}}^2]}{1+v_{ds}/\mathcal{E}_c\,L}\,, \tag{10.48}$$

where $K' = \mu_e C_{ox}$. Apart from the $(1+v_{ds}/\mathcal{E}_c\,L)$ term in the denominator, this is the elementary expression for resistive-mode drain current.

Despite a somewhat overused terminology, velocity "saturation" effects should not be confused with drain-current saturation, which occurs when $\partial i_d/\partial v_{ds} = 0$. When the latter definition is applied to Eq. 10.48, we find

$$v_{ds,sat} = (v_{gs}-V_T)\left[\frac{\sqrt{1+2\zeta}-1}{\zeta}\right]\,, \tag{10.49}$$

where

$$\zeta = \frac{v_{gs}-V_T}{\mathcal{E}_c\,L}\,. \tag{10.50}$$

For long-channel devices, $\zeta \ll 1$, the term in brackets in Eq. 10.49 is reduced to unity, and we recover the familiar expression for the saturation voltage. However, for very-short-channel devices with $\zeta \gg 1$, the bracketed term asymptotically approaches $\sqrt{2/\zeta}$, and

$$i_{d,sat} \approx C_{ox}W(v_{gs}-V_T)\,v_{max} \tag{10.51}$$

(see Problem 10.57). The result reflects a strip of channel charge that moves at the saturation velocity. As such, there is no L dependence.

Unfortunately, practical MOSFET circuits seldom experience $\zeta \gg 1$. For example, an n-channel device with $\mu_e = 300$ cm^2/V-s, $v_{max} = 1.5\times10^7$ cm/s, and $L = 0.2$ μm is characterized by $\mathcal{E}_c L = 1$ V. So at the rather large $v_{gs}-V_T = 4$ V, $\zeta = 4$, and the actual and asymptotic $v_{ds,sat}/(v_{gs}-V_T)$ values are 0.5 and 0.707, respectively. The difference is substantial for $i_{d,sat}$.

Our starting expression for effective mobility was an empirical relation, so we should not hesitate to make further approximations that make design calculations involving velocity saturation more palatable. If we restrict our attention to the range $0 < \zeta < 4$, some numerical experimentation reveals

$$v_{ds,sat} \approx (v_{gs} - V_T) \left[1 - \frac{1}{4} \sqrt{\frac{v_{gs} - V_T}{\mathcal{E}_c L}} \right] \tag{10.52}$$

as an asymptotic relation with less than 5% error in the range of interest. Equation 10.52 shows a fractional decrease in proportion to $\sqrt{v_{gs} - V_T}$.

To determine the saturated drain current, we need to substitute Eq. 10.49 back into Eq. 10.48. The result is a somewhat dyspeptic expression:

$$i_{d,sat} = \frac{1}{2} K' \frac{W}{L} (v_{gs} - V_T)^2 \underbrace{\frac{2\,(1 + \zeta - \sqrt{1 + 2\zeta}\,)}{\zeta^2}}_{F(\zeta)} , \tag{10.53}$$

which is barely tractable for design (although quite acceptable for analysis). Nevertheless, we move forward by postulating an empirical expression for $F(\zeta)$ in Eq. 10.53 of the form

$$F(\zeta) \approx \frac{1}{1 + a\,\zeta} , \tag{10.54}$$

where a is an adjustable constant. Subsequent curve-fitting analysis shows that $a = 0.8$ yields less than 5 % error over $0 \leq \zeta \leq 4$ (see Problem 10.58). In turn,

$$i_{d,sat} \approx \frac{1}{2} K' \frac{W}{L} (v_{gs} - V_T)^2 \left[1 + \frac{0.8\,(v_{gs} - V_T)}{\mathcal{E}_c L} \right]^{-1} . \tag{10.55}$$

For the case of the numbers cited previously with $\zeta = 4$, the effect of velocity saturation is to reduce the saturated drain current by a factor of about 4.2 with respect to the long-chanel value. (Compare: $1/F = 4.0$ in Eq. 10.53.) Moreover, the square-law relationship between $i_{d,sat}$ and $v_{gs} - V_T$ is moved closer to a linear dependence. This has prompted some designers to use

$$i_{d,sat} \approx P \frac{W}{L} (v_{gs} - V_T)^{\alpha} , \tag{10.56}$$

where P is constant and $1 < \alpha < 2$. However, Eq. 10.55 is more versatile.

Parameter measurements related to velocity saturation are demanding. One must first determine accurate values for W and L, which are invariably smaller than those that define MOSFET size during the fabrication process. Then one finds $\mathcal{E}_c L$ through a curve fit to the resistive-mode drain current (Eq. 10.48) subject to varying v_{ds}. In turn, $v_{max} = \mathcal{E}_c / \mu_o = \mathcal{E}_c\, C_{ox} / K'$.

As noted previously, channel-length modulation does little to vary V_m. Not so for velocity saturation effects—both the electron and hole mobilities are ultimately compromised by a velocity limit, and the effective electron mobility is the first to suffer as a consequence of its larger low-field value. Thus, K_R decreases for a fixed $(W/L)_p/(W/L)_n$ ratio, and V_m increases. One can effect a happy compensation for this increase by *reducing* $(W/L)_p$, thereby achieving area savings for the CMOS inverter.

To demonstrate the influence of velocity "saturation" with regard to the switching threshold, we rewrite the expression for saturated (non-resistive) drain current (Eq. 10.53). Specifically,

$$i_{d,sat} = K' \left(\frac{W}{L}\right) (\mathcal{E}_c L)^2 \left(1 + \zeta - \sqrt{1+2\zeta}\right), \tag{10.57}$$

where $\zeta = (v_{gs} - V_T)/\mathcal{E}_c L$ and $\mathcal{E}_c = v_{max}/\mu$. To design for V_m, we equate the drain currents of the n- and p-channel MOSFETs:

$$\begin{aligned} K_n{}' \left(\frac{W}{L}\right)_n (\mathcal{E}_{cn} L_n)^2 \left(1 + \zeta_n - \sqrt{1+2\zeta_n}\right) = \\ K_p{}' \left(\frac{W}{L}\right)_p (\mathcal{E}_{cp} L_p)^2 \left(1 + |\zeta_p| - \sqrt{1+2|\zeta_p|}\right). \end{aligned} \tag{10.58}$$

In turn,

$$K_R = \left(\frac{\mathcal{E}_{cp} L_p}{\mathcal{E}_{cn} L_n}\right)^2 \frac{1 + |\zeta_p| - \sqrt{1+2|\zeta_p|}}{1 + \zeta_n - \sqrt{1+2\zeta_n}}. \tag{10.59}$$

While less memorable than $K_R = 1$, this design expression is easily applied.

Exercise 10.4 A particular CMOS technology features $V_{DD} = 3.3$ V, $V_{Tn} = +0.6$ V, $V_{Tp} = -0.6$ V, $\mu_e = 500$ cm^2/V-s, and $\mu_h = 200$ cm^2/V-s. The saturation velocity for electrons or holes is 8×10^6 cm/s. Determine the K_R factor and the p/n ratio of W values that establishes $V_m = V_{DD}/2$ with: (a) $L_n = L_p = 2.0$ μm; (b) $L_n = L_p = 0.4$ μm.

Ans: (a) $K_R = 1.16$, $W_p/W_n = 2.16$

(b) $K_R = 1.48$, $W_p/W_n = 1.69$

We pause to note that velocity saturation can be important for analog MOSFET circuits. However, economic and technical pressures that force reduced channel lengths are less prevalent, and there is no compelling reason to seek confounding device behavior—daring and adventure are ill advised. Short-channel analog issues typically reflect channel-length modulation.

Example 10.5

A 0.25-μm CMOS process supported by the TSMC corporation features the SPICE .model parameters below (available to the public at www.mosis.org). Design an inverter for this process subject to $V_m = V_{DD}/2 = 1.25$ V.

(These data reflect the TSMC SCN025 CMOS technology, run #T51V.)

.model CMOSN NMOS (LEVEL=7 VERSION=3.1 TNOM=27 **TOX=5.7E-9** XJ=1E-7 NCH=2.3549E17 **VTH0=0.3635554** K1=0.470132 K2=1.129691E-3 K3=1E-3 K3B=2.5800217 W0=1E-7 NLX=1.931237E-7 DVT0W=0 DVT1W=0 DVT2W=0 DVT0=0.357209 DVT1=0.393627 DVT2=-0.5 **U0=306.116512** UA=-1.296859E-9 UB=2.710698E-18 UC=5.809324E-11 **VSAT=1.412988E5** A0=1.8397149 AGS=0.3636088 B0=-4.218354E-9 B1=-1E-7 KETA=-7.361109E-3 A1=3.479364E-4 A2=0.4444299 RDSW=157.9959537 PRWG=0.5 PRWB=0.0670201 WR=1 WINT=0 LINT=0 XL=0 XW=-4E-8 DWG=-1.715449E-8 DWB=4.995911E-9 VOFF=0.0986976 NFACTOR=1.332776 CIT=0 CDSC=2.4E-4 CDSCD=0 CDSCB=0 ETA0=5.532952E-3 ETAB=4.461256E-4 DSUB=0.0388315 PCLM=1.7948938 PDIBLC1=0.8172639 PDIBLC2=2.655381E-3 PDIBLCB=-0.1 DROUT=1 PSCBE1=7.566033E10 PSCBE2=5.241901E-10 PVAG=0 DELTA=0.01 RSH=4.3 MOBMOD=1 PRT=0 UTE=-1.5 KT1=-0.11 KT1L=0 KT2=0.022 UA1=4.31E-9 UB1=-7.61E-18 UC1=-5.6E-11 AT=3.3E4 WL=0 WLN=1 WW=0 WWN=1 WWL=0 LL=0 LLN=1 LW=0 LWN=1 LWL=0 CAPMOD=2 XPART=0.5 CGDO=5.69E-10 CGSO=5.69E-10 CGBO=1E-12 CJ=1.712621E-3 PB=0.9893888 MJ=0.461402 CJSW=4.036457E-10 PBSW=0.99 MJSW=0.318511 CJSWG=3.29E-10 PBSWG=0.99 MJSWG=0.318511 CF=0 PVTH0=-8.821935E-3 PRDSW=-10 PK2=3.733519E-3 WKETA=7.630624E-3 LKETA=-4.174821E-3)

.model CMOSP PMOS (LEVEL=7 VERSION=3.1 TNOM=27 **TOX=5.7E-9** XJ=1E-7 NCH=4.1589E17 **VTH0=-0.5615767** K1=0.6277876 K2=6.04884E-3 K3=0 K3B=11.23612 W0=1E-6 NLX=4.366735E-9 DVT0W=0 DVT1W=0 DVT2W=0 DVT0=3.1553567 DVT1=0.8622271 DVT2=-0.1296736 **U0=106.4695414** UA=1.206201E-9 UB=1E-21 UC=-9.65823E-11 **VSAT=2E5** A0=1.1521919 AGS=0.1997754 B0=8.346514E-7 B1=5E-6 KETA=8.162172E-3 A1=0.0294564 A2=0.3670593 RDSW=1.075749E3 PRWG=0.2127437 PRWB=-0.2443102 WR=1 WINT=0 LINT=2.119569E-8 XL=0 XW=-4E-8 DWG=-3.761827E-8 DWB=2.479315E-9 VOFF=-0.15 NFACTOR=0.7468997 CIT=0 CDSC=2.4E-4 CDSCD=0 CDSCB=0 ETA0=0.1143872 ETAB=-0.0263705 DSUB=0.72148 PCLM=1.1285403 PDIBLC1=2.091296E-3 PDIBLC2=1.136509E-6 PDIBLCB=-1E-3 DROUT=4.0027E-3 PSCBE1=8E10 PSCBE2=5.711638E-10 PVAG=7.917891E-3 DELTA=0.01 RSH=3.3 MOBMOD=1 PRT=0 UTE=-1.5 KT1=-0.11 KT1L=0 KT2=0.022 UA1=4.31E-9 UB1=-7.61E-18 UC1=-5.6E-11 AT=3.3E4 WL=0 WLN=1 WW=0 WWN=1 WWL=0 LL=0 LLN=1 LW=0 LWN=1 LWL=0 CAPMOD=2 XPART=0.5 CGDO=6.91E-10 CGSO=6.91E-10 CGBO=1E-12 CJ=1.925168E-3 PB=0.99 MJ=0.466148 CJSW=3.421303E-10 PBSW=0.525292 MJSW=0.287428 CJSWG=2.5E-10 PBSWG=0.525292 MJSWG=0.287428 CF=0 PVTH0=5.020164E-3 PRDSW=2.2959891 PK2=3.008098E-3 WKETA=0.0311891 LKETA=-7.792658E-3)

Don't panic. With subsequent understanding, a secure career awaits.

Solution

As noted in Chapter 5, an entire text is needed to do justice to the so-called BSIM3 Level-7 SPICE parameters. Nevertheless, four of them are familiar and sufficient to support Eq. 10.59 in a first-order design for a particular switching threshold voltage. These parameters are highlighted in **bold** and rounded from their generally shameless precision as follows:

MOSFET Parameter	Level-7 Keyword	n-channel Value	p-channel Value
Gate Oxide Thickness (m)	TOX	5.7×10^{-9}	5.7×10^{-9}
Threshold Voltage (V) (Long, Wide Device, $v_{bs} = 0$)	VTH0	0.364	-0.562
Low-field Mobility (cm^2/V-s)	U0	306.1	106.5
Carrier Saturation Velocity (m/s)	VSAT	1.413×10^5	2×10^5

The "critical" electric field ($\mathcal{E}_c$) applicable to short-channel MOSFETs is the ratio of the carrier saturation velocity to the low-field mobility. Thus, with $L_n = L_p = 0.25$ μm, we obtain $\mathcal{E}_{cn}L_n = 1.15$ V and $\mathcal{E}_{cp}L_p = 4.70$ V. Then Eq. 10.59 yields $K_R = 0.885$. The n- and p-channel transistors have the same TOX, so $(W/L)_p/(W/L)_n = \mu_e/(\mu_h K_R) = 3.25$.

Fig. 10.37 shows the simulated inverter transfer characteristic subject to $W_n = 1.00$ μm, $W_p = 3.25$ μm, and 0.25-μm minimum channel lengths. The extracted V_m value is 1.22 V, just under the 1.25-V target.

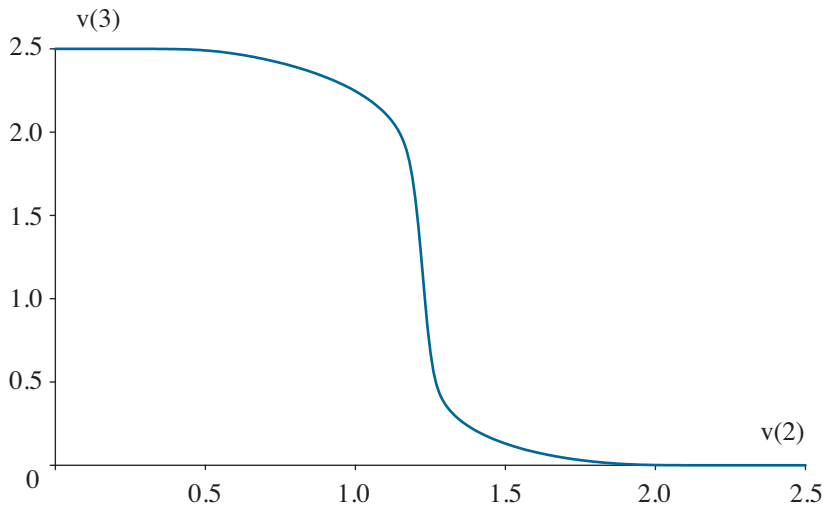


Figure 10.37: Inverter characteristic for Example 10.5 ($V_{DD} = 2.5$ V).

10.4 Dynamic CMOS Inverter Design

Our discussion of speed begins with the **ring oscillator** shown in Fig. 10.38. This pathetic little circuit attempts to establish a set of consistent logical conditions at the nodes around the ring, and the effort is endlessly repetitive —recall Sisyphus moving his stone—since the number of inverters is odd. Nevertheless, the ring oscillator demonstrates the dynamics of individual inverter switching in a practical environment: each inverter responds to the output of a predecessor, whereupon it "drives" the input of a successor.

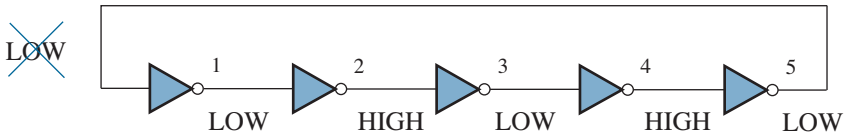


Figure 10.38: Five-stage ring oscillator. The logical conditions of the open ring are never consistent upon closure.

Figure 10.39 shows a set of voltage waveforms for a five-stage oscillator. Inverters are often designed subject to switching thresholds at $V_m = V_{DD}/2$, so it is reasonable to focus on two characteristic times:

- The **high-to-low propagation delay** (t_{PHL}) is the time difference between the instance when an output falls below 50% V_{DD} and that when the corresponding input rises above 50% V_{DD}.
- The **low-to-high propagation delay** (t_{PLH}) is the time difference between the instance when an output rises above 50% V_{DD} and that when the corresponding input falls below 50% V_{DD}.

The so-called **propagation delay** (t_P) is the average of t_{PHL} and t_{PLH}.

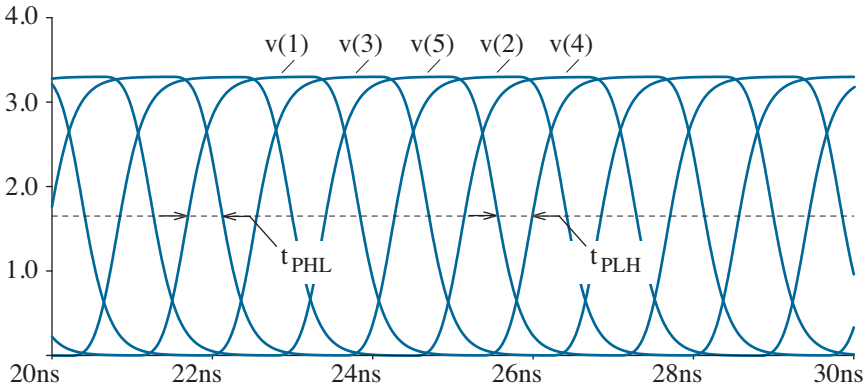


Figure 10.39: Node-voltage waveforms for a five-stage ring oscillator.

Pedants beware. Propagation delays are measured, NEVER calculated. For example, an inspection of the input/output transitions that apply to one of the inverters highlighted in Fig. 10.39 yields $t_{PHL} = t_{PLH} = 430$ ps. When confining interest to the average propagation delay,

$$t_P = \frac{\tau_o}{2N}\,, \tag{10.60}$$

where τ_o is the ring-oscillator period and N is the number of inverter stages. The 2 factor reflects two N-fold sets of logical transitions during a period. As expected, t_P is consistent with $\tau_o = 4.30$ ns in the preceding example.

Exercise 10.5 The TSMC #T51V CMOS process of Example 10.5 cites 31-stage ring oscillators with (a) 193.11 MHz and (b) 252.84 MHz operating frequencies given supplies $V_{DD} = 3.3$ V and $V_{DD} = 2.5$ V, respectively. Determine the average propagation delays.

Ans: (a) $t_P = 83.5$ ps (b) $t_P = 63.8$ ps

In what follows, we treat the ring oscillator as a vehicle for the fulfillment of three design objectives relating to speed:

- Balanced high-to-low and low-to-high propagation delays.

CMOS inverters are highly non-linear, so it is impossible to formulate useful analytic expressions for t_{PHL} or t_{PLH} subject to realistic input conditions. In particular, we avoid the ponderous exercise of determining the unit-step inverter response (despite an existing closed-form solution) since step inputs are seldom available and the results have misleading design implications. Nevertheless, simulation experiments applied to the ring oscillator provide a rough guide for relating K_R to the propagation delays so that adjustments can be made in consideration of a prior selection for V_m.

- Optimized n- and p-channel transistor widths for minimum delay.

The K_R value that forces balanced propagation delays does not generally imply minimum average delay, circuit area, or dynamic power consumption. Further deliberations are required to guide a more complicated tradeoff.

- Proper logic-element sizing for specific driver/load combinations.

Having established inverter sizing for optimum ring-oscillator performance, one can show that the measured average propagation delay determines a **delay unit** that applies in predictable degrees to arbitrary logical circuits. The formalism of "logical effort" is a means for determining appropriate transistor sizes in higher-level digital circuits (NAND, NOR, XOR, etc.). Logical-effort procedures go beyond the scope of this text.

Our ring oscillator is assumed to have a layout consistent with Fig. I.22 as well as the capacitance model of Fig. 10.40. Apart from C_{bsn1} and C_{bsp1}, which are subject to short circuits, all of the capacitances combine to form a single capacitance at the output of the first inverter. Specifically,

$$C_1 = \underbrace{C_{i2n} + C_{i2p} + C_{wire}}_{\text{load}} + \underbrace{C_{bdn1} + C_{bdp1}}_{\text{parasitic}} . \tag{10.61}$$

Here, $C_{i2n} + C_{i2p}$ reflects the total input capacitance of the second inverter. It includes gate-to-source and gate-to-drain contributions that indicate the extent of the poly excursions over the n and p active MOSFET regions. Both factors depend on the transistor operational mode (see Table 5.3), and the latter are enhanced through an output-to-input feedback process that is comparable to the Miller effect discussed in Chapter 8. Nevertheless, we assume that these complexities average to make $C_{i2n} + C_{i2p}$ constant. Parameter C_{wire} is a small interconnect capacitance with contributions from Metal 1 and the poly "wire" leading to the gates in the second inverter. Together with $C_{i2n}+C_{i2p}$, it presents a load to the first inverter—the added connection of similar inverters is presumed to increase the load for the first inverter in proportion to $C_{i2n} + C_{i2p} + C_{wire}$. Parameters C_{bdn1} and C_{bdp1} are drain-to-body capacitances that relate to the sizing of the active regions. Both are parasitic since they limit the inverter speed independently from what is connected to it, and both are highly non-linear with output voltage. We treat them as constant for now and correct for the non-linearities later.

For the layout of Fig. I.22, the inverters are closely packed and C_{wire} is small compared to the other capacitances, which approximately scale with MOSFET width. Thus, in consideration of Eq. 10.31,

$$C_1 \approx C_n \, (1 + \frac{K_n{}'/K_p{}'}{K_R}) \, , \tag{10.62}$$

where C_n associates with the n-channel MOSFETs. A similar capacitance applies at each inverter output node in a ring-oscillator loop.

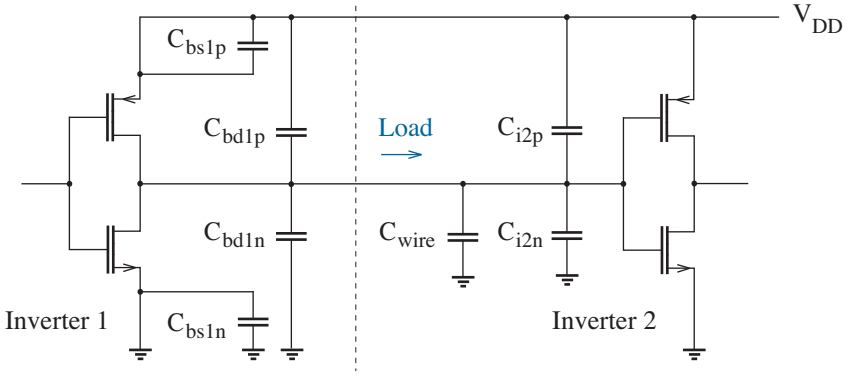


Figure 10.40: Capacitance model for the CMOS inverter layout of Fig. 9.18

Relative Propagation Delays

Figure 10.41 shows a five-stage CMOS ring oscillator circuit for simulation. Although many more stages are typically used in actual process-evaluation experiments, five stages (but not three) assure sufficiently accurate results. The equivalent C_x components combine capacitive attributes at the output of a particular stage, the input of the following stage, and the "wire" that connects them. Equations 10.61 and 10.62 give the specific form.

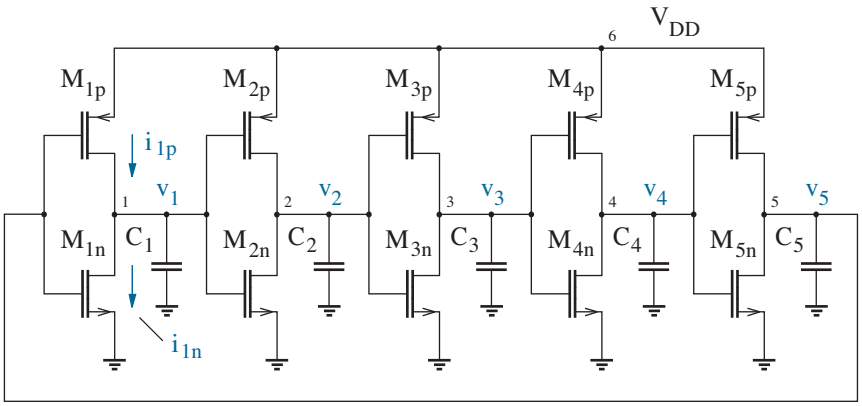


Figure 10.41: Five-stage ring oscillator circuit for simulation experiments.

Before we turn to SPICE, it will be constructive to examine the time scale for the two propagation delays and their principal governing factors. The output voltage for the first inverter satisfies the equation

$$C_n \, (1 + \frac{K_n{}'/K_p{}'}{K_R}) \frac{dv_1}{dt} = i_{1p} \, (V_{DD} - v_5, V_{DD} - v_1) - i_{1n} \, (v_5, v_1) \, , \quad (10.63)$$

where i_{1p} and i_{1n} reflect the drain currents for M_{1p} and M_{1n}, respectively. The functional arguments distinctly involve the gate-to-source and gate-to-drain voltages for the two MOSFETs. An arbitrary N-stage ring oscillator has inverter output voltages that lag behind their corresponding inputs by the fractional period

$$\zeta = \frac{N+1}{2N} \quad (10.64)$$

so that the lag delays are $0.6\,\tau_o$ for $N = 5$. In turn, with v_1 lagging v_5,

$$C_n \frac{dv_1}{dt} = \frac{i_{1p} \, [\, V_{DD} - v_1(t - \zeta\tau_o), V_{DD} - v_1(t) \,] - i_{1n} \, [\, v_1(t - \zeta\tau_o), v_1(t) \,]}{1 + \dfrac{K_n{}'/K_p{}'}{K_R}} \, . \quad (10.65)$$

This first-order non-linear discrete-delay differential equation looks benign, but it is impossible to solve in closed form.

Equation 10.65 assumes a form with *dimensionless* variables as follows: Let $i_{1p}(\ ) = K_p{}'(W/L)_{1p}V_{DD}{}^2\,\mathbf{F_{1p}}(\ )$, $i_{1n}(\ ) = K_n{}'(W/L)_{1n}V_{DD}{}^2\,\mathbf{F_{1n}}(\ )$, $\mathbf{v_1} = v_1/V_{DD}$, and

$$\mathbf{t} = \frac{t}{K_n{}'(W/L)_{1n}V_{DD}/C_n}\,. \tag{10.66}$$

Then upon substitution,

$$\frac{d\mathbf{v_1}}{d\mathbf{t}} = \frac{\mathbf{F_{1p}}\,[\,1-\mathbf{v_1}(\mathbf{t}-\zeta\tau_o), 1-\mathbf{v_1}(\mathbf{t})\,] - K_R\,\mathbf{F_{1n}}\,[\,\mathbf{v_1}(\mathbf{t}-\zeta\tau_o), \mathbf{v_1}(\mathbf{t})\,]}{K_R + K_n{}'/K_p{}'}\,. \tag{10.67}$$

Equation 10.67 tells us that ring-oscillator timing results that are scaled according to Eq. 10.66 depend only on the right-hand-side terms K_R and ζ, the functionality of $\mathbf{F_{1p}}$ and $\mathbf{F_{1n}}$ (off-state, resistive, and saturation modes), and scaled MOSFET parameters such as V_T/V_{DD}, λV_{DD}, and $\mathcal{E}_c L/V_{DD}$.

On to the simulations.

```
* Five-Stage Ring Oscillator Circuit

VDD     6       0       3.3
M1P     1       5       6       6       MOSP
M1N     1       5       0       0       MOSN
C1      1       0       Cmodel  1p
M2P     2       1       6       6       MOSP
M2N     2       1       0       0       MOSN
C2      2       0       Cmodel  1p
M3P     3       2       6       6       MOSP
M3N     3       2       0       0       MOSN
C3      3       0       Cmodel  1p
M4P     4       3       6       6       MOSP
M4N     4       3       0       0       MOSN
C4      4       0       Cmodel  1p
M5P     5       4       6       6       MOSP
M5N     5       4       0       0       MOSN
C5      5       0       Cmodel  1p

.param  KR=1
.model  MOSP    PMOS    (KP={1m/KR}, VTO=-0.5)
.model  MOSN    NMOS    (KP=1m, VTO=+0.5)
.model  Cmodel  CAP     (C={1 + 2.5/KR})

.ic     v(1)=0
.tran   1p      550n    500n    10p
.probe

.end
```

The SPICE code has *arbitrary* values for $K_n'(W/L)_n$ (1 mA/V^2), V_{DD} (3.3 V), and C_n (1 pF) that scale the time results according to Eq. 10.66—the τ_{PLH}/τ_{PHL} ratio is immune. In contrast, the dimensionless MOSFET thresholds of $\pm 0.5\ \text{V}/V_{DD}$ are buried on the right-hand side of Eq. 10.67 and exhibit particular influence. Capacitors C_1 through C_5 all vary according to Eq. 10.62 with K_R defined using the .param statement and $K_n'/K_p' = 2.5$. The factor C in the CAP .model statement multiplies the 1-pF value for C_n. Note the K_R dependence of $K_p'(W/L)_p$ in the PMOS .model statement.

The .ic statement sets v(1) = 0 as an initial condition prior to oscillation. The .tran statement specifies a 1-ps initial step size and a 10-ps maximum step size for numerical time integration. Whereas the oscillator needs time for steady-state behavior, the first 500 ns of operation are not displayed. The results from 500 ns to 550 ns are indicative.

Figure 10.42 shows simulated τ_{PLH}/τ_{PHL} ratios as a function of $\sqrt{K_R}$ for balanced threshold voltages (± 0.25 V, ± 0.5 V, and ± 0.75 V). The unity ratio for $K_R = 1$ (as when $V_m = V_{DD}/2$) indicates equal LOW-HIGH and HIGH-LOW propagation delays, a reflection of similar pull-up and pull-down MOSFET capability over an oscillation cycle. The data also suggest

$$\frac{\tau_{PLH}}{\tau_{PHL}} \approx \sqrt{K_R} \tag{10.68}$$

in the neighborhood of $K_R = 1$. This is a useful *empirical* result.

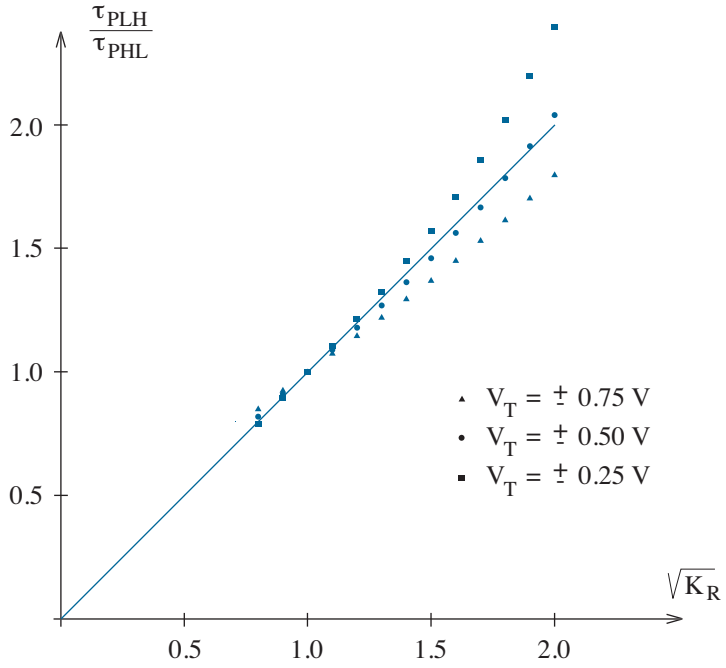


Figure 10.42: Simulation data for an elementary five-stage ring oscillator.

Design for Balanced Propagation Delays

So far, the unity-K_R condition for balanced propagation delays is restricted. It does not allow for unequal threshold-voltage magnitudes, departures from the elementary theory of MOSFET operation (as with velocity saturation), or non-linear capacitive effects. Extensions are welcome.

A simple model for inverter timing has the HIGH-to-LOW propagation delay inversely proportional to ${K_n}'(W/L)_n$, the "pull-down" capability of the n-channel MOSFET. An additional dependence on some power of K_R accounts for the gradual turn-off of the p-channel MOSFET. Thus,

$$\tau_{PHL} \sim \frac{{K_R}^{\alpha}}{{K_n}'(W/L)_n} \,. \tag{10.69}$$

Similarly, for the LOW-to-HIGH propagation delay,

$$\tau_{PLH} \sim \frac{{K_R}^{-\alpha}}{{K_p}'(W/L)_p} \,. \tag{10.70}$$

In turn,

$$\frac{\tau_{PLH}}{\tau_{PHL}} \sim \frac{{K_n}'(W/L)_n}{{K_p}'(W/L)_p}\,{K_R}^{-2\alpha} = {K_R}^{1-2\alpha} \,. \tag{10.71}$$

The empirical τ_{PLH}/τ_{PHL} ratio near $K_R = 1$ varies as ${K_R}^{1/2}$, so $\alpha = 1/4$. It follows that

$$\tau_{PHL} \sim [\,{K_n}'(W/L)_n\,]^{-3/4}\,[\,{K_p}'(W/L)_p\,]^{-1/4} \tag{10.72}$$

and

$$\tau_{PLH} \sim [\,{K_p}'(W/L)_p\,]^{-3/4}\,[\,{K_n}'(W/L)_n\,]^{-1/4} \,. \tag{10.73}$$

For an inverter with $v_{in} = V_{DD}/2$ and $v_{out} > V_{DD}/2$ ($v_{out} < V_{DD}/2$), the elementary n-channel (p-channel) drain current varies in proportion to $(V_{DD}/2 - |V_T|)^2$. Then with a conjecture that focuses on the -3/4 exponent in Eqs. 10.72 and 10.73 to suggest a "net" pulling power, we propose

$$K_R \left(\frac{V_{DD}/2 - V_{Tn}}{V_{DD}/2 - |V_{Tp}|} \right)^{3/2} = 1 \tag{10.74}$$

as a modified condition for balanced propagation delays when $V_{Tn} \neq |V_{Tp}|$. The adjustment is limited to one dominant device in deference to tests.

Test 1: $V_{Tn} = 0.4$ V, $V_{Tp} = -0.6$ V, $V_{DD}/2 = 1.65$ V.

$$K_R = (1.05/1.25)^{3/2} = 0.77. \text{ SPICE: } \tau_{PLH}/\tau_{PHL} = 1.005.$$

Test 2: $V_{Tn} = 0.7$ V, $V_{Tp} = -0.3$ V, $V_{DD}/2 = 1.65$ V.

$$K_R = (1.35/0.95)^{3/2} = 1.69. \text{ SPICE: } \tau_{PLH}/\tau_{PHL} = 0.991.$$

While two tests hardly cover all possible conditions, the correction factor in Eq. 10.73 does remarkably well for reasonable threshold-voltage pairings. An effective adjustment for the effects of velocity saturation is more elusive, and we are arguably naive in seeking it. Nevertheless, $v_{in} = V_{DD}/2$ implies

$$i_{d,sat} = K' \left(\frac{W}{L}\right) (\mathcal{E}_c L)^2 \left(1 + \zeta - \sqrt{1+2\zeta}\right) , \tag{10.75}$$

where

$$\zeta = \frac{V_{DD}/2 - |V_T|}{\mathcal{E}_c L} . \tag{10.76}$$

The proposed empirical condition for balanced propagation delays is

$$K_R \left(\frac{\mathcal{E}_{cn} L_n}{\mathcal{E}_{cp} L_p}\right)^{3/2} \left(\frac{1 + \zeta_n - \sqrt{1+2\zeta_n}}{1 + \zeta_p - \sqrt{1+2\zeta_p}}\right)^{3/4} = 1 . \tag{10.77}$$

Given the form of Eq. 10.75, the K_R condition of Eq. 10.77 incorporates the effect of unequal threshold voltages.

Test 3: $\mathcal{E}_{cn} L_n = 2$ V, $\mathcal{E}_{cp} L_p = 5$ V, Test-1 conditions

$K_R = 0.94$. SPICE: $\tau_{PLH}/\tau_{PHL} = 1.023$.

Test 4: $\mathcal{E}_{cn} L_n = 1$ V, $\mathcal{E}_{cp} L_p = 2.5$ V, Test-2 conditions

$K_R = 1.97$. SPICE: $\tau_{PLH}/\tau_{PHL} = 1.045$.

We celebrate with moderation, as the mathematics is somewhat contrived. It merely extends the "three-quarters" conjecture of Eqs. 10.72 and 10.73 with an effective $K'(W/L)$ evaluated at $v_{in} = V_{DD}/2$. The most balanced results tend to reflect small values for ζ_n and ζ_p.

A third and last adjustment concerns the voltage dependence of the pn junction capacitances that make areal and sidewall contributions to C_{bdn1} and C_{bdp1} in Fig. 10.40. And while the theory is physically straightforward, the non-linearites involved upset the oscillator model under development. As noted in Chapter 5, body-to-drain capacitances vary according to the (modified) relation

$$C_{bd} = C_o \left(1 + \frac{v}{\phi}\right)^{-m} , \tag{10.78}$$

where C_o is the zero-bias junction capacitance, v is the *reverse* bias voltage, ϕ is the built-in junction voltage, and m is the junction grading coefficient. The areal and sidewall specifications for C_j, ϕ, and m usually differ.

In what follows, we put aside areal and sidewall distinctions to assume that Eq. 10.78 adequately describes the combination of both capacitances. We further assume that the nMOS and pMOS body-to-drain pn junctions have comparable ϕ and m parameters. Then with v_{out} varying from zero to $V_{DD}/2$ prior to a measurement for τ_{PLH}, the *average* capacitance is

$$C\uparrow = A\,C_{no} + B\,\frac{K_n{}'/K_p{}'}{K_R}C_{no}\,, \tag{10.79}$$

where

$$\begin{aligned} A &= \frac{2}{V_{DD}}\int_0^{V_{DD}/2}\left(1+\frac{v}{\phi}\right)^{-m}dv \\ &= \frac{2}{1-m}\left(\frac{\phi}{V_{DD}}\right)\left[\left(1+\frac{V_{DD}}{2\phi}\right)^{1-m}-1\right] \end{aligned} \tag{10.80}$$

and

$$\begin{aligned} B &= \frac{2}{V_{DD}}\int_{V_{DD}/2}^{V_{DD}}\left(1+\frac{v}{\phi}\right)^{-m}dv \\ &= \frac{2}{1-m}\left(\frac{\phi}{V_{DD}}\right)\left[\left(1+\frac{V_{DD}}{\phi}\right)^{1-m}-\left(1+\frac{V_{DD}}{2\phi}\right)^{1-m}\right]. \end{aligned} \tag{10.81}$$

Similarly, with v_{out} varying from V_{DD} to $V_{DD}/2$ prior to a measurement for τ_{PHL}, the *average* capacitance is

$$C\downarrow = B\,C_{no} + A\,\frac{K_n{}'/K_p{}'}{K_R}C_{no}\,. \tag{10.82}$$

Equations 10.79 and 10.82 support a final empirical condition for balanced propagation delays:

$$K_R\left(\frac{\mathcal{E}_{cn}L_n}{\mathcal{E}_{cp}L_p}\right)^{3/2}\left(\frac{1+\zeta_n-\sqrt{1+2\zeta_n}}{1+\zeta_p-\sqrt{1+2\zeta_p}}\right)^{3/4}\frac{C\uparrow}{C\downarrow} = 1\,. \tag{10.83}$$

Note that

$$\begin{aligned} C\uparrow &= A\,C_{no}\left(1+\frac{K_n{}'/K_p{}'}{K_R}\right)-(A-B)\frac{K_n{}'/K_p{}'}{K_R}\,C_{no}\,, \\ C\downarrow &= A\,C_{no}\left(1+\frac{K_n{}'/K_p{}'}{K_R}\right)-(A-B)\,C_{no}\,. \end{aligned} \tag{10.84}$$

Subject to typical $V_{DD} = 3.3$ V, $\phi = 0.75$ V, and $m = 0.33$, one obtains $A = 0.807$, $B = 0.624$, and $(A-B)/A = 0.227$. Thus, an approximation ignores the terms that are proportional to $(A-B)$ so that $C\uparrow/C\downarrow \approx 1$. Alternatively, one solves Eqs. 10.79, 10.82, and 10.83 together for K_R.

Example 10.6

The n- and p-channel MOSFETs used in the inverter layout of Fig. I.22 ($\lambda = 0.125\ \mu$m) have the SPICE parameters indicated in Example 10.5. Determine $C\uparrow$ and $C\downarrow$. Assume $V_{DD} = 2.5$ V and $C_{wire} = 0.2$ fF.

Solution

Fortunately, computers can be assigned much of the following work.

The $C\uparrow$ capacitance has junction, sidewall, load, and wire components. SPICE parameters for the areal portions of the drain-to-body junctions are CJ (1.713×10^{-3} F/m^2 and 1.925×10^{-3} F/m^2 for n- and p-channel devices), PB $= \phi$ and MJ $= m$ (0.99 V and 0.46 for both MOSFETs). Subject to

$$C\uparrow \text{ (junction)} = A_j\, C_{njo} + B_j\, C_{pjo}\ ,$$

we find (Eq. 10.80)

$$A_j = \frac{2}{0.54}\left(\frac{0.99}{2.5}\right)\left[\left(1+\frac{2.5}{1.98}\right)^{0.54} - 1\right] = 0.813$$

and (Eq. 10.81)

$$B_j = \frac{2}{0.54}\left(\frac{0.99}{2.5}\right)\left[\left(1+\frac{2.5}{0.99}\right)^{0.54} - \left(1+\frac{2.5}{1.98}\right)^{0.54}\right] = 0.617\ .$$

Values for C_{njo} and C_{pjo} are extracted from the inverter layout of Fig. 9.18. The n-channel drain has area $6\lambda\times6\lambda$. Thus, $C_{jno} = 36\,\lambda^2\times\text{CJ} = 0.964$ fF. The p-channel drain has the same area scaled by the factor $(K_n{}'/K_p{}')/K_R$. So with $K_n{}'/K_p{}' = \mu_n/\mu_p = 2.874$ (see Example 10.5), $C_{jpo} = 36\,\lambda^2\times\text{CJ}\times 2.874/K_R = 3.113$ fF$/K_R$. In turn,

$$C\uparrow \text{ (junction)} = 0.784 + \frac{1.921}{K_R}\quad \text{fF}\ .$$

Calculations for the sidewall portions of the drain-to-body capacitances proceed in the same way but with different governing SPICE parameters. We use n-channel PBSW $= 0.99$ V and MJSW $= 0.32$ in Eq. 10.80 to find $A_{sw} = 0.865$, and we use p-channel PBSW $= 0.53$ V and MJSW $= 0.29$ in Eq. 10.81 to find $B_{sw} = 0.647$. The n-channel drain has a perimeter of $24\,\lambda$. Thus with CJSW $= 4.036\times10^{-10}$ F/m, $C_{nswo} = 24\,\lambda\times\text{CJSW} = 1.211$ fF. The p-channel drain has a modified perimeter: $12\,\lambda+12\,\lambda\times(K_n{}'/K_p{}')/K_R$. Thus with CJSW $= 3.421\times10^{-10}$ F/m, $C_{pswo} = 12\,\lambda\times(1+2.874/K_R)\times$ CJSW $= 0.513 + 1.475/K_R$ fF. In turn,

$$C\uparrow \text{ (sidewall)} = 1.379 + \frac{0.954}{K_R}\quad \text{fF}\ .$$

The capacitance that reflects the input to another inverter has the form

$$C\uparrow\ (\text{load}) \approx 2\,\text{CGDO}\,W + \text{CGSO}\,W + \frac{1}{2}C_{ox}WL$$

for the n- and p-channel load devices. The factor of two that multiplies the CGDO overlap parameter accounts for output-to-input capacitive coupling (Miller effect). The MOSFETs have CGDO = CGSO $= 5.69 \times 10^{-10}$ F/m and CGDO = CGSO $= 6.91 \times 10^{-10}$ F/m for n and p channels, respectively. The device dimensions are $W_n = 0.75\ \mu$m, $W_p = 0.75\ \mu\text{m} \times 2.874/K_R$, and $L_n = L_p = 0.25\ \mu$m. Then with $C_{ox} = \epsilon_{ox}/\text{TOX} = 6.06\ \text{fF}/\mu\text{m}^2$,

$$C\uparrow\ (\text{load}) = 1.848 + \frac{6.100}{K_R}\quad \text{fF}\,.$$

Finally, we add the preceding three results and $C_{wire} = 0.2$ fF to obtain

$$C\uparrow\ (\text{total}) = 4.211 + \frac{8.975}{K_R}\quad \text{fF}\,.$$

The $C\downarrow$ calculation grinds along in much the same way except for the A and B factors, which are interchanged for the n- and p-channel MOSFETs. In the case of the "junction" capacitance component, the device parameters that contribute to A_j and B_j are the same for both devices. Thus,

$$C\downarrow\ (\text{junction}) = 0.595 + \frac{2.531}{K_R}\quad \text{fF}\,.$$

In the case of the "sidewall" capacitance component, the contributing device parameters are different. The new factors are

$$A_{sw} = \frac{2}{0.71}\left(\frac{0.53}{2.5}\right)\left[\left(1+\frac{2.5}{1.06}\right)^{0.71} - 1\right] = 0.814$$

and

$$B_{sw} = \frac{2}{0.68}\left(\frac{0.99}{2.5}\right)\left[\left(1+\frac{2.5}{0.99}\right)^{0.68} - \left(1+\frac{2.5}{1.98}\right)^{0.68}\right] = 0.715\,.$$

In turn,

$$C\downarrow\ (\text{sidewall}) = 1.283 + \frac{1.200}{K_R}\quad \text{fF}\,.$$

The "load" and "wire" capacitance components are unchanged. So finally,

$$C\downarrow\ (\text{total}) = 3.926 + \frac{9.831}{K_R}\quad \text{fF}\,.$$

Note that $C\uparrow/C\downarrow$ is K_R dependent.

Example 10.7

The n- and p-channel MOSFETs used in the inverter layout of Fig. I.22 ($\lambda = 0.125\ \mu$m) have the SPICE parameters indicated in Example 10.5. Estimate K_R for balanced τ_{PLH} and τ_{PHL}. $V_{DD} = 2.5$ V, $C_{wire} = 0.2$ fF.

Solution

From Example 10.5, $V_{Tn} = 0.364$ V, $V_{Tp} = -0.562$ V, $\mathcal{E}_{cn}L_n = 1.15$ V, and $\mathcal{E}_{cp}L_p = 4.70$ V. Then $\zeta_n = 0.770$ and $\zeta_p = 0.146$ (Eq. 10.76). In turn,

$$\left(\frac{\mathcal{E}_{cn}L_n}{\mathcal{E}_{cp}L_p}\right)^{3/2}\left(\frac{1+\zeta_n-\sqrt{1+2\zeta_n}}{1+\zeta_p-\sqrt{1+2\zeta_p}}\right)^{3/4} = 1.096\,.$$

With the help of the $C\uparrow$ and $C\downarrow$ results from Example 10.6, the condition for balanced τ_{PLH} and τ_{PHL} is

$$K_R\,(1.096)\,\frac{4.211 + 8.795/K_R}{3.926 + 9.831/K_R} = 1\,.$$

This yields a quadratic equation with the solution $K_R = 0.97$.

Figure 10.43 shows the results of a SPICE simulation for a five-stage ring oscillator featuring the device parameters of Example 10.5. The MOSFETs are described with area and perimeter values that are consistent with the layout of Fig. 9.18, and additional 0.2-fF load capacitors account for C_{wire}. The measured τ_{PLH}/τ_{PHL} ratio is 0.957 —not bad.

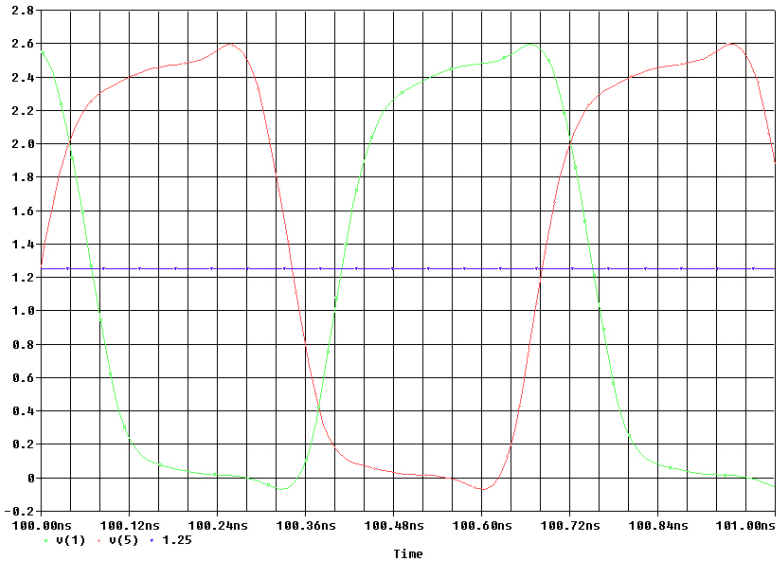


Figure 10.43: Simulation data for Example 10.7.

Relative Sizing for Optimized Delay

Balanced propagation delays are appropriate for clock drivers that provide regular and impartial upward or downward transitions. In contrast, digital subsystems are often designed to have minimum average propagation delay. This typically implies near-minimum n- and p-channel circuit capacitance and near-minimum dynamic power requirements.

In what follows, we assume that n- and p-channel MOSFETs have the same gate capacitance (C_{ox}) so that

$$\frac{W_p}{W_n} = \frac{\mu}{K_R} \tag{10.85}$$

with $\mu = \mu_e/\mu_h$. The HIGH-to-LOW output transition has the empirical K_R dependence of Eq. 10.69 ($\alpha = 1/4$), and it is proportional to the sum of the layout capacitances for the n- and p-channel MOSFETs, which roughly scale with W_n and W_p, respectively. Thus,

$$\tau_{PHL} \sim \frac{{K_R}^{1/4}}{{K_n}'W_n}(W_n + W_p)\,. \tag{10.86}$$

Similarly, the LOW-to-HIGH output transition has (Eq. 10.70)

$$\tau_{PLH} \sim \frac{{K_R}^{-1/4}}{{K_p}'W_p}(W_n + W_p)\,. \tag{10.87}$$

So with the help of Eq. 10.85 (Eq. 10.31 for K_R), the *average* propagation delay has the form

$$\tau_P \sim \frac{1}{{K_n}'}\left[{K_R}^{1/4}\left(1+\sqrt{K_R}\right)\left(1+\frac{\mu}{K_R}\right)\right]\,. \tag{10.88}$$

To find the optimum K_R, we differentiate Eq. 10.88 and equate the result to zero to obtain

$$3{K_R}^{3/2} + K_R - \mu{K_R}^{1/2} - 3\mu = 0\,. \tag{10.89}$$

Example 10.8

Find the K_R value that minimizes τ_P for the inverter of Example 10.7.

Solution

With $\mu = 306.1/106.5 = 2.874$ (see Example 10.5), the solution to Eq. 10.89 is $K_R = 2.34$. Then we apply the method of Example 10.7 with the relation

$$K_R\,(1.096)\,\frac{4.211 + 8.795/K_R}{3.926 + 9.831/K_R} = 2.34\,.$$

The modified result is $K_R = 2.19$ so that $W_p/W_n \approx 1.3$.

Concept Summary

The well-designed CMOS inverter has numerous requirements.

- The transfer characteristic, which relates the output to input voltage,
  - Has a switching threshold V_m midway between V_{DD} and ground;
    - For an inverter containing MOSFETs with equal-magnitude threshold voltages, this requires
    $$K_R = \frac{K_n'(W/L)_n}{K_p'(W/L)_p} = 1\,.$$
    - A consistent $(W/L)_p/(W/L)_n$ ratio is typically $\mu_e/\mu_p \approx 2.5$.
    - Velocity saturation effects tend to lower the preceding ratio.
  - Has a maximum low-level input voltage V_{IL} just below V_m;
  - Has a minimum high-level input voltage V_{IH} just above V_m;
  - Has evenly balanced noise margins slightly less than $V_{DD}/2$:
    - $NM_L = V_{IL} - V_{GND}$,
    - $NM_H = V_{DD} - V_{IH}$.
- The high-to-low (τ_{PHL}) and low-to-high (τ_{PLH}) propagation delays, which relate the time difference between input and output transitions with respect to $V_{DD}/2$,
  - Are ideally balanced;
    - For an inverter containing MOSFETs with equal-magnitude threshold voltages, this requires $K_R = 1$.
    - A consistent $(W/L)_p/(W/L)_n$ ratio is typically $\mu_e/\mu_p \approx 2.5$.
    - Velocity saturation effects tend to raise the preceding ratio.
  - Vary with $\tau_{PLH}/\tau_{PHL} \approx \sqrt{K_R}$;
  - Are complicated through the non-linearity of capacitive loading.
- Capacitive loading is kept small to minimize dynamic power loss, which varies in proportion to ${V_{DD}}^2$ and the switching frequency.
- Inputs that extend to the IC periphery are protected from electrostatic discharge and latch-up with diode clamps to V_{DD} and GND.
- Outputs that extend to the IC periphery are available as true CMOS, open drain, and tri-state (for non-interfering output connections).

Problems

Section 10.1

10.1 Justify Eqs. 10.3 and 10.4.

10.2 The MOSFETs in a particular CMOS inverter have the following specifications:

n-channel — $K_n'W/L = 100\ \mu\text{A/V}^2$, $V_{Tn} = 0.4$ V
p-channel — $K_p'W/L = 220\ \mu\text{A/V}^2$, $V_{Tp} = -0.6$ V

Use SPICE to determine V_m, V_{IL}, V_{IH}, and the noise margins subject to $V_{DD} = 3.3$ V.

10.3 Repeat Problem 10.2, but apply the following MOSFET specifications:

n-channel — $K_n'W/L = 160\ \mu\text{A/V}^2$, $V_{Tn} = 0.7$ V
p-channel — $K_p'W/L = 100\ \mu\text{A/V}^2$, $V_{Tp} = -0.3$ V

10.4 Repeat Problem 10.2, but apply the following MOSFET specifications:

n-channel — $K_n'W/L = 120\ \mu\text{A/V}^2$, $V_{Tn} = 0.5$ V
p-channel — $K_p'W/L = 500\ \mu\text{A/V}^2$, $V_{Tp} = -0.5$ V

10.5 A CMOS inverter has MOSFET parameters described in Problem 10.2, $\lambda_n = \lambda_p = 0.1\ \text{V}^{-1}$, and $V_{DD} = 3.3$ V. Use SPICE to determine the HIGH and LOW noise margins with and without channel-length modulation. Is the difference significant?

10.6 Use SPICE to show that $v_{out} = V_m$ when the inverter of Problem 10.2 has its output connected to the input. What V_{DD} requirement is fulfilled for this condition to be valid?

10.7 The input to the inverter of Problem 10.2 ramps slowly between zero and $V_{DD} = 3.3$ V over 100 μs. Neglect any input or output capacitance.

(a) Use SPICE to plot the crowbar current ($i_d \neq 0$) during the ramp period.

(b) Estimate the average power dissipated during the ramp period.

(c) Repeat part b for a ramp with 1-μs duration.

10.8 The n- and p-channel MOSFETs in the CMOS inverter of Fig. P10.8 have their body potentials at ground and V_{DD}, respectively, so that the body-to-drain capacitances C_{bdn} and C_{bdp} have connections as shown. Assume $C_{bdn} = C_{bdp} = 5$ fF.

(a) Describe the charging/discharging processes for C_{bdn} and C_{bdp} as the inverter output moves from ground to V_{DD} and back to ground.

(b) Consider a VLSI circuit featuring 20,000 similar inverters, all of which change state with every clock pulse. The inverters collectively dissipate 1 W of dynamic power subject to $V_{DD} = 3.3$ V. Determine the clock frequency.

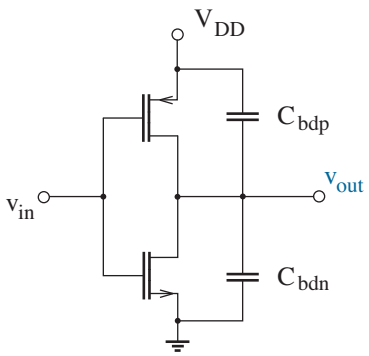


Figure P10.8

10.9 Consider the human body model of Fig. 10.5. Prior to $t = 0$, the 100-pF "body" capacitor (C_1) is charged to 3000 V. The human model suddenly connects to a 0.1-pF gate capacitance (C_2) at $t = 0$.

(a) Determine the final voltage across C_2 (assuming no breakdown).

(b) Find the maximum charging current for C_2.

(c) Estimate the time when the C_2 voltage reaches 90% of its final value.

10.10 Consider the human body model of Fig. 10.5. Prior to $t = 0$, the 100-pF "body" capacitor (C_1) is charged to 20,000 V. The human model suddenly connects to an integrated-circuit input with diode protection shown in Fig. 10.10. Find the maximum discharge current for C_1, then modify the protection circuit to limit the discharge current to 100 mA under worst-case (20,000-V) conditions.

10.11 Provide an expression for the worst-case input voltage v_x subject to negative i_{esd} in the protection circuit of Fig. 10.10.

10.12 In the so-called **machine model** for electrostatic discharge (appropriate for automated circuit-board assembly systems), a 200-pF capacitor (C_1) discharges through a 500-nH inductor (L) into an integrated-circuit input. The input has capacitance $C_2 = 0.1$ pF, and the resistance of the series input wire is $R = 20\ \Omega$. Capacitor C_1 is initially at 400 V.

(a) Use SPICE to find the peak discharge current from C_1 into the integrated circuit.

(b) Show how the results of part a change when the input is protected by two diodes as in Fig. 10.10. Assume IS=10p for each diode.

10.13 The latch-up circuit model shown in Fig. 10.7 features $R_t = 0.1\ \Omega$, $R_w = 2$ kΩ, $R_s = 100\ \Omega$, and $V_{DD} = 5$ V. The BJTs have the following SPICE parameters:

Q_1 IS=1f, BF=2, RB=10, CJE=50f, CJC=20f

Q_2 IS=10f, BF=100, RB=10, CJE=0.2p, CJC=0.1p

(a) Use SPICE to perform a transient simulation that allows for the possibility of latch-up with the application of a 200-ns current pulse i_x at the base of Q_2, and find the minimum i_x pulse magnitude that results in latch-up.

Note: To ensure good simulation behavior, use the .ic command to provide 5-V and 0-V initial conditions at the bases of Q_1 and Q_2.

(b) Show the time variations of the BJT collector currents if i_x has the value obtained in part a. Discuss the results.

(c) Specify the holding current and holding voltage that applies after the circuit enters latch-up.

10.14 Repeat the SPICE analysis of Problem 10.13, but let the i_x pulse have 50-mA amplitude. Find the *maximum* R_2 value that avoids latch-up.

10.15 Consider the input circuit of Fig. 10.10 with $V_{DD} = 3.3$ V, $R_t = 10\ \Omega$, $C_b = 10\ \mu$F, and $L_1 = L_2 = 0.1$ nH. The diodes feature the SPICE parameters IS=100f, CJO=0.1p, RS=5, and the connection to internal CMOS logic can be modeled as a 0.5-pF capacitance to ground. Use SPICE to demonstrate input behavior when the inputs are floating (not connected externally). Assume a short current-pulse perturbation.

10.16 Repeat the SPICE analysis of Problem 10.15, but connect the input to a 1-kΩ resistor that is tied to ground.

10.17 The CMOS inverter of Fig. 10.11 has the specifications given in Problem 10.2. Use SPICE to find I_{OL} subject to $V_{DD} = 3.3$ V.

10.18 The CMOS inverter of Fig. 10.12 has the specifications given in Problem 10.2. Use SPICE to find I_{OH} subject to $V_{DD} = 3.3$ V.

10.19 In the wired-OR circuit of Fig. 10.13 the n-channel MOSFETs have $K_n'W/L = 200\ \mu$A/V^2 and $V_{Tn} = 0.4$ V. Find R_{common} so that $V_{OL} = 0.2$ V if only one connected gate demands a LOW output, then (with the same R_{common}) determine v_{out} if two connected gates demand a LOW output.

10.20 The tri-state inverter of Fig. 10.14 has $K_n'W/L = 120\ \mu$A/V^2, $V_{Tn} = 0.4$ V for the n-channel MOSFET, $K_p'W/L = 120\ \mu$A/V^2, $V_{Tp} = -0.6$ V for the p-channel MOSFET. Let $V_{DD} = 5$ V. Use SPICE to find the high and low noise margins, then repeat with $\gamma = 0.2$ V$^{1/2}$ for all devices.

Section 10.2

10.21 Design a combinational CMOS circuit that implements

$$OUT = \overline{A}B(C + \overline{D})E$$

and indicate the relative device widths for optimum output transitions. Assume a p/n width ratio of 2/1 for the inverter standard.

10.22 Repeat Problem 10.21: $OUT = AB\overline{C} + D\overline{E}$.

10.23 Repeat Problem 10.21: $OUT = A\overline{B}C\overline{D} + E$.

10.24 Repeat Problem 10.21: $OUT = \overline{A}(\overline{B}C + D\overline{E})$.

10.25 MOSFETs in the NOR gate of Fig. 10.15 have the specifications of Problem 10.2. Use SPICE to plot v_{out} as input B varies from 0 to $V_{DD} = 3.3$ V subject to HIGH input A. Perform the simulation with and without the body effect ($\gamma = 0.5$ V$^{1/2}$).

10.26 MOSFETs in the NAND gate of Fig. 10.15 have the specifications of Problem 10.2. Use SPICE to plot v_{out} as input A varies from 0 to $V_{DD} = 3.3$ V subject to HIGH input B. Perform the simulation with and without the body effect ($\gamma = 0.5$ V$^{1/2}$).

10.27 Consider the pseudo nMOS circuit of Fig. 10.20b. The n-channel MOSFETs have $K_n{}' = 50\ \mu$A/V^2 and $V_{Tn} = 0.5$ V; the p-channel MOSFET has $K_p{}' = 20\ \mu$A/V^2 and $V_{Tp} = -0.5$ V. Assume $(W/L)_n = 2$, and let $V_{DD} = 3.3$ V.

(a) Apply Eq. 10.26 to find $(W/L)_p$ such that $V_{OL} = 0.2$ V (worst case).

(b) Use SPICE to determine the worst-case V_{OL} subject to the design of part a and show that it agrees with theoretical expectations.

(c) Use SPICE to find the output voltage subject to the design of part a when only two inputs are HIGH, and show that the result agrees with theoretical expectations.

10.28 Design a pseudo-nMOS gate that implements $OUT = A + B + C\overline{D}$ subject to the transistor and power-supply specifications of Problem 10.27 and worst-case $V_{OL} = 0.3$ V. Verify with SPICE.

10.29

(a) Design a pseudo-nMOS inverter subject to the transistor and power-supply specifications of Problem 10.27 and $V_{OL} = 0.2$ V.

(b) Use SPICE to determine the HIGH and LOW noise margins.

(c) The inverter of part a experiences an input that abruptly steps from zero to V_{DD}. Use SPICE to determine the time needed for the output to fall to $V_{DD}/2$. Assume a 1-pF capacitive load.

(d) The inverter of part a experiences an input that abruptly steps from V_{DD} to zero. Use SPICE to determine the time needed for the output to rise to $V_{DD}/2$. Assume a 1-pF capacitive load.

(e) Show how the results of parts c and d compare to a conventional inverter for which $(W/L)_p = 5$.

10.30 Show how you would design a **pseudo pMOS** four-input NAND gate and explain whether it is an attractive option in relation to conventional CMOS.

10.31 A dynamic logic inverter has the transistor and power-supply specifications of Problem 10.27, and all devices have $W/L = 4$. The clock waveform varies between 0 and 3.3 V with 200-MHz frequency and 0.1-ns rise/fall times.

(a) Use SPICE to demonstrate the inverter function and input timing constraints if $C = 0.1$ pF.

(b) The capacitor has a parasitic parallel leakage path with resistance R. Use SPICE to find the worst-case value for R.

10.32 Design a dynamic logic circuit that implements $OUT = A + BC + DE$.

10.33 Consider the cascade of identical dynamic inverters shown in Fig. P10.33. Let input A be HIGH. Sketch the time dependence of outputs $OUT1$ and $OUT2$ in relation to the clock signal ϕ to show that $OUT2$ misbehaves during the evaluation period. Allow for modest clock and output transition times.

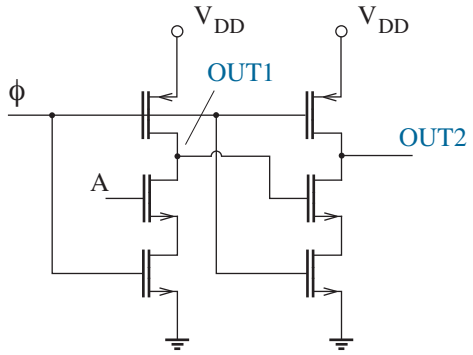


Figure P10.33

10.34 Show that dynamic-logic-inverter cascade in Problem 10.31 functions as intended if each dynamic gate is followed by a conventional CMOS inverter.

This revised circuit is an example of **domino logic**.

10.35 The MOSFETs in the latch circuit of Fig. P10.35 both feature $K'W/L = 1$ mA/V^2 and $|V_T| = 0.5$ V. Nodes A and B are LOW and HIGH, respectively. Assume $V_{DD} = 3.3$ V.

(a) Apply a transient SPICE simulation to find the magnitude of a ground-referenced voltage pulse at node A sufficient to change the latch state. Assume a 1-μs pulse width.

(b) Repeat part a, but include a V_{DD}-referenced pulse generator at node B with the same pulse amplitude as the one at node A.

10.36 The MOSFETs in the latch portion of the sRAM cell of Fig. 10.22 have $K'W/L = 1$ mA/V^2 and $|V_T| = 0.5$ V. Supply voltage is $V_{DD} = 3.3$ V. During a read operation, the BL and $\overline{BL}$ lines can be modeled as 500-fF capacitances that "float" at $V_{DD}/2$. Use SPICE to simulate the read operation

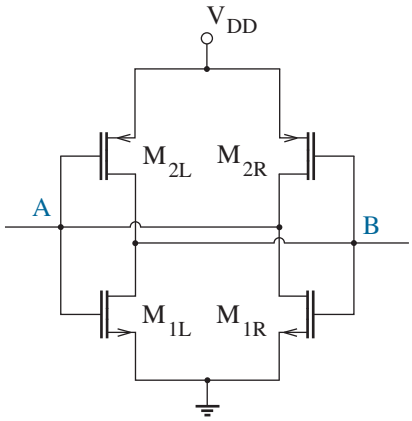


Figure P10.35

and find the maximum $K'W/L$ for M_{3L} and M_{3R} that preserves a particular latch state.

10.37 The MOSFETs in the latch portion of the sRAM cell of Fig. 10.22 have $K'W/L = 1$ mA/V^2 and $|V_T| = 0.5$ V. Supply voltage is $V_{DD} = 3.3$ V. During a particular write operation, the BL and $\overline{BL}$ lines are fixed at V_{DD} and ground, respectively. Use SPICE to simulate the write operation and find the minimum $K'W/L$ for M_{3L} and M_{3R} that alters an opposing latch state.

10.38 Consider the latch circuit of Problem 10.35 with the p-channel MOSFETs replaced by equal resistors. Use SPICE to find the maximum resistance value that allows latch behavior. Let $V_{DD} = 3.3$ V.

Note: A resistively loaded sRAM latch invites three-dimensional integration with each polysilicon load placed on top of an n-channel MOSFET.

10.39 Verify Eq. 10.28.

10.40 The floating-gate MOSFET in Fig. 10.28 has capacitance C_1 between the floating gate and the channel region, capacitance C_2 between the select gate and the floating gate. Estimate the change in threshold voltage if C_1 receives negative charge $-Q_x$. Assume $2\phi_F$ for the surface potential at inversion.

Section 10.3

10.41 The MOSFETs in a particular CMOS inverter have the following specifications:

n-channel — $K_n' = 45\ \mu\text{A/V}^2$, $V_{Tn} = 0.4$ V
p-channel — $K_p' = 18\ \mu\text{A/V}^2$, $V_{Tp} = -0.6$ V

Determine the K_R factor and the p/n ratio of (W/L) that establishes $V_M = 2$ V subject to $V_{DD} = 5$ V.

10.42 The MOSFETs in a particular CMOS inverter have the following specifications:

n-channel — $K_n' = 36\ \mu\text{A/V}^2$, $V_{Tn} = 0.7$ V
p-channel — $K_p' = 22\ \mu\text{A/V}^2$, $V_{Tp} = -0.3$ V

Determine the K_R factor and the p/n ratio of (W/L) that establishes $V_M = 2$ V subject to $V_{DD} = 3.3$ V.

10.43 Repeat Problem 10.33 with $v_{sat} = 8 \times 10^6$ cm/s for electrons or holes, $C_{ox} = 8 \times 10^{-8}$ F/cm^2, and $L_n = L_p = 0.5\ \mu$m.

10.44 Repeat Problem 10.34 with $v_{sat} = 8 \times 10^6$ cm/s for electrons or holes, $C_{ox} = 6 \times 10^{-8}$ F/cm^2, and $L_n = L_p = 0.35\ \mu$m.

10.45 A CMOS inverter subject to $K_R = 1.2$ has the following MOSFET specifications:

n-channel — $K_n' = 42\ \mu\text{A/V}^2$, $V_{Tn} = 0.5$ V
p-channel — $K_p' = 24\ \mu\text{A/V}^2$, $V_{Tp} = -0.5$ V

(a) Use SPICE to find the HIGH and LOW noise margins.

(b) Repeat part a, but let $v_{sat} = 8 \times 10^6$ cm/s for electrons or holes, $C_{ox} = 5 \times 10^{-8}$ F/cm^2, and $L_n = L_p = 0.4\ \mu$m.

10.46 A CMOS inverter has the specifications given in Problem 10.33 with $K_R = 1.5$. Use SPICE to show how V_m varies for 2 V $\leq V_{DD} \leq$ 5 V.

10.47 Repeat Example 10.3 subject to the MOSFET specifications of Problem 10.2 and $V_{DD} = 3.3$ V.

10.48 Repeat Example 10.3 subject to the MOSFET specifications of Problem 10.3 and $V_{DD} = 3.3$ V.

10.49 Apply the procedures of Example 10.4 to design for $NM_L = 1.5$ V.

10.50 Apply the procedures of Example 10.4 to design for $V_m = 2$ V.

10.51 Apply the procedures of Example 10.4 in conjunction with the SPICE .op command to find V_{IL} and V_{IH} for a CMOS inverter with the MOSFETs of Problem 10.2 and $K_R = 1.4$. Let $V_{DD} = 3.3$ V.

10.52 Download the SPICE parameters for an IBM 0.25-μm SCN018 process at

http://www.mosis.com/cgi-bin/cgiwrap/umosis/swp/params/ibm-018/t7cw_7rf_6lm_ml-params.txt.

Design a CMOS inverter for this process subject to $V_m = V_{DD}/2 = 1.25$ V.

Note: Your instructor will provide an appropriate link in the event that the one provided is out of date.

Section 10.4

10.53 Use SPICE to determine τ_{PLH}, τ_{PHL}, and τ_P for a five-stage ring oscillator with the MOSFETs of Problem 10.2. Assume that each inverter has a 1-pF output load.

10.54 Use SPICE to determine τ_{PLH}, τ_{PHL}, and τ_P for a five-stage ring oscillator with the MOSFETs of Problem 10.3. Assume that each inverter has a 1-pF output load.

10.55 Determine the K_R factor and the p/n ratio of (W/L) that establishes balanced propagation delays for the inverter of Problem 10.33. Verify with SPICE. Assume that each inverter has a 1-pF output load.

10.56 Determine the K_R factor and the p/n ratio of (W/L) that establishes balanced propagation delays for the inverter of Problem 10.34. Verify with SPICE. Assume that each inverter has a 1-pF output load.

10.57 Repeat Problem 10.47 under velocity saturation with $\mathcal{E}_{cn}L_n = 1.2$ V and $\mathcal{E}_{cp}L_p = 3.8$ V.

10.58 Repeat Problem 10.48 under velocity saturation with $\mathcal{E}_{cn}L_n = 1.8$ V and $\mathcal{E}_{cp}L_p = 5.2$ V.

10.59 Repeat Example 10.6 with SPICE parameters given in Problem 10.44, and assume that all sources and drains have a minimum 8-λ extension from a nearby poly gate (see Fig. 10.28).

10.60 Repeat Example 10.7, but assume that each of the inverters in the five-stage ring oscillator is loaded by pn junction capacitance subject to CJO=0.5 pF, VJ=0.8 V, and M=0.5. Other SPICE parameters are given in Problem 10.44. Evaluate with SPICE.

10.61 Determine the K_R value that minimizes τ_P for the inverter of Problem 10.41.

10.62 Determine the K_R value that minimizes τ_P for the inverter of Problem 10.50.

Perspective: Feedback

To ensure a proper perspective as Chapters 11 and 12 unfold, it will be helpful to consider four amplifier classifications that relate the type of input excitation (voltage or current) to the output response (voltage or current). As depicted in Fig. D1, these are:

- The (trans)**voltage amplifier**, for which an input voltage produces an output voltage according to the relation

$$\frac{v_{out}}{v_t} = A_v \, . \tag{D1}$$

A well-designed voltage amplifier features a *large* input resistance (so that the signal source can be fully effected with negligible loading across series Thevenin resistance) and a *small* output resistance (so that the available output voltage can be fully delivered across an arbitrary load resistance). Example — a pre-amplifier in a typical stereo system.

(The voltage amplifier was our primary focus in Chapters 7 and 8.)

- The **transresistance amplifier**, for which an input current produces an output voltage according to the relation

$$\frac{v_{out}}{i_n} = R_m \, . \tag{D2}$$

A well-designed transresistance amplifier features a *small* input resistance (so that the signal source can be fully effected with negligible loading through shunt Norton resistance) and a *small* output resistance (so that the available output voltage fully appears across an arbitrary load resistance). Example — a circuit that provides an output voltage in proportion to a controlling current from a photodetector.

- The (trans)**current amplifier**, for which an input current produces an output current according to the relation

$$\frac{i_{out}}{i_n} = A_i \, . \tag{D3}$$

A well-designed current amplifier features a *small* input resistance (so that the signal source can be fully effected with negligible loading through shunt Norton resistance) and a *large* output resistance (so that the available output current can be fully delivered through an arbitrary load resistance). Example — a circuit that "boosts" a current level by a constant factor.

- The **transconductance amplifier**, for which an input voltage produces an output current according to the relation

$$\frac{i_{out}}{v_t} = G_m \, . \tag{D4}$$

A well-designed transconductance amplifier features a *large* input resistance (so that the signal source can be fully effected with negligible loading across series Thevenin resistance) and a *large* output resistance (so that the available output current fully delivers through an arbitrary load resistance). Example — a circuit that provides output current to a coil that induces mechanical force in proportion to a controlling voltage.

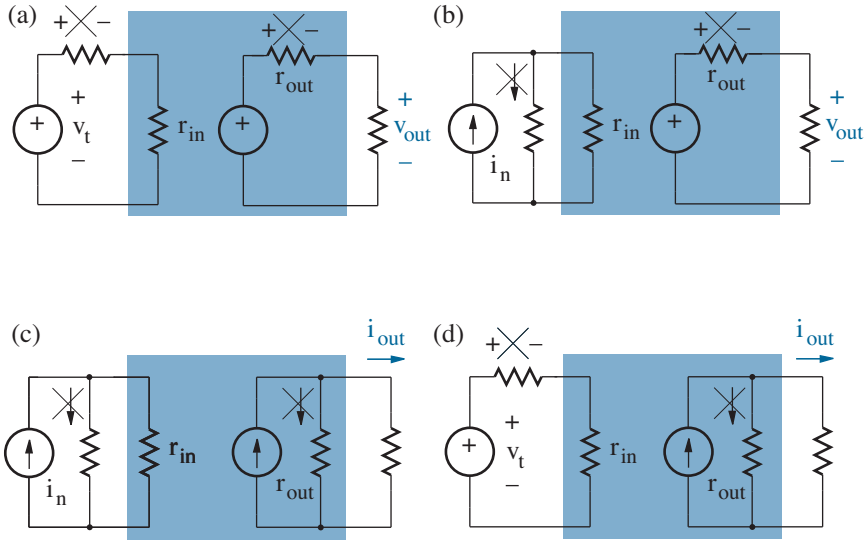


Figure D1: Basic amplifier classifications: (a) voltage; (b) transresistance; (c); current (d) transconductance.

One way to promote particular amplifier behavior is the application of a **feedback** process shown in the generic block diagram of Fig. D2. Here, one "samples" the output y (voltage or current) to obtain a signal Fy with the same character as the input x (voltage or current). If the sampled signal is declared consistent with an ideal input-output relation of the form $x = Fy$, then a subtractive combination with x produces an "error" signal $x - Fy$. But whereas this is also the input to amplifier A, we have $x - Fy = y/A$, which tends to zero for large A. The exact input-output relationship is

$$y = \frac{A\,x}{1 + AF}\,. \tag{D5}$$

In turn, the ideal $y \approx x/F$ relationship requires $1/F \ll A$ and a sacrifice of available signal transfer capability without feedback ($F = 0$).

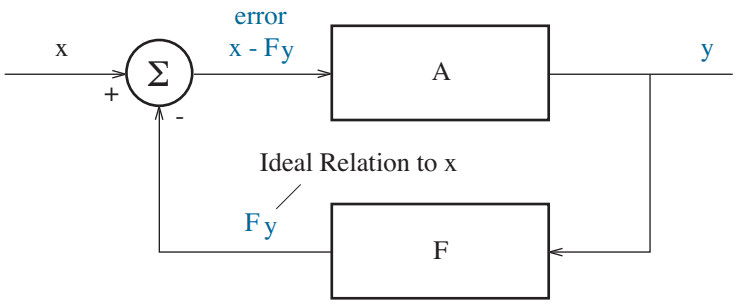
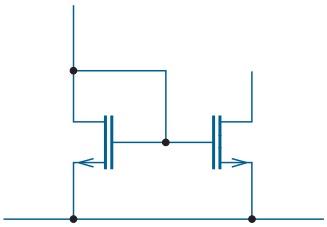


Figure D2: Block diagram for a generic feedback process.

The application of feedback causes the system output to resist change. Thus, for the case of voltage, an output stimulus Δi induces a weak Δv, so $\Delta v/\Delta i$ and the output resistance are small (as desired). Conversely, for current, an output stimulus Δv induces a weak Δi, so $\Delta v/\Delta i$ and the output resistance are large (as desired). While harder to visualize, subtractive conditions at the input tend to increase or decrease the apparent input resistance for the case of voltage or current, respectively.

Chapter 11 moves from the abstract visualization of Fig. D2 to specific circuit formats that help to achieve the amplifier behavioral classifications. Each of the four cases requires two interconnected "A" and "F" circuits, the latter typically passive through a combination of well-defined resistors. The system will prove insensitive to A-circuit uncertainties.

Chapter 12 explores dynamic feedback effects. Since feedback reduces the overall system gain, we expect increased bandwidth as in the tradeoffs encountered in Chapter 8. Nevertheless, frequency-dependent phase shifts can lead to effective *positive* feedback and unstable system performance. While undesirable at first glance, this is also the basis for oscillator design.



