Analysis and Design of Elementary MOS Amplifier Stages

Boris Murmann

Roger T. Howe

Charles G. Sodini

- Draft Rev. 1.0-

(January 2012)

Contents

1 CHAPTER

Introduction

With the development of the integrated circuit, the semiconductor industry is undoubtedly the most influential industry to appear in our society. Its impact on almost every person in the world exceeds that of any other industry since the beginning of the Industrial Revolution. The reasons for its success are:

- Exponential growth of the number of functions on a single integrated circuit.
- Exponential reduction in the cost per function.
- Exponential growth in sales (economic importance) for approximately forty years.

This growth has led to ever-increasing performance at lower prices for consumer electronics such as cellular phones, personal computers, audio players, etc. The computational power available to the individual has increased to the point that it has changed the way we think about problem solving. Communication technology including wired and wireless networks have fundamentally changed the way we live and communicate.

The innovation responsible for these impressive results is the integration of electronic circuit components fabricated in silicon **integrated circuit** (IC) technology. Today, many of the ICs shaping new applications contain both analog and digital circuitry, and are therefore often called **mixed-signal integrated circuits**. In mixed-signal ICs, the analog circuitry is typically responsible for interfacing with physical signals, and concerned for example with the amplification of a weak signal from an antenna, or driving a sound signal into a loudspeaker. On the other hand, digital circuitry is primarily used for computing, enabling powerful functions such as Fast Fourier Transforms or floating point multiplication.

This module was written as an introduction to the analysis and design of analog integrated circuits in **complementary metal-oxide-semiconductor** (CMOS) technology. In this first chapter, we will motivate this subject by looking at an example of a mixed-signal IC and by highlighting the need for a systematic study of the fundamental principles and proper engineering approximations in analog design.

Chapter Objectives

- Provide a motivation for the study of elementary analog integrated circuits.
- Provide a roadmap for the subjects that will be covered throughout this module.
- Review fundamental concepts for the construction of two-port circuit models.

1-1 Mixed-Signal Integrated Circuits

Figure 1-1 shows a generic diagram of a mixed-signal system, incorporating a mixed-signal integrated circuit. To the left of this diagram are the transducers and media that represent the sources and sinks of the information processed by the system. Examples of input transducers include microphones or photo-

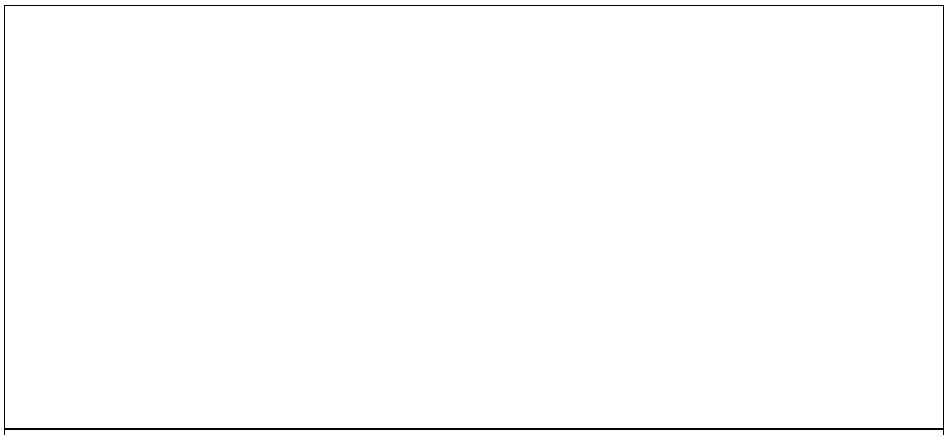


Figure 1-1: Block diagram of a mixed-signal system.

diodes used to receive communication signals from an optical fiber. Likewise, the output of the system may drive an antenna for radio-frequency communication or a mechanical actuator that controls the zoom of a digital camera. At the boundary between the media and transducers are typically signal conditioning circuits that translate the incoming and outgoing signals to the proper signal strength and physical format. For instance, an amplifier is usually needed to increase the strength of the receive signal from a radio antenna, so that it can be more easily processed by the subsequent system components. In most systems, the signal conditioning circuitry interfaces to analog-to-digital and digital-to-analog converters, which provide the link between analog quantities and their digital representation in the computing back-end of the system.

1-1-1 Example: Single-Chip Radio

The block diagram of a modern mixed-signal integrated circuit is shown in Figure 1-2 (see Reference 1). This design incorporates most of the circuitry needed to realize a modern cellular phone. For instance, it contains a front-end low-noise amplifier (LNA) to condition the incoming antenna signal. The amplified signal is subsequently frequency shifted, converted into digital format and fed into a digital processor. Even though the block diagram in Figure 1-2 looks quite complex, all of its elements can be mapped into one of the main blocks of the generic diagram of Figure 1-1.

Figure 1-3 shows the chip photo of the single-chip radio, with some of the system's key building blocks annotated. As evident from this diagram, the digital logic dominates the area of this particular IC. This situation is not uncommon in modern mixed-signal ICs and due to the fact that the utilized digital algorithms have reached an enormous complexity, requiring millions or tens of millions of logic gates.

Despite the dominance of digital logic within most systems, the analog interface components are equally important, as they determine how and how much information can be communicated between the physical world and the digital processing backbone. In many cases, the performance of the signal conditioning and data conversion circuity ultimately determines the performance of the overall system.

1-1-2 Example: Photodiode Interface Circuit

Figure 1-4 shows an example of a signal conditioning circuit that plays a critical role in fiber-optic communication systems. In such a system, a photodiode is used to convert light intensity to electrical current (i_{IN}). In order to condition the signal for further processing, the diode current is converted into a voltage (v_{OUT}) by a so-called **transimpedance amplifier**. This amplifier must be fast enough to process the incoming light pulses, which often occur at frequencies of multiple gigahertz. In addition, the amplifier must obey certain limits on power dissipation, or else the system may become impractical in terms of heat management or power supply requirements.

Limitations in speed and power dissipation are, in general, among the main concerns in the interface circuitry of mixed-signal systems. Since new products tend to demand higher performance, the analog designer is constantly concerned with the design and optimization of system-critical building blocks, aiming for the best possible performance that can be achieved within the framework of the target application and process technology.

A specific example for the circuit realization of a transimpedance amplifier is shown in Figure 1-5. It consists of three transistor stages, each of which serves a specific purpose and design intent. This is true for most amplifier circuits; even though the full schematic of a particular realization may be

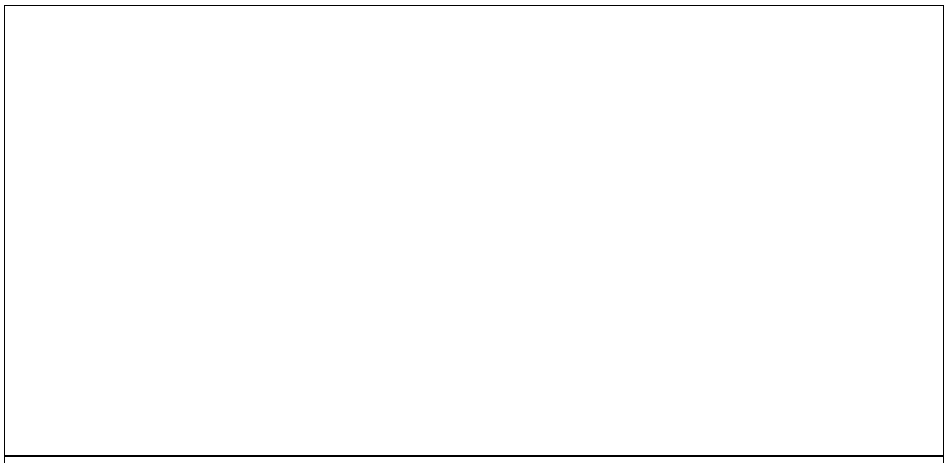


Figure 1-2: Block diagram of a single-chip radio (Reference 1).

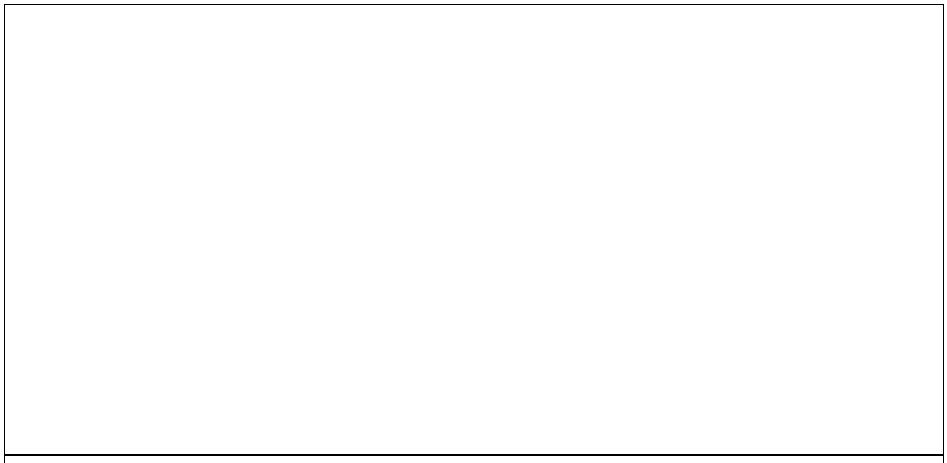


Figure 1-3: Chip photo of a single-chip radio (Reference 1).