Chapter 11

Steady-State Feedback

Way back in Chapter 1, we applied feedback with operational ampliers to support a simple form of black-box behavior. We now examine feedback as a process that helps to improve amplifier effectiveness under particular signal-source and load conditions (while, among other things, reducing the influence of amplifier performance factors that tend to vary significantly). Our "steady-state" focus ignores all aspects of time dependence when predicting or designing for feedback benefits. Chapter 12 considers dynamic feedback effects, which have the potential to upset steady-state results.

The Perspective described four amplifier classifications. In turn, we will show how each amplifier can be combined with a separate feedback circuit to improve performance. Formal analysis procedures are straightforward. Nevertheless, we look for informal shortcuts that are convenient for design.

Objectives

After studying this chapter and working through the problems, you should be able to perform each of the following:

- Identify four basic feedback circuit configurations in the context of two interconnected two-ports (Sections 11.1 - 11.4).
- Form an appropriate "A" circuit that accounts for feedback loading effects without backward information exchange (Sections 11.1 - 11.4).
- Use simple rules to find the feedback factor F (Sections 11.1 - 11.4).
- Determine the feedback-stabilized transfer characteristic and input / output impedances using an A-circuit and F (Sections 11.1 - 11.4).

11.1 Series-Shunt Feedback

The voltage amplifier wants to have both large input resistance and small output resistance. So with this in mind, it seems worthwhile to examine an "ordinary" two-port amplifier and a *passive* two-port network that are interconnected as shown in Fig. 11.1. Subject to positive resistance looking into any two-port terminal pair, the left-side series combination increases the input resistance from the perspective of the Thevenin source, and the right-side parallel combination decreases the output resistance from the perspective of the load. The overall circuit configuration is an example of **series-shunt feedback**—the passive network samples v_2, a voltage shared at the amplifier output, and the information is "fed back" to influence i_1, a loop current presumably shared at the input sides of the amplifier and feedback two-ports. (The latter sharing is open to question.)

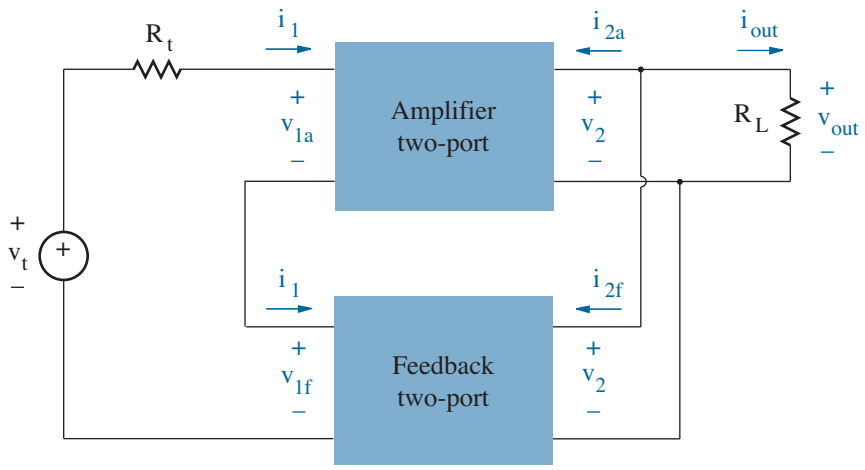


Figure 11.1: Series-shunt two-port feedback configuration.

It should not be difficult to solve for i_1 and v_2 in the series-shunt feedback configuration if the input terminal voltages and the output terminal currents are expressed in terms of the shared variables of interest. Thus, we look to the two-port h**-parameter** representation in which

$$v_1 = h_{11}i_1 + h_{12}v_2\,, \tag{11.1a}$$
$$i_2 = h_{21}i_1 + h_{22}v_2\,. \tag{11.1b}$$

The individual h parameters are easily determined. For example, one finds h_{11} by measuring v_1 subject to an i_1 current-source excitation while the two-port output terminal pair is shorted ($v_2 = 0$). (See Section 7.2.)

Exercise 11.1 Determine the h parameters for the two-port of Fig. 11.2.

Ans: $h_{11} = 0.66\ \text{k}\Omega$ $h_{12} = 0.33$ $h_{21} = -0.33$ $h_{22} = 0.66\ \text{m}\mho$

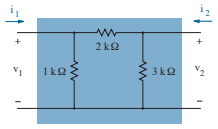


Figure 11.2: Circuit for Exercise 11.1.

Returning to the two-port arrangement of Fig. 11.1, we write Kirchhoff's voltage law around the input loop,

$$R_t i_1 + \underbrace{h_{11a} i_1 + h_{12a} v_2}_{v_{1a}} + \underbrace{h_{11f} i_1 + h_{12f} v_2}_{v_{1f}} = v_t \,, \tag{11.2}$$

and we write Kirchhoff's current law at the output node,

$$G_L v_2 + \underbrace{h_{21a} i_1 + h_{22a} v_2}_{i_{2a}} + \underbrace{h_{21f} i_1 + h_{22f} v_2}_{i_{2f}} = 0 \,. \tag{11.3}$$

So in matrix form, we have

$$\begin{pmatrix} z_i & h_{12a}+h_{12f} \\ h_{21a}+h_{21f} & y_o \end{pmatrix} \begin{pmatrix} i_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} v_t \\ 0 \end{pmatrix}, \tag{11.4}$$

where

$$z_i = R_t + h_{11a} + h_{11f} \tag{11.5}$$

is the sum of the key impedance values in the input loop, and

$$y_o = G_L + h_{22a} + h_{22f} \tag{11.6}$$

is the sum of the key admittance values at the output node.

Now consider an approximation. Passive feedback networks propagate signals from input to output or vice versa with gains that are less than unity. In contrast, amplifiers tend to propagate signals from input to output with large gain, while tending not to propagate any signal from output to input. Thus, we assume

$$h_{21a} \gg h_{21f} \,, \tag{11.7a}$$
$$h_{12a} \ll h_{12f} \,. \tag{11.7b}$$

In turn,

$$\begin{pmatrix} z_i & h_{12f} \\ h_{21a} & y_o \end{pmatrix} \begin{pmatrix} i_1 \\ v_2 \end{pmatrix} \approx \begin{pmatrix} v_t \\ 0 \end{pmatrix} \,. \tag{11.8}$$

Subject to $v_{out} = v_2$, the solution is

$$\frac{v_{out}}{v_t} = \frac{A}{1 + AF} \,, \tag{11.9}$$

where

$$A = \frac{-h_{21a}}{z_i \, y_o} \tag{11.10}$$

and

$$F = h_{12f} \,. \tag{11.11}$$

Equation 11.9 is the fundamental feedback relationship.

It is little extra work to solve Eq. 11.8 for v_t/i_1. Then subtracting R_t, we obtain

$$z_{in} = z_i \, (1 + AF) - R_t \tag{11.12}$$

as the input impedance "seen" by the Thevenin circuit featuring v_t and R_t. This is larger than h_{11a}, the input impedance in the absence of feedback, and it can be very large if $AF \gg 1$.

If (with $v_t = 0$) we apply a test current source in place of R_L and evaluate the voltage produced by this source—see Problem 11.3—we obtain

$$y_{out} = y_o \, (1 + AF) - G_L \tag{11.13}$$

as the output admittance "seen" by the load. This is larger than h_{22a}, the output admittance in the absence of feedback, and it can be very large if $AF \gg 1$. The reciprocal output impedance is accordingly small.

More-than-anticipated improvements in input and output impedance are welcome. Nevertheless, feedback provides a particularly special benefit through the form of Eq. 11.9. When re-expressed in terms of differentials,

$$\frac{\Delta\left(\frac{v_{out}}{v_t}\right)}{\left(\frac{v_{out}}{v_t}\right)} = \frac{\Delta A}{A}\left(\frac{1}{1+AF}\right) . \tag{11.14}$$

Thus, *the fractional change in the voltage gain is made insensitive to the fractional change in A by the factor* $1/(1+AF)$. The reduced sensitivity is significant if $AF \gg 1$. Of course we pay a price—the voltage gain in the absence of feedback is decreased by the same factor $1/(1+AF)$.

Exercise 11.2 Without feedback, an amplifier features a voltage gain of (a) 200 for $|v_t| < 10$ mV and (b) 2000 for $|v_t| > 10$ mV. Determine the amplifier voltage gain with feedback when $F = 0.05$.

Ans: (a) 18.2 (b) 19.8

Figure 11.3 shows the input-output characteristics for the amplifier of Exercise 11.2. Feedback tends to compress the scale of the input variable by the factor $(1+AF)$. In turn, there is decreased non-linearity or reduced **distortion** if $AF \gg 1$.

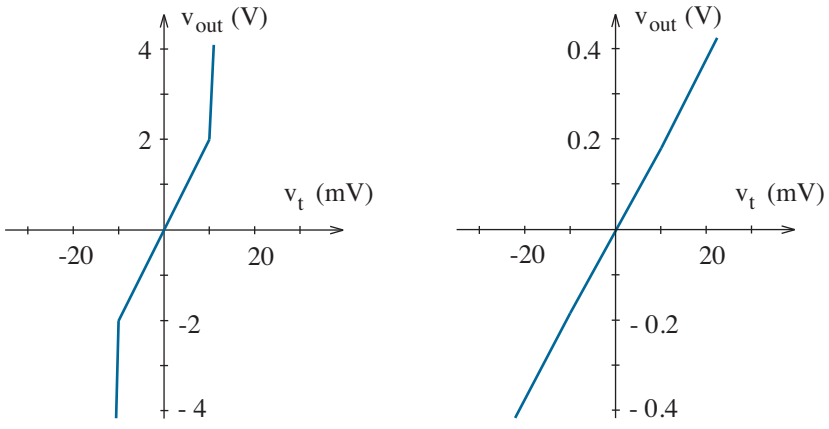


Figure 11.3: Transfer behavior for the voltage amplifier of Exercise 11.2: (a) without feedback; (b) with feedback.

Example 11.1

The op-amp in the circuit of Fig. 11.4 features a differential voltage gain of 20,000, an input resistance of 500 kΩ, and an output resistance of 300 Ω. Determine v_{out}/v_t.

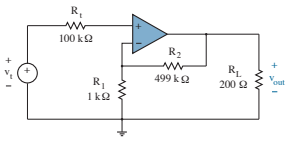


Figure 11.4: Circuit for Example 11.1.

Solution

We redraw the circuit as a pair of interconnected two-ports in the series-shunt configuration as shown in Fig. 11.5.

Note the positions of the ground connections.

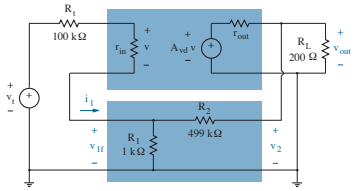


Figure 11.5: Series-shunt circuit for Example 11.1.

Next, we make a clever observation.

The quantity A in Eq. 11.10 is only influenced by the feedback two-port through the terms z_i and y_o, which partially feature $h_{11f} + R_t$ and $h_{22f} + G_L$, respectively. Thus, by defining

$$R_t{}' = R_t + h_{11f} \tag{11.15}$$

and

$$R_L{}' = R_L \parallel \frac{1}{h_{22f}} \tag{11.16}$$

(so that $G_L{}' = G_L + h_{22f}$), we can calculate A directly from the "A" circuit of Fig. 11.6. Apart from the revised loading, the "A" circuit exchanges no information between output and input. Thus, $F = 0$ and $v_{out}/v_t = A$.

The preceding observation is motivated by a desire to avoid h-parameter calculations, particularly those related to h_{21a}. And even though h_{11f} and h_{22f} require little effort, it seems a nuisance to look to Eq. 11.1 to specify whether a terminal pair is open ($i = 0$) or shorted ($v = 0$) when determining the effective resistance or conductance over at the opposite terminal pair. So h parameters begone. The series-shunt feedback configuration obviously

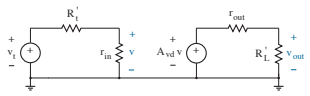


Figure 11.6: "A" circuit for Example 11.1.

increases R_t and G_L with *additive* resistance in the input loop and *additive* conductance at the output node, respectively. Moreover, when calculating the loading effect at either end of the feedback two-port, *it is appropriate to null the shared feedback information at the other end.*

Back to Fig. 11.5. To determine $R_t{}'$, we add the input resistance of the feedback two-port to R_t. The feedback process at the opposite (output) terminal pair is sampling of a shared voltage, so we null that information by making $v_2 = 0$. In turn, $R_t{}' = 100\ \text{k}\Omega + (1\ \text{k}\Omega \parallel 499\ \text{k}\Omega) \approx 101\ \text{k}\Omega$. To determine $R_L{}'$, we take the parallel combination of R_L and the output resistance of the feedback two-port. The feedback process at the opposite (input) terminal pair is the influence of a shared current, so we make $i_1 = 0$. In turn, $R_L{}' = 200\ \Omega \parallel (1\ \text{k}\Omega + 499\ \text{k}\Omega) \approx 200\ \Omega$.

Returning to Fig. 11.6, we find

$$\frac{v_{out}}{v_t} = \left(\frac{r_{in}}{R_t{}' + r_{in}}\right) A_{vd} \left(\frac{R_L{}'}{r_{out} + R_L{}'}\right) = 6660 \tag{11.17}$$

as the voltage gain for the "A" circuit. —That was easy.

Once again, we go back to the feedback two-port to determine $F = h_{12f}$. And although this requires little effort, it would be nice to avoid thinking about h parameters. Thus, we apply some simple rules:

- The test source for the F measurement is applied at the output side of the feedback two-port, *with the character of the shared information.*
- The response at the input side of the feedback two-port is observed *with the shared circuit variable under feedback influence set to zero.*

In the series-shunt configuration, the feedback two-port samples the shared output voltage. Thus, we remove the sampling connections and apply a test *voltage* source with value v_2. The feedback two-port influences the shared input current. So we complete the set-up for the F measurement by making $i_1 = 0$ and observing v_{1f}. In turn, we have $F = v_{1f}/v_2 = 0.002$. Note that F is consistent with dimensionless AF.

Having found A and F from separate measurements, we find that the overall amplifier voltage gain is $v_{out}/v_t = A/(1 + AF) = 465$ (about 7 % lower than the ideal value of 500 that is obtained by circuit inspection).

While we are at it—

To determine the effective input resistance of the feedback amplifier, we evaluate z_i as the sum of $R_t{}'$ (101 kΩ) and the input resistance of the two-port op-amp (500 kΩ). Then with $1 + AF = 14.3$, we apply Eq. 11.12 to find $z_{in} = 601\ \text{k}\Omega \times 14.3 - 100\ \text{k}\Omega = 8.5\ \text{M}\Omega$.

To determine the effective output resistance of the feedback amplifier, we evaluate y_o as the sum of $G_L{}'$ (5×10^{-3} ℧) and the output conductance of the two-port op-amp (3.33×10^{-3} ℧). Then we apply Eq. 11.13 to find $y_{out} = 8.33 \times 10^{-3}\ \text{℧} \times 14.3 - 5 \times 10^{-3}\ \text{℧} = 0.114\ \text{℧}$. Thus, $z_{out} = 8.8\ \Omega$.

Task completed, and hardly an h parameter in sight.

In part, the intent of the preceding example was to show that feedback calculations can be easily performed without the use of formal two-port parametric representations. Nevertheless, some individuals may prefer the comfort of well-defined quantities appearing in simple relations for A and F (that will maintain a particular form for other feedback configurations).

Example 11.2

The two-stage amplifier of Fig. 11.7 has a midfrequency voltage gain of 133 when $R_f \to \infty$ (as determined with SPICE). But it is highly g_m dependent, and the output resistance is large (2.7 kΩ). Complete the feedback design to stabilize the voltage gain at 10 while lowering the output resistance.

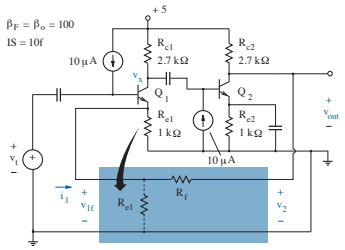


Figure 11.7: Circuit for Example 11.1.

Solution

Without feedback loading, the A-circuit voltage gain is moderately large. Thus, in a first approximation, we use $A = 133$ in the expression

$$\frac{v_{out}}{v_t} = \frac{A}{1 + AF} \,. \tag{11.18}$$

The feedback-stabilized gain objective of 10 is consistent with $F = 0.0925$. To find an expression for F, we observe that the arrangement of the A and F two-ports in the series-shunt configuration compels R_{e1} to participate with the latter circuit (for non-trivial F). Then with reference to the F two-port, we connect a voltage source at the output side where a *voltage* is shared, we make $i_1 = 0$ at the input side where *current* is shared, and we find v_{1f}. This process yields

$$F = \frac{v_{1f}}{v_2} = \frac{R_{e1}}{R_{e1} + R_f} \, . \tag{11.19}$$

So with $R_{e1} = 1$ kΩ, the previous F requirement demands $R_f = 9.81$ kΩ.

If this were an op-amp circuit with nearly infinite A, we would be done. For the circuit at hand, the preceding R_f value serves as a starting point for finding A subject to feedback loading effects. Once A has been modified, Eqs. 11.18 and 11.19 can be used to find a better R_f solution.

Transistors Q_1 and Q_2 are similarly biased with $i_{c1}|_Q = i_{c2}|_Q = 1$ mA, $g_{m1} = g_{m2} = 38.6 \times 10^{-3}$ ℧, and $r_{\pi 1} = r_{\pi 2} = 2.59$ kΩ. The loading effect on the input side of the feedback two-port produces an effective ${R_{e1}}' = R_{e1} \parallel R_f = 0.9$ kΩ (with shared $v_2 \to 0$ on the output side). Meanwhile, the loading effect on the output side of the feedback two-port produces an effective ${R_{c2}}' = R_{c2} \parallel (R_f + R_{e1}) = 2.13$ kΩ (with shared $i_1 \to 0$ on the input side). Together, these considerations establish (see Table 7.4)

$$A = \frac{v_{out}}{v_t} = \frac{-g_{m1}(R_{c1} \parallel r_{\pi 2})}{1 + g_{m1}{R_{e1}}'(1 + 1/\beta_o)} (-g_{m2}){R_{c2}}' = 116 \,. \tag{11.20}$$

In turn, we use Eqs. 11.18 and 11.19 to find a revised $R_f = 9.94$ kΩ ≈ 10 kΩ. The consistent feedback loading conditions are ${R_{e1}}' = 0.909$ kΩ and ${R_{c2}}' = 2.17$ kΩ so that $A = 117$ (an insignificant change). Finally, the voltage gain with feedback is $v_{out}/v_t = A/(1 + AF) = 10.1$. SPICE has the same result.

To find the new r_{out}, it is sufficient to evaluate

$$r_{out} = \frac{{R_{c2}}'}{1 + AF} = 187 \ \Omega \,. \tag{11.21}$$

We check with SPICE by replacing v_t with a short circuit, connecting an ac current source through a capacitor to the amplifier output, and examining the ratio of the the ac voltage produced by the test source to the ac current. (The capacitor preserves bias currents.) The SPICE result is $r_{out} = 188$ Ω, which is in good agreement.

Contemporary series-shunt feedback applications usually have a voltage output returned through a tapped divider circuit to a well-defined negative input of an op-amp (Example 11.1) or a differential amplifier (Chapter 9). Example 11.2 also demonstrates a circuit with positive and negative inputs. The input from v_t is positive because Q_1 has an output v_x with opposite sign (as for a common-emitter amplifier) that experiences another sign reversal when subjected to Q_2. The feedback-derived input is negative because Q_1 imposes the same sign on the output v_x (as for a common-base amplifier). The negative input helps us to understand a reduction in overall output resistance under feedback conditions. If we apply a test current at the Q_2 collector, part of the current is absorbed by R_{c2}, and the output node rises in relation to ac ground. Part of the test current also flows to the negative input through the feedback network, and the output node attempts to fall. Counteractive node behavior is consistent with reduced output resistance.

Example 11.2 successfully used the concepts of shared current or voltage when calculating the feedback factor F and the effects of feedback loading. Nevertheless, the cautious reader may have been disturbed when we cited "shared" i_1 on the input sides of the A and F two-ports. When R_{e1} moves to the F circuit as shown in Fig. 11.7, the input currents are not even close —they differ by a factor of $\beta_o + 1$.

We explore this difficulty by considering the voltage amplifier of Fig. 11.8. Here, the series-shunt feedback network influences the output circuit through the dependent current source with value $h_{21f}i_1$ and the loading of ${h_{22f}}^{-1}$. Nevertheless, the input and output relationships that apply to series-shunt feedback are based on an approximation in which h_{21f} is neglected relative to h_{21a} (Eq. 11.7a). In turn, we are free to neglect the dependent current source whether i_1 is properly shared or not.

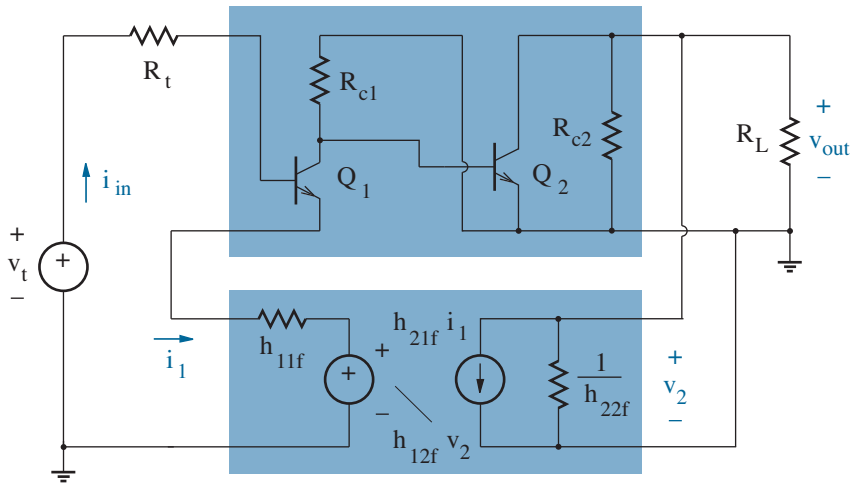


Figure 11.8: Small-signal voltage amplifier with series-shunt feedback.

The feedback network affects the input circuit through the dependent voltage source with value $h_{12f}v_2$ and the loading of h_{11f}. If we express Kirchhoff's voltage law around the input loop,

$$v_t = R_t i_{in} + r_{\pi 1} i_{in} + h_{11f} i_1 + h_{12f} v_2 \,, \tag{11.22}$$

we find that the dependent voltage source effectively acts in series with v_t —the related term can be moved over to the left-hand side of Eq. 11.22. Thus, the actual transfer of feedback information does not depend upon the current through the dependent voltage source.

We conclude that shared *input* current is not a feedback requirement.

11.2 Shunt-Shunt Feedback

We now examine the transresistance amplifier, which wants to have both small input resistance and small output resistance. With the hope of promoting these objectives, we examine an "ordinary" two-port amplifier and a *passive* two-port network that are interconnected as shown in Fig. 11.9. Here, the left-side parallel combination decreases the input resistance from the perspective of the Norton source, and the right-side parallel combination decreases the output resistance from the perspective of the connected load. The overall circuit configuration is an example of **shunt-shunt feedback**—the passive network samples v_2, the shared voltage at the amplifier output, and this information is "fed back" to influence v_1, the shared voltage at the amplifier input.

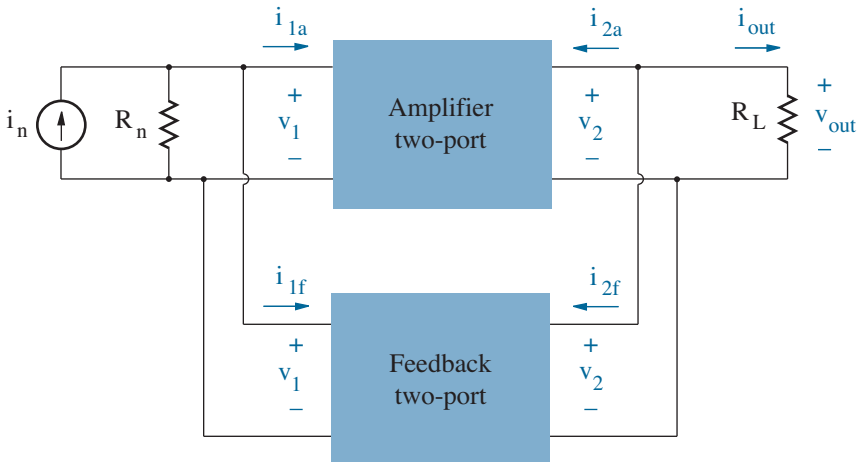


Figure 11.9: Shunt-shunt two-port feedback configuration.

The feedback analysis will require us to express the input and output two-port terminal currents in terms of the two node voltages. Thus, we look to the two-port ***y*-parameter** representation in which

$$i_1 = y_{11}v_1 + y_{12}v_2 \,, \qquad (11.23)\text{a}$$

$$i_2 = y_{21}v_1 + y_{22}v_2 \,. \qquad (11.23)\text{b}$$

The individual y parameters are easily determined. For example, one finds y_{11} by measuring i_1 subject to a v_1 voltage-source excitation while the two-port output terminal pair is shorted ($v_2 = 0$).

Exercise 11.3 Determine the y parameters for the two-port of Fig. 11.10.

Ans: $y_{11} = 1.5$ m℧ $y_{12} = -0.5$ m℧ $y_{21} = -0.5$ m℧ $y_{22} = 0.833$ m℧

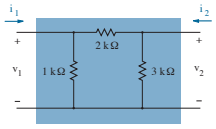


Figure 11.10: Circuit for Exercise 11.3.

Returning to the two-port arrangement of Fig. 11.9, we write Kirchhoff's current law at the input and output nodes. Specifically,

$$G_n v_1 + \underbrace{y_{11a} v_1 + y_{12a} v_2}_{i_{1a}} + \underbrace{y_{11f} v_1 + y_{12f} v_2}_{i_{1f}} = i_n \tag{11.24}$$

and

$$G_L v_2 + \underbrace{y_{21a} v_1 + y_{22a} v_2}_{i_{2a}} + \underbrace{y_{21f} v_1 + y_{22f} v_2}_{i_{2f}} = 0\,. \tag{11.25}$$

So in matrix form, we have

$$\begin{pmatrix} y_i & y_{12a}+y_{12f} \\ y_{21a}+y_{21f} & y_o \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} i_n \\ 0 \end{pmatrix}, \tag{11.26}$$

where

$$y_i = G_n + y_{11a} + y_{11f} \tag{11.27}$$

is the sum of the key admittance values at the input node, and

$$y_o = G_L + y_{22a} + y_{22f} \tag{11.28}$$

is the sum of the key admittance values at the output node.

Again, we make an approximation that acknowledges the tendency for amplifiers to propagate signals from input to output with large gain and from output to input with nearly zero gain. Specifically, in relation to the feedback two-port,

$$y_{21a} \gg y_{21f} \,, \tag{11.29a}$$
$$y_{12a} \ll y_{12f} \,. \tag{11.29b}$$

In turn,

$$\begin{pmatrix} y_i & y_{12f} \\ y_{21a} & y_o \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \approx \begin{pmatrix} i_n \\ 0 \end{pmatrix} \,. \tag{11.30}$$

Subject to $v_{out} = v_2$, the solution is

$$\frac{v_{out}}{i_n} = \frac{A}{1 + AF} \,, \tag{11.31}$$

where

$$A = \frac{-y_{21a}}{y_i y_o} \tag{11.32}$$

and

$$F = y_{12f} \,. \tag{11.33}$$

Equation 11.31 is the familiar feedback relationship.

It is little extra work to solve Eq. 11.30 for i_n/v_1. Then subtracting G_n, we obtain

$$y_{in} = y_i \,(1 + AF) - G_n \tag{11.34}$$

as the input admittance "seen" by the Norton circuit featuring i_n and R_n. This is larger than y_{11a}, the input admittance in the absence of feedback, and it can be very large if $AF \gg 1$. The reciprocal input impedance is accordingly small.

If (with $i_n = 0$) we apply a test current source in place of R_L and evaluate the voltage produced by this source—see Problem 11.2.15—we find

$$y_{out} = y_o \,(1 + AF) - G_L \tag{11.35}$$

as the output admittance "seen" by the load. This is larger than y_{22a}, the output admittance in the absence of feedback, and it can be very large if $AF \gg 1$. The reciprocal output impedance is accordingly small.

Example 11.3

The op-amp in the circuit of Fig. 11.11 features a differential voltage gain of 20,000, an input resistance of 500 kΩ, and an output resistance of 300 Ω. Determine v_{out}/i_n.

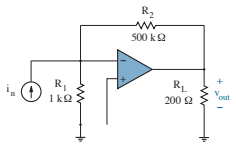


Figure 11.11: Circuit for Example 11.3.

Solution

We redraw the circuit as a pair of interconnected two-ports in the shunt-shunt configuration as shown in Fig. 11.12.

Note the positions of the ground connections.

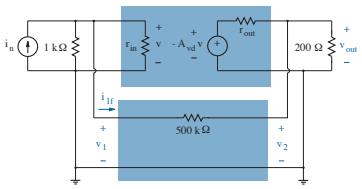


Figure 11.12: Shunt-shunt circuit for Example 11.3.

As in Example 11.1, we incorporate feedback loading by defining

$$R_n{}' = R_n \parallel \frac{1}{y_{11f}} \tag{11.36}$$

(so that $G_n{}' = G_n + y_{11f}$) and

$$R_L{}' = R_L \parallel \frac{1}{y_{22f}} \tag{11.37}$$

(so that $G_L{}' = G_L + y_{22f}$). Then we calculate A from the "A" circuit of Fig. 11.13. There is no feedback in this circuit ($F = 0$), and $v_{out}/i_n = A$.

Forget about y parameters— Resistance $R_n{}'$ is simply the parallel combination of R_n and the proper input resistance of the feedback two-port. The feedback process at the opposite (output) terminal pair is sampling of a voltage, so we null that shared information by making $v_2 = 0$. In turn, $R_n{}' = 1\ \text{k}\Omega \parallel 500\ \text{k}\Omega = 998\ \Omega$. Resistance $R_L{}'$ is simply the parallel combination of R_L and the proper output resistance of the feedback two-port. The feedback process at the opposite (input) terminal pair is the influence of a shared voltage, so we null that information by making $v_1 = 0$. In turn, $R_L{}' = 200\ \Omega \parallel 500\ \text{k}\Omega \approx 200\ \Omega$.

The "A"-circuit voltage gain is $v_{out}/v = -A_{vd}\, R_L{}'/(R_L{}' + r_{out}) =$

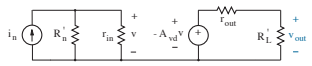


Figure 11.13: "A" circuit for Example 11.3.

-8000. Thus, $A = v_{out}/i_n = (v_{out}/v)(v/i_n) = (v_{out}/v)(R_n' \parallel r_{in}) = -7.98 \times 10^6\ \Omega$.

To find F, we note that the feedback two-port samples output voltage. Thus, we remove the sampling connections and apply a test *voltage* source with value v_2. The feedback two-port influences the shared input voltage, so we complete the f measurement by making $v_1 = 0$ and observing i_{1f}. In turn, $F = i_{1f}/v_2 = -2 \times 10^{-6}\ \mho$.

Having found A and F from separate measurements, we find that the overall amplifier transresistance is $v_{out}/i_n = A/(1 + AF) = -470\ \text{k}\Omega$.

Example 11.4

The circuit of Fig. 11.7 is intended to exhibit a transresistance of 200 kΩ. Complete the feedback design to achieve this behavior, and determine the change in performance over the temperature range -55 °C $\leq T \leq$ +125 °C.

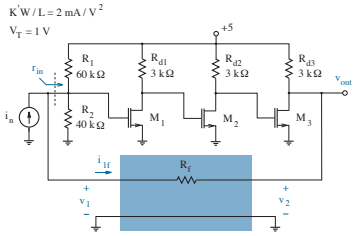


Figure 11.14: Circuit for Example 11.4.

Solution

We leave it as an exercise to show that the MOSFETs are similarly biased such that $i_d|_Q = 1$ mA and $g_m = 2 \times 10^{-3}$ ℧. Excluding feedback, each common-source amplifier has $A_{vm} = -2 \times 10^{-3}$ ℧ $\times\, 3 \times 10^3\ \Omega = -6$. Thus with $r_{in} = R_1 \parallel R_2 = 24$ kΩ, $A = v_{out}/i_n = (-6)^3 r_{in} = -5180$ kΩ. When feedback is present, however, we desire

$$\frac{v_{out}}{i_n} = \frac{A}{1 + AF} = -200\ \text{k}\Omega\,, \tag{11.38}$$

which is consistent with $F^{-1} = -208$ kΩ.

To find an expression for F, we connect a voltage source at the output side of the F two-port where a *voltage* is shared, we make $v_1 = 0$ at the input side where a *voltage* is shared, and we find i_{1f}. So we have

$$F = \frac{i_{1f}}{v_2} = \frac{-1}{R_f}\,. \tag{11.39}$$

In turn, $R_f = 208$ kΩ. As in Example 11.2, this value is a starting point for a new estimate of A subject to feedback loading effects.

The left- or right-side loading resistance that the F two-port presents is found by nulling the shared *voltage* at the opposite end. The result is R_f for each side. So $r_{in} = R_1 \parallel R_2 \parallel R_f = 21.5$ kΩ, $R_{d3}{}' = 3$ kΩ $\parallel R_f = 2.96$ kΩ, and $A = (-6)^2 r_{in} \times -2 \times 10^{-3}$ ℧ $\times\ 2.96 \times 10^3\ \Omega = -4580$ kΩ. We return to Eqs. 11.38 and 11.39 to find $R_f = 209$ kΩ, a value not far from the original. A quick SPICE simulation reveals $R_m = 199.9$ kΩ as the transresistance, and the design is complete.

To check the design performance over temperature, we apply a SPICE simulation similar to that in Example 7.8. Manufacturer's data show that the resistors have +100 ppm/°C linear temperature variation, so the circuit netlist has device statements of the form

```
Rd1        1          3          RModel     3k
```

with the corresponding .model statement

```
.model     Rmodel     RES        (TC1=100E-6)
```

The simulation command statements are

```
.ac        LIN        1          1k         1k
.temp      -55   -35  -15   5    25   45    65   85    105   125
.print     ac         v(5)       I(In)
```

subject to v_{out} at node 5 and "In" as the ac current source at the input. The amplifier transresistance is the ratio of the v(5) and I(In) ac amplitudes listed in the simulation output file for the various temperatures specified in the .temp statement.

Here are the results of the simulations:

T °C	R_m (kΩ, Feedback)	R_m (kΩ, No Feedback)
-55	202.5	—
-35	202.9	14.47
-15	202.1	51.43
5	201.1	172.3
25	200.0	5674
45	198.7	1343
65	197.1	—
85	195.3	—
105	193.3	—
125	191.0	—

The performance with feedback is excellent except for large temperatures, while the circuit behavior is horrible without feedback. Indeed, the missing results reflect unacceptable biasing. Apart from stabilizing transresistance, R_f helps to ensure that the MOSFETs have steady Q points.

11.3 Shunt-Series Feedback

Unlike the voltage amplifier, the current amplifier wants to have small input resistance and large output resistance. And in view of our successful two-port manipulations in preceding sections, it seems worthwhile to examine an "ordinary" two-port amplifier and a *passive* two-port network that are interconnected as shown in Fig. 11.15. The left-side parallel combination decreases the input resistance from the perspective of the Norton source, and the right-side series combination increases output resistance from the perspective of the load. The overall circuit configuration is an example of **shunt-series feedback**—the passive network samples i_2, a loop current presumably shared at the output sides of the amplifier / feedback two-ports, and the information is "fed back" to influence v_1, a voltage shared at the amplifier input. (The former sharing is open to question.)

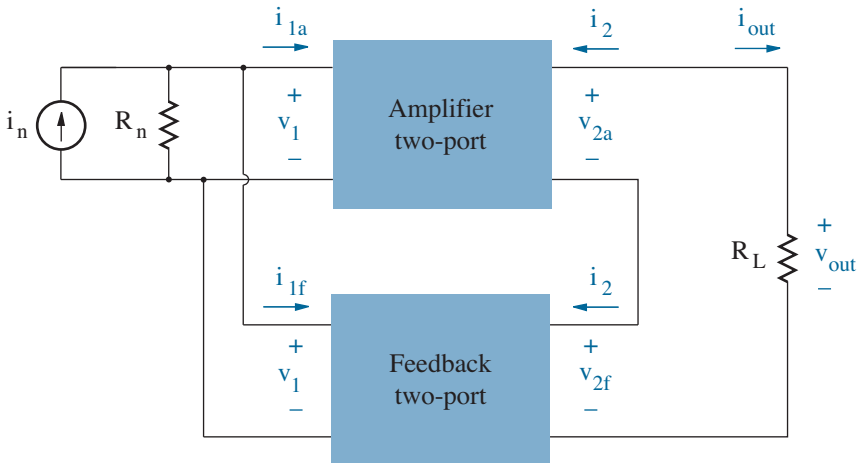


Figure 11.15: Shunt-series two-port feedback configuration.

The forthcoming feedback analysis will require us to express the two-port input currents and output voltages in terms of the shared input voltage and output current. Thus, we use a two-port ***g*-parameter** representation in which

$$i_1 = g_{11}v_1 + g_{12}i_2 \,, \qquad (11.40)\text{a}$$

$$v_2 = g_{21}v_1 + g_{22}i_2 \,. \qquad (11.40)\text{b}$$

The individual g parameters are easily determined. For example, one finds g_{11} by measuring i_1 subject to a v_1 voltage-source excitation while the two-port output terminal pair is open ($i_2 = 0$).

Exercise 11.4 Determine the g parameters for the two-port of Fig. 11.16.

Ans: $g_{11} = 1.2$ m℧ $g_{12} = -0.6$ $g_{21} = 0.6$ $g_{22} = 1.2$ kΩ

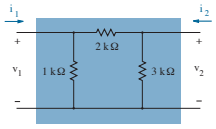


Figure 11.16: Circuit for Exercise 11.4.

Returning to the two-port arrangement of Fig. 11.15, we write Kirchhoff's current law at the input node,

$$G_n v_1 + \underbrace{g_{11a} v_1 + g_{12a} i_2}_{i_{1a}} + \underbrace{g_{11f} v_1 + g_{12f} i_2}_{i_{1f}} = i_n \,, \tag{11.41}$$

and we write Kirchhoff's voltage law around the output loop,

$$R_L i_2 + \underbrace{g_{21a} v_1 + g_{22a} i_2}_{v_{2a}} + \underbrace{g_{21f} v_1 + g_{22f} i_2}_{v_{2f}} = 0 \,. \tag{11.42}$$

So in matrix form, we have

$$\begin{pmatrix} y_i & g_{12a} + g_{12f} \\ g_{21a} + g_{21f} & z_o \end{pmatrix} \begin{pmatrix} v_1 \\ i_2 \end{pmatrix} = \begin{pmatrix} i_n \\ 0 \end{pmatrix}, \tag{11.43}$$

where

$$y_i = G_n + g_{11a} + g_{11f} \tag{11.44}$$

is the sum of the key admittance values at the input node, and

$$z_o = R_L + g_{22a} + g_{22f} \tag{11.45}$$

is the sum of the key impedance values in the output loop.

Again, we make an approximation that acknowledges the tendency for amplifiers to propagate signals from input to output with large gain and from output to input with nearly zero gain. Specifically, in relation to the feedback two-port,

$$g_{21a} \gg g_{21f} \,, \tag{11.46a}$$

$$g_{12a} \ll g_{12f} \,. \tag{11.46b}$$

In turn,

$$\begin{pmatrix} y_i & g_{12f} \\ g_{21a} & z_o \end{pmatrix} \begin{pmatrix} v_1 \\ i_2 \end{pmatrix} \approx \begin{pmatrix} i_n \\ 0 \end{pmatrix} \,. \tag{11.47}$$

Subject to $-i_{out} = i_2$, the solution is

$$\frac{-i_{out}}{i_n} = \frac{A}{1+AF} \,, \tag{11.48}$$

where

$$A = \frac{-g_{21a}}{y_i z_o} \tag{11.49}$$

and

$$F = g_{12f} \,. \tag{11.50}$$

Equation 11.48 is the familiar feedback relationship.

It is little extra work to solve Eq. 11.47 for i_n/v_1. Then subtracting G_n, we obtain

$$y_{in} = y_i \,(1+AF) - G_n \tag{11.51}$$

as the input admittance "seen" by the Norton circuit featuring i_n and R_n. This is larger than g_{11a}, the input admittance in the absence of feedback, and it can be very large if $AF \gg 1$. The reciprocal input impedance is accordingly small.

If (with $i_n = 0$) we apply a test voltage source in place of R_L and evaluate the current produced by this source—see Problem 11.27—we find

$$z_{out} = z_o \,(1+AF) - R_L \tag{11.52}$$

as the output impedance "seen" by the load. This is also larger than g_{22a}, the output impedance in the absence of feedback, and it can be very large if $AF \gg 1$.

Example 11.5

Determine midfrequency $-i_{out}/i_n$ in the circuit of Fig. 11.17.

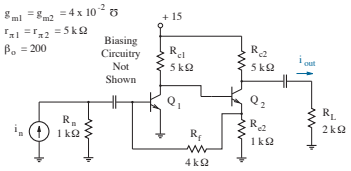


Figure 11.17: Circuit for Example 11.5.

Solution

We redraw the mid-frequency circuit as a pair of interconnected two-ports in the shunt-series configuration as shown in Fig. 11.18.

Note the positions of the ground connections.

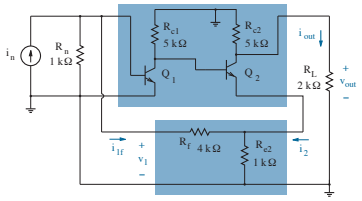


Figure 11.18: Shunt-series circuit for Example 11.5.

Caution: The formalism that led to the fundamental feedback relations was based upon a simple two-port interconnection rule:

The two-ports *share* current or voltage at input and output.

This rule is easily violated when series connections are casually applied in conjunction with multiple grounds as in the circuit of Fig. 11.18. If we eliminate the ground connection that is internal to the amplifier two-port, we obtain Fig. 11.19 in which $i_2 \neq -i_{out}$. And even after this modification, the two ground connections ensure that i_2 differs from $-i_{out}{}'$ by a factor of $(1 + 1/\beta_o)$. Nevertheless, the remaining difference is minor if $\beta_o \gg 1$.

(Shared node voltages tend not to suffer from multiple ground connections.)

So our actual task at hand is to determine the feedback stabilization that applies to $-i_{out}{}'/i_n$. In turn, we evaluate $-i_{out}/i_n$ by applying the current divider rule.

As in Example 11.1, we can save some time by observing that quantity A in Eq. 11.49 is only influenced by the feedback two-port through the terms y_i and z_o, which partially feature $g_{11f} + G_n$ and $g_{22f} + R_L$, respectively. Thus, by defining

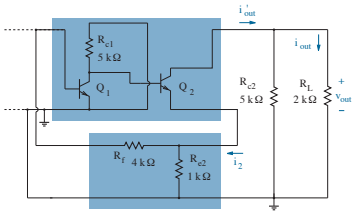


Figure 11.19: Modified shunt-series circuit for Example 11.5.

$$R_n{}' = R_n \parallel \frac{1}{g_{11f}} \tag{11.53}$$

(so that $G_n{}' = G_n + g_{11f}$), we simplify the input side of an appropriate "A" circuit. Shunt loading can always be treated this way.

We are similarly inclined to simplify the output side of the "A" circuit by defining $R_L{}' = (R_{c2} \parallel R_L) + g_{22f}$. However, as shown in Fig. 11.20, the process is complicated through the presence of the right-side ground connection at the node common to $R_{c2} \parallel R_L$ and g_{22f}. Series combinations require shared current, which is disrupted here (to account for $i_2 \neq -i_{out}{}'$).

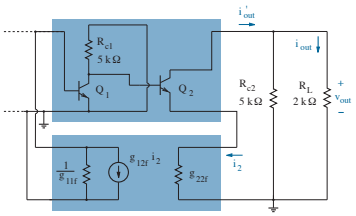


Figure 11.20: Shunt-series circuit for Example 11.5 with general feedback. The feedback two-port model assumes $g_{21} \to 0$.

A less perilous simplification procedure is to move g_{22f} up into the amplifier two-port so that it is in series with the Q_2 emitter. This leaves behind a short circuit at the output side of the feedback two-port, but the sampled loop current (i_2) is unaffected. The complete "A" circuit now has the form shown in Fig. 11.21 (with $R_{e2}{}' = g_{22f}$).

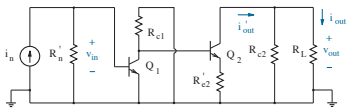


Figure 11.21: "A" circuit for Example 11.5.

Forget about g parameters— Resistance ${R_n}'$ is the parallel combination of R_n and the input resistance of the feedback two-port. The feedback process at the opposite (output) terminal pair is sampling of a shared current, so we make $i_2 = 0$. In turn, ${R_n}' = 1\ \text{k}\Omega \parallel (4\ \text{k}\Omega + 1\ \text{k}\Omega) = 0.833\ \text{k}\Omega$. Emitter resistance ${R_{e2}}'$ is the output resistance of the feedback two-port. The feedback process at the opposite (input) terminal pair is the influence of a shared voltage, so we make $v_1 = 0$, and ${R_{e2}}' = 4\ \text{k}\Omega \parallel 1\ \text{k}\Omega = 0.8\ \text{k}\Omega$.

The voltage gain for the "A" circuit is

$$\frac{v_{out}}{v_{in}} = \frac{g_{m1}\left\{R_{c1} \parallel [r_{\pi 2} + (\beta_o + 1){R_{e2}}']\right\} g_{m2}(R_{c2} \parallel R_L)}{1 + g_{m2}{R_{e2}}'(1 + 1/\beta_o)} = 334\,. \quad (11.54)$$

So with

$$A = -{i_{out}}'/i_n = (v_{out}/v_{in}) \times (v_{in}/i_n) \times (-{i_{out}}'/v_{out})\,, \quad (11.55)$$

we evaluate

$$\frac{v_{in}}{i_n} = {R_n}' \parallel r_{\pi 1} = 0.714\ \text{k}\Omega \quad (11.56)$$

and

$$\frac{-i_{out}{}'}{v_{out}} = \frac{-1}{R_{c2} \parallel R_L} = -1.43 \text{ m}\mho \tag{11.57}$$

to obtain $A = -167$.

To find F, we note that the feedback two-port samples current in the output loop. Thus, we break that loop and apply a test *current* source with value i_2. The feedback two-port influences an input voltage, so we complete the set-up for the F measurement by making $v_1 = 0$ (with a short circuit) and observing i_{1f}. In turn, $F = i_{1f}/i_2 = -0.2$.

Having found A and F from separate measurements, we find that the feedback-stablized amplifier current gain is $-i_{out}{}'/i_n = A/(1 + AF) = -4.85$. Then $-i_{out}/i_n = -i_{out}{}'/i_n \times G_L/(G_L + G_{c2}) = -3.46$.

To determine the effective input resistance of the feedback amplifier, we evaluate ${y_i}^{-1}$ as the parallel combination of ${R_n}'$ and the input resistance of the two-port amplifier (5 kΩ). Then with $y_i = 1.40$ m$\mho$, we use Eq. 11.51 to find $z_{in} = 21$ Ω. The effective output resistance is R_{c2} in parallel with something much larger (per Eq. 11.52), so it is approximately 5 kΩ.

11.4 Series-Series Feedback

Finally, we consider the transconductance amplifier, which wants to have both large input resistance and large output resistance. With hope of promoting both objectives, we examine an "ordinary" two-port amplifier and a *passive* two-port network that are interconnected as shown in Fig. 11.22. The left-side series combination increases the input resistance from the perspective of the Thevenin source, and the right-side series combination increases the output resistance from the perspective of the load. The overall circuit configuration is generally an example of **series-series feedback**—the passive network samples i_2, the loop current at the amplifier output, and the information is "fed back" to influence i_1, the loop current at the amplifier input. The loop-current sharing is presumed with caution.

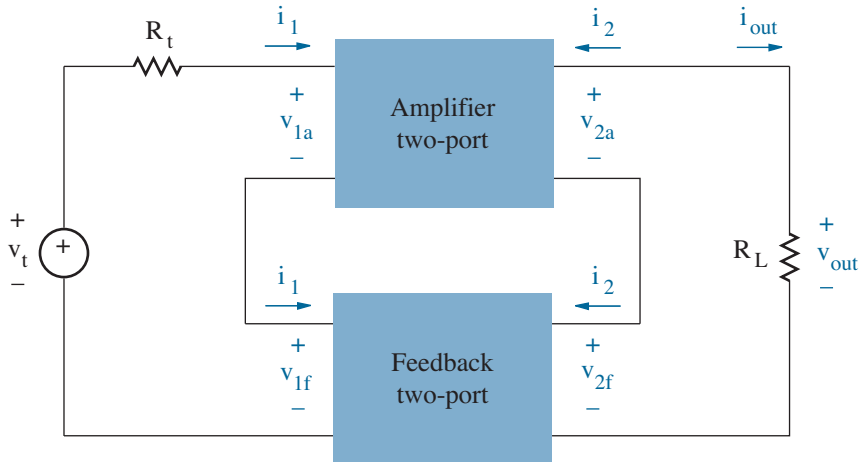


Figure 11.22: Series-series two-port feedback configuration.

The feedback analysis will require us to express the two-port input and output voltages in terms of the two loop currents. Thus, we look to the two-port z-**parameter** representation in which

$$v_1 = z_{11}i_1 + z_{12}i_2\,, \qquad (11.58)\text{a}$$

$$v_2 = z_{21}i_1 + z_{22}i_2\,. \qquad (11.58)\text{b}$$

The individual z parameters are easily determined. For example, one finds z_{11} by measuring v_1 subject to an i_1 current-source excitation while the two-port output terminal pair is open ($i_2 = 0$).

Exercise 11.5 Determine the z parameters for the two-port of Fig. 11.23.

Ans: $z_{11} = 0.833\ \text{k}\Omega$ $z_{12} = 0.5\ \text{k}\Omega$ $z_{21} = 0.5\ \text{k}\Omega$ $z_{22} = 1.5\ \text{k}\Omega$

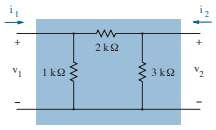


Figure 11.23: Circuit for Exercise 11.5.

Returning to the two-port arrangement of Fig. 11.22, we write Kirchhoff's voltage law around the input and output loops. Specifically,

$$R_t i_1 + \underbrace{z_{11a} i_1 + z_{12a} i_2}_{v_{1a}} + \underbrace{z_{11f} i_1 + z_{12f} i_2}_{v_{1f}} = v_t \tag{11.59}$$

and

$$R_L i_2 + \underbrace{z_{21a} i_1 + z_{22a} i_2}_{v_{2a}} + \underbrace{z_{21f} i_1 + z_{22f} i_2}_{v_{2f}} = 0\,. \tag{11.60}$$

So in matrix form, we have

$$\begin{pmatrix} z_i & z_{12a} + z_{12f} \\ z_{21a} + z_{21f} & z_o \end{pmatrix} \begin{pmatrix} i_1 \\ i_2 \end{pmatrix} = \begin{pmatrix} v_t \\ 0 \end{pmatrix}, \tag{11.61}$$

where

$$z_i = R_t + z_{11a} + z_{11f} \tag{11.62}$$

is the sum of the key impedance values in the input loop, and

$$z_o = R_L + z_{22a} + z_{22f} \tag{11.63}$$

is the sum of the key impedance values in the output loop.

Again, we make an approximation that acknowledges the tendency for amplifiers to propagate signals from input to output with large gain and from output to input with nearly zero gain. Specifically, in relation to the feedback two-port,

$$z_{21a} \gg z_{21f} \,, \tag{11.64a}$$
$$z_{12a} \ll z_{12f} \,. \tag{11.64b}$$

In turn,

$$\begin{pmatrix} z_i & z_{12f} \\ z_{21a} & z_o \end{pmatrix} \begin{pmatrix} i_1 \\ i_2 \end{pmatrix} \approx \begin{pmatrix} v_t \\ 0 \end{pmatrix} \,. \tag{11.65}$$

Subject to $-i_{out} = i_2$, the solution is

$$\frac{-i_{out}}{v_t} = \frac{A}{1+AF} \,, \tag{11.66}$$

where

$$A = \frac{-z_{21a}}{z_i z_o} \tag{11.67}$$

and

$$F = z_{12f} \,. \tag{11.68}$$

Equation 11.66 is the familiar feedback relationship.

It is little extra work to solve Eq. 11.65 for v_t/i_1. Then subtracting R_t, we obtain

$$z_{in} = z_i \,(1+AF) - R_t \tag{11.69}$$

as the input impedance "seen" by the Thevenin circuit featuring v_t and R_t. This is indeed larger than z_{11a}, the input impedance in the absence of feedback, and it can be very large if $AF \gg 1$.

If (with $v_t = 0$) we apply a test voltage source in place of R_L and find the current supplied from this source—see Problem 11.31—we obtain

$$z_{out} = z_o \,(1+AF) - R_L \tag{11.70}$$

as the output impedance "seen" by the load. This is also larger than z_{22a}, the output impedance in the absence of feedback, and it can be very large if $AF \gg 1$.

Perhaps we did not appreciate it at the time, but we used series-series feedback in Chapter 7 when we chose not to remove a bypass capacitor to obtain a g_m-insensitive "common emitter" or "common source" amplifier. As shown in Fig. 11.24, the small-signal circuit for the BJT case features a true common-emitter amplifier (with the emitter node shared between input and output) and a feedback two-port containing R_e. To simplify the analysis that follows, we leave out R_t to avoid a loading factor.

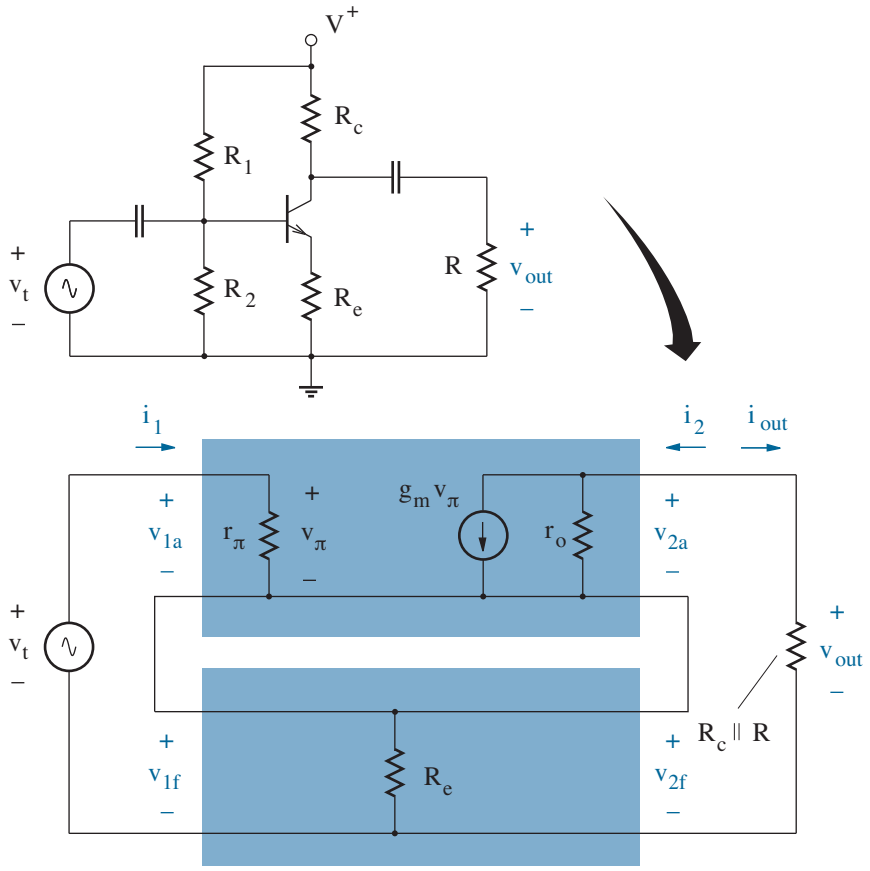


Figure 11.24: Common-emitter amplifier with feedback.

For a change, we apply the two-port parameters. Along the input loop,

$$z_i = z_{11a} + z_{11f} + R_t = r_\pi + R_e + 0\,. \tag{11.71}$$

And along the output loop,

$$z_o = z_{22a} + z_{22f} + R_L = r_0 + R_e + R_c \parallel R\,. \tag{11.72}$$

Then with $g_m r_\pi = \beta_o$,

$$z_{21a} = \left.\frac{v_{2a}}{i_1}\right|_{i_2=0} = -\beta_o r_o \,, \tag{11.73}$$

and

$$A = \frac{-z_{21a}}{z_i z_o} = \frac{\beta_o r_o}{(r_\pi + R_e)(r_o + R_e + R_c \parallel R)} \,. \tag{11.74}$$

The r_o output resistance for a BJT is generally large. Thus,

$$A \approx \frac{\beta_o}{r_\pi + R_e} \,. \tag{11.75}$$

Meanwhile,

$$F = z_{12f} = \left.\frac{v_{1f}}{i_2}\right|_{i_1=0} = R_e \,. \tag{11.76}$$

In turn, (after some algebraic manipulation),

$$\frac{-i_{out}}{v_t} = \frac{A}{1 + AF} = \frac{\beta_o}{r_\pi + (\beta_o + 1)R_e} \,. \tag{11.77}$$

The more familiar voltage gain is

$$\frac{v_{out}}{v_t} = \frac{-g_m(R_c \parallel R)}{1 + g_m R_e(1 + 1/\beta_o)} \,, \tag{11.78}$$

since $v_{out} = i_{out}(R_c \parallel R)$. Finally, the input resistance is given by

$$r_{in} = z_i(1 + AF) - R_t = r_\pi + (\beta_o + 1)R_e \tag{11.79}$$

from the perspective of v_t, and the output resistance is given by

$$r_{out} = z_o(1 + AF) - R_L \approx r_o\,(1 + \frac{g_m R_e}{1 + R_e/r_\pi}) \tag{11.80}$$

from the perspective of $R_c \parallel R$, the effective "load" for feedback analysis. Nevertheless, it is common practice to lump R_c with the internal amplifier so that

$$r_{out} = R_c \tag{11.81}$$

from the perspective of R, the load for the amplifier.

Feedback makes the BJT a better transconductance device that draws current through $R_c \parallel R$ in response to a voltage change. This is a natural extension, given the form of the small-signal transistor model.

Example 11.6

Determine midfrequency $-i_{out}/v_t$ in the circuit of Fig. 11.25.

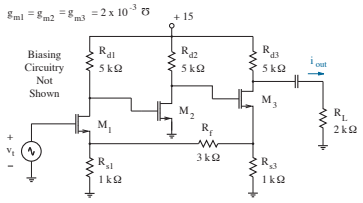


Figure 11.25: Circuit for Example 11.6.

Solution

We redraw the mid-frequency circuit as a pair of interconnected two-ports in the series-series configuration as shown in Fig. 11.26.

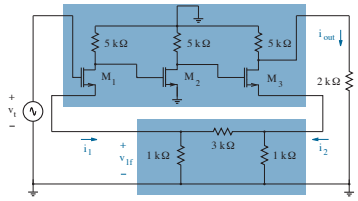


Figure 11.26: Series-series circuit for Example 11.6.

Despite ground worries, we observe that the output side of the circuit is similar to the shunt-series configuration of Example 11.4. So we look to $-i_{out}{}'$ as a shared current whose negative feeds $R_{d3} \parallel R_L$. Current is not shared on the input side of the circuit, but this is not a strict requirement (as demonstrated in Section 11.1 for series-shunt feedback).

We complete the "A" circuit of Fig. 11.27 by including feedback loading as effective source resistors for M_1 and M_3. To find ${R_{s1}}'$, we determine the input resistance of the feedback two-port while nulling shared feedback information at the opposite (output) terminal pair so that $i_2 = 0$. In turn, ${R_{s1}}' = 1\ \text{k}\Omega \parallel (3\ \text{k}\Omega + 1\text{k}\Omega) = 800\ \Omega$. A similar process at the output side of the feedback two-port yields ${R_{s3}}' = 800\ \Omega$.

The voltage gain for the 3-stage "A" circuit is

$$\frac{v_{out}}{v_t} = \left(\frac{-g_{m1}R_{d1}}{1 + g_{m1}{R_{s1}}'}\right)(-g_{m2}R_{d2})\left[\frac{-g_{m3}(R_{d3} \parallel R_L)}{1 + g_{m3}{R_{s3}}'}\right] = -42.3\,. \quad (11.82)$$

Thus, $A = -i_{out}{}'/v_t = (-i_{out}{}'/v_{out})(v_{out}/v_t) = 29.6 \times 10^{-3}\ \mho$.

In the series-series configuration, the feedback two-port samples current in the output loop. Thus, we break that loop and apply a test *current* source with value i_2. The shared input variable is a loop current, so we complete the set-up for the F measurement by making $i_1 = 0$ and observing v_{1f}.

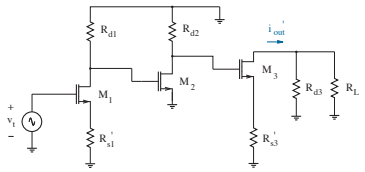


Figure 11.27: "A" circuit for Example 11.6.

In turn, $F = v_{1f}/i_2 = 200\ \Omega$.

Having found A and F from separate measurements, we find that the amplifier transconductance is $-i_{out}{}'/v_t = A/(1 + AF) = 4.28 \times 10^{-3}\ \mho$. Then with the help of the current divider rule, $i_{out}/v_t = -3.1 \times 10^{-3}\ \mho$.

Concept Summary

A two-port amplifier can be combined with a two-port passive feedback network to promote a particular performance attribute.

- Two-port series connections yield *shared* current and increased combined resistance; whereas, two-port shunt (parallel) connections yield *shared* current and decreased combined resistance.
- With respect to the input and output sides of the amplifier two-port,
  - Series-shunt feedback improves a voltage amplifier;
  - Shunt-shunt feedback improves a transresistance amplifier;
  - Shunt-series feedback improves a current amplifier;
  - Series-series feedback improves a transconductance amplifier.
- The overall system response has the form

$$H = \frac{\text{shared output variable}}{\text{unshared input source}} = \frac{A}{1 + AF} \ ,$$

  where typically $A \gg 1/F$. In turn, H is insensitive to A variations, an indication of reduced distortion.
- To find A:
  - Construct an "A" circuit with feedback loading —
    - The loading from the input side of the feedback two-port is the input resistance with the shared output set to zero.
    - The loading from the output side of the feedback two-port is the output resistance with the shared input set to zero.
  - With no other feedback influence, $H = A$.
- To find F:
  - Apply a test source at the output side of the feedback two-port with the character of the shared information.
  - Find the response at the input side of the feedback two-port subject to zero value for the shared input variable.

  (Zero-value current implies an open circuit.)

  (Zero-value voltage implies a short circuit.)

Problems

Perspective

11.1 Find an application for each of the following: voltage amplifier, transresistance amplifier, current amplifier, and transconductance amplifier. Try not to duplicate the examples in the text.

11.2 Specify the type of amplifier that is most likely appropriate for the following (Thevenin or Norton) source and load conditions. Explain your reasoning.

(a) R_t (R_n) = 20 kΩ, R_L = 20 kΩ.

(b) R_t (R_n) = 100 Ω, R_L = 10 Ω.

(c) R_t (R_n) = 80 kΩ, R_L = 50 Ω.

(d) R_t (R_n) = 50 Ω, R_L = 8 kΩ.

Section 11.1

11.3 Prove the expression for output admittance in Eq. 11.13.

11.4 The op-amp in the circuit of Fig. P11.4 has a differential voltage gain of 50,000, an input resistance of 800 kΩ, and an output resistance of 200 Ω. Determine v_{out}/v_t, z_{in}, and z_{out}.

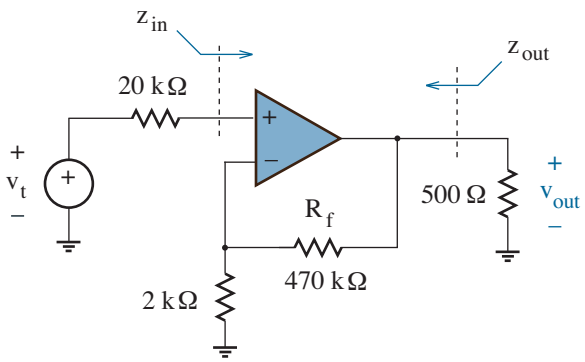


Figure P11.4

11.5 Repeat Problem P11.4, but assume a poor op-amp with a differential voltage gain of 5000, an input resistance of 80 kΩ, and 2-kΩ output resistance.

11.6 Change R_f in the circuit of Problem P11.4 so that $v_{out}/v_t = 800$.

11.7 Change R_f in the circuit of Problem P11.4 so that $v_{out}/v_t = 2800$.

11.8 The op-amp in the circuit of Fig. P11.8 has a differential voltage gain of 100,000, an input resistance of 400 kΩ, and an output resistance of 80 Ω. Determine v_{out}/v_t.

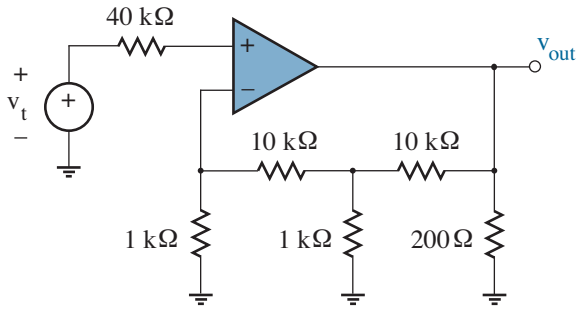


Figure P11.8

11.9 Repeat Example 11.2 with $R_{c1} = R_{c2} = 3.3$ kΩ, $R_{e1} = R_{e2} = 470$ Ω, and $i_{b1}|_Q = i_{b2}|_Q = 4$ μA. Verify with SPICE.

11.10 The MOSFET in the circuit of Fig. P11.10 has $K'W/L = 2$ mA/V^2 and $V_T = 1$ V.

(a) Determine $i_d|_Q$.

(b) Redraw the circuit as a pair of interconnected two-ports.

(c) Determine v_{out}/v_t.

(d) Use SPICE to verify the result of part c.

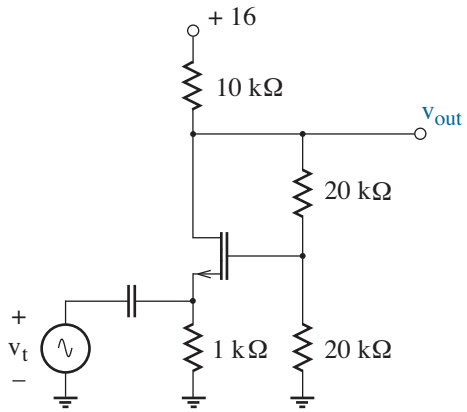


Figure P11.10

11.11 The MOSFETs in the circuit of Fig. P11.11 have $K'W/L = 2$ mA/V^2 and $V_T = 1$ V.

(a) Determine $i_{d1}|_Q$ and $i_{d2}|_Q$.

(b) Redraw the circuit as a pair of interconnected two-ports.

(c) Determine v_{out}/v_t.

(d) Use SPICE to verify the result of part c.

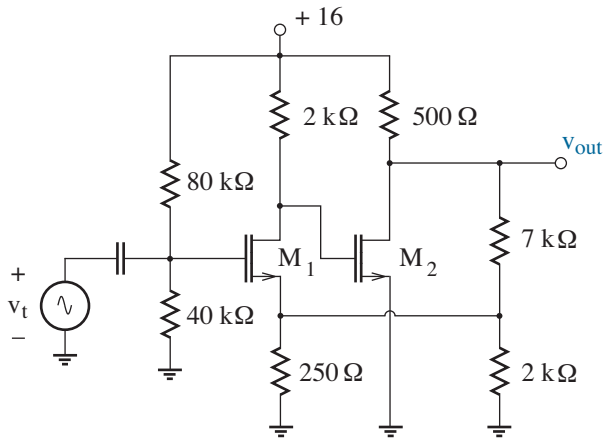


Figure P11.11

11.12 The BJTs in the circuit of Fig. P11.12 have $\beta_F = \beta_o = 200$.

(a) Use SPICE to find v_{out}/v_t and the quiescent BJT collector currents. Assume IS=10f.

(b) Redraw the circuit as a pair of interconnected two-ports.

(c) Apply the SPICE biasing results to calculate v_{out}/v_t by hand.

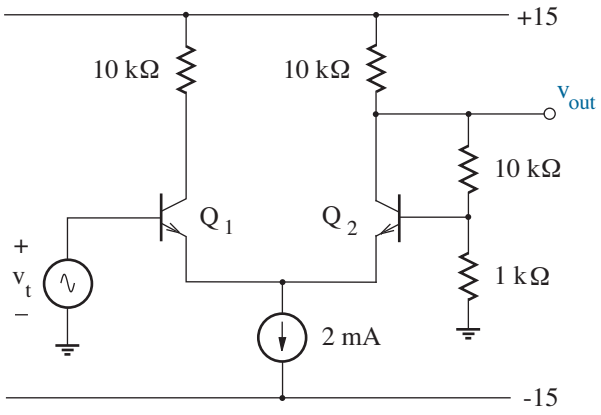


Figure P11.12

11.13 The MOSFETs of Fig. P11.13 have

$K_n' = 50\ \mu$A/V^2, $V_{Tn} = +0.5$ V, $\lambda_n = 0.05$ V^{-1}
$K_p' = 20\ \mu$A/V^2, $V_{Tp} = -0.5$ V, $\lambda_p = 0.1$ V^{-1}

Complete the feedback design for $A_{vm} = 50$.

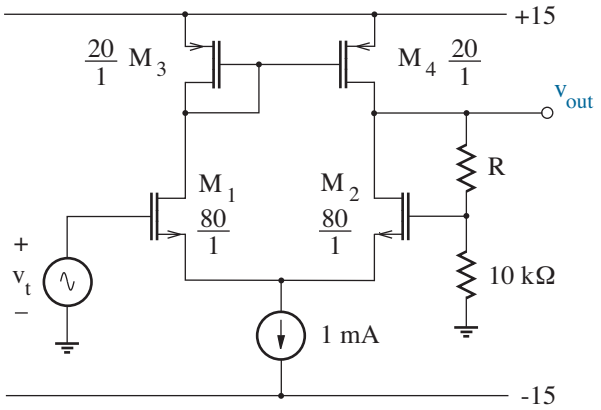


Figure P11.13

11.14 The MOSFETs of Fig. P11.14 have

$K_n' = 50\ \mu\text{A/V}^2$, $V_{Tn} = +0.5$ V, $\lambda_n = 0.05\ \text{V}^{-1}$
$K_p' = 20\ \mu\text{A/V}^2$, $V_{Tp} = -0.5$ V, $\lambda_p = 0.1\ \text{V}^{-1}$

Determine v_{out}/v_{in}.

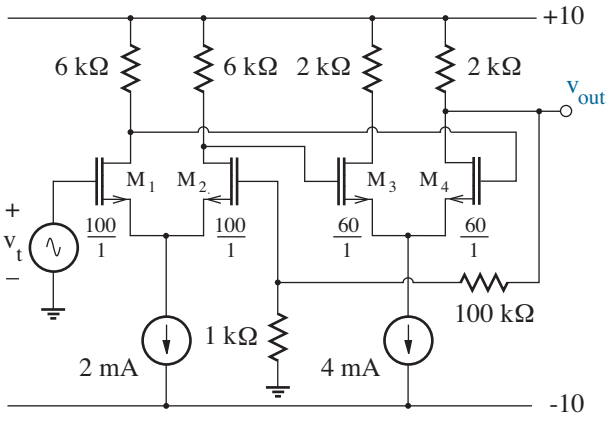


Figure P11.14

Section 11.2

11.15 Prove the expression for output admittance in Eq. 11.35.

11.16 The op-amp in the circuit of Fig. P11.16 has a differential voltage gain of 50,000, an input resistance of 800 kΩ, and an output resistance of 200 Ω. Determine v_{out}/i_n, z_{in}, and z_{out}.

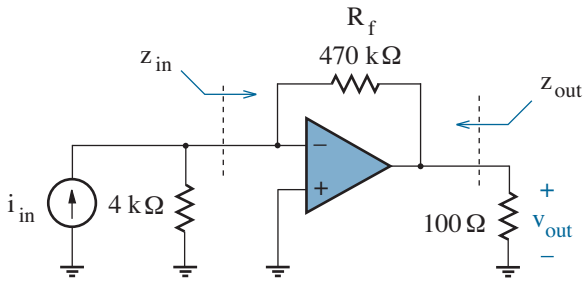


Figure P11.16

11.17 Repeat Problem P11.16, but assume an op-amp with a differential voltage gain of 5000, an input resistance of 80 kΩ, and 2-kΩ output resistance.

11.18 Change R_f in the circuit of Problem P11.16 so that $v_{out}/i_n = -800$ kΩ.

11.19 Change R_f in the circuit of Problem P11.16 so that $v_{out}/i_n = -2800$ kΩ.

11.20 The op-amp in the circuit of Fig. P11.20 has a differential voltage gain of 100,000, an input resistance of 400 kΩ, and an output resistance of 80 Ω. Determine v_{out}/i_n.

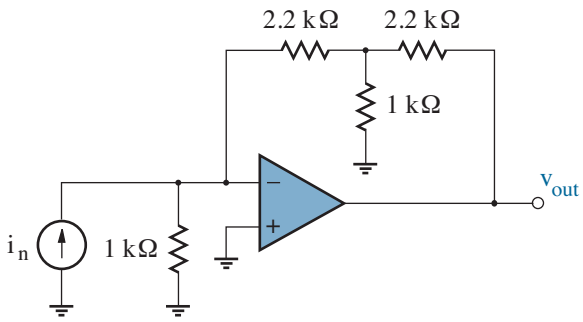


Figure P11.20

11.21 Repeat Example 11.4, but let the circuit have identical BJTs for which $\beta_F = \beta_o = 200$. Redesign the circuit (emitter resistors and bypass capacitors) so that the quiescent collector currents are 1 mA. Do not change R_1 or R_2. For the SPICE simulation, let IS=10f, XTB=1.5, and let the resistors exhibit a temperature variation of +100 ppm/°C.

11.22 The BJT in the circuit of Fig. P11.22 has $\beta_F = \beta_o = 200$ and $I_s = 10^{-14}$ A.

(a) Determine $i_c|_Q$ (subject to $i_n = 0$).

(b) Redraw the circuit as a pair of interconnected two-ports.

(c) Determine v_{out}/i_n.

(d) Use SPICE to verify the result of part c.

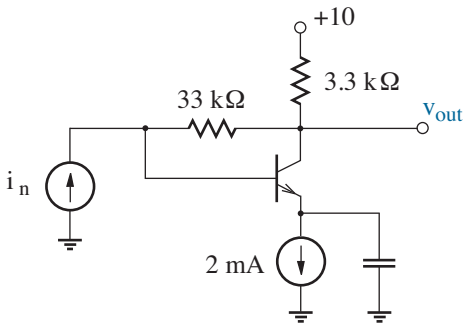


Figure P11.22

11.23 The BJTs in the circuit of Fig. P11.23 have $\beta_F = \beta_o = 200$ and $I_s = 10^{-14}$ A.

(a) Determine $i_{c1}|_Q$ and $i_{c2}|_Q$ (subject to $i_n = 0$).
Hint: Mark up the circuit diagram in terms of current i.

(b) Redraw the circuit as a pair of interconnected two-ports.

(c) Determine v_{out}/i_n.

(d) Use SPICE to verify the result of part c.

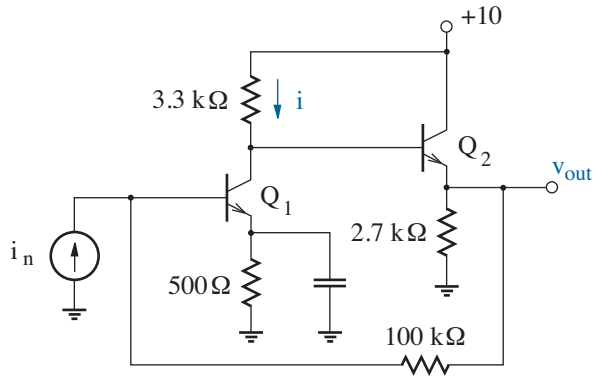


Figure P11.23

11.24 The MOSFETs in the circuit of Fig. P11.24 have $K'W/L = 2$ mA/V^2 and $V_T = \pm 1$ V.

(a) Determine $i_{d1}|_Q$ and $i_{d2}|_Q$ (subject to $i_n = 0$).

(b) Redraw the circuit as a pair of interconnected two-ports.

(c) Determine v_{out}/i_n.

(d) Use SPICE to verify the result of part c.

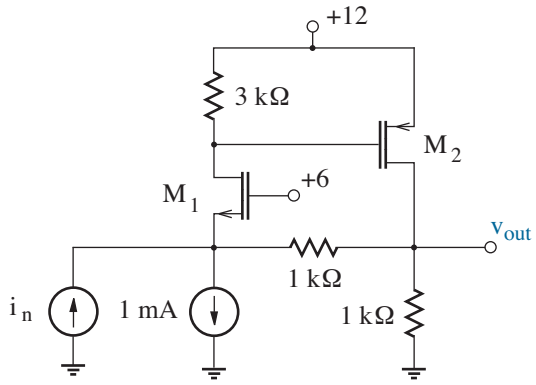


Figure P11.24

11.25 The BJTs of Fig. P11.25 have $\beta_F = \beta_o = 200$.

(a) Use SPICE to find v_{out}/i_n and the quiescent collector currents. Let IS=10f and VAF=40.

(b) Redraw the circuit as a pair of interconnected two-ports.

(c) Apply the SPICE biasing results to calculate v_{out}/i_n by hand.

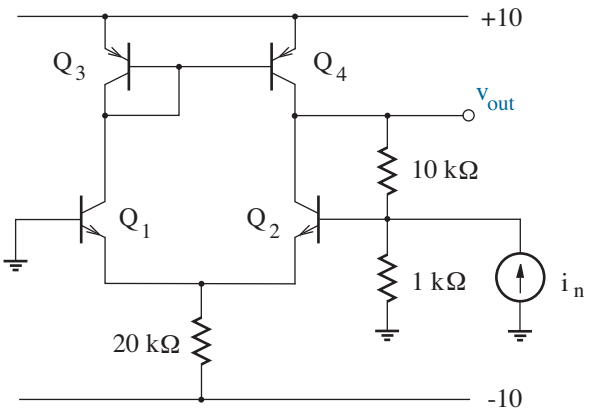


Figure P11.25

11.26 The MOSFETs of Fig. P11.26 have

$K_n' = 50\ \mu\text{A/V}^2$, $V_{Tn} = +0.5$ V, $\lambda_n = 0.05\ \text{V}^{-1}$
$K_p' = 20\ \mu\text{A/V}^2$, $V_{Tp} = -0.5$ V, $\lambda_p = 0.1\ \text{V}^{-1}$

Determine v_{out}/i_n.

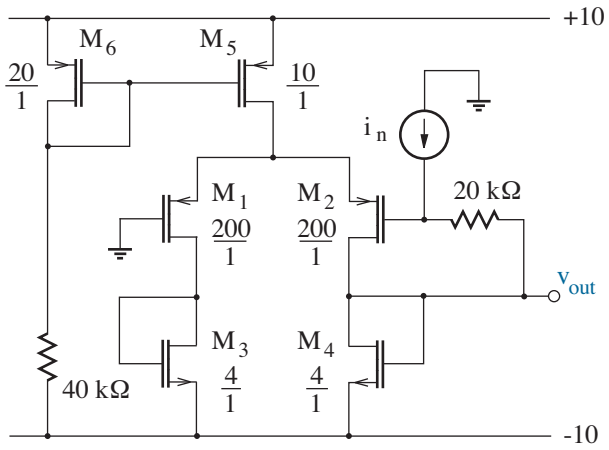


Figure P11.26

Section 11.3

11.27 Prove the expression for output impedance in Eq. 11.52.

11.28 The op-amp in the circuit of Fig. P11.28 has a differential voltage gain of 50,000, an input resistance of 800 kΩ, and an output resistance of 200 Ω. Determine i_{out}/i_n, z_{in}, and z_{out}.

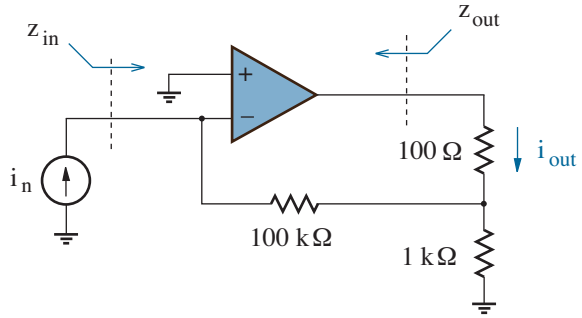


Figure P11.28

11.29 Repeat Problem P11.28, but use a poor op-amp with a differential voltage gain of 5000, an input resistance of 80 kΩ, and 2-kΩ output resistance.

11.30 The BJTs in the circuit of Fig. P11.30 have $\beta_F = \beta_o = 200$ and $I_s = 10^{-14}$ A.

(a) Design for $i_{c1}|_Q = i_{c2}|_Q = 1$ mA subject to $i_n = 0$, $R_f \to \infty$, and $R_1 \parallel R_2 = 100$ kΩ.

(b) Find R_f for $i_{out}/i_n = 5$, and verify with SPICE.

Section 11.4

11.31 Prove the expression for output impedance in Eq. 11.70.

11.32 The op-amp in the circuit of Fig. P11.32 has a differential voltage gain of 50,000, an input resistance of 800 kΩ, and an output resistance of 200 Ω. Determine i_{out}/v_t, z_{in}, and z_{out}.

11.33 Repeat Problem P11.32, but use a poor op-amp with a differential voltage gain of 5000, an input resistance of 80 kΩ, and 2-kΩ output resistance.

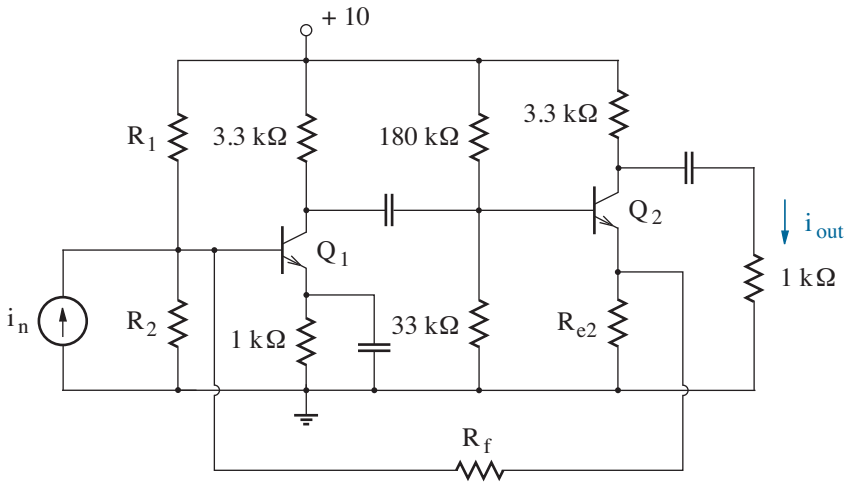


Figure P11.30

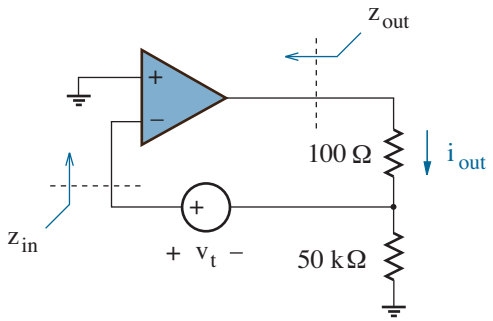


Figure P11.32

11.34 The circuit of Fig. P11.34 is intended to control ac fluctuations in the light output from an LED. When the circuit is properly designed, a 1-V change at the input produces a 1-mA change in LED current. The BJTs feature $\beta_F = \beta_o = 200$.

Note: For simplicity, assume the LED has an "on" voltage of about 1.0 V.

(a) Estimate $i_{c1}|_Q$.

(b) Redraw the circuit as a pair of interconnected two-ports.

(c) Complete the design by specifying R.

(d) Use SPICE to verify your design. Let IS=10f for the BJT and IS=1E-20 for the LED.

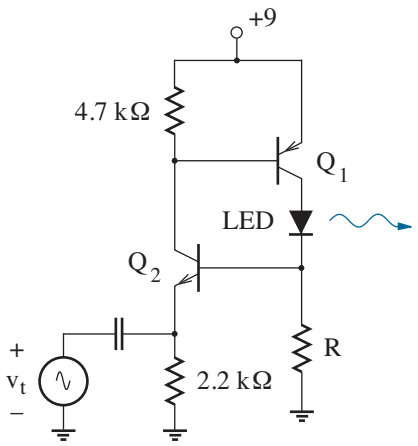
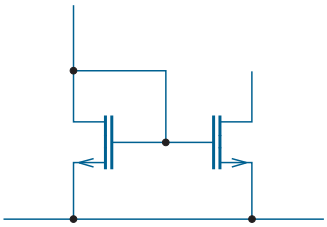


Figure P11.34



Chapter 12

Dynamic Feedback Effects

As demonstrated in the preceding chapter, four different types of feedback can be used to improve amplifier behavior in relation to particular signal sources and loads while decreasing sensitivity to performance factors that tend to vary significantly. Our focus was static—the behavior at dc applies equally well for some arbitrary signal frequency. We now examine dynamic feedback effects that result when system components change their character with increasing frequency. Numerous hazards await.

Objectives

After studying this chapter and working through the problems, you should be able to perform each of the following:

- Describe the evolution of circuit behavior in the time and frequency domain as A_oF increases in one-, two-, or three-pole feedback circuits (Section 12.1).
- Perform magnitude- and phase-related tests to predict circuit stability (Section 12.1).
- Find the effective bandwidth of an op-amp amplifier (Section 12.1).
- Compensate a circuit by adding an appropriately positioned pole to achieve a desired phase margin (Section 12.2).
- Compensate a simple two-stage CMOS op-amp (Section 12.3).
- Design a Colpitts oscillator (Section 12.4).
- Explain the operation of a crystal oscillator (Section 12.4).

12.1 Wandering Poles

Feedback typically yields a circuit transfer characteristic of the form

$$H = \frac{A}{1 + AF}\,, \tag{12.1}$$

for any one of four modes of output/input. In general, the A factor reflects the frequency-dependent behavior of an amplifier two-port with feedback-modified input and output loading, and the F (feedback) factor is constant—at least in the examples presented in Chapter 11. Given A and F, we want to understand the dynamic consequences for H.

One-Pole Behavior

Suppose the A factor features a single pole that induces high-frequency cutoff at angular frequency ω_{p1}. Specifically,

$$A = \frac{A_o}{1 + \dfrac{s}{\omega_{p1}}}\,, \tag{12.2}$$

where A_o applies at the dc limit and $s = \jmath\omega$. In the presence of feedback, we look to Eq. 12.1 to find (after some algebraic rearrangement)

$$H = \frac{A_o}{1 + A_oF}\left[\frac{1}{1 + \dfrac{s}{\omega_{p1}(1 + A_oF)}}\right]\,. \tag{12.3}$$

The effect of feedback is to cause the pole to "wander" to a more negative position on the real axis of the s plane such that the high-frequency cutoff *increases* by a factor of $(1 + A_oF)$. Meanwhile, the penalty is to *decrease* the dc limit by the same factor as shown in the Bode diagram of Fig. 12.1.

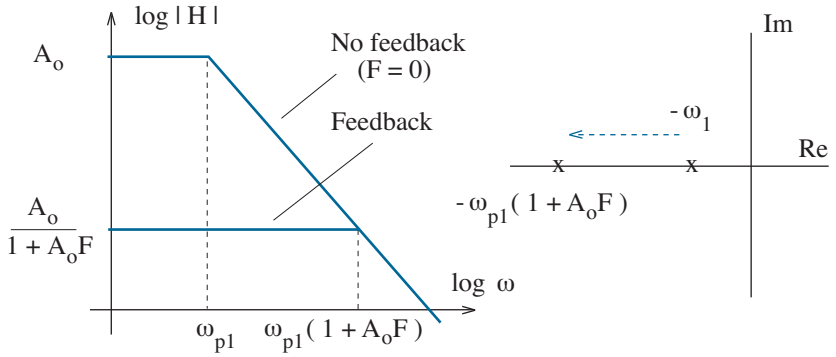


Figure 12.1: Bode diagram applied to a one-pole amplifier with feedback. The adjacent figure shows pole migration in the s plane as A_oF increases.

For reasons to be addressed in Section 12.2, many op-amps behave as if they have a single low-frequency pole that wanders from ω_{p1} as in Fig. 12.1. Applications are subject to a constant **gain bandwidth product**.

Example 12.1

Figure 12.2 features an op-amp with a gain-bandwidth product of 800 kHz. Determine the circuit bandwidth, then check with SPICE.

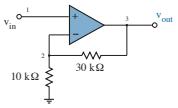


Figure 12.2: Circuit for Example 12.1.

Solution

We construct the so-called "A" circuit that includes feedback loading without actual feedback using the methods of Chapter 11. Given series-shunt feedback with *voltage* sharing on the output side, the loading on the input side is found by disconnecting the 30-kΩ feedback resistor from the output, reconnecting it to a *voltage* source, and setting that voltage source to zero. Similarly, with *current* sharing on the input side, the output-side loading is found by disconnecting the wire leading to the inverting op-amp input, reconnecting it to a *current* source, and setting that current source to zero. Figure 12.3 shows the results of these actions.

It is little work to show that the loaded A circuit yields $v_{out}/v_{in} = A_{vd}$, the same differential gain that applies to the gain-bandwidth specification.

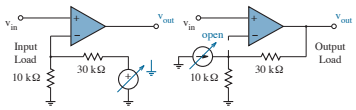


Figure 12.3: "A"-circuit loading for Example 12.1.

The non-inverting feedback amplifier has $v_{out}/v_{in} \approx (1+30/10) = 4$. Thus, the circuit bandwidth is 800 kHz /4 = 200 kHz.

The SPICE code for ac analysis has the following form:

```
* Non-Inverting Amplifier Frequency Response

Vin        1       0       ac      1m
R1         2       0       10k
Rf         3       2       30k
X1         3       0       1       2       OpAmp

.ac        dec     50      100     1MEG
.probe

.subckt    OpAmp           out     com     in_p    in_n
Rin        in_p    in_n    1000MEG
Eout       out     com     LAPLACE
+          { v(in_p , in_n) } { 800k*(2*3.14159 / (s + 2*3.14159)) }
.ends

.end
```

Here, v(in_p , in_n) is the voltage difference between inputs in_p and in_n. The expression in the second set of curly brackets is a system function for A_{vd} with a dc gain of 800,000 and an effective dominant pole at $f_{p1} = 1$ Hz. Note the 2π factor for angular frequency. Figure 12.4 shows .probe results.

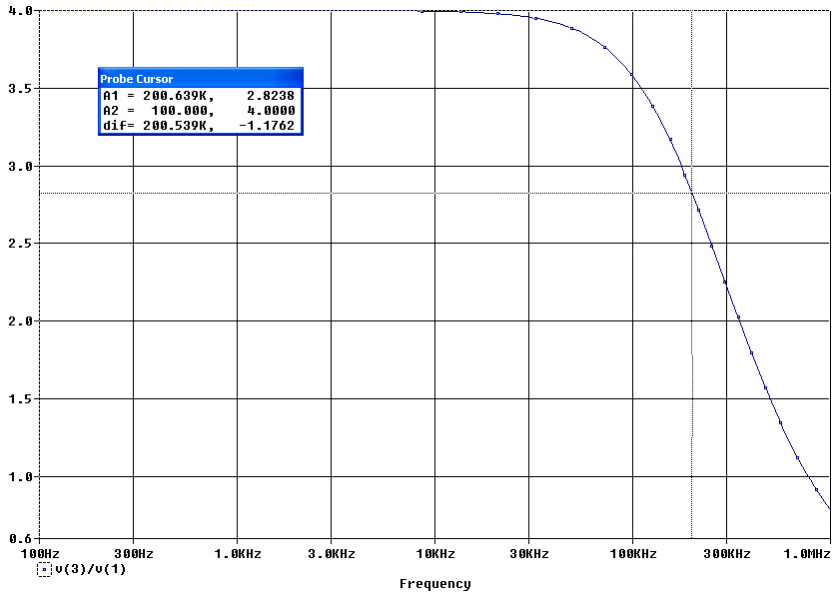


Figure 12.4: .probe SPICE results for the circuit of Example 12.1.

Example 12.2

Figure 12.5 features an op-amp with a gain-bandwidth product of 800 kHz. Determine the circuit bandwidth, then check with SPICE.

Solution

We construct an "A" circuit with feedback loading but no actual feedback. Given shunt-shunt feedback with *voltage* sharing on either side of the 40-kΩ feedback resistor, we separately disconnect each resistor end, reconnect to a *voltage* source, and set that voltage source to zero. Figure 12.6 shows the results of these actions.

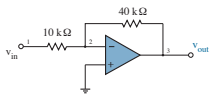


Figure 12.5: Circuit for Example 12.2.

The A circuit of Fig. 12.6 yields

$$\frac{v_{out}}{v_{in}} = -\left(\frac{40\,\text{k}\Omega}{40\,\text{k}\Omega + 10\,\text{k}\Omega}\right) A_{vd} = -0.8\,A_{vd}\,.$$

In turn, the gain-bandwidth product reduces to 0.8×800 kHz $= 640$ kHz. Subject to feedback, the inverting amplifier has $v_{out}/v_{in} \approx -40/10 = -4$. Thus, the circuit bandwidth is 640 kHz $/4 = 160$ kHz.

The SPICE simulation uses the same op-amp model of Example 12.1. For a portion of the netlist, we have,

```
* Inverting Amplifier Frequency Response

Vin     1     0     ac     1m
R1      1     2     10k
Rf      3     2     40k
X1      3     0     0      2      OpAmp
```

Figure 12.7 shows .probe results.

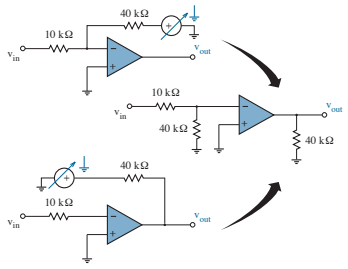


Figure 12.6: "A"-circuit loading for Example 12.2.

We defer the consideration of frequency dependence in more complicated op-amp circuits to the end-of-chapter problems.

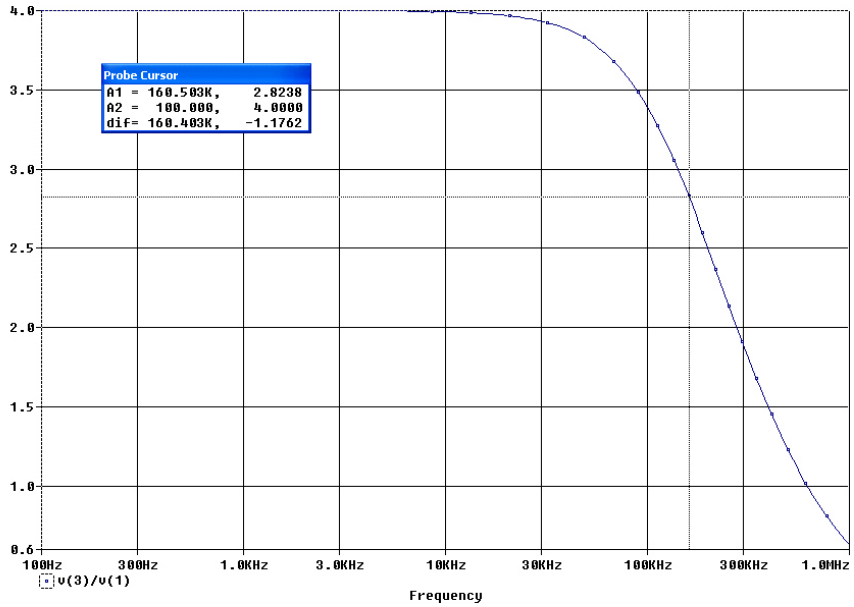


Figure 12.7: .probe SPICE results for the circuit of Example 12.2.

Two-Pole Behavior

Now let the A factor feature two poles at angular frequencies ω_{p1} and ω_{p2}, and let $\omega_{p2} \gg \omega_{p1}$. In this case,

$$A = \frac{A_o}{\left(1 + \dfrac{s}{\omega_{p1}}\right)\left(1 + \dfrac{s}{\omega_{p2}}\right)} . \tag{12.4}$$

Then with feedback,

$$H = \frac{\dfrac{A_o}{1 + A_o F}}{1 + \dfrac{s}{1 + A_o F}\left(\dfrac{1}{\omega_{p1}} + \dfrac{1}{\omega_{p2}}\right) + \dfrac{s^2}{\omega_{p1}\omega_{p2}(1 + A_o F)}} . \tag{12.5}$$

The new pole positions are determined by setting the denominator of this expression to zero and finding the roots of the resulting quadratic equation. So we obtain,

$$s_1 = \frac{-(\omega_{p1} + \omega_{p2})}{2} + \frac{(\omega_{p1} + \omega_{p2})}{2}\sqrt{1 - \frac{4\omega_{p1}\omega_{p2}(1 + A_o F)}{(\omega_{p1} + \omega_{p2})^2}} \tag{12.6}$$

and

$$s_2 = \frac{-(\omega_{p1} + \omega_{p2})}{2} - \frac{(\omega_{p1} + \omega_{p2})}{2}\sqrt{1 - \frac{4\omega_{p1}\omega_{p2}(1 + A_oF)}{(\omega_{p1} + \omega_{p2})^2}}\,. \qquad (12.7)$$

Equations 12.6 and 12.7 are beginning to show signs of algebraic haze. Nevertheless, if $1 + A_oF$ is sufficiently small,

$$\sqrt{1 - \frac{4\omega_{p1}\omega_{p2}(1 + A_oF)}{(\omega_{p1} + \omega_{p2})^2}} \;\approx\; 1 - \frac{2\omega_{p1}\omega_{p2}(1 + A_oF)}{(\omega_{p1} + \omega_{p2})^2}\,. \qquad (12.8)$$

In turn, with $\omega_{p2} \gg \omega_{p1}$, $s_1 \approx -\omega_{p1} - A_oF\omega_{p1}$ and $s_2 \approx -\omega_{p2} + A_oF\omega_{p1}$. Thus, the high-frequency cutoff begins to increase by the same $(1 + A_oF)$ factor encountered for the one-pole amplifier. Meanwhile, the second pole becomes less negative, but it is still far removed. As A_oF becomes larger, s_1 and s_2 continue to wander to $-(\omega_1 + \omega_2)/2$, but from opposite directions. Eventually, the two poles overlap and then split away from the negative real axis as a complex conjugate pair with a constant negative real component. The condition for complex poles is that the argument of the square-root term in Eqs. 12.6 and 12.7 is negative. Specifically,

$$1 + A_oF \;>\; \frac{(\omega_{p2} + \omega_{p1})^2}{4\,\omega_{p1}\omega_{p2}} \approx \frac{\omega_{p2}}{4\,\omega_{p1}}\,. \qquad (12.9)$$

Figure 12.8 shows the pole trajectories.

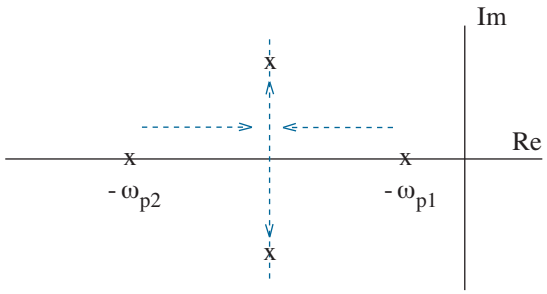


Figure 12.8: Pole trajectories for a two-pole amplifier with feedback.

Imaginary pole components promote an oscillatory transient response and sometimes peaking of the steady-state amplitude just prior to cutoff (see Problem 12.10). Fortunately, the real pole components are negative in a two-pole feedback amplifier, so natural oscillations eventually decay.

We pause to observe that the pole trajectories of Fig. 12.8 obey two general rules:

Rule 1: Pole trajectories along the real axis proceed to the left of an odd number of stationary AF poles when no AF zeros are present.

Proof: At new poles that derive from increasing A_oF, $1+AF = 0$, $|AF| = 1$, and $\angle AF = \pi$ (less some multiple of 2π). The angle function is calculated by drawing vectors from each stationary AF pole to one of the wandering feedback poles and subtracting the individual angles subtended with respect to the positive real axis. This process yields 0 to the right of the negative stationary poles and $-\pi$ to the left of the least negative stationary pole. Thus, the wandering pole trajectory proceeds to the left in relation to p_1. The individual trajectories alter rightward and leftward thereafter.

Rule 2: The pole "center of gravity" is constant as A_oF increases.

Proof: The denominator of AF is $(s+p_1)(s+p_2)(s+p_3)\ldots(s+p_n)$, and the highest-order terms in the expansion are s^n and $(p_1+p_2+p_3+\ldots)s^{n-1}$. So the sum of the stationary poles is the ratio of the polynomial coefficients of order $n-1$ and n . The wandering feedback poles $(p_1{}', p_2{}', p_3{}', \ldots)$ are the roots of the expression $(s+p_1)(s+p_2)(s+p_3)\ldots(s+p_n)+A_oF = 0$. Since the constant A_oF factor does not affect the two highest-order terms in the expansion, the sum of the poles is unchanged. The same applies to the pole average or center of gravity.

Be comforted. Our objective is not to provide a catalog of pole behavior in the presence of feedback. But armed with the preceding trajectory rules, we are ready to make an important observation.

Three-Pole Behavior

In this next level of complexity, we let A feature three poles at angular frequencies ω_{p1}, ω_{p2}, and ω_{p3} (subject to $\omega_{p3} \gg \omega_{p2} \gg \omega_{p1}$). Specifically,

$$A = \frac{A_o}{\left(1 + \dfrac{s}{\omega_{p1}}\right)\left(1 + \dfrac{s}{\omega_{p2}}\right)\left(1 + \dfrac{s}{\omega_{p3}}\right)} . \tag{12.10}$$

Instead of writing out an expression for $H = 1/(1 + AF)$ and attempting to evaluate the roots of the denominator—horrific algebra for a cubic equation—we examine the pole trajectories that satisfy the real-axis and center-of-gravity rules. The former rule requires that the two poles commencing at $-\omega_{p1}$ and $-\omega_{p2}$ move to the left and right, respectively, as shown in Fig. 12.9. These poles eventually overlap and split away from the real axis as a complex conjugate pair, Meanwhile, the pole commencing at $-\omega_{p3}$ moves steadily to the left along the real axis. In turn, the center-of-gravity rule requires that the common real component of the complex conjugate pair become less negative. Indeed, the two poles are made to wander off into the right half of the complex plane if $A_o F$ is sufficiently large. The result is an *unstable* system: transient oscillations grow rather than decay, a condition that is best avoided (except for intentional oscillator circuits).

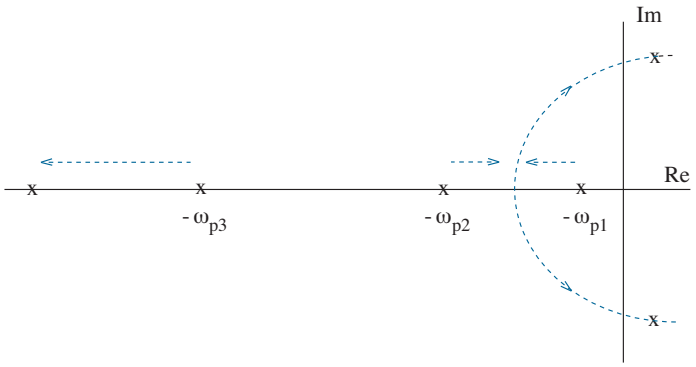


Figure 12.9: Pole trajectories for a two-pole amplifier with feedback.

In cases of three or more poles, control engineers apply **root locus** rules to determine pole trajectories in order to elicit a specific system response. We are mostly content to recognize and avoid unstable conditions.

Tests for Stability

We look for some simple tests that recognize unstable feedback conditions (in anticipation of prompt correction procedures for better behavior).

> Caution: It is assumed that readers are familiar with the construction of magnitude and phase-angle Bode plots for arbitrary system functions. Those in need of remedial help may wish to consult a textbook on elementary circuit analysis.

Consider an amplifier exhibiting flat response over a broad midfrequency range that extends toward dc —we ignore any pole-zero pairs that impose low-frequency cutoff with little impact at moderate frequencies and beyond. The amplifier gain tends to roll off at high frequencies, so we expect an AF product of the form

$$AF = \frac{A_o F}{\left(1 + \frac{\jmath\omega}{\omega_{p1}}\right)\left(1 + \frac{\jmath\omega}{\omega_{p2}}\right)\left(1 + \frac{\jmath\omega}{\omega_{p3}}\right)\ldots}\,, \tag{12.11}$$

where $\omega_{p1}, \omega_{p2}, \omega_{p3}\ldots$ are pole angular frequencies, and A_oF is constant. The rolloff is continuous (monotonic) as ω increases, and

$$|AF| \to 0 \qquad \text{as } \omega \to \infty\,. \tag{12.12}$$

Meanwhile, for the phase angle,

$$\angle AF \to \frac{-n\pi}{2} \qquad \text{as } \omega \to \infty\,, \tag{12.13}$$

where n is the number of poles. It follows that for $n \geq 3$, there exists some angular frequency—call it ω_{180} —such that $\angle AF = -\pi = -180°$. In turn,

$$AF = |AF|e^{-\jmath\pi} = -|AF| \tag{12.14}$$

and

$$H = \frac{A}{1 - |AF|}\,. \tag{12.15}$$

If $|AF| = 1$ at ω_{180}, H is clearly undefined. Moreover,

H promotes inherent system instability if $|AF| > 1$ at ω_{180}

(despite benign mathematical appearance). In light of this test condition, it may be tempting just to promise never to operate a potentially unstable feedback circuit using high-frequency signals. Yet, promises aside, there is no safeguard against start-up transients or high-frequency electrical noise.

Perhaps we can better understand the rationale for the ω_{180} test with the help of a crude argument applied to the feedback block diagram of Fig. 12.10. Here, the intermediate signal z is $x - Fy$. Then with $y = Az$, z conveniently drops out to yield $H = y/x = A/(1 + AF)$.

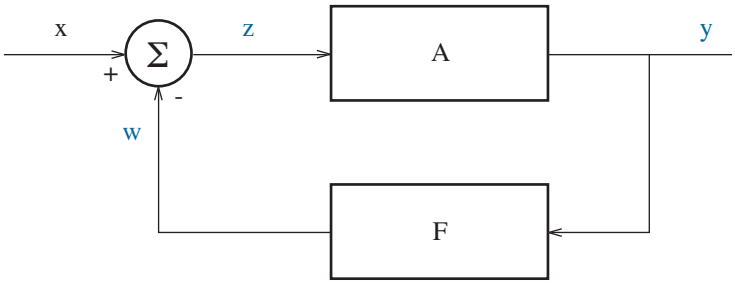


Figure 12.10: Feedback block diagram for $H = y/x = A/(1 + AF)$.

Now suppose $x(t)$ represents a steady supply of sinusoidal "packets", each with the time dependence $\sin \omega_{180} t$ over a small range of t, and let the feedback system be initially void of any packets. Shortly after the arrival of the first packet, the A-circuit input is $z(t) = \sin \omega_{180} t$. Following a delay, and given the definition for ω_{180}, the signal at the negative input to the summing node becomes $w(t) = |AF| \sin(\omega_{180} - 180°)t = -|AF| \sin \omega_{180} t$. So soon thereafter,

$$z(t) = \underbrace{\sin \omega_{180} t}_{\text{New packet 2}} + \underbrace{|AF| \sin \omega_{180} t}_{\text{Feedback packet 1}} . \tag{12.16}$$

After two delay periods,

$$z(t) = \underbrace{\sin \omega_{180} t}_{\text{New packet 3}} + \underbrace{|AF| \sin \omega_{180} t}_{\text{Feedback packet 2}} + \underbrace{|AF|^2 \sin \omega_{180} t}_{\text{Feedback packet 1}} . \tag{12.17}$$

And eventually,

$$z(t) = \sin \omega_{180} t \, (1 + |AF| + |AF|^2 + |AF|^3 + \dots \,) . \tag{12.18}$$

If $|AF| > 1$, the series diverges in agreement with the test for instability. However, if $|AF| < 1$, the series converges. Whereas $AF = -|AF|$ at ω_{180}, the series is geometric such that

$$y(t) = \frac{A}{1 + AF} \, x(t) . \tag{12.19}$$

This is the original feedback relationship.

Example 12.3

A feedback system features an "A" circuit with $A = 1000$ in the dc limit and poles at 100 kHz, 500 kHz, 2 MHz, and 5 MHz. Use the "ω_{180}" test to determine whether the system is stable for the feedback factor $F = 0.1$.

Solution

The loop gain is

$$AF = \frac{1000 \times 0.1}{\left(1 + \frac{\jmath f}{100\ \text{kHz}}\right)\left(1 + \frac{\jmath f}{500\ \text{kHz}}\right)\left(1 + \frac{\jmath f}{2\ \text{MHz}}\right)\left(1 + \frac{\jmath f}{5\ \text{MHz}}\right)},$$

where the right side f is frequency—not to be confused with the left-side feedback factor. The corresponding phase angle is given by

$$\angle AF = 0 - \tan^{-1}\frac{f}{100\ \text{kHz}} - \tan^{-1}\frac{f}{500\ \text{kHz}} - \tan^{-1}\frac{f}{2\ \text{MHz}} - \tan^{-1}\frac{f}{5\ \text{MHz}}.$$

After trial and error (and the help of a programmable calculator), we find $\angle AF = -180°$ at $f = 914$ kHz. Then we evaluate

$$|AF| = \frac{100}{\sqrt{1 + \left(\frac{f}{100\ \text{kHz}}\right)^2}\sqrt{1 + \left(\frac{f}{500\ \text{kHz}}\right)^2}\sqrt{1 + \left(\frac{f}{2\ \text{MHz}}\right)^2}\sqrt{1 + \left(\frac{f}{5\ \text{MHz}}\right)^2}}$$

at the same frequency to find $|AF| = 4.67$. The result is greater than unity, so the system is unstable with a **gain margin** of $20 \log 4.67 = 12.4$ dB. Stable systems exhibit gain margins in negative dB.

Another test for stability is to determine the angular frequency ω_T for which $|AF| = 1$. One then defines a **phase margin**

$$\phi = \angle AF|_{\omega_T} + 180° \tag{12.20}$$

as a conditioning factor for the **Nyquist test:**

H promotes inherent system instability if $\phi < 0$.

We offer no proof—the rationale should be evident from the ω_{180} test.

In many cases, the Nyquist test is more easily applied than the ω_{180} test. Recall that a magnitude Bode plot relates the logarithm of some quantity to the logarithm of frequency. At ω_T, $\log AF = 0$, and $\log A = \log(1/F)$. So ω_T is simply obtained from the intersection of superimposed magnitude Bode plots for A and $1/F$. The phase angle satisfies $\angle AF = \angle A - \angle(1/F)$, and the phase margin at ω_T is ready for evaluation.

Example 12.4

A feedback system features an "A" circuit with $A_o = 1000$ in the dc limit and poles at 100 kHz, 500 kHz, 2 MHz, and 5 MHz. Use the Nyquist test to determine whether the system is stable for the feedback factor $F = 0.5$.

Solution

We have a yes or no decision, which we hope to resolve with the aid of very rough Bode plots of magnitude and phase. The plots can always be refined if the decision is a close call.

Figure 12.11 shows straight-line magnitude asymptotes for A and $1/F$. For A, we start with a flat response with magnitude $1000 = 60$ dB. Then we reach a first breakpoint at the 100-kHz pole where the response begins to fall off at a rate of -20 dB (one order of magnitude) per decade. Next we reach a second breakpoint at the 500-kHz pole, and the response begins to fall off at a rate of -40 dB (two orders of magnitude) per decade. And so on ... With $1/F = 2 = 6$ dB, the A plot intersects at $f_T = \omega_T/2\pi \approx 3.3$ MHz.

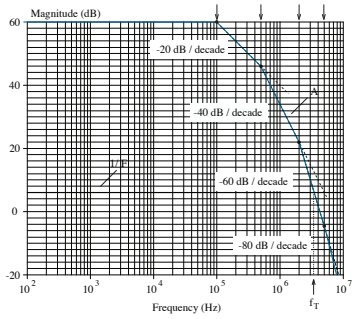


Figure 12.11: Magnitude Bode plot for Example 12.4.

We now construct the rough phase-angle Bode plots of Fig. 12.12 subject to the observation that the poles for for A are widely separated in frequency. Thus, with each pole contributing a -90° phase shift, the phase angles at the

pole frequencies are successively -45°, -135°, -225°, and -315°. A free-hand sketch starting from zero and linking these points ought to be sufficient. The F factor is constant, so it has zero contribution to $\angle AF$.

At 3.3 MHz, $\angle AF$ is about $-267°$. In turn, the phase margin is $-87°$, and the system is unstable. Had we used SPICE to prepare the Bode plots (see Problem 12.17), we would obtain $\phi = -81°$ and the same conclusion.

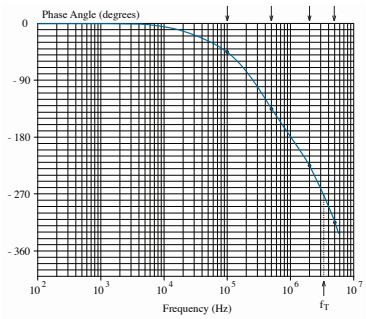


Figure 12.12: Phase-angle Bode plot for Example 12.4.

Exercise 12.1 A feedback system features $A_o = 100$ in the dc limit and poles at 1 MHz, 20 MHz, and 100 MHz. Determine the phase margin and stability condition for $F = 1$.

Ans: $\phi = 16°$ Stable

Example 12.5

Use SPICE to determine the phase margin for the amplifier of Fig. 12.13. The transresistance is 200 kΩ (see Example 12.4).

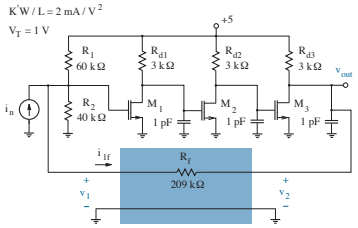


Figure 12.13: Circuit for Example 12.5.

Solution

We are interested in the magnitude and phase for the AF cascade without actual feedback. Thus, we modify the circuit for simulation by breaking the feedback loop on the input side of the F two-port and setting v_1 to zero, as it is no longer shared. The easiest way to do this is a simple grounding. Nevertheless, in anticipation of measuring i_{1f}, we add a voltage source V_m with zero value and a polarity that is consistent with i_{1f} flowing from plus to minus so that the .probe expression i(Vm) reads an ammeter in SPICE.

Setting $v_1 = 0$ does not change the loading effect that the F two port presents to the output side of the A two-port. But the break in the feedback loop destroys the input-side loading, and it must be restored if we are to obtain an accurate simulation for A. Once again, we are tempted to make an easy fix by connecting a 209-kΩ resistor in parallel with R_2. Yet doing so disrupts the A-circuit biasing, which was formerly stabilized with feedback. The solution is to trick the circuit by inserting a huge (say 1-F) capacitor in series with R_f and presenting the combination in parallel with R_2.

Figure 12.14 shows the circuit modifications that have been described.

We apply a standard SPICE code and ac simulation to find $v_{out}/i_n = A$, $i_{1f}/v_{out} = F$, or $i_{1f}/i_n = AF$.

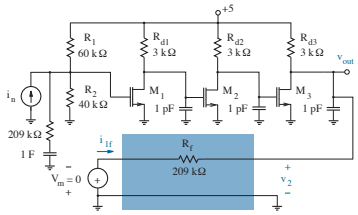


Figure 12.14: Modifications for determining A and F with proper loading.

Figure 12.15 shows .probe simulation results with the AF magnitude and phase plots in the same window to facilitate the cursor application. The AF magnitude is unity at about 138 MHz, and the corresponding AF phase angle is $-207°$. Thus, the phase margin is $-207° + 180° = -27°$. The transresistance amplifier is unstable.

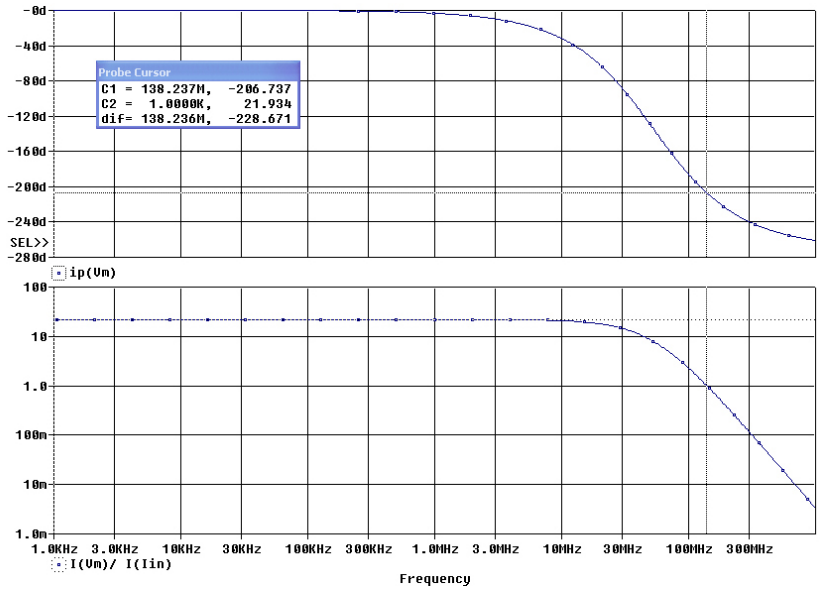


Figure 12.15: .probe AF magnitude and phase results for Example 12.5.

12.2 Additive-Pole Compensation

The previous section showed that feedback circuits have the potential for unstable operation, especially for three or more poles and substantial A_oF. When ω_{180} or Nyquist tests show signs of trouble, one seeks improved stable performance through a form of system **compensation** such as

- Modifying A (by making it frequency dependent or changing A_o),
- Modifying the effective Z_t or Z_n for the input signal source,
- Modifying F (by making it frequency dependent).

This section examines a restricted but nonetheless powerful compensation method in which a feedback circuit engages a single additive pole to assume stable dynamic behavior while maintaining its low-frequency performance. Section 12.3 examines a compensation method involving a pole and zero.

> Warning: Compensation methods are notoriously intimidating. Forge ahead. Don't panic.

In the additive-pole compensation process, the new pole frequency ω_{p0} is typically much less than ω_{p1}, ω_{p2}, ... So we make two observations:

Observation 1: The asymptotic $|AF|$ experiences a supplemental decline of -20 dB/decade for $\omega > \omega_{p0}$. It is otherwise unchanged.

Observation 2: The approximate $\angle AF$ is 90° more negative for $\omega \gg \omega_{p0}$, 45° more negative at $\omega = \omega_{p0}$, and unchanged for $\omega \ll \omega_{p0}$.

The magnitude and phase plots in Fig. 12.16 support these observations, which are convenient as guides for achieving a phase margin of 45°.

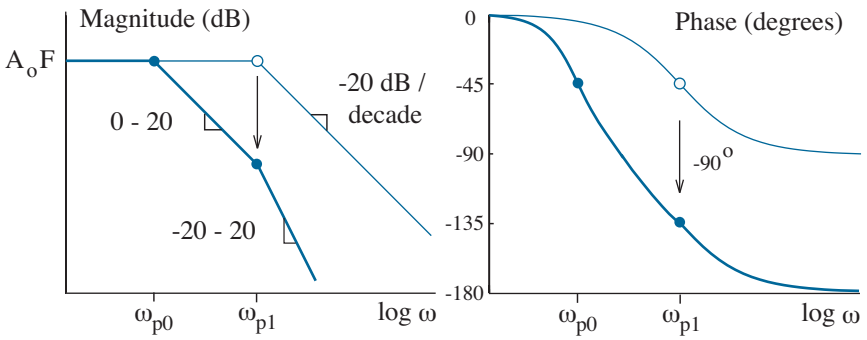


Figure 12.16: Influence of a new pole ω_{p0} on $|AF|$ and $\angle AF$ for $\omega_{p0} \ll \omega_{p1}$.

Example 12.6

A feedback system features an "A" circuit with $A_o = 1000$ in the dc limit and poles at 100 kHz, 500 kHz, 2 MHz, and 5 MHz. The F factor is 0.5. Add a pole to compensate the system so that the phase margin is 45°.

Solution

We require a magnitude Bode plot consistent with $\phi = 45°$. For reference, we draw horizontal lines that reflect $A_o = 60$ dB and $1/F = 6$ dB. Then we let $f_{p1} = 100$ kHz assume a distinction as the point where the A and $1/F$ curves intersect. Subject to $f_{p0} \ll f_{p1}$ for the new pole, this f_{p1} designation establishes $\angle AF = -45° - 90° = -135°$ where $|AF| = 1$ so that $\phi = 45°$. The f_{p1} pole marks the end of a -20 dB/decade decline from a breakpoint, so we extend a same-slope line upwards from f_{p1} as an asymptote for A. In turn, $f_{p0} = 200$ Hz at the intersection with $A_o = 60$ dB.

Figure 12.17 shows the details for this graphical solution. In general, adding a pole to make $\phi = 45°$ requires

$$f_{p0} = \frac{f_{p1}}{A_o F} \tag{12.21}$$

(see Problem 12.29). The special-case additive-pole compensation is trivial.

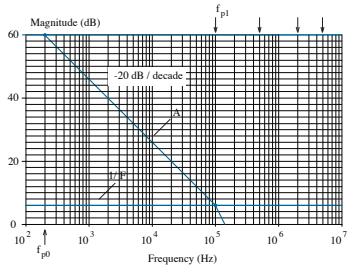


Figure 12.17: Graphical solution for Example 12.6.

Graphical procedures are less helpful for phase margins other than 45°. Some analytic treatment is needed to guide the compensation process.

Consider a feedback system containing three or more poles and no zeros. The addition of a pole at ω_{p0} presumably yields a phase margin somewhere between zero and 90°. Thus, ω_{p0} and ω_{p1} guide the behavior of an effective two-pole system—the other poles are pushed aside, as they are only evident at inconsequential frequencies where $|AF| \ll 1$. Then we have

$$AF \approx \frac{A_o F}{\left(1 + \frac{\jmath\omega}{\omega_{p0}}\right)\left(1 + \frac{\jmath\omega}{\omega_{p1}}\right)} . \tag{12.22}$$

And with typical $\omega_{p0} \ll \omega_{p1}$,

$$AF \approx \frac{A_o F}{\frac{\jmath\omega}{\omega_{p0}}\left(1 + \frac{\jmath\omega}{\omega_{p1}}\right)} . \tag{12.23}$$

The phase angle for the latter expression is given by

$$\angle AF = -90° - \tan^{-1}\frac{\omega}{\omega_{p1}} . \tag{12.24}$$

So at ω_T, we have the phase margin

$$\phi = 180° - 90° - \tan^{-1}\frac{\omega_T}{\omega_{p1}} . \tag{12.25}$$

In turn,

$$\omega_T = \omega_{p1} \tan\left(90° - \phi\right) . \tag{12.26}$$

For $\phi = 45°$, Eq. 12.26 reduces to $\omega_T = \omega_{p1}$, the special-case compensation condition used in Example 12.6. For $\phi > 45°$, we require $\omega_T < \omega_{p1}$, and for $\phi < 45°$, we require $\omega_T > \omega_{p1}$. Meanwhile, by definition,

$$|AF|_{\omega_T} = \frac{A_o F}{\frac{\omega_T}{\omega_{p0}}\sqrt{1 + \left(\frac{\omega_T}{\omega_{p1}}\right)^2}} = 1 . \tag{12.27}$$

Thus, and in consideration of Eq. 12.26,

$$\omega_{p0} = \frac{\omega_T}{A_o F \cos\left(90° - \phi\right)} . \tag{12.28}$$

Equations 12.26 and 12.28 guide the compensation process. For a desired ϕ, we use the former to find ω_T and then the latter to find ω_{p0} for the new pole. For $\phi = 45°$, the special-case compensation requires

$$f_{p0} = \frac{\sqrt{2}\, f_{p1}}{A_o F} . \tag{12.29}$$

With reference to Eq. 12.21, this indicates a -3-dB adjustment for $A_o F$.

Choosing the Phase Margin

Once a particular phase margin has been established through additive-pole compensation, the transfer characteristic for the effective two-pole system under feedback takes the form

$$H = \frac{A_o}{1 + A_o F} \frac{{\omega_o}^2}{s^2 + s\omega_o/Q + {\omega_o}^2} , \tag{12.30}$$

where

$$\omega_o = \sqrt{\omega_{p0}\, \omega_{p1}(1 + A_o F)} \tag{12.31}$$

and

$$Q \approx \frac{\sqrt{\omega_{p0}\, \omega_{p1} A_o F}}{\omega_{p0} + \omega_{p1}} . \tag{12.32}$$

We are interested in the relationship between the phase margin ϕ and the quality factor Q. Thus, we engage Eq. 12.32 to express $A_o F$ in terms of Q, we substitute the result into Eq. 12.28, and we solve for Q. This yields

$$Q^2 = \frac{1}{(1 + \omega_{p0}/\omega_{p1})^2} \left(\frac{\omega_T}{\omega_{p1}}\right) \frac{1}{\cos(90° - \phi)} . \tag{12.33}$$

Typically, $\omega_{p0} \ll \omega_{p1}$. Then again in consideration of Eq. 12.26,

$$Q \approx \sqrt{\frac{\sin(90° - \phi)}{1 - \sin^2(90° - \phi)}} . \tag{12.34}$$

Equation 12.34 is accurate to within 3.5% for $\phi \geq 30°$ and $|A_o F| \geq 100$. The accuracy greatly improves with larger ϕ and varies inversely with $A_o F$. Figure 12.18 shows a plot for the relation.

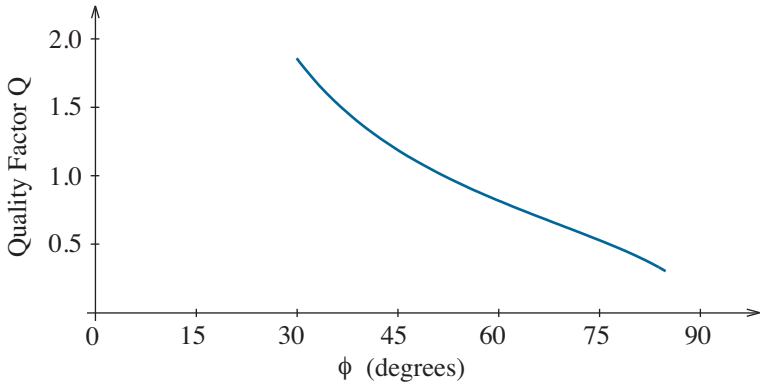


Figure 12.18: Quality factor Q as a function of phase margin ϕ.

Figures 12.19 and 12.20 show frequency and step response, respectively, for $H = A/(1+AF)$ subject to $\phi = 45°$ and $\phi = 66°$. Alas, the trivial-case 45° phase margin has peaking and ringing. With $\phi = 66°$ ($Q = 1/\sqrt{2}$), there is significant improvement at the expense of bandwidth and delay. Phase margins of about 70° give the best frequency and step responses.

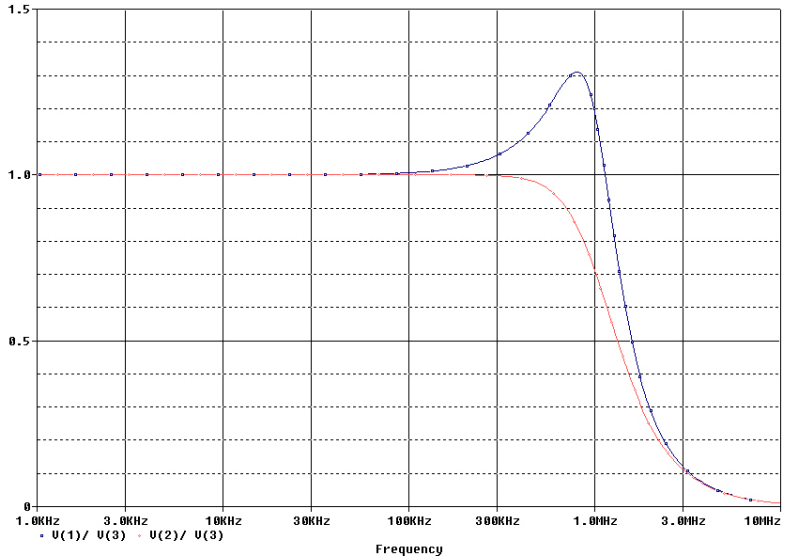


Figure 12.19: Typical feedback frequency response: $\phi = 45°$ and $\phi = 66°$.

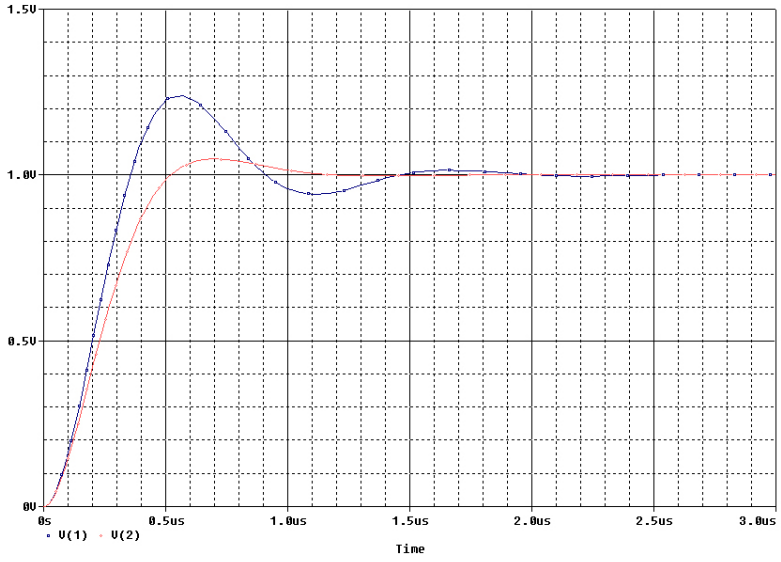


Figure 12.20: Typical feedback step response: $\phi = 45°$ and $\phi = 66°$.

Example 12.7

A feedback system features an "A" circuit with $A_o = 1000$ in the dc limit and poles at 100 kHz, 500 kHz, 2 MHz, and 5 MHz. The F factor is 0.5. Add a pole to compensate the system so that the phase margin is 70°.

Solution

In what follows, it is straightforward to work with frequency f as opposed to angular frequency ω since the 2π conversion factors cancel out in the design equations for additive-pole compensation.

With $f_{p1} = 100$ kHz, we use Eq. 12.26 to find

$$f_T = 100 \text{ kHz } \tan(90^\circ - 70^\circ) = 36.4 \text{ kHz}\,.$$

Then with $A_oF = 500$, we use Eq. 12.28 to obtain

$$f_{p0} = \frac{36.4 \text{ kHz}}{500 \cos(90^\circ - 70^\circ)} = 77 \text{ Hz}\,.$$

Example 12.8

A feedback system features an "A" circuit with $A_o = 1000$ in the dc limit and poles at 100 kHz, 500 kHz, 2 MHz, and 5 MHz. The F factor is 0.5. Add a pole to compensate the system so that its response reflects $Q = 1$.

Solution

Since A_oF is large, we set $Q = 1$ in Eq. 12.34 and solve a quadratic equation to find $\sin(90^\circ - \phi) = (\sqrt{5} - 1)/2$ (a famous number) and $90^\circ - \phi = 38.2^\circ$. Then Eq. 12.26 yields $f_T = 78.7$ kHz and Eq. 12.28 yields $f_{p0} = 200$ Hz.

Exercise 12.2 A feedback system features $A_o = 12,000$ in the dc limit, poles at 5 MHz, 40 MHz, and 100 MHz, and $F = 0.2$. Determine the new pole frequency that compensates the system with $\phi = 60^\circ$.

Ans: $f_{p0} = 1.4$ kHz

Exercise 12.3 The feedback system of Exercise 12.2 is compensated with a new pole at $f_{p0} = 4$ kHz. Estimate f_T, ϕ, and Q (without Bode plots).

Ans: $f_T = 6.1$ MHz $\quad \phi = 39^\circ \quad Q = 1.9$

Example 12.9

Use SPICE to compensate the circuit of Example 12.5 so that $\phi = 70°$.

Solution

We perform an open-loop SPICE simulation as in Example 12.5 to evaluate $\angle AF$ vs. frequency, and we *estimate* $f_{p1} = 14.3$ MHz where $\angle AF = -45°$. Then with $A_oF = 21.9$, we apply Eqs. 12.26 and 12.28 to find

$$f_T = 14.3 \text{ MHz } \tan(90° - 70°) = 5.21 \text{ MHz}$$

and

$$f_{p0} = \frac{5210 \text{ kHz}}{21.9 \cos(90° - 20°)} = 250 \text{ kHz}\,.$$

To realize the new pole frequency, we consider three options:

Option 1: Provide C_c to ground at the input.

(This option modifies both the "A" circuit and Z_n.)

The small-signal resistance at the input node is $R_1 \parallel R_2 \parallel R_f = 21.5$ kΩ. In turn, $C_c = 1/(2\pi \times 250 \times 10^3 \times 21.5 \times 10^3) = 30$ pF. The simulated phase margin that results with this added capacitor is $\phi = 76°$, which is too large. So we apply a trial-and-error process to lower it to 70° with $C_c = 22$ pF. When we simulate the closed-loop feedback circuit of Fig. 12.13 with the added capacitor, the circuit bandwidth (f_h) is 12.8 MHz.

Note: The open-circuit time constants for the circuit of Fig. 12.13 are

$$\tau_1 = R_{d1} \times 1 \text{ pF} = 3 \text{ ns} \qquad \tau_2 = R_{d2} \times 1 \text{ pF} = 3 \text{ ns}$$

$$\tau_3 = R_{d3} \parallel R_f \times 1 \text{ pF} = 2.96 \text{ ns}$$

at the drain nodes for M_1, M_2, and M_3, respectively. Thus, f_{p1}, f_{p2}, and f_{p3} overlap, and improved design equations for f_T and f_{p0} can be derived (see Problem 12.37). The 30-pF C_c value is a "ballpark" estimate at best.

Option 2: Provide C_c to ground at the output).

(This option modifies A by *moving* one of the original poles.)

The small-signal resistance looking into the output is $R_{d3} \parallel R_f = 2.96$ kΩ. In turn, $C_c = 1/(2\pi \times 250 \times 10^3 \times 2.96 \times 10^3) = 215$ pF. The simulated phase margin is now 81°, which is much greater than the desired value. However, SPICE simulations are cheap, so we hack away to determine $C_c = 105$ pF. The circuit bandwidth for this compensation method is 18.3 MHz (better). Design equations that shrink the ballpark are available (see Problem 12.38).

Option 3: Add a pole to the "F" circuit as shown in Fig. 12.21.

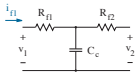


Figure 12.21: Modified "F" circuit for Example 12.9.

In consideration of Fig. 12.21 and the rules for finding feedback factor F, we apply a voltage source for v_2, we let $v_1 = 0$, and we evaluate $F = i_{1f}/v_2$. In turn, we have

$$F = \frac{-1}{R_{f1} + R_{f2}} \left\{ \frac{[(R_{f1} \parallel R_{f2})C_c]^{-1}}{s + [(R_{f1} \parallel R_{f2})C_c]^{-1}} \right\} .$$

We arbitrarily choose $R_{f1} = R_{f2} = R_f/2$ so that $C_c = 1/(2\pi \times 250 \times 10^3 \times 52.25 \times 10^3) = 12$ pF. The simulated phase margin is 75°. So with further trial-and-error readjustment, we eventually find $C_c = 9.1$ pF for $\phi = 70°$. When we look for the bandwidth, we obtain the dramatic result in Fig. 12.22. Adding frequency dependence to F has major impact, since the feedback response varies like $1/F$ when A is large. Option 3 is questionable.

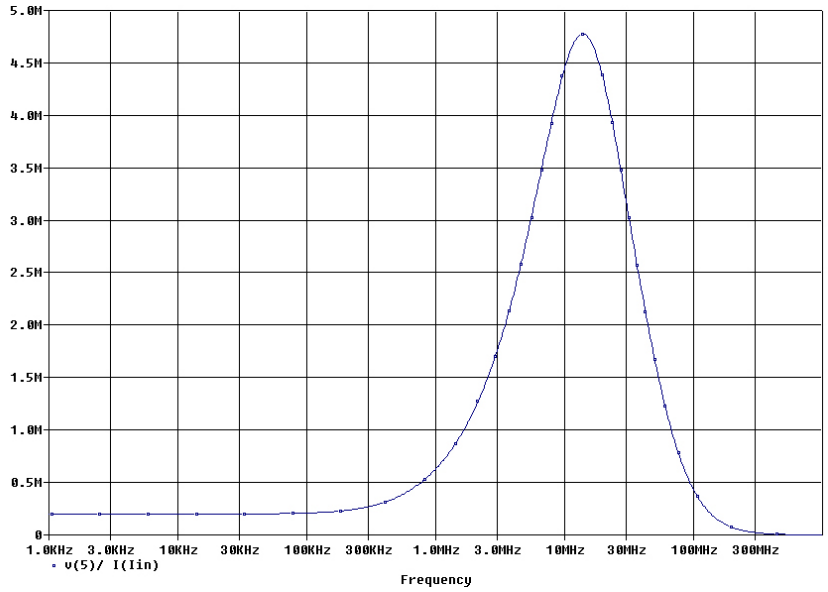


Figure 12.22: Feedback circuit behavior with F-circuit compensation.

12.3 Internal Op-Amp Compensation

We begin this section by considering the stability of the CMOS differential amplifier of Example 10.4 (Fig. 12.23). The W/L ratios that produce $A_{dd} = 20,000$, $|i_{d5}|_Q| = 200$ μA, and $v_{out}|_Q \approx 0$ are beside the various MOSFETs, and the device parameters are those of Example 10.2. For argument's sake, we assume that each source and drain contributes 1 fF for each unit of width to an effective capacitance to ground—the actual contributions depend on the layout details for the integrated circuit. In turn, $C_7 = 320$ fF (node 7), $C_8 = 320$ fF (node 8), and $C_9 = 161$ fF (node 9). The capacitance values at the other nodes are less important because the nodes are either at ac ground or connected to voltage sources with small Thevenin resistances.

SPICE analysis reveals unity gain and -198° phase shift at 835 MHz. Thus, the phase margin is $\phi = -18°$, and the amplifier is unstable ($F = 1$). The SPICE results also indicate that $f_{p1} \approx 2.66$ MHz where $\angle v_o = -45°$. So for additive-pole compensation that achieves $\phi = 70°$, we require

$$f_T = 2.66 \text{ MHz } \tan(90° - 70°) = 968 \text{ kHz}$$

and

$$f_{p0} = \frac{968 \text{ kHz}}{20,000 \cos(90° - 20°)} = 51.5 \text{ Hz}\,.$$

If we choose to effect a new dominant pole by adding a single compensation capacitor C_c to ground at node 8, we find

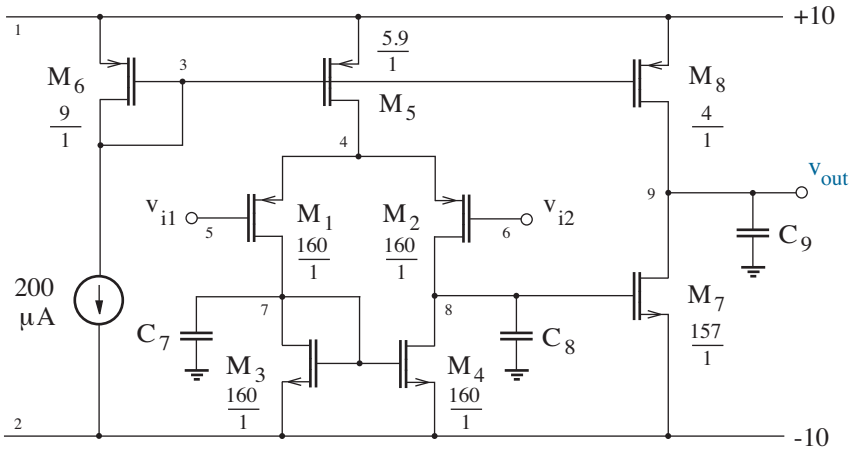


Figure 12.23: CMOS amplifier for stability analysis and compensation.

$$C_c{}^{-1} = 2\pi f_{po}(r_{o2} \parallel r_{o4}) \,. \tag{12.35}$$

Then with $r_{o2} \parallel r_{o4} = 112$ kΩ (from the SPICE output file), $C_c = 27.6$ nF. But this value is HUGE for our integrated circuit. What to do?

Back in Chapter 8, we made the unhappy discovery that any capacitance coupling an amplifier output to the input is effectively enhanced at the input by a "Miller" factor of one plus the magnitude of the forward voltage gain. Our present predicament allows a favorable application. From SPICE data, the gain of the second-stage common-source amplifier involving M_7 is -167. Thus, the effect of a 27.6-nF capacitor connected from node 8 to ground ought to be the same as a capacitor with value 27.6 nF/(1 + 167) = 164 pF between nodes 8 and 9. This capacitor is also large, but it is manageable.

Figure 12.24 shows SPICE simulation results for the new phase margin, which is 55°. We want $\phi = 70°$, so try doubling C_c to find 55° once more. Double again — 55°. And again — 55°. Trial and error is not working. What to do?

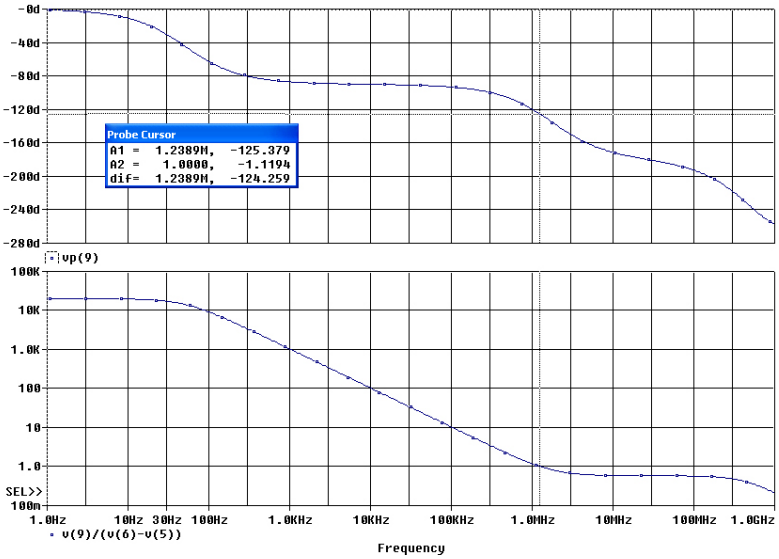


Figure 12.24: SPICE results for compensation with a feedback capacitor.

As we will demonstrate shortly, the feedback capacitor introduces a zero in addition to a low-frequency pole such that

$$AF = \frac{A_o F \left(1 + \dfrac{j\omega}{\omega_z}\right)}{\left(1 + \dfrac{j\omega}{\omega_{p0}}\right)\left(1 + \dfrac{j\omega}{\omega_{p1}}\right)\left(1 + \dfrac{j\omega}{\omega_{p2}}\right)\ldots} . \tag{12.36}$$

Moreover, the zero is in the right half of the complex plane ($\omega_z < 0$). Thus, it degrades the $|AF|$ attenuation rate while providing *negative* phase shift. The latter attribute is a killer for phase margin. The presence of the zero is evident in Fig. 12.24 in the frequency range between 3 MHz and 300 MHz. Consider the much worse phase margin were $|AF| > 1$ over that range.

A detailed pole-zero analysis uses the small-signal circuit of Fig. 12.25. The portion to the left of the dashed line reflects the differential amplifier front end with output resistance $r_{o1}{}' = r_{o2} \parallel r_{04}$. The remaining portion has M_7 as a common-source amplifier with load resistance $r_{o2}{}' = r_{o7} \parallel r_{08}$.

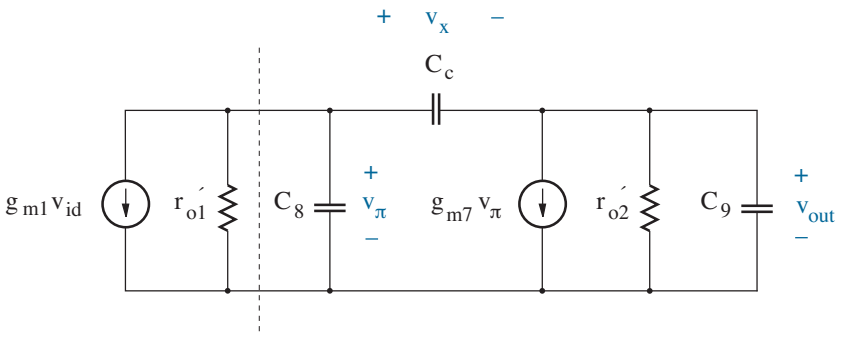


Figure 12.25: Small-signal circuit for feedback compensation.

Finding the zero is easy. If $v_{out} = 0$, $v_x = v_\pi$, $sC_c v_\pi = g_{m7} v_\pi$, and

$$\omega_z = \frac{-g_{m7}}{C_c} \, . \tag{12.37}$$

To determine the poles, we need equations for node voltages v_π and v_{out}. In matrix form, we have

$$\underbrace{\begin{pmatrix} g_{o1}{}' + sC_8 + sC_c & -sC_c \\ g_{m7} - sC_c & sC_c + g_{o2}{}' + sC_9 \end{pmatrix}}_{\mathbf{G}} \begin{pmatrix} v_\pi \\ v_{out} \end{pmatrix} = \begin{pmatrix} g_{m1} v_{id} \\ 0 \end{pmatrix} \, . \tag{12.38}$$

The poles are the s values that make the determinant of matrix $\mathbf{G}$ vanish. In turn, with $g_{m7} \gg g_{o1}{}', g_{o2}{}'$,

$$\omega_{p0} \approx \frac{1}{C_C \, g_{m7} \, r_{o1}{}' r_{o2}{}'} \tag{12.39}$$

and

$$\omega_{p1} \approx \frac{g_{m7}}{C_8 + C_9} \, . \tag{12.40}$$

The relative pole and zero positions typically satisfy

$$\omega_{po} \ll |\omega_z| \ll \omega_{p1} \, . \tag{12.41}$$

One can also show that $AF > 1$ at ω_z (a poor prospect for phase margin) if $g_{m7} < g_{m1}$ (see Problem 12.40).

Fortunately, an additional degree of freedom is available for design by including a resistor R_c in series with C_c. The same argument that produced the right-half-plane zero of Eq. 12.37 now yields

$$\omega_z = \frac{-g_{m7}}{C_c}\left(\frac{1}{1-g_{m7}R_c}\right) . \tag{12.42}$$

Thus, the zero can be positioned favorably with a suitable choice for R_c.

To find the new pole positions, we need to write three node equations, set the **G**-matrix determinant to zero, and solve for the consistent s values. The algebra involved is not attractive, as it involves the roots of a third-order polynomial equation. Thus, we are content to look to Eqs. 12.39 and 12.40 as rough estimates for ω_{p0} and ω_{p1}, respectively.

What follows is a simple three-step compensation procedure.

- Move f_z to infinity where it is out of the way.

For the design at hand, $g_{m7} = 1.81 \times 10^{-3}$ ℧. Then $R_c = 1/g_{m7} = 552\ \Omega$.

- Choose C_c to obtain a phase margin that is 15° less than desired.

In the spirt of estimation, we apply Eq. 12.40 to find

$$f_{p1} \approx \frac{g_{m7}}{2\pi(C_8 + C_9)} = 599\ \text{MHz} .$$

In turn, with $\phi = 55°$ as our temporary objective,

$$f_T = 599\ \text{MHz}\ \tan(90° - 55°) = 419\ \text{MHz}$$

and

$$f_{p0} = \frac{419\ \text{MHz}}{20,000\ \cos(90° - 55°)} = 25.6\ \text{kHz} .$$

Note that f_{p1} and f_{p0} are much higher than earlier SPICE-derived values. We use Eq. 12.39 to estimate C_c. Specifically,

$$C_c \approx \frac{1}{2\pi f_{p0}\, g_{m7}\, r_{o1}{}' r_{o2}{}'} = 0.33\ \text{pF} .$$

Trial and error SPICE simulations produce $C_c = 0.46$ pF as a correction to the preceding estimate. The C_c capacitance would have been larger had we opted to complete the design at this point with $\phi = 70°$.

- Move f_z to the left-half plane to obtain the desired phase margin.

SPICE simulations yield $\phi = 70°$ when $R_c = 910\ \Omega$. This corresponds to $f_z \approx 1.6 f_{p1}$. No all-purpose guideline is available.

Figure 12.26 shows a portion of the compensation circuit in which M_9 operates in the resistive mode such that

$$R_c{}^{-1} = r_{ds}{}^{-1} = K_n{}'(W/L)_9(v_{gs9} - V_{Tn}) \, . \tag{12.43}$$

Whereas $v_o \approx 0$, $v_{gs9} = 10$ V. However, we also have $v_{bs9} = -10$ V so that V_{Tn} adjusts through the body effect. With $\gamma = 0.2\ V^{1/2}$ and $\phi_f = 0.6$ V, $V_{Tn} = 0.5\text{ V} + 0.2\,(\sqrt{10 + 0.6} - \sqrt{0.6})\text{ V} = 0.996$ V. In turn, $(W/L)_9 = 2.4$. Trial and error with SPICE yields $(W/L)_9 = 1.8$. Figure 12.27 shows the final compensation results (without corrections for M_9 capacitance).

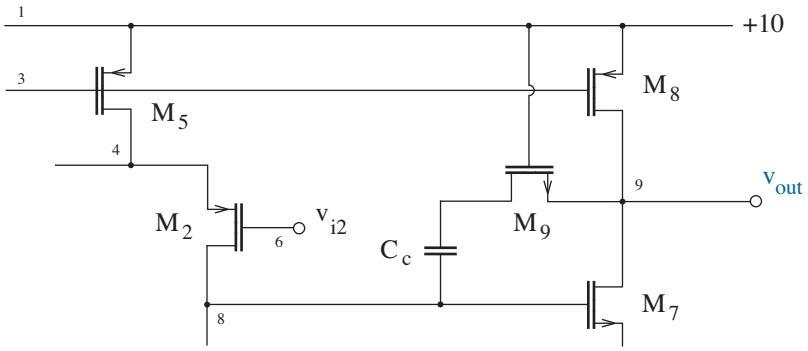


Figure 12.26: CMOS op-amp compensation with M_9 in the resistive mode.

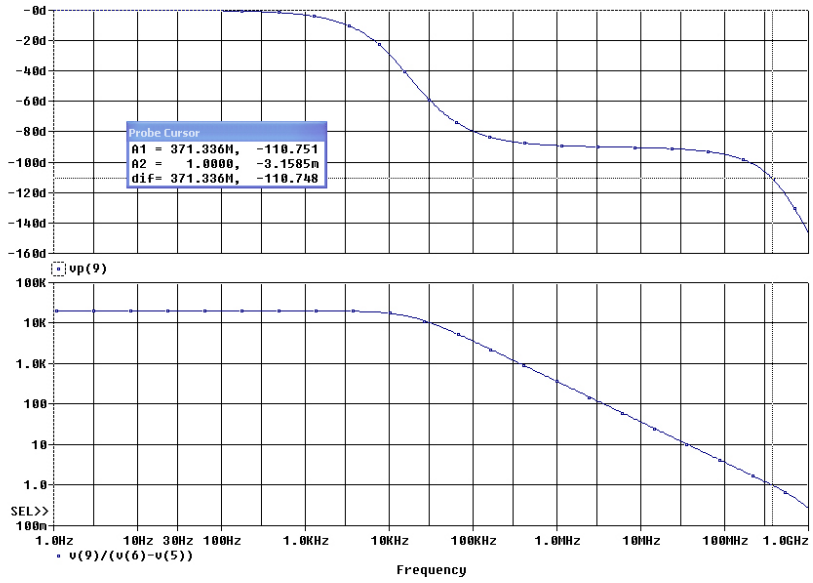


Figure 12.27: SPICE compensation results with the circuit of Fig. 12.26.

Slew Rate

Unfortunately, compensation that is achieved with the help of a feedback capacitor comes at a price—the op-amp exhibits limited output **slew rate**. Figure 12.28 shows the particular behavior when the compensated op-amp of the preceding discussion is configured as a unity gain buffer with a square-wave input of 10-V peak-to-peak amplitude. The downward transitions are nearly linear with a slew rate of about 0.5 V/ns. The upward transitions are somewhat slower.

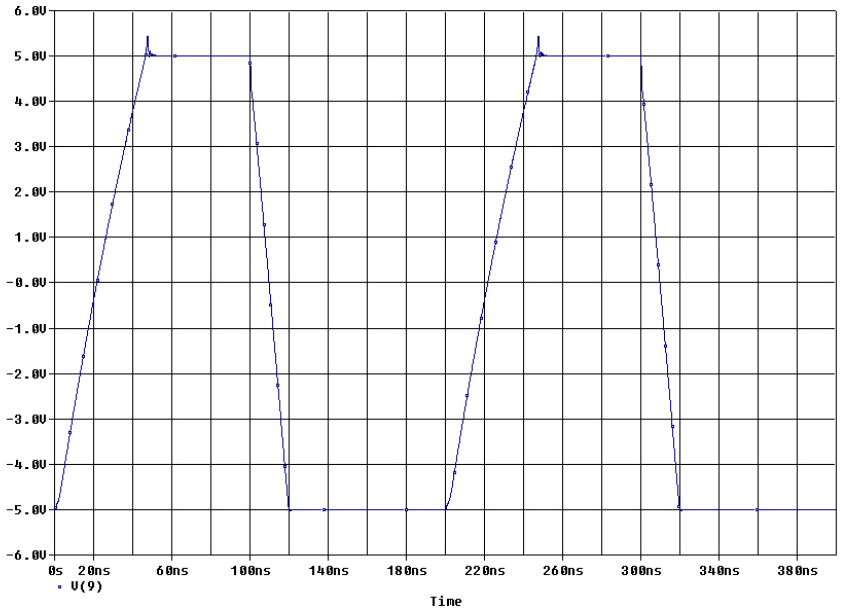


Figure 12.28: Demonstration of op-amp slew rate.

We can understand the downward slew rate with the help of Figs. 12.23 and 12.26. When the signal voltage at the non-inverting op-amp input (v_{i2}) suddenly decreases from 5 V to -5V, the voltage at the inverting input ($v_{i1} = v_{out}$) lags behind, and the differential amplifier is forced into extreme imbalance with M_2 on and M_1 off. Thus, M_4 is also off so that the entire 238-μA drain current for M_5 flows through M_2 to C_c. In turn,

$$-i_{d2} = C_c \frac{d}{dt}\left[\frac{-v_{out}}{A_2} - v_{out}\right] , \tag{12.44}$$

where A_2 is the voltage gain for the op-amp stage featuring M_7 and M_8. Then with A_2 large, the downward slew rate is approximately $-i_{d2}/C_c =$ 238 μA / 0.46 fF = 0.52 V/ns. The rationale for the upward slew rate is more complicated as a consequence of resistive-mode behavior for M_4 when M_1 is on and M_2 is off (see Problem 12.49).

Corruptive slew-rate penalties are foreign to "lawfully" varying signals. Consider a sinusoidal op-amp output signal with the form $v_{out} = \tilde{v} \sin \omega t$. The rate of change is

$$\frac{dv_{out}}{dt} = \tilde{v}\, \omega \cos \omega t \, . \tag{12.45}$$

Whereas the cosine function has unity maximum amplitude, and $\omega = 2\pi f$, the slew-rate police look the other way when

$$f < \frac{S}{2\pi \tilde{v}} \, . \tag{12.46}$$

Here, S is the posted maximum slew rate. In terms of amplitude,

$$\tilde{v} < \frac{S}{2\pi f} \, . \tag{12.47}$$

Large-amplitude signals suffer the greatest degree of corruption.

Example 12.10

A non-inverting amplifier with a voltage gain of 4 features an op-amp with a gain-bandwidth product of 800 kHz and a maximum slew rate of 1 V/μs. The maximum amplitude of any sinusoidal input is 0.5 V peak-to-peak. Determine the effective circuit bandwidth.

Solution

We follow Example 12.1 to determine a bandwidth of 800 kHz/4 = 200 kHz for signals with small amplitude. Then subject to maximum $\tilde{v} = 1.0$ V, we apply Eq. 12.46 to find $f = 159$ kHz as the upper frequency limit that avoids slew-rate-induced corruption. This upper limit is less than 200 kHz, so it is the effective circuit bandwidth.

The feedback capacitor used for compensation is the physical basis for the "sluggish response" that accompanies the "suppression of unwanted dynamic behavior" cited in Chapter 1. Comparators do not have feedback, they do not require compensation, and slew rate is not a problem.

Nearly all commercial op-amps have an internal feedback capacitor for compensation. The typical maximum slew rate and a phase or gain margin are generally specified on the manufacturer's data sheet. Notwithstanding, some op-amps are available without the internal compensation capacitor. Package pins are provided for the inclusion of an external capacitor so that an optimum combination of phase margin and slew rate can be obtained. The design of high-performance current-to-voltage converters for detecting fast optical signals is one application that warrants this tradeoff.

12.4 Oscillators

Sometimes instability is desirable. Oscillator circuits that produce periodic waveforms are important for communication systems, for the regulation of finite-state machines and computers, and for numerous test applications. This section examines a category of oscillator circuits that feature a single BJT or MOSFET for oscillation frequencies extending into the GHz range. Oscillators involving op-amps for relatively low frequencies are considered in the end-of-chapter problems.

Colpitts Oscillator

Figure 12.29 shows a **Colpitts oscillator**. The supply voltage V^+ is at ac ground, so the resonant circuit with L, the C_1 and C_2 series capacitance, and R has its output voltage sampled through a capacitive voltage divider. The amplifier accepting the divider voltage is configured as common base. In turn, the BJT collector voltage is in phase with the divider voltage, and resonator oscillations are sustained by **positive feedback**.

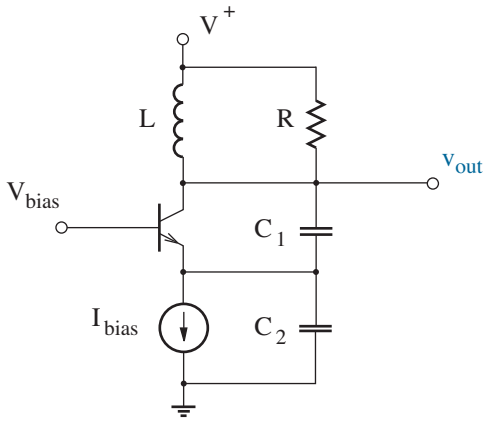


Figure 12.29: Colpitts oscillator circuit.

In what follows, we show that the Colpitts oscillator is governed by a system function of the form

$$\frac{v_{out}}{v_t} = \frac{A}{1 + AF} \tag{12.48}$$

and that instability (oscillation) results when $AF = -1$. A similar analysis can be applied to to a system defined through the relation

$$\frac{v_{out}}{v_t} = \frac{A}{1 - AF'}\,, \tag{12.49}$$

and instability results with $AF' = 1$, the so-called **Barkhausen** condition. We take the former approach, as it involves the methods of Chapter 11.