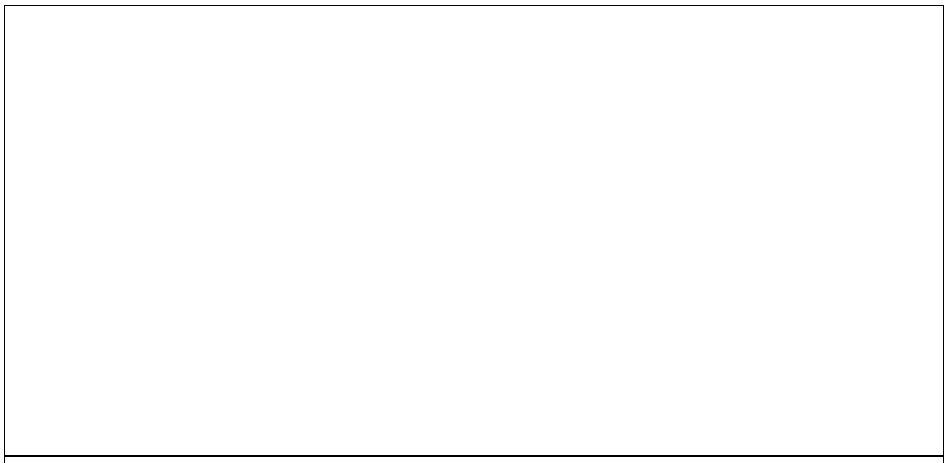


Figure 1-4: Photodetector circuit for fiber-optic communication.

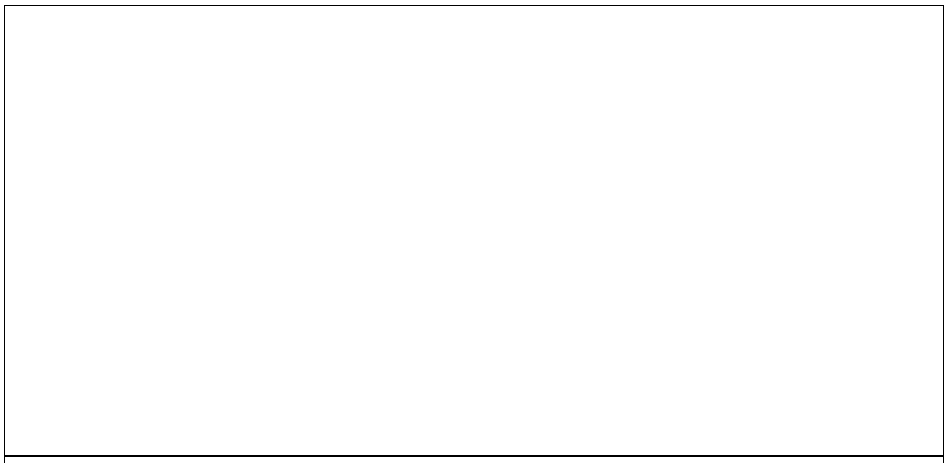


Figure 1-5: Example realization of a transimpedance amplifier.

complex, it can usually be broken up into smaller sub-blocks that are more easily understood. Specifically, for the amplifier of Figure 1-5, the experienced designer will recognize that the circuit consists of a cascade connection of a **common-gate**, **common-source**, and **common-drain** stage. These sub-blocks form the basis for a large number of analog circuits, and be viewed as the "atoms" or fundamental building blocks of analog design. In this module, you will learn to analyze these blocks from first principles, and re-use the gained knowledge for the design of more complex circuits. The circuit of Figure 1-5 will be analyzed in detail in Chapter 6 of this module, building upon the principles covered in Chapters 2-5.

1-2 Managing Complexity

As evident from the example of Section 1-1-1, modern integrated circuits are highly complex and require a hierarchical approach in design and analysis. That is, a modern integrated circuit is typically far too complex to be fully understood and analyzed in a single sheet schematic at the transistor level. Typically, a mixed-signal IC is represented by a block diagram as the one shown in Figure 1-2. At the level of this description, suitable specifications are derived for each block, which itself may contain several sub-blocks. The blocks and sub-blocks are then designed and optimized until they meet the desired target specifications.

Figure 1-6 illustrates examples of the various levels of abstraction that come into play in the design of a modern integrated circuit. At the highest level, the constituent elements can be partitioned into analog and digital blocks. An example of a high-level analog block is an analog-to-digital converter; whereas a microprocessor is an example of a large digital block. These blocks themselves contain smaller functional units, as for example operational amplifiers in the case of an analog-to-digital converter. The operational amplifiers themselves contain the aforementioned elementary transistor stages, which are the main subject of this module.

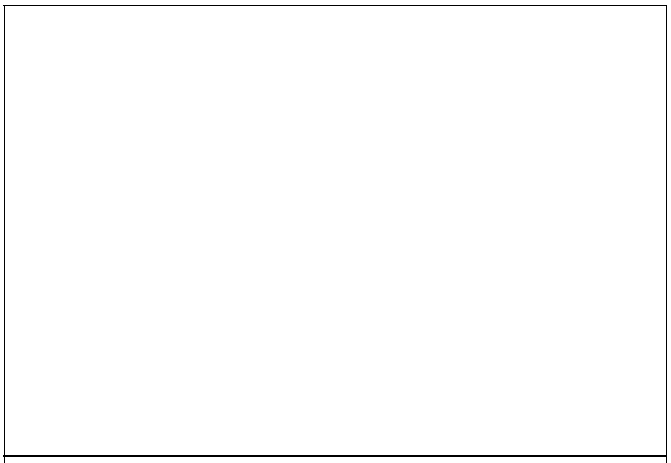


Figure 1-6: Levels of abstraction in integrated circuit design.

Interestingly, even at the level of elementary transistor stages, is often not possible to work with a perfect model or description of the circuit. This is particularly so because the physical effects in the constituent transistors are highly complex and often impossible to capture perfectly with a tractable set of equations for hand analysis. Therefore, making proper engineering approximations in transistor modeling is an important aspect in maintaining a systematic design methodology.

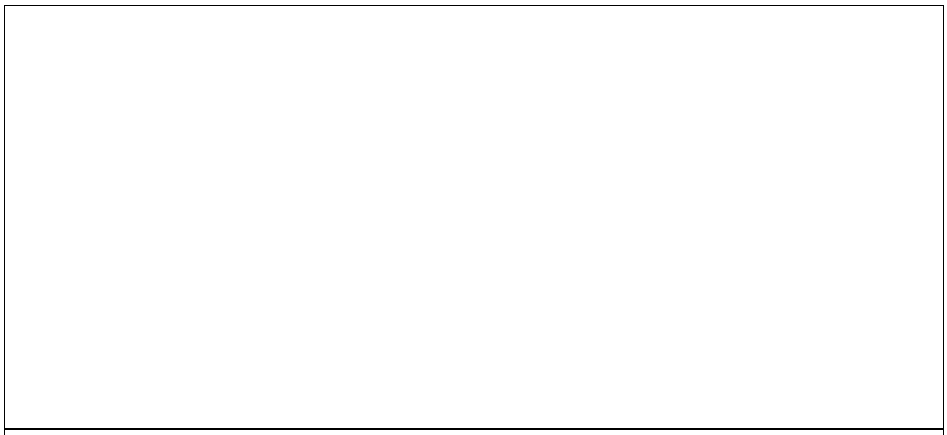


Figure 1-7: Two-port model of the transimpedance amplifier circuit in Figure 1-5.

For this particular reason, the presentation in this module follows a “just-in-time” approach for the modeling of transistor behavior. Rather than deriving a complete transistor model in an isolated chapter (as done in most texts), we begin with only the basic device properties and increase complexity throughout the module upon demand, and where needed to gain further insight and accuracy. With this approach, the reader learns to appreciate the complexity of a refined model, and will be able to assess and track potential limitations of working with simplified models.

1-3 Two-Port Abstraction for Amplifiers

High-level system block diagrams, such as Figure 1-2, are typically drawn as unidirectional flowcharts and do not capture details about the electrical behavior of each connection port and how certain blocks may interact once they are connected. Unfortunately, electrical signals are not unidirectional, and connecting two blocks always means that there is some level of interaction through the voltages and currents at the connection points.

The commonly used linear two-port modeling abstraction for amplifiers and amplifier stages allows the designer to take these effects into account while maintaining a high level of abstraction. For instance, the circuit of Figure 1-5 can be approximately modeled as shown in Figure 1-7 (the details on obtaining this model are discussed later in this module). Each stage of the overall amplifier is represented via a simplified circuit model that captures its essential features. Once this model is created, the interaction among stages can be analyzed at this high level of abstraction, without requiring detailed insight on how each stage is implemented. The two-port modeling approach is particularly useful in the design of amplifiers, as it can help shape the thought process on how the various stages should be configured to optimize performance. In the following sub-sections, we will review some of the basic concepts of amplifier two-port modeling used in this module.