Figure 12.30 shows the Colpitts oscillator as interconnected two-ports. For the moment, we are only interested in circuit conditions that promote instability and a specific oscillation frequency. Thus, the small-signal BJT model is adequate (C_π, $C_\mu \ll C_1$, C_2), pending non-linear considerations. Although absent in the large-signal circuit, the v_t source completes a series configuration on the left sides of the two-ports while providing stimulation. The right-side shunt connection is more readily apparent.

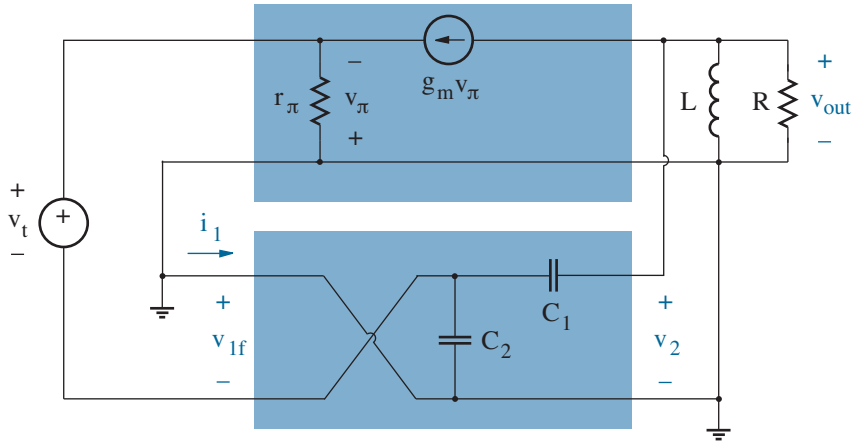


Figure 12.30: Colpitts oscillator circuit as two interconnected two-ports. Note the ground connections.

To find the feedback factor F that is applicable to the lower two-port, we apply a test voltage source v_2 at the right side —voltage is shared there, we break the current loop on the left side so $i_1 = 0$, and we measure v_{1f}. In turn,

$$F = \frac{v_{1f}}{v_2} = \frac{\dfrac{-1}{j\omega C_2}}{\dfrac{1}{j\omega C_1} + \dfrac{1}{j\omega C_2}} \tag{12.50}$$

or

$$F = \frac{-C_1}{C_1 + C_2}\,. \tag{12.51}$$

The feedback factor is independent of frequency.

We pause to note that the two-ports could have been configured with left-side shunt connections and a current source as the means of stimulation. This would not have been appropriate, since the oscillation frequency would no longer incorporate C_2 (see Problem 12.51).

To find the feedback loading on the left side of the so-called "A" circuit, we set $v_2 = 0$ (a shared voltage) so that the equivalent capacitance in the input loop is $C_1 + C_2$. On the right side, we set $i_1 = 0$ (a shared current) so that the series combination $C' = C_1C_2/(C_1 + C_2)$ parallels L and R. Figure 12.31 shows the "A" circuit with a redefined v_π.

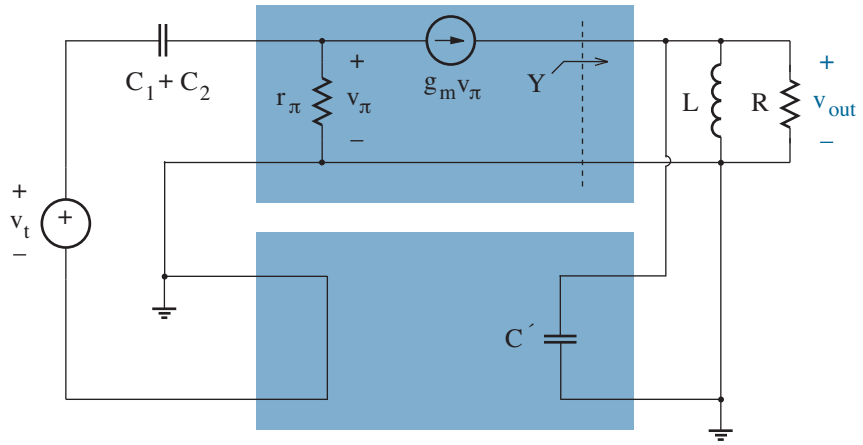


Figure 12.31: "A" circuit for the Colpitts oscillator with feedback loading.

The governing equations for node voltages v_π and v_{out} are

$$(v_t - v_\pi)\jmath\omega(C_1 + C_2) = g_m v_\pi + g_\pi v_\pi \,, \tag{12.52}$$

where $g_\pi = 1/r_\pi$, and

$$g_m v_\pi = Y v_{out} \,, \tag{12.53}$$

where $Y = \jmath\omega C' + 1/\jmath\omega L + G$ is the admittance of the resonant "tank" to the right of the dependent source. Then solving for v_{out}, we find

$$A = \frac{v_{out}}{v_t} = \frac{\jmath\omega(C_1 + C_2)\, g_m}{[\, g_m + g_\pi + \jmath\omega(C_1 + C_2)]\, Y} \,. \tag{12.54}$$

Now multiply Eq. 12.54 by F (Eq. 12.51), set $AF = -1$ (for instability), and rearrange so that real and imaginary terms separately combine to zero. For the real terms,

$$\frac{C_1 + C_2}{L} - \omega^2(C_1 + C_2)C' + (g_m + g_\pi)G = 0 \,. \tag{12.55}$$

And for the imaginary terms,

$$(g_m + g_\pi)\left(\omega C' - \frac{1}{\omega L}\right) + \omega(C_1 + C_2)G - \omega g_m C_1 = 0 \,. \tag{12.56}$$

Typically, $g_\pi \ll g_m$ (as for a MOSFET).

Assume that the $(g_m+g_\pi)G$ term in Eq. 12.55 is smaller than the others. Then the angular frequency of oscillation is

$$\omega_o = \frac{1}{\sqrt{LC'}} \,. \tag{12.57}$$

We insert this result into Eq. 12.56 so that the first term vanishes and

$$g_m R \left(\frac{C_1}{C_1 + C_2} \right) > 1 \tag{12.58}$$

is the condition for sustained oscillations. Since the resonant circuit looks like resistance R at $\omega = \omega_o$ —inductive and capacitive admittances cancel, Eq. 12.58 states that the open-loop gain $(-g_m R \times F)$ must exceed unity. Keeping the $(g_m + g_\pi)G$ term in Eq. 12.55 typically yields a minor shift in ω_o and the gain requirement.

The remaining design issues go a bit beyond the scope of this text.[1]

- The oscillation amplitude is limited by non-linear transistor behavior.

Let $v_{out}(t) = v_{out}|_Q + \tilde{v} \sin \omega_o t$. As amplitude $\tilde{v}$ gets larger, the effective transconductance of the BJT becomes smaller, and conditions promoting signal growth are relaxed. In one approximation,

$$\tilde{v}_{out} \approx 2\, I_{\text{bias}} R \left(\frac{C_2}{C_1 + C_2} \right) \,. \tag{12.59}$$

More accurate, but complicated expressions are available.

- The C_2/C_1 ratio governs the oscillator **phase noise**.

Phase noise reflects the tendency of an oscillator circuit to exhibit small random variations in output-voltage zero crossings as shown in Fig. 12.32. To minimize these fluctuations while preserving a reasonable F, we choose

$$\frac{C_2}{C_1} \sim 4 \tag{12.60}$$

as a convenient guide for design.

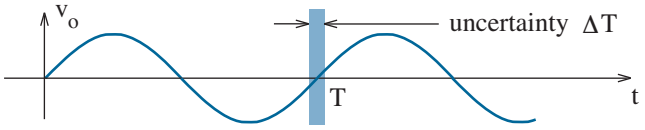


Figure 12.32: Oscillator fluctuations indicative of phase noise.

[1] See, for example, Thomas H. Lee, *The Design of CMOS Radio Frequency Integrated Circuits*, 2^{nd} Edition, Cambridge University Press, 2004.

Example 12.11

Design a Colpitts oscillator with an output amplitude of 5 V at 50 MHz. The supply voltage is +5 V and the BJT has $\beta_F = 200$.

Solution

We use the circuit of Fig. 12.33 in which I_{bias} is set through V_{bias} and R_e. Let $R = 1\ \text{k}\Omega$ and $C_1 = 20$ pF. Then with Eq. 12.60 to guide, $C_2 = 80$ pF. In turn, Eq. 12.59 requires

$$I_{bias} = \frac{5\ \text{V}}{2 \times 1\ \text{k}\Omega \times 0.8} = 3.13\ \text{mA}\ .$$

The consistent g_m is 3.13 mA/25.9 mV = 0.121 $\mho$ so that the left side of Eq. 12.58 is 24. The unity threshold condition for oscillation is satisfied. To establish I_{bias}, we arbitrarily set $V_{bias} = 2.5$ V. Thus, the node voltage at the BJT emitter is approximately 1.8 V and

$$R_e = \frac{1.8\ \text{V}}{3.13\ \text{mA}} = 580\ \Omega\ .$$

Finally, with $C' = C_1C_2/(C_1 + C_2) = 16$ pF, we apply Eq. 12.57 to find

$$L = \frac{1}{(2\pi \times 50 \times 10^6)^2 \times 16 \times 10^{-12}} = 630\ \text{nH}\ .$$

This represents a first approximation for the desired oscillation frequency.

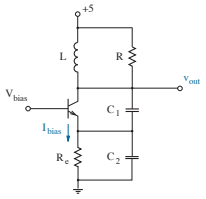


Figure 12.33: Colpitts circuit for Example 12.11.

The SPICE code that tests the design thus far is unremarkable except for the statements

```
Ipulse      2          0          PULSE ( 1   1m   0   0   0   1p   1)
```

to provide a pulse with 1-mA amplitude and 1-ps duration at the output node as an initial stimulus and

```
.tran       1p         2500n      2400n      10p
```

to order data between 2400ns and 2500ns subject to a 10-ps step-size limit. The BJT is adequately modeled with IS=10f.

SPICE .probe results indicate an oscillation amplitude of about 5.7 V. This can be reduced by increasing R_e. Nevertheless, we are content to leave the amplitude alone with the expectation that parasitic circuit losses will make up for the difference when the circuit is built. The oscillation period is $T = 17.76$ ns, so the frequency is $f_o = 1/T = 56.3$ MHz, which is high. Trial and error leads to $L = 800$ nH for the desired f_o. However, further adjustments may eventually be needed to offset unaccounted parasitics.

Figure 12.34 shows the final output waveform for the Colpitts oscillator. A non-sinusoidal character is clearly evident.

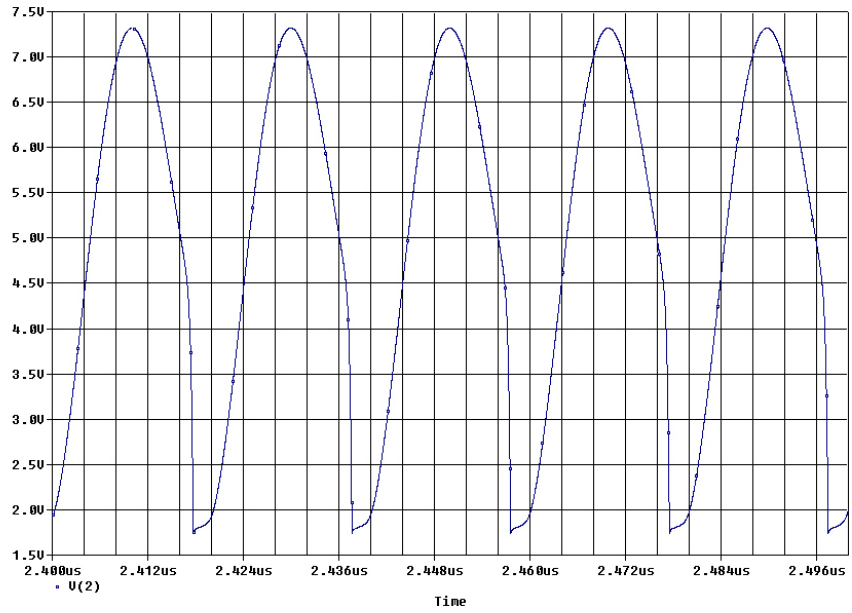


Figure 12.34: .probe results for the Colpitts oscillator of Example 12.11.

Crystal Oscillator

A crystal oscillator typically uses a quartz piezoelectric resonator to define a precise frequency of oscillation in circuit subject to feedback instability. Figure 12.35 shows an electrical model for a quartz resonator. The series combination of capacitor C_x, inductor L_x, and resistor R_x actually reflects a *mechanical* system of effective masses, springs, and dashpots that model the displacement of atoms when the quartz crystal is electrically stimulated. Capacitor C_p is a purely electrical parasitic element.

The value of C_x is small, of the order of 20 fF. Thus with

$$L_x \approx \frac{1}{\omega_o{}^2 C_x} \,, \tag{12.61}$$

a 10-MHz resonant frequency has L_x of the order of 10 mH, a HUGE value. Meanwhile, R_x is about 20 Ω. The quality factor of the series resonator is

$$Q = \omega_o \frac{L_x}{R_x} \,, \tag{12.62}$$

so Q is of the order of 30,000, and the characteristic resonance is very sharp. The ω_o temperature variation is of the order of 1 ppm/°C.

As a practical matter, the thickness of a quartz film that is fabricated to serve as a resonator is inversely proportional to ω_o. Very thin films are difficult to handle, so quartz crystals tend to be expensive for $f_o > 30$ MHz. To get around this problem, the quartz resonator can be made to operate at $N\omega_o$, where N is odd. Unfortunately,

$$R_x(N) \sim \frac{R_{xo}}{N^2} \,, \tag{12.63}$$

where R_{xo} is the series resistance effected at the fundamental resonance ω_o. Thus, the $N\omega_o$ **overtones** are realized with a loss of quality factor Q.

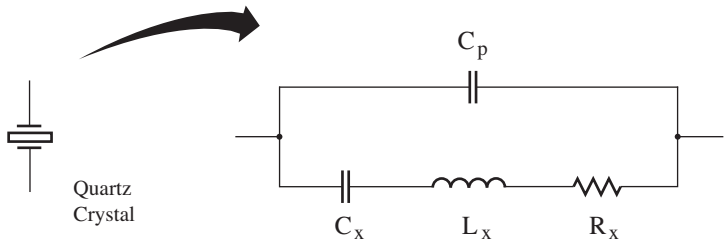


Figure 12.35: Electrical model for a quartz crystal resonator.

Crystal oscillator circuits are easy to understand if we remember that a series RLC resonant circuit looks like resistor R when operated at $\omega = \omega_o$. The equivalent series impedance is typically substantially higher for $\omega \neq \omega_o$. Thus, we use a crystal resonator to degrade feedback away from resonance as shown in Fig. 12.36. Both of the circuits have the Colpitts configuration. In Circuit A, the feedback coupling from the capacitor divider is available in the vicinity of the crystal resonant frequency. In Circuit B, the BJT base connection is at ac ground (for common-base gain) in the same ω_o vicinity. The much broader (lower-Q) resonance specifications for the RLC circuit external to the crystal resonator are assumed to show a comparable ω_o.

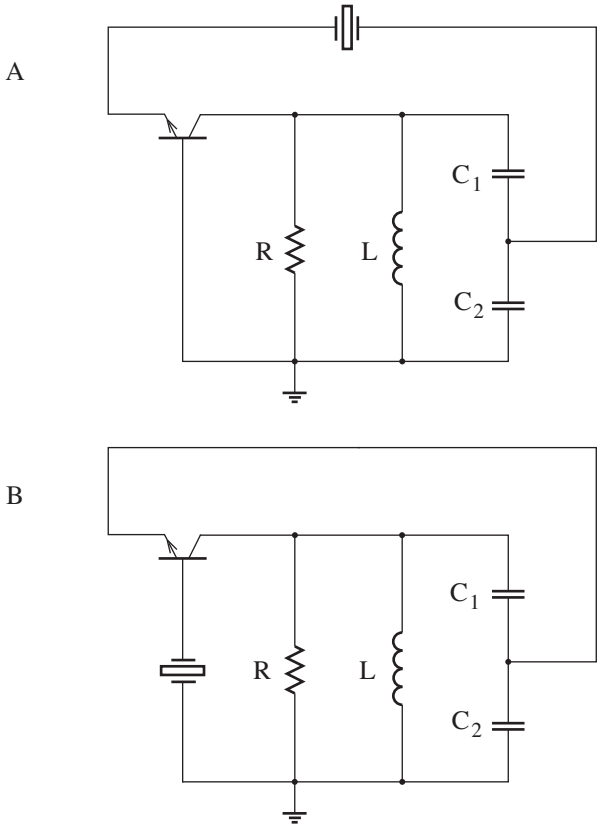


Figure 12.36: Crystal oscillator circuits (biasing not shown).

Another way to realize a crystal oscillator circuit is to allow the crystal resonator to operate above ω_o so that the device has inductive character. The selection of a higher operating frequency is made through resonance with an external capacitance (see Problem 12.58).

Concept Summary

Feedback alters the pole (and zero) characteristics of amplifier circuits.

- In the case of a single dominant pole, feedback increases amplifier bandwidth while decreasing amplifier gain.
  - The gain-bandwidth product is essentially constant.
  - This behavior typically applies to non-inverting and inverting op-amp amplifiers.
- In the case of two poles, feedback can change the step response from one with pure exponential decay to one with oscillatory decay.
- In the case of three poles, feedback promotes an unstable system if
  - $|AF| > 1$ at the frequency corresponding to 180° phase shift;
  - $\angle AF + 180° = \phi < 0$ (negative phase margin) when $|AF| = 1$.
- Unstable feedback systems usually need some form of compensation. The simplest procedure is to add a new pole at f_{p0}.
  - An optimum phase margin is about 70°.
  - Given a system with its lowest pole frequency at f_{p1},
    - The required unity-gain frequency is

$$f_T = f_{p1} \tan(90° - \phi) .$$

    - Let $AF = A_o F$ in the dc limit. The new pole frequency is

$$f_{p0} = \frac{f_T}{A_o F \cos(90° - \phi)} .$$

  - An alternative procedure modifies f_{p1} in relation to f_{p2}.
  - Internal compensation for CMOS op-amps generally introduces a zero in the right-half plane that compromises the phase margin. The zero must be forced to a non-interfering frequency.
  - Op-amp compensation tends to decrease slew rate.
- Oscillator circuits are deliberately unstable.
  - The popular Colpitts oscillator requires a single transistor.
  - Crystal oscillators offer precise resonance with high Q.

Problems

Section 12.1

12.1 An amplifier has a dc gain of 400 and a single pole at 250 kHz. Determine the new cutoff frequency subject to feedback with $F = 0.08$.

12.2 An amplifier with a single pole at 800 kHz participates in a feedback circuit for which $F = 0.32$. Find the consistent dc gain if $f_o = 2$ MHz.

12.3 An amplifier has a dc gain of 20 and two poles at 600 kHz and 4 MHz. Find the new pole positions subject to feedback with $F = 0.25$.

12.4 Repeat Problem 12.3, but with a dc gain of 50.

12.5 An amplifier has poles at 1.2 MHz and 20 MHz. Subject to feedback with $F = 0.12$, the amplifier response indicates poles at 8 MHz and 13.2 MHz. Determine the dc gain.

12.6 An amplifier has poles at 1.2 MHz and 20 MHz. Subject to feedback with $F = 0.12$, a step excitation yields a 5-MHz oscillatory response with an envelope that decays with $\tau = 94.3$ ns. Determine the dc gain.

12.7 A feedback system features an "A" circuit with $A = 400$ as $f \to 0$ and poles at 1 MHz, 40 MHz, and 200 MHz. Use the ω_{180} test to determine if the system is stable for the feedback factor $F = 0.5$.

12.8 A feedback system features an "A" circuit with $A = 100$ as $f \to 0$ and poles at 2 MHz, 10 MHz, and 80 MHz. Use the ω_{180} test to determine if the system is stable for the feedback factor $F = 0.2$.

12.9 A feedback system features an "A" circuit with $A = 1200$ as $f \to 0$ and poles at 1 MHz, 120 MHz, and 800 MHz. Use the ω_{180} test to determine if the system is stable for the feedback factor $F = 1$.

12.10 Repeat Problem 12.7 using the Nyquist test and specify the phase margin.

12.11 Repeat Problem 12.8 using the Nyquist test and specify the phase margin.

12.12 Repeat Problem 12.9 using the Nyquist test and specify the phase margin.

12.13 The op-amp in the circuit of Fig. P12.13 has a gain-bandwidth product of 500 kHz. Determine the circuit bandwidth, then check with SPICE.

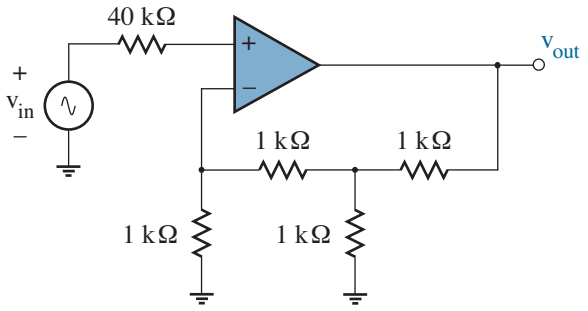


Figure P12.13

12.14 The op-amp in the circuit of Fig. P12.14 has a gain-bandwidth product of 500 kHz. Determine the circuit bandwidth, then check with SPICE.

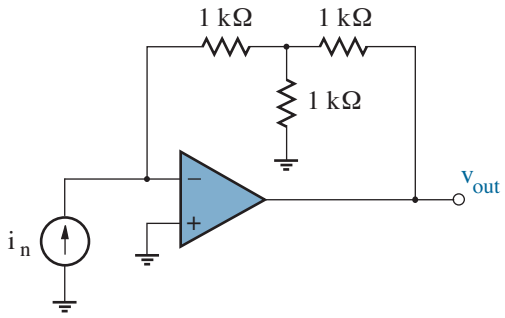


Figure P12.14

12.15 The op-amp in the circuit of Fig. P12.15 has a gain-bandwidth product of 1 MHz. Determine the circuit bandwidth, then check with SPICE.

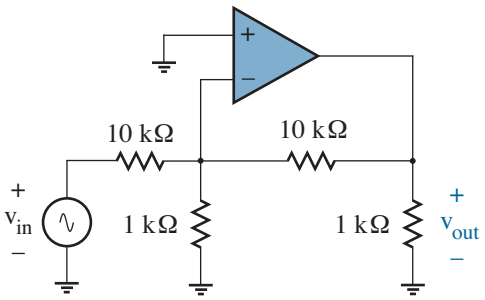


Figure P12.15

12.16 The op-amp in the circuit of Fig. P12.16 has a gain-bandwidth product of 1 MHz. Determine the circuit bandwidth, then check with SPICE.

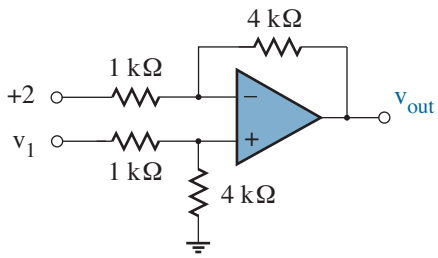


Figure P12.16

12.17 Figure P12.17 shows a "Pedagogical Circuit" that is useful for finding phase-margin and stability conditions with the help of SPICE. Box zero uses a voltage-dependent voltage source (E0) to set the AF product, while boxes 1, 2, and 3 use a similar source and an RC network to set pole frequencies f_{p1}, f_{p2}, and f_{p3}. These dependent sources are described with statements of the form

E0 a b c d g

where a and b denote positive and negative nodes for the dependent output, c and d denote positive and negative nodes for the controlling input, and g is the gain (unity for E1, E2, and E3). Subject to $C = 159.2$ pF, a pole frequency is $1\,\text{GHz}/R$, with R expressed in ohms.

Let $f_{p1} = 10$ kHz, $f_{p2} = 1$ MHz, and $f_{p3} = 100$ MHz. Simulate the pedagogical circuit with SPICE to find the phase margin for feedback systems with $AF = 5000$, $AF = 10,000$, and $AF = 20,000$.

12.18 The pedagogical circuit of Problem 12.17 is used as a unity-gain feedback amplifier with its output tied directly to the negative input of Box 0. Use SPICE to demonstrate the step response for $AF = 5000$, $AF = 10,000$, and $AF = 20,000$.

12.19 The pedagogical circuit of Problem 12.17 is used in place of the operational amplifier shown in Example 12.1. The pole frequencies are at 1 MHz, 12 MHz, and 50 MHz. Determine the phase margin.

12.20 The pedagogical circuit of Problem 12.17 is used in place of the operational amplifier shown in Example 12.1. The pole frequencies are at 1 MHz, 12 MHz, and 50 MHz. Determine the maximum A gain for stable operation ($\phi > 0$).

12.21 Use SPICE to determine the phase margin for the circuit of Example 11.2. Assume $C_\pi = 5$ pF and $C_\mu = 2$ pF for both transistors.

12.22 The pedagogical circuit of Problem 12.17 is used in place of the operational amplifier shown in Example 12.3. The pole frequencies are at 1 MHz, 12 MHz, and 50 MHz. Determine the phase margin.

12.23 The pedagogical circuit of Problem 12.17 is used in place of the operational amplifier shown in Example 12.3. The pole frequencies are at 1 MHz, 12 MHz, and 50 MHz. Determine the maximum A gain for stable operation ($\phi > 0$).

12.24 Use SPICE to determine the phase margin for the circuit of Example 12.5 when the drain-to-ground capacitors for M_1, M_2, and M_3, are 1 pF, 2 pF, and 3 pF, respectively.

12.25 Use SPICE to determine the phase margin for the circuit of Example 11.5. Provide current sources to bias Q_1 and Q_2. Assume IS=10f and BF=200. Let $C_\pi = 5$ pF and $C_\mu = 2$ pF for both transistors.

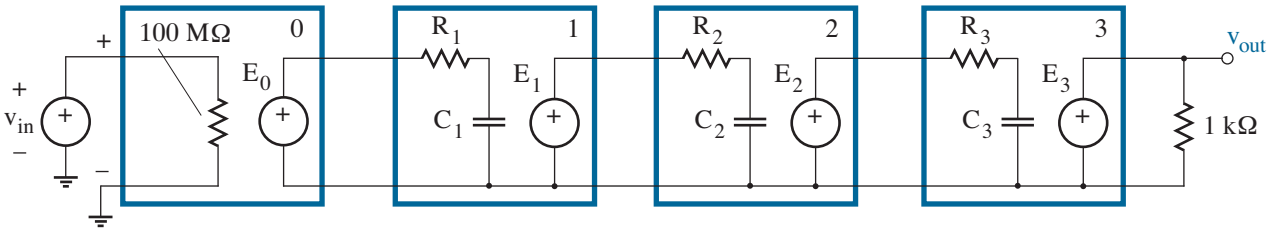


Figure P12.17

12.26 Use SPICE to determine the phase margin for the circuit of Example 11.6. Assume that the MOSFETs feature $K'W/L = 2$ mA/V^2 and $V_T = 1$ V. To establish the proper biasing, it will be helpful to provide a 1-kΩ source resistor with bypass capacitor for M_2. Let $C_{gs} = C_{gd}$ for each device, and assume 1-pF capacitive loading to ground at each drain.

12.27 Use SPICE to determine the phase margin for the circuit of Problem 11.11. Let $C_{gs} = C_{gd}$ for each MOSFET, and assume 1-pF capacitive loading to ground at each drain.

12.28 Use SPICE to determine the phase margin for the circuit of Problem 11.21 (revised Example 11.4). Let $C_\pi = 5$ pF and $C_\mu = 2$ pF for both transistors.

Section 12.2

12.29 Apply the logarithmic characteristics of the magnitude Bode plot to prove Eq. 12.21.

12.30 A feedback system features an "A" circuit with $A = 40,000$ in the dc limit and poles at 1 MHz, 40 MHz, and 200 MHz. Find the new pole frequency that establishes $\phi = 45°$.

12.31 A feedback system features an "A" circuit with $A = 25,000$ in the dc limit and poles at 2 MHz, 10 MHz, and 80 MHz. Find the new pole frequency that establishes $\phi = 45°$.

12.32 Repeat Problem 12.30, but let $\phi = 70°$. Demonstrate with SPICE (see Problem 12.17).

12.33 Repeat Problem 12.31, but let $\phi = 60°$. Demonstrate with SPICE (see Problem 12.17).

12.34 A feedback system features an "A" circuit with $A = 100,000$ in the dc limit and poles at 1 MHz, 20 MHz, and 120 MHz. Establish $\phi = 65°$ by moving the dominant pole frequency. Demonstrate with SPICE (see Problem 12.17).

12.35 Establish $\phi = 70°$ for the circuit of Problem 12.25 by adding a new pole. Use SPICE to demonstrate your compensation design.

12.36 Establish $\phi = 70°$ for the circuit of Problem 12.26 by adding a new pole. Use SPICE to demonstrate your compensation design.

12.37 An amplifier with three overlapping poles at ω_p is compensated with a new pole at $\omega_{p0} \ll \omega_p$.

(a) Show that the compensation design equations take the form

$$\omega_T = \omega_p \tan\left(\frac{90° - \phi}{3}\right)$$

and

$$\omega_{p0} = \frac{\omega_T}{A_o F} \cos^{-3}\left(\frac{90° - \phi}{3}\right).$$

(b) Apply the results to Example 12.9 (Option 1).

12.38 An amplifier with three overlapping poles at ω_p is compensated by moving one of the poles to $\omega_{p0} \ll \omega_p$.

(a) Show that the compensation design equations take the form

$$\omega_T = \omega_p \tan\left(\frac{90° - \phi}{2}\right)$$

and

$$\omega_{p0} = \frac{\omega_T}{A_o F} \cos^{-2}\left(\frac{90° - \phi}{2}\right) .$$

(b) Apply the results to Example 12.9 (Option 2).

Section 12.3

12.39 Derive Eqs. 12.39 and 12.40 subject to appropriate approximations.

12.40 Show that the circuit of Fig. 12.23 has $v_{out}/v_{id} > 1$ at ω_z when $g_{m7} < g_{m1}$.

12.41 Repeat the compensation design of Section 12.3 with $\phi = 60°$. Demonstrate with SPICE.

12.42 Repeat the compensation design of Section 12.3 with $\phi = 70°$, but include an additional 0.5-pF load at the output. Demonstrate with SPICE.

12.43 Repeat the compensation design of Section 12.3 with $\phi = 70°$, but double the value for C_8. Demonstrate with SPICE.

12.44 Repeat the compensation design of Section 12.3 with $\phi = 70°$, but double the quiescent drain current for M_7 with $v_{out}|_Q = 0$. Make consistent capacitance adjustments. Demonstrate with SPICE.

12.45 Figure P12.45 has a common-emitter circuit to be used as the second gain stage in an op-amp. The BJT has transconductance g_m and small-signal resistance r_π. Ignore other small-signal components. Capacitor C_c is intended for compensation.

(a) Determine an expression for the dominant pole.

(b) Discuss the influence of the circuit zero, and compare with that for a MOSFET amplifier.

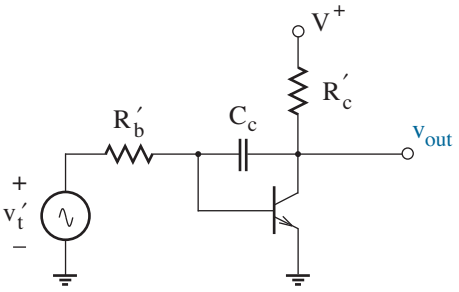


Figure P12.45

12.46 The LM741 operational amplifier features a 30-pF compensation capacitance (see the schematic diagram at the end of Chapter 10). Determine the dominant pole frequency that results.

12.47 An op-amp with a slew rate of 2 V/μs is used to amplify a 75-kHz sinusoidal input. Determine the maximum output amplitude that avoids slew-rate problems. Assume acceptable bandwidth.

12.48 A non-inverting op-amp amplifier with a gain of +5 is used to process a rectangular pulse train with the form shown in Fig. P12.48. Pulse heights of 1 V, 0.5 V, -0.5 V, or -1 V are randomly distributed. A pulse detector at the output correctly assesses the pulse value if the associated voltage is within 1.0 V of a legitimate 5X level for 0.2 μs. Find an acceptable maximum slew rate and operating frequency.

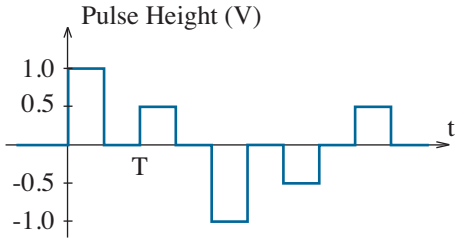


Figure P12.48

12.49 Use SPICE to replicate Fig. 12.28 for the compensated op-amp of Section 12.3. Then investigate the circuit behavior that produces upward slew rate.

12.50 Estimate the slew rate that applies to the LM741 op-amp shown at the end of Chapter 9.

Section 12.4

12.51 Consider the Colpitts oscillator of Fig. 12.29.

(a) Draw the circuit as two interconnected two-ports in the shunt-shunt configuration with a current-source small-signal excitation.

(b) Show that feedback analysis misrepresents C_2.

12.52 Design a Colpitts oscillator with an output amplitude of 4 V at 100 MHz. The supply voltage is +10 V and the BJT has $\beta_F = 160$. Let $R = 2$ kΩ and $C_1 = 15$ pF (see Fig. 12.29). Demonstrate with SPICE.

12.53 Repeat Problem 12.52, but use a MOSFET with $K_n' = 50\ \mu\text{A/V}^2$ and $V_T = 0.5$ V.

12.54 Figure P12.54 shows a **Hartley** oscillator.

(a) Draw the circuit as two interconnected two-ports with an appropriate source of excitation.

(b) Derive appropriate design equations.

12.55 Figure P12.55 shows a **Clapp** oscillator (biasing excluded).

(a) Draw the circuit as two interconnected two-ports with an appropriate source of excitation.

(b) Derive appropriate design equations.

12.56 Figure P12.56 shows a **Pierce** oscillator.

(a) Draw the circuit as two interconnected two-ports with an appropriate source of excitation.

(b) Derive appropriate design equations.

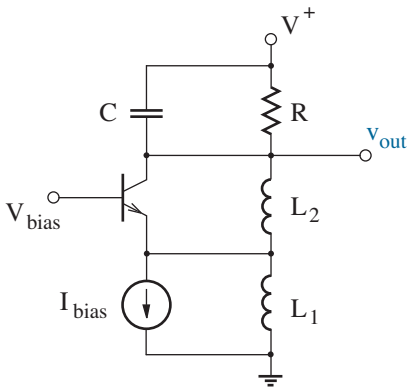


Figure P12.54

12.57 Consider a crystal resonator for which $f_o = 4.12$ MHz and $Q = 25,000$.

(a) Determine L_x and R_x if $C_x = 12$ fF.

(b) Consider a Colpitts circuit with the crystal resonator in the feedback loop as shown in Fig. 12.36a. Complete an oscillator design with 2-V output amplitude. Assume $R = 1$ kΩ, $C_1 = 33$ pF, and $C_p = 2$ pF. Use SPICE to demonstrate your design.

12.58 Consider a crystal resonator for which $f_o = 6.42$ MHz and $Q = 22,000$.

(a) Determine L_x and R_x if $C_x = 18$ fF.

(b) Use SPICE to plot the impedance of the crystal resonator over the range from 6 MHz to 7 MHz. Assume $C_p = 3.2$ pF.

(c) Let the crystal resonator take the place of L in the Pierce oscillator circuit of Fig. P12.56. Determine values for C_1 and C_2 that allow for an oscillation frequency of 6.2 MHz.

12.59 Figure P12.59 is a **Wien-Bridge** oscillator.

(a) Draw the circuit as two interconnected two-ports with an appropriate source of excitation.

(b) Derive appropriate design equations.

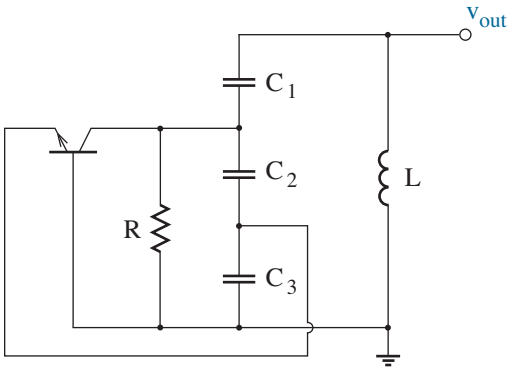


Figure P12.55

12.60 Figure P12.60 shows a **phase-shift** oscillator.

(a) Draw the circuit as two interconnected two-ports with an appropriate source of excitation.

(b) Derive appropriate design equations.

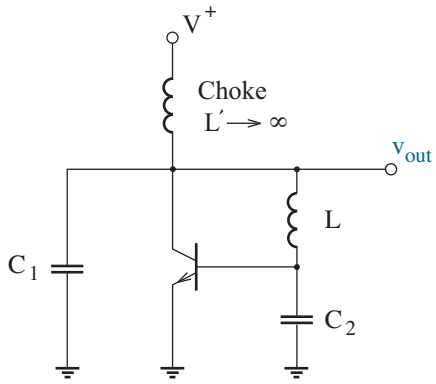


Figure P12.56

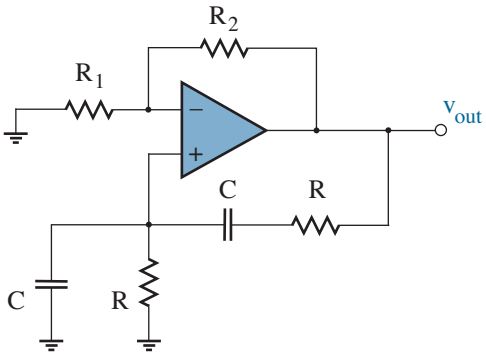


Figure P12.59

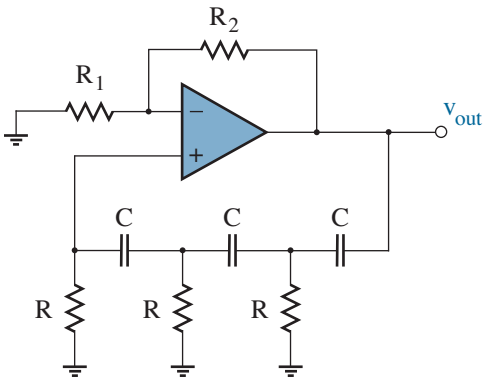
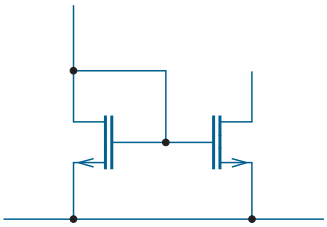


Figure P12.60



Chapter 13

Conditionings and Corruptions

Apart from amplifiers considered in previous chapters, electronic circuits are useful for signal conditioning, especially as filters that selectively remove or possibly enhance signal components over a particular range of frequency. This chapter explores the basic principles of filter design using op-amps or integrated circuits. We also explore unintentional signal corruption through the addition of noise or distortion.

Objectives

After studying this chapter and working through the problems, you should be able to perform each of the following:

- Design a first-order active low- or high-pass filter (Section 13.1).
- Design a Butterworth, Chebyshev, or Bessel filter with desired stop-band and passband characteristics (Section 13.1).
- Design a state-variable integrated-circuit filter using one of several integrator functions (Section 13.2).
- Calculate the root-mean-square noise voltage that is produced by a simple resistive circuit (Section 13.3).
- Characterize an op-amp amplifier circuit in terms of noise figure and other noise performance indicators (Section 13.3).
- Characterize an op-amp amplifier circuit in terms of the distortion products that are introduced (Section 13.4).
- Demonstrate noise and distortion circuit characteristics using SPICE (Sections 13.3 and 13.4).

13.1 Analog Filters

Amplification is a signal conditioning process that ideally imparts uniform scaling over an unlimited spectrum of constituent frequencies. In contrast, filtering provides frequency-selective scaling, as in the modification of base, midrange, and treble audio components to conform with individual tastes. This section provides an overview of analog filter characteristics that target a desired response. Section 13.2 examines contemporary circuit designs.

Filters are generally classified in terms of a **pass-band** wherein all signal frequency components are unmodified apart from a constant gain factor K and a **stop-band** wherein all frequency components are fully suppressed. Figure 13.1 shows the four major classifications in terms of a characteristic (cutoff or center) angular frequency ω_o and filter **bandwidth** B.

- **Low-pass:** Passband below ω_o.
- **High-pass:** Passband above ω_o.
- **Bandpass:** Passband inclusive of $\omega_o \pm B/2$.
- **Bandstop:** Passband exclusive of $\omega_o \pm B/2$ (notch filter).

These filters are further classified as **passive** ($K \leq 1$) or **active** ($K > 1$). We are comfortable with op-amps, so we focus on the latter.

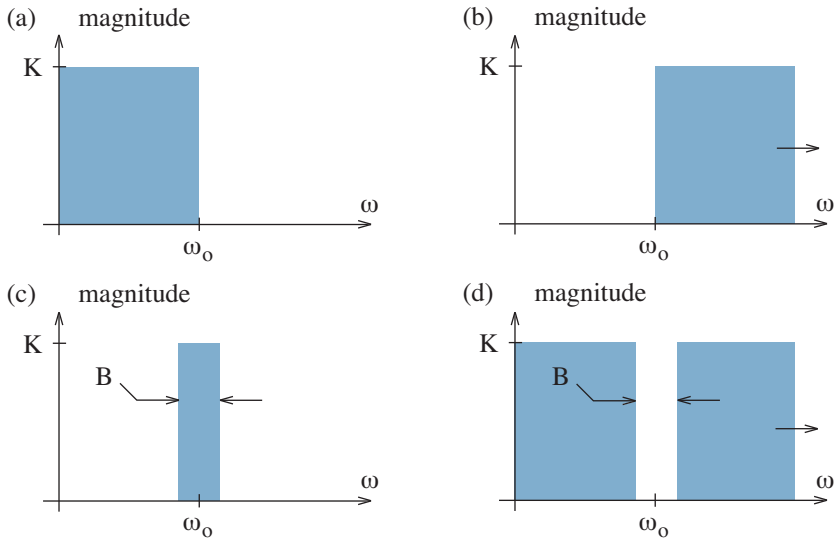


Figure 13.1: Major analog filter classifications: (a) low-pass; (b) high-pass; (c) bandpass; (d) bandstop (notch). Passbands are shaded.

The "brick wall" filters of Fig. 13.1 are ideal in the sense that they have abrupt pass-band/stop-band transitions at or near ω_o. Consider the ideal low-pass case for which the transfer function has the form

$$H(s) = \begin{cases} K & \omega \le \omega_o \\ \\ 0 & \omega > \omega_o \end{cases} \tag{13.1}$$

subject to $s = \jmath\omega$. From our study of circuit theory, we know that $H(s)$ is the Laplace transform of the system impulse response $h(t)$. Many readers are probably accustomed to finding inverse Laplace transforms from tables. Nevertheless, in general

$$h(t) = \frac{1}{2\pi\jmath} \int_{-\jmath\infty}^{\jmath\infty} H(s)\, e^{st}\, ds\,. \tag{13.2}$$

Real signals require $H(s) = H(-s)$. Thus, for the brick-wall filter at hand,

$$h(t) = \frac{1}{2\pi} \int_{-\omega_o}^{\omega_o} K\, e^{\,\jmath\omega t}\, d\omega = \frac{K}{\pi t} \sin\omega_o t\,. \tag{13.3}$$

Figure 13.2 shows a representative plot of this result.

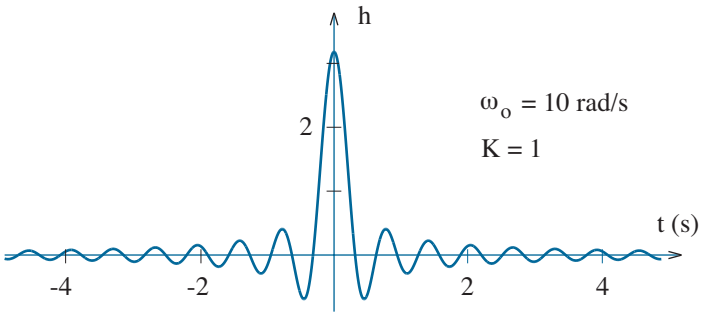


Figure 13.2: Ideal low-pass impulse response (Eq. 13.3).

The impulse excitation commences at $t = 0$. So the response of Fig. 13.2 should give rise to some distress since it exhibits time dependence for $t < 0$. Absence of **causality** shows the impossibility of an ideal filter characteristic —one must be content to realize *approximate* behavior.

By convention, the filter transition point now implies a half-power index, an angular frequency where the magnitude response is diminished by $1/\sqrt{2}$. It is also common practice to express magnitudes in dB (decibels) through the relation

$$|H|_{\,\mathrm{dB}} = 20 \log_{10} |H|\,. \tag{13.4}$$

A factor of $1/\sqrt{2}$ corresponds to -3 dB.

First-Order Filters

The first-order approximation to the ideal low-pass filter characteristic has a transfer function of the form

$$H(s) = \frac{K\,\omega_o}{s + \omega_o}\,. \tag{13.5}$$

At high frequencies ($\omega \gg \omega_o$), $|H| \approx K/\omega$. Thus, the magnitude response decreases by -20 dB/decade above the ω_o breakpoint as shown in Fig. 13.3. Note the phase shift of $-45°$ at the breakpoint (cutoff) angular frequency.

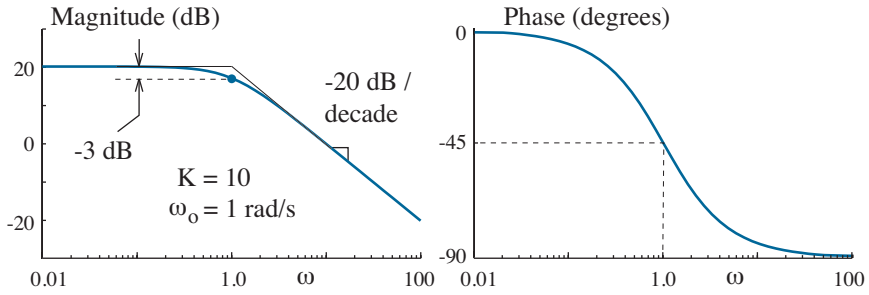


Figure 13.3: Magnitude/phase Bode plots for a first-order low-pass filter.

Figure 13.4 illustrates two first-order low-pass filter circuits. Option (a) is a passive low-pass filter ($K = 1$) followed by a non-inverting amplifier. Option (b) is an inverting amplifier with gain $K = -R/R_1$ in the dc limit where the capacitor functions as an open circuit. The parallel combination of R and C becomes a short circuit as frequency increases, and the gain is reduced accordingly.

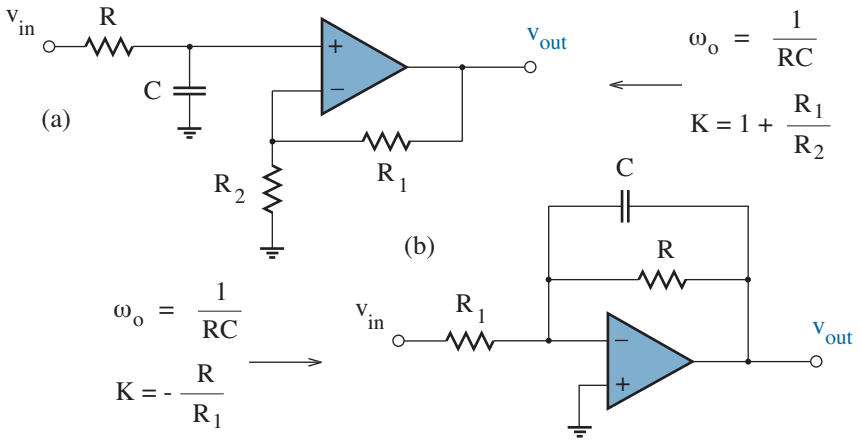


Figure 13.4: Circuit implementations for a first-order active low-pass filter.

Apart from the numerator of $H(s)$, the first-order approximation to the ideal high-pass filter characteristic is similar to that for the low-pass filter. Specifically,

$$H(s) = \frac{Ks}{s + \omega_o} \,. \tag{13.6}$$

At low frequencies ($\omega \gg \omega_o$), $|H| \approx K\omega$. Thus, the magnitude response increases by +20 dB/decade below the ω_o breakpoint as shown in Fig. 13.5. Note the phase shift of +45° at the breakpoint (cutoff) angular frequency.

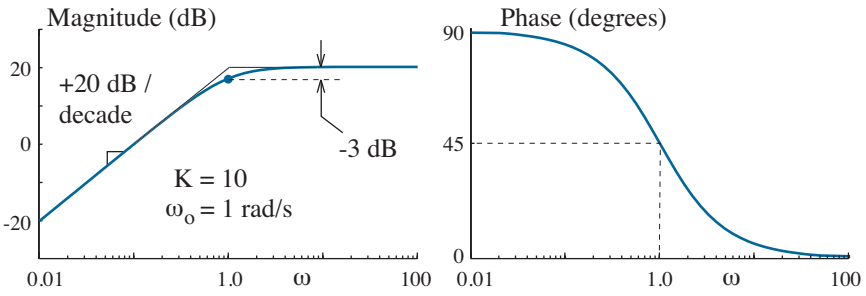


Figure 13.5: Magnitude/phase Bode plots for a first-order high-pass filter.

Figure 13.6 illustrates two first-order high-pass filter circuits. Option (a) is a passive high-pass filter ($K = 1$) followed by a non-inverting amplifier. Option (b) is an inverting amplifier with gain $K = -R_1/R$ in the high-frequency limit where the capacitor functions as a short circuit. The series combination of R and C becomes an open circuit as frequency decreases, and the gain is diminished.

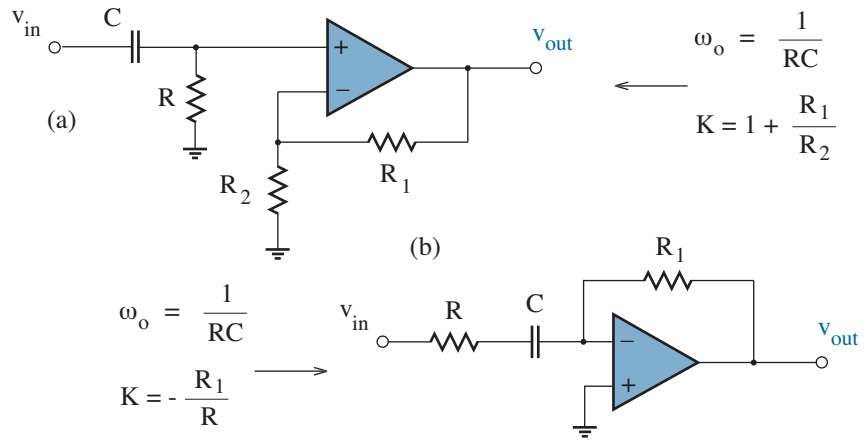


Figure 13.6: Circuit implementations for a first-order active high-pass filter.

Example 13.1

Design a first-order active low-pass filter with a response of +12 dB in the dc limit and -25 dB at 10 kHz.

Solution

We adopt a circuit design of the form in Fig. 13.4a. The K (gain) factor is

$$K = 10^{12/20} = 3.98\,.$$

So with $K = 1 + R_1/R_2$, we are content with $R_1 = 3$ kΩ and $R_2 = 1$ kΩ. The filter response at 10 kHz is $12-(-25) = 37$ dB below that at the cutoff frequency f_o, and the characteristic decline is -20 dB/decade. Thus,

$$f_o = 10\text{ kHz} \times 10^{-37/20} = 141\text{ Hz}\,.$$

Then with $f_o = \omega_o/2\pi = 1/2\pi RC$ and arbitrary $C = 1$ μF,

$$R = \frac{1}{2\pi \times 141 \times 1 \times 10^{-6}} = 1.1\text{ k}\Omega\,.$$

Figure 13.7 shows the completed design.

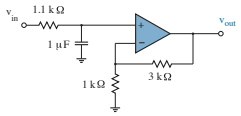


Figure 13.7: Circuit design for Example 13.1.

Exercise 13.1 Design a first-order active low-pass filter with a response of +28 dB in the dc limit and -12 dB at 6 kHz (Fig. 13.4b, $C = 220$ nF).

Ans: $R = 12$ kΩ, $R_1 = 480$ Ω

Exercise 13.2 Design a first-order active high-pass filter with +16 dB in the high-frequency limit and -30 dB at 100 Hz (Fig. 13.6b, $C = 4.7$ nF).

Ans: $R = 1.7$ kΩ, $R_1 = 11$ kΩ

Example 13.2

You require a high-pass filter with 10-kHz cutoff and 20-dB passband gain. A member of your design team has graciously taken time from monitoring personal stock options to create an "awesome" first-order circuit (Fig. 13.8). Use SPICE to check the design.

Note: Your company has not been doing well lately.

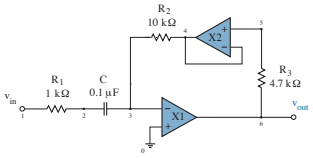


Figure 13.8: Circuit for Example 13.2.

Solution

An appropriate SPICE code takes the following form:

```
* High-Pass Filter (Revision #7)

Vin     1     0     ac    1m
R1      1     2     1k
C1      2     3     0.1u
X1      6     0     0     3     OpAmp
R2      3     4     10k
X2      4     0     5     4     OpAmp
R3      5     6     4.7k
```

```
.ac         dec     50      100     100k
.probe

.subckt     OpAmp           out     com     in_p    in_n
Rin         in_p    in_n    1000MEG
Eout        out     com     in_p    in_n    1E6
.ends

.end
```

As in Chapter 7, the input signal source has *arbitrary* 1-mV amplitude. After a simulation, the various circuit nodes have the character of phasors with particular amplitude and phase. Thus, the ratio of the amplitudes at nodes 6 and 1 is the transfer function v_{out}/v_{in}. If the v_{in} amplitude had been 100 V, SPICE would give the same ac ratio despite obvious concerns. Power supply limits do not apply to ac SPICE simulations.

The companion .ac command sweeps v_{in} by decades (dec) from 100 Hz to 100 kHz with 50 data points per decade. An alternative linear (lin) sweep is inappropriate if the frequency range spans several orders of magnitude—the data points will have unequal spacing on a logarithmic scale.

Figure 13.9 shows the .probe results. Since we are working with decibels, we express the v_{out}/v_{in} ratio as VDB(6) - VDB(1).

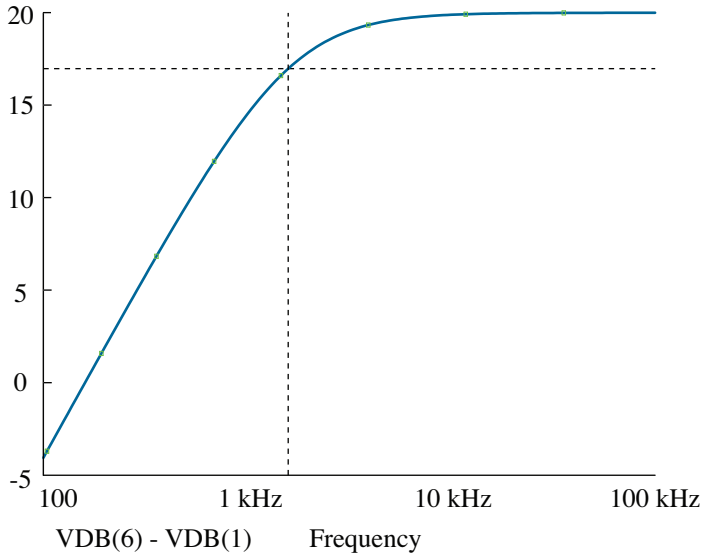


Figure 13.9: .probe results for the simulation of Example 13.2.

The passband gain is indeed 20 dB. To determine the cutoff frequency, we use the .probe cursor to find the frequency at which the filter response has diminished by 3 dB in relation to the passband. Here, the 17-dB point occurs at 1.59 kHz, which is a factor of 2π lower than the design objective. So it appears that a common calculation error strikes again: $\omega_o \neq f_o$.

As a manager, you would likely have found the R_1C error by inspection. And you would not be impressed by the X2 op-amp and 4.7-kΩ resistor, which add cost without benefit. Perhaps there is a particular set of stock options that should not be allowed to vest.

Second-Order Filters

Second-order filter approximations feature transfer functions of the form

$$H(s) = \frac{K\,N}{s^2 + s\,\omega_o/Q + {\omega_o}^2}\,, \tag{13.7}$$

where N is one (or a combination) of the three terms in the denominator. Parameter Q is called the **quality factor**—of more later. The options—

Low-Pass:

$$H(s) = \frac{K\,{\omega_o}^2}{s^2 + s\,\omega_o/Q + {\omega_o}^2} \tag{13.8}$$

Figure 13.10 shows the magnitude and phase characteristics. Note the -90° phase shift at ω_o where $|H| = KQ$. (This is the -3 dB point for $Q = 1/\sqrt{2}$).

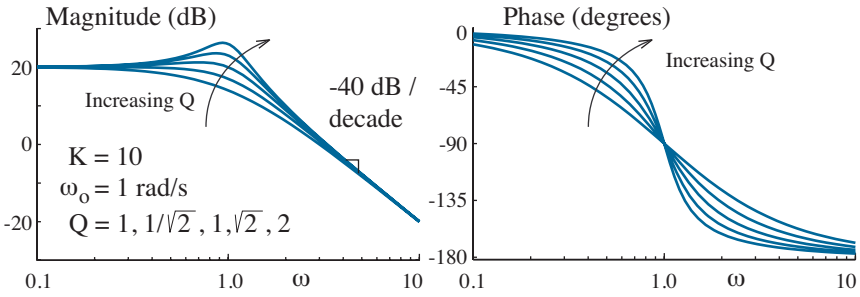


Figure 13.10: Second-order low-pass magnitude/phase Bode plots.

High-Pass:

$$H(s) = \frac{K\,s^2}{s^2 + s\,\omega_o/Q + {\omega_o}^2} \tag{13.9}$$

Figure 13.11 shows the magnitude and phase characteristics. Note the +90° phase shift at ω_o where $|H| = KQ$. (This is the -3 dB point for $Q = 1/\sqrt{2}$).

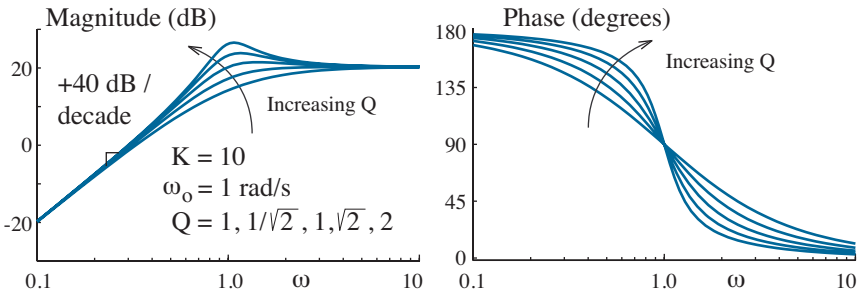


Figure 13.11: Second-order high-pass magnitude/phase Bode plots.

Bandpass:

$$H(s) = \frac{K\, s\, \omega_o/Q}{s^2 + s\, \omega_o/Q + {\omega_o}^2} \tag{13.10}$$

Figure 13.12 shows the magnitude and phase characteristics. Note the zero phase shift at ω_o where $|H| = K$. The roll-off is ± 20 dB/decade for $Q = 1$.

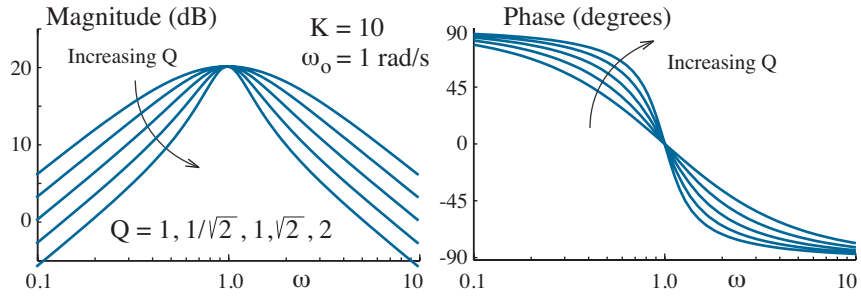


Figure 13.12: Second-order bandpass magnitude/phase Bode plots.

Bandstop:

$$H(s) = \frac{K\, (s^2 + {\omega_o}^2)}{s^2 + s\, \omega_o/Q + {\omega_o}^2} \tag{13.11}$$

This is simply the bandpass characteristic subtracted from constant K.

Caution: The preceding bandpass and bandstop transfer characteristics typically apply to highly selective **narrowband** filters with very large Q. **Wideband** filters with significant bandwidth require a cascade of low- and high-pass filters (see Problem 13.10).

By now it should be apparent that ω_o is actually a *resonant* angular frequency at which the denominator is purely imaginary for any one of the preceding transfer characteristics. The quality factor Q is also associated with resonance—the system poles are a complex conjugate pair for $Q > 1/2$, and large Q promotes minimal amplitude decay at the resonant frequency. When designing second-order low- and high-pass filters, choosing $Q = 1/\sqrt{2}$ ensures maximum "flatness" without peaking within the passband.

Different considerations determine the choice of Q for a second-order bandpass filter. The two -3-dB points occur at

$$\omega = \omega_o \left(\sqrt{1 + \frac{1}{4Q^2}} \pm \frac{1}{2Q} \right) \tag{13.12}$$

(see Problem 13.11). Thus, the filter bandwidth is given by

$$B = \frac{\omega_o}{Q}\,. \tag{13.13}$$

Equation 13.13 establishes Q for most designs.

We defer a discussion of second-order filter circuits to Section 13.2.

Higher-Order Filters

Higher-order filters are nothing more than first- and second-order cascades. For example, a filter of order $n = 7$ is realized with one first-order and three second-order circuits in succession. However, the overall filter performance is critically dependent upon the ω_o and Q values for the individual stages. Three options concern us:

- **Butterworth** filters have maximally flat magnitude in the passband. In the case of a low-pass filter of order n, the first n derivatives of the magnitude characteristic all vanish in the dc limit.
- **Chebyshev** filters have sharper transitions than Butterworth filters, but at the expense of ripples with equal variation in the passband. For example, 1-dB Chebyshev implies 1-dB (12.2-%) ripple behavior.
- **Bessel** filters have more gradual transitions than Butterworth filters, but with the advantage of maximally flat delay over the passband. This behavior is desirable for a "clean" transient response.

Tables 13.1, 13.2, and 13.3 provide $f_o{}'$ and Q values for Butterworth, 1-dB Chebyshev, and Bessel filters, respectively, with filter orders $n \leq 10$. Parameter $f_o{}'$ is a stage cutoff frequency in relation to $f_o = 1$ Hz overall. To apply the tables at $f_o = x$ Hz, one simply scales $f_o{}'$ by the factor x. Circuit implementations require $\omega_o{}' = 2\pi f_o{}'$.

Figure 13.13 shows even-order low-pass filter stopbands with $f_o = 1$ Hz —we omit the odd-order curves to avoid clutter. Thus, a filter that achieves -60 dB at 5 Hz has five Butterworth stages, three 1-dB Chebyshev stages, or seven Bessel stages. Let α represent the difference in dB between the passband maximum and a desired stopband minimum at f_{min}. If this were a text on circuit theory, we could derive

$$n > \frac{\ln \sqrt{10^{\,\alpha/10} - 1}}{\ln \left(f_{min}/f_o \right)} \tag{13.14}$$

for Butterworth filters, and

$$n > \frac{\cosh^{-1} \sqrt{10^{\,\alpha/10} - 1}}{\cosh^{-1} \left(f_{min}/f_o \right)} \tag{13.15}$$

for Chebyshev filters. No such relation applies to Bessel filters, so one must estimate n from the characteristics of Fig. 13.13.

Figure 13.14 shows even-order low-pass unit-step response ($f_o = 1$ Hz). As advertised, the Bessel filter exhibits the least amount of delay with negligible overshoot. The 1-dB Chebyshev filter offers the worst behavior. Filters with unequal delays for different frequency components within the passband promote signal **distortion**, a topic for Section 13.4.

Filter Order	Stage 1		Stage 2		Stage 3		Stage 4		Stage 5	
	$f_o{}'$	Q	$f_o{}'$	Q	$f_o{}'$	Q	$f_o{}'$	Q	$f_o{}'$	Q
2	1.0000	0.7071								
3	1.0000	1.0000	1.0000							
4	1.0000	0.5412	1.0000	1.3065						
5	1.0000	0.6180	1.0000	1.6181	1.0000					
6	1.0000	0.5177	1.0000	0.7071	1.0000	1.9320				
7	1.0000	0.5549	1.0000	0.8019	1.0000	2.2472	1.0000			
8	1.0000	0.5098	1.0000	0.6013	1.0000	0.8999	1.0000	2.5628		
9	1.0000	0.5321	1.0000	0.6527	1.0000	1.0000	1.0000	2.8802	1.0000	
10	1.0000	0.5062	1.0000	0.5612	1.0000	0.7071	1.0000	1.1013	1.0000	3.1969

Table 13.1: Butterworth filter specifications ($f_o = 1$ Hz).

Filter Order	Stage 1		Stage 2		Stage 3		Stage 4		Stage 5	
	$f_o{}'$	Q	$f_o{}'$	Q	$f_o{}'$	Q	$f_o{}'$	Q	$f_o{}'$	Q
2	1.0500	0.9565								
3	0.9971	2.1076	0.4942							
4	0.5286	0.7845	0.9932	3.5600						
5	0.6552	1.3988	0.9941	5.5538	0.2895					
6	0.3532	0.7608	0.7468	2.1977	0.9953	8.0012				
7	0.4800	1.2967	0.8084	3.1554	0.9963	10.901	0.2054			
8	0.2651	0.7530	0.5838	1.9564	0.8506	4.2661	0.9971	14.245		
9	0.3812	1.1964	0.6623	2.7119	0.8805	5.5239	0.9976	18.007	0.5193	
10	0.2121	0.7495	0.4760	1.8639	0.7214	3.5609	0.9024	6.9419	0.9981	22.278

Table 13.2: 1-dB Chebyshev filter specifications ($f_o = 1$ Hz).

Filter Order	Stage 1		Stage 2		Stage 3		Stage 4		Stage 5	
	$f_o{}'$	Q	$f_o{}'$	Q	$f_o{}'$	Q	$f_o{}'$	Q	$f_o{}'$	Q
2	1.2736	0.5773								
3	1.4524	0.6910	1.3270							
4	1.4192	0.5219	1.5912	0.8055						
5	1.5611	0.5635	1.7607	0.9165	1.5069					
6	1.6060	0.5103	1.6913	0.6112	1.9071	1.0234				
7	1.7174	0.5324	1.8235	0.6608	2.0507	1.1262	1.6853			
8	1.7837	0.5060	1.8376	0.5596	1.9591	0.7109	2.1953	1.2259		
9	1.8794	0.5197	1.9488	0.5894	2.0815	0.7606	2.3235	1.3220	1.8575	
10	1.9490	0.5040	1.9870	0.5380	2.0680	0.6200	2.2110	0.8100	2.4850	1.4150

Table 13.3: Bessel filter specifications ($f_o = 1$ Hz).

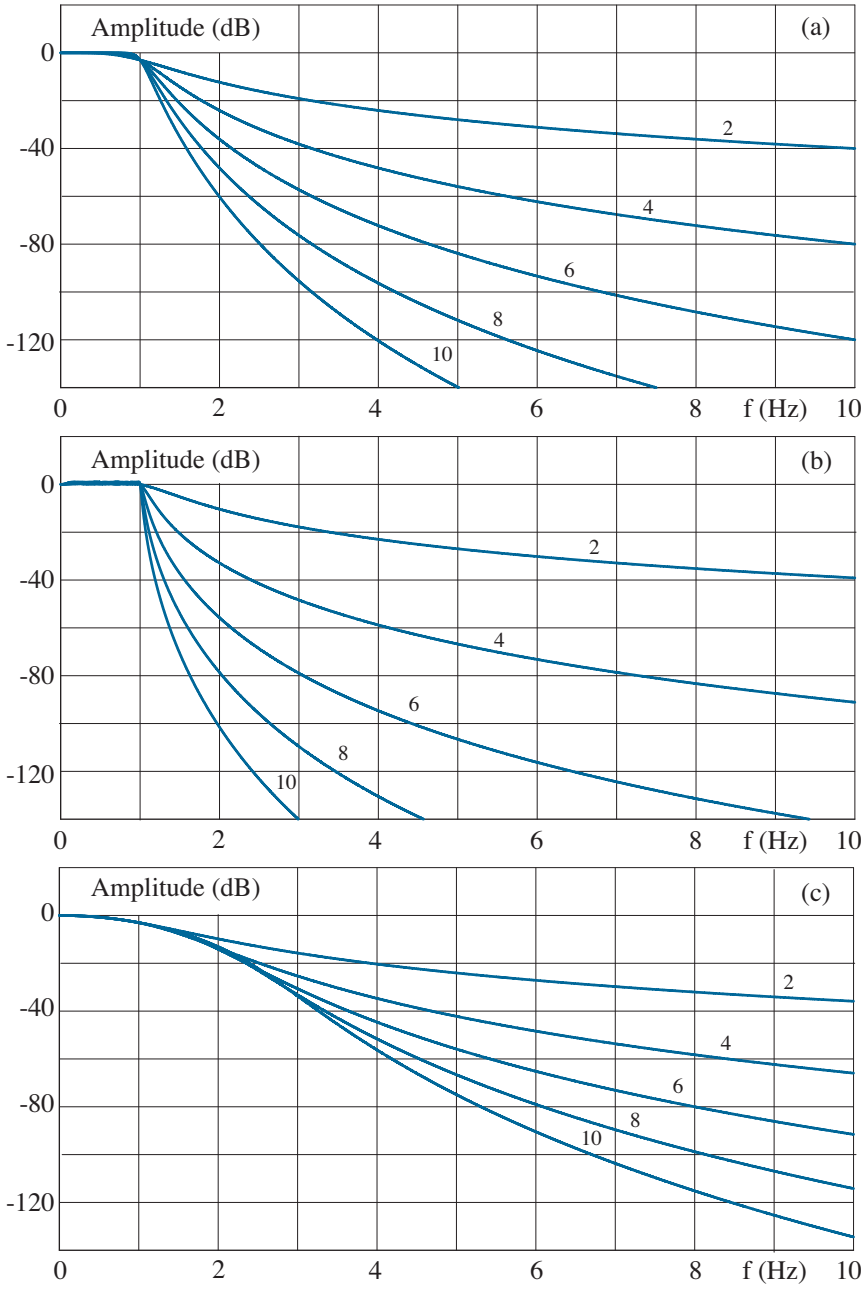


Figure 13.13: Low-pass stopband behavior with filter order (f_o = 1 Hz): (a) Butterworth; (b) 1-dB Chebyshev; (c) Bessel.

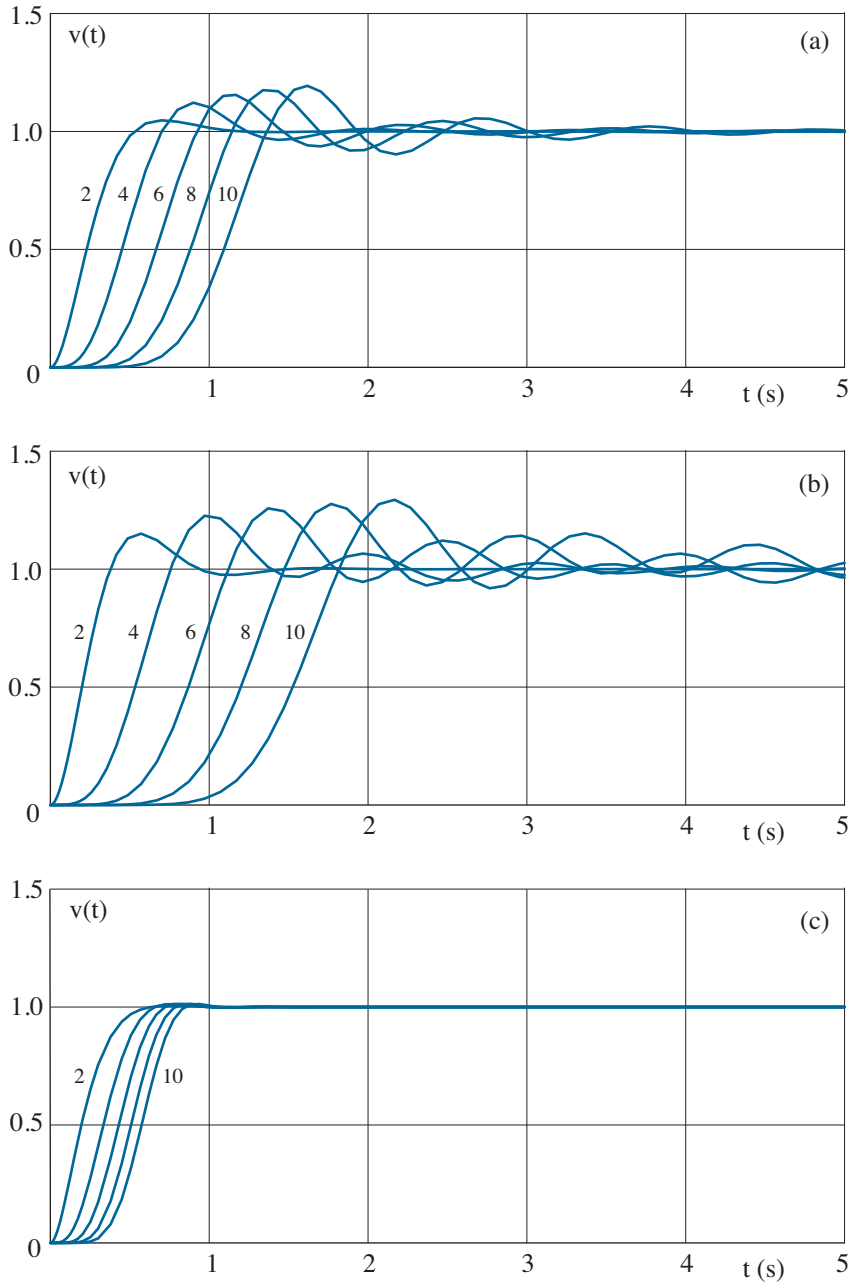


Figure 13.14: Low-pass unit-step responses with filter order (f_o = 1 Hz): (a) Butterworth; (b) 1-dB Chebyshev; (c) Bessel.

Example 13.3

Design an active low-pass Butterworth filter with +8 dB in the dc limit, -72 dB at 40 kHz, and $f_o = 8$ kHz.

Solution

The requisite decline in the stopband is $\alpha = 8 - (-72) = 80$ dB. Then with $f_{min} = 40$ kHz, we apply Eq. 13.14 to find

$$n > \frac{\ln \sqrt{10^8 - 1}}{\ln (40/8)} = 5.76 \, .$$

The filter order n must be an integer, so $n = 6$. For an alternate procedure, we observe that a factor of 8000 normalizes f_o to 1 Hz and f_{min} to 5 Hz. The point on the graph of Fig. 13.13a that corresponds to -80 dB at 5 Hz is just above the line for which $n = 6$.

With $n = 6$, we need three second-order stages. As shown in Table 13.1, each stage requires $f_o{}' = 1.0000$. Thus, all applicable f_o values are 8 kHz ($\omega_o = 5.03 \times 10^4$ rad/s). The Q factors are 0.5177, 0.7071, and 1.9320. Finally, we note that 8 dB is a passband gain of 2.512. So we arbitrarily choose $K = \sqrt[3]{2.512} = 1.36$ for each stage.

Example 13.4

Design an active high-pass Bessel filter with +25 dB at high frequencies, -65 dB at 5 kHz, and $f_o = 50$ kHz.

Solution

It is straightforward to show that the substitution $s/\omega_o \rightarrow \omega_o/s$ transforms a high-pass filter characteristic into the low-pass form (see Problem 13.15). Whereas 50,000 normalizes f_o to 1 Hz and f_{min} to 0.1 Hz, the preceding transform yields an effective $f_{min} = 10$ Hz in relation to low-pass stopbands. There is no formula that determines the Bessel filter order. So we look to the Fig. 13.13c graph to find that a reduction of $-(25 - (-65)) = -90$ dB requires a filter with $n = 7$. From Table 13.3, the stage specifications are:

Stage 1	$f_o = 50$ kHz $/1.7174 = 29.11$ kHz	$Q = 0.5324$
Stage 2	$f_o = 50$ kHz $/1.8235 = 27.42$ kHz	$Q = 0.6608$
Stage 3	$f_o = 50$ kHz $/2.0507 = 24.38$ kHz	$Q = 1.1262$
Stage 4	$f_o = 50$ kHz $/1.6853 = 29.67$ kHz	

Note the downward frequency scaling (as opposed to upward for low pass). The 25-dB passband gain requires an overall $K = 17.8$.

13.2 Integrated Filters

Second-order filter implementations with one or two discrete op-amps and external resistor/capacitor connections have mostly been relegated to the dustbin of classic circuits used to promote one's intellectual development (see Problems 13.23 - 13.36). Contemporary filter designs take advantage of full integration with limited concern over the number of active components. In what follows, we demonstrate how integrator (antiderivative) circuits are used as building blocks for integrated (black-box) analog conditioning.

Signal Flow Graphs

Our development of a second-order fully-integrated filter will require the management of several intermediate signals and operations, so we introduce the concept of a **signal flow graph** to keep track of electrical developments. Figure 13.15a shows circular nodes representing separate signals x_1 and x_2. The arrow on the line joining the nodes shows the direction of signal flow, and the adjoining operational expression indicates how an input signal (x_1) is modified as it contributes to an output signal (x_2). In this case, $x_2 = ax_1$. Figure 13.15b shows that the signal at any node derives from the sum of the signals entering that node: $x_2 = ax_1 + bx_1 = (a+b)x_1$. Figure 13.15c shows the result of cascaded operations: $x_2 = ax_1$ and $x_3 = bx_2$, so $x_3 = abx_1$. Here, the quantity ab is a **direct path gain** from x_1 to x_3.

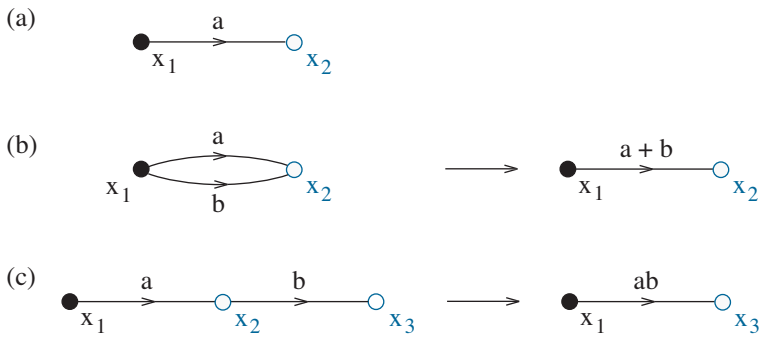


Figure 13.15: Elementary signal flow graphs.

An input node has only outgoing signals, and it is shaded for distinction. Strictly speaking, an output node has only incoming signals. Nevertheless, any node that is not an input can be viewed as a particular form of output. Note that signals can change attributes from voltage to current to whatever (as in the case of an electromechanical system). The operational expressions necessarily have dimensions consistent with indicated transformations.

Exercise 13.3 Find the relationship between output x_4 and input x_1 that is established in the signal flow graph of Fig. 13.16.

Ans: $x_4 = (abc + ad + ec)\, x_1$

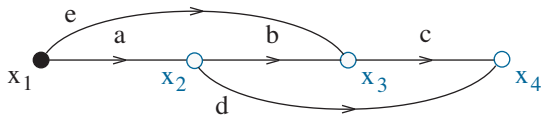


Figure 13.16: Signal flow graph for Exercise 13.3.

Signal flow graphs are a bit more complicated when feedback is present. For example, the graph of Fig. 13.17a has a direct path between x_1 and x_3, but there is feedback from x_3 to the intermediate node x_2. Here, we write $x_2 = ax_1 + cx_3$ and $x_3 = bx_2$. Then we solve for x_3 to find

$$x_3 = \frac{ab\, x_1}{1 - bc}\,. \tag{13.16}$$

The quantity bc appearing in the denominator of Eq. 13.16 is a **loop gain**. Figure 13.17b shows a similar signal flow graph; however, the feedback loop coincides with a direct path between x_1 and output x_4. Whereas, $x_4 = dx_3$, we now have a system of three equations that reduce to

$$x_4 = \frac{abd\, x_1}{1 - bc}\,. \tag{13.17}$$

In Eqs. 13.16 and 13.17, the factor that multiplies input x_1 in the numerator is the direct path gain from input to output, so a pattern appears to emerge.

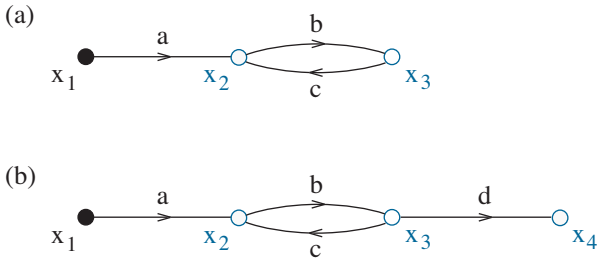


Figure 13.17: Signal flow graphs with feedback.

Exercise 13.4 Find the relationship between output x_3 and input x_1 that is established in the signal flow graph of Fig. 13.18.

Ans: $x_3 = \dfrac{ab\,x_1}{1 - bd - c}$

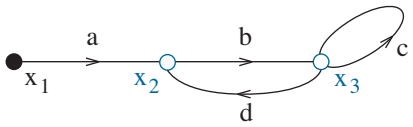


Figure 13.18: Signal flow graph for Exercise 13.4.

It is sometimes convenient to apply a restricted form of **Mason's rule**[1] to relate graphical nodes. If all feedback loops touch every forward path, and no more than two loops fail to touch one another, the relation is

$$T = \frac{\Sigma\, P_{nm}}{1 - \Sigma\, L(1) + \Sigma\, L(2)}\,. \tag{13.18}$$

In this expression, P_{nm} is a direct path gain between nodes n and m, $L(1)$ is a first-order (individual) loop gain, and $L(2)$ is a second-order product of two loop gains that do not touch. Note the applicability to Exercise 13.4.

Exercise 13.5 Use Mason's rule to find the relationship between output x_6 and input x_1 that is established in the signal flow graph of Fig. 13.19.

Ans: $\dfrac{x_6}{x_1} = \dfrac{(2/s)\,s + 2}{1 - (-2/s) - (-2s) + (-2/s)(-2s)} = \dfrac{4s}{2s^2 + 5s + 2}$

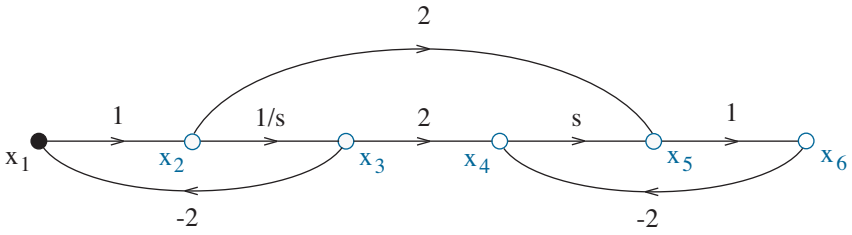


Figure 13.19: Signal flow graph for Exercise 13.5.

[1] For a treatment with fewer restrictions, see S. J. Mason and H. J. Zimmermann, *Electronic Circuits, Signals, and Systems*, John Wiley & Sons, New York, 1960.

Now consider a second-order low-pass filter that produces output voltage v_a in response to input voltage v_{in} while subject to a gain factor of $-K$. Our design proceeds in seven steps:

Step 1 - Characterize a "building-block" integrator circuit.

For the moment, the integrator circuit most easily obtained consists of an op-amp, a capacitor, and a resistor as in Fig. 13.20. The input-output relationship is

$$v_{out}(t) = \frac{-1}{RC}\int_{-\infty}^{t} v_{in}(t')\,dt' , \tag{13.19}$$

which you were asked to derive in Problem 1.39. In the s domain,

$$v_{out}(s) = \left(\frac{-\omega_o}{s}\right) v_{in}(s) , \tag{13.20}$$

where $\omega_o = 1/RC$. Thus, $(-\omega_o/s)$ is the integration operator at hand.

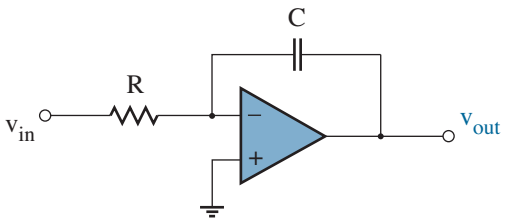


Figure 13.20: Op-amp "building-block" integrator circuit.

Step 2 - Express the transfer function in terms of the integration operator.

The low-pass transfer characteristic of Eq. 13.8 takes the form ($K < 0$)

$$H_a(s) = \frac{v_a}{v_{in}} = \frac{-K\left(\dfrac{-\omega_o}{s}\right)^2}{1 + \left(\dfrac{-1}{Q}\right)\left(\dfrac{-\omega_o}{s}\right) + \left(\dfrac{-\omega_o}{s}\right)^2} . \tag{13.21}$$

Step 3 - Express the output as the sum of input and feedback operations.

We cross multiply Eq. 13.21 and rearrange to find

$$v_a = -K\left(\frac{-\omega_o}{s}\right)^2 v_{in} - \underbrace{\left(\frac{-\omega_o}{s}\right)^2 v_a}_{\text{Feedback term 1}} - \underbrace{\left(\frac{-1}{Q}\right)\left(\frac{-\omega_o}{s}\right) v_a}_{\text{Feedback term 2}} . \tag{13.22}$$

Step 4 - Construct a signal flow graph.

We construct a signal flow graph by treating the right-hand-side terms of Eq. 13.19 one by one as shown in Fig. 13.21. The first iteration (Fig. 13.21a) establishes a direct signal path from input to output. In the process at hand, input v_{in} experiences a sequence of three operations: multiplication by $-K$, and two integrations. Thus, we require the intermediate voltages v_b and v_c. The $-K$ multiplication is not used again, so it is first. The second iteration (Fig. 13.21b) tackles feedback term 1. Here, v_a undergoes multiplication by -1 followed by two integrations. The latter already occur along the signal path from v_c to v_a, so we direct the feedback to v_c with a -1 scale factor. The third iteration (Fig. 13.21c) treats feedback term 2, but with difficulty. The straightforward approach is to scale v_a by $(1/Q)$ while feeding back to v_b for a single integration in the return path. However, when working with op-amp circuits, it is convenient to require that the summation processes at any node maintain the same operational character (see Problem 13.40). The alternative approach in which a scaled v_a feeds back to v_c fails because of the two integrations in the return path. But $v_b = (-s/\omega_o)v_a$. In turn, feedback from v_b to v_c yields $(-\omega_o/s)^2(1/Q)(-s/\omega_o)v_a = (1/Q)(-\omega_o/s)v_a$. Nevertheless, we will also find it convenient to split the $1/Q$ multiplication into two stages $(-1/Q \times -1)$, so we create another intermediate voltage v_d. *The completed signal flow graph is a roadmap for circuit implementation.*

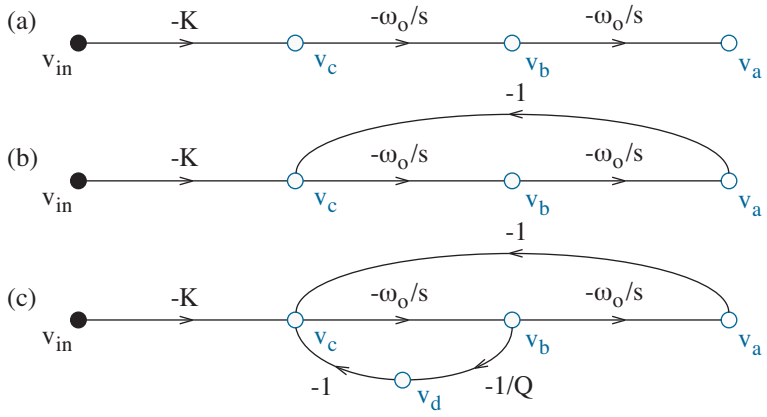


Figure 13.21: Iterative signal flow graphs for a second-order low-pass filter.

We pause to note that Fig. 13.21c can feature v_b or v_c as output nodes. The applicable analysis reveals a happy outcome: v_b reflects the bandpass characteristic (scaled by $-Q$), and v_c reflects the high-pass characteristic. Obtaining all three of the major filter characteristics from one circuit is a property of a **state-variable filter**.

Step 5 - Implement circuits for the signal flow graph without feedback.

Figure 13.22 shows op-amp circuits that implement the required signal-flow functions along all paths excluding those from v_a to v_c and v_d to v_c. Inverting amplifiers perform the $-K$ and $-1/Q$ multiplications.

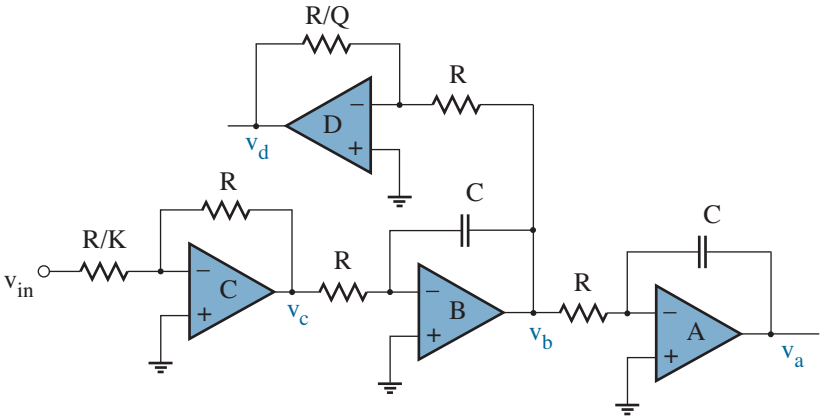


Figure 13.22: Signal-flow-graph circuits for Fig. 13.21c without feedback.

Step 6 - Add feedback.

Figure 13.23 illustrates feedback through current summation at node X, a virtual ground. Two resistors establish v_a- and v_d-proportionate currents. We leave it as an exercise to prove that $v_c = -Kv_{in} - v_a - v_d$ as required.

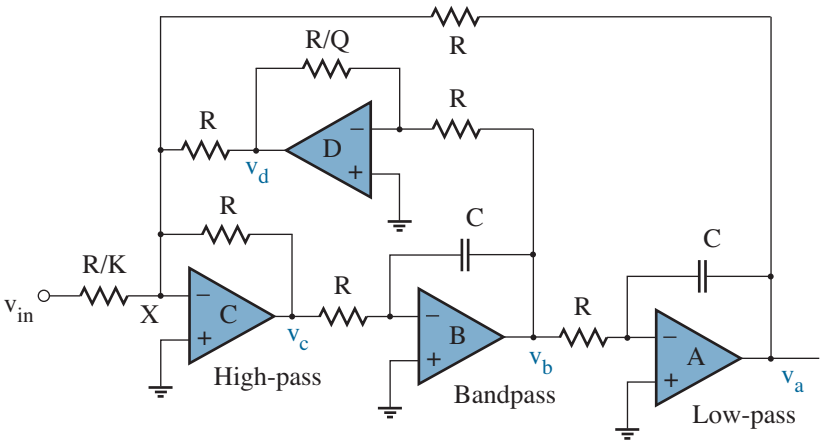


Figure 13.23: Four-op-amp state variable filter.

Step 7 - Look for circuit efficiencies.

When we constructed the signal flow graph in Step 4, we split the $1/Q$ multiplication into two stages ($-1/Q \times -1$) because one of us anticipated a simple op-amp circuit for the former and a simple negative feedback process that effected the latter. (The same feedback process was used with node v_a.) An option eliminates op-amp D in Fig. 13.23 in favor of feedback to the non-inverting input of op-amp C. Accordingly, define v_x as that terminal voltage. The inverting input is forced to have the same value, so we have

$$\frac{v_{in} - v_x}{R/K} + \frac{v_a - v_x}{R} + \frac{v_c - v_x}{R} = 0\,. \tag{13.23}$$

In turn,

$$v_c = -Kv_{in} - v_a + (K+2)v_x\,. \tag{13.24}$$

Equation 13.24 conforms with the signal flow graph of Fig. 13.21c if the last right-hand term is replaced with v_b/Q. The circuit substitution is achieved with the help of a voltage divider between v_b and ground as in Fig. 13.24. Then with $v_x = v_b/Q(K+2) = v_b\,R/(R+R_1)$, resistor R_1 has the value

$$R_1 = [\,Q(K+2) - 1\,]\,R\,. \tag{13.25}$$

Thus, a change in K affects the circuit value for Q when R_1 is constant. Note that the subcircuit with op-amp C is a differential amplifier (Fig. 1.7). With acquired design experience, one might have anticipated this option.

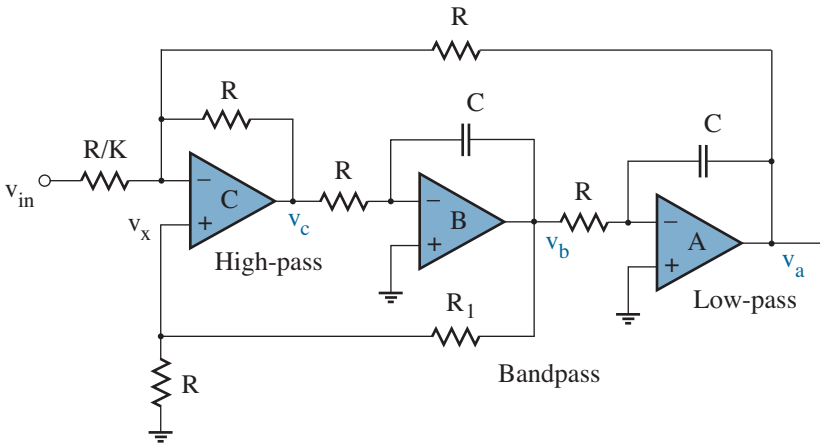


Figure 13.24: Three-op-amp state variable filter.

The UAF42 (Texas Instruments, formerly Burr-Brown) is an integrated circuit with the state-variable filter design of Fig. 13.24 (see Problem 13.42). The MAX274 has a modified design (see Problems 13.43 and 13.44).

Alternative Integrators

One of the recurring themes in integrated as opposed to discrete-component electronics is the avoidance of resistors that consume large physical area. Figure 13.25a shows one popular method. Suppose switch S_1 is closed with S_2 open. The capacitor voltage is V_1 and the stored charge is $Q_1 = CV_1$. Now open S_1 and close S_2. The capacitor voltage changes to V_2 and the stored charge changes to $Q_2 = CV_2$. Thus,

$$\Delta Q = Q_1 - Q_2 = C(V_1 - V_2)\,. \tag{13.26}$$

The charge transfer process occurs over a time interval $\Delta t = 1/f$, where f is the clock frequency that controls the switch action. In turn, the current delivered to voltage source V_2 at the conclusion of one clock cycle is

$$i = \frac{\Delta Q}{\Delta t} = Cf(V_1 - V_2)\,. \tag{13.27}$$

The comparable (but continuous) current in the resistive circuit of Fig. 13.25b is $i = (V_1 - V_2)/R$. So the switched capacitor has an effective $R = 1/Cf$.

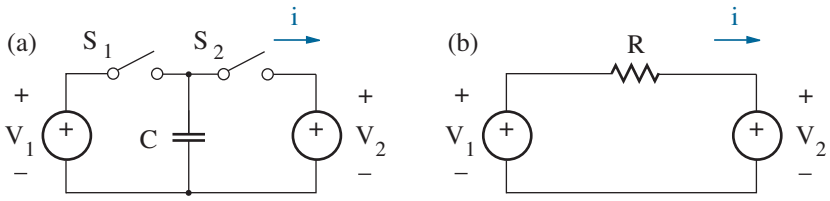


Figure 13.25: Switched capacitor circuit and resistor equivalent. Switch S_1 is closed when S_2 is open, and conversely.

Figure 13.26 shows a switched capacitor integrator circuit. With the help of Eq. 13.18 it is easy to deduce

$$v_{out}\,(t) = -\left(\frac{C_1}{C_2}\right) f \int_{-\infty}^{t} v_{in}\,(t')\,dt'\,. \tag{13.28}$$

Thus, $-(C_1 f/C_2 s)$ is the characteristic integration operator.

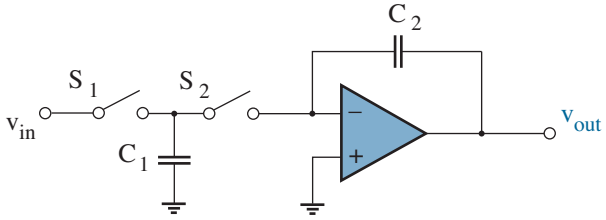


Figure 13.26: Switched capacitor integrator.

In practice, switched capacitor integrators become complicated by the requirement of two additional switches to mitigate the parasitic capacitance that otherwise contributes to an inaccurate value for C (see Problem 13.48). Moreover, switched capacitor circuits function as **discrete-time systems** and are subject to discrete-time (z-domain) analysis, especially when signal frequencies of interest approach the clock frequency for switching action. On a more favorable note, switched capacitor integrators can be compact since the characteristic integration operator has a *ratio* of capacitor values that individually scale by the same factor as area decreases.

Another more modern method of signal integration that avoids switches as well as resistors features the **transconductor** of Fig. 13.27a for which

$$i_{out}(t) = G_m(v^+ - v^-)\,. \tag{13.29}$$

Parameter G_m is a **transconductance**. The symbol for the transconductor resembles that of an op-amp, but the output end of the triangular symbol is truncated. It is easy to show that a resistive load leads to the relation $v_{out} = G_m R\,(v^+ - v^-)$, the same as that for an op-amp with $A_{vd} = G_m R$. Figure 13.27b shows a transconductor and capacitor in use as an integrator. Here, we have

$$i_{out} = G_m(v_{in} - 0) = C\,\frac{dv_{out}}{dt}\,. \tag{13.30}$$

In turn,

$$v_{out}(t) = \frac{G_m}{C}\int_{-\infty}^{t} v_{in}(t')\,dt'\,. \tag{13.31}$$

Thus, (G_m/Cs) is the characteristic integration operator. So-called G_m-C filters are **continuous time systems** that are receptive to compactness but relatively prone to tuning errors since G_m is difficult to control.

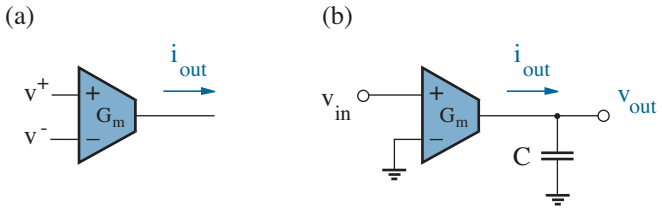


Figure 13.27: Transconductor and "building-block" integrator.

Design procedures for first-order switched-capacitor and G_m-C filters are examined in Problems 13.49 - 13.52. Second-order design procedures follow the signal-flow-graph and integrator methods developed previously. Nevertheless, we place them beyond the expectations of this text.

13.3 Digital Conditioning for Serial Links

This section examines digital signal conditioning processes and protocols that promote effective communications over serial links. Special attributes of information exchange in corruptive environments emerge shortly.

Figure 13.28 shows a **parallel communication** process that transfers an n-bit digital "word" between registers in blocks 1 and 2. With n links—one for each bit—there is nearly simultaneous information exchange after an appropriately designated clock pulse. Thus, the simple parallel process is commonly used over short distances within an integrated circuit or board. Parallel communication over long distances tends be expensive, requiring the overhead of multiple lines and elaborate connectors. Seasoned readers will recall SCSI (Small Computer System Interface) "Scuzzy" connections between their personal computer and a peripheral printer or hard drive. Such cumbersome parallel connections have gradually become less common.

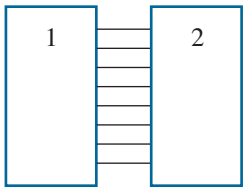


Figure 13.28: Parallel (8-bit) data communication process.

Figure 13.29 shows a **serial communication** process with a single link between transducers that unpack, exchange, and repack n-bit words over the course of n or more clock cycles. While relatively slow and complex, the serial process is usually cost effective when long distances are involved. In what follows, we restrict attention to **asynchronous** processes in which the clients at each end of the communication link maintain separate clocks with an agreed-upon frequency. **Synchronous** processes are also available, but they require a separate dedicated link for the shared clock.

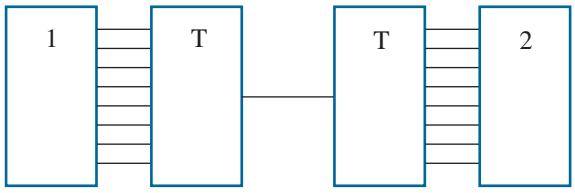


Figure 13.29: Serial (8-bit) data communication process.

The UART

A UART (*you*-art), short for Universal Asynchronous Receiver/Transmitter, is an integrated circuit responsible for data manipulation at each end of a serial link. When presented with 8- or 16-bit words, a UART prepares bit sequences for transmission by **framing** 8-bit segments or **bytes**:

- Begin with a **start bit**. By convention, idle communication channels are assumed to be logically HIGH. Thus, the start bit is always LOW. The commencement of a HIGH-to-LOW transition at the receiving end of a serial link reflects the arrival of a bit sequence.
- Add eight **data bits**, beginning with the least significant bit (LSB). The data can reflect numerical information such as that from a sensor, a memory address, or code for a displayed character. Table 13.4 shows ASCII codes for alphanumeric symbols common in English messages. This is a subset of the ubiquitous UTF-8 Unicode character set.
- Add an optional **parity bit** for error-checking received data.
  - E - Even parity — LOW for an even number of bits
  - O - Odd parity — LOW for an odd number of bits
  - N - No parity bit
- End with one or two **stop bits** (always HIGH). A single stop bit is generally sufficient for contemporary data links.

Figure 13.30 shows an example of the bit sequence for the character string S (53h) C (43h) for an E-8-1 framing process: even parity, eight data bits, one stop bit. With LSB ordering, the data bits appear reversed.

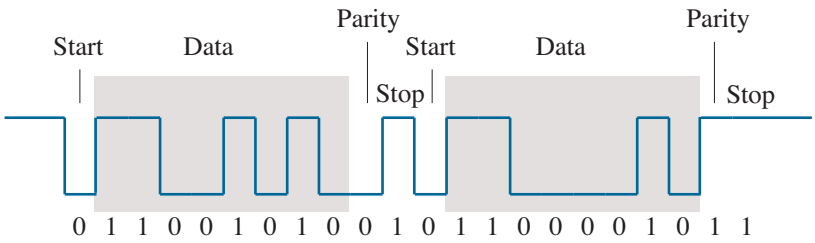


Figure 13.30: Bit sequence for the string SC (E-8-1 framing format).

Exercise 13.6 Specify the bit transmission sequence for the character string j2 in an O-8-2 framing process.

Ans: 0 0 1 0 1 0 1 1 0 1 1 0 0 0 1 0 1 1 0 0 0 1

Dec.	Hex	Char.	Dec.	Hex	Char.	Dec.	Hex	Char.	Dec.	Hex	Char.
00	00	End of String	32	20	Space	64	40	@	96	60	‘
01	01		33	21	!	65	41	A	97	61	a
02	02		34	22	"	66	42	B	98	62	b
03	03		35	23	#	67	43	C	99	63	c
04	04		36	24	$	68	44	D	100	64	d
05	05		37	25	%	69	45	E	101	65	e
06	06		38	26	&	70	46	F	102	66	f
07	07	Bell	39	27	’	71	47	G	103	67	g
08	08	Backspace	40	28	(	72	48	H	104	68	h
09	09	Tab	41	29	)	73	49	I	105	69	i
10	0A	Line Feed	42	2A	*	74	4A	J	106	6A	j
11	0B		43	2B	+	75	4B	K	107	6B	k
12	0C	Form Feed	44	2C	,	76	4C	L	108	6C	l
13	0D	Return	45	2D	-	77	4D	M	109	6D	m
14	0E		46	2E	.	78	4E	N	110	6E	n
15	0F		47	2F	/	79	4F	O	111	6F	o
16	10		48	30	0	80	50	P	112	70	p
17	11	XON	49	31	1	81	51	Q	113	71	q
18	12		50	32	2	82	52	R	114	72	r
19	13	XOFF	51	33	3	83	53	S	115	73	s
20	14		52	34	4	84	54	T	116	74	t
21	15		53	35	5	85	55	U	117	75	u
22	16		54	36	6	86	56	V	118	76	v
23	17		55	37	7	87	57	W	119	77	w
24	18		56	38	8	88	58	X	120	78	x
25	19		57	39	9	89	59	Y	121	79	y
26	1A		58	3A	:	90	5A	Z	122	7A	z
27	1B	Escape	59	3B	;	91	5B	[	123	7B	{
28	1C		60	3C	<	92	5C	\	124	7C	|
29	1D		61	3D	=	93	5D	]	125	7D	}
30	1E		62	3E	>	94	5E	^	126	7E	~
31	1F		63	3F	?	95	5F	_	127	7F	Delete

Table 13.4: ASCII character codes. The table excludes many 0-31 codes, which are mostly control characters for outdated teletype equipment.

In the early days of serial data communications, E-7-1 framing was often used since ASCII data require only seven bits. The need for larger character sets such as UTF-8 Unicode to accommodate complex and especially non-English messages has tended to promote N-8-1 framing. Another factor that supports the N-8-1 format is the feeble benefit of parity-bit error detection. Modern error-*correcting* codes are far more reliable, but at the expense of additional bytes of transmitted data. A nearly fail-safe option is to transmit 8-bit data twice and compare the results. While clearly slow and inefficient, this is acceptable for short bursts of data, as with infrared remote controls.

In the absence of a shared clock, the UART on the receiving end of a serial link has to determine appropriate times for sampling incoming bits. To do this, the UART maintains a clock that operates at 16 times the clock frequency for bit transmission. When the UART senses a HIGH-to-LOW transition that reflects the arrival of the start bit, an internal 4-bit counter is set to zero, and counting proceeds at a pace governed by the 16X clock. The detected bit state is that for the majority of three samples taken at counts seven, eight, and nine (roughly the halfway point for bit duration). Subsequent sampling points are easily determined as the counter rolls over to zero with every bit transition (see Fig. 13.31). When the complete bit sequence has been detected, a byte of data is recovered according to the rules of a particular frame format. Some UARTs can perform a parity check and signal violations. Others leave error checking to separate circuitry.

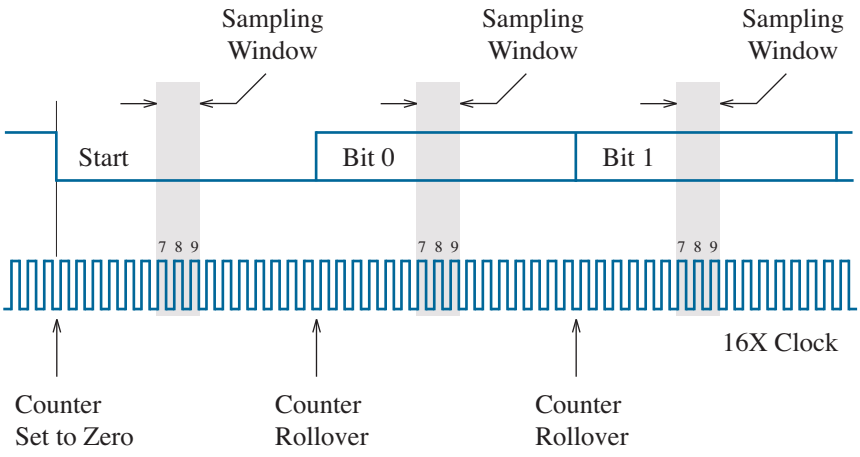


Figure 13.31: UART sampling process for incoming bit streams.

To ensure smooth data flow, UARTs are designed to follow **flow control** or **handshaking** protocols. If a device is ready to receive bits, it imparts a HIGH logic level on an RTS (Request to Send) output. The transmitting device holds off until it senses this level at a CTS (Clear to Send) input. Thus, RTS and CTS connections require two links, one for each direction. In an alternate handshaking process, the receiving UART sends an "XON" (ASCII 11h) control character back to the transmitting UART when it is ready to accept bits. The receiver sends "XOFF" (ASCII 13h) otherwise. Most UARTs feature several **FIFO** (*fi*-fo) First-in-First-out buffer registers that store unprocessed or reconstructed data to promote regular flow.

UARTs typically operate with a 3.3-V or 5-V HIGH logic level and a 0-V LOW logic level. Unfortunately, these levels are inappropriate for practical long-distance communication channels. Further conditioning is required.

RS-232

While somewhat old (1962), RS-232 is an industry standard that remains in common use for data communications over wired links. For the moment, we assume that wires allow rapid signal transfer with minor attenuation. Section 13.6 examines electrical details and corruptive modes of behavior.

The RS-232 standard calls for digital conditioning of transmitted bits. Specifically,

LOW	$\rightarrow$	Output between +5 V and +15 V
HIGH	$\rightarrow$	Output between -5 V and -15 V

Thus, the RS-232 transmitter or **line driver** has the form of the inverter shown in Fig. 13.32. The input is a typical UART-compatible CMOS level. The conditioned output is governed by the V^+ and V^- supply voltages. A 2500-pF maximum capacitive load limits line lengths to about 100 feet.

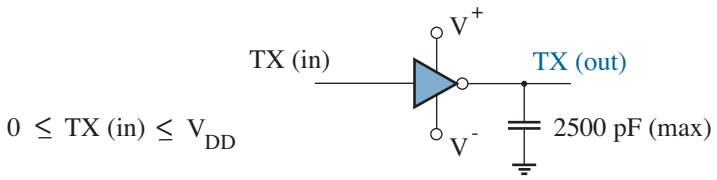


Figure 13.32: RS-232-transmitter circuit with logic-level conditioning.

The RS-232 standard also calls for digital conditioning of received bits. Specifically,

Input between +3 V and +15 V	$\rightarrow$	LOW
Input between -3 V and -15 V	$\rightarrow$	HIGH

Thus, the RS-232 receiver has the form of the inverter shown in Fig. 13.33. The input is an RS-232 level that is slightly less positive or negative to allow for signal loss on the line. The conditioned output is a typical CMOS level. Note the usual 5-kΩ input impedance, another RS-232 requirement.

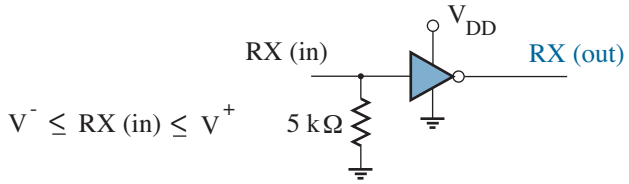


Figure 13.33: RS-232-receiver circuit with logic-level conditioning.

The MAX232 is noteworthy as the first integrated circuit to combine transmitter (TX) and receiver (RX) functions using one 5-V power supply. The IC package manages a **charge pump** featuring four external 1-μF capacitors and the four-step periodic charge sequence shown in Fig. 13.34. In Step 1, C_1 and C_2 connect in parallel with V_{DD}. In Step 2, the capacitors are reconfigured in series and then parallel-connected to C_3 to obtain $2V_{DD}$. In Step 3, the Step-1 process repeats. Finally, in Step 4, a series C_1-and-C_2 combination is inverted and parallel-connected to C_4 to produce $-2V_{DD}$. RS-232 levels are typically ± 8 V —the inherent limitations of the MOSFET switches used in the charge pump and circuitry for output regulation make them lower than $\pm 2V_{DD}$. Successors such as the MAX3232 achieve typical RS-232 levels of ± 5.4 V with one 3.3-V supply and four 0.1-μF capacitors.

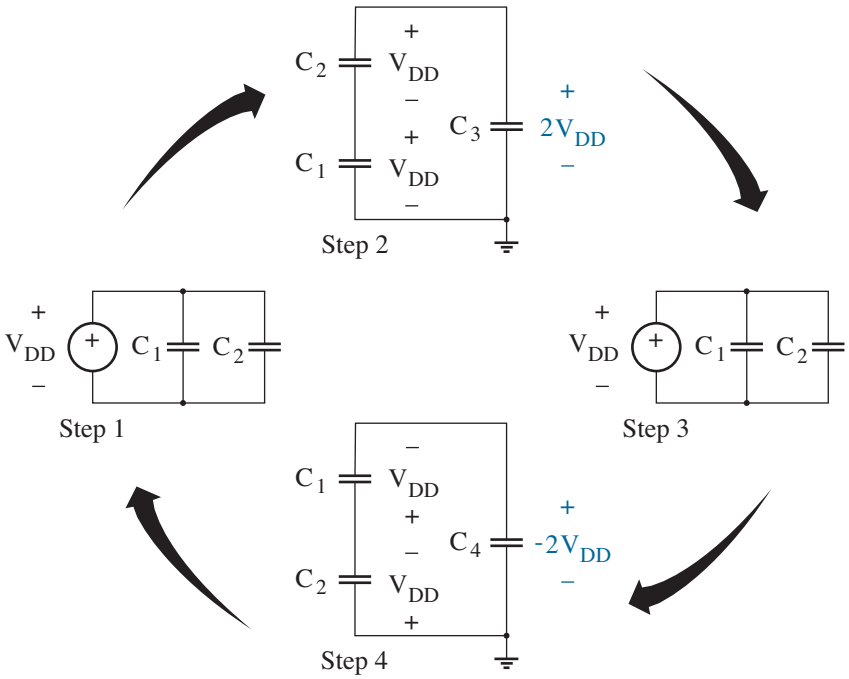


Figure 13.34: Charge-pump sequence for the MAX232.

To avoid line reflections and other Section-13.6 corruptions, the RS-232 standard limits maximum output slew rates to 30 V/μs. This implies a transition time of 0.33 μs for the case of a +5 V to -5 V level adjustment. Meanwhile, strict adherence to the RS-232 standard requires a bit rate no more than 20 kbps, which corresponds to a period no less than 50 μs/bit. Modern RS-232 circuits easily tolerate less conservative ratios of transition time to bit period. Thus, RS-232 bit-rate extensions as high as 250 kbps have become available. Compare with the bit rate of 300 bps in the 1970s.

The original RS-232 standard required 25 separate lines to provide data and control functions. Only nine have retained importance, and only three (TX, RX, and GND) are vital. Table 13.5 lists the line types in the order of RS-232-specified pin numbers for a 9-pin "D-sub" connector (Fig 13.35). The DTR and DSR lines assert receiver and transmitter operational status, the RI signal instructs a receiver to answer the phone, and the DCD signal indicates an active phone line. The latter two functions are nearly obsolete. Apart from ground, all of the lines warrant RS-232 logic levels.

Pin #	Symbol Line	Name	Transmitter Perspective	Receiver Perspective
1	DCD	Data Carrier Detect	Output	Input
2	RX	Receive Data	Input	Output
3	TX	Transmit Data	Output	Input
4	DTR	Data Terminal Ready	Input	Output
5	GND	Signal Ground	Ground	Ground
6	DSR	Data Set Ready	Output	Input
7	RTS	Request to Send	Input	Output
8	CTS	Clear to Send	Output	Input
9	RI	Ring Indicator	Output	Input

Table 13.5: RS-232-specified line types and nine-pin "D-sub" connections.

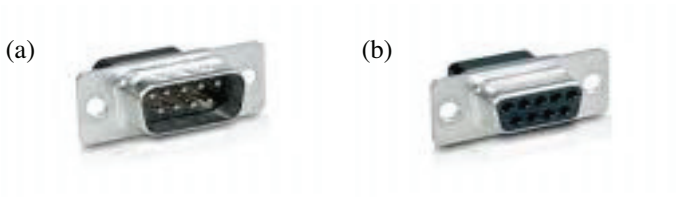


Figure 13.35: Nine-pin "D-sub" RS-232 connectors: (a) male; (b) female. Top-row pins number 1 to 5 (LR), bottom-row pins number 6 to 9 (LR).

Lest we feel starved for RS-232-related acronyms, the lines of Table 13.5 have historically been assigned to DCE (Data Communications Equipment) such as a computer and DTE (Data Terminal Equipment) such as a printer. Confusing statements are inevitable:

"The DTE's DTR disestablished the DCE's DSR ...

We reserve the terminology for socially challenged engineers. Nevertheless, it addresses the issue of a dominant "transmitter" and a submissive receiver so that proper cable connections can be established. Such distinctions are difficult when two computers talk to each other. In this common situation, one of the computers needs to connect to a **null modem** that interchanges the incoming RX/TX, RTS/CTS, and DSR/DTR line pairs.

Differential Data Transmission

Warning! Circuits function (or not) in a rich electromagnectic environment with extraneous currents, voltages, and radiation throbbing here and there. In the network of Fig. 13.36, for example, capacitors C_{1a} and C_{2a} reflect electric-field-induced coupling between trace a at voltage v_a and signal lines 1 and 2, respectively. The current injected into a quiescent line 1 is

$$i_1 = C_{1a}\frac{dv_a}{dt}\,. \tag{13.32}$$

Meanwhile, mutual inductances M_{1b} and M_{2b} reflect magnetic-field-induced coupling between trace b carrying current i_b and the same two signal lines. The voltage introduced in a mesh containing line 1 is

$$v_1 = M_{1b}\frac{di_b}{dt}\,. \tag{13.33}$$

Capacitive or inductive **crosstalk** is never zero. And notwithstanding our best efforts to minimize EMI effects through careful circuit-board layout, we can only hope for magnitudes that are less than upsetting thresholds. High frequencies diminish our prospects for success.

Despite these concerns, we note that lines 1 and 2 are equally affected by electromagnetic disturbances when they are placed in close proximity. Thus, adverse interference is significantly reduced if the signal of interest is the *difference* between the voltages on lines 1 and 2 so that crosstalks cancel. In what follows, we examine op-amp circuits for differential processing.

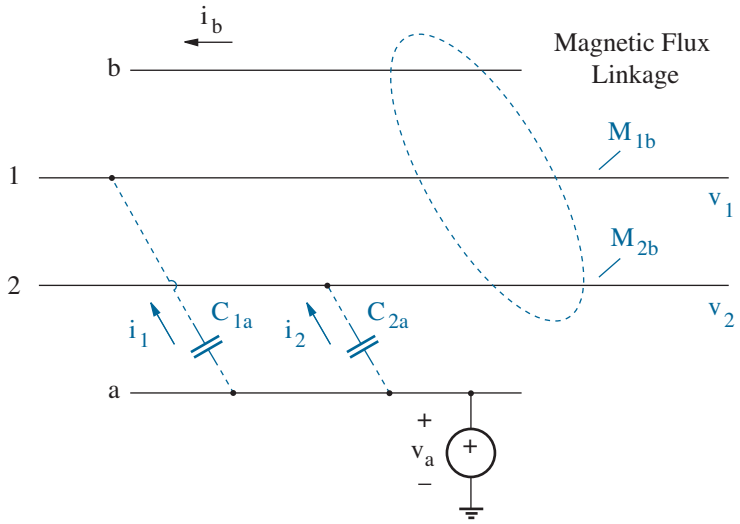


Figure 13.36: Capacitive and inductive coupling.

An EMI-resistant circuit of primary interest is the differential amplifier of Fig. 13.37 in which op-amps 1 and 2 present their high-impedance non-inverting inputs to differential signal components v_{i1} and v_{i2}, respectively. Negative feedback ensures that the node voltages at the inverting op-amp inputs are separately tracked to v_{i1} and v_{i2}. Thus, for op-amp 1,

$$\frac{v_{o1} - v_{i1}}{R_f} + \frac{v_{i2} - v_{i1}}{R_x} = 0 \tag{13.34}$$

so that

$$v_{o1} = \left(1 + \frac{R_f}{R_x}\right) v_{i1} - \frac{R_f}{R_x} v_{i2} \,. \tag{13.35}$$

Similarly, for op-amp 2,

$$\frac{v_{o2} - v_{i2}}{R_f} + \frac{v_{i1} - v_{i2}}{R_x} = 0 \tag{13.36}$$

so that

$$v_{o2} = \left(1 + \frac{R_f}{R_x}\right) v_{i2} - \frac{R_f}{R_x} v_{i1} \,. \tag{13.37}$$

We subtract Eqs. 13.35 and 13.37 to find

$$v_{o1} - v_{o2} = \left(1 + \frac{2R_f}{R_x}\right) (v_{i1} - v_{i2}) \,. \tag{13.38}$$

Here, the factor of $1 + 2R_f/R_x$ is A_{dd}, the differential voltage gain. For the common-to-differential-mode voltage gain, we have $A_{cd} = 0$ with $v_{i1} = v_{i2}$. In turn, the common-mode rejection ratio (CMRR) is infinite. These latter results are less favorable if the R_f resistors and op-amps are mismatched.

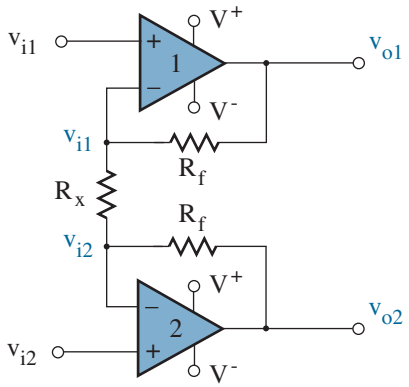


Figure 13.37: Two-op-amp differential amplifier.

Even with infinite CMRR, the circuit of Fig. 13.37 produces a common-mode output signal that is not necessarily zero. When we add Eqs. 13.35 and 13.37 and divide by two, we find

$$\left(\frac{v_{o1}+v_{o2}}{2}\right)=\left(\frac{v_{i1}+v_{i2}}{2}\right)\,. \tag{13.39}$$

Thus, the differential amplifier has unity common-mode voltage gain (A_{cc}). With $v_{i1}=-v_{i2}$, the differential-to-common-mode voltage gain is $A_{dc}=0$.

Chapter 9 showed that the two outputs v_{o1} and v_{o2} could be expressed in terms of a differential component $v_{od}=v_{o1}-v_{o2}$ and a common-mode component $v_{oc}=(v_{o1}+v_{o2})/2$. Specifically,

$$v_{o1}=\frac{+v_{od}}{2}+v_{oc} \tag{13.40}$$

and

$$v_{o2}=\frac{-v_{od}}{2}+v_{oc}\,. \tag{13.41}$$

Whereas the outputs are constrained by the op-amp power supplies,

$$V^{-}\le\frac{\pm v_{od}}{2}+v_{oc}\le V^{+}\,. \tag{13.42}$$

In turn,

$$2V^{-}\le v_{od}\le 2\,V^{+} \tag{13.43}$$

for $v_{oc}=v_{ic}=0$. A differential output signal with zero common mode has *twice* the dynamic range of a single-ended (zero-referenced) signal if both are limited by the op-amp supplies. This implies reduced distortion when v_{o1} and v_{o2} do not need to make excursions close to the power-supply rails. In any case, even distortion products cancel when forming v_{od}. If $v_{oc}\neq 0$, Eq. 13.40 yields

$$2\,[\,max(V^{-}-v_{oc}\,,V^{-})]\le v_{od}\le 2\,[\,min(V^{+}-v_{oc}\,,V^{+})]\,. \tag{13.44}$$

The available v_{od} dynamic range is diminished as v_{oc} nears one of the two power-supply levels.

Exercise 13.7 Let the differential amplifier of Fig. 13.37 have a single-ended input $v_{i1}=\pm 1$ V with $v_{i2}=0$. The op-amps have $\pm$5-V supplies. Determine the maximum acceptable differential gain if v_{o1} and v_{o2} are free to make rail-to-rail excursions.

Ans: $A_{dd}=9$

One way to impose $v_{oc} = 0$ for maximum dynamic range regardless of v_{ic} is to provide **common-mode feedback** as shown in the three-op-amp circuit of Fig. 13.38. By superposition,

$$v_c = v_{o1}\left(\frac{R}{R+R}\right) + v_{o2}\left(\frac{R}{R+R}\right) = \frac{v_{o1}+v_{o2}}{2}\,. \tag{13.45}$$

Notwithstanding the v_c sampling at the non-inverting input to op-amp 3, the overall feedback is *negative* with loops that tie to the inverting terminals of op-amps 1 and 2. Thus, op-amp 3 is forced to have equal input voltages, and $v_c = v_{oc} = 0$. Meanwhile, the op-amp-3 output is at node voltage v_x. Then node-voltage analysis reveals

$$v_{o1} = \left(1 + \frac{2R_f}{R_x}\right) v_{i1} - \frac{2R_f}{R_x}\, v_x \tag{13.46}$$

and

$$v_{o2} = \left(1 + \frac{2R_f}{R_x}\right) v_{i2} - \frac{2R_f}{R_x}\, v_x\,. \tag{13.47}$$

The average of Eqs. 13.46 and 13.47 is v_{oc}, which has been forced to zero. In turn, some further algebra provides the relation

$$v_x = \left(1 + \frac{R_x}{2R_f}\right) v_{ic}\,. \tag{13.48}$$

Proper circuit operation requires $V^- \le v_x \le V^+$ for non-zero v_{ic}.

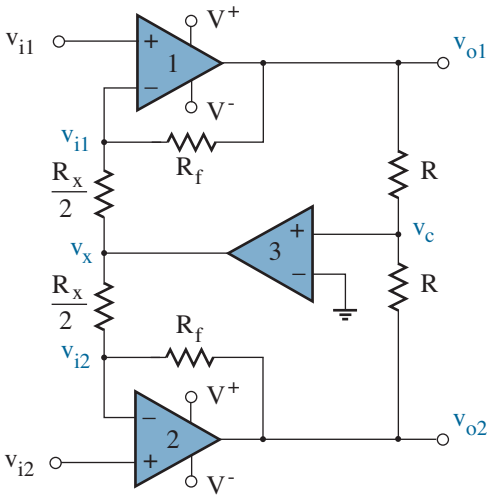


Figure 13.38: Three-op-amp differential amplifier with v_{oc} control.

Fully Differential Op-Amps

When we examined the source- or emitter-coupled pair as an op-amp front end in Chapter 9, we extracted a fully differential output from two node voltages that made equal and opposite Q-point excursions. Thus, it should come as no surprise that fully differential op-amps are available with two complementary outputs. The "traditional" non-inverting output satisfies

$$v_{o+} = A_{vd}\left(v^+ - v^-\right). \tag{13.49}$$

The inverting output, marked with a bubble on the op-amp symbol, satisfies

$$v_{o-} = -A_{vd}\left(v^+ - v^-\right). \tag{13.50}$$

Fully differential op-amps are used in *symmetric* circuits where the non-inverting output feeds back to the inverting input and the inverting output feeds back to the non-inverting input. Figure 13.39 shows a simple example.

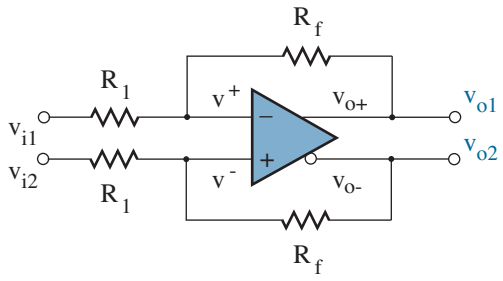


Figure 13.39: Fully differential inverting op-amp amplifier.

The analysis of the circuit of Fig. 13.39 is straightforward. We apply superposition to find v^+ subject to zero op-amp input current. Specifically,

$$v^+ = v_{i2}\left(\frac{R_f}{R_1 + R_f}\right) + v_{o2}\left(\frac{R_1}{R_1 + R_f}\right). \tag{13.51}$$

Similarly,

$$v^- = v_{i1}\left(\frac{R_f}{R_1 + R_f}\right) + v_{o1}\left(\frac{R_1}{R_1 + R_f}\right). \tag{13.52}$$

Then with $v^+ = v^-$,

$$v_{o1} - v_{o2} = -\left(\frac{R_f}{R_1}\right)\left(v_{i1} - v_{i2}\right). \tag{13.53}$$

So we have an inverting differential amplifier for which $A_{dd} = -R_f/R_1$. The amplifier also features $A_{cd} = 0$ and infinite CMRR as long as the resistors are symmetrically matched.

Unfortunately, the preceding analysis says nothing about the individual values for v_{o1} and v_{o2}, nor does it establish a common-mode average value. Additional common-mode feedback is required.

Figure 13.40 shows a simplified internal diagram of an op-amp featuring a fully differential folded cascode amplifier with common-mode feedback. The bias voltages V_{b1} and V_{b2} are constant. The internal op-amp samples v_c, the common-mode output voltage. Then it imposes a controlling voltage v_x so that v_c is made equal to v_{ocm} (derived from an external voltage source). The v_{o1} and v_{o2} outputs relate to v_x in a common-source and common-gate amplifier cascade. Thus, $v_c > v_{ocm}$ and increased v_x causes v_{o1} and v_{o2} to decrease in response to a positive common-mode perturbation.

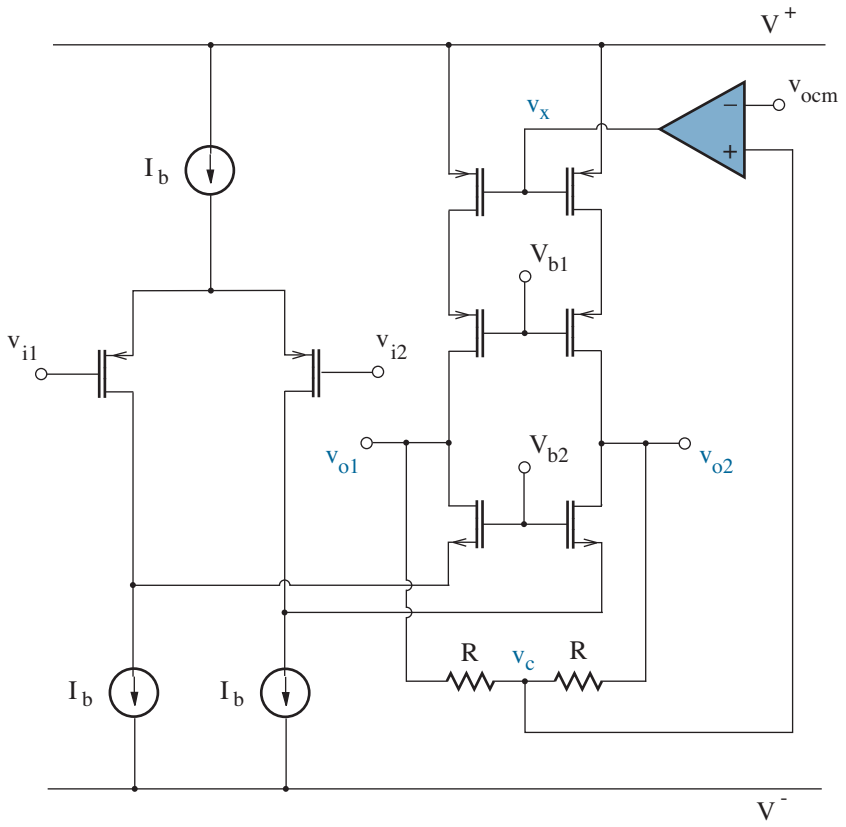


Figure 13.40: Fully differential amplifier featuring common-mode feedback. The common-mode voltage is controlled by v_{ocm} at an external pin.

Fully differential op-amps are often used as front-end amplifiers for some analog-to-digital converters (Chapter 14), and they come as discrete parts (from Texas Instruments, for example). Nevertheless, they also serve as **line drivers** for RS-485 data communications, our next consideration.

RS-485

In an RS-485 data communications system, transmitters have differential outputs Z and Y superimposed over a 2.5-V common-mode level for the case of a 5-V power supply. Typically, a HIGH state has $Z = 4$ V and $Y = 1$ V ($\Delta V = +3$ V), while a LOW state has $Z = 1$ V and $Y = 4$ V ($\Delta V = -3$ V). Receivers interpret $\Delta V > 0.2$ V as HIGH and $\Delta V < -0.2$ V as LOW, and the effects of corruptive noise are suppressed through differential sensing. Thus, long communication lines are permitted. The RS-485 standard limits the product of data rate (≤ 10 Mbps) and line length to about 10^8 bps-m.

Figure 13.41 shows a typical **full-duplex** RS-485 system that allows simultaneous transmit and receive operations. A centrally located master unit transmits data independently, and up to 255 slave units transmit data when the master indicates acceptable conditions. Four wires are required. A simpler **half-duplex** configuration requires only two wires, but only one transmit or receive operation is permitted at any time. Special hardware and protocols are needed to avoid data collisions. In either case, a set of two wires comprise a twisted-pair transmission line with 120-Ω impedance. Properly terminated lines promote signal integrity (see Section 13.6).

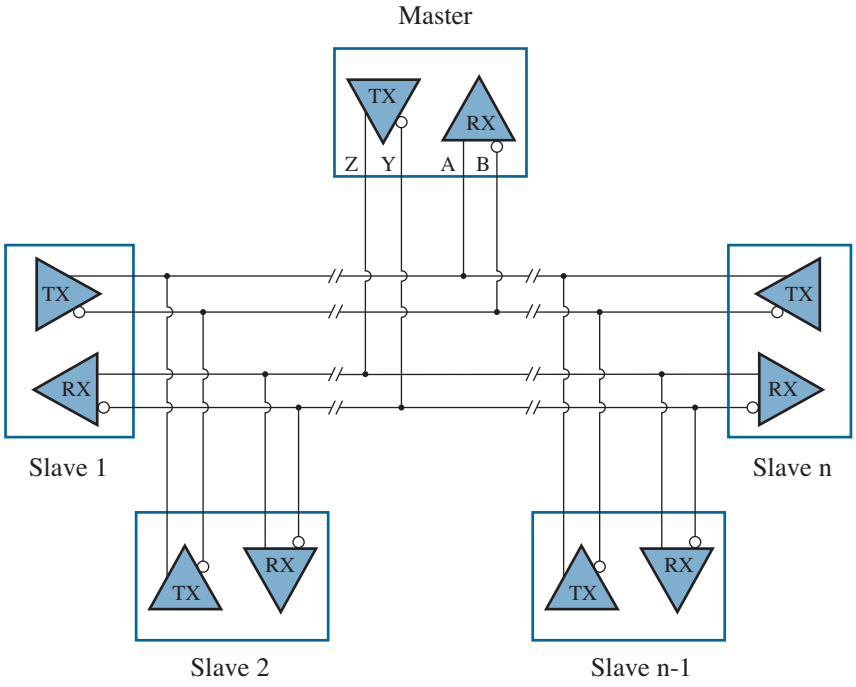


Figure 13.41: Typical RS-485 full-duplex system for simultaneous transmit and receive operations. A more restrictive half-duplex system shares the TX outputs and RX inputs.

13.4 Noise

Warning! Analog and digital signals are subject to the corrupting influence of electrical noise that is inherent to every circuit despite assiduous design. This section examines important noise sources, the characterization of noise at the black-box level, and design practices that promote noise mitigation. The physical basis of transistor-level noise is briefly considered.

One way to hear electrical noise is to crank up the volume of your stereo when no input source is connected and listen to the familiar hissing sound. The best way to *see* noise requires an instrument called a spectrum analyzer, which measures signal power density over a specified range of frequencies. Figure 13.42 shows a typical measurement. The large peak towards the left of the frequency span corresponds to a 120-kHz sinusoidal signal of interest. Everything else is noise.

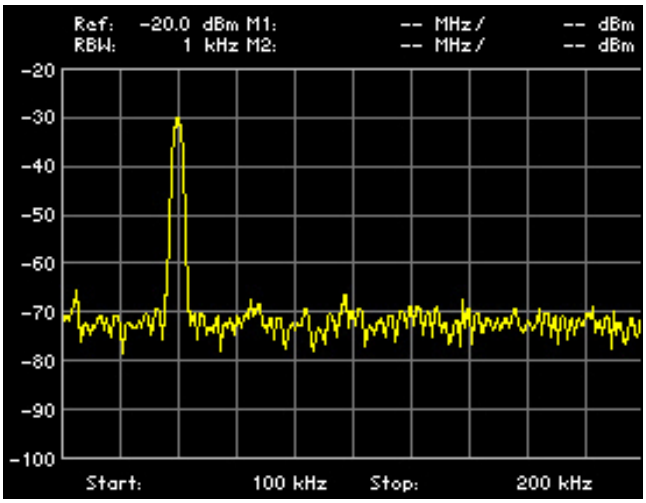


Figure 13.42: Spectrum analyzer measurement (dBm vertical scale).

Signal and noise powers tend to vary over many orders of magnitude, so the preferred measure of power is in dBm:

$$P\ (\text{dBm}) = 10\ \log_{10}\left(\frac{\text{Signal Power}}{1\ \text{mW}}\right)\ . \tag{13.54}$$

A spectrum analyzer displays the total power in dBm that passes through an effective bandpass filter featuring a specified resolution bandwidth and a center frequency that changes to match a frequency sweep within the span. The resolution bandwidth is 1 kHz for the measurement of Fig. 13.42. Thus, the vertical scale indicates power density in units of dBm/kHz.

Forget about signals for the moment. We can determine the total noise power that applies to the measurement of Fig. 13.42 by summing the power contributions to the various frequency bands: 100 kHz to 110 kHz, 110 kHz to 120 kHz, etc. The integration of power extends from 0 kHz to infinity. But noise spectra are commonly invariant with frequency. Thus, the infinite limit of integration leads to a divergent condition with infinite noise power. Perhaps we should forget about electrical signals altogether.

Fortunately, many electrical systems process noise as shown in Fig. 13.43. Here, internal network noise sources are referred to an external "Thevenin" noise source that produces identical noise behavior at the output terminals. The network is further subdivided into a cascade featuring an ideal network with no frequency dependence and a first-order low-pass filter that accounts for the generally dominant character of performance degradation.

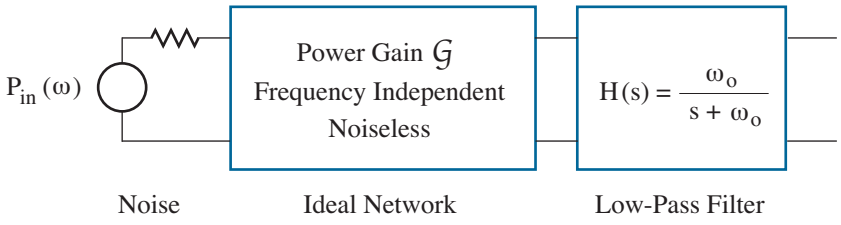


Figure 13.43: Noise processing in a typical electrical system.

In the system of Fig. 13.43, let $P_{in}(\omega)$ represent the power density of the input noise source, and let $\mathcal{G}$ equal the power gain of the "ideal" network. The noise power at the system output is

$$P_{out} = \int_0^\infty \mathcal{G} P_{in}(\omega)\,|H|^2 d\omega = \int_0^\infty \frac{\mathcal{G} P_{in}(\omega)\,{\omega_o}^2\,d\omega}{\omega^2 + {\omega_o}^2}\,. \tag{13.55}$$

Subject to uniform $P_{in}(\omega) = P_{in}$, the integration yields $P_{out} = \mathcal{G} P_{in}\,\omega_o \pi/2$. The same output noise power is achieved with a brick-wall low-pass filter with cutoff at B_n so that $P_{out} = \mathcal{G} P_{in} B_n$. Thus, the **noise bandwidth** is

$$B_n = \omega_o \frac{\pi}{2}\,. \tag{13.56}$$

Back to the measurement of Fig. 13.42. The circuit under test is an op-amp amplifier with 45-dB voltage gain. The frequency at which the voltage gain is reduced by $1/\sqrt{2}$ (-3 dB) is 165 kHz, so the noise bandwidth is 259 kHz. The average of the noise "floor" is about -73 dBm/kHz. In turn, the total noise power becomes $-73 + 10\log(259) = -49$ dBm (see Problem 13.69). To convert to an rms (root-mean-square) voltage amplitude, we must know the dissipating load resistance, which is usually 50 Ω. For the case at hand, -49 dBm is $\sqrt{50 \times 10^{-49/10} \times 10^{-3}} = 0.79$ mV, a somewhat raucous result.

Elementary Noise Sources

The lowly resistor is a source of electrical noise that arises from thermally induced and randomly directed electron currents in a conductive material. Also known as **Johnson noise**, the **thermal noise** spectrum is "white" in the sense that all frequencies make equal contributions (as for white light). In a Thevenin representation, a resistor features a series noise-voltage source with mean-square value

$$\overline{v_n{}^2} = 4kTRB_n \,, \tag{13.57}$$

where k is Boltzmann's constant (here taken as 1.381×10^{-23} V^2/K Ω Hz), T is absolute temperature (K), and B_n is an applicable noise bandwidth. In an alternate Norton representation, a resistor (conductance G) features a parallel noise-current source with mean-square value

$$\overline{i_n{}^2} = 4kTGB_n \,. \tag{13.58}$$

Figure 13.44 shows the equivalent noise models. In either case, the average value of the noise-source random variable vanishes ($\overline{v_n} = 0$ or $\overline{i_n} = 0$). Thus, a particular polarity or direction is meaningless. An rms noise value is obtained by taking the square root of Eq. 13.57 or 13.58. A specification that does not include B_n is V/$\sqrt{\text{Hz}}$ or A/$\sqrt{\text{Hz}}$ (V or A per "root Hz").

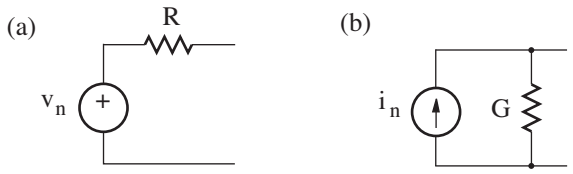


Figure 13.44: Resistor noise models: (a) Thevenin; (b) Norton.

Capacitors and inductors do not generate electrical noise. Nevertheless, real components are subject to parasitic resistance and associated noise. Both components alter noise spectra through filter action.

Exercise 13.8 Determine the Thevenin rms noise voltage for a 47-kΩ resistor at 290 K subject to $B_n = 50$ MHz.

Ans: v_n (rms) = 0.19 mV

Exercise 13.9 A specific Norton noise current is 2.7 pA/$\sqrt{\text{Hz}}$ at 290 K. Determine the corresponding resistor value.

Ans: $R = 2.2$ kΩ

It is constructive to consider the total noise produced by two resistors that are connected in parallel as shown in Fig. 13.45. By superposition,

$$v_n = v_{n1}\,\frac{R_2}{R_1+R_2} + v_{n2}\,\frac{R_1}{R_1+R_2}\,. \tag{13.59}$$

Then taking the square of Eq. 13.59, we find

$$v_n{}^2 = v_{n1}{}^2\left(\frac{R_2}{R_1+R_2}\right)^2 + 2\,v_{n1}v_{n2}\frac{R_1R_2}{(R_1+R_2)^2} + v_{n2}{}^2\left(\frac{R_1}{R_1+R_2}\right)^2. \tag{13.60}$$

In turn, the time average is

$$\overline{v_n{}^2} = \overline{v_{n1}{}^2}\left(\frac{R_2}{R_1+R_2}\right)^2 + 2\,\overline{v_{n1}v_{n2}}\frac{R_1R_2}{(R_1+R_2)^2} + \overline{v_{n2}{}^2}\left(\frac{R_1}{R_1+R_2}\right)^2. \tag{13.61}$$

We pause to observe that the v_{n1} and v_{n2} noise sources are **uncorrelated**—the random variations of one does not affect the other. Thus,

$$\overline{v_{n1}v_{n2}} = \overline{v_{n1}}\;\overline{v_{n2}} = 0\,, \tag{13.62}$$

since v_{n1} and v_{n2} separately average to zero. It follows that *cross-product terms in calculations featuring uncorrelated noise sources always vanish.* This has significant algebraic benefit when many noise sources are present. Returning to Eq. 13.61, we substitute $\overline{v_{n1}{}^2} = 4kTR_1B_n$ and $\overline{v_{n2}{}^2} = 4kTR_2B_n$ as expressions for the averages of the squared resistor noise voltages:

$$\begin{aligned}\overline{v_n{}^2} &= 4kTR_1B_n\left(\frac{R_2}{R_1+R_2}\right)^2 + 4kTR_2B_n\left(\frac{R_1}{R_1+R_2}\right)^2 \\ &= 4kT\left(\frac{R_1R_2}{R_1+R_2}\right)\,.\end{aligned} \tag{13.63}$$

The noise voltage resulting from the parallel combination of two resistors is the same as that for the equivalent resistance $R = R_1 \parallel R_2$.

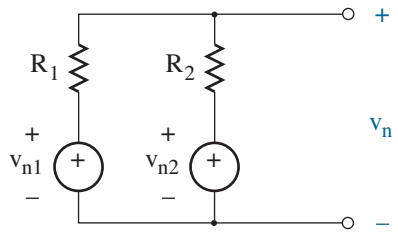


Figure 13.45: Noise circuit for two resistors connected in parallel.

Now consider the noise behavior arising from a "black-box" amplifier with voltage gain A_v and noise bandwidth B_n. In the model of Fig. 13.46a, a set of *internal* noise sources generate measured $\overline{v_{out,short}{}^2}$ and $\overline{v_{out,open}{}^2}$ when the input terminals connect to short and open circuits, respectively. The same noise measurements apply to the equivalent model of Fig. 13.46b in which *external* noise sources e_n and i_n excite an otherwise noiseless box. Both noise sources have uniform (white) spectral characteristics over B_n as a consequence of the averaging effect of the output noise measurement. When the input terminals connect to a short circuit, we have

$$\overline{v_{out,short}{}^2} = \overline{e_n{}^2} A_v{}^2 B_n \; . \tag{13.64}$$

Thus,

$$e_n = \sqrt{\frac{\overline{v_{out,short}{}^2}}{A_v{}^2 B_n}} \tag{13.65}$$

is the equivalent **input voltage noise density**. Similarly, when the input terminals connect to an open circuit, we have

$$\overline{v_{out,open}{}^2} = \overline{i_n{}^2} R_{in}{}^2 A_v{}^2 B_n \; , \tag{13.66}$$

where R_{in} is the amplifier input resistance. Thus,

$$i_n = \sqrt{\frac{\overline{v_{out,open}{}^2}}{R_{in}{}^2 A_v{}^2 B_n}} \tag{13.67}$$

is the equivalent **input current noise density**.

In the case of an operational amplifier, effective e_n and i_n sources apply exclusively at the non-inverting input (see Problem 13.79). The equivalent noise sources are often partially correlated since they reflect the aggregate behavior of overlapping sets of internal op-amp noise sources.

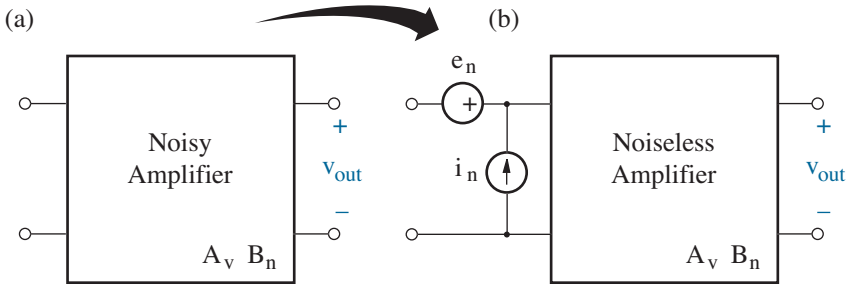


Figure 13.46: Equivalent average noise models for a two-port amplifier.

Figures of Merit

A widely used, although deceptive figure-of-merit for noise characterization is the **signal-to-noise ratio** defined as

$$\text{SNR} = \frac{\text{Signal Power}}{\text{Noise Power}} \,. \tag{13.68}$$

While this metric is indicative of prospects for obtaining useful information from a time-dependent voltage or current, it says nothing about the quality of circuits that participate in the signal detection and conditioning process. To describe circuit noise performance, one provides an average **noise factor**

$$F = \frac{\text{Total Output Noise Power}}{\text{Output Noise Power Due to the Input Source}} \,, \tag{13.69}$$

subject to the requirement that the input source operates at $T_o = 290$ K. An equivalent definition for the average noise factor is

$$F = \frac{\text{SNR (Input)}}{\text{SNR (Output)}} \,, \tag{13.70}$$

a reflection of degraded of signal-to-noise ratios between input and output. Restriction over a narrow frequency interval warrants a **spot noise factor**, and expression in decibels denotes a **(spot) noise figure**:

$$NF = 10 \, \log_{10} F \,. \tag{13.71}$$

The minimum noise factor (unity) and noise figure (0 dB) suggest an ideal circuit that adds no output noise. They do not indicate the absence of noise at the output. Both depend on the source admittance (see Problem 13.82). Measurements with special source admittance conditions are common.

When several two-port amplifiers connect in cascade, the noise factor is given by **Friis' formula** (see Problem 13.84). Specifically,

$$F = F_1 + \frac{F_2 - 1}{\mathcal{G}_1} + \frac{F_3 - 1}{\mathcal{G}_1 \mathcal{G}_2} + \, \ldots \, + \frac{F_n - 1}{\mathcal{G}_1 \mathcal{G}_2 \ldots \mathcal{G}_{n-1}} \,, \tag{13.72}$$

where F_i and $\mathcal{G}_i$ are the noise factor and power gain for the i-th two-port. Friis' formula shows that the first two-port (closest to the input) tends to make the greatest contribution to the overall noise factor. For this reason, amplifier front ends must be designed with careful attention to noise.

Systems such as those for satellite communications tend to have very-low-noise hardware with noise factors close to unity. Thus, it is sometimes convenient to specify a **noise temperature** that shows the relative increase in source temperature needed to account for the observed total output noise. For a given noise factor,

$$T_N = T_o \, (F - 1) \,. \tag{13.73}$$

Noise temperatures are additive in cascaded systems (see Problem 13.87).

Example 13.5

Figure 13.47 features an op-amp with $e_n = 35$ nV/$\sqrt{\text{Hz}}$, $i_n = 1$ fA/$\sqrt{\text{Hz}}$, and an 800-kHz gain-bandwidth product. Find the rms output noise voltage and the amplifier noise figure. Assume uncorrelated op-amp noise sources.

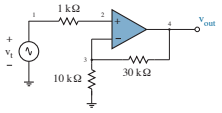


Figure 13.47: Circuit for Example 13.5.

Solution

The dc amplifier gain is +4, so the bandwidth is 800 kHz/4 = 200 kHz (see Example 12.1). In turn, the noise bandwidth is $200\,\pi/2 = 314$ kHz. Figure 13.48 shows the circuit for noise analysis. Then at $T_o = 290$ K,

- $\overline{v_{nt}{}^2} = 4kTR_tB_n = 5.03 \times 10^{-12}$ V^2
- $\overline{v_{n1}{}^2} = 4kTR_1B_n = 5.03 \times 10^{-11}$ V^2
- $\overline{v_{nf}{}^2} = 4kTR_fB_n = 1.51 \times 10^{-10}$ V^2

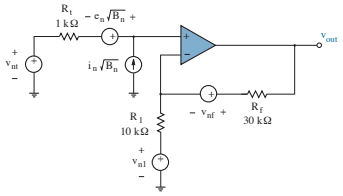


Figure 13.48: Noise circuit for Example 13.5.

When v_{nt} acts alone and the other sources are turned off, the network reduces to a non-inverting amplifier with

$$v_{out} = v_{nt}\left(1 + \frac{R_f}{R_1}\right) .$$

Similar expressions for v_{out} apply with $e_n\sqrt{B_n}$ or $i_n\sqrt{B_n}R_t$ replacing v_{nt}. When v_{n1} acts alone, we have an inverting amplifier with

$$v_{out} = -v_{n1}\left(\frac{R_f}{R_1}\right) .$$

And when v_{nf} acts alone, there is a virtual ground at the inverting terminal, no current flows through R_f, and

$$v_{out} = v_{nf} .$$

Then by superposition,

$$v_{out} = v_{nt}\left(1 + \frac{R_f}{R_1}\right) + e_n\sqrt{B_n}\left(1 + \frac{R_f}{R_1}\right) + i_n\sqrt{B_n}R_t\left(1 + \frac{R_f}{R_1}\right) - v_{n1}\left(\frac{R_f}{R_1}\right) + v_{nf}.$$

We square both sides of this expression and perform the time average while ignoring uncorrelated cross-product terms. Specifically,

$$\overline{v_{out}{}^2} = \underbrace{\overline{v_{nt}{}^2}\left(1+\frac{R_f}{R_1}\right)^2}_{\text{Term 1}} + \underbrace{\overline{e_n{}^2}\,B_n\left(1+\frac{R_f}{R_1}\right)^2}_{\text{Term 2}} + \underbrace{\overline{i_n{}^2}\,B_n R_t{}^2\left(1+\frac{R_f}{R_1}\right)^2}_{\text{Term 3}} + \underbrace{\overline{v_{n1}{}^2}\left(\frac{R_f}{R_1}\right)^2}_{\text{Term 4}} + \underbrace{\overline{v_{nf}{}^2}}_{\text{Term 5}}\,.$$

Next, we evaluate and compare the individual terms.

- Term 1: 8.05×10^{-11} V^2
- Term 2: 6.15×10^{-9} V^2 (Largest contribution)
- Term 3: 5.02×10^{-18} V^2
- Term 4: 4.53×10^{-10} V^2
- Term 5: 1.51×10^{-10} V^2

$$6.84 \times 10^{-9} \text{ V}^2$$

The square root of the sum yields an rms output noise voltage of 83 μV.

Whereas all of the $v_{out}{}^2$ terms in the preceding sum change by the same proportion when converted to output powers, the noise factor is the sum divided by Term 1. The result of 85.0 translates to a noise figure of 19 dB. The quickest path for improvement seeks an op-amp with smaller e_n.

Noise Models for Electronic Devices

Figure 13.49 presents the small-signal noise model for a pn junction diode. The current source adds **shot noise**—it sounds like lead shot dropped onto a tin roof—that accounts for fluctuating electron transport over a barrier. The mean-square noise current has the value

$$\overline{i_{nx}{}^2} = 2q\, i_d|_Q B_n \,, \tag{13.74}$$

where $i_d|_Q$ is the quiescent bias current. The shot-noise spectrum is white. It is important to note that resistor $r_d = (kT/q)/\, i_d|_Q$ is an artifact of the small-signal diode model—*it produces no noise.* The model of Fig. 13.49 neglects thermal noise that is generated in a series parasitic resistance.

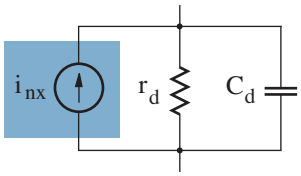


Figure 13.49: Small-signal diode noise model. Subject to 1-mA bias current, the i_{nx} noise-current spectrum resembles that for a 50-Ω resistor at 290 K.

Whereas BJT operation is based on the interaction of two closely spaced pn junctions, it should come as no surprise that the small-signal noise model includes shot-noise current sources as shown in Fig. 13.50. Here, we have

$$\overline{i_{nb}{}^2} = 2q\, i_b|_Q B_n \tag{13.75}$$

and

$$\overline{i_{nc}{}^2} = 2q\, i_c|_Q B_n \,, \tag{13.76}$$

given quiescent base ($i_b|_Q$) and collector ($i_c|_Q$) bias currents, respectively. Small-signal resistors r_π and r_o do not produce noise. However, the usually large parasitic base resistance adds thermal noise with $\overline{v_{nb}{}^2} = 4kTr_bB_n$.

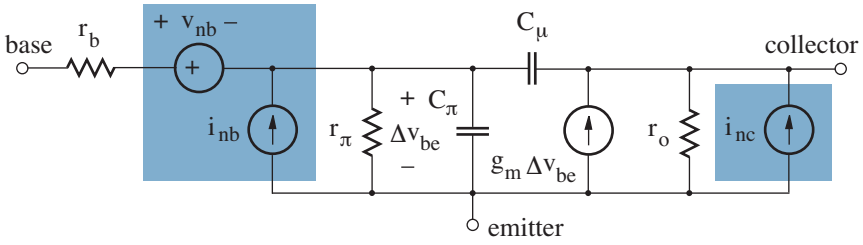


Figure 13.50: Small-signal noise model for a bipolar junction transistor.

Figure 13.49 shows a simple noise model for a three-terminal MOSFET. The model contains a single noise-current source with mean square value

$$\overline{i_{nd}{}^2} = \frac{8}{3}\, kT g_m B_n \,, \tag{13.77}$$

an indication of thermal noise from the channel given transconductance g_m. At very high frequencies (such that $\omega C_{gs} > g_m$), the noise model requires modifications that go beyond the scope of this text.

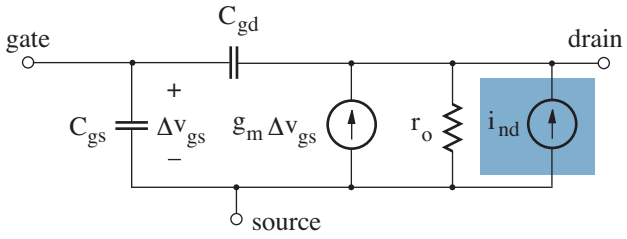


Figure 13.51: Small-signal noise model for a three-terminal MOSFET.

Another type of noise in electronic devices and resistors is **flicker noise** or $1/f$ (one over "f") noise. As the name implies, $1/f$ noise has a spectral density that increases inversely with frequency. Thus, it is problematic at low frequencies (< 1 kHz), and averaging is difficult (see Problem 13.99). For the case of a MOSFET, inclusion of $1/f$ noise in the small-signal model adds a current source in parallel with i_{nd} with mean square value

$$\overline{i_{ndf}{}^2} = \frac{K_f\, {g_m}^2}{C_{ox} W L}\, \frac{\Delta f}{f} \tag{13.78}$$

over a narrow frequency interval Δf. The process-dependent parameter K_f increases linearly with absolute temperature, and $C_{ox}WL$ reflects the total gate-to-channel capacitance. The transconductance g_m varies with $\sqrt{W/L}$. It follows that $1/f$ noise is minimized in circuit designs that incorporate MOSFETs with very large gate areas. Whereas the physical mechanisms for $1/f$ noise in MOSFETs involve carrier exchange between the channel and "traps" within the gate dielectric, more massive holes are less likely participants than electrons. In turn, p-channel MOSFETs tend to exhibit less $1/f$ noise than comparably sized n-channel devices. Low-noise CMOS op-amps thus have front-end differential stages with p-channel MOSFETs. As noted, diodes, BJTs, and resistors also produce $1/f$ noise, and the level reflects a power of the dc bias current (see Problems 13.100 and 13.101).

The intricacies of low-noise amplifier design at the transistor level is not a subject for this text. Nevertheless, it should be evident that keeping bias currents and resistors small is always beneficial for low-noise operation.

Example 13.6

Use SPICE to verify the output noise level of Example 13.5.

Solution

The simulation requires the e_n and i_n noise sources that accompany the otherwise *noiseless* op-amp in Fig. 13.48. To account for noiseless op-amp input resistance (or any special noiseless resistance), we apply a voltage-dependent current source (G) whereby the current between two connecting nodes is made proportional to the voltage between the same two nodes. The "gain" for the source is the reciprocal of the noiseless resistance.

To imitate the influence of e_n, we apply the circuit of Fig. 13.52 in which identical diodes are separately biased with equal current sources ($I_a = I_b$). The node voltages at a and b are about 0.7 V relative to *com*, but the node-voltage difference is subject to fluctuations within the diode noise model. Specifically,

$$ {e_n}^2 B_n = \overline{(v_a - v_b)^2} \,. \tag{13.79} $$

Then with $v_a = v_a|_Q + v_{an}$ (noise) and $v_b = v_b|_Q + v_{bn}$ (noise),

$$ {e_n}^2 B_n = \overline{(v_{an} - v_{bn})^2} = \overline{{v_{an}}^2} + \overline{{v_{bn}}^2} = 2\,\overline{{i_{nx}}^2}\,{r_d}^2 \,. \tag{13.80} $$

In consideration of Eq. 13.74 for the diode mean-square noise current $\overline{{i_{nx}}^2}$, the bias currents needed to produce noise-voltage density e_n are

$$ I_a = I_b = \frac{4q\,(kT/q)^2}{{e_n}^2} \,. \tag{13.81} $$

For Example 13.1, $e_n = 35$ nV/$\sqrt{\text{Hz}}$, and $I_a = I_b = 328$ nA ($T_o = 290$ K). The noise transfers to the op-amp input with the help of a voltage-dependent voltage source:

```
En      x       y       a       b       1
```

The voltage between nodes x and y has the same noise dependence as that between nodes a and b since the end-of-statement gain factor is unity.