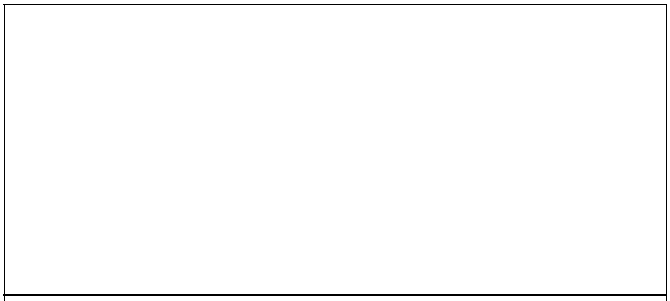


Figure 1-8: General amplifier two-port.

1-3-1 Amplifier Types

In this module, we model amplifier circuits as blocks that have an input and output port, where the term “port” refers to a pair of terminals. For each port, we can define input and output currents and voltages as shown in Figure 1-8. Depending on the intended function, we distinguish between the four possible amplifier types listed in Table 1-1. For example, an amplifier that takes an input current and amplifies this current to produce a proportional output voltage is called a transresistance amplifier. In this context, it is important to emphasize that in a general practical amplifier circuit, the input and output ports will

Table 1-1:

Amplifier Type	Input Quantity	Output Quantity
Voltage Amplifier	Voltage	Voltage
Current Amplifier	Current	Current
Transconductance Amplifier	Voltage	Current
Transresistance Amplifier	Current	Voltage

always carry both non-zero voltages *and* currents, and there exist transfer functions between all possible combinations of input/output variables. What truly defines the type of an amplifier is what the circuit designer deems as the main quantities of interest in the amplifier's application.

Now, in order to model the inner workings of each amplifier type, we can invoke the four corresponding two-port amplifier models shown in Figure 1-9. Each amplifier model has an input and output resistance (or more generally, a frequency dependent impedance) and a controlled source to model the amplification.

- In the **voltage amplifier model**, the controlled source is a voltage-controlled voltage source. Ideally, the input resistance is infinite (open circuit, no current flow). The ideal output resistance is zero (ideal voltage source).
- The **current amplifier model** has a current-controlled current source. Ideally, the input resistance is zero (short circuit, no voltage across the input port) and the output resistance is infinite (ideal current source).
- The **transconductance amplifier model** has a voltage-controlled current source. Ideally, the input resistance is infinite (open circuit, no current flow) and the output resistance is infinite (ideal current source).
- The **transresistance or transimpedance**[1] **amplifier model** has a current-controlled voltage source. Ideally, the input resistance is zero (short circuit, no voltage across the input port) and the output resistance is zero (ideal voltage source).

From these four models and their ideal behavior, we note that the two-ports containing a voltage-controlled source should ideally have large input resistance (R_{in}). This minimizes the signal loss due to resistive voltage division between the source voltage (v_s) and the control voltage (v_{in}). In contrast, the two-ports that use a current-controlled source should have small input resistance to minimize the signal loss due to current division between the source current (i_s) and the control current (i_{in}). In this context, "large" and "small" refer to the value of R_{in} relative to the source resistance (R_s).

On the output side, if the variable of interest is a voltage, the output resistance (R_{out}) should be small so that only a small amount of the amplified voltage is lost through the division with the load (R_L). Conversely, for a current output, the output resistance should be large to minimize current division losses. Again, "large" and "small" are taken as relative measures comparing R_{out} to R_L.

1. The term transimpedance is sometimes used to refer to an amplifier that is primarily meant to realize a transresistance. Referring to "impedance" highlights the fact that the transfer function will usually be frequency dependent.

Figure 1-9: Two-port amplifier models with input source and load: (a) voltage amplifier, (b) current amplifier, (c) transconductance amplifier, and (d) transresistance amplifier.

Consider for example the voltage amplifier of Figure 1-9(a). To calculate the transfer function of the overall circuit (v_{out}/v_s), the input voltage, including its source resistance, is connected to the input of the two-port model and the load resistance is connected to the output. The full circuit is shown in Figure 1-10.

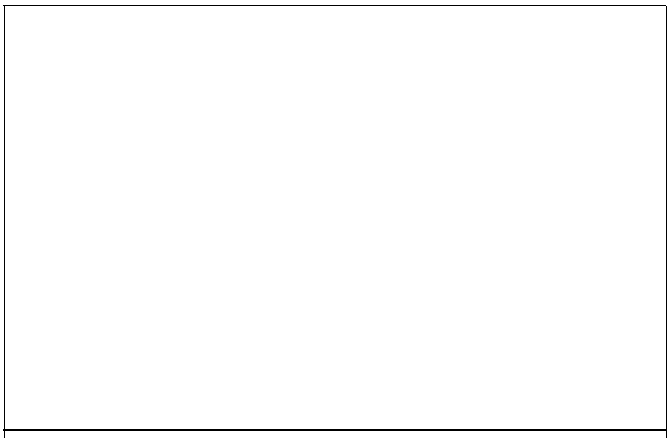


Figure 1-10: Voltage amplifier with connected source and load resistances.

Applying the voltage divider rule at the input and output of the circuit gives

$$\frac{v_{out}}{v_s} = \left(\frac{v_{in}}{v_s}\right) \cdot A_v \cdot \left(\frac{v_{out}}{v_x}\right) = \left(\frac{R_{in}}{R_{in} + R_s}\right) \cdot A_v \cdot \left(\frac{R_L}{R_L + R_{out}}\right) \tag{1.1}$$

As we can see from this expression, the overall voltage gain is maximized when the amplifier has a large input resistance (relative to R_s) and a small output resistance (relative to R_L). For the ideal case of infinite input resistance and zero output resistance, v_{out}/v_s becomes equal to A_v.

For the sake of compact notation, we will often want to use a symbol for the overall circuit gain. The notation used in this module uses primed variables to distinguish between the gain of the controlled source and the gain of the overall amplifier circuit. For example, for the above-discussed voltage amplifier we define $A'_v = v_{out}/v_s$. This notation is meant to emphasize the connection between the two symbols. A'_v is usually smaller than A_v, but can approach A_v for ideal source and load configurations.

Example 1-1: Transfer Function of a Transconductance Amplifier.

For the transconductance amplifier circuit in Figure Ex1-1, calculate the overall transconductance $G'_m = i_{out}/v_s$.

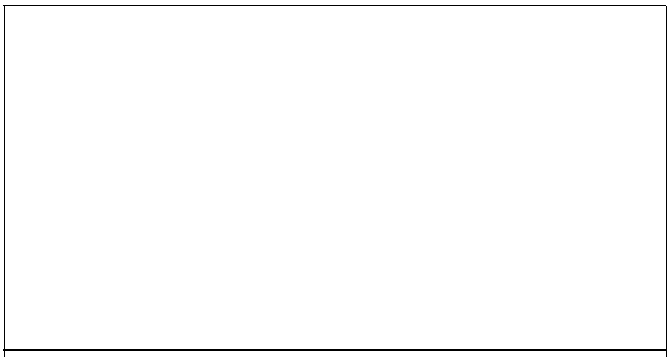


Figure Ex1-1

SOLUTION

Applying the voltage divider rule at the input and the current divider rule at the output yields the following result

$$G'_m = \frac{i_{out}}{v_s} = \left(\frac{v_{in}}{v_s}\right) \cdot G_m \cdot \left(\frac{i_{out}}{i_x}\right) = \left(\frac{R_{in}}{R_{in} + R_s}\right) \cdot G_m \cdot \left(\frac{R_L}{R_L + R_{out}}\right)$$

Thus, the overall transconductance gain is maximized when the amplifier has a large input resistance (relative to R_s) and a large output resistance (relative to R_L). For the ideal case of infinite input and output resistances, G'_m becomes equal to G_m.

As a final remark for this sub-section, it is important to recognize that all of the models in Figure 1-9 can be used interchangeability to describe the exact same electrical behavior (see Problem P1-1). For instance, a voltage amplifier model can be converted into a transconductance amplifier model by applying a Thevénin to Norton transformation for the controlled source.

A corollary to this equivalence is that we can for example use a transconductance amplifier model to describe a voltage amplifier circuit. This is illustrated through the circuit of Figure 1-11, which is electrically equivalent to that of Figure 1-10 (see Problem P1-2). Note that the output is taken as the voltage across the output port instead of the output current; this indicates that the circuit is viewed as a voltage amplifier. Just as in the original circuit of Figure 1-10, we require a large input resistance and small output resistance for this circuit to maximize the overall voltage gain.

The choice of amplifier model depends on several factors. At first glance, it seems natural to model each amplifier type

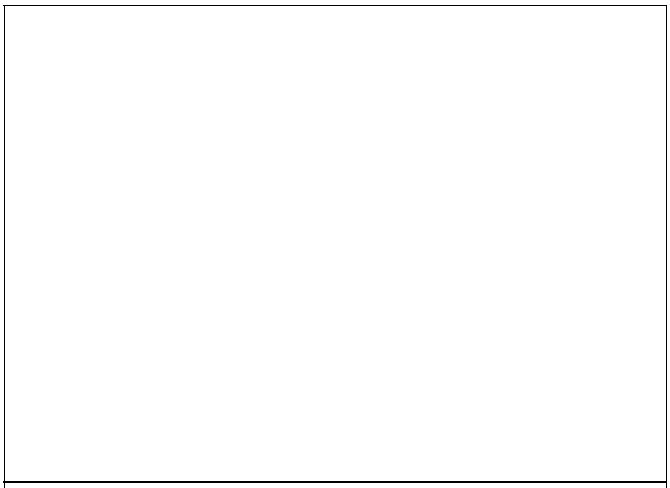


Figure 1-11: Voltage amplifier with an underlying transconductance amplifier model.

using its "native" model that directly corresponds to the intended function. For example, we could always describe a voltage amplifier using the corresponding voltage amplifier model that contains a voltage controlled voltage source. However, as we shall see throughout this module, it is sometimes more convenient to align the amplifier model with the physical amplification mechanism or a structural feature of the underlying transistor circuit. For instance, the common-source voltage amplifier discussed in Chapter 2 naturally invokes a transconductance-based model due to the physical model of the employed transistor.