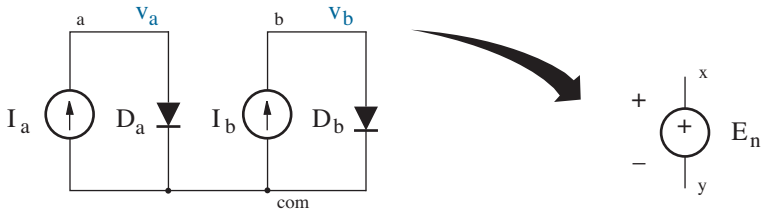


Figure 13.52: Artificial SPICE circuit for noise source e_n.

To imitate the influence of i_n, we apply the circuit of Fig. 13.53 in which a single diode is biased with current source I_c. Here, we have

$$i_n{}^2 B_n = \overline{i_{nx}{}^2}\,. \tag{13.82}$$

Thus, the bias current needed to produce noise-current density i_n is

$$I_c = \frac{i_n{}^2}{2q}\,. \tag{13.83}$$

For Example 13.1, $i_n = 1$ fA/$\sqrt{\text{Hz}}$, and $I_c = 3.12$ pA. The noise transfers to the op-amp input with the help of a current-dependent current source that monitors the current through the zero-value voltage source V_m:

```
Fn      x       y       Vm      1
```

The end-of-statement unity gain factor ensures one-to-one tracking.

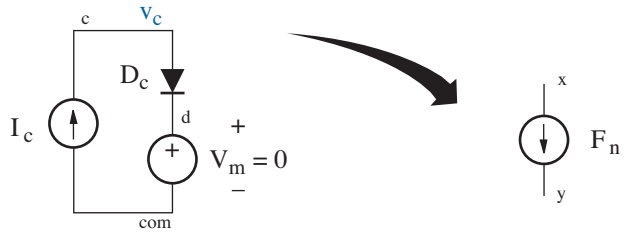


Figure 13.53: Artificial SPICE circuit for noise source i_n.

The preceding deliberations suggest the following SPICE netlist:

```
* Noise simulation for Example 13.6.

Vin     1       0       ac      1m
Rt      1       2       1k
R1      3       0       10k
Rf      3       4       30k
X1      4       0       2       3       OpAmp

.temp   17
.ac     dec     50      1       1MEG
.noise  v(4)    Vin
.probe

.subckt OpAmp   out     com     in_p    in_n
Gin     rin_p   in_n    rin_p   in_n    1n
Eout    out     com     LAPLACE
+       { v(rin_p , in_n) } { 800k*(2*3.14159 / (s + 2*3.14159)) }
```

```
* Noise-source generators

En        rin_p  in_p   a       b       1
Fn        com    rin_p  Vm      1
Ia        com    a      328n
Da        a      com    Diode
Ib        com    b      328n
Db        b      com    Diode
Ic        com    c      3.12p
Dc        c      com    Diode
Vm        d      com    0
.model    Diode  D
.ends

.end
```

Note the .temp statement for analysis at 290 K. The .noise statement calls for noise analysis with Vin as the input signal source and node-4 output.

Figure 13.54 shows .probe results. VONOISE is node-4 output noise in units of $\mathrm{V}/\sqrt{\mathrm{Hz}}$. The amplifier high-frequency cutoff where the noise (or applied signal) falls to 0.707 times the low-frequency value is 200 kHz. In turn, the noise bandwidth is $200\ \mathrm{kHz} \times \pi/2 = 314\ \mathrm{kHz}$. The output noise level is $147.4\ \mathrm{nV}/\sqrt{\mathrm{Hz}} \times \sqrt{314 \times 10^3\ \mathrm{Hz}} = 83\ \mu\mathrm{V}$.

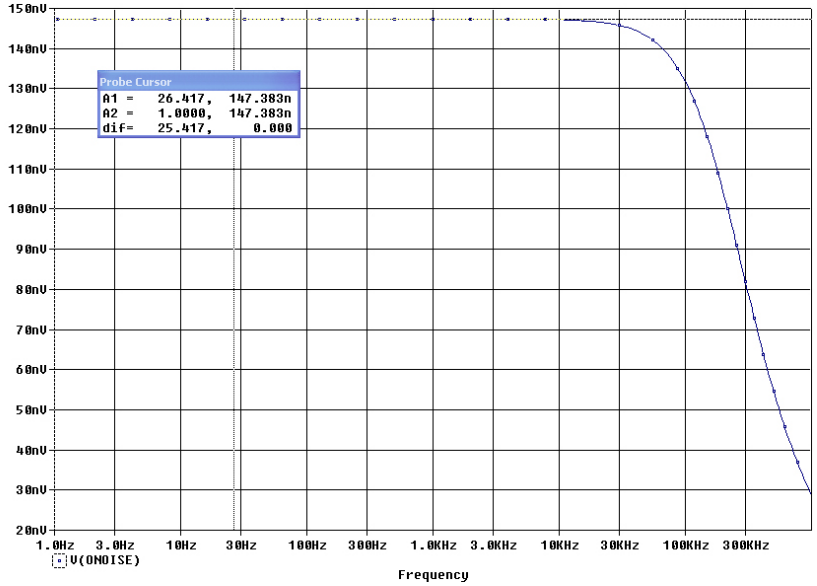


Figure 13.54: .probe results for the simulation of Example 13.6.

13.5 Distortion

Warning! Analog signals are subject to distortion from amplifiers and other conditioning equipment. The amplitude-dependent corruption may reflect the inherent non-linearity of constituent devices or power-supply bounds. Whatever the cause, we need some metrics for predicting circuit operating conditions that favor distortion so that it can be minimized.

Consider a voltage amplifier with the input-output relation

$$v_{out} = \underbrace{K_1\, v_{in}}_{e_1} + \underbrace{K_2\, {v_{in}}^2}_{e_2} + \underbrace{K_3\, {v_{in}}^3}_{e_3} + \ldots \tag{13.84}$$

subject to $v_{in} = a\cos\omega_1 t$ and *small* values for K_2, K_3, etc. A common distortion measure is **total harmonic distortion (THD)**, given by

$$\text{THD} = \frac{\sqrt{{e_2}^2 + {e_3}^2 + \ldots}}{e_1} \times 100\% \,. \tag{13.85}$$

Nevertheless, a more useful metric is the **1-dB compression point** found by plotting output power vs. input power (in dBm) as shown in Fig. 13.55. The low-level relationship is linear with unity slope. Eventually, the output power falls short of this behavior with limiting distortion. Thus, the 1-dB compression point is the input power for which the output power falls 1 dB below the projected slope of unity. In turn, the amplifier **dynamic range** is the difference between the 1-dB compression point and the input power needed for the output to be distinguished from the noise floor. Signals with power levels outside the dynamic range cannot be processed reliably.

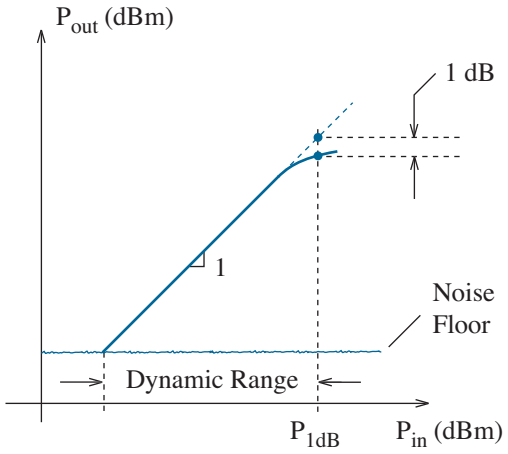


Figure 13.55: Measurement for the 1-dB compression point.

Intercepts

Another type of distortion-related plot can reveal the level of an undesired harmonic given a particular amplitude for a fundamental signal frequency. Let $v_{in} = a\cos\omega_1 t$ in Eq. 13.84, and neglect third- and higher-order terms. Then

$$v_{out} = K_1\, a\cos\omega_1 t + \frac{1}{2}K_2\, a^2 + \frac{1}{2}K_2\, a^2\cos 2\omega_1 t\,, \tag{13.86}$$

for which we have used the identity

$$\cos^2\omega_1 t = \frac{1+\cos 2\omega_1 t}{2}\,. \tag{13.87}$$

With the help of a spectrum analyzer we plot the output power in dBm at the ω_1 fundamental against the total input signal power. The relationship is linear with unity slope at low power levels, but it eventually falls short of this behavior as in Fig. 13.55. We also plot the output power at the $2\omega_1$ second harmonic against the total input signal power. The a^2 dependence in Eq. 13.86 predicts that the second-harmonic power increases twice as fast as that for the fundamental before falling short. Thus, the plotted segments with the slopes of one and two extrapolate to a common **intercept point** as shown in Fig. 13.56. When referred to the output power, we specify the second-order-harmonic output intercept point or OIP2_H. When referred to the input power, we specify IIP2_H. A similar plot applies to the evolution of the third harmonic and an intercept point denoted as OIP3_H or IIP3_H. In this case, the third-harmonic power increases three times as fast as that for the fundamental.

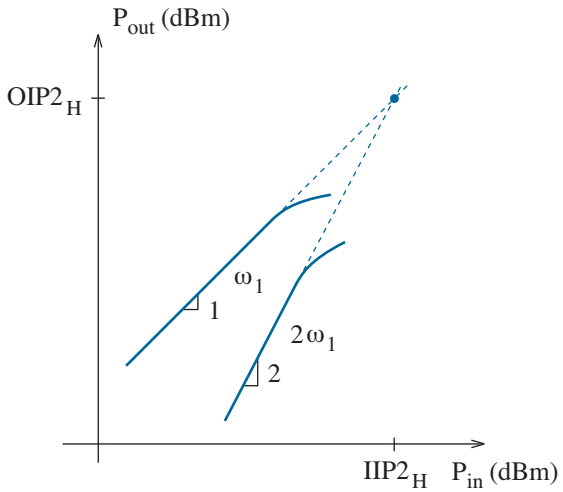


Figure 13.56: Measurement for the second-order harmonic intercept point.

Example 13.7

An amplifier features $\text{OIP2}_H = 12$ dBm. The fundamental output signal power at 20 kHz is -8 dBm. Determine the output signal power for the second harmonic at 40 kHz.

Solution

The 20-kHz signal power is $12-(-8) = 20$ dB below the OIP2_H intercept. Thus, the second-harmonic signal power is 40 dB below at -28 dBm.

Exercise 13.10 An amplifier with $\text{OIP3}_H = 8$ dBm is to provide no more than -40 dBm of output signal power at a 24-kHz third harmonic. Find the maximum acceptable output signal power at the 8-kHz fundamental.

Ans: -8 dBm

An opportunity for **two-tone intermodulation distortion** arises if $v_{in} = a\cos\omega_1 t + b\cos\omega_2 t$, where ω_2 typically represents a competing signal. When v_{in} is applied to Eq. 13.84, numerous harmonic and intermodulation signal components are produced with the amplitudes specified in Table 13.1. Here, we have made frequent use of the relation

$$\cos\omega_1 t\,\cos\omega_2 t = \frac{1}{2}\left[\cos(\omega_1-\omega_2)t + \cos(\omega_1+\omega_2)t\right]. \tag{13.88}$$

The fundamental amplitude at ω_1 is limited for large a, so $K_3 < 0$. In turn, large b can cause the ω_1 fundamental to *drop out completely*. Apart from the spectral components at $2\omega_1-\omega_2$ and $\omega_1-2\omega_2$, the distortion byproducts are mostly subject to elimination through filtering.

The intercepts that reflect the evils of K_3 intermodulation distortion are denoted as OIP3_{IM} or IIP3_{IM}. These are typically evaluated with $a = b$ so that the spurious signal components increase in proportion to a^3.

To minimize distortion effects in amplifiers, bias conditions should be selected so that quiescent operating points are at the midpoints of pertinent transistor load lines. Avoid early departures from saturation (MOSFETs) or the forward active mode (BJTs). When multiple gain stages are used, take particular care near the amplifier output. Attenuate the expression of device non-linearities by restricting drain- or collector-current changes to sub-quiescent levels (see Problem 13.113).

Negative feedback is invariably helpful. Examples include the insertion of a source or emitter resistor in the common-source or common-emitter amplifier and the use of series-shunt feedback for the mitigation of cross-over distortion in the BJT push-pull amplifier (see Problem 13.118).

Spectral Component	Amplitude	Easily Filtered? ($\omega_2 \sim \omega_1$)
ω_1	$K_1 a + \frac{3}{4} K_3 (a^3 + 2ab^2)$	—
ω_2	$K_1 b + \frac{3}{4} K_3 (b^3 + 2a^2 b)$	No
$\omega_1 - \omega_2$	$K_2\, ab$	Yes
$\omega_1 + \omega_2$	$K_2\, ab$	Yes
$2\omega_1$	$\frac{1}{2} K_2\, a^2$	Yes
$2\omega_2$	$\frac{1}{2} K_2\, b^2$	Yes
$2\omega_1 - \omega_2$	$\frac{3}{4} K_3\, a^2 b$	No
$\omega_1 - 2\omega_2$	$\frac{3}{4} K_3\, ab^2$	No
$2\omega_1 + \omega_2$	$\frac{3}{4} K_3\, a^2 b$	Yes
$\omega_1 + 2\omega_2$	$\frac{3}{4} K_3\, ab^2$	Yes
$3\omega_1$	$\frac{1}{4} K_3\, a^3$	Yes
$3\omega_2$	$\frac{1}{4} K_3\, b^3$	Yes

Table 13.6: Two-tone intermodulation distortion products.

Distortion is not always bad. **Mixers** are non-linear circuits that are designed to produce $\omega_1 - \omega_2$ or $\omega_1 + \omega_2$ signals in communication systems. Advocates of "heavy-metal" guitar music often like to enhance distortion with the help of diode clipping. Indeed, the unique distortion characteristics of vacuum tubes has kept this technology alive in the audio community.

Example 13.8

A 10-kHz sinusoidal signal applies to the amplifier of Fig. 13.57. Use SPICE to find the relative amplitudes of the second and third output harmonics with 0.1-V and 0.4-V input amplitudes. $K'W/L = 2$ mA/V^2, $V_T = 0.5$ V.

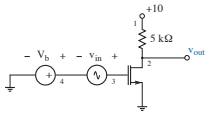


Figure 13.57: Circuit for Example 13.8.

Solution

The dc bias voltage at the MOSFET gate establishes $i_d|_Q = 1$ mA and $v_{out}|_Q = 5$ V. Then $g_m = 2 \times 10^{-3}$ ℧ and $A_{vm} = -g_m R_d{}' = -10$.

SPICE coding is routine apart from the control statements

```
.tran    1n     500u   0      1u
.four    10k    3      v(2)
```

The first statement makes a transient analysis over five periods of duration $T = 100$ μs subject to a maximum step size of $T/100$. The second statement completes a Fourier analysis of the first three harmonics of 10 kHz with node 2 as the designated voltage output. For the input signal, we write

```
Vin      3      4      SIN    (0     0.1    10k)
```

The zero in the SIN argument reflects zero signal delay.

After running the simulation, we look to the output file to find

DC COMPONENT = 4.975000E+00

HARMONIC NO	FREQUENCY (HZ)	FOURIER COMPONENT	NORMALIZED COMPONENT
1	1.000E+04	9.997E-01	1.000E+00
2	2.000E+04	2.497E-02	2.498E-02
3	3.000E+04	5.838E-07	5.840E-07

TOTAL HARMONIC DISTORTION = 2.497583E+00 PERCENT

When we repeat the simulation subject to an input amplitude of 0.4 V, the effects of distortion are clearly evident in the .probe plot of Fig. 13.58. For the first ten harmonics we find

DC COMPONENT = 4.702489E+00

HARMONIC NO	FREQUENCY (HZ)	FOURIER COMPONENT	NORMALIZED COMPONENT
1	1.000E+04	3.801E+00	1.000E+00
2	2.000E+04	2.240E-01	5.893E-02
3	3.000E+04	1.430E-01	3.763E-02
4	4.000E+04	1.051E-01	2.764E-02
5	5.000E+04	6.699E-02	1.762E-02
6	6.000E+04	3.370E-02	8.864E-03
7	7.000E+04	8.563E-03	2.253E-03
8	8.000E+04	7.025E-03	1.848E-03
9	9.000E+04	1.362E-02	3.582E-03
10	1.000E+05	1.342E-02	3.482E-03

TOTAL HARMONIC DISTORTION = 7.794486E+00 PERCENT

The fundamental amplitude of 3.8 (not 4.0) implies -0.44-dB compression. Thus, the amplifier is close to the limit of its dynamic range.

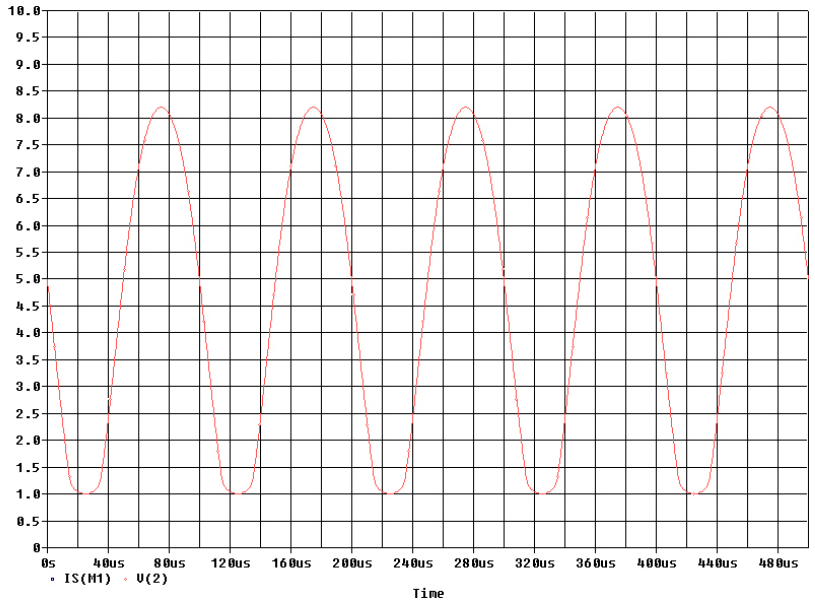


Figure 13.58: .probe plot for Example 13.8 (0.4-V input amplitude).

13.6 Transmission-Line Effects

Warning! Wires are electrical conduits for the information superhighway. And while they may not appear to be particularly exciting circuit elements, wires impose hazardous conditions when treated indifferently. This section examines electrical transport processes on long wires and concurrent forms of mild signal conditioning that are undesirable but easily accommodated. Thoroughly corruptive wire behavior is a topic for circumventing design.

Short wires that locally connect R_1 to R_2 and C_3 on a printed circuit board are like little side streets for signal flow and do not concern us here. Rather, we focus on two-wire **transmission lines** that function more like freeways for long-distance travel. In the electrical sense, "long" implies a length l that is at least comparable to the spatial period of a moving signal. Specifically,

$$l \sim \hat{v}\,T\,, \tag{13.89}$$

where $\hat{v}$ is the velocity of signal propagation and T is the signal time period. Subject to $T = 5$ ns (200 MHz) and typical $\hat{v} = 2 \times 10^8$ m/s, $l \sim 1$ m. Figure 13.59 shows three common forms of transmission line.

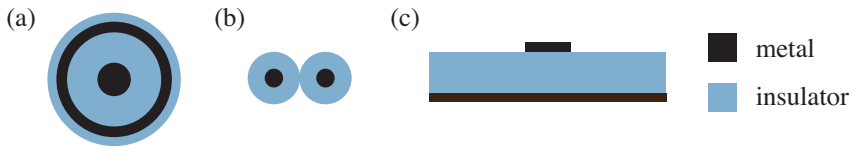


Figure 13.59: Transmission lines: (a) coaxial; (b) twisted pair; (c) strip-line.

Our focus is further limited to transmission lines that maintain uniform cross section. Thus, per-unit-length line specifications for inductance (L), capacitance (C), series resistance (R), and shunt conductance (G) reflect a transverse materials geometry in x and y, not the propagation direction z. Figure 13.60 shows a circuit model for a transmission-line segment in terms of these specificatons. The segment length is Δz.

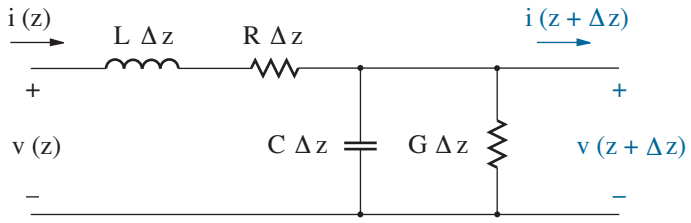


Figure 13.60: Electrical model for a transmission line of length Δz.

A preliminary objective is to derive the impedance Z_o looking into the end of a semi-infinite transmission line. For this task, we apply a neat trick: we terminate the right side of the circuit of Fig. 13.60 with impedance Z_o. A line with length Δz + infinity remains infinite, so the impedance looking into the left side of the terminated circuit is also Z_o. Thus,

$$Z_o = (\jmath\omega L + R)\Delta z + \underbrace{\frac{1}{(\jmath\omega C + G)\Delta z} \parallel Z_o}_{Z'} \quad . \tag{13.90}$$

The underbraced parallel combination in Eq. 13.90 takes the form

$$\begin{aligned} Z' &= \frac{Z_o}{1 + Z_o(\jmath\omega C + G)\Delta z} \\ &\approx Z_o \left\{1 - Z_o(\jmath\omega C + G)\Delta z + [Z_o(\jmath\omega C + G)\Delta z]^2 - \dots\right\} \end{aligned} \tag{13.91}$$

following expansion as a geometric series. Since Δz is small, we can discard terms of order $(\Delta z)^2$ or higher. Then returning to Eq. 13.90, we have

$$Z_o \approx (\jmath\omega L + R)\Delta z + Z_o - {Z_o}^2(\jmath\omega C + G)\Delta z \, . \tag{13.92}$$

Some further algebra yields a general result:

$$Z_o = \sqrt{\frac{\jmath\omega L + R}{\jmath\omega C + G}} \, . \tag{13.93}$$

Nevertheless, practical transmission lines feature $R \ll \jmath\omega L$ and $G \ll \jmath\omega C$. In turn,

$$Z_o \approx \sqrt{\frac{L}{C}} \, . \tag{13.94}$$

The impedance of a semi-infinite transmission line is resistive.

The resistive nature of Eq. 13.94 may seem odd since $R \to 0$ and $G \to 0$ leaves only inductive and capacitive components in the model of Fig. 13.60. However, the resistive impedance is actually an indication of similar phase demands that place current and voltage in lockstep as they enter the line. Like a resistor, the semi-infinite line absorbs incident power.

Equation 13.94 also suggests that a transmission line with finite length can be made to appear like a semi-infinite line if the far end terminates with a resistor of value Z_o. Transmission lines tend to be intellectually challenged —they are thoroughly clueless regarding the difference between a Z_o load resistance and the alluring notion of trailing off into the sunset.

Various applications have settled on specific industry standards for Z_o: 50 Ω (radio-frequency communications hardware), 75 Ω (video, cable TV), and 100 Ω (ethernet data transfers). (See Problems 13.122 and 13.123.)

Now consider signal propagation along the transmission line of Fig. 13.60. From Kirchhoff's voltage law, we have

$$v(z+\Delta z) = v(z) - L\Delta z\,\frac{\partial i}{\partial t} - R\Delta z\, i\,. \tag{13.95}$$

In the limit as $\Delta z \to 0$, $[\,v(z+\Delta z) - v(z)\,]/\Delta z$ is a spatial derivative in z. Thus,

$$\frac{\partial v}{\partial z} = -L\,\frac{\partial i}{\partial t} - R\, i\,. \tag{13.96}$$

Similarly, from Kirchhoff's current law,

$$i(z+\Delta z) = i(z) - C\Delta z\,\frac{\partial v}{\partial t} - G\Delta z\, v\,. \tag{13.97}$$

In turn,

$$\frac{\partial i}{\partial z} = -C\,\frac{\partial v}{\partial t} - G\, v\,. \tag{13.98}$$

Equations 13.96 and 13.98 are **Telegrapher relations**, which are famous, having been used to model the propagation of telegraph signals on the first transatlantic cable. (The theory was problematic —see Problem 13.124.)

For the moment, we restrict our attention to a lossless transmission line with $R = G = 0$. Then it is easy to eliminate the current variable from the Telegrapher relations as follows: Differentiate both sides of Eq. 13.96 with respect to z, differentiate both sides of Eq. 13.98 with respect to t, determine separate expressions for $\partial^2 i/\partial z\partial t$ and equate them. This process yields

$$\frac{\partial^2 v}{\partial z^2} = LC\,\frac{\partial^2 v}{\partial t^2} \tag{13.99}$$

as one form of the **wave equation**. Readers who are getting nervous about a profusion of partial derivatives will be relieved to know that substitution verifies the solution

$$v(z,t) = v_+(z - \hat{v}\, t)\,, \tag{13.100}$$

subject to a velocity

$$\hat{v} = \frac{1}{\sqrt{LC}}\,. \tag{13.101}$$

If we eliminate the voltage variable from the lossless Telegrapher relations, we find

$$\frac{\partial^2 i}{\partial z^2} = LC\,\frac{\partial^2 i}{\partial t^2} \tag{13.102}$$

with the readily verified solution

$$i(z,t) = i_+(z - \hat{v}\, t)\,. \tag{13.103}$$

Although related, the v_+ and i_+ functions are arbitrary.

Let $t = 0$, and let $v\,(z, 0) = v_+\,(z)$ assume a pulse-like form (Fig. 13.61a). At $t = t_1$, the v_+ functional argument is $x - \hat{v}\,t_1$, so the spatial distribution of the voltage pulse is unchanged apart from a translational shift by $\hat{v}\,t_1$ (Fig. 13.61b). At $t = t_2$, the pulse is shifted by $\hat{v}\,t_2$ (Fig. 13.61c).

The shape of any voltage pulse is unmodified as it moves with velocity $\hat{v}$ along a lossless semi-infinite or Z_o-terminated transmission line.

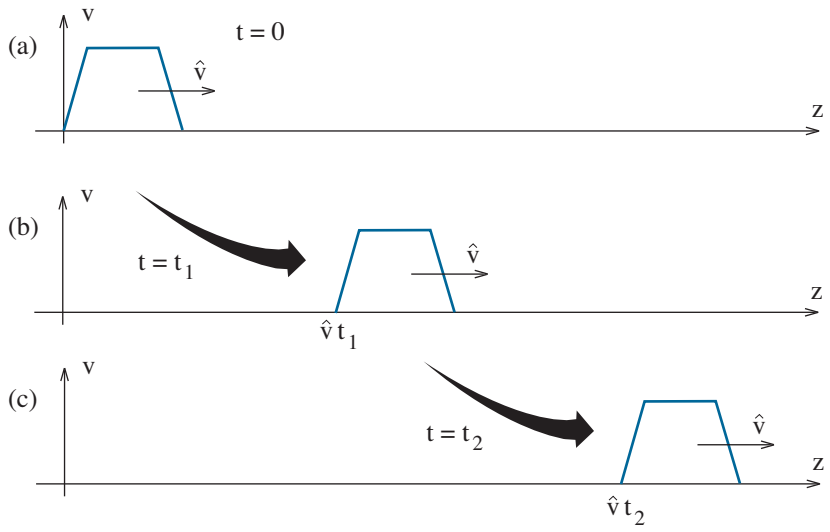


Figure 13.61: Voltage-pulse evolution along a lossless transmission line.

One finds the same propagation behavior with $i_+\,(x - \hat{v}\,t)$ line currents. Since voltage and current pulses travel with the same velocity $\hat{v}$ and $v = Z_o i$ at $z = 0$, we require

$$v_+(z - \hat{v}\,t) = Z_o\,i_+(z - \hat{v}\,t) \tag{13.104}$$

along the entire transmission line.

Exercise 13.11 RG-58C cable features $Z_o = 50\ \Omega$ and $C = 101$ pF/m. Determine L.

Ans: $L = 253$ nH/m

Exercise 13.12 RG-59B cable features $Z_o = 75\ \Omega$ and $C = 67.6$ pF/m. An 80-m line is terminated with 75 Ω. Determine the pulse delay time.

Ans: delay = 410 ns

Unmodified propagation along a lossless line is more than encouraging, but real transmission lines feature attenuation through $R \neq 0$ and $G \neq 0$. To work out the lossy-line problem, we return to the Telegrapher relations in the frequency domain. Specifically, we assume a voltage or current response that varies as $\exp \jmath\omega t$ so that time derivatives are reduced to $\jmath\omega$ operators. Upon elimination of the current variable, the line voltage satisfies

$$\frac{\partial^2 v}{\partial z^2} = \gamma^2 v \,, \tag{13.105}$$

where

$$\gamma = \sqrt{(\jmath\omega L + R)(\jmath\omega C + G)} = \alpha + \jmath\beta \tag{13.106}$$

is a complex **propagation factor**. A solution to Eq. 13.105 is proportional to $\exp \pm\gamma z$. However, with $\alpha > 0$, we are only interested in the negative exponent since the voltage amplitude must be finite in the limit as $z \to \infty$. Thus,

$$v(z,t) = \tilde{v}_+ \, e^{-\alpha z} \, e^{\,\jmath(\omega t - \beta z)} \,, \tag{13.107}$$

where $\tilde{v}_+$ is a phasor constant. Equation 13.107 suggests a **traveling wave**. Figure 13.62 shows a typical "snapshot" of the real cosine-dependent part.

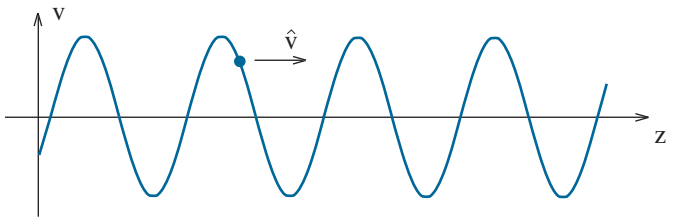


Figure 13.62: Forward traveling wave on a lossy transmission line.

Imagine you are a surfer who rides the wave on the indicated blue dot. Your position with respect to the wave is fixed, so $\omega t - \beta z$ is also constant. Then differentiation with respect to t yields

$$\omega - \beta \frac{dz}{dt} = 0 \,. \tag{13.108}$$

In turn, the **phase velocity** is

$$\frac{dz}{dt} = \hat{v} = \frac{\omega}{\beta} \,. \tag{13.109}$$

Subject to practical $R \ll \omega L$ and $G \ll \omega C$, Eq. 13.106 yields

$$\beta \approx \omega\sqrt{LC} \,. \tag{13.110}$$

Thus, $\hat{v} = 1/\sqrt{LC}$ (in agreement with Eq. 13.101).

The practical $R \ll \omega L$ and $G \ll \omega C$ conditions also support a Taylor series expansion of Eq. 13.106 such that

$$\alpha \approx \frac{1}{2}\left(R\sqrt{\frac{C}{L}} + G\sqrt{\frac{L}{C}}\right) \tag{13.111}$$

(see Problem 13.126). Typically, $G \ll R$. Then the α attenuation factor is

$$\alpha \approx \frac{R}{2Z_o}\,. \tag{13.112}$$

The attenuation on a transmission line of length l is proportional to $\exp(-\alpha l)$. In increasingly rare settings, αl is expressed in **Nepers** after John Napier (Neper) (1550-1617), the Scottish mathematician who invented logarithms. The more common measure of attenuation is in decibels (dB) with 1 Neper = 8.686 dB. Attenuations in dB are additive: if a 20-m line has 0.6 dB loss, then a 40-m line has 1.2 dB loss.

The same attenuated traveling-wave behavior applies to current $i_+(z,t)$. Specifically,

$$i(z,t) = \tilde{i}_+\, e^{-\alpha z}\, e^{\jmath(\omega t - \beta z)}\,, \tag{13.113}$$

where $\tilde{i}_+ = \tilde{v}_+/Z_o$ is a phasor constant. We also acknowledge -z-directed traveling waves as valid solutions to the wave equation:

$$v(z,t) = \tilde{v}_-\, e^{+\alpha z}\, e^{\jmath(\omega t + \beta z)}\,, \tag{13.114}$$

$$i(z,t) = \tilde{i}_-\, e^{+\alpha z}\, e^{\jmath(\omega t + \beta z)}\,. \tag{13.115}$$

Nevertheless, we put them aside for the moment, as there is no source for their existence on a semi-infinite or Z_o-terminated transmission line.

Note: Any voltage or current pulse is a superposition of traveling waves at different frequencies. And whereas the wave velocity and the attenuation factor are constant, *pulses are unmodified apart from gradual attenuation.* Corruptions are possible when $\hat{v}$ and α have frequency dependence.

Exercise 13.13 A particular cable has an attenuation of 5.7 dB/100 ft. Determine the percent pulse attenuation when the cable has 240-ft length.

Ans: % attenuation = 21

Exercise 13.14 A particular cable has an attenuation of 28 dB/100 m. Determine R, the cable resistance per meter.

Ans: $R = 3.3\ \Omega/\text{m}$

Example 13.9

Manufacturer's data for RG-119 coaxial cable specifies $Z_o = 50\ \Omega$, $C = 29.4$ pF/ft, and attenuation = 3.8 dB/100 ft. A 40-ft line connects to a 50-Ω load and a signal source with 50-Ω Thevenin resistance as shown in Fig. 13.63. The source pulses between 0 and 5V over 10 ns with 1-ns rise and fall times. Find the pulse characteristics at the load and verify with SPICE.

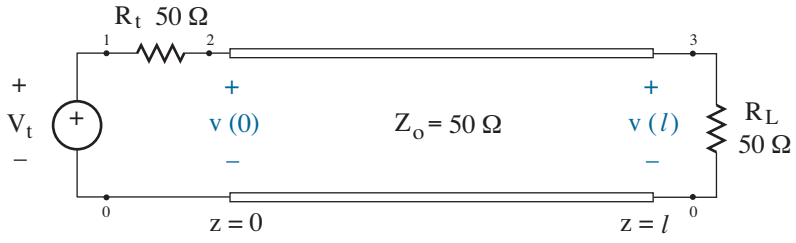


Figure 13.63: Circuit for Example 13.9.

Solution

Since the signal source has Thevenin resistance, the pulse characteristics at $z = 0$ derive from a voltage divider relation. Specifically,

$$v_{max}\,(0) = V_{t,max}\,\frac{Z_o}{Z_o + R_t} = 2.5\ \text{V}\,.$$

The attenuation over 40 ft is $3.8 \times 40/100 = 1.52$ dB, a factor of 0.840. Thus, the pulse amplitude at $z = l$ is $0.840 \times 2.5 = 2.1$ V. The propagation velocity is $\hat{v} = 1/CZ_o = 6.80 \times 10^8$ ft/s, so the line delay is $l/\hat{v} = 59$ ns.

For the SPICE simulation, we have $L = {Z_o}^2 C = 73.5$ nH/ft. Then with $\alpha l = -\ln 0.840 = 0.174$, $\alpha = 4.35 \times 10^{-3}$ ft^{-1} and $R = 2Z_o\alpha = 0.435\ \Omega$/ft. The SPICE code takes the form ...

```
* Lossy-Line Circuit

Vt      1     0     PULSE     (0 5 0 1n 1n 8n 1)
Rt      1     2     50
RL      3     0     50
Tline   2     0     3     0     TMOD     LEN=40

.model  TMOD        TRN   (L=73.5n, C=29.4p, R=0.435, G=0)

.tran   0.1n  100n
.probe

.end
```

The SPICE code has three noteworthy features:

- A circuit element name beginning with "T" followed by a pair of input nodes and and pair of output nodes corresponds to a transmission line. TMOD cites a particular model, and LEN is the length parameter.

- The .model statement provides parametric details for citing elements. TRN is a transmission-line model with parameters L, C, R, and G. The parameters have per-unit-length values subject to the condition that the length dimension is consistent with the LEN specification. (Students who dislike the English unit of feet will want to revise LEN in meters, chains, cubits, light-years or whatever else is palatable.)

- PULSE calls for a trapezoidal voltage waveform governed by a set of seven parenthetical parameters: initial voltage (0), pulsed voltage (5), delay (0), rise time (1n), fall time (1n), pulse width (8n), period (1). The pulse time is the duration of the pulsed voltage level, whereas the sum of the rise, pulse, and fall times is the duration of the complete pulse waveform. The pulse period is the start time for another pulse—it is set to a large value when only a single short pulse has interest.

Figure 13.64 shows the simulation results. Traces v(2) and v(3) apply to the source and load ends of the line, respectively. The pulse amplitude and delay at the load end is in good agreement with hand calculations.

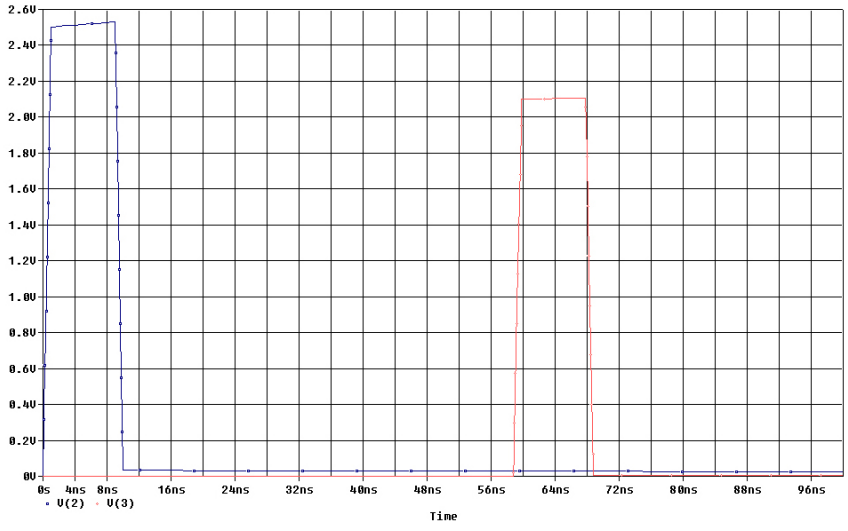


Figure 13.64: SPICE .probe plot for the circuit of Example 13.9.

Reflections

As noted previously, transmission lines support both $+z$ and $-z$-directed traveling waves that move with velocity $\hat{v}$. By superposition, the voltage at any point along the line is given by

$$v(z,t) = v_+ + v_- \,, \tag{13.116}$$

where v_+ and v_- are arbitrary functions of $z - \hat{v}t$ and $z + \hat{v}t$, respectively. With negligible line loss, substitution into Eq. 13.96 ($R = 0$) or Eq. 13.98 ($G = 0$) yields a companion expression for current. Specifically,

$$Z_o\, i(z,t) = v_+ - v_- \,. \tag{13.117}$$

Then at $z = l$, where $v(l,t)/i(l,t) = Z_L$ (a particular load impedance),

$$\Gamma_L = \left.\frac{v_-}{v_+}\right|_{z=l} = \frac{Z_L - Z_o}{Z_L + Z_o} \,. \tag{13.118}$$

Similarly, at $z = 0$, where $v(0,t)/i(0,t) = Z_t$ (a particular source impedance),

$$\Gamma_S = \left.\frac{v_+}{v_-}\right|_{z=0} = \frac{Z_t - Z_o}{Z_t + Z_o} \,. \tag{13.119}$$

Here, Γ_L and Γ_S are **reflection coefficients** that apply at opposite ends of the transmission line.

Reflection coefficients are zero for lines terminating with Z_o.

Non-zero reflection coefficients allow for unruly behavior.

Example 13.10

The transmission line of Fig. 13.65 is subjected to a step input excitation of the form $V_t(t) = 6\,u(t)$. Find $v(0,t)$ and $v(l,t)$, and verify with SPICE. Assume lossless conditions and a transit time of $T = 10$ ns.

Solution

At $t = 0+$, the previously unexcited transmission line looks like an impedance Z_o from the perspective of the Thevenin circuit. Thus,

$$v(0, 0+) = 6 \left(\frac{50}{10 + 50} \right) = 5 \text{ V} .$$

The immediate v_+ signal takes the form of a rectangular wavefront that propagates toward $z = l$ with velocity $\hat{v}$. Alternatively, one can envision a steady procession of 5-V impulses that march rightward from the V_t source. Figure 13.66 shows a snapshot of applicable line conditions for $0 < t < T$. There is no v_- signal as yet.

Whereas Z_L is infinite—it is an open circuit, the reflection coefficient at $z = l$ is $\Gamma_L = (\infty - Z_o)/(\infty + Z_o) = 1$. So the arrival of the v_+ signal at $t = T$ induces a reflected v_- signal with the same wavefront shape and magnitude that moves to the left with velocity $\hat{v}$. Figure 13.67 shows a snapshot of applicable line conditions for $T < t < 2T$. In view of Eq. 13.116, the v_+ and v_- signals add at $z = l$ so that $v(l, t) = 5 + 5 = 10$ V. This value holds over the subsequent interval $2T < t < 3T$ since it takes at least two line transits for the arrival of new information with potential for readjustment.

The reflection coefficient at $z = 0$ is $\Gamma_S = (10 - 50)/(10 + 50) = -2/3$. So the arrival of the v_- signal at $t = 2T$ induces an *incremental* $\Delta v_+ = -2/3 \times 5 = -10/3$ that travels rightward in the form of a step discontinuity. We calculate the new $v(0, t)$ voltage as follows:

$$\begin{aligned} v(0,t) &= \text{old value} + \text{updated } \Delta v_+ + \text{continuing } v_- \\ &= \quad 5 \quad + \quad -10/3 \quad + \quad 5 \quad = 20/3 \text{ V} . \end{aligned}$$

This voltage will hold over the subsequent interval $3T < t < 4T$.

The remaining analysis is best accomplished with the help of a Table. The v_+ and v_- entries are taken as incremental line adjustments, and the $v(0,t)$ and $v(l,t)$ calculations have the "old" + "updated" + "continuing" format used most recently. Over the interval $(0, 6T)$, we have

Time Interval	Δv_+		Δv_-	$v(0,t)$	$v(l,t)$
$0 < t < T$	5		0	5	0
	↓			↓	
$T < t < 2T$	5	→	5	5	0 + 5 + 5
			↓		↓
$2T < t < 3T$	-3.33	←	5	5 - 3.33 + 5	10
	↓			↓	
$3T < t < 4T$	-3.33	→	-3.33	6.67	10 - 3.33 - 3.33
			↓		↓
$4T < t < 5T$	2.22	←	-3.33	6.67 + 2.22 - 3.33	3.33
	↓			↓	
$5T < t < 6T$	2.22	→	2.22	5.56	3.33 + 2.22 + 2.22

The SPICE simulation is straightforward apart from Z0=50 and TD=10n (the delay time) in the transmission-line .model statement and the use of a 1-MΩ resistor in place of the right-side open circuit. Figure 13.68 shows .probe results. Both $v(0,t)$ and $v(l,t)$ eventually settle at 6V for large t. Nevertheless, the transmission line imparts substantial ringing.

In digital systems, unterminated signal sources and loads generally have low- and high-impedance characteristics, respectively. Thus, there are three effective termination arrangements. Figure 13.69a shows a termination with a resistor of value Z_o in series with the Thevenin source so that reflections induced at the load end of the line are totally absorbed at the source end. Figure 13.69b shows a termination at the load so that reflections do not propagate back to the source where they would be reflected once again. Note that the input to the load ties to V_{DD} and ground through separate resistors with value $2Z_o$. While this promotes balanced LOW-to-HIGH and HIGH-to-LOW transition times, the termination is made at the expense of continuous dc power dissipation. Figure 13.69c shows a fail-safe approach that suppresses reflections at both ends of the transmission line.

Analog systems typically feature Z_o terminations at both ends of the transmission line. Since analog circuits are relatively susceptible to noise, resistive terminations tend to be avoided. Inductor/capacitor combinations or transformers offer noiseless terminations.

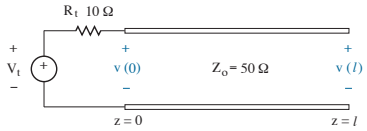


Figure 13.65: Circuit for Example 13.10.

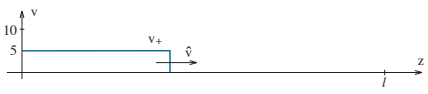


Figure 13.66: Transmission-line conditions for $0 < t < T$ (Example 13.10).

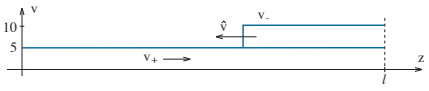


Figure 13.67: Transmission-line conditions for $T < t < 2T$ (Example 13.10).

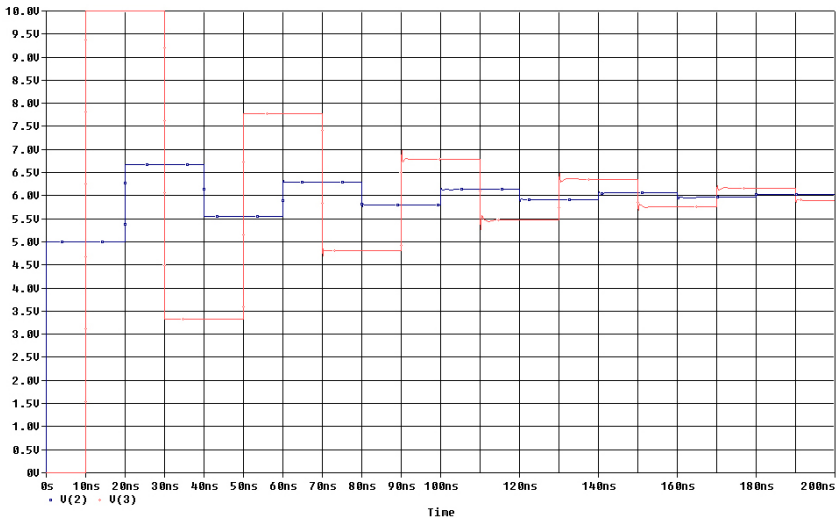


Figure 13.68: SPICE results for Example 13.10.

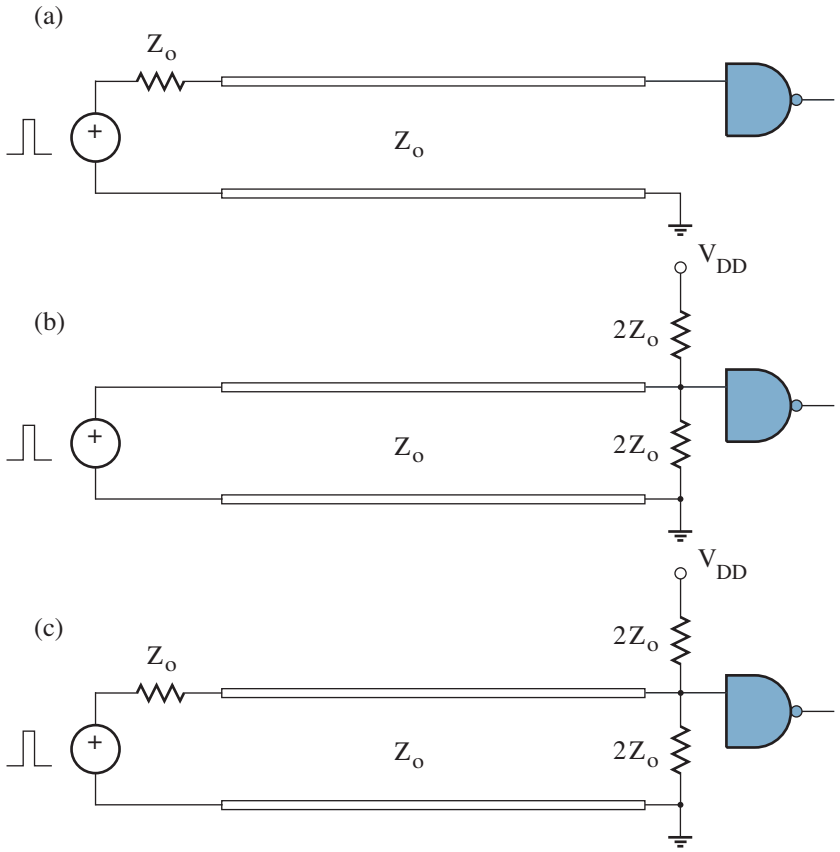


Figure 13.69: Transmission-line terminations for digital systems.

Problems

Section 13.1

13.1 Derive the expressions for ω_o and K shown in Fig. 13.4 for two types of first-order low-pass filter.

13.2 Derive the expressions for ω_o and K shown in Fig. 13.6 for two types of first-order high-pass filter.

13.3 An inverting first-order low-pass filter ($K < 0$) has an ideal op-amp, an inductor, and two resistors.

(a) Show the general form of the circuit design, and determine expressions for ω_o and K.

(b) Complete a design with $f_o = 50$ kHz and 6-dB gain magnitude at dc. Let $L = 10$ mH.

(c) Use SPICE to demonstrate the changes in f_o and K when the inductor of part b has a parasitic series resistance of 5 Ω.

13.4 An inverting first-order high-pass filter ($K < 0$) has an ideal op-amp, an inductor, and two resistors.

(a) Show the general form of the circuit design, and determine expressions for ω_o and K.

(b) Complete a design with $f_o = 20$ kHz and 8-dB gain magnitude in the limit of high frequency. Let $L = 10$ mH.

(c) Use SPICE to demonstrate the changes in f_o and K when the inductor of part b has a parasitic series resistance of 5 Ω.

13.5 Design a first-order active low-pass filter with a response of +18 dB in the dc limit and -18 dB at 25 kHz. Let $C = 220$ nF.

13.6 Design a first-order active high-pass filter with a response of +12 dB in the high-frequency limit and -20 dB at 1.2 kHz. Let $C = 1$ nF.

13.7 Design a first-order active low-pass filter with a response of +16 dB in the dc limit and -60° of phase shift at 6 kHz. Let $C = 1$ nF.

13.8 Consider the second-order low- and high-pass transfer functions of Eqs. 13.8 and 13.9. Find the angular frequency at which $|H|$ is maximum in terms of ω_o and Q. Specify any restrictions on Q.

13.9 Consider the second-order low- and high-pass transfer functions of Eqs. 13.8 and 13.9. Find the angular frequency at which $|H|$ is -3 dB down from the low- and high-frequency limiting values in terms of ω_o and Q.

13.10 An active filter is to be designed such that the maximum passband gain is 12 dB and the bandwidth is 80 kHz in relation to a 50-kHz center frequency. A maximum degree of passband "flatness" is desired.

(a) Find K, ω_o, and Q for a single second-order bandpass filter that satisfies the requirements.

(a) Find K and ω_o values for a cascade of first-order low- and high-pass filters that together satisfy the requirements.

(c) Make an argument for the better design.

13.11 Prove the -3-dB specifications in Eq. 13.12.

13.12 Design an active Butterworth low-pass filter with +16 dB in the dc limit, -40 dB at 20 kHz, and $f_o = 4$ kHz.

13.13 Repeat Problem 13.12 for an active 1-dB-Chebyshev low-pass filter.

13.14 Repeat Problem 13.12 for an active Bessel low-pass filter.

13.15 Show that the substitution $s/\omega_o \rightarrow \omega_o/s$ transforms a second-order high-pass filter characteristic into the low-pass form with unmodified Q.

13.16 Design an active Butterworth high-pass filter with +24 dB in the high-frequency limit, -24 dB at 20 kHz, and $f_o = 100$ kHz.

13.17 Repeat Problem 13.16 for an active 1-dB-Chebyshev high-pass filter.

13.18 Repeat Problem 13.16 for an active Bessel high-pass filter.

13.19 Show that a filter with order $n \geq 3$ is invariant to the sequence of stages.

13.20 Consider a bandpass filter of order $2m$, where m is an integer. The filter is to be realized through a cascade of second-order bandpass filters having the same ω_o and Q values.

(a) In terms of m, determine the rate of amplitude decrease for angular frequencies much less or much greater than ω_o.

(b) In terms of m and Q, determine the cascaded filter bandwidth.

13.21 A first-order **all-pass** filter exhibits a transfer function of the form

$$H(s) = \frac{s - \omega_o}{s + \omega_o}.$$

(a) Demonstrate the all-pass capability in terms of output magnitude.

(b) Determine the frequency dependence of the all-pass phase response.

13.22

(a) Consider a capacitor C with parasitic parallel conductance G. Find the applicable Q factor in terms of C and G if the elements are in a parallel resonator with inductor L.

(b) Let all of the dimensions for C and G in part a scale by a factor x. Make appropriate physical arguments to show how Q scales with x.

(c) Consider an inductor L with parasitic series resistance R. Find the applicable Q factor in terms of L and R if the elements are in a series resonator with capacitor C.

(d) Let all of the dimensions for L and R in part c scale by a factor x. Make appropriate physical arguments to show how Q scales with x.

(e) Discuss the implications for integrated circuits.

Section 13.2

13.23 Figure P13.23 shows a **Salen-Key** second-order low-pass filter with an op-amp (no inductors). The position of the capacitors and their open-circuit behavior at low frequencies clearly supports the low-pass function.

(a) Show that the input-output transfer characteristic has the low-pass form of Eq. 13.8 with $K = 1 + R_3/R_4$,

$$\omega_o = \frac{1}{\sqrt{R_1 R_2 C_1 C_2}},$$

and

$$Q = \frac{1}{(1-K)\sqrt{\frac{R_1 C_1}{R_2 C_2}} + \sqrt{\frac{R_2 C_2}{R_1 C_1}} + \sqrt{\frac{R_1 C_2}{R_2 C_1}}}.$$

(b) Complete a filter design that establishes $f_o = 2$ kHz and $Q = 0.707$ subject to $R_1 = R_2 = R$ and $C_1 = C_2 = 100$ pF. Specify the K value that is required for this design.

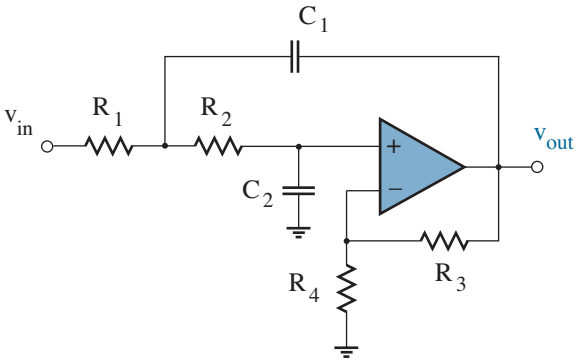


Figure P13.23

13.24 Figure P13.24 shows a Salen-Key second-order high-pass filter with an op-amp (no inductors). The position of the capacitors and their short-circuit behavior at high frequencies clearly supports the high-pass function.

(a) Show that the input-output transfer characteristic has the high-pass form of Eq. 13.8 with $K = 1 + R_3/R_4$,

$$\omega_o = \frac{1}{\sqrt{R_1 R_2 C_1 C_2}},$$

and

$$Q = \frac{1}{(1-K)\sqrt{\frac{R_2 C_2}{R_1 C_1}} + \sqrt{\frac{R_1 C_1}{R_2 C_2}} + \sqrt{\frac{R_1 C_2}{R_2 C_1}}}.$$

(b) Complete a filter design that establishes $f_o = 1$ kHz and $Q = 0.707$ subject to $R_1 = R_2 = R$ and $C_1 = C_2 = 1$ nF. Specify the K value that is required for this design.

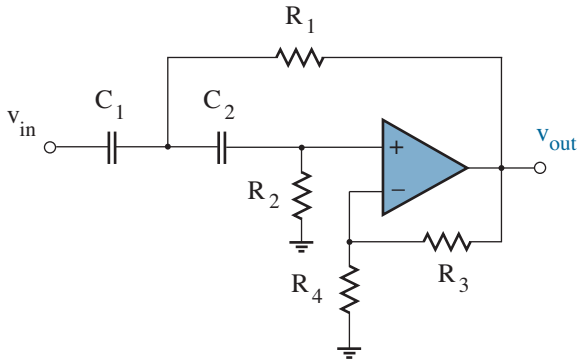


Figure P13.24

13.25 Figure P13.25 shows the general Salen-Key circuit format in terms of admittances Y_1, Y_2, Y_3, and Y_4, and an amplifier with gain K.

(a) Determine the types of admittance needed (resistive or capacitive) to support a bandpass filter characteristic with the form of Eq. 13.10.

(b) Show that the Salen-Key bandpass filter can only be realized if $K < 1$.

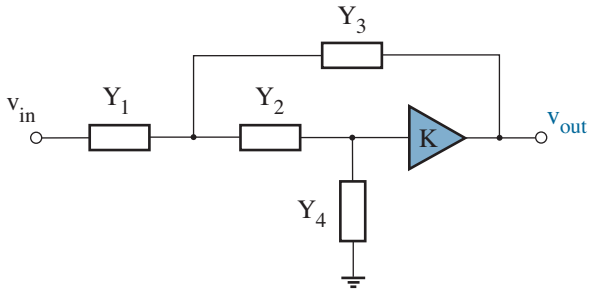


Figure P13.25

13.26 This problem explores the "resonance" that a Salen-Key circuit achieves without an inductor.

Consider the low-pass Salen-Key filter of Problem 13.23 with v_{in} at ground, and break the feedback loop to the right of C_1. For simplicity, let $R_1 = R_2 = R$ and $C_1 = C_2 = C$. These changes yield the circuit of Fig. P13.26. Signal source $v_{in}{}'$ stimulates the circuit at one end of the former loop.

(a) Find the function $H(s)$ that relates v_{out} to $v_{in}{}'$.

(b) Show that an angular frequency exists such that $|H|$ is maximum and $\angle H = 0$, then specify the value of this angular frequency. Conclude that with v_{out} comparable to and in phase with $v_{in}{}'$, closing the feedback loop leads to oscillatory transient behavior.

(c) Provide a similar qualitative argument involving first-order low- and high-pass filters that relate to the circuit of Fig. P13.26.

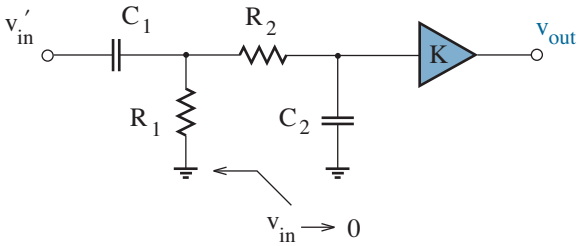


Figure P13.26

13.27 Figure P13.27 has a low-pass **Kundert** filter. It is similar to the low-pass Salen-Key filter apart from the unity-gain buffer to the left of R_2.

(a) Show that the input-output transfer characteristic exhibits the low-pass form of Eq. 13.8, and determine expressions for ω_o, Q, and K.

(b) Complete a filter design that establishes $f_o =$ 2 kHz and $Q = 0.707$ subject to $R_1 = R_2 = R$ and $C_1 = C_2 = 1$ nF.

(c) How do the Kundert design constraints compare to those for the corresponding Salen-Key filter implementation for large Q?

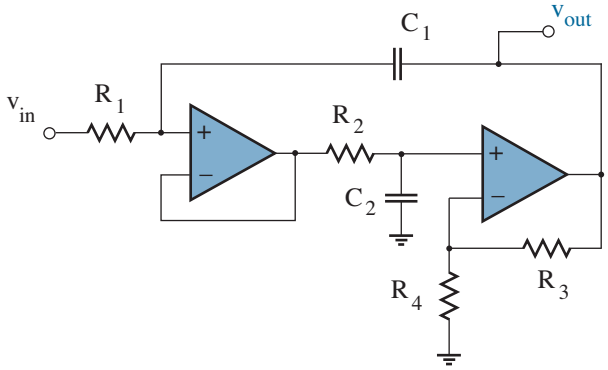


Figure P13.27

13.28 Figure P13.28 has a low-pass **Rausch** filter with multiple-loop feedback. The circuit is relatively insensitive to changes in component values when compared with the Salen-Key low-pass filter.

(a) Show that the input-output transfer characteristic exhibits the low-pass form of Eq. 13.8, and determine expressions for ω_o, Q, and K.

(b) Complete a filter design that establishes $f_o =$ 1 kHz and $Q = 0.707$ subject to $R_1 = R_2 = R$ and $C_1 = 1$ nF.

(c) What component modifications lead to the high-pass characteristic?

13.29 Figure P13.29 shows a band-pass Rausch filter with multiple-loop feedback. Unlike the Salen-Key implementation, the K parameter can exceed unity.

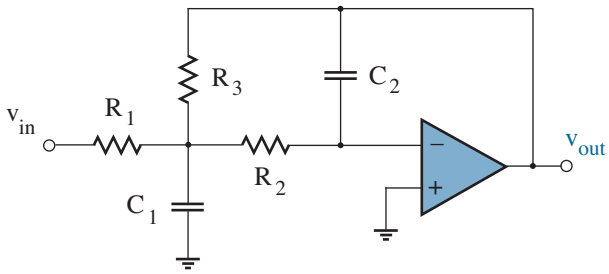


Figure P13.28

(a) Show that the input-output transfer characteristic has the bandpass form of Eq. 13.10 with

$$K = \frac{-R_3}{R_1}\frac{C_1}{C_1 + C_2},$$

$$\omega_o = \sqrt{\frac{1}{R_2 R_3 C_1 C_2}\left(1 + \frac{R_2}{R_1}\right)},$$

and

$$Q = \frac{\sqrt{\frac{R_3}{R_2}\left(1 + \frac{R_2}{R_1}\right)}}{\sqrt{\frac{C_1}{C_2}} + \sqrt{\frac{C_2}{C_1}}}.$$

(b) Complete a filter design that establishes $f_o =$ 1 kHz and $Q = 4$ subject to $R_2 = R_3 = R$ and $C_1 = C_2 = 100$ pF. Specify the K value that is required, and use SPICE to verify your design.

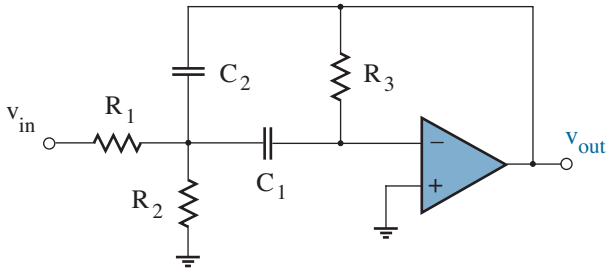


Figure P13.29

13.30 A **Frequency-Dependent Negative Resistor** (FDNR) is a two-terminal element that exhibits a current-voltage relation of the form

$$v = \frac{D}{(j\omega)^2}\, i$$

(see Fig. P13.30a).

(a) Show that the Riordan circuit of Fig. P13.30b satisfies the FDNR relationship and evaluate D. (See Problem 1.43.)

(b) Show that any second-order low-pass, bandpass, or high-pass filter can be realized with an appropriate FDNR, capacitor, and resistor. Hint: Apply a **Bruton transformation** that divides the numerator and the denominator of the filter transfer characteristic by s^2.

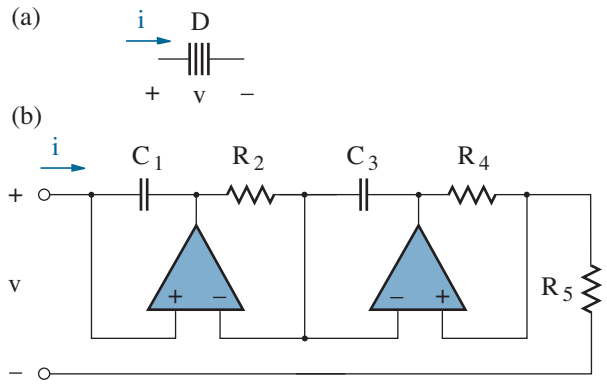


Figure P13.30

13.31 A second-order low-pass filter is to be designed with $f_o = 50$ kHz and $Q = 1/\sqrt{2}$.

(a) Use the results of Problem 13.30 to implement the filter using a FDNR circuit ($C_1 = C_3 = C$, $R_2 = R_4 = R_5 = 1$ kΩ). Use SPICE to verify your design. Assume ideal op-amps.

(b) Use the results of Problem 13.23 to design the filter in the Salen-Key form. Verify with SPICE. Assume an ideal op-amp.

(c) Use SPICE to compare the preceding designs when all resistors and capacitors are subject to *random* 5-% variation. (See Example 7.11.)

13.32 Repeat Problem 13.31 for a second-order high-pass filter with $f_o = 50$ kHz and $Q = 1/\sqrt{2}$.

13.33 Show that the two-op-amp filter circuit of Fig. P13.33 has the second-order bandpass form with $\omega_o = 1/R_2C$, $Q = R_1/R_2$, and $K = 2$.

(This filter is easily tuned.)

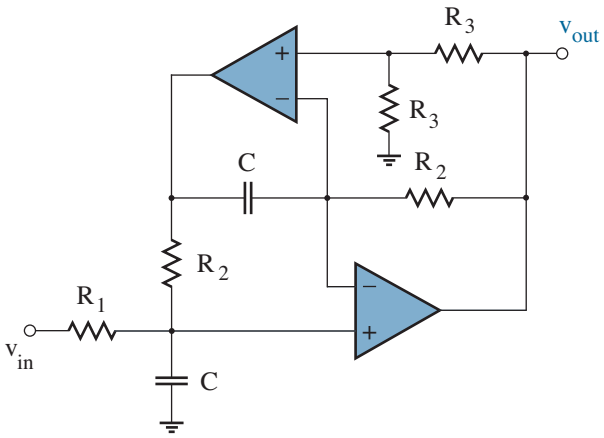


Figure P13.33

13.34 Consider the passive **Twin-T** bandstop (notch) filter of Fig. P13.34.

(a) Show that $\omega_o = 1/RC$, $Q = 1/4$, and $K = 1$.

(b) Let the circuit be subject to positive feedback so that node voltage v_x is $\alpha\, v_{out}$, where $\alpha \leq 1$. Show that ω_o is unchanged, but $Q = 1/4(1-\alpha)$.

(c) Use one or more ideal op-amps and the Twin-T network to design a bandstop filter for which $f_o = 20$ kHz and $Q = 20$. Verify with SPICE.

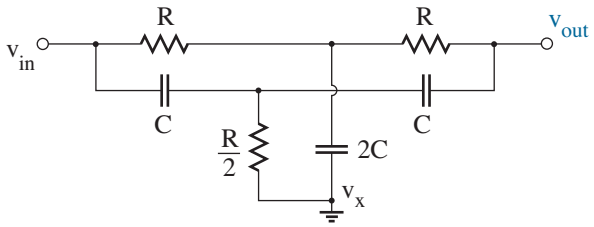


Figure P13.34

13.35 The **Bainter** bandstop (notch) filter of Fig. P13.35 has the characteristic

$$\frac{v_{out}}{v_{in}} = \frac{K\left(s^2 + {\omega_z}^2\right)}{s^2 + s\omega_o/Q + {\omega_o}^2} .$$

The circuit is called a low-pass bandstop filter for $\omega_o < \omega_z$, a high-pass bandstop filter for $\omega_o > \omega_z$, and an ordinary bandstop filter for $\omega_o = \omega_z$.

(a) Show that $\omega_z = \sqrt{\alpha_1}/RC$, $\omega_o = \sqrt{\alpha_2}/RC$, $Q = \sqrt{\alpha_2}/2$, and $K = \alpha_2$, where $\alpha_1 = R_1/R_2$ and $\alpha_2 = 1 + R_3/R_4$.

(b) Complete a design for which $f_z = f_o = 60$ Hz and $Q = 20$ subject to $C = 1\ \mu$F and $R_2 = R_4 = 1$ kΩ. Use SPICE to verify your design.

(c) Use SPICE to demonstrate the change in the filter behavior when f_z is doubled (low-pass) or halved (high-pass).

13.36 The **Boctor** bandstop (notch) filter of Fig. P13.36 has the characteristic

$$\frac{v_{out}}{v_{in}} = \frac{K\left(s^2 + {\omega_z}^2\right)}{s^2 + s\omega_o/Q + {\omega_o}^2} .$$

The one-op-amp circuit is a high-pass bandstop filter ($\omega_o < \omega_z$).

Determine expressions for ω_z, ω_o, Q, and K.

13.37 A particular system is characterized by the following equations:

$$\begin{aligned} x_2 &= 2\,x_1 \\ x_3 &= x_1 + x_2 \\ x_4 &= 3\,x_3 + 2\,x_4 \\ x_5 &= x_2 + x_4 \end{aligned}$$

(a) Construct an appropriate signal flow graph.

(b) Determine an expression that relates x_5 to x_1.

13.38 Find a relation between x_5 and x_1 in the signal flow graph of Fig. P13.38.

13.39 Find a relation between x_5 and x_1 in the signal flow graph of Fig. P13.39.

13.40 In the development of the signal flow graph of Fig. 13.21, it was argued that scaling v_a by $1/Q$ and feeding back to v_b was not a viable option because the summation process at v_b would involve two different types of operation.

(a) Show that the circuits of Fig. P13.40a and Fig. P13.40b implement scaled sums and integrating sums, respectively.

(b) Demonstrate the relative complexity of a circuit that implements a mixed summation process in which

$$v_b = \frac{1}{Q}v_a + \left(\frac{-\omega_o}{s}\right)v_c .$$

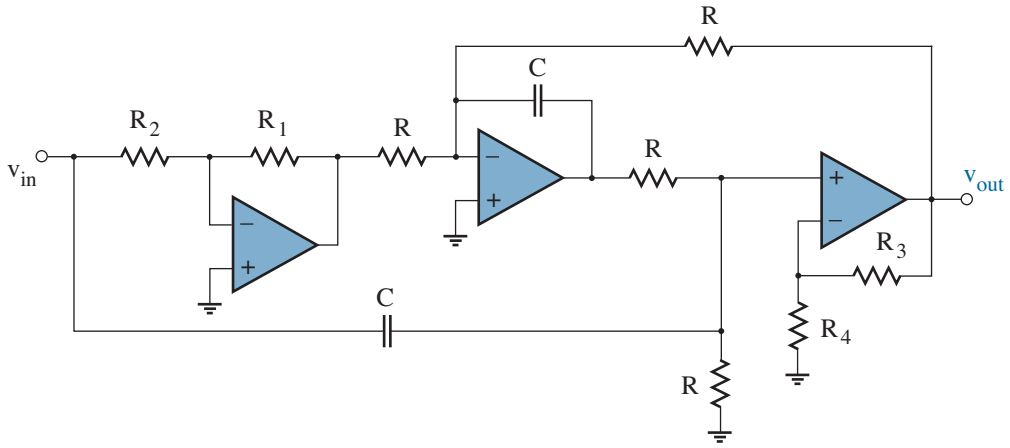


Figure P13.35

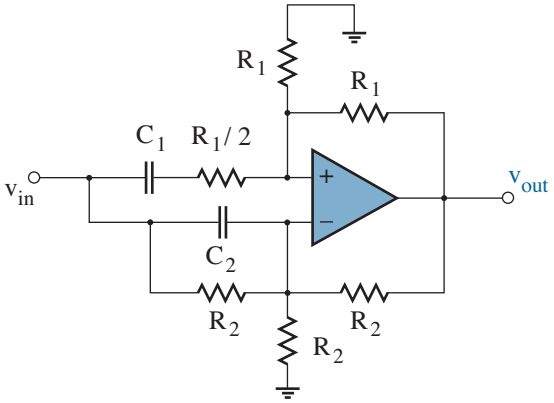


Figure P13.36

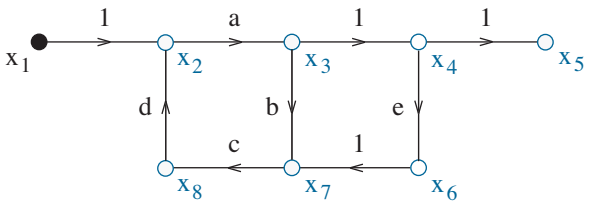


Figure P13.38

13.41 Prove that the outputs at nodes v_b and v_c in the signal flow graph of Fig. 13.21c reflect bandpass and high-pass filter characteristics, respectively.

13.42 Download the datasheet for the UAF42 from Texas Instruments (www.ti.com), then complete a second-order bandpass design that establishes $f_o = 5$ kHz and $Q = 10$.

13.43 Download the datasheet for the MAX274 from Maxim Integrated Products (www.maxim-ic.com). Draw the signal flow graph that applies to the circuit on page 11, and show that it provides the low-pass, bandpass, and high-pass filter functions.

13.44 Download the datasheet for the MAX274 from Maxim Integrated Products (www.maxim-ic.com), then complete a second-order low-pass design that establishes $f_o = 10$ kHz and $Q = 1/\sqrt{2}$.

13.45 Derive a signal flow graph that implements a first-order low-pass filter characteristic using integration ($-\omega_o/s$) and other operations.

13.46 Derive a signal flow graph that implements a first-order high-pass filter characteristic using integration ($-\omega_o/s$) and other operations.

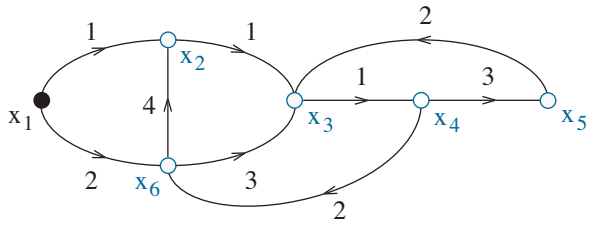


Figure P13.39

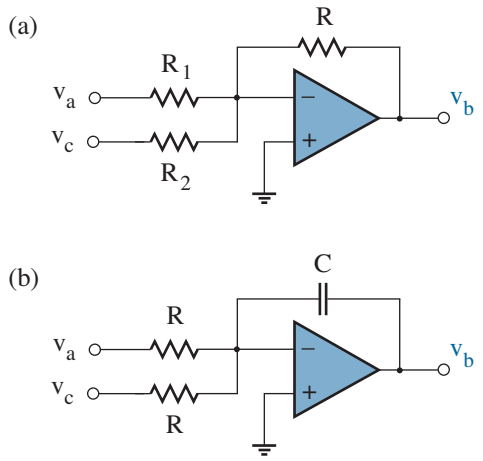


Figure P13.40

13.47 Derive a signal flow graph that implements the first-order all-pass filter characteristic

$$H(s) = \frac{s - \omega_o}{s + \omega_o}$$

using integration $(-\omega_o/s)$ and other operations.

13.48 Integrated-circuit capacitors generally have parasitic capacitance between each plate and the substrate (ground) as shown in Fig. P13.48a.

The switched-capacitor circuit of Fig. P13.48b has switch pairs controlled by clock signals ϕ_1 and ϕ_2. When one clock signal is HIGH, the other is LOW. Let n denote the completion of n clock cycles involving switch pairs 1 and 2 in succession. Show that the circuit implements the integration function through the relation

$$v_{out}(n) = v_{out}(n-1) - \frac{C_1}{C_2} v_{in}(n)$$

with no parasitic capacitive influence.

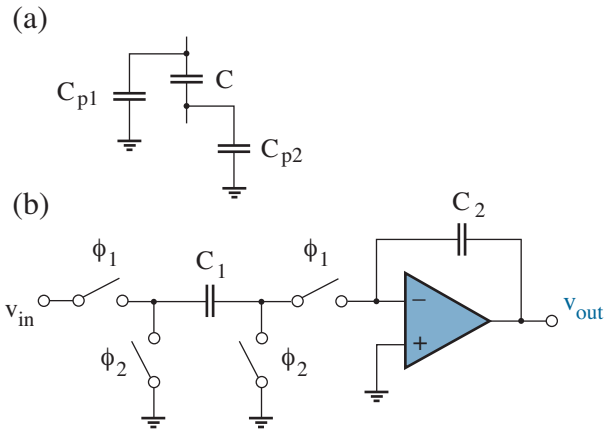


Figure P13.48

13.49 Use the integrator of Fig. P13.48 to implement a low-pass filter with the characteristic

$$\frac{v_{out}}{v_{in}} = \frac{-\omega_o}{s + \omega_o}.$$

Identify (and remove) any redundant switches in your design.

13.50 Use the integrator of Fig. P13.48 to implement a high-pass filter with the characteristic

$$\frac{v_{out}}{v_{in}} = \frac{-s}{s+\omega_o}\,.$$

Identify (and remove) any redundant switches in your design.

Hint: You may want to consider the operation that results when the ϕ_1 switches are replaced with short circuits and the ϕ_2 switches are removed.

13.51 Show that the G_m-C circuit of Fig. P13.51 implements the first-order low-pass filter function, and identify ω_o and K.

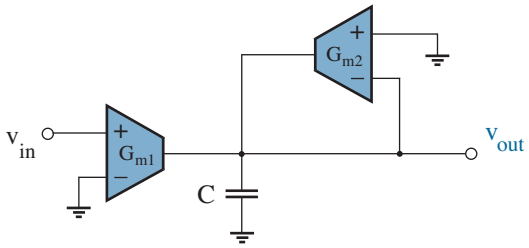


Figure P13.51

13.52 Show that the G_m-C circuit of Fig. P13.52 implements the first-order high-pass filter function, and identify ω_o and K.

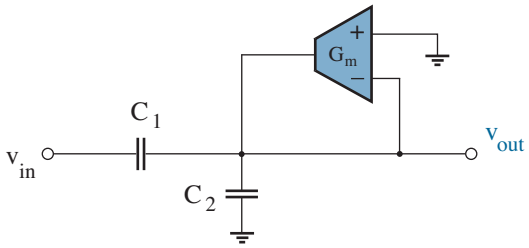


Figure P13.52

Section 13.3

13.53 Specify the bit transmission sequence for the character string R$ in an E-8-2 framing process. Compare with an N-8-1 process.

13.54 Use half-circuit analysis to verify Eqs. 13.38 and 13.39 for the differential amplifier of Fig. 13.37.

13.55 Find the common-mode rejection ratio for the differential amplifier of Fig. 13.37 when the feedback resistor for op-amp 1 is $R_f + \Delta R_f$.

13.56 Let the differential amplifier of Fig. 13.37 have matched R_f resistors but unmatched op-amp common-mode rejection ratios (CMRRs). Find the overall CMRR, and show that it becomes infinite in the limit of matched CMRR values.

13.57 Let the op-amps in the differential amplifier of Fig. 13.37 have 20-MHz gain-bandwidth product.

(a) Complete a design with a differential gain of 10. Let $R_x = 10$ kΩ.

(b) Use SPICE to find the overall bandwidth.

13.58 The op-amps in the differential amplifier of Fig. 13.37 have ±10-V supplies. Complete a design with a differential gain of 40, and specify the input range for v_{i1} when $v_{i2} = 0$. Let $R_x = 10$ kΩ.

13.59 The op-amps in the differential amplifier of Fig. 13.38 have ±10-V supplies. Complete a design with a differential gain of 40, and specify the input range for v_{i1} when $v_{i2} = 0$. Let $R_x = 10$ kΩ.

13.60 The fully differential op-amp of Fig. 13.39 has ±10-V supplies, and $v_{ocm} = 2$ V. Complete a design with a differential gain of -40, and specify the input range for v_{i1} when $v_{i2} = 0$.

13.61 Find the common-mode rejection ratio for the fully differential amplifier of Fig. 13.39 when the feedback resistor connecting to the non-inverting input is $R_f + \Delta R_f$.

13.62 Find $(v_{o1} - v_{o2})/v_{in}$ for the fully differential circuit of Fig. P13.62.

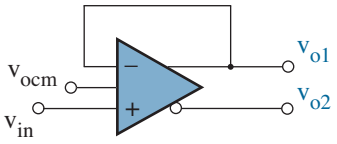


Figure P13.62

13.63 Figure P13.63 has a fully differential amplifier connecting to a differential signal source with 50-Ω Thevenin resistance.

(a) Find R_x so that the resistance looking into the amplifier is 50 Ω.

(b) Find R_f so that $A_{dd} = -4$.