1-3-2 Unilateral versus Bilateral Two-Ports

All of the two-port models shown in Figure 1-9 are called **unilateral**, because they can only propagate a signal from the input port to the output port and not the other way around. For instance, injecting a current into the output port of the current amplifier of Figure 1-9(b) will not induce a current at the input port. Unfortunately, many practical transistor circuits are not unilateral, and exhibit **bilateral** behavior when analyzed in detail, and especially at high frequencies.

An example of a bilateral current amplifier is shown Figure 1-12(a). Note that in this circuit, resistor R_2 couples the input and output networks and it can therefore transfer currents in both directions. Consequently, the unilateral model of Figure 1-9(b) cannot perfectly represent this circuit. When it is desired to capture the bilateral behavior, the two-port model in Figure 1-12(b) could be employed in principle. Here, the controlled source A_{ir} models the reverse current transfer from the output back to the input. Alternatively, one could employ other bilateral and more general two-port models based on admittance parameters (Y), impedance parameters (Z), and hybrid or inverse-hybrid parameters (H or G); see advanced circuit design texts such as Reference 2.These models are particularly useful when reverse transmission (i.e., feedback from the output to the input) is incorporated in the circuit as part of the intended design.

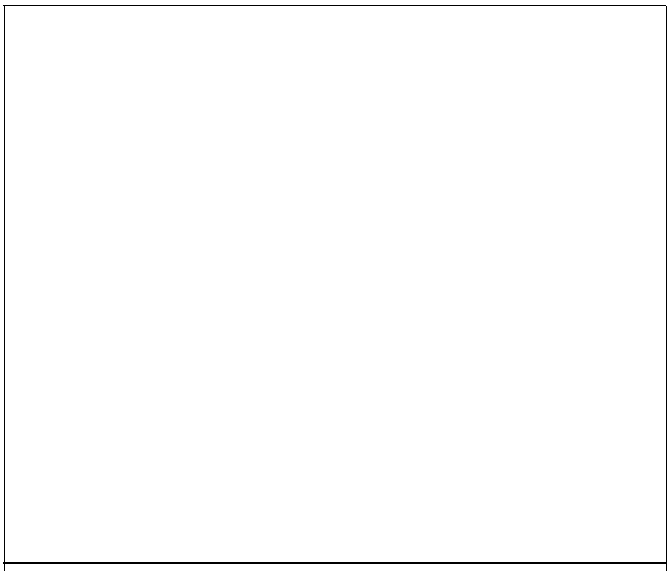


Figure 1-12: Example of a bilateral current amplifier. (b) Bilateral current amplifier two-port model.

There are two reasons why we will work exclusively with unilateral two-port approximations in this module. First, the circuits considered are designed primarily to implement forward gain rather than reverse gain; feedback circuits are not treated in this module. For example, referring to the model of Figure 1-12(b), the reverse gain A_{ir} will be negligibly small in any current amplifier circuit that we will consider. Second, a clear drawback of working with bilateral two-port models would be a significant increase in analysis complexity. As we have seen in Example 1-1, the overall transfer function of a unilateral two-port circuit can be written by applying simple voltage and current divider rules. This also extends to cascade connections of multiple two-ports. For example, the overall transfer function of the circuit in Figure 1-7 is easily written by inspection, without requiring extensive algebra. With reverse transmission included, the transfer function analysis will generally require solving a linear system of equations. In light of the fact that we do not intend to design circuits in this module that have significant reverse transmission, this increase in complexity is not welcome, and would also hinder us from developing intuition from inspection-driven analysis.

1-3-3 Construction of Unilateral Two-Port Models

We will now describe the general procedures to calculate the controlled sources, as well as the input and output resistances, for the unilateral two-port models of Figure 1-9. The approach is based on applying test voltages and currents to find the desired model parameters.

The most important parameter of any amplifier circuit is its gain. To identify the gain parameters for the models of Figure 1-9(a)-(d), we apply the tests shown in Figure 1-13(a)-(d), respectively.

- To calculate the gain term A_v of a voltage amplifier model, we apply a test voltage at the input with zero source resistance and measure the open-circuit output voltage. $A_v = v_{oc}/v_t$ is therefore also called the **open-circuit voltage gain**.
- To calculate the gain term A_i of a current amplifier model, we apply a test current at the input with infinite source resistance and measure the short-circuit output current. $A_i = i_{sc}/i_t$ is therefore also called the **short-circuit current gain**.
- To calculate the gain term G_m of a transconductance amplifier model, we apply a test voltage at the input with zero source resistance and measure the short-circuit output current to find $G_m = i_{sc}/v_t$.
- To calculate the gain term R_m of a transresistance amplifier model, we apply a test current at the input with infinite source resistance and measure the open-circuit output voltage to find $R_m = v_{oc}/i_t$.

The rationale behind these tests can be understand by considering for example the case of the voltage amplifier model of Figure 1-9(a). When driven with an ideal voltage source, the effect of any resistance at the input port is eliminated, and the controlled source is directly stimulated by the applied test source (without any voltage division). Likewise, by measuring the resulting output voltage open-circuited, any resistance in series with the controlled source has no effect and the measurement therefore accurately extracts the parameter A_v. Similar explanations apply to the test cases for the remaining amplifier models.

The test setup for extracting the input and output resistances for all amplifier models is shown in Figure 1-13(e), and (f), respectively.

- To calculate the **input resistance** R_{in} we apply a test voltage and measure the current coming from the test source, or apply a test current and measure the voltage across the test source. In this test, the load resistance (R_L) must be connected to the output port as shown in Figure 1-13(e).
- To calculate the **output resistance** R_{out}, we apply either a test voltage or a test current source at the output port and measure the respective current or voltage from the source. Here, the input source must be set equal to zero. this means that input voltage sources are shorted and input current sources are open-circuited. This means that only the source resistance (R_S) is left across the input terminals as shown in Figure 1-13(f).

The above procedures extract the input and output resistances perfectly and without any approximations, even if the circuit is bilateral. As we shall see through the examples below, R_{in} and R_{out} do not depend on R_L and R_S, respectively, in a perfectly unilateral amplifier. However, this is not the case in a bilateral amplifier, and therefore the general procedure includes R_L and R_S in the test setup.

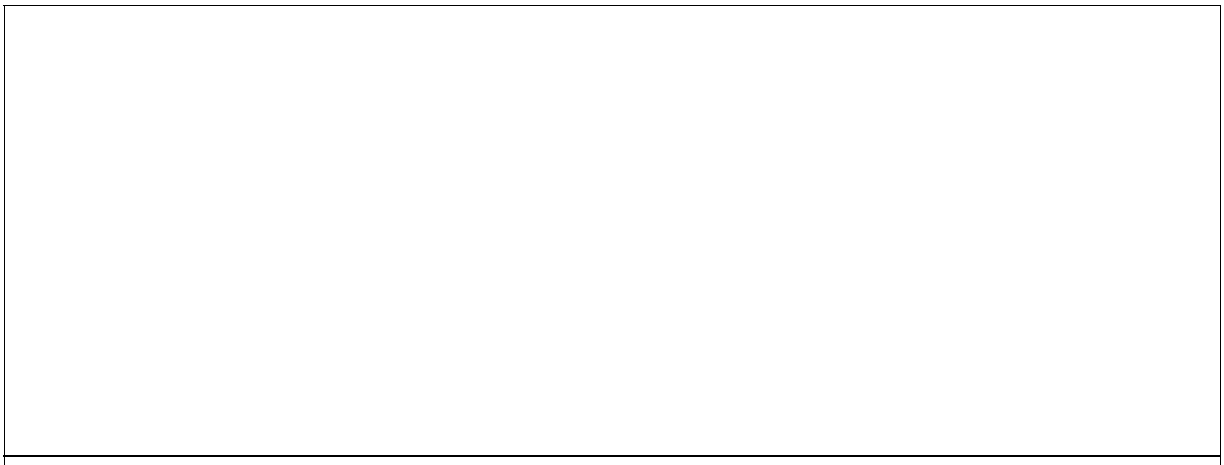


Figure 1-13: Method to calculate two-port amplifier model parameters: (a) voltage gain A_v, (b) current gain A_i, (c) transconductance G_m, (d) transresistance R_m, (e) input resistance R_{in}, and (f) output resistance R_{out}.

In summary, the above procedures for measuring unilateral two-port model parameters aim at finding the best possible unilateral representation of an arbitrary amplifier circuit, which itself may or may not be unilateral. The obtained models are approximate when the amplifier is bilateral, since they do not include a controlled source that captures reverse transmission from the output back to the input. In most cases considered in this module, the reverse transmission term is negligible. Exceptions will be highlighted and treated as appropriate.

Example 1-2: Two-Port Model Calculations for a Unilateral Amplifier.

For the transconductance amplifier in Figure Ex1-2, calculate the following two-port model parameters: the transconductance G_m, the input resistance R_{in}, and the output resistance R_{out}. Also, compute the overall transfer function $G'_m = i_{out}/v_s$.

Figure Ex1-2

SOLUTION

To find the transconductance, we short the output port and apply an ideal test voltage source (v_t) at the input. From this circuit, we see that

$$G_m = \frac{i_{sc}}{v_t} = \frac{g_m v_x \cdot \dfrac{R_3}{R_3 + R_4}}{v_t} = \frac{g_m v_t \cdot \dfrac{R_2}{R_1 + R_2} \cdot \dfrac{R_3}{R_3 + R_4}}{v_t}$$

$$= g_m \cdot \frac{R_2}{R_1 + R_2} \cdot \frac{R_3}{R_3 + R_4}$$

Next, to find R_{in}, we apply a test voltage at the input and connect the load resistance R_L at the output. From this circuit, we find that the input resistance is simply the series connection of R_1 and R_2, i.e., $R_{in} = R_1 + R_2$. Note that the output network does not influence this result.

Finally, to find R_{out}, we apply a test voltage at the output and connect the source resistance R_S across the input port (the source v_s is replaced by a short). In the resulting circuit, v_x must be zero because no current is flowing in the input network. Thus, the controlled source carries no current and we conclude that $R_{out} = R_3 + R_4$.

In order to compute the transfer function of the complete circuit, we can re-use the result obtained in Example 1-1.

$$G'_m = \frac{i_{out}}{v_s} = \left(\frac{R_{in}}{R_{in} + R_S}\right) G_m \left(\frac{R_{out}}{R_L + R_{out}}\right)$$

Substituting G_m, R_{in} and R_{out} from the above calculation yields the final result.

$$G'_m = \frac{i_{out}}{v_s} = \frac{g_m R_2 R_3}{(R_1 + R_2 + R_S)(R_L + R_3 + R_4)}$$

In the preceding example, we have seen that the source and load resistances have no effect on the extracted two-port parameters. In the following example, we will investigate a bilateral circuit to show that in general, the input and output resistances depend on R_S and R_L, which must therefore always be included in the general two-port modeling calculations.

Example 1-3: Two-Port Model Calculations for a Bilateral Amplifier.

For the current amplifier in Figure 1-12(a), calculate the fol-

lowing unilateral two-port model parameters: the current gain A_i, the input resistance R_{in}, and the output resistance R_{out}. Also, compute the overall transfer function i_{out}/v_s using the obtained unilateral two-port model. Compare the result to a direct KCL-based analysis of the transfer function. Assume that the circuit is driven by a current source with resistance R_S and loaded by a resistance R_L. For algebraic simplicity, assume $R_1 = 1/g_m$ (this case also corresponds to the common-gate amplifier circuit covered in Chapter 4).

SOLUTION

To find the current gain A_i, we short the output port and apply an ideal test current source (i_t) at the input From this circuit, we see that

$$v_{in} = \left(g_m + \frac{1}{R_2}\right)^{-1} \cdot i_t$$

and

$$i_{sc} = -\left(g_m + \frac{1}{R_2}\right) \cdot v_{in}$$

Thus, $A_i = i_{sc}/i_t = -1$.

Next, to find R_{in}, we apply a test voltage at the input and connect the load resistance R_L at the output. From this circuit, we note that the input resistance is not easily identified by inspection. Hence we write KCL for the two nodes of the circuit (v_t and v_{out}).

$$0 = -i_t + g_m v_t + \frac{v_t - v_{out}}{R_2}$$

$$0 = -g_m v_t + \frac{v_{out}}{R_L} + \frac{v_{out} - v_t}{R_2}$$

Solving this system of equations yields

$$R_{in} = \frac{v_t}{i_t} = \frac{R_2 + R_L}{1 + g_m R_2} = \frac{1 + \frac{R_L}{R_2}}{g_m + \frac{1}{R_2}}$$

Note from this result that R_{in} depends on R_L, as mentioned previously; this dependency stems from the bilateral structure of the circuit. Also note that R_{in} approaches $1/g_m$ when R_2 is large compared to R_L and $1/g_m$. We will re-visit this important point in Chapter 4, in the context of a common-gate amplifier circuit.

Now, to find R_{out}, we apply a test voltage at the output and connect the source resistance R_S across the input port. Again, we must write KCL at the two circuit nodes and solve the resulting system of equations. This yields

$$R_{out} = \frac{v_t}{i_t} = R_2 + R_S + g_m R_2 R_S$$

Again, note that R_{out} is a function of R_S; this is the case for any bilateral circuit.

Finally, to compute the transfer function based on the obtained unilateral model, we consider the circuit shown below. By inspection, we see that

$$A'_i = \frac{i_{out}}{i_s} = \left(\frac{R_S}{R_{in} + R_S}\right) A_i \left(\frac{R_{out}}{R_L + R_{out}}\right)$$

Substituting A_i, R_{in} and R_{out} from the above calculation into this expression yields

$$A'_{i,\,Two-Port} = \frac{i_{out}}{i_s} = \frac{R_S(1 + g_m R_2)(R_2 + R_S + g_m R_2 R_S)}{(R_L + R_2 + R_S + g_m R_2 R_S)^2}$$

We now wish to compare this result to the accurate transfer function of the circuit, obtained by direct calculation and without approximating the circuit as a unilateral two-port. For this purpose, we consider the full circuit shown below and write KCL for its two nodes.

$$0 = -i_s + g_m v_{in} + \frac{v_{in}}{R_S} + \frac{v_{in} - v_{out}}{R_2}$$

.

$$0 = -g_m v_{in} + \frac{v_{out}}{R_L} + \frac{v_{out} - v_{in}}{R_2}$$

Solving this system of equations for v_{out} and substituting $i_{out} = -v_{out}/R_L$ yields

$$A'_{i,\,Exact} = \frac{i_{out}}{i_s} = \frac{R_S(1 + g_m R_2)}{R_L + R_2 + R_S + g_m R_2 R_S}$$

The discrepancy factor between the two results is given by

$$\frac{A'_{i,\,Two-Port}}{A'_{i,\,Exact}} = \frac{R_2 + R_S + g_m R_2 R_S}{R_L + R_2 + R_S + g_m R_2 R_S} = \frac{1 + R_S\left(\frac{1}{R_2} + g_m\right)}{1 + \frac{R_L}{R_2} + R_S\left(\frac{1}{R_2} + g_m\right)}$$

From this result, we see that the discrepancy factor approaches unity (no error) when R_2 is much larger than R_L, a condition that is often satisfied in practice (the ideal load for a current amplifier is a short circuit). In this case, the unilateral two-port model will accurately capture the behavior of the circuit.

The outcome of the above example captures the main spirit in which we justify relying on unilateral two-port models in this module. Even though the considered amplifier is strictly speaking bilateral, a unilateral model describes its behavior to within the desired engineering accuracy, provided that reasonable boundary conditions hold.

1-4 Integrated Circuit Design versus Printed Circuit Board Design

In the design of analog circuits, the underlying technology has a significant impact on the choice of architecture because it tends to restrict the availability and specification range of the underlying active and passive components. For instance, a designer working with discrete components on a printed circuit board may be subjected to the following constraints:

- Limit the component count below 100 elements to achieve a small board area.
- Resistors can be chosen in the range of 1Ω–10 MΩ.
- Capacitors can be chosen in the range of 1 pF–10,000 μF.
- The resistor and capacitor values match to within 1–10%.
- The available (discrete) bipolar junction transistors match to within 20% in their critical parameters.

In contrast, the designer of a CMOS system-on-chip may face the constraints summarized below:

- Avoid using resistors, use as many MOSFET transistors as needed (within reasonable limits, on the order of hundreds to several thousands) to realize the best possible circuit implementation.
- Capacitors can be chosen in the range of 10 fF–100 pF.
- The critical parameters in the MOSFET transistors can be made to match to within 1%, but vary by more than 30% for different fabrication runs.
- Capacitors of similar size can match to within 0.1%, but vary by more than 10% for different fabrication runs.

As a consequence of the vastly different constraints that apply to the design of analog circuits in CMOS technology, the resulting practical and preferred circuit architectures differ substantially from the ones that would be used in a printed circuit board design. For example, a discrete voltage amplifier may utilize large AC coupling capacitors to simplify and decouple the biasing of the individual gain stages (see example in Figure 1-14). In contrast, it is typically not possible to use AC coupling techniques (except for very high-frequency designs) in integrated circuits, primarily due to the restriction on maximum capacitor size.

The material covered in this module is primarily concerned with analog integrated circuit design. While this choice does not affect many of the key principles used in the analysis and the discussed circuits, it does affect the architectural choices made in arriving at a practical design. For instance, large AC coupling capacitors are not used throughout the discussion. Also, where appropriate, we will invoke certain assumptions about the typical matching of component parameters in CMOS to eliminate impractical design choices.