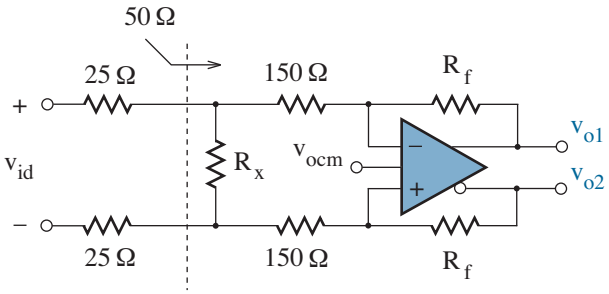


Figure P13.63

13.64 Figure P13.64 shows a fully differential first-order low-pass filter. Find expressions for the cutoff frequency and the pass-band gain.

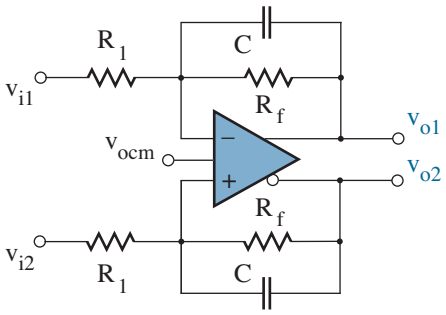


Figure P13.64

13.65 Figure P13.65 shows a fully differential second-order low-pass filter. Find expressions for the cutoff frequency, the pass-band gain, and Q.

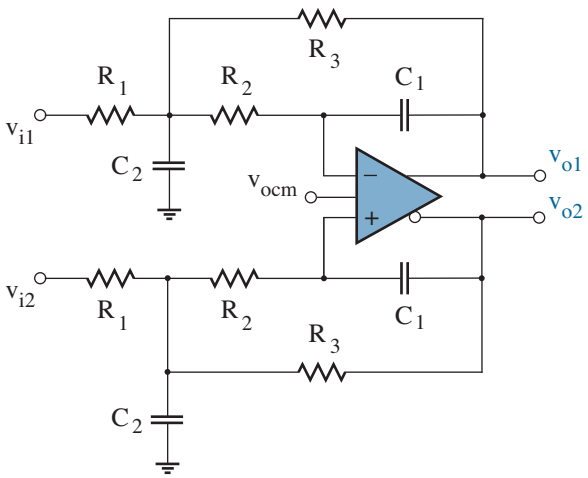


Figure P13.65

13.66 A differential amplifier having a single-ended input is to be realized through the cascade of two inverting op-amp amplifiers.

(a) Show that the differential gain must be unity, and complete a design with 10-kΩ resistors.

(b) Each op-amp features 20-MHz gain-bandwidth. Use SPICE to evaluate the relative signal delays at both sides of the differential output.

13.67 A differential amplifier having a single-ended input is to be realized through the parallel action of non-inverting and inverting op-amp amplifiers.

(a) Complete a design with $A_{vd} = 5$ and 100-kΩ feedback resistors between the op-amp outputs and the corresponding inverting terminals.

(b) Each op-amp features 20-MHz gain-bandwidth. Use SPICE to evaluate the relative signal delays at both sides of the differential output.

(c) Show how the non-inverting side of the circuit can be modified to achieve equal signal delays.

Section 13.4

13.68 A sinusoidal signal with 10-mV peak-to-peak amplitude is applied to an amplifier with a power gain of 12 dB. The amplifier loading factor is 0.8. Find the dBm power absorbed by a 50-Ω load that connects to the amplifier output.

13.69 Justify the calculation of the -51-dBm total noise power that applies to the spectrum analyzer measurement of Fig. 13.42.

13.70 Figure P13.70 has a spectrum-analyzer display that applies to the output of a noisy amplifier with no input signal. The resolution bandwidth is 5 kHz. Estimate the total output noise power.

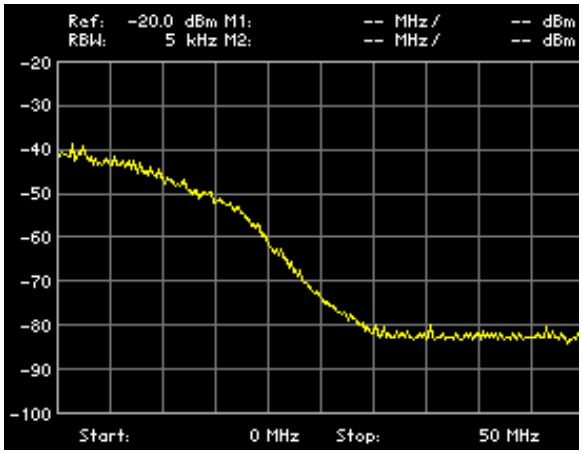


Figure P13.70

13.71 A 68-kΩ resistor yields an rms noise voltage of 0.155 mV subject to a 20-MHz noise bandwidth. Determine the measurement temperature.

13.72 A resistive circuit connects to a *noiseless* load (with real Z_L) so that the noise power delivered to the load is maximized. Show that the noise power is

$$P_n = kTB_n$$

regardless of the Thevenin source resistance.

13.73 Find the output rms noise voltage for the circuit of Fig. P13.73 at $T_o = 290$ K. $B_n = 100$ MHz.

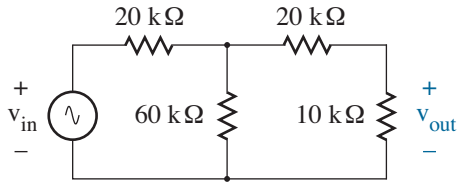


Figure P13.73

13.74 Find the output rms noise voltage for the circuit of Fig. P13.74 at $T_o = 290$ K.

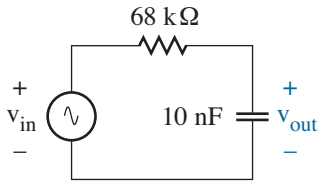


Figure P13.74

13.75 Find the output rms noise voltage for the circuit of Fig. P13.75 at $T_o = 290$ K.

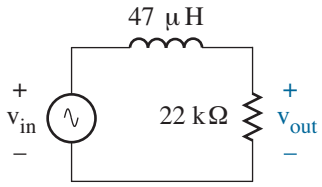


Figure P13.75

13.76 Find the output rms noise voltage for the circuit of Fig. P13.76 at $T_o = 290$ K. Hint: $Q = 1/\sqrt{2}$.

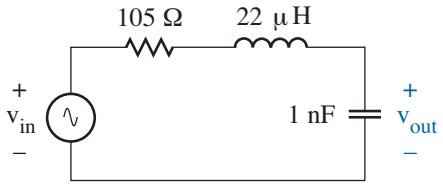


Figure P13.76

13.77 Show that a tuned amplifier with bandwidth B has a noise bandwidth $B_n = \pi B/2$.

13.78 Find the equivalent noise sources e_n and i_n relating to the circuit of Fig. P13.78 at $T_o = 290$ K.

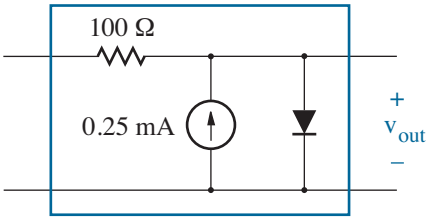


Figure P13.78

13.79 Figure P13.79 shows an op-amp noise model with equivalent noise sources e_{n+}, i_{n+} and e_{n-}, i_{n-} at non-inverting and inverting inputs, respectively. The op-amp has noiseless input resistance r_{in}.

(a) How are the noise sources related?

(b) Show how e_{n+}, i_{n+} are modified with e_{n-}, i_{n-} turned off to achieve the same noise behavior.

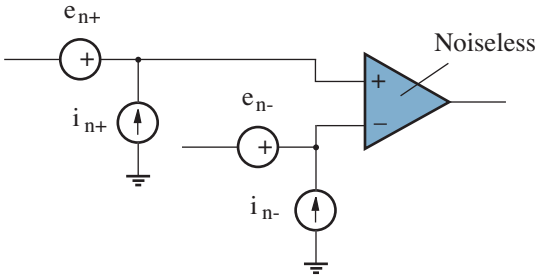


Figure P13.79

13.80 Show that Eqs. 13.69 and 13.70 are equivalent expressions for the average noise factor.

13.81 A particular amplifier with 12-dB power gain and 8-dB noise figure connects to a -20-dBm signal source with 50-Ω Thevenin resistance. ($T_o = 290$ K)

(a) Find the total output noise power.

(b) Find the signal-to-noise ratio at the input.

(c) Find the signal-to-noise ratio at the output.

(d) Find the output noise power from the amplifier when the signal source is disconnected.

13.82 This problem concerns the optimum source admittance that minimizes amplifier noise factor.

Figure P13.82 shows a two-port amplifier model with voltage gain A_v, input admittance Y_i, and equivalent noise sources e_n and i_n. The Norton equivalent of the signal source has admittance Y_s with an associated noise current source i_{sn}

(a) Show that the amplifier noise factor is given by

$$F = 1 + \frac{\overline{(i_n + Y_s\, e_n)^2}}{\overline{i_{sn}^{\,2}}}\,.$$

(b) The e_n and i_n sources may be correlated. So let

$$i_n = i_c + i_u = Y_e\, e_n + i_u\,.$$

Here, Y_e is a **noise correlation admittance**.

All of the admittances can be expressed in the form $Y = G + \jmath B$, where G is a conductance and B is a **susceptance**. Assume $\overline{e_n^{\,2}} = 4kTR_nB_n$, $\overline{i_u^{\,2}} = 4kTG_uB_n$, and $\overline{i_{sn}^{\,2}} = 4kTG_sB_n$. Prove

$$F = 1 + \frac{G_u + [(G_e + G_s)^2 + (B_e + B_s)^2]\, R_n}{G_s}\,.$$

(c) Find G_{so} and B_{so} values that minimize F.

(d) Demonstrate that curves of constant noise factor satisfy the relation

$$(G_s - G_{so})^2 + (B_s - B_{so})^2 = \left(\frac{G_s}{R_n}\right)(F - F_{min}).$$

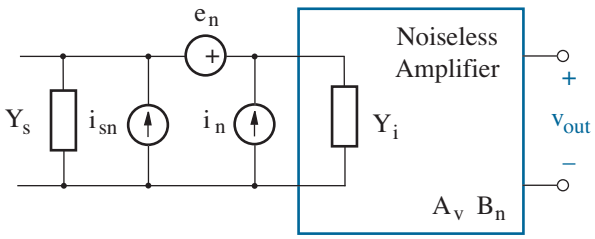


Figure P13.82

13.83 The two-port of Fig. P13.78 makes an input (left-side) connection to a signal source with 100-Ω Thevenin resistance. Determine the noise figure.

13.84 Consider a cascade of three two-ports, each with power gain $\mathcal{G}_i$ and noise factor F_i $(i = 1, 2, 3)$. The input signal source has Thevenin resistance R_t. Show that Friis' formula (Eq. 13.72) applies.

13.85 An amplifier with power gain $\mathcal{G}$ and noise factor F can be described with a **noise measure**

$$M = \frac{F - 1}{1 - \dfrac{1}{\mathcal{G}}} .$$

Show that a cascade of two amplifiers has optimum noise factor when the amplifier with the smaller noise measure is first in the cascade.

13.86 An amplifier has a noise temperature of 80 K. Determine the corresponding noise figure.

13.87 Use Friis' formula (Eq. 13.72) to show that noise temperatures are additive in cascaded systems.

13.88 The op-amp in the circuit of Fig. P13.88 has $e_n = 50$ nV/$\sqrt{\text{Hz}}$, $i_n = 5$ fA/$\sqrt{\text{Hz}}$, and a 10-MHz gain bandwidth product. e_n and i_n are uncorrelated.

(a) Determine the output rms noise voltage and the amplifier noise figure. $(T_o = 290$ K$)$

(b) Repeat part a, but increase both resistors by a factor of ten.

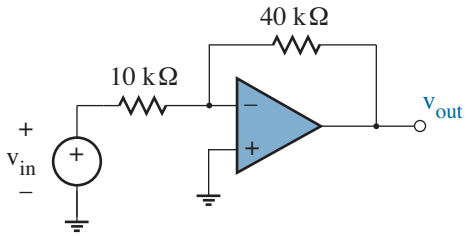


Figure P13.88

13.89 The op-amp in the circuit of Fig. P13.89a has $e_n = 80$ nV/$\sqrt{\text{Hz}}$, $i_n = 0.5$ pA/$\sqrt{\text{Hz}}$, and a 5-MHz gain bandwidth product. e_n and i_n are uncorrelated.

(a) Determine the output rms noise voltage and the amplifier noise figure. $(T_o = 290$ K$)$

(b) Redesign the amplifier to exhibit the same voltage gain, but replace the feedback resistor with a "T" network (Fig. P13.89b). Specify the revised noise figure.

(c) Repeat part b, but let $R_x = 10$ kΩ.

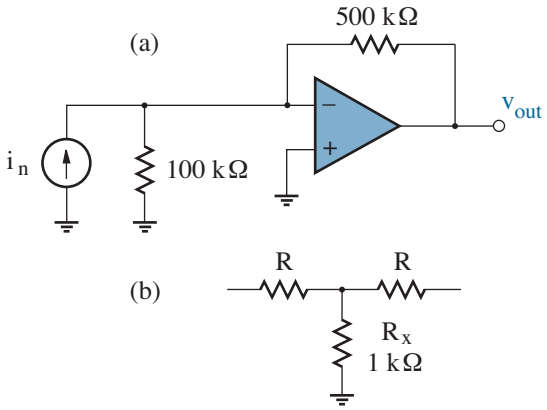


Figure P13.89

13.90 The op-amp in the circuit of Fig. P13.90a has $e_n = 65$ nV/$\sqrt{\text{Hz}}$, $i_n = 120$ fA/$\sqrt{\text{Hz}}$, and a 20-MHz gain bandwidth product. e_n and i_n are uncorrelated.

(a) Determine the output rms noise voltage and the amplifier noise figure. $(T_o = 290$ K$)$

(b) Repeat part a, but increase all of the resistors by a factor of ten.

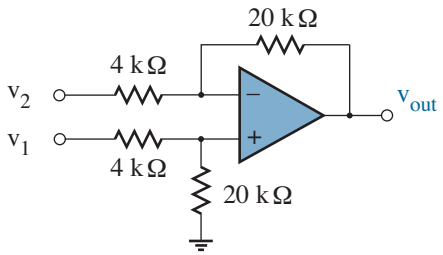


Figure P13.90

13.91 The op-amp in the circuit of Fig. P13.91a has $e_n = 35$ nV/$\sqrt{\text{Hz}}$, $i_n = 470$ fA/$\sqrt{\text{Hz}}$, and a 2-MHz gain bandwidth product. e_n and i_n are uncorrelated. (The circuit is a dependent current source for R_L.)

(a) Find the output rms noise voltage. ($T_o = 290$ K)

(b) Determine the Norton equivalent of the circuit connected to R_L. Assume $A_{vd} = 100,000$, $r_{in} = 100$ MΩ, and $r_{out} = 10$ Ω for the op-amp.

(c) Compare the noise voltage of part a with that induced by the Norton resistance of part b.

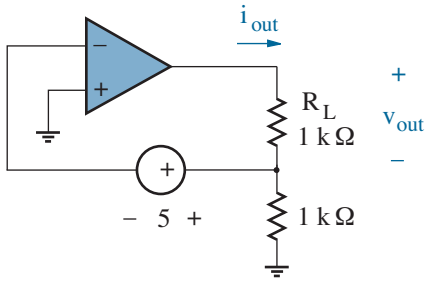


Figure P13.91

13.92 The BJT in the biasing circuit of Fig. P13.92 has $\beta_F = \beta_o = 100$. A signal source with 50-Ω Thevenin resistance capacitively couples to the base ($C_c \to \infty$) to form a common-emitter amplifier with output at the BJT collector. Find the output rms noise voltage and the amplifier noise figure subject to $f_h = 1$ MHz. Assume $f_l \approx 0$ Hz and $T_o = 290$ K.

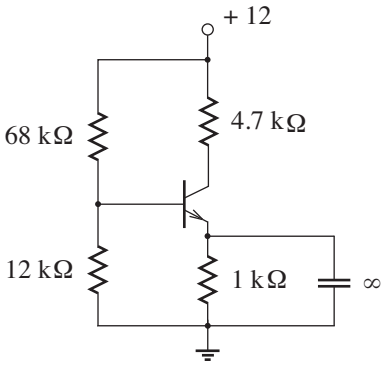


Figure P13.92

13.93 Repeat Problem 13.92, but take the output at the BJT emitter as in a common-collector amplifier. Remove the emitter bypass capacitor and replace the 4.7-kΩ resistor with a short circuit.

13.94 Repeat Problem 13.92, but capacitively couple the Thevenin source to the emitter ($C_c \to \infty$) to form a common-base amplifier with BJT collector output. Provide an ac ground at the base ($C_b \to \infty$).

13.95 The MOSFET in the circuit of Fig. P13.95 has $K'W/L = 2$ mA/V^2 and $V_T = 1$ V. A signal source with 50-Ω Thevenin resistance capacitively couples to the gate ($C_c \to \infty$) to realize a common-source amplifier with output at the drain. Find the output rms noise voltage and the amplifier noise figure when $f_h = 1$ MHz. Assume $f_l \approx 0$ Hz and $T_o = 290$ K.

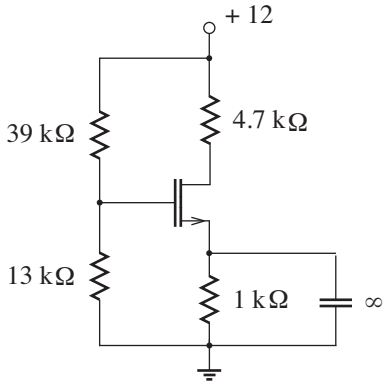


Figure P13.95

13.96 Repeat Problem 13.95, but take the output at the MOSFET source as in a common-drain amplifier. Remove the source bypass capacitor and replace the 4.7-kΩ resistor with a short circuit.

13.97 Repeat Problem 13.95, but capacitively couple the input circuit to the MOSFET source ($C_c \to \infty$) to form a common-gate amplifier. Provide an ac ground at the gate ($C_b \to \infty$).

13.98 The MOSFETs in the differential amplifier of Fig. P13.98 have $K'W/L = 4$ mA/V^2, $V_T = 0.5$ V. Determine the equivalent noise sources e_n and i_n at the input when $T_o = 290$ K. Neglect $1/f$ noise.

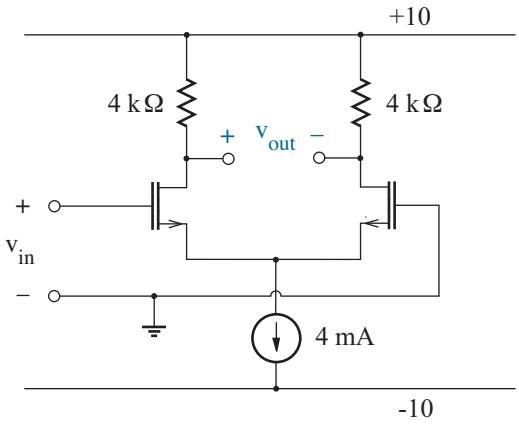


Figure P13.98

13.99 The MOSFETs in the differential amplifier of Fig. P13.83 have $K' = 50$ μA/V^2, $V_T = 0.5$ V, $W = 20$ μm, $L = 0.5$ μm, $t_{ox} = 40$ nm, $K_f = 10^{-25}$ V^2-F. Assume 20-MHz noise bandwidth and $T_o = 290$ K.

(a) Let $v_{in} = 0$. At what "corner" frequency does the $1/f$ noise-density contribution at the output equal the contributions from other sources?

(b) Most listeners will be unaffected by audio noise below 10 Hz. Find the total rms output noise voltage subject to this lower frequency limit.

13.100 Experiments show that diodes have additive $1/f$ noise density according to the relation

$$\overline{i_{nx}^2} = 2q\, i_d|_Q \Delta f + K_f\, i_d|_Q \frac{\Delta f}{f} .$$

A diode exhibits a "corner" frequency of 3 kHz below which $1/f$ noise is dominant. Determine K_f.

Note: BJTs display $1/f$ (flicker) noise in their base current with similar functionality. The mean-square level is proportional to the emitter junction area. The collector current has no $1/f$ noise contribution.

13.101 An integrated polysilicon resistor features additive $1/f$ noise density according to the relation

$$\overline{i_n^2} = 4kTG\Delta f + K_f\, i|_Q{}^2 \frac{\Delta f}{f^\gamma} ,$$

where typically $0.8 < \gamma < 1.1$. Parameter K_f varies with geometry, and it is highly process dependent. Consider a 4-kΩ resistor for which $K_f = 5 \times 10^{-11}$ and $\gamma = 1$. Find the maximum bias current so that $1/f$ and thermal noise contribute equally to overall noise at a 10-Hz "corner" frequency. ($T_o = 290$ K)

Note: Discrete carbon resistors have the greatest K_f values and $1/f$ noise levels.

13.102

(a) Use SPICE to find the output noise level for the circuit of Problem 13.88. ($T_o = 290$ K)

(b) Show that the inclusion of a capacitor in parallel with the 40-kΩ resistor can reduce the output noise level when the useful bandwidth is 20 kHz (as for an audio application).

13.103 Use SPICE to find the output noise level for the circuit of Problem 13.90. ($T_o = 290$ K)

13.104 Use SPICE to find the output noise level for the circuit of Problem 13.92. ($T_o = 290$ K)

13.105 The MOSFET in the circuit of Fig. P13.105 is biased so that $g_m = 4 \times 10^{-3}$ ℧ and $C_{gs} = 0.2$ pF. Find a value for L so that the small-signal resistance looking into the MOSFET gate is 50 Ω at 100 MHz.

Note: This impedance match does not add noise.

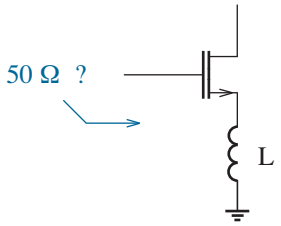


Figure P13.105

Section 13.5

13.106 Find the total harmonic distortion applicable to the spectrum analyzer display of Fig. P13.106.

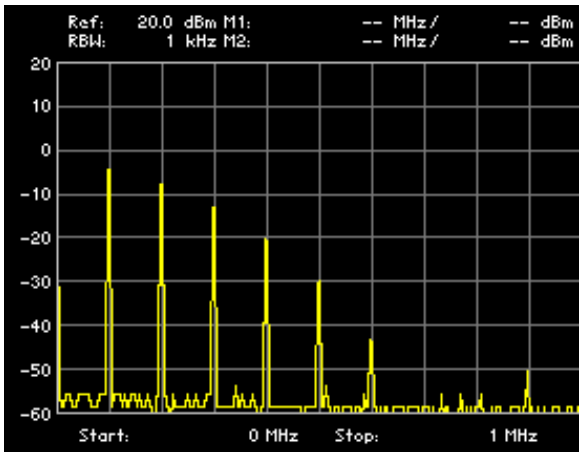


Figure P13.106

13.107 Find the 1-dB compression point that applies to the following input-output power data (dBm):

P_{in}	P_{out}	P_{in}	P_{out}	P_{in}	P_{out}
-40.0	-21.4	-25.0	-8.5	-15.0	-4.6
-35.0	-16.4	-22.5	-7.1	-12.5	-4.3
-30.0	-11.9	-20.0	-5.9	-10.0	-4.0
-27.5	-9.9	-17.5	-5.2	-7.5	-3.7

13.108 A 100-kHz signal drives a certain amplifier. Find the OIP2_H intercept point that applies to the following output power data (dBm):

P_{out} 100 kHz	P_{out} 200 kHz	P_{out} 100 kHz	P_{out} 200 kHz
-16.4	-38.4	-5.9	-12.7
-11.9	-27.2	-4.6	-8.8
-8.5	-19.4	-4.0	-6.3

13.109 An amplifier is characterized with $\text{OIP2}_H = 8$ dBm and $\text{OIP3}_H = 14$ dBm. Find the power levels of the second and third harmonics if the fundamental output power is -12 dBm.

13.110 Prove the validity of the amplitude given for the signal component at ω_1 in Table 13.6.

13.111 Use SPICE to find a 1-dB compression point for the common-emitter amplifier of Problem 13.92.

13.112 Use SPICE to find OIP2_H for the common-source amplifier of Problem 13.95.

13.113

(a) Starting with the elementary expression for the MOSFET drain current ($\lambda = 0$), show what is meant by a "small" signal in relation to $i_d|_Q$.

(b) Starting with the v_{be}-dependent expression for the BJT collector current, redefine "small."

13.114

(a) Change the collector resistor in the circuit of Problem 13.92 so that the voltage gain is -20 for a common-emitter amplifier. Assume $R_t \approx 0$. Demonstrate with SPICE.

(b) Apply SPICE to find total harmonic distortion when the common-emitter amplifier has 0.5-V peak-to-peak output magnitude.

(c) Redesign the circuit of Problem 13.95 so that the voltage gain is -20 for a common-source amplifier with the same quiescent bias current of part a. Assume $R_t \approx 0$, and demonstrate with SPICE. (You are free to scale $K'W/L$, but not V_T.)

(d) Apply SPICE to find total harmonic distortion when the common-source amplifier has 0.5-V peak-to-peak output magnitude. Make sure that the output goes no closer to the power-supply rail than in part b.

13.115 Use SPICE to find total harmonic distortion for the differential amplifier of Problem 13.98 when the output has 1-V peak-to-peak amplitude.

13.116 Repeat Problem 13.115 to show the effect of a small resistor in series with each MOSFET source. Be sure to maintain amplifier symmetry, and make comparisons subject to constant output magnitude.

13.117 A 10-kHz input signal with 2-V peak-to-peak amplitude connects to the two-transistor push-pull amplifier of Fig. P13.117. The BJTs have $I_s = 10$ fA and $\beta_F = 100$. Use SPICE to find the amplitudes of the first five output harmonics.

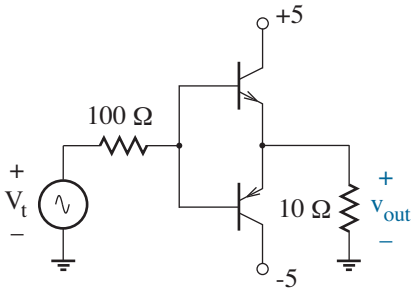


Figure P13.117

13.118 The circuit of Problem 13.117 reconfigures to include an op-amp with negative feedback as shown in Fig. P13.118. The op-amp has $A_{vd} = 100,000$, $r_{in} = 400$ kΩ, and $r_{out} = 100$ Ω. Use SPICE to find the amplitudes of the first five output harmonics.

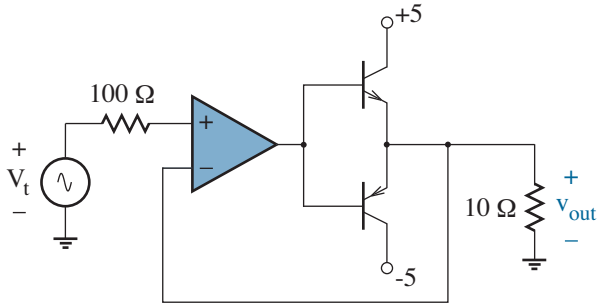


Figure P13.118

13.119 A particular op-amp has $A_{vd} = 200,000$, $r_{in} = 100$ kΩ, $r_{out} = 10$ Ω, and ±10-V supplies.

(a) Design a non-inverting amplifier with $A_v = 10$. Then use SPICE to investigate distortion for an input signal with 1-V peak-to-peak amplitude. Assume a 100-Ω load resistance.

(b) Design an inverting amplifier with $A_v = -10$. Then use SPICE to investigate distortion for the same input and load conditions of part a.

13.120 Figure P13.120 shows a simple diode **mixer** that applies distortion in support of demodulation—with $v_1 = \tilde{v_1}\cos\omega_1 t$ and $v_2 = \tilde{v_2}\cos\omega_2 t$, the output contains a component with angular frequency $\omega_1 - \omega_2$ (amongst others). Mixers are used in radio receivers.

(a) Use SPICE to find the amplitudes of the first ten output components (lowest frequency first). Assume $\tilde{v_1} = \tilde{v_2} = 0.1$ V, $f_1 = 100$ MHz, $f_2 = 90$ MHz, and let $I_s = 10$ fA for the diode.

(b) The desired output voltage signal is at 10 MHz. Demonstrate how the circuit can be improved so that other signal components are suppressed.

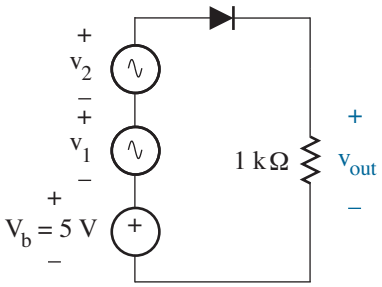


Figure P13.120

13.121 Figure P13.121 shows a Gilbert multiplier that functions as a mixer (Problem 9.23). Use SPICE to demonstrate this operation for the output $i_L - i_R$. Let v_1 and v_2 have the conditions of Problem 14.52, and let $I_s = 10$ fA, $\beta_F = 200$ for the BJTs.

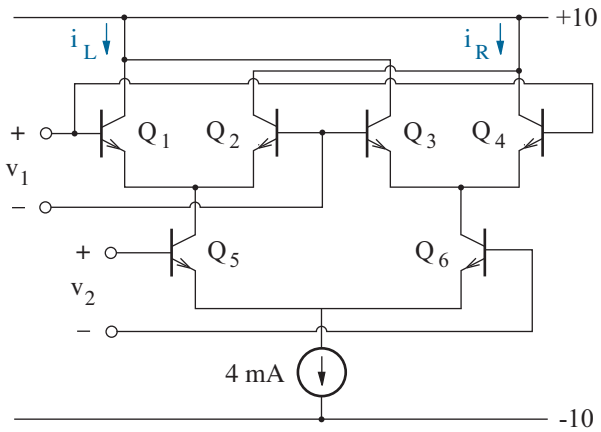


Figure P13.121

Section 13.6

13.122 Consider a uniform coaxial transmission line with the cross-sectional geometry of Fig. P13.122. The inner and outer conductors are separated by an insulator with dielectric constant ϵ. The magnetic permeability is $\mu_o = 4\pi \times 10^{-7}$ H/m.

(a) Determine an expression for the capacitance per unit length in terms of a, b, and other factors.

(b) Determine an expression for the inductance per unit length in terms of a, b, and other factors.

(c) Find Z_o using the results of parts a and b.

(d) Find the voltage v_{max} between the conductors when the electric field is $\mathcal{E}_{max}$, a maximum value prior to dielectric breakdown.

(e) Find the b/a value that maximizes ${v_{max}}^2/Z_o$, the power-handling capability for the line.

(f) Find the value of Z_o that is consistent with the result of part e.

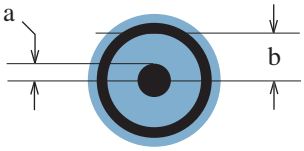


Figure P13.122

13.123 Consider a uniform coaxial transmission line with the structural conditions of Problem 14.122. As a consequence of the skin effect (see text), current in each conductor flows in a thin shell of thickness δ.

(a) Show that the line resistance per unit length is

$$R = \frac{1}{2\pi\delta\sigma}\left[\frac{1}{a} + \frac{1}{b}\right] ,$$

where σ is the metal conductivity.

(b) Find the b/a value that minimizes $\alpha = R/2Z_o$, the attenuation factor for the line.

(c) Find the value of Z_o that is consistent with the result of part b.

13.124 The concept of transmission-line inductance was poorly understood at the completion of the first transatlantic cable in 1858. As a result, the presumed telegrapher relations took the form

$$\frac{\partial v}{\partial z} = -R\, i$$

and

$$\frac{\partial i}{\partial z} = -C\,\frac{\partial v}{\partial t} ,$$

so that current or voltage obey a diffusion equation:

$$\frac{\partial^2 i}{\partial z^2} = RC\,\frac{\partial i}{\partial t} .$$

Subject to a step input $v(0,t) = V_o\, u(t)$,

$$i(z,t) = V_o\sqrt{\frac{C}{\pi R t}}\,\exp\left(\frac{-z^2 RC}{4t}\right) .$$

(a) Determine the time needed to achieve maximum current at the end of a line with 1852-mile length (Newfoundland to Ireland). Let $R \approx 10$ Ω/mile and $C \approx 0.2$ μF/mile.

(b) Find the V_o step amplitude that is consistent with $i_{max} = 1$ mA (for an insensitive detector).

Alas . . . Dr. Edward Orange Wildman Whitehouse, chief electrician for the Atlantic Telegraph Company, cared little for theory, no matter how encouraging. Input amplitudes "obviously" had to be of the order of 2000 V to propagate signals across the Atlantic. This soon led to cable failure. Whitehouse was fired, and a second cable was implemented in 1866.

13.125 Show that Eq. 13.100 is a solution to the wave equation (Eq. 13.99).

13.126 Verify Eq. 13.111 for attenuation factor α.

13.127 RG-8 cable has $Z_o = 50$ Ω, $\hat{v} = 0.78c$, and 1.8-dB loss/100 ft. Find R, C, and L (per foot).

13.128 A cable has $Z_o = 75$ Ω and $R = 40$ Ω/100 ft. Find the cable length that sustains -20-dB loss.

13.129 A cable has $C = 30$ pF/100 ft, $\hat{v} = 0.78c$. Find the end termination that yields $\Gamma = -0.5$.

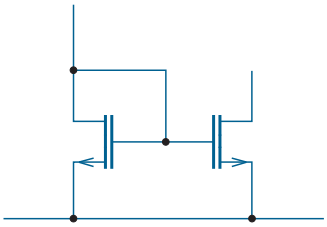
13.130 Repeat Example 13.9 with $R_t = 12.5\ \Omega$, attenuation = 2.5 dB/100 ft, and $l = 120$ ft.

13.131 Repeat Example 13.10 with $R_t = 150\ \Omega$.

13.132 Repeat Example 13.10, but terminate the line with a short circuit at $z = l$.

13.133 Repeat Example 13.10, but terminate the line with a short circuit at $z = l$ and let $R_t = 150\ \Omega$

13.134 Repeat Example 13.10 when the input provides a single 5-V pulse at $t = 0$ with 1-ns duration.



Chapter 14

Conversions

Have you ever been at work or school and faced with a problem that could only be resolved at home? Most would say often, and just a few the reverse. So you boarded a plane, or jumped on your motorcycle, or perhaps walked to the more accommodating place. Maybe your feelings were "converted."

Now this is admittedly a stretch, but somewhat less personable electrical signals often have similar difficulties—powerful processing techniques that offer enticing or opportune benefits are restricted to analog or digital realms. And a signal in the wrong "place" takes no advantage without conversion. Rapid and minimally disruptive forms of transport are needed.

After a brief motivating discussion and the specification of important performance measures, this chapter examines a wide variety of mixed-signal circuits that implement analog-to-digital or digital-to-analog conversions. The circuits are easily understood with little more than the black-box rules developed in Chapter 1 and elementary network theory. Notwithstanding, basic signal concepts in the time and frequency domains are helpful.

Objectives

After studying this chapter and working through the problems, you should be able to perform each of the following:

- Identify quantization error and other quasistatic conversion errors — offset, gain, integral and differential non-linearity (Section 14.2).
- Determine the signal-to-noise ratio (SNR) and the effective number of bits (ENOB) for Nyquist-rate dynamic conversions (Section 14.2).
- Describe the features and relative merits of flash, pipeline, successive-approximation, oversampling, and dual-slope integrating analog-to-digital converters (Section 14.3 - 14.5).
- Describe the features and relative merits of string, current-mode, and charge-mode digital-to-analog converters (Section 14.6).

14.1 Two Electrical Worlds

Nearly everyone finds the time to enjoy recorded music. Once an exclusive technology of analog electrical and mechanical processes, audio engineering exploits ever more digital conditioning for sound reproduction of the highest quality via compact disc (CD), digital video disc (DVD), and other media. Modern audio systems partition tasks between two electrical worlds.

Begin in the recording studio, where individual microphones convert sound to a small analog voltage that is proportional to the instantaneous pressure on an internal diaphragm. Each microphone is positioned to have a special purpose—one may focus on a particular singer, a second on the lead guitar, a third on drums—so that the recording engineer is presented with a set of analog signals having different degrees of momentary interest. The signals are separately scaled (faded) by appropriate factors, filtered (tone set) to adjust bass-midrange-treble, and divided (balanced) between left and right stereo channels. Finally, the channel components are added together (mixed) to yield blends that will hopefully suit the average listener. Figure 14.1 summarizes these audio processes.

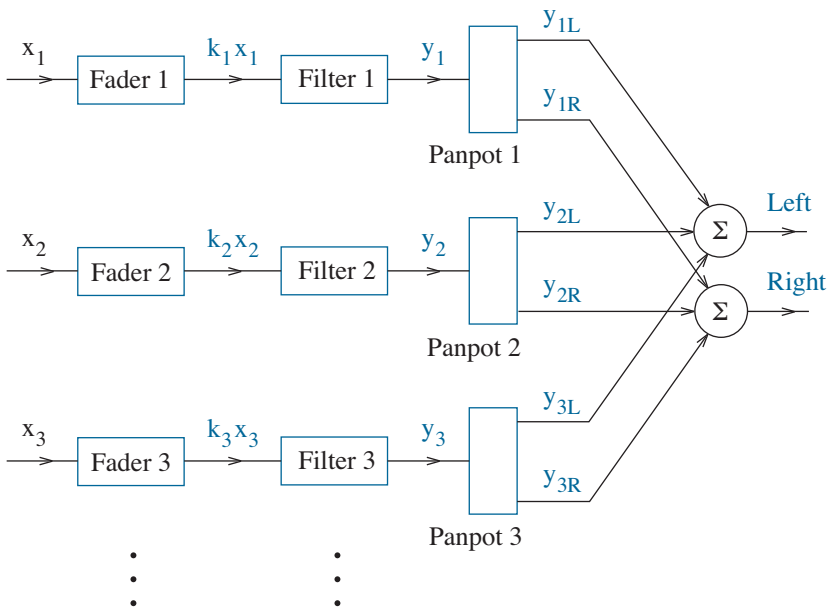


Figure 14.1: Audio signal processing in a studio. Faded and filtered signals are diverted into left and right mixers through "panpot" balance controls.

Which electrical world best accommodates the recording studio?

Figure 14.2 shows an analog implementation of the preceding processes. Three dual-op-amp panpots 1, 2, and 3 (see Problem 14.1) follow attenuated and filtered input signals from potentiometers R_1, R_2, and R_3, respectively. The left and right panpot outputs are then mixed with op-amps L and R.

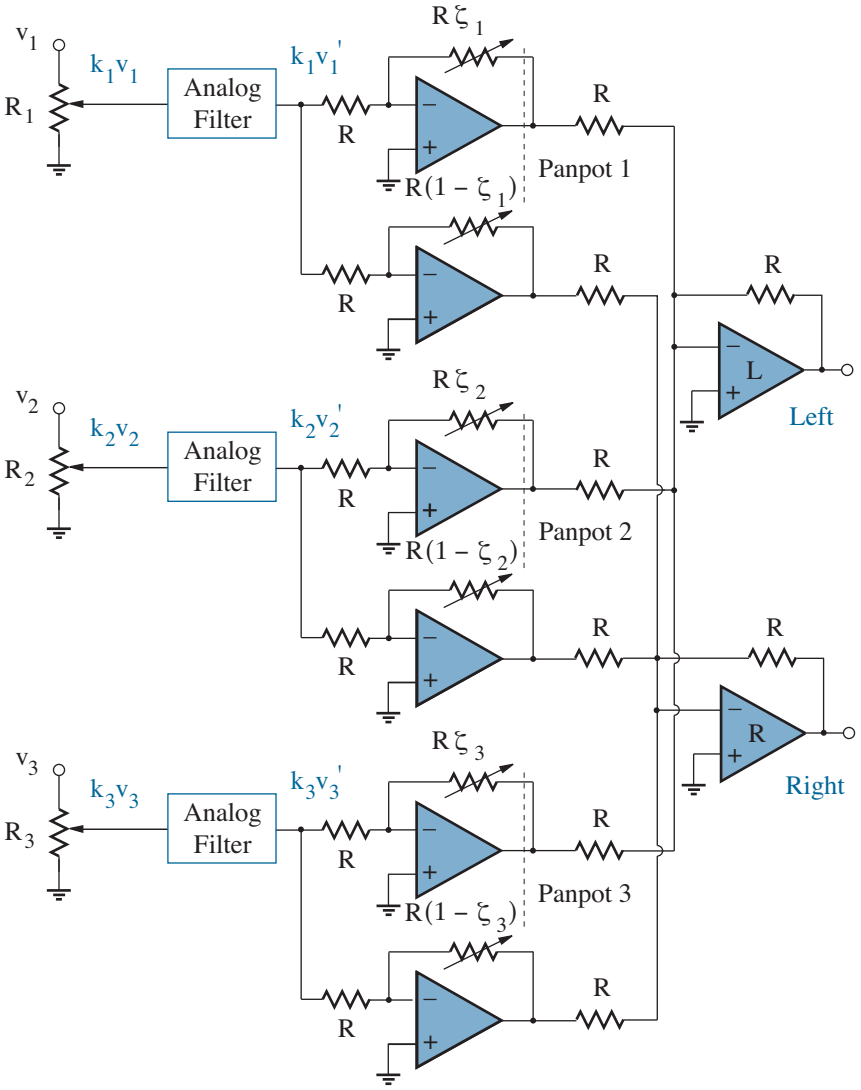


Figure 14.2: Partial analog implementation of audio processes in Fig. 14.1. The active panpots feature mechanically linked potentiometers that yield resistances $R\zeta$ and $R(1-\zeta)$, where $0 < \zeta < 1$.

Now consider Fig. 14.3, a digital implementation of the same processes. Here, the three input signals are immediately converted to digital format, and fader action is achieved through multiplication by a preset constant (that is not necessarily less than unity). The result is digitally filtered and split into left and right channels through a pair of multiplication actions. Adding the left and right signal components achieves final mixing.

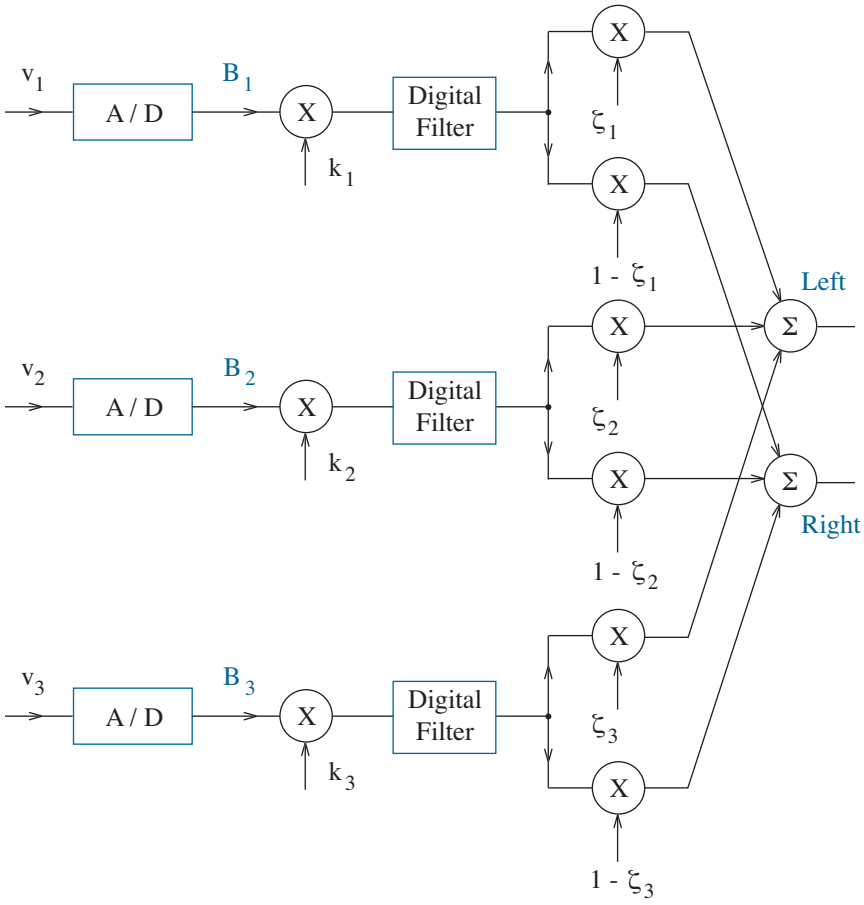


Figure 14.3: Partial digital implementation of audio processes in Fig. 14.1. Following analog-to-digital conversion, the fader, filter, panpot, and mixer functions are performed through separate numerical operations.

Compare: The analog system has five adjustments for each input signal (one fader, one panpot, and three filter range controls), and all of them are manually configured at the recording console. In contrast, the equivalent digital operations are numerical and well suited for computer interaction, thereby allowing almost immediate response to changing musical scenes.

Were we to continue, we would find that analog alternatives are simply blown away by digital processes subsequent to the left/right channel mixing. For example:

- Digital audio information is easily stored and then retrieved from a computer or an archival medium, possibly with efficient intermediate data compression. Analog information is relatively difficult to store, and the storage is exclusively long-term (tape or vinyl LP records).

- As a steady sequence of numbers, digital audio provides a faithful representation of sounds with marginal need for "quality" assessment. Analog audio signals are degraded to some extent at every stage of the recording and playback process. Quality demands expensive design.

- In the unlikely event of failing to distinguish between a one or a zero, specially coded digital audio signals can usually be restored through sophisticated error-correction algorithms. Corrupted analog signals offer no identification that invites particular correction.

Despite these disadvantages, the analog world is the alpha and omega—the beginning and end—of audio and numerous other electrical systems. The initial microphone signals are analog. The final output signals to a set of loudspeakers are also analog. And if you are thinking that everything in between has to be digital, consider the need for intermediate analog control operations such as optical groove tracking when playing a CD.

The moral of our story ...

Waste no time in jumping over to the digital world.

One usually finds some advantage, and this has become common practice. As noted, there is frequent need to return eventually to the analog world, thereby suggesting the system of Fig. 14.4. The key to success is easy inter-world passage through analog-to-digital and digital-to-analog conversions. Let the journey begin.

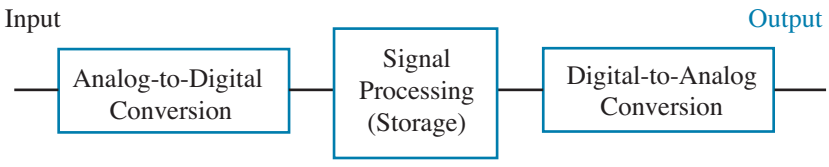


Figure 14.4: Common approach to electrical system design.

14.2 Rites of Passage

It should come as no surprise that passage to and from the digital world enlists the aid of an analog-to-digital converter (ADC) or a digital-to-analog converter (DAC), respectively. Like most journeys, the perils of interworld transit depend upon speed. If analog or digital information changes slowly, we are content to execute transitions at the leisurely pace of a cruise ship with results subject to nearly fixed **quasistatic** measures of performance. In the more likely event that electrical information bears temporal interest, analog-to-digital conversions transpire with the **pulse-code modulation** process of Fig. 14.5 whereby sampling and coding occur at regular intervals. The resulting stream of digital data can be manipulated before passing back to the analog world, very often at the same rate as the original sampling. Transitions may require the pace of a racecar or jet fighter, and results are subject to ever-more-stringent **dynamic** measures of performance.

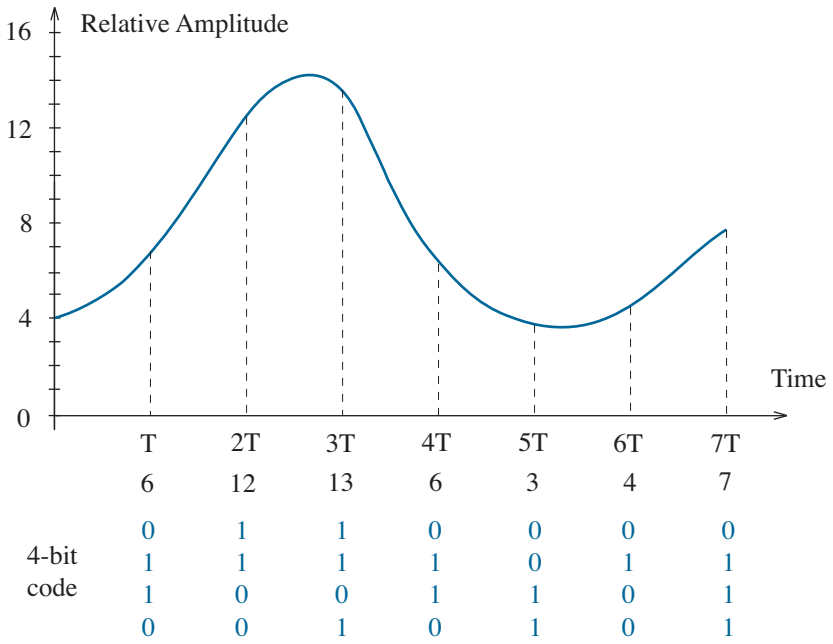


Figure 14.5: Pulse-code modulation. The sampling and conversion process is equivalent to multiplying a steady sequence of equal-strength impulses by a time-dependent amplitude and rounding to the nearest digital number.

In order to appreciate the major converter circuits (Sections 14.3 -14.6), we need to understand the consequences of digital roundoff quantization and discrete-time sampling at different rates and with potential irregularity. In part, our objective is to determine appropriate bit lengths and dynamic specifications that will ensure reliable conversions.

Quantization

Analog-to-digital conversion begins with an understanding that any analog voltage of interest lies within some range, say $0 < v_{in} < V_{ref}$, where V_{ref} is a rock-solid dc reference. If intended digital representations have N bits, V_{ref} can be divided into 2^N equal partitions with binary code assignments. The partition "width" or, equivalently, the voltage change that corresponds to a binary increment of one **least significant bit** (LSB) is given by

$$v_{LSB} = \frac{V_{ref}}{2^N} \,. \tag{14.1}$$

We thus obtain a set of voltage "bins" that conveniently center on integer multiples of v_{LSB} as shown in Fig. 14.6. This is a metric for v_{in} evaluation. The remaining analog-to-digital conversion process is much like a game: "Place" the input voltage into the proper bin. Score points for number of attempts, cost, and effort. Low score wins. Strategies are forthcoming.

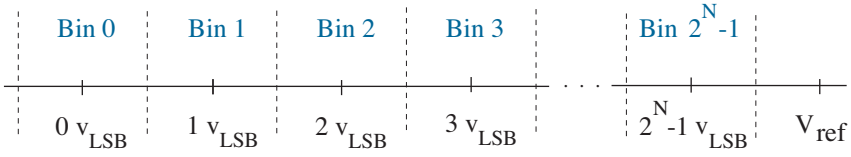


Figure 14.6: V_{ref} voltage-bin partitions for analog-to-digital conversion. Partition boundaries are offset from the integer scale by $v_{LSB}/2$.

Binary-coded bin assignments are usually made to facilitate the inverse digital-to-analog conversion process. The simplest **natural binary** method counts upwards in binary beginning with zero for Bin 0. Thus, the word $B = b_1, b_2, \ldots b_N$ corresponds to voltage

$$v_{out} = V_{ref} \left(b_1\, 2^{-1} + b_2\, 2^{-2} + \ldots + b_N\, 2^{-N} \right) \,. \tag{14.2}$$

We pause to note that v_{in} is often negative (as for a sinusoidal signal). In this case, one chooses V_{ref} to ensure that $-V_{ref}/2 < v_{in} < +V_{ref}/2$. The voltage span of total width V_{ref} partitions as before so that Bin 0 corresponds to $(-V_{ref} \pm v_{LSB})\,/\,2$. Counting-upwards-from-zero bin assignments yield an **offset binary** representation with positive and negative voltages featuring 1 and 0, respectively, as the first or **most significant bit** (MSB). The code word $B = b_1, b_2, \ldots b_N$ now corresponds to voltage

$$v_{out} = V_{ref} \left(b_1\, 2^{-1} + b_2\, 2^{-2} + \ldots + b_N\, 2^{-N} \right) - V_{ref}/2 \,. \tag{14.3}$$

Offset binary representation is easily converted to **2's complement** code —one changes the value of the most significant bit from 0 to 1 or vice versa. The result is convenient for computer arithmetic (see Problem 14.5).

Exercise 14.1 Consider a 6-bit DAC with $V_{ref} = 4.096$ V. Find v_{LSB}, then determine the voltage that reflects 1 1 0 1 0 1 for natural binary and offset binary coding.

Ans: $v_{LSB} = 64$ mV, $v = 3.392$ V (natural), $v = 1.344$ V (offset)

Exercise 14.2 Consider an 8-bit ADC with $V_{ref} = 5$ V. Determine the natural binary code that corresponds to $v = 2.18$ V.

Ans: 0 1 1 1 0 0 0 0 (How did *you* play the game?)

Exercise 14.3 Consider an 8-bit ADC with $V_{ref} = 5$ V. Determine the offset binary code that corresponds to $v = -1.06$ V.

Ans: 0 1 0 0 1 0 1 0

Quasistatic Error

Analog-to-digital conversions that are immediately followed by digital-to-analog reconversions should yield initial voltage levels (if within bound). Nevertheless, the finite width of each V_{ref} voltage-bin partition produces **quantization error** and the non-linear transfer characteristic of Fig. 14.7. The errors are decreased as the number of bits increases—more steps with proportionately reduced height improves the staircase approximation to the straight-line ideal.

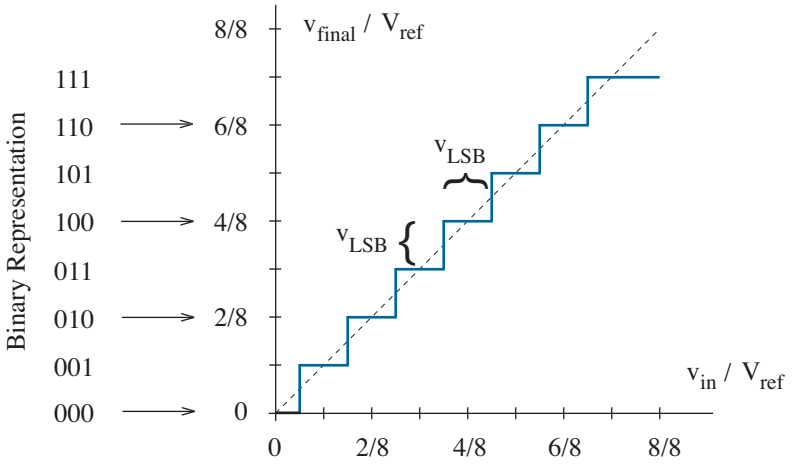


Figure 14.7: 3-bit analog-digital-analog conversion transfer characterisitic.

To make matters worse, four modes of imperfection distort the transfer characteristic of Fig. 14.7 so that the quantization errors are non-uniform. The offending contributors—

- **Offset error** implies a non-zero output for the case of zero input. An analog-to-digital converter has offset error if the center of Bin 0 shifts from zero (Fig. 14.8a). In contrast, a digital-to-analog converter has offset error if code word $B = 0 \ldots 0$ yields $v_{out} \neq 0$ (Fig. 14.8b).

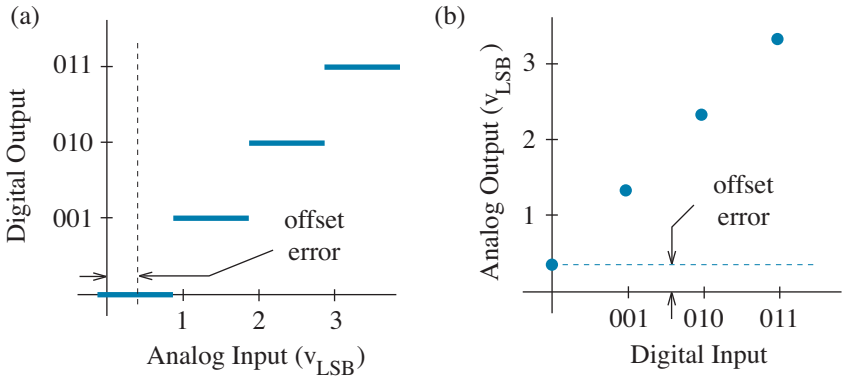


Figure 14.8: Offset errors for a 3-bit converter: (a) ADC; (b) DAC.

- **Gain error** implies a non-full-scale ouput for the case of full-scale input after correction for offset error. An analog-to-digital converter has gain error if the center of Bin $(2^N - 1)$ shifts from $(2^N - 1)\, v_{LSB}$ (Fig. 14.9a). In contrast, a digital-to-analog converter has gain error if code word $B = 1 \ldots 1$ yields $v_{out} \neq (2^N - 1)\, v_{LSB}$ (Fig. 14.9b).

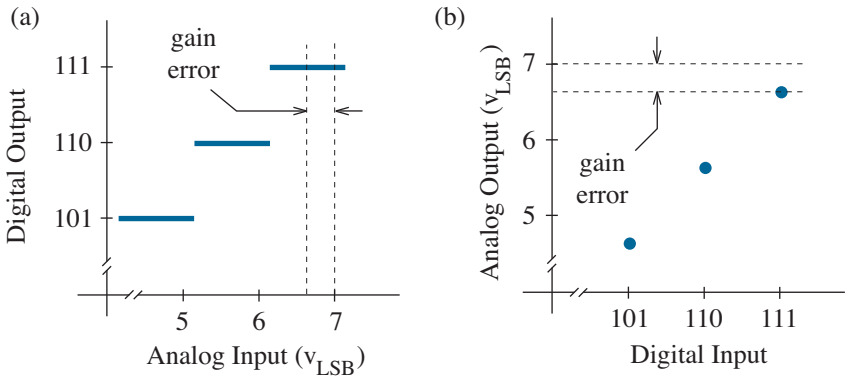


Figure 14.9: Gain errors for a 3-bit converter: (a) ADC; (b) DAC.

- **Integral nonlinearity (INL) error** reflects transfer-characteristic deviations from a progressive linear average. For analog-to-digital and digital-to-analog converters, individual errors measure at voltage-bin boundaries (Fig. 14.10a) and step levels (Fig. 14.10b), respectively. In each case, errors are cumulative from step to step.

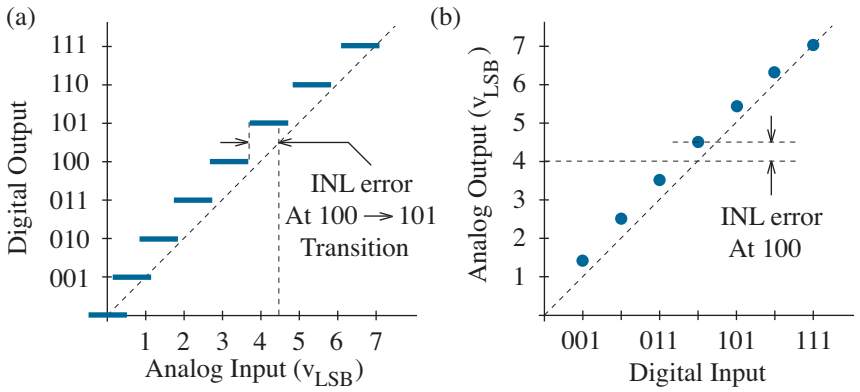


Figure 14.10: INL errors for a 3-bit converter: (a) ADC; (b) DAC.

- **Differential nonlinearity (DNL) error** reflects step distortions in the transfer characteristic. For analog-to-digital and digital-to-analog converters, individual errors are measured as bin-width (Fig. 14.11a) and step-height (Fig. 14.11b) changes from a v_{LSB} ideal, respectively. If an error exceeds one v_{LSB} in an ADC, there may be **missing codes** —one or more digital outputs will never appear. The same condition in a DAC can lead to a loss of **monotonicity** or steady code increase.

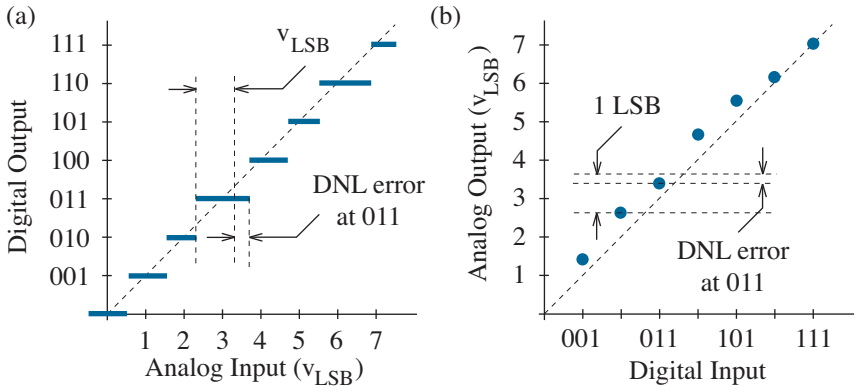


Figure 14.11: DNL errors for a 3-bit converter: (a) ADC; (b) DAC.

Example 14.1

A 3-bit ADC features $V_{ref} = 4$ V. Beginning with the 000 → 001 transition, the voltage-bin boundaries occur at 0.239, 0.746, 1.268, 1.756, 2.271, 2.742, and 3.237 V. Determine the applicable INL and DNL values.

Assume precalibration so that offset and gain errors are zero.

Solution

We make a table. Columns one and two contain transition specifications and the corresponding as-measured voltage-bin boundaries, respectively. The third column contains ideal boundary voltages, which start at $v_{LSB}/2$ and increment in v_{LSB} multiples. Here, we have $v_{LSB} = V_{ref}/8 = 0.500$ V. Column four contains INL, which is the difference between column-two (measured) and column-three (ideal) voltages. Column five contains DNL, which is the difference between column-two voltages at the next and current transitions, less v_{LSB}.

Binary Transition	Measured Boundary Voltage	Ideal Boundary Voltage	INL (V)	DNL (V)
000 → 001	0.239	0.250	-0.011	+0.007
001 → 010	0.746	0.750	-0.004	+0.022
010 → 011	1.268	1.250	+0.018	-0.012
011 → 100	1.756	1.750	+0.006	+0.015
100 → 101	2.271	2.250	+0.021	-0.029
101 → 110	2.742	2.750	-0.008	-0.005
110 → 111	3.237	3.250	-0.013	—

Note: INL and DNL are often expressed in units of v_{LSB}.

As assumed, ADC and DAC manufacturers precalibrate their products so that gain and offset errors are close to zero at some nominal temperature. INL and DNL characteristics must be minimized through careful design. Errors are temperature dependent.

Figs. 14.12 and 14.13 show typical INL and DNL data. The INL error can often exceed 1 LSB in high-speed converters. Barring gross INL errors, DNL error tends to warrant greater concern for most applications, since associated discontinuities such as missing codes are easily detected.

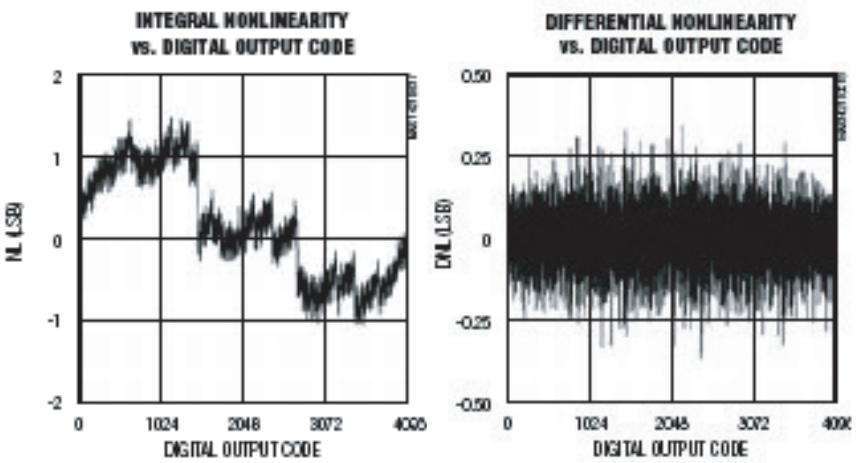


Figure 14.12: Typical INL and DNL data for the 12-bit MAX1421 ADC (Copyright Maxim Integrated Products. Used by permission.)

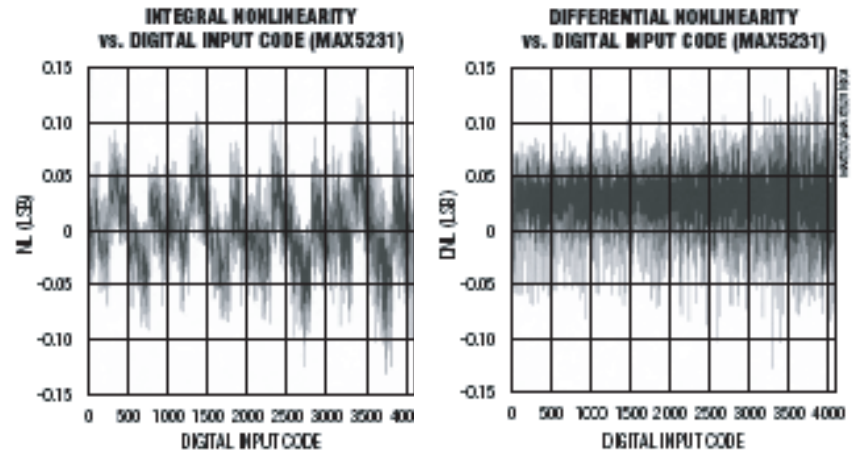


Figure 14.13: Typical INL and DNL data for the 12-bit MAX5231 DAC (Copyright Maxim Integrated Products. Used by permission.)

Discrete-Time Sampling

No doubt many readers are unfamiliar with the fundamentals of analog and digital signal processing, an important engineering discipline that warrants special consideration elsewhere. Nevertheless, some of the consequences of discrete-time sampling for analog-to-digital conversion can be understood with little more than intuitive common sense.

Figure 14.14 shows a time-dependent signal $v(t)$ and a representation as a continuous spectrum of frequency-dependent $\tilde{v}(\omega)$ amplitudes that multiply superimposed $\exp(\jmath\omega t)$ oscillatory components. In this example, we assume that the angular frequency spectrum ends abruptly at ω_c.

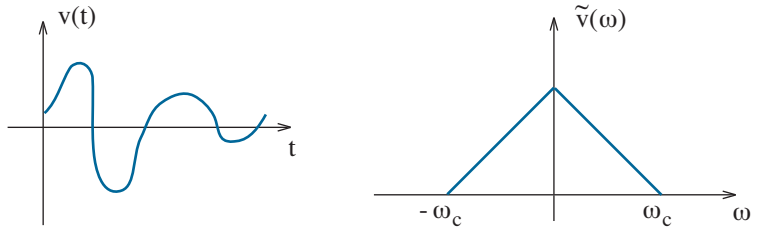


Figure 14.14: Particular signal representations in both time and frequency. Note that real signals require $\tilde{v}(\omega)$ symmetry with respect to $\omega = 0$.

When $v(t)$ is sampled at ω_s, the overall effect in the angular frequency domain is to create $\tilde{v}(\omega)$ duplicates with respect to integer multiples of ω_s. This yields the spectrum of Fig. 14.15. After digital-to-analog conversion, the original signal is recovered by removing the higher-order harmonics.

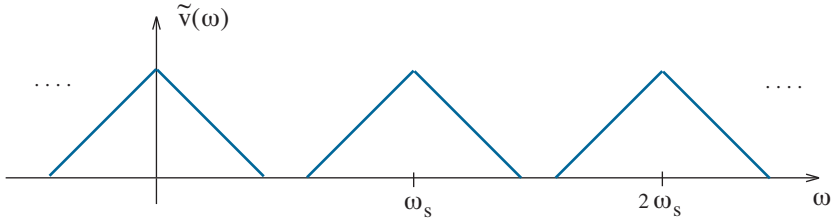


Figure 14.15: Effect of discrete-time sampling in the frequency domain.

Everything went smoothly in the preceding example as a result of the gaps between the $\tilde{v}(\omega)$ duplicates—filtering leaves behind a clean original. It is easy to see that the signal recovery process would have been corrupted had the sampling frequency been too low or, equivalently, had the abrupt ω_c cutoff for $\tilde{v}(\omega)$ been too high so that the post-sampling duplicates overlap. The unhappy overlapping condition is called **aliasing**.

To prevent aliasing, sampling must exceed the **Nyquist rate** given by

$$\omega_s = 2\,\omega_c\,. \tag{14.4}$$

The maximum ω_c angular frequency component of the signal to be sampled is often established with the help of a special antialiasing low-pass filter. This filtering is nessessarily initiated in the analog world.

Unfortunately, cruel reality disallows sampling at perfect time intervals. Figure 14.16 shows how conversion error can be introduced when sampling is subjected to time uncertainty or **jitter**.

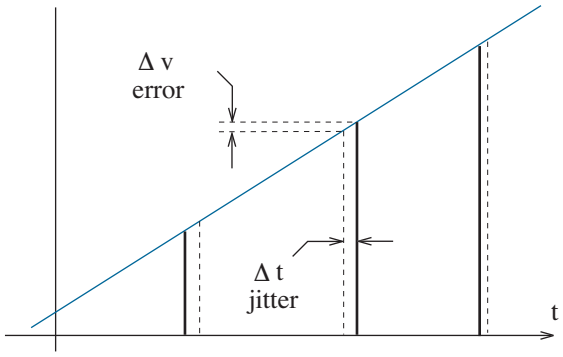


Figure 14.16: Sampling jitter and its effect on conversion error.

The influence of jitter is most pronounced when signals have maximum rate of change. Consider a worst-case signal

$$v(t) = V_{ref}\left[\frac{1+\sin(\omega_s t/2)}{2}\right], \tag{14.5}$$

which fully spans V_{ref} at angular frequency $\omega_c = \omega_s/2$. The maximum rate of change is

$$\left.\frac{\Delta v}{\Delta t}\right|_{\max} = \frac{V_{ref}\,\omega_s}{4}\,. \tag{14.6}$$

We interpret Δv as the error that arises through sampling uncertainty Δt, and a reasonable demand is $\Delta v < v_{LSB}/2$. Then with the help of Eq. 14.1,

$$\Delta t < \left(\frac{T_s}{\pi}\right) 2^{-N}\,, \tag{14.7}$$

where $T_s = 2\pi/\omega_s$ is the Nyquist sampling period. Upper-limit jitter-time specifications impose particular **phase noise** requirements on oscillators used to generate clock signals for sampling (see Problem 14.15).

Dynamic Error

From a dynamic perspective, we view the quantization error as electrical "noise" that is superimposed over an ac signal. Consider a "round-trip" excursion that applies to a sinusoidal input signal bounded by 0 and V_{ref}. The ac and dc components are $(V_{ref}/2)\sin\omega_o t$ and $V_{ref}/2$, respectively. Subject to $\omega_o \leq \omega_c$ and $T = 2\pi/\omega_o$, the input signal power is

$$P_{in} = \frac{1}{T}\int_0^T \left(\frac{V_{ref}}{2}\right)^2 \sin^2\omega_o t\, dt = \frac{{V_{ref}}^2}{8}\,, \tag{14.8}$$

and it is concentrated at angular frequency ω_o. Then with Eq. 14.1,

$$P_{in} = \frac{2^{2N}{v_{LSB}}^2}{8}\,. \tag{14.9}$$

The effect of quantization in the analog-to-digital conversion process adds a *random* error between $-v_{LSB}/2$ and $+v_{LSB}/2$ at every sampling interval. The quantization-error noise power is (see Problem 14.16)

$$P_Q = \frac{{v_{LSB}}^2}{12}\,, \tag{14.10}$$

and it is uniformly spread over the angular-frequency interval $[-\omega_s/2, \omega_s/2]$ (as verified experimentally). Thus, the "ideal" **signal-to-noise ratio** is

$$\text{SNR} = (3/2)\,2^{2N} \tag{14.11}$$

when $\omega_s/2 = \omega_c$. In terms of decibels (dB),

$$\text{SNR} = 10\,\log\left[(3/2)\,2^{2N}\right] = 6.02\,N + 1.76\ \text{dB}\,. \tag{14.12}$$

For the case of $N = 12$ bits, the ideal SNR is 74 dB.

When measuring signal-to-noise ratio for data converters, it is common practice to combine *all* sources of noise and non-linear behavior (distortion) by specifying **SINAD** (Signal-to-Noise and Distortion):

$$\text{SINAD} = 10\,\log\left(\frac{\text{Signal Power}}{\text{Noise and Distortion Power}}\right)\,. \tag{14.13}$$

Here, "power" reflects an average as in Eq. 14.8. Given a particular SINAD, we rearrange Eq. 14.12 to determine **ENOB**, an Effective Number of Bits. Specifically,

$$\text{ENOB} = \frac{\text{SINAD} - 1.76}{6.02}\,. \tag{14.14}$$

ENOB characteristics tend to be applied to analog-to-digital converters as opposed to DACs, which are more likely evaluated on the basis of distortion. Chapter 13 has considered some consequences of signal distortion.

Example 14.2

Figure 14.17 shows an 8192-point FFT (**Fast-Fourier-Transform**) of the output of a 10-bit ADC subject to a single 39.991-MHz tone with -0.5-dB amplitude in relation to full scale (V_{ref}).[1] The average noise "floor" has -98.4-dB amplitude in relation to full scale, and the sampling frequency is 82.345 MHz. Determine the ideal SNR, actual SNR, SINAD, and ENOB.

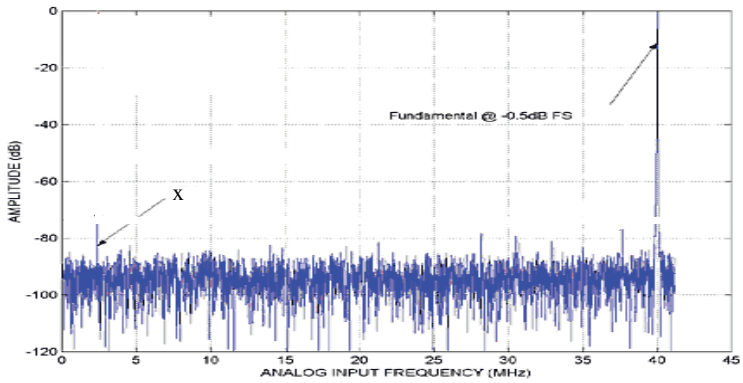


Figure 14.17: FFT spectrum for Example 14.2.

Solution

- Ideal $SNR = 6.02 \times 10 + 1.76 = 62.0$ dB (Eq. 14.12).
- Actual SNR involves the logarithm of a fraction, so estimates suffice. The SNR would appear to be the difference between -0.5 dB (related to signal power) and -98.4 dB (related to noise power). But Fig. 14.17 is a spectral *density*. Thus, the signal power reflects an integration over 1 frequency "bin" — 82.345 MHz / 8192 samples = 10.052 kHz, and the noise power reflects an integration over 8192 bins. So SNR = -0.5 dB - (-98.4 dB) + $10 \log 1 - 10 \log 8192 = 58.8$ dB.
- The inclusion of distortion peaks makes the SINAD slightly smaller than SNR, say 58.7 dB.
- ENOB = (58.7 - 1.76) / 6.02 = 9.46 (Eq. 14.14).

Note: Another quantity of interest is the **Spurious-Free Dynamic Range** (SFDR), which is the difference between 0 dB (full-scale amplitude) and the distortion "spur" with largest amplitude (marked with an x in Fig. 14.17). This particular SFDR is 74.2 dBc (dB from center frequency).

[1] Data from *Defining and Testing Dynamic Parameters in High-Speed ADCs, Part 1* (available at www.maxim-ic.com, Tutorial 728). Copyright Maxim Integrated Products. Used by permission.

14.3 Nyquist-Rate A/D Converters

Having considered the effects of quantization and other non-linear errors, we are prepared to examine specific circuits for analog-to-digital conversion. Our discussion moves in descending order of speed and overall simplicity, yet ascending order of economy of scale. Such is the everpresent trade-off. Each implementation finds optimum merit for particular applications.

Flash Converters

N-bit **flash** or **parallel** ADCs feature 2^N-1 comparators that mark the upper edges of voltage bins defined by 2^N resistors in a simple divider chain (see Fig. 14.18 for $N = 3$). Thus, an input voltage induces HIGH and LOW outputs for those comparators marking lower- and higher-level voltage bins, respectively. The 3-bit flash ADC functions as follows:

Bin #	Lower Limit	Upper Limit	Comparator Outputs
0	—	$V_{ref}/16$	0 0 0 0 0 0 0
1	$V_{ref}/16$	$3V_{ref}/16$	0 0 0 0 0 0 1
2	$3V_{ref}/16$	$5V_{ref}/16$	0 0 0 0 0 1 1
3	$5V_{ref}/16$	$7V_{ref}/16$	0 0 0 0 1 1 1
4	$7V_{ref}/16$	$9V_{ref}/16$	0 0 0 1 1 1 1
5	$9V_{ref}/16$	$11V_{ref}/16$	0 0 1 1 1 1 1
6	$11V_{ref}/16$	$13V_{ref}/16$	0 1 1 1 1 1 1
7	$13V_{ref}/16$	—	1 1 1 1 1 1 1

So we have a string of ones growing longer with increasing input voltage. This **thermometer code** —consider the mercury in a real thermometer— is readily changed to binary form with an encoder.

The flash converter is a jet-engine ADC. Nevertheless, when N is large, it suffers from size, cost, power consumption, and capacitive input loading. The flash ADC must also contend with so-called **bubble errors** that arise when mismatched comparator delays produce thermometer code containing a one embedded within a string of zeros or a zero within a string of ones. Unless recognized (see Problem 14.24), transient encoder-confusing bubble errors yield erratic ADC outputs known as **sparkle codes**.

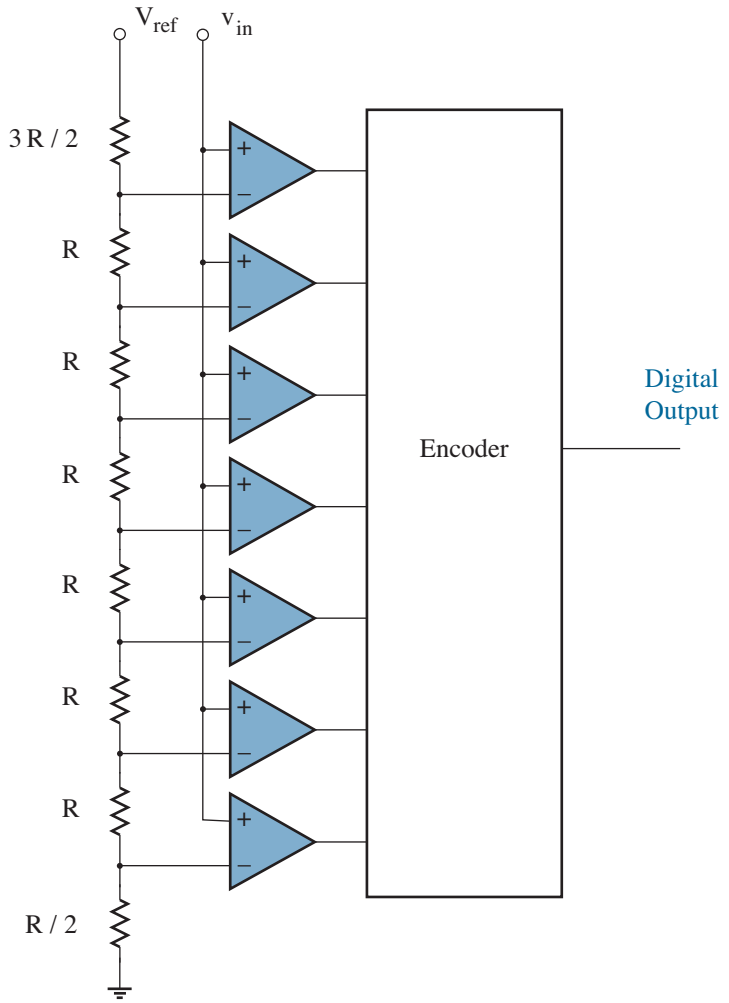


Figure 14.18: 3-bit flash (parallel) analog-to-digital converter.