1-5 Prerequisites and Advanced Material

The reader of this module is expected to be familiar with the basis concepts of linear circuit analysis (see Reference 3), including

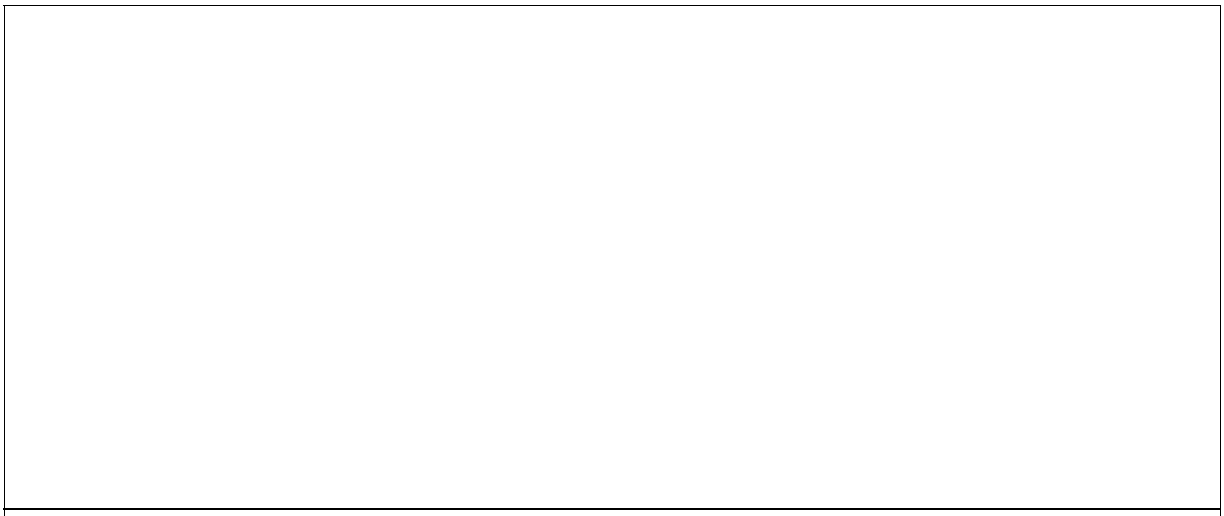


Figure 1-14: Example of a discrete amplifier circuit using bipolar junction transistors (BJTs).

- Passive components (resistors, capacitors)
- Kirchhoff's voltage and current laws (KVL and KCL)
- Independent and dependent voltage and current sources; Thevénin and Norton representation of controlled sources
- Two-port representation of circuits; calculation of port resistances and frequency dependent impedances
- Manipulation of complex variables and numbers
- Phasor analysis and Laplace domain representation of passive circuit elements
- Bode plots

The derivations of device models in this module assume familiarity with basic solid-state physics and electrostatics as treated in introductory texts on solid-state device physics (see Reference 4).

A few sections of this module are marked with an asterisk (*) to indicate advanced material that may in some cases go beyond the learning goals of an introductory course. These sections can be skipped at the instructor's discretion without affecting the overall flow and context.

1-6 Notation

This module follows the notation for signal variables as standardized by the IEEE. Total signals are composed of the sum of DC quantities and small signals. For example, a total input voltage v_{IN} is the sum of a DC input voltage V_{IN} and a small-signal voltage v_{in}. The notation is summarized below.

- Total quantity has a lower case variable name and upper case subscript
- DC quantity has an upper case variable name and upper case subscript
- Small-signal quantity has a lower case variable name and lower case subscript

Summary

This chapter offered a brief motivation for the topics covered in this module, which focuses on the analysis and design of elementary amplifier stages in CMOS technology. These elementary stages can be viewed as the "atoms" of analog circuit design and a thorough understanding of the blocks is a necessary prerequisite for the design of advanced analog circuits design, as for instance in the context of large systems-on-chip. At all levels of circuit design, complexity is managed using hierarchical abstraction and model simplification using proper engineering approximations. The unilateral port-models reviewed in Section 1-3 and used throughout this module, are an example of such abstractions.

References

1. R. B. Staszewski, et al., "A 24 mm^2 Quad-Band Single-Chip GSM Radio with Transmitter Calibration in 90 nm Digital CMOS," *IEEE Solid-State Circuits Conference*, Digest of Technical Papers, pp.208-209, Feb. 2008.

2. P. R. Gray, P. J. Hurst, S. H. Lewis, and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 5th Edition, Wiley, 2008.
3. F. T. Ulaby and M. M. Maharbiz, *Circuits*, NTS Press, 2009.
4. S. Bhave, R. T. Howe and C. G. Sodini, *TBD*, NTS Press 2011.

Problems

P1.1 Given the amplifier circuit in Figure P1-1

(a) Find the input and output resistance.

(b) Construct an equivalent circuit using a voltage amplifier two-port model and determine all model parameters symbolically.

(c) Repeat part (b) for a current amplifier model.

(d) Repeat part (b) for a transconductance amplifier model.

(e) Repeat part (b) for a transresistance amplifier model.

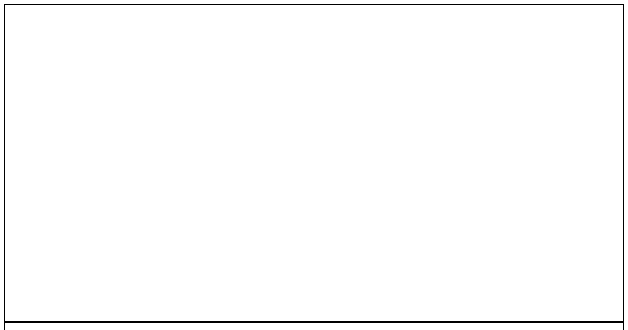


Figure P1-1:

P1.2 Convince yourself that the circuits of Figure 1-10 and Figure 1-11 are equivalent by showing symbolically that both circuits have the same overall voltage gain $A'_v = v_{out}/v_s$.

P1.3 You are given an input voltage source with a source resistance, R_S.

(a) Use the unilateral voltage amplifier two-port model found in P1.1 to find the overall voltage gain when the amplifier is driving a load resistor R_L.

(b) Specify whether the resistances r_1, r_i, r_o, r_2 in the small signal model should be increased, decreased or remain the same to improve the overall voltage gain.

P1.4 You are given an input current source with a source resistance, R_S.

(a) Use the unilateral current amplifier two-port model found in P1.1 to find the overall current gain when the amplifier is driving a load resistor R_L.

(b) Specify whether the resistances r_1, r_i, r_o, r_2 in the small-signal model should be increased, decreased or remain the same to improve the overall current gain.

P1.5 Given the small-signal model in Figure P1-4 for an amplifier circuit

(a) Find the input and output resistance.

(b) Construct a two-port model for a unilateral voltage amplifier.

(c) Construct a two-port model for a unilateral current amplifier.

(d) Construct a two-port model for a unilateral transconductance amplifier.

(e) Construct a two-port model for a unilateral transresistance amplifier.

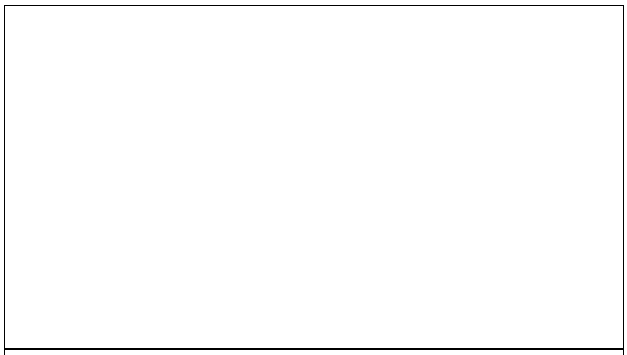


Figure P1-5:

P1.6 Consider the two-port model of a voltage amplifier as shown in Figure 1-9(a) with the following parameters: $A_v = 10$, $R_{in} = 5$ kΩ, and $R_{out} = 100$ Ω.

(a) Draw the two-port model for a transresistance amplifier by conversion from the voltage amplifier model.

(b) Draw the two-port model for a transconductance amplifier by conversion from the voltage amplifier model

(c) Draw the two-port model for a current amplifier by conversion from the voltage amplifier model.

P1.7 Derive an expression for the transresistance v_{out}/i_{in} for the circuit of Figure 1-7 using the following parameters: $A_{i1} = 1$, $G_{m2} = 10$ mS, $A_{v3} = 0.8$, $R_{in1} = 50$ Ω, $R_{out1} = 500$ Ω, $R_{out2} = 1$ kΩ, and $R_{out3} = 100$ Ω. Using this result, lump the entire circuit into a single transresistance amplifier as shown in Figure 1-9(d). Draw the resulting model including R_{in} and R_{out}.

P1.8 Consider the amplifier circuit of Figure 1-12(a) with $R_1 = 1/g_m = 1$ kΩ and $R_2 = 100$ kΩ. Compute all component values for the bilateral two-port current amplifier model of Figure 1-12(b). Note that A_{if}, R_{in}, and R_{out} can be described as explained in Section 1-3-3. Similar to A_{if}, A_{ir} is found by short-circuiting the input port and by injecting a test current into the output port. Compare the relative magnitude of A_{if} and A_{ir}.