Exercise 14.4 A 4-bit flash ADC features $V_{ref} = 1.4$ V. Determine the thermometer code that corresponds to $v_{in} = 0.538$ V.

Ans: 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1

Exercise 14.5 An 8-bit flash ADC features $V_{ref} = 4.096$ V. Find the necessary voltage-reference supply current if $R = 500\ \Omega$.

Ans: $i_{ref} = 32\ \mu$A

Pipeline Converters

Consider two black boxes: Box A achieves a desired process result after a single clock pulse. Box B achieves the same result after 12 clock pulses. Which box is preferable? At first glance, A, especially for intermittent use. But B can produce a second result on the 13th pulse, a third on the 14th, and so on. (We also forgot to mention that B costs a quarter that for A.) The elusively meritorious Box B is an example of a **pipeline processor**—provide a steady flow of input data at one end of the "pipe" and collect a similar flow of processed data at the other end with some delay or **latency**. Expect a benefit such as reduced cost or power consumption in return for increased delay (hopefully somewhat less than 12 clock periods).

The two-step $N + M$ **subranging** process of Fig. 14.19 is the simplest pipeline ADC architecture. Here, a "coarse" flash ADC resolves the most significant N bits that reflect a particular input voltage v_{in}. An N-bit DAC transforms the result to an analog voltage, which is subsequently subtracted from the input. Thereafter, a "fine" ADC examines the difference voltage to resolve the least significant M bits that apply to v_{in}. Typically, $N = M$.

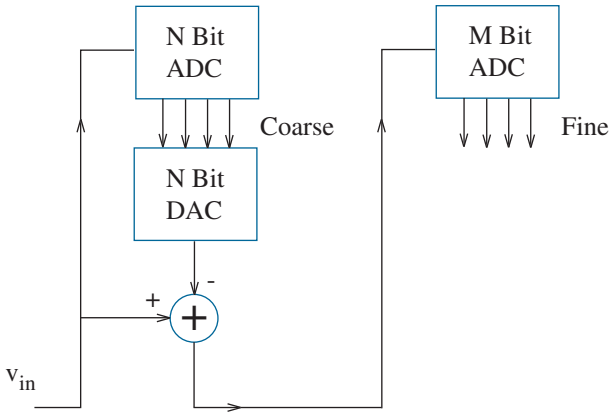


Figure 14.19: Two-step analog-to-digital subranging conversion process.

The reward? Consider a 12-bit flash ADC, which requires $2^{12} - 1 = 4095$ comparators (amongst other expensive requirements). With $N = M = 6$, the two-step ADC subranging process uses $2 \times (2^6 - 1) = 126$ comparators, a savings factor of about 32. As indicated in Fig. 14.19, the $6 + 6$ pipeline ADC also requires a 6-bit digital-to-analog converter. Yet this demands relatively little hardware—see Section-14.6 details. Indeed, the DAC can share the resistor divider that services the coarse ADC (see Problem 14.27) if comparators are appropriately designed to assist the difference operation that precedes the fine ADC process (see Problem 14.28).

The three-step pipeline algorithm of Fig. 14.20 is all digital for clarity. Assume that the voltage for conversion reflects 12 specific bits partitioned into three sets of four. The most significant first set of bits can be recognized with a coarse 4-bit flash ADC, so subtract this binary number followed by eight zeros to obtain an 8-bit remainder. Now multiply by 16, which is the same as a 4-bit shift to the left and promotion of the second set of bits into most significant territory recognized by the coarse 4-bit flash ADC. Finally, subtract, multiply by 16, and recognize the third set of bits.

```
 1 0 1 1   0 1 1 0   1 0 0 1
-1 0 1 1   0 0 0 0   0 0 0 0
-----------------------------
           0 1 1 0   1 0 0 1
  x 16
-----------------------------
 0 1 1 0   1 0 0 1   0 0 0 0
-0 1 1 0   0 0 0 0   0 0 0 0
-----------------------------
           1 0 0 1   0 0 0 0
  x 16
-----------------------------
 1 0 0 1   0 0 0 0   0 0 0 0
```

Figure 14.20: Three-step pipeline algorithm for a 12-bit conversion.

Figure 14.21 shows the signal block diagram that implements the pipeline ADC process of Fig. 14.20. Only $3\times(2^4-1) = 45$ comparators are required.

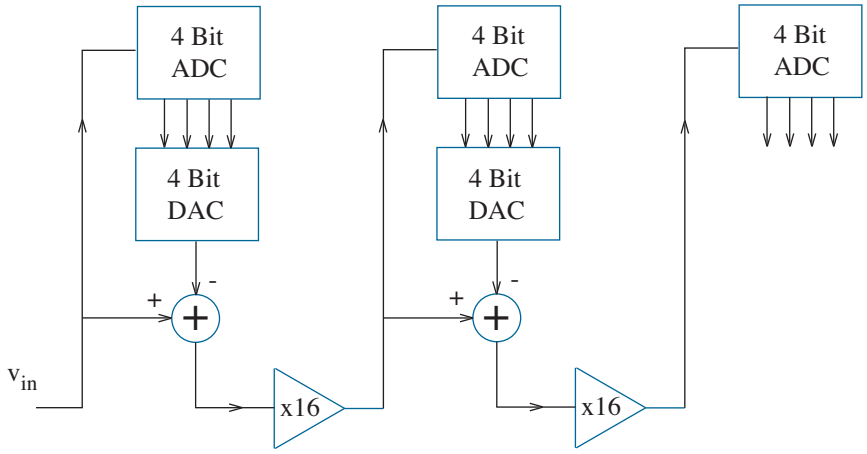


Figure 14.21: Signal block diagram for the pipeline process of Fig. 14.20.

Portions of the 12-bit pipeline converter that subtract and multiply by 16 have simple analog implementations. Figure 14.22 shows one circuit ...

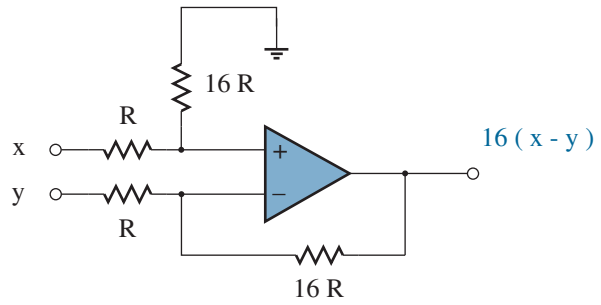


Figure 14.22: Analog circuit that subtracts and multiplies by 16.

... which, without care, provides a potential pitfall for ADC performance. The op-amp gain factor must be precisely 16, and the op-amp output must be exactly zero (a condition of zero **offset**) when both input signals are zero. No problem for the ideal circuits of Chapter 1. But in practice?

In practice, ADC manufacturers have an economic incentive to produce robust designs that withstand a certain degree of sloppy functionality in the presence of uncertainty. Thus, it is a welcome observation that the rules of digital arithmetic tend to suppress conversion errors if the exemplary 12-bit pipeline process of Figs. 14.20 and 14.21 takes an alternate form:

- Make a 3-bit analog conversion, reconvert the result with a 3-bit DAC, find the difference, and propagate a 4-X multiple forward (not 8-X). Do this four times to obtain 101, 010, 011, and 001.
- Perform a final 4-bit ADC. Do this with relative care to obtain 0001.
- Add the results with one-bit overlap:

```
  1 0 1
    0 1 0
      0 1 1
        0 0 1
          0 0 0 1
-----------------
  1 0 1 1 0 1 1 0 1 0 0 1
```

So we obtain the correct digital word. The more general pipeline algorithm with similar **digital error correction** has performance implications for the ADC, DAC, and analog multiplier (see Problem 14.34).

Decision-Based Converters

A simple analog-to-digital conversion algorithm—worthy of a six-year-old —goes something like this:

> Is it less than $(0+1/2)\times v_{LSB}$? [0000] (Are we there yet?) No.
> Is it less than $(1+1/2)\times v_{LSB}$? [0001] (Are we there yet?) No.
> Is it less than $(2+1/2)\times v_{LSB}$? [0010] (Are we there yet?) No.
> . . . another ten negative test results . . .
> Is it less than $(13+1/2)\times v_{LSB}$? [1101] . . . Yes! (Finally)
>
> Then the conversion result is 1101 (4 bits, $v_{in} = 13.2\, v_{LSB}$).

Thus, we have an algorithm that counts upwards from zero so that higher edges of prospective voltage bins are successively specified (see Fig. 14.6). Counting stops when the input voltage under consideration is less than a particular edge voltage.

The simple counting algorithm is guaranteed to determine the correct voltage bin. Nevertheless, the procedure is grossly inefficient, since it can require up to 2^N trial comparisons for N-bit conversions. This is potentially time consuming when N is large.

Figure 14.23 shows the system implementation of a 4-bit counting ADC. The output from a binary counter connects to a DAC whose analog output is offset by $v_{LSB}/2$ and fed to the non-inverting input of a comparator. The comparator transitions from LOW to HIGH to stop counting when the adjusted DAC output exceeds v_{in}.

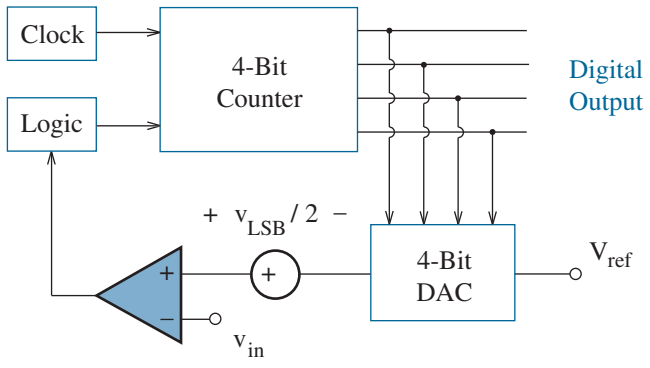


Figure 14.23: 4-bit counting ADC (almost never used).

Surely we can find a more efficient algorithm.

The **successive approximation** algorithm depicted in Fig. 14.24 for the case of four bits has efficiency to satisfy all but the chronically impatient. Suppose $v_{in} = 13.2\, v_{LSB}$ (as in the previous 14-step counting conversion). From the figure, we see that v_{in} lies within the voltage bin that corresponds to the digital word 1101. To establish this fact systematically, we embark upon a set of comparisons between v_{in} and a trial voltage v_x that is a fixed function of the bits b_1, b_2, b_3, b_4. Start with the digital word $B = 0000\ldots$

- Let $b_1 = 1$ and $v_x = (b_1\, 2^{-1} + b_2\, 2^{-2} + b_3\, 2^{-3} + b_4\, 2^{-4})\, V_{ref}$.
  $v_{in} > v_x$? Yes. Then $b_1 = 1$, $B = 1000$. End of trial 1.
- Let $b_2 = 1$ and $v_x = (b_1\, 2^{-1} + b_2\, 2^{-2} + b_3\, 2^{-3} + b_4\, 2^{-4})\, V_{ref}$.
  $v_{in} > v_x$? Yes. Then $b_2 = 1$, $B = 1100$. End of trial 2.
- Let $b_3 = 1$ and $v_x = (b_1\, 2^{-1} + b_2\, 2^{-2} + b_3\, 2^{-3} + b_4\, 2^{-4})\, V_{ref}$.
  $v_{in} > v_x$? No. Then $b_3 = 0$, $B = 1100$. End of trial 3.
- Let $b_4 = 1$ and $v_x = (b_1\, 2^{-1} + b_2\, 2^{-2} + b_3\, 2^{-3} + b_4\, 2^{-4})\, V_{ref}$.
  $v_{in} > v_x$? Yes. Then $b_4 = 1$, $B = 1101$. End of trial 4.

End of process, and only four steps—one for each bit.

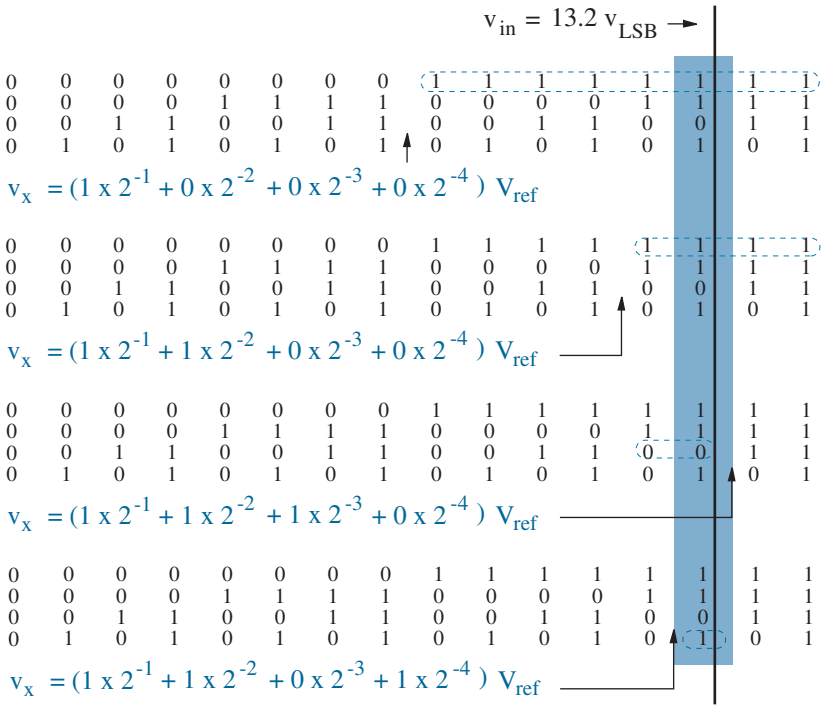


Figure 14.24: Successive approximation algorithm for a 4-bit conversion.

An electronic implementation of the successive approximation algorithm is easier than it may seem given the aid of a comparator, some appropriately scaled capacitors, and a set of switches that control a decision process in response to charge redistribution. Figure 14.25 shows one such circuit that executes a four-bit conversion. The relative capacitor values add up to $2C$, and the actual C value is arbitrary within reason.

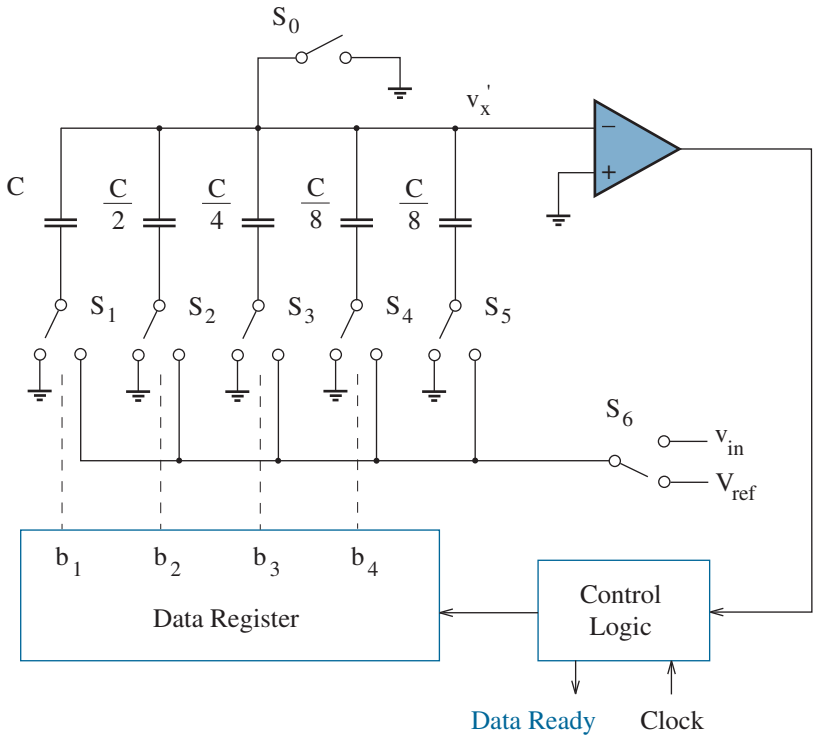


Figure 14.25: Successive-approximation charge-redistribution ADC.

Before the circuit of Fig. 14.25 can make meaningful decisions, it needs to be initialized so that $v_x{}'$, the voltage at the inverting comparator input, has the value $-v_{in}$. This is accomplished in two steps:

- Close S_0, toggle S_1-S_5 in unison so that the bottom capacitor plates connect to S_6, and force S_6 to v_{in}. The node voltages at the top and bottom capacitor plates are now 0 and v_{in}, respectively.
- Open S_0 and set S_1-S_5 in unison so that the bottom capacitor plates connect to ground. The total capacitor charge remains unchanged—there is no discharge path. Thus, $v_x{}'$ has the desired initial value.

The conversion process is ready to proceed after S_6 connects to V_{ref}.

To effect the first bit trial, toggle S_1 to V_{ref} so that Fig. 14.26 applies. The initial charge stored in the capacitors is $-2\,Cv_{in}$; whereas, the final charge stored after S_1 toggles is $C(v_x{}' - V_{ref})$ in the S_1-connected capacitor and $Cv_x{}'$ in the other capacitors. But total storage is invariant to S_1 action. So equate the initial and final total charge stored and solve for $v_x{}'$ to find

$$v_x{}' = -v_{in} + (b_1\,2^{-1} + b_2\,2^{-2} + b_3\,2^{-3} + b_4\,2^{-4})\,V_{ref} \tag{14.15}$$

subject to $b_1 = 1$ and $b_2 = b_3 = b_4 = 0$.

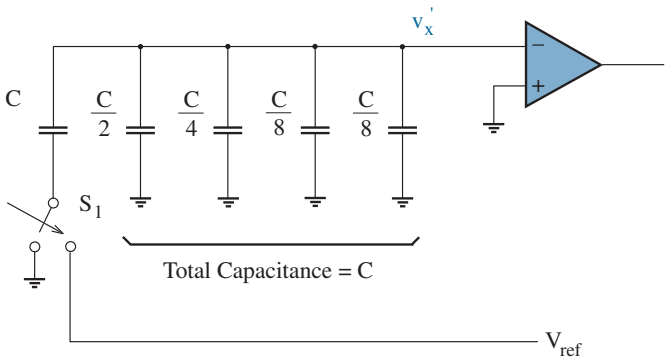


Figure 14.26: Charge-redistribution circuit for the first bit trial.

We now have two possible outcomes:

- If the comparator output is HIGH, then $v_x{}' < 0$ and $v_{in} > V_{ref}/2$. Conclude that $b_1 = 1$, and maintain the S_1 connection to V_{ref}.
- If the comparator output is LOW, then $v_x{}' > 0$ and $v_{in} < V_{ref}/2$. Conclude that $b_1 = 0$, and restore the S_1 connection to ground.

Maintaining the S_1 connection to V_{ref} preserves $v_x{}'$, and restoring the S_1 connection to ground redistributes capacitor charge to the pre-trial state. Equation 14.15 applies to both cases subject to the particular b_1 conclusion.

Now effect the second bit trial by toggling S_2 to V_{ref}. One can determine the new $v_x{}'$ by invoking charge conservation as before. However, one obtains the same result by invoking superposition and a capacitor divider relation. The total capacitance positioned between S_2 and $v_x{}'$ is $C/2$, and the total capacitance positioned between $v_x{}'$ and other nodes is $3C/2$. In turn,

$$\Delta v_x{}' = \left(\frac{2/3C}{2/C + 2/3C}\right) V_{ref} = 2^{-2}\,V_{ref}\,. \tag{14.16}$$

This change in $v_x{}'$ is equivalent to setting $b_2 = 1$ in Eq. 14.15.

Once again, we have two possible outcomes based on comparator output:

- If the output is HIGH: $v_{in} > (b_1\,2^{-1} + b_2\,2^{-2} + b_3\,2^{-3} + b_4\,2^{-4})\,V_{ref}$. Conclude that $b_2 = 1$, and maintain the S_2 connection to V_{ref}.
- If the output is LOW: $v_{in} < (b_1\,2^{-1} + b_2\,2^{-2} + b_3\,2^{-3} + b_4\,2^{-4})\,V_{ref}$. Conclude that $b_2 = 0$, and restore the S_2 connection to ground.

Two similar processes determine the third and fourth bits.

Exercise 14.6 Consider a 6-bit successive-approximation ADC with $V_{ref} = 2.048$ V. Find the $v_x{}'$ trial values and output code if $v_{in} = 0.900$ V.

Ans: $v_x{}' = 0.124$ V, -0.388 V, -0.132 V, -0.004 V, 0.060 V, 0.028 V

Output code = 0 1 1 1 0 0

The shrewd reader may be wondering why the circuit of Fig. 14.25 is not configured with v_{in} at the non-inverting comparator input. This would have eliminated S_5 (always ground-connected) and S_6 (always V_{ref}-connected), and it would have simplified the initialization process (close S_0 with S_1 - S_4 at ground to discharge the capacitors). The modified circuit also establishes $v_x{}'$ as v_x in the process of Fig. 14.24.

Recall from Chapter 1 that a non-ideal op-amp or comparator suffers from an input offset voltage—the component behaves as if its non-inverting input has a series-connected voltage source with value v_{os}. Although seldom more than a few mV, an input offset voltage disrupts the conversion process, especially at or near the LSB range. Thus, it is common practice to employ a comparator with an **auto-zeroing** mechanism such as that in Fig. 14.27. Here, S_a closes with S_b at ground so that C_x charges to v_{os}. Then with S_a open and S_b at $v_x{}'$, the voltage at the inverting comparator input is $v_x{}'+v_{os}$. This voltage is compared with v_{os}, so the effect is the same as a comparison between $v_x{}'$ and zero. The auto-zeroing process does not work if the non-inverting comparator input is connected to any voltage other than ground.

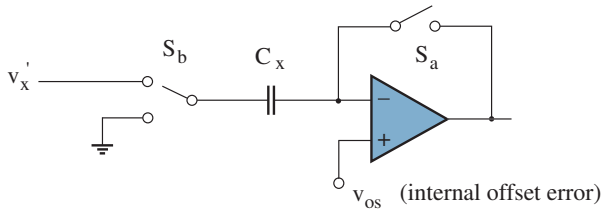


Figure 14.27: Auto-zeroing comparator circuit.

14.4 Oversampling A/D Converters

The preceding ADCs typically operate at or just above the Nyquist rate in which sampling frequency (ω_s) is twice the maximum signal frequency (ω_c). When sampling exceeds the Nyquist rate, the quantization-error noise power is uniformly distributed within the angular frequency interval $[-\omega_s/2, \omega_s/2]$ as shown in Fig. 14.28. Notwithstanding, the filtering process that recovers the original signal removes "out-of-band" higher-order harmonics and noise over the range $|\omega| > \omega_c$. Thus, the "in-band" quantization noise power is

$$P_Q = \int_{-\omega_c}^{\omega_c} \frac{1}{\omega_s} \frac{{v_{LSB}}^2}{12}\, d\omega \;=\; \frac{{v_{LSB}}^2}{12} \left(\frac{\omega_c}{\omega_s/2} \right) . \tag{14.17}$$

In turn, we use Eq. 14.12 to obtain an oversampling signal-to-noise ratio:

$$\text{SNR} = 6.02\, N + 1.76 + 10 \log \text{OSR} \quad \text{dB} . \tag{14.18}$$

Here,

$$\text{OSR} = \frac{\omega_s/2}{\omega_c} \tag{14.19}$$

is the **oversampling ratio**. Every factor-of-two (octave) increase in OSR improves the signal-to-noise ratio by about 3 dB.

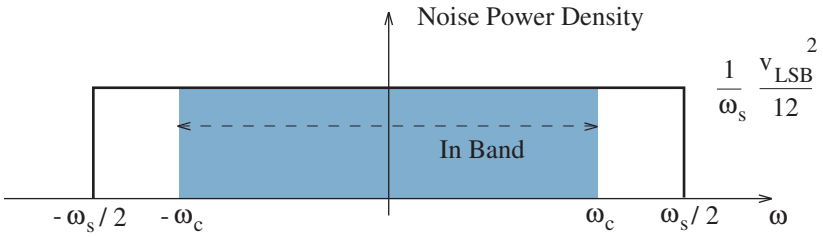


Figure 14.28: Noise spectrum that applies to oversampling conditions.

Chapter 13 established the impossibility of a "brick-wall" anti-aliasing or reconstruction filter that passes all or nothing in relation to a desired ω_c. Thus, an engineering rule-of-thumb is to perform ADC sampling at a rate roughly 10 % higher than the Nyquist specification such that $\omega_s/2 = 1.1\,\omega_c$. This adds 0.8 dB to the ideal SNR.

Exercise 14.7 Estimate the OSR that is required to emulate the noise performance of a 12-bit ADC with: a) an 8-bit ADC; b) a 1-bit ADC.

Ans: a) $4^4 = 256$ b) $4^{11} = 4,194,304$ (not very practical)

The second part of Exercise 14.7 may strike you as pedantic nonsense. Nevertheless, we are about to demonstrate the practical utility of a one-bit ADC when it is part of an oversampling *system* with several components. The one-bit converter features simplicity and inherent "linearity" in the sense of lacking INL or DNL—two output points determine a straight line.

Our starting point is the **Delta modulator** (ΔM) system of Fig. 14.29, which compares an input voltage with its moving average of one-bit round-trip conversions (v_x) as established with the help of an integrator.

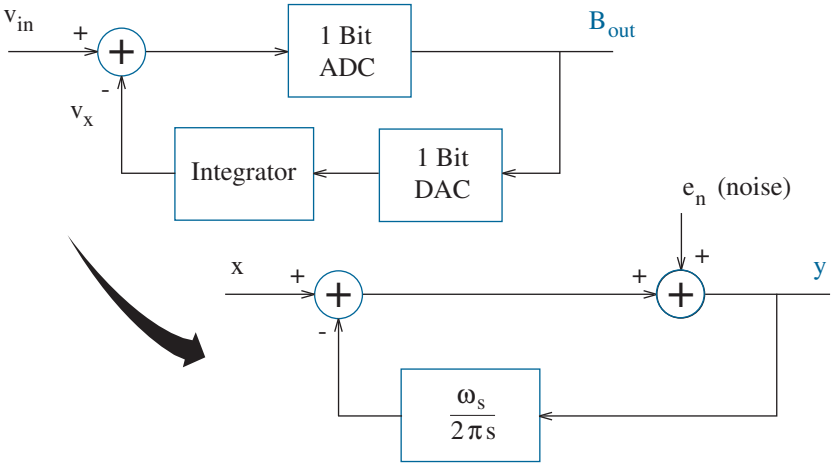


Figure 14.29: Delta modulator and continuous-time system diagram.

The ΔM output bit sequence is a clear guide for analog signal recovery. For example, the sequence

$$1\,1\,0\,1\,0\,0\,0\,0\,1\,0\,1\,1\,0\,1\,0\,1\,0\,1\,0\,1 \;=\; \uparrow\uparrow\downarrow\uparrow\downarrow\downarrow\downarrow\downarrow\uparrow\downarrow\uparrow\uparrow\downarrow\uparrow\downarrow\uparrow\downarrow\uparrow\downarrow\uparrow$$

provides directions for ramping up ($\uparrow$) or down ($\downarrow$) over short time intervals of duration $T_s = \omega_s/2\pi$. Figure 14.30 shows the corresponding waveforms.

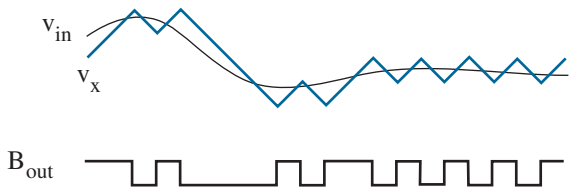


Figure 14.30: Analog signal tracking and recovery for a Delta modulator.

The Delta-modulator system diagram presented in Fig. 14.29 determines the overall response to input signals as well as noise. Here, the 1-bit ADC and DAC operations effectively cancel each other apart from the addition of quantization noise (e_n). The integration bears the signature $1/s$ operation, where $s = \jmath\omega$, and the $\omega_s/2\pi$ integration constant reflects the input sampling frequency.

Warning: It is generally preferable to employ the mathematical tools of discrete-time signal processing when considering oversampling systems. Nevertheless, the continuous-time s-variable approach fully supports the limited set of conclusions that follow, and it is more likely familiar to readers with a rudimentary background in circuit analysis.

To find the response y to noise-signal e_n, we let $x = 0$. Then we express y as the difference between e_n and an integration over y. In turn,

$$H_n = \frac{y}{e_n} = \frac{s}{s + \omega_s/2\pi} \,. \tag{14.20}$$

This is a first-order high-pass response that favorably limits in-band noise. To find the response y to input-signal x, we let $e_n = 0$. Then

$$H_x = \frac{y}{x} = \frac{s}{s + \omega_s/2\pi} \,, \tag{14.21}$$

which is also a high-pass response. The input signal is presumably in-band, so low-pass behavior is preferable. Yet if we first integrate the input signal, the added $\omega_s/2\pi s$ operation transforms Eq. 14.20 to the low-pass response. Integrating both inputs on the left summing node in Fig. 14.29 is the same as performing a single integration on the summed output. We thus obtain the first-order **Sigma-Delta** ($\Sigma\Delta$) modulator of Fig. 14.31.

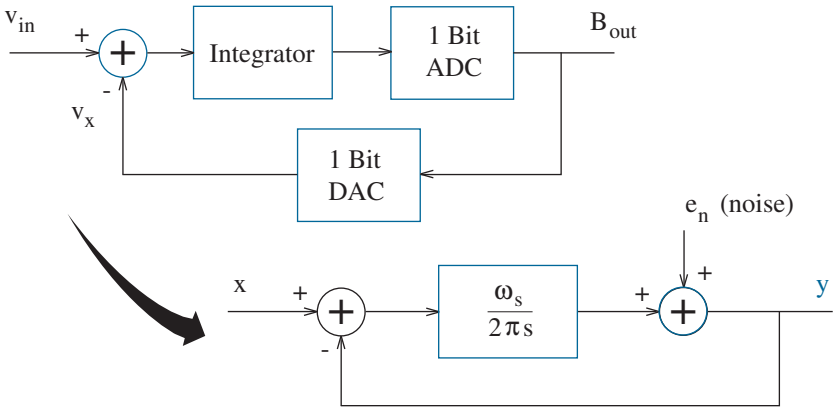


Figure 14.31: First-order Sigma-Delta modulator.

Back to noise. It is little effort to show that the Sigma-Delta modulator exhibits the noise response of Eq. 14.20. Thus, a formerly uniform spectral density of quantization noise power is scaled or "shaped" by the factor ${H_n}^2 = H_n {H_n}^*$. The total in-band quantization noise power is

$$P_Q = \int_{-\omega_c}^{\omega_c} \frac{1}{\omega_s} \frac{{v_{LSB}}^2}{12} \frac{\omega^2}{\omega^2 + {\omega_s}^2/4\pi^2} \, . \tag{14.22}$$

Subject to the substitution $\tan\theta = 2\pi\omega/\omega_s$, the integration of $\tan^2\theta \, d\theta$, and the Taylor expansion of $\tan^{-1}\xi$ for small ξ (see Problem 14.42),

$$P_Q \approx \frac{{v_{LSB}}^2}{12} \frac{\pi^2}{3} \left(\frac{\omega_c}{\omega_s/2} \right)^3 . \tag{14.23}$$

In turn, the Sigma-Delta signal-to-noise ratio is given by

$$\text{SNR} = 6.02\, N + 6.93 + 30 \log \text{OSR} \quad \text{dB} \, , \tag{14.24}$$

where OSR is the familiar oversampling ratio. The SNR improves by about 9 dB for every octave increase in OSR, a three-fold improvement in relation to conventional oversampling. Such is the benefit of **noise shaping**.

We can do much better. The second-order Sigma-Delta modulator of Fig. 14.32 is characterized by

$$P_Q \approx \frac{{v_{LSB}}^2}{12} \frac{\pi^4}{15} \left(\frac{\omega_c}{\omega_s/2} \right)^5 . \tag{14.25}$$

and

$$\text{SNR} = 6.02\, N + 9.88 + 50 \log \text{OSR} \quad \text{dB} \, . \tag{14.26}$$

This implies effective 12-bit performance with (practical) OSR = 14.5.

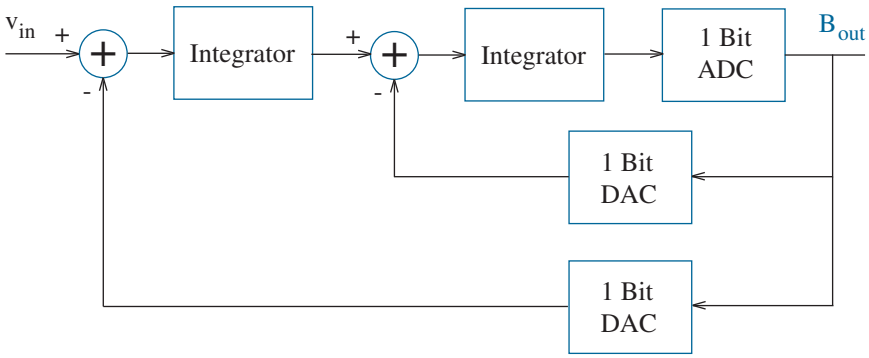


Figure 14.32: Second-order Sigma-Delta modulator.

An Example beckons, pending some preliminary practical details.

In a 1-bit offset binary system, Bins 0 and 1 correspond to the input voltage intervals $[-V_{ref}/2, 0]$ and $[0, V_{ref}/2]$, respectively. Thus, a suitable 1-bit ADC is a comparator that determines the sign of the input voltage. And we will do little to spoil the forthcoming DAC discussion (Section 14.6) by revealing that a 1-bit offset-binary DAC is merely a switch that connects to $-V_{ref}/2$ when offered a 0 or $V_{ref}/2$ when offered a 1.

The integrator is less trivial and is usually implemented in discrete time. Consider the sequence of discrete-time voltages v_{n-1}, v_n, v_{n+1}, v_{n+2} ... According to the trapezoidal rule for numerical integration,

$$\int v\,dt = \left[\ldots \frac{v_{n-1}+v_n}{2} + \frac{v_n+v_{n+1}}{2} + \frac{v_{n+1}+v_{n+2}}{2} + \ldots\right] T_s\,. \quad (14.27)$$

It follows that discrete-time integration is achieved by adding a present and delayed sample as shown in the process of Fig. 14.33. Circuit realizations feature capacitors and an arrangement of switches (see Fig. 13.26).

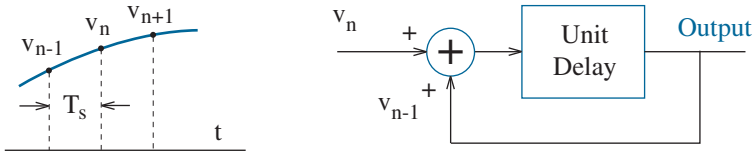


Figure 14.33: Discrete-time integration.

The first-order Sigma-Delta modulator now takes the form of Fig. 14.34. Everything is in place to determine a specific output bit sequence.

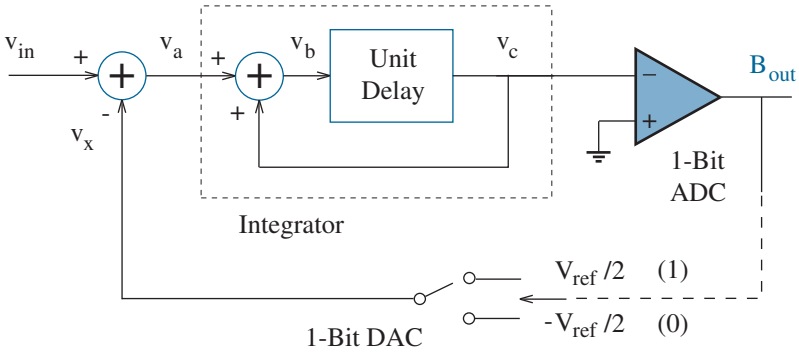


Figure 14.34: Circuit details for a first-order Sigma-Delta modulator.

Example 14.3

A first-order Sigma-Delta modulator with $V_{ref} = 2$ V has $v_{in} = 0.38$ V. Determine the first 16 bits of the output stream.

Solution

We make a table for the analog voltages and binary output in Fig. 14.34. Voltage v_a is the difference between v_{in} and the *previous* DAC output v_x; v_b is the sum of v_a and the *previous* v_c, which, in turn, is the *previous* v_b. If $v_c < 0$, $B_{out} = 0$, and $v_x = -1$ V. But if $v_c > 0$, $B_{out} = 0$, and $v_x = 1$ V. The first row provides initial conditions. Note the "0+" (1E-10) entries.

A spreadsheet program helps to avoid propagating errors.

v_{in}	v_a	v_b	v_c	B_{out}	v_x
0.38	1E-10	1E-10	1E-10	0	0
0.38	0.38	0.38	1E-10	1	1
0.38	-0.62	-0.62	0.38	1	1
0.38	-0.62	-0.24	-0.62	0	-1
0.38	1.38	0.76	-0.24	0	-1
0.38	1.38	1.14	0.76	1	1
0.38	-0.62	0.14	1.14	1	1
0.38	-0.62	0.52	0.14	1	1
0.38	-0.62	-0.48	0.52	1	1
0.38	-0.62	-0.10	-0.48	0	-1
0.38	1.38	0.90	-0.10	0	-1
0.38	1.38	1.28	0.90	1	1
0.38	-0.62	0.28	1.28	1	1
0.38	-0.62	0.66	0.28	1	1
0.38	-0.62	-0.34	0.66	1	1
0.38	-0.62	0.04	-0.34	0	-1
0.38	1.38	1.04	0.04	1	1

Note: The v_x average is 0.375 V (and close to v_{in}).

With all this talk of Sigma-Delta modulators, where is the actual ADC? The preceding Example suggests that the time density of 1's is proportional to the input voltage. This applies, provided the suppression of "favored" numerical results or tones. The conversion of high-rate 1-bit streams to lower-rate N-bit Nyquist data requires **decimation filters** and special digital signal processing—issues best reserved for special texts.

14.5 Integrating A/D Converters

We now arrive at the slowest, but perhaps most accurate analog-to-digital converter, an implementation easily capable of 16 or more bits of resolution. The **single-slope** integrating converter flaunts an op-amp, a comparator, a digital counter, and control logic as shown in Fig. 14.35.

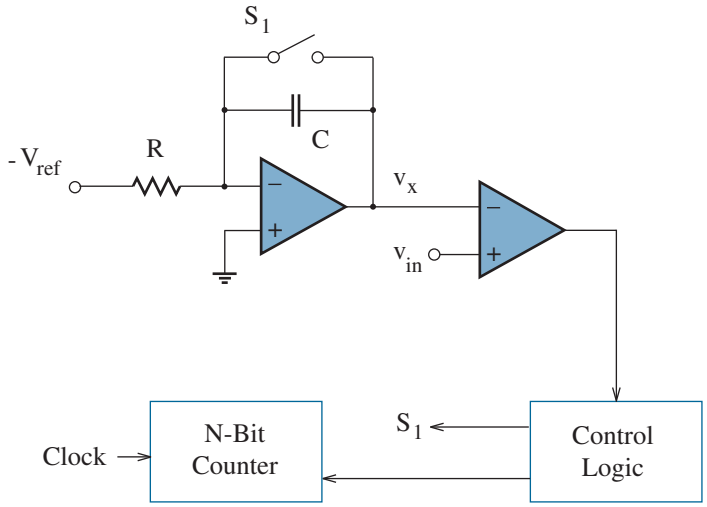


Figure 14.35: Single-slope integrating ADC.

Our focus is v_x, the node voltage at the inverting comparator input. With S_1 closed for all $t < 0$, $v_x = 0$. Now open S_1 so that v_x can change, and start counting. It is easy to show that voltage v_x assumes the form

$$v_x = \frac{-1}{RC} \int_0^t (-V_{ref})\, dt' \tag{14.28}$$

(as you were asked to demonstrate in Problem 1.39). The reference voltage is constant. Thus, the integration produces a v_x ramp starting from zero. Leave it to the comparator to shut down counting at $t = t_1$ when

$$v_{in} = \frac{V_{ref}\, t_1}{RC}\,. \tag{14.29}$$

But $t_1 = nT$, where T is the clock period for the counter. Thus,

$$n = \left(\frac{v_{in}}{V_{ref}}\right) \frac{RC}{T} \tag{14.30}$$

for the binary count. Unfortunately, the n result depends on R, C, and T, all of which are uncertain. And frankly, the circuit is just another rendition of the inefficient counting ADC of Fig. 14.23. Flaunting indeed.

Something has to justify the integrating procedure, and this will emerge as we analyze the **dual-slope** integrating converter of Fig. 14.36.

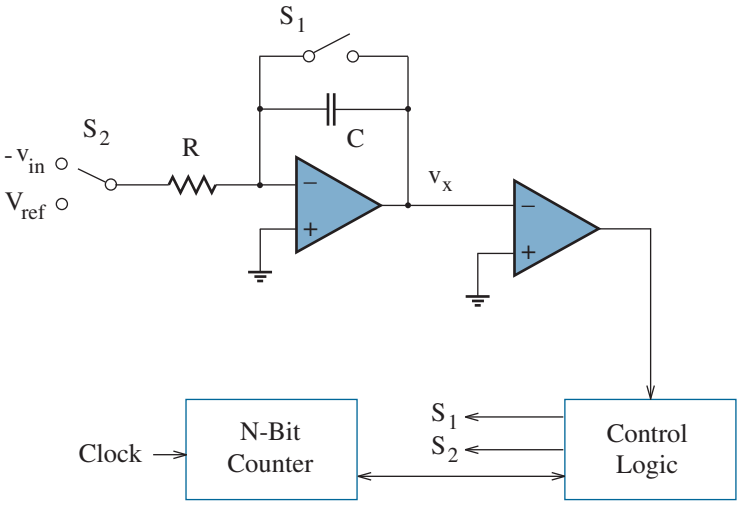


Figure 14.36: Dual-slope integrating ADC.

In Step 1, S_2 connects to $-v_{in}$. Then S_1 opens so that v_x ramps upward with a slope that is proportional to v_{in}. After 2^N clock cycles, v_x assumes a maximum value given by

$$v_{max} = \frac{v_{in}\, 2^N T}{RC}\,, \tag{14.31}$$

where T is the clock period. Figure 14.37 shows the v_x time dependence.

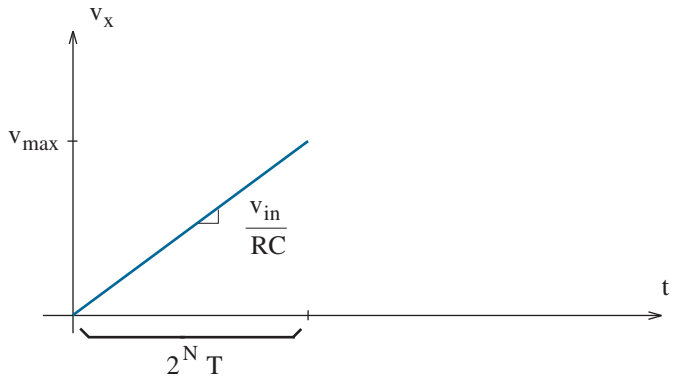


Figure 14.37: v_x time dependence during the first integration.

The completion of 2^N clock cycles—no comparator action is needed—immediately launches Step 2 in which S_2 transitions to V_{ref}. Voltage v_x now ramps *downward* with a relatively steep slope in proportion to V_{ref}. Meanwhile, the counter engages in a second round of counting from zero, and it is quick to cease effort as soon as the comparator detects $v_x < 0$. Figure 14.38 shows the Step-2 v_x time dependence.

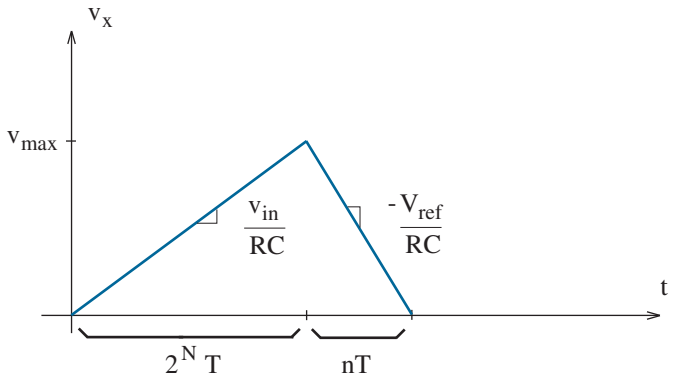


Figure 14.38: v_x time dependence during the second integration.

In Step 1, v_x reaches v_{max} through a ramp with v_{in}/RC slope and $2^N T$ duration. And in Step 2, v_x returns to zero through a ramp with V_{ref}/RC slope magnitude and nT duration. Thus,

$$\left(\frac{v_{in}}{RC}\right) 2^N T = \left(\frac{V_{ref}}{RC}\right) nT \,. \tag{14.32}$$

The T/RC factors cancel in this expression, and

$$n = \left(\frac{v_{in}}{V_{ref}}\right) 2^N \,. \tag{14.33}$$

No uncertainty here (when compared with Eq. 14.30). Be sure to remember the dual-slope integrating process the next time you use a digital voltmeter.

Exercise 14.7 A 16-bit dual-slope integrating ADC features $R = 10$ kΩ, $V_{ref} = 2.048$ V, and $T = 50$ ns (20 MHz). Find C so that $v_{max} \leq 4.8$ V.

Ans: $C = 280$ nF

Exercise 14.8 Let $v_{in} = 3$ V, as applied to the ADC of Exercise 14.7. Estimate the required conversion time.

Ans: Conversion time = 5.68 ms

14.6 Digital-to-Analog Converters

In comparison with the opposite journey, digital-to-analog conversion is a straightforward procedure. The binary voltage-bin assignments are known —one need only provide appropriate mapping to a specific analog voltage. As always, speed, complexity, and cost are crucial factors.

String Converters

The simplest, fastest DAC has the string configuration shown in Fig. 14.39 for the case $N = 3$ bits. With $2^N = 8$ elements, the resistor chain is similar to that used for the flash ADC. However, the node voltages along the chain should relate to bin *midpoints* instead of edges, so the R values are equal. An N-stage set of bit-controlled switches connects a selected node voltage to the DAC output through an op-amp buffer that isolates resistive loads.

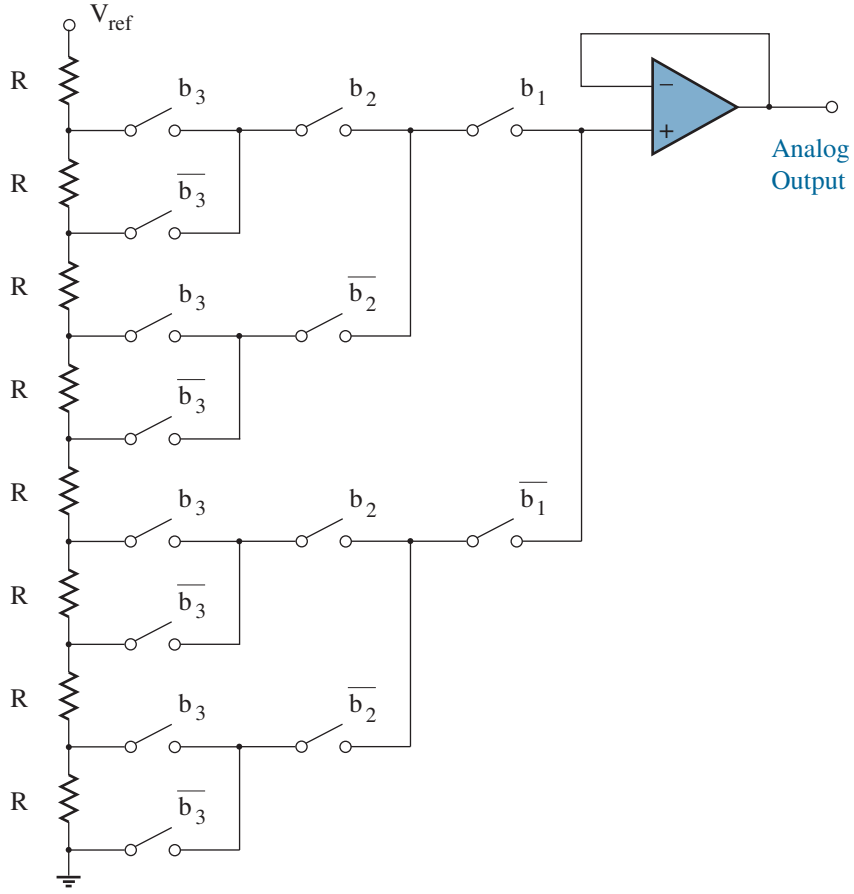


Figure 14.39: 3-bit string digital-to-analog converter.

The N-bit string DAC requires $2 \times (2^N - 1)$ switches and 2^N resistors. Apart from exponential size requirements and cost, large-N string DACs suffer N-fold delays through the cascaded switches that separate resistor-chain nodes from the output buffer. Delays are significantly reduced with the help of an N-to-2^N decoder that controls a single switch at every node. Nevertheless, the dynamic improvement is achieved at the expense of even greater circuit complexity.

As for the flash ADC, component count decreases when DAC operations are performed in two steps. Specifically, one uses an N_1 row by N_2 column architecture, where $N_1 + N_2 = N$. Figure 14.40 shows a 4-bit 2×2 design. Four "coarse" resistors (R_c) establish voltage-bin midpoints that apply to each combination of the first two bits, and four sets of four "fine" resistors (R_f) establish appropriate subranges in relation to the second two bits. Two 2-to-4 decoders determine a row and column. This small-N example achieves a savings of ten switches, two resistors, and two switch delay times. Expect far greater savings for larger N.

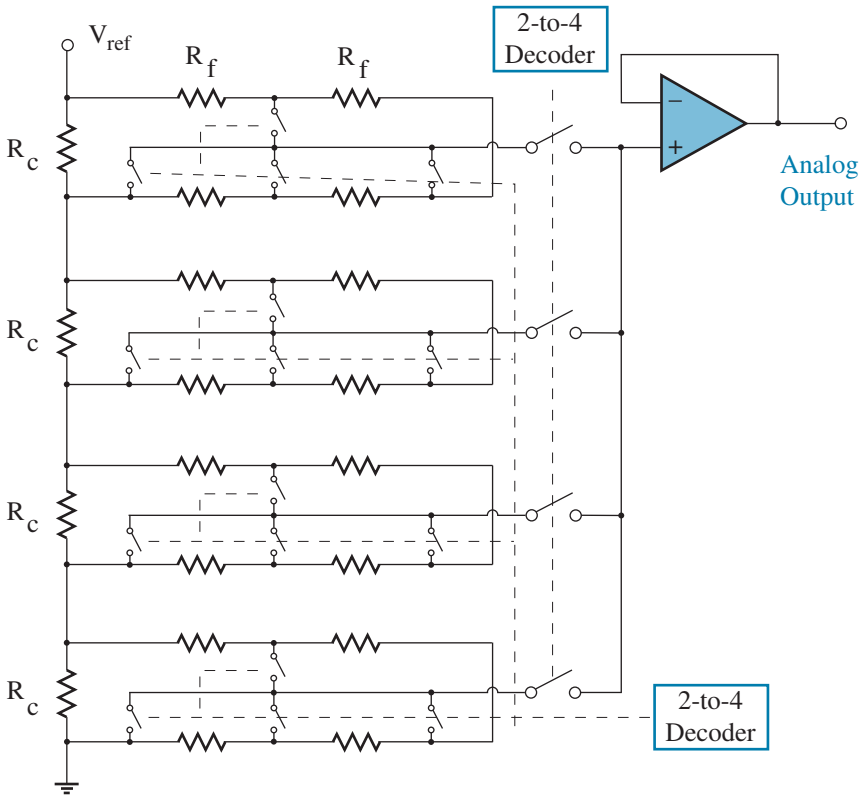


Figure 14.40: 4-bit 2×2 DAC configuration.

Current-Mode Converters

The preceding digital-to-analog converters function by mapping a digital word to an analog voltage located at a particular node in a resistor string. A different process constructs the output as a sum of voltages that are either weighted in proportion to bit significance or taken in separate units of v_{LSB}. Adding voltages is relatively difficult compared to the addition of currents, which need only share a node for KCL. Thus, it is common practice to design current-mode processes that are followed by a current-to-voltage conversion.

Figure 14.41 shows a simple 4-bit current-mode DAC. Through feedback, the op-amp sustains its inverting input at the same ground potential as the non-inverting input. Thus, the current i_1 takes the form

$$i_1 = b_1 \frac{0 - (-V_{ref})}{2R} , \tag{14.34}$$

where $b_1 = 1$ if the series-connected switch is closed and $b_1 = 0$ otherwise. Similar expressions apply to currents i_2, i_3, and i_4. The sum of all of these currents flows through R, since the ideal op-amp draws no current. Thus,

$$v_{out} = R\left(i_1 + i_2 + i_3 + i_4\right) , \tag{14.35}$$

as for a current-to-voltage converter (with a first appearance in Chapter 1). Then with the help of Eq. 14.34 and others like it,

$$v_{out} = V_{ref}\left(b_1\, 2^{-1} + b_2\, 2^{-2} + b_3\, 2^{-3} + b_4\, 2^{-4}\right) . \tag{14.36}$$

The form of this expression reflects standard binary coding (Eq. 14.2).

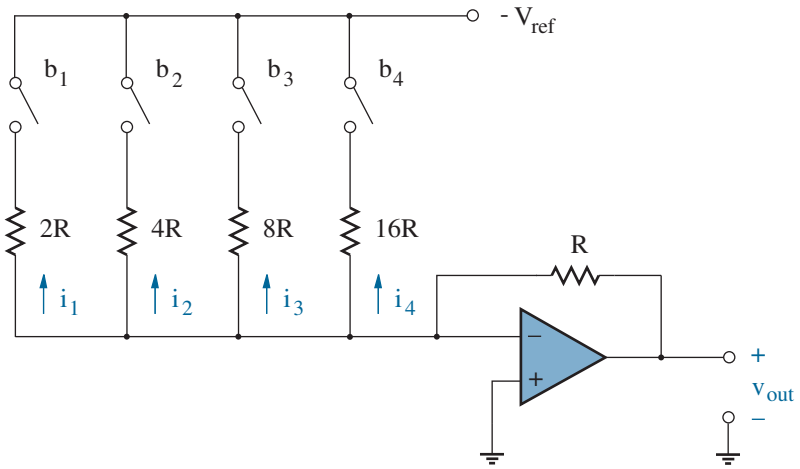


Figure 14.41: Current-mode DAC configuration.

Current-mode DACs with the general form of Fig. 14.41 often suffer an impractical spread in component values when the number of bits is large. For example, a 10-bit DAC accommodates resistors in the range R - $1024\,R$. Fortunately, the spread is greatly reduced with the not-just-for-textbooks "R-$2R$" network of Fig. 14.42. The $2R$ resistors either terminate at ground or virtual ground depending upon the positions of the bit-controlled switches, so the resistance looking to the right of any one of the dashed vertical lines is always $2R$. In turn,

$$i_{ref} = \frac{V_{ref}}{2R \parallel 2R} = \frac{V_{ref}}{R} \,. \tag{14.37}$$

This current divides equally between two $2R$ paths. Specifically,

$$i_1 = \frac{V_{ref}}{2R} = i_{x1} \,. \tag{14.38}$$

Similar divisions scale the remaining currents by successive factors of two. One thus obtains the desired output characteristic of Eq. 14.36.

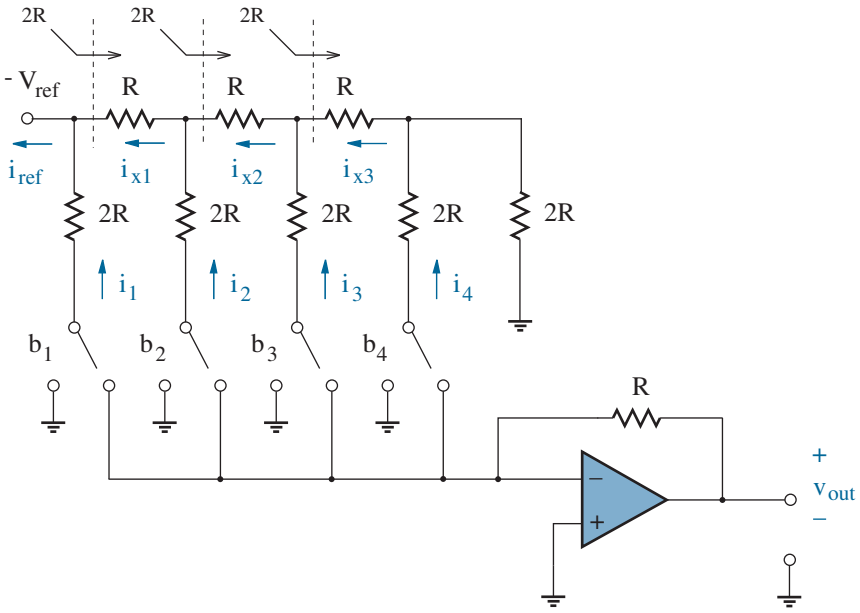


Figure 14.42: Current-mode DAC with "R-$2R$" current scaling.

The two resistor networks in Figs. 14.41 and 14.42 are essentially current-source realizations. Source-coupled MOSFET pairs can be used to divert specific currents between ground and virtual ground with very high speed. Current-level spread remains a problem.

Charge-Mode Converters

Another DAC option processes buckets of charge in geometric sizes rather than steady flows of charge (currents) in rivers, streams, and bare trickles. Figure 14.43 shows the enabling switched-capacitor charge-packet amplifier.

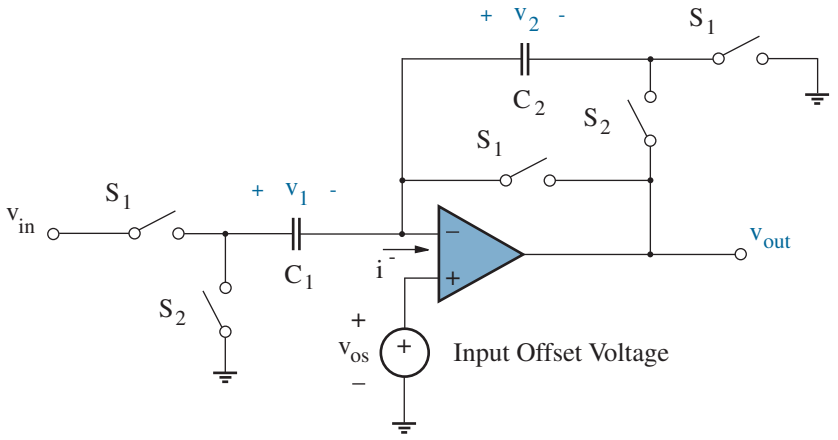


Figure 14.43: Switched-capacitor charge-packet amplifier.

In the first phase, the S_1 switches are closed and the S_2 switches are open. The charge stored in capacitor C_1 is

$$Q_1 = C_1 \left(v_{in} - v_{os} \right), \tag{14.39}$$

and the charge stored in capacitor C_2 is

$$Q_2 = C_2 \left(0 - v_{os} \right), \tag{14.40}$$

where v_{os} is the input offset voltage for the op-amp. In the second phase, the S_1 switches are open and the S_2 switches are closed. In turn,

$$Q_1 = C_1 \left(0 - v_{os} \right), \tag{14.41}$$

and

$$Q_2 = C_2 \left(v_{out} - v_{os} \right). \tag{14.42}$$

Now let ΔQ denote the difference between second- and first-phase charge. KCL at the non-inverting node requires $\Delta Q_1 + \Delta Q_2 = 0$ subject to $i^- = 0$. So after some algebra, we find

$$v_{out} = \frac{C_1}{C_2} v_{in} \tag{14.43}$$

after the second clock cycle. Note that v_{os} does not influence the result.

Some simple changes produce a 4-bit charge-mode DAC. Let $v_{in} = V_{ref}$, let $C_2 = 16\,C$, and replace C_1 with a capacitor network shown in Fig. 14.44. Here, the bit-controlled switches connect the bottom capacitor plates to ground ($b_x = 0$) or virtual ground ($b_x = 1$). Thus, the total C_1 capacitance connected to the latter is

$$C_1 = 8\,b_1 C + 4\,b_2 C + 2\,b_3 C + b_4 C\,. \tag{14.44}$$

In turn, with the particular choice for C_2,

$$v_{out} = V_{ref}\left(b_1\,2^{-1} + b_2\,2^{-2} + b_3\,2^{-3} + b_4\,2^{-4}\right)\,. \tag{14.45}$$

at the end of the second clock cycle. The actual C value does not matter (within reason), so the circuit functions with small component dimensions.

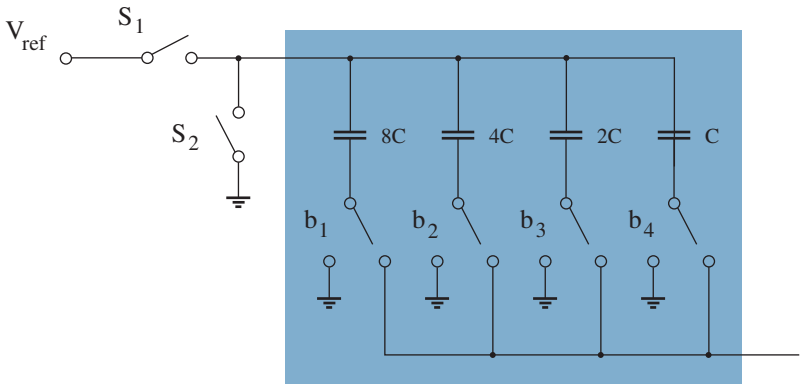


Figure 14.44: Scaled capacitor network that replaces C_1 in Fig. 14.43.

A final remark—

It should be obvious that the many switches in the preceding ADC and DAC circuits are not of the toggle variety that require mechanical action (perhaps effected through a supporting cast of well-conditioned gremlins). Instead, the switches are:

- Small, and fully integrated with other components;
- Characterized by near-infinite "off" and near-zero "on" resistances;
- Controlled by HIGH and LOW digital signals, but able to pass analog signals over the full voltage range between 0 and V_{ref};
- Able to support charge-mode processes without corruptive addition or subtraction of parasitic charge packets.

Chapter 5 addressed the confrontation of these objectives with MOSFETs.

Problems

Section 14.1

14.1 Show that the two-op-amp pan-pot circuits in Fig. 14.2 have the intended function for a particular parameter ζ.

14.2 Show that the left and right mixers in Fig. 14.2 have the intended function.

Section 14.2

14.3 Consider a 1.55-V level with $V_{ref} = 4$ V.

(a) Determine the 8-bit digital representation in natural binary code.

(b) Determine the 8-bit digital representation in offset binary code.

14.4 Consider the code 01010111 with $V_{ref} = 4$ V.

(a) Determine the corresponding analog voltage if the code is natural binary.

(b) Determine the corresponding analog voltage if the code is offset binary.

14.5 In two's-complement subtraction, the negative binary $-y$ to be added to a positive x is formed by complementing each bit in y and adding one. For the case of n bits, carry-bit $n + 1$ is ignored.

(a) Consider the interval [-8,7]. Show that the two's-complement binary code is equivalent to offset binary code apart from the complementary condition of the most significant bit.

(b) Apply two's-complement coding to show that 7 + (-5) = 2.

(c) Apply two's-complement coding to show that 5 + (-7) = -2.

14.6 The sign-plus-magnitude code uses the most significant bit to indicate a positive (0) or negative (1) condition. The remaining bits reflect the natural binary code for an absolute numerical value.

(a) Determine the sign-plus-magnitude codes over the interval [-7,7].

(b) Consider an analog voltage that changes from $+v_{LSB}/4$ to $-v_{LSB}/4$. Determine the bit changes that correspond to sign-plus-magnitude and offset binary coding.

(c) Sign-plus-magnitude coding is useful for instruments such as a digital voltmeter. Explain.

14.7 A three-bit ADC with $V_{ref} = 5$ V exhibits digital transitions at the following input voltages: 0.591, 1.212, 1.860, 2.542, 3.157, 3.714, 4.342. Determine the applicable offset error, gain error, INL, and DNL (all expressed in fractional v_{LSB}).

14.8 A three-bit DAC with $V_{ref} = 5$ V exhibits the following sequence of output voltages: 0.631, 1.258, 1.896, 2.492, 3.131, 3.782, 4.363. Determine the applicable offset error, gain error, INL, and DNL (all expressed in fractional v_{LSB}).

14.9 A three-bit ADC with $V_{ref} = 4$ V has been adjusted so that the offset and gain errors are zero. The successive INL voltages are 0.486, 0.492, 0.528, 0.516, 0.496, 0.533, 0.501. Determine the input voltages that produce step transitions.

14.10 A three-bit DAC with $V_{ref} = 4$ V has been adjusted so that the offset and gain errors are zero. The successive DNL voltages are -0.015, -0.026, +0.018, +0.003, -0.012, -0.021. Determine the output voltages as v_{in} increases from zero.

14.11 Show that an ADC can exhibit a missing code when one of its DNL errors exceeds one v_{LSB}.

14.12 Show a DAC can lose output monotonicity when one of its DNL errors exceeds one v_{LSB}.

14.13 Use a figure similar to Fig. 14.15 to prove the Nyquist rate condition: $\omega_s = 2\omega_c$.

14.14 A 12-bit ADC is operates at 2×10^8 samples per second. Determine the upper limit of acceptable jitter that ensures accuracy to within $\pm v_{LSB}/4$.

14.15 An oscillator produces a voltage of the form

$$v_c(t) = A\ \cos(\omega t + \Delta\phi)\ ,$$

where ω is the angular oscillation frequency and $\Delta\phi$ is a randomly varying phase-noise component. For simplicity, assume that this signal causes a 10-bit ADC to make a conversion whenever $v_c = 0$. Determine the maximum acceptable $\Delta\phi$ if the ADC is to perform 5×10^7 conversions per second with $\pm v_{LSB}/2$ accuracy.

14.16 Analog-to-digital quantization noise can be modeled as a sawtooth signal that varies between $\pm v_{LSB}/2$ over a sampling period T (Fig. P14.16). Use this model to justify Eq. 14.10.

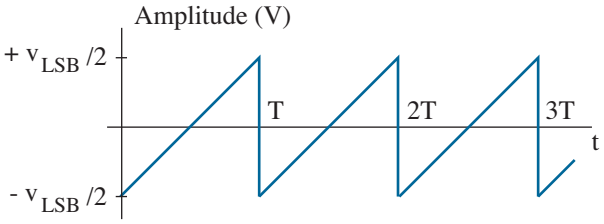


Figure P14.16

14.17 Figure P14.17 provides FFT data for a MAX1214 12-bit ADC. Explain the basis for the SNR and SINAD specifications, and determine the effective number of bits (ENOB). Data: Copyright Maxim Integrated Products. Used by permission.

14.18 Figure P14.18 provides FFT data for a MAX12555 14-bit ADC. Explain the basis for the SNR and SINAD specifications, and determine the effective number of bits (ENOB). Data: Copyright Maxim Integrated Products. Used by permission.

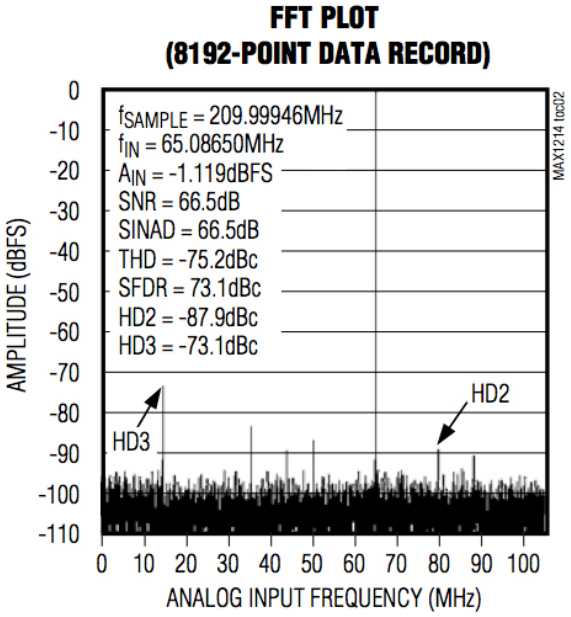


Figure P14.17

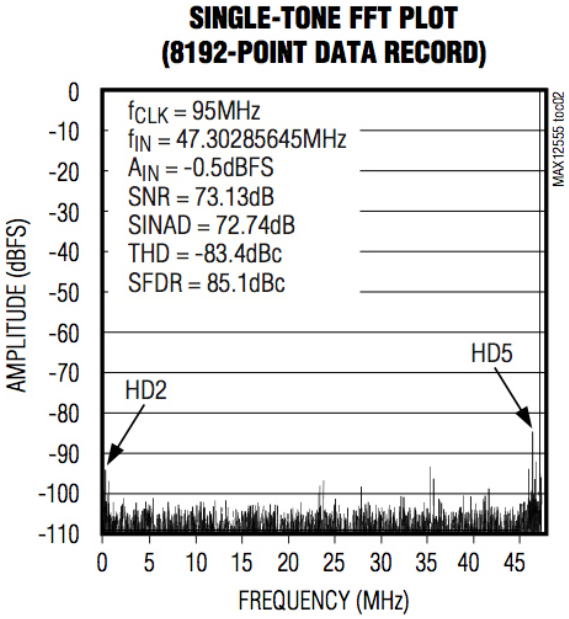


Figure P14.18

Section 14.3

14.19 The 3-bit flash ADC of Fig. 14.18 has the following resistor values (in kΩ) from the top to the bottom of the chain: 15.23, 10.14, 9.85, 10.02, 10.21, 9.97, 10.15, 4.89. Determine the INL and DNL.

14.20 An 8-bit flash ADC is similar to the form of Fig. 14.18 with $V_{ref} = 5$ V and $R = 1$ kΩ. Each nonideal comparator draws current i into its negative input from the resistor chain.

(a) Describe the qualitative effect on INL.

(b) Determine the value for i that is consistent with a maximum INL of $v_{LSB}/2$.

Hint: This is one of the very few problems for which mesh equations are essential.

14.21 Flash ADCs with large numbers of bits often utilize simple comparator designs to minimize cost. In the circuit of Fig. P14.21, the CMOS inverter has $K_R = 1$ and $V_{Tn} = |V_{Tp}|$.

(a) Determine v_c at $t = 0-$ when S_1 is closed and S_2 is open.

(b) Determine v' at $t = 0+$ when S_2 is closed and S_1 is open.

(c) Show that v_{out} is HIGH for $v_{in} > v_x$.

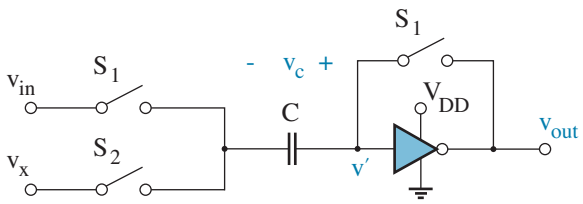


Figure P14.21

14.22 Discuss the noise-margin requirements for the CMOS inverter of Problem 14.21.

14.23 Design a simple encoder for a 3-bit flash ADC and indicate how the number of gates scales with N, the number of conversion bits.

14.24 Design a simple digital circuit that eliminates the effect of a single bubble error in a flash ADC.

14.25 Look up the data sheet for the MAX108 (flash) ADC and describe the timing operations required for a single conversion. (www.maxim-ic.com)

14.26 Prove that an $N + M$ subranging ADC minimizes the number of comparators when $N = M$.

14.27 Figure P14.27 shows a simple architecture for a 4-bit $(2 + 2)$ subranging ADC. For simplicity, voltage-bin boundaries center on $(n + 1/2)\ v_{LSB}$, where $n = 0, 1, \ldots$ The special *three-input* comparators produce HIGH or LOW outputs that depend on the sign following subtraction of two - inputs from a + input (see Problem 14.28). The comparator outputs control encoders that yield separate 2-bit results for coarse and fine conversions.

(a) Specify the algebraic value for R'.

(b) Specify the assignments for voltages v_{xc} and v_{xf}, and indicate how the circuit should be modified so that these voltages are properly derived.

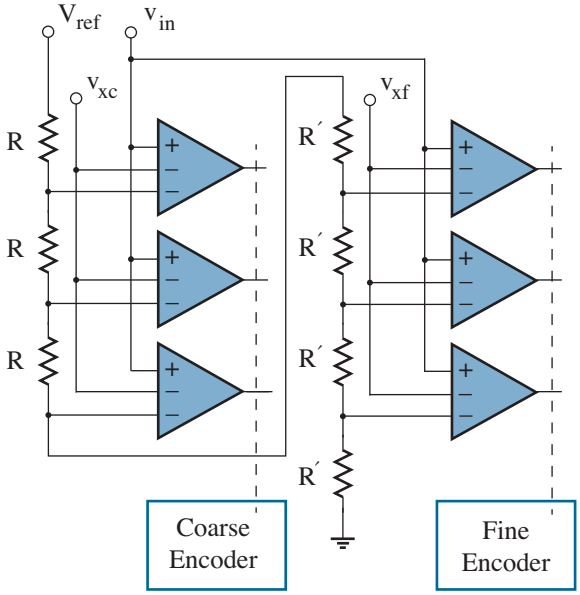


Figure P14.27

14.28 Figure P14.28 shows a circuit that implements the three-input comparator function described in the preceding problem. The CMOS inverter has $K_R = 1$ and $V_{Tn} = |V_{Tp}|$.

(a) Determine v' with S_1 closed and S_2 open.

(b) Now let S_1 open while S_2 closes. Apply charge conservation to show that v_{out} goes HIGH for $v_{in} - v_x - v_y > 0$.

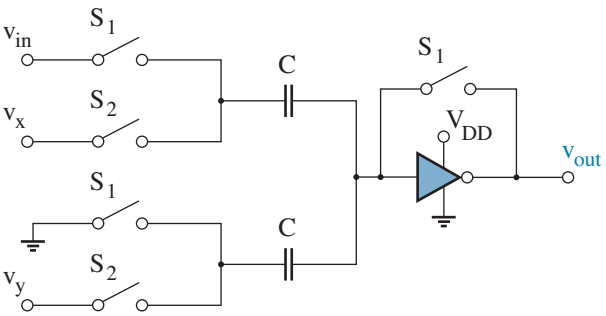


Figure P14.28

14.29 Figure P14.29 shows a coarse resistor string (R_c) with intermeshed fine resistor strings (R_f) for a proposed 4-bit (2+2) subranging ADC. The voltage-bin boundaries are centered as in Problem 14.27.

(a) Show how two sets of three comparators and a set of switches can be used to make the coarse and fine conversions in two steps.

(b) What circuit modifications allow only one set of three comparators?

14.30 Consider a 12-bit pipeline ADC. Determine the comparator count for a process that requires: a) 4×3 bits, b) 3×4-bits, c) 2×6-bits.

14.31 Consider the 3×4 pipeline converter in the example of Figs. 14.20 and 14.21. Discuss the effects of bit errors in the 4-bit ADC, bit errors in the 4-bit DAC, and multiplication errors in the $\times 16$ operation.

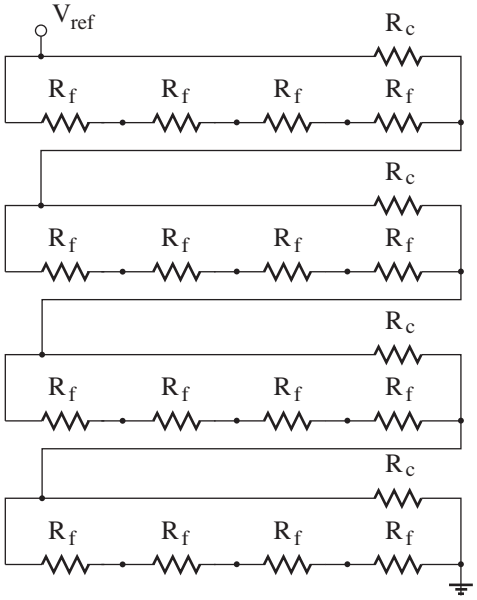


Figure P14.29

14.32 Consider the 1-bit-overlap error-correction process in the discussion that follows Fig. 14.22. Determine the intermediate numerical results if the correct output is 110001011010.

14.33 Consider the 1-bit-overlap error-correction process in the discussion that follows Fig. 14.22. Determine the intermediate numerical results if the algorithm is modified as:

(a) a series of 4-bit conversions followed by a single 5-bit conversion,

(b) a series of 2-bit conversions followed by a single 3-bit conversion.

14.34 Consider the 1-bit-overlap error-correction process in the discussion that follows Fig. 14.22. For some reason, the least significant bit in the 3-bit ADC always produces a "0" regardless of input.

(a) Show that the error-correction process preserves the correct output for the case 110001011010.

(b) What happens if the second-least significant bit is stuck at zero? Why?

14.35 A six-bit successive-approximation ADC has $V_{ref} = 4.096$ V. Determine the intermediate trial voltages and output code if $v_{in} = 3.145$ V.

14.36 Show how the 4-bit successive-approximation process of Fig. 14.24 is modified if the voltage-bin boundaries are offset by $v_{LSB}/2$.

14.37 A 4-bit successive-approximation ADC has the form of Fig. 14.25 with $C/4 \rightarrow 3C/8$. Let $v_{in} = 1.47$ V and $V_{ref} = 4$ V. Determine the output code.

14.38 A 4-bit successive-approximation ADC has the form of Fig. 14.25 with $C = 1$ pF. Let $v_{in} = 2.89$ V and $V_{ref} = 4$ V. Determine the output code if the comparator has 50-fF parasitic input capacitance.

14.39 A 4-bit successive-approximation ADC has the form of Fig. 14.25 with $C = 1$ pF. Let $v_{in} = 2.89$ V and $V_{ref} = 4$ V. Determine the output code if the comparator has $v_{os} = 50$ mV.

14.40 A 4-bit successive-approximation ADC has the form of Fig. 14.25 with $C = 1$ pF. Let $v_{in} = 1.74$ V and $V_{ref} = 4$ V. The "on" resistance of any switch is 50 Ω. Use SPICE to examine the settling time required for an accurate result following the first bit trial in which S_1 toggles to V_{ref}.

Section 14.4

14.41 An 8-bit ADC samples signals at ten times the Nyquist rate. Determine the improvement in SNR.

14.42 Work out the details leading to Eq. 14.23.

14.43 Use the block diagram of Fig. 14.32 to find the corresponding second-order noise-shaping function, then use this result to derive Eq. 14.25.

14.44 Given the first- and second-order Sigma-Delta P_Q results, determine a sequence to find nth-order expressions for P_Q and SNR.

14.45 Repeat Example 14.3 with $v_{in} = 1.24$ V.

Section 14.5

14.46 The op-amp for the single-slope integrating ADC in Fig. 14.35 has an input offset voltage v_{os}. Determine the error that results for n.

14.47 The op-amp for the dual-slope integrating ADC in Fig. 14.36 has an input offset voltage v_{os}. Determine the error that results for n.

Section 14.6

14.48 Show the architecture of a 4-bit string DAC.

14.49 Consider the design of an 8-bit string DAC. Determine the resistor and switch count for a process that requires: a) 1×8 bits; b) 2×4 bits.

14.50 The 3-bit string DAC of Fig. 14.39 has the following resistor values (in Ω) from the top to the bottom of the chain: 992, 1012, 1008, 986, 995, 1015, 1027, 978. Determine the INL and DNL.

14.51 The 4-bit current-mode DAC of Fig. 14.41 has altered left-to-right switch-connected resistor values: $2.04R$, $3.95R$, $8.11R$, $15.92R$. Determine the offset error, gain error, INL and DNL. Let $V_{ref} = 4$ V.

14.52 The op-amp in the 4-bit current-mode DAC of Fig. 14.42 has a 15-mV offset voltage (v_{os}). Determine the offset error, gain error, INL and DNL. Let $V_{ref} = 4$ V.

14.53 The 4-bit charge-mode DAC of Fig. 14.43 has altered left-to-right capacitor values in the switching network of Fig 14.44: $7.92C$, $4.12C$, $2.02C$, $0.97C$. Determine the offset error, gain error, INL and DNL. Let $V_{ref} = 4$ V.

14.54 The 4-bit charge-mode DAC of Fig. 14.43 has parasitic capacitance $C/10$ to ground at the negative op-amp input. Determine the offset error, gain error, INL and DNL. Let $V_{ref} = 4$ V.