2

CHAPTER

Transfer Characteristic of the Common-Source Voltage Amplifier

In this chapter we will review the first-order characteristics of the Metal-Oxide-Semiconductor Field-Effect Transistor (MOSFET) and show how this device can be used to build a common-source (CS) voltage amplifier. Using the first-order I-V laws of the MOSFET, we begin by deriving the transfer characteristic of the CS stage in terms of its input and output voltages. As we shall see in this process, it is useful to apply a linear approximation to the characteristic for small incremental voltage changes around a operating point. This simplification, called small-signal approximation, is a generally useful and broadly applicable tool that lets us analyze complex transistor stages using simple and intuitive linear methods.

The overall goal of this chapter is to develop a feel for the first-order behavior and modeling of a MOSFET using the CS stage as a first application. In essence, the presentation in this chapter (and the remainder of this module) follows a "just-in-time" modeling approach; i.e., the transistor model is gradually modified and augmented to answer questions that arise from a circuit design perspective. Rather than deriving a complete model in one step, we begin with the most basic properties and increase complexity only upon demand, where needed to gain further insight and accuracy. This approach was chosen primarily because analog circuit design relies on a solid understanding of device models and their limitations. A combined treatment of modeling and its relation to circuit behavior is therefore useful in developing intuition about the relevant issues in practical design problems.

Chapter Objectives

- Review the MOSFET device structure and basic operation as described by the square-law model.
- Introduce large- and small-signal analysis techniques using the common-source voltage amplifier as a motivating example.
- Derive a small-signal model for the MOSFET device, consisting of a transconductance and output resistance element.
- Provide a feel for potential inaccuracies and range limitations of simple modeling expressions.

2-1 First-Order MOSFET Model

The device-level derivations of this section assume familiarity with basic solid-state physics and electrostatics. For a ground-up treatment from first principles, the reader is referred to introductory solid-state device material (see Reference 1).

2-1-1 Derivation of I-V Characteristics

The basic structure of an **enhancement mode n-channel** MOSFET is shown in Figure 2-1(a). It consists of a lightly doped p-substrate (**bulk**), two heavily doped n-type regions (**source** and **drain**) and a conductive gate electrode that is iso-

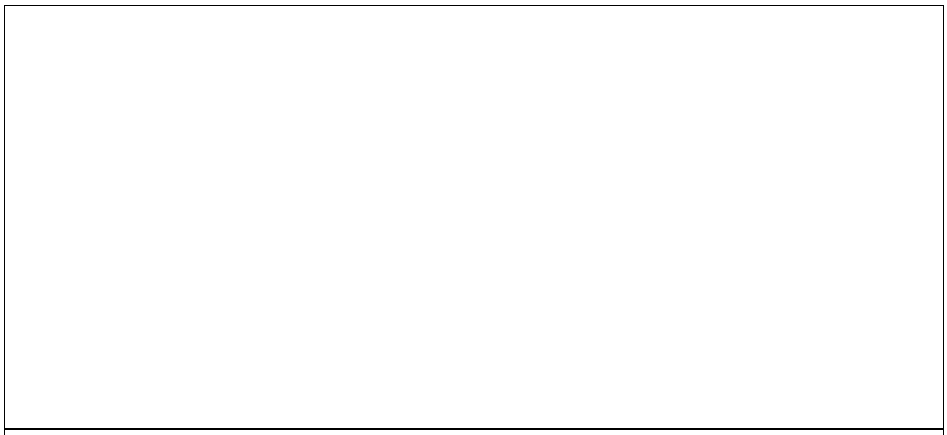


Figure 2-1: (a) Cross-section of an n-channel MOSFET. (b) Schematic symbol.

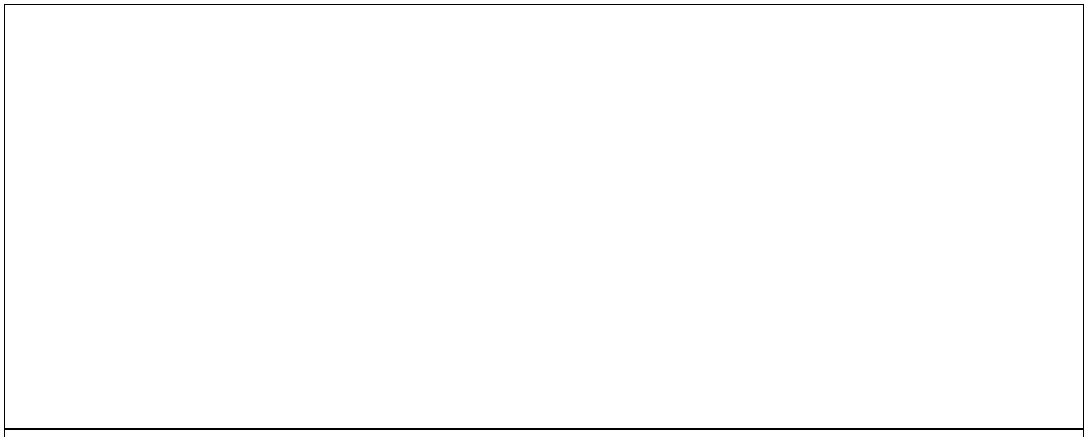


Figure 2-2: (a) n-channel MOSFET with $V_{GS} = 0$ and (b) $V_{GS} > V_{Tn}$.

lated from the substrate using a thin silicon dioxide layer of thickness t_{ox}. Other important geometry parameters of this device include the **channel length** L (distance between the source and drain) and the **channel width** W. As we shall see, the name "n-channel" stems from the fact that this device conducts current by forming an n-type layer underneath the gate. A **p-channel** device can be constructed similarly using an n-type bulk and p-type source/drain regions. The differentiating details between n- and p-channel devices are summarized in Section 2-1-2. For the time being, we will use the n-channel device to discuss the basic principles.

In order to study the electrical behavior of a MOSFET, it is useful to define a schematic symbol and conventions for electrical variables as shown in Figure 2-1(b). The variables V_{GS}, V_{DS}, and V_{BS} describe the voltages between the respective terminals using the commonly used ordered subscript convention $V_{XY} = V_X - V_Y$. The current flowing into the drain node is labeled I_D.

It is important to note that the MOSFET device considered here is perfectly symmetric; i.e., the drain and source terminal labels can be interchanged. It is a common convention to assign the source to the lower potential of these two terminals, since this terminal is the source of electrons that enable the flow of current. We will see later that this convention, together with the arrow that marks the source (and the direction of current flow), provides useful intuition when reading a larger circuit schematic.

We now begin our analysis of the MOSFET device by considering the condition shown in Figure 2-2(a), where the bulk and source are connected to a reference potential (GND), $V_{GS} = 0$ V and $V_{DS} \geq 0$ V. Under this condition, the drain and source terminals are isolated by two reverse-biased pn-junctions and their **depletion regions**, which prevents any significant flow of current. Applying a positive voltage at the drain ($V_{DS} > 0$) increases the reverse-bias at the drain-bulk junction and will

only increase the width of the depletion region at the drain, while $I_D = 0$ is still maintained (to first order).

Consider now $V_{GS} > 0$, as shown in Figure 2-2(b). This positive voltage at the gate attracts electrons from the source. With increasing V_{GS}, a larger amount of electrons is supplied by the source, and ultimately, a so-called **inversion layer** forms underneath the gate. The voltage V_{GS} at which a significant number of mobile electrons underneath the gate become available is called the **threshold voltage** of the transistor, or V_T. In order to differentiate the threshold voltages and other device parameters of n- and p-channel devices, we will utilize the subscripts *n* and *p* throughout this module. E.g., we denote the threshold voltage for n-channels and p-channels as V_{Tn} and V_{Tp}, respectively.

With the inversion layer under the gate, the drain and source regions are now "connected" through a conductive path and any voltage between these terminals ($V_{DS} > 0$) will result in a flow of drain current. How can we calculate this current? In order to answer this question, the following approximations are useful

1. The current primarily depends on the number of mobile electrons in the channel times their velocity.
2. The number of mobile electrons in the channel is set by the vertical electric field from the gate to the conductive channel (**gradual channel approximation**).
3. The threshold voltage is constant along the channel; this assumption neglects the so-called body effect.
4. The velocity of the electrons traveling from the source to the drain is proportional to the lateral electric field in the channel.

Figure 2-2(b) establishes relevant variables for further analysis. The auxiliary variable y ranges from 0 to L and is used to express electrical quantities as a function of the distance from the source. The inversion layer charge density (per unit area) and voltage at position y in the channel are denoted as $Q_n(y)$ and $V(y)$, respectively. With these conventions in place, we can translate the above-listed assumptions into the following equations

$$I_D(y) = W \cdot Q_n(y) \cdot v(y) \tag{2.1}$$

$$Q_n(y) = -C_{ox}(V_{GS} - V(y) - V_{Tn}) \tag{2.2}$$

$$v(y) = -\mu_n \cdot E(y) \tag{2.3}$$

In these expressions, v is the velocity of the carriers, C_{ox} is the **gate capacitance** per unit area (between the gate electrode and the conductive channel). The term μ_n is called **mobility**, and it relates the drift velocity of the carriers to the local electric field.

As indicated in Eq. (2.2), the mobile charge density at coordinate y depends on the local potential, since the voltage across the oxide is given by $V_{GS} - V(y)$. An inversion layer is present at any location under the gate where this voltage difference is larger than the threshold (V_{Tn}). Assuming that the inversion layer extends from source to drain [as drawn in Figure 2-2(b)], we have $V(L) = V_{DS}$ and $Q_n(L) = -C_{ox}(V_{GS} - V_{Tn} - V_{DS})$. This implies that V_{DS} cannot exceed $V_{GS} - V_{Tn}$ for an inversion layer that extends across the entire channel. For the time being, we will solve for the drain current for this condition and later extend the obtained result for the case of $V_{DS} \geq V_{GS} - V_{Tn}$.

Now, by combining Eq. (2.1) through Eq. (2.3) and noting that the electric field is given by $E(y) = dV(y) / dy$, we can write

$$I_D(y) = \mu_n C_{ox} W (V_{GS} - V(y) - V_{Tn}) \frac{dV(y)}{dy} \tag{2.4}$$

This result describes the current density profile along the channel. The terminal current, I_D, can be found by separating the variables and integrating along the direction of y

$$\int_0^L I_D(y)dy = \mu_n C_{ox} W \int_0^{V_{DS}} (V_{GS} - V(y) - V_{Tn})dV \tag{2.5}$$

which yields a closed-form solution for the drain current

$$\boxed{I_D = \mu_n C_{ox} \frac{W}{L}\left(V_{GS} - V_{Tn} - \frac{V_{DS}}{2}\right)V_{DS}} \tag{2.6}$$

Note that this expression is valid for $V_{DS} < V_{GS} - V_{Tn}$, as assumed above.

In order to extend the obtained result for $V_{DS} \geq V_{GS} - V_{Tn}$, we continue by inspecting the shape of the inversion layer for various V_{DS} (see Figure 2-3). For $V_{DS} = 0$ V [case (a)], no current flows and $V(y) = 0$ for all y. Provided that $V_{GS} > V_{Tn}$, a uniform inversion layer exists underneath the gate. For small $V_{DS} > 0$, a current flows in the inversion layer, which causes increasing $V(y)$ and decreasing inversion layer charge along the channel. As V_{DS} approaches $V_{GS} - V_{Tn}$, $Q_n(L)$ approaches zero with a point of diminishing charge at the drain. This effect is called **pinch-off**.

What happens when we increase V_{DS} beyond the point of pinch-off? Further analysis based on solving the two-dimensional **Poisson Equation** at the drain predicts that the pinch-off point will move from L to $L - \Delta L$, where ΔL is small relative to L. Even though no inversion layer exists in the region from $L - \Delta L$ to L, the device still conducts current. The charges arriving at $y = L - \Delta L$ are being swept to the drain by the electric field present in the depletion region of the surrounding pn junction.

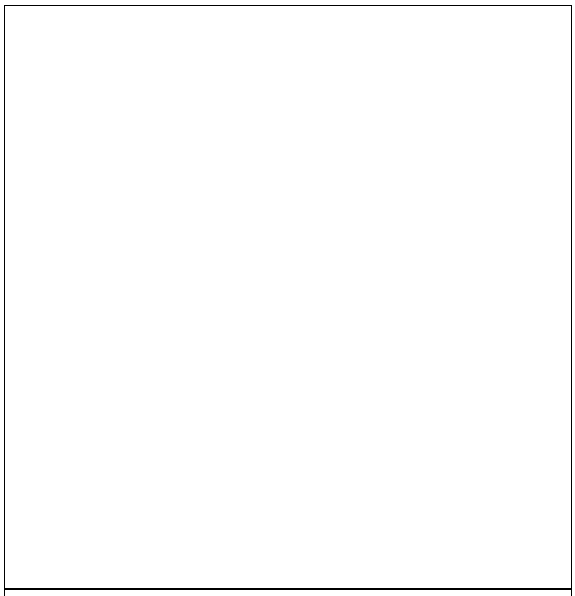


Figure 2-3: Channel profile for varying V_{DS}.

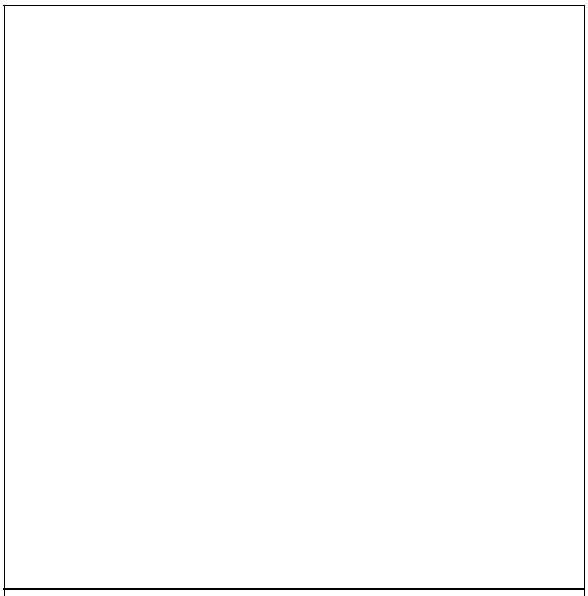


Figure 2-4: (a) n-channel I-V characteristic for a fixed value of $V_{GS} - V_{Tn} = 2$ V. (b) I-V plots with varying $V_{GS} - V_{Tn}$ (drain characteristic). Parameters: $\mu C_{ox} = 50\ \mu A/V^2$ and $W/L = 10$.

To first order, and neglecting the small change in channel length ΔL, the current becomes independent of V_{DS} and is approximately given by the current at the onset of pinch-off, i.e., at $V_{DS} = V_{GS} - V_{Tn}$. Substituting this condition into Eq. (2.6), we obtain

$$I_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{GS} - V_{Tn})^2 \tag{2.7}$$

for $V_{DS} \geq V_{GS} - V_{Tn}$.

Eq. (2.6) and Eq. (2.7) are plotted in Figure 2-4(a) as a function of V_{DS} and some fixed $V_{GS} > V_{Tn}$. The operating region for $V_{DS} < V_{GS} - V_{Tn}$ is commonly called the **triode region**. This name stems from the direct dependence of the drain current on the drain-source voltage, which is qualitatively similar to the behavior of vacuum tube "triodes". The region $V_{DS} \geq V_{GS} - V_{Tn}$ is called the **saturation region** due to the saturation in current at large V_{DS}. In this region, the device operates essentially like a current source; the current is (to first order) independent of the applied V_{DS} and $I_D = I_{Dsat}$ = constant. The quantity $V_{GS} - V_{Tn}$ is often called **gate overdrive**.

The drain-source voltage at which the drain current saturates is called V_{DSsat}. From the above first-order analysis, it is clear that $V_{DSsat} = V_{GS} - V_{Tn}$. Nonetheless, it is useful to distinguish between these two quantities because they may differ significantly when a more elaborate device model is used. V_{DSsat} is generally not exactly equal to $V_{GS} - V_{Tn}$ when second-order effects, for example related to small geometries and modern device structures, are considered.

From a circuit perspective, the device's behavior in the triode region is similar to a resistor: the current increases monotonically with increasing terminal voltage. Even though the dependence of I_D on V_{DS} is nonlinear [as seen from Eq. (2.6)], it is sometimes useful to approximate the characteristic using a linear I-V law, shown as a dashed line in Figure 2-4(a). For $V_{DS} << V_{GS} - V_{Tn}$, we can approximate Eq. (2.6) as

$$I_D \cong \mu_n C_{ox}\frac{W}{L}(V_{GS} - V_{Tn})V_{DS} \tag{2.8}$$

Under this approximation, I_D depends linearly on V_{DS} and we can define the so-called **on-resistance** of the device as

$$R_{on} = \frac{V_{DS}}{I_D} \cong \frac{1}{\mu_n C_{ox}\frac{W}{L}(V_{GS} - V_{Tn})} \tag{2.9}$$

It is interesting to interpret the dependencies in this expression using basic intuition. Increasing the aspect ratio W/L decreases R_{on} since the conductive path becomes shorter and/or wider; this is a basic property of any conductor. The on-resistance also decreases with increasing C_{ox} and $V_{GS} - V_{Tn}$; this is because the inversion charge increases with these quantities ($Q = CV$). Larger mobility (μ_n) means that the carriers travel faster for the same applied voltages (electric field). This increases the current (charge per unit of time) and therefore also results in smaller R_{on}.

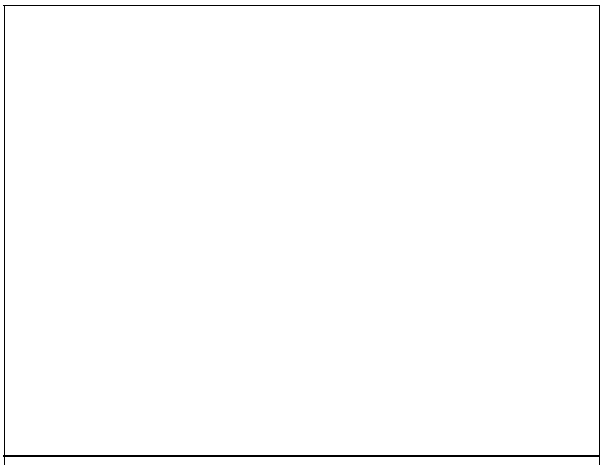


Figure 2-5: Plot of n-channel drain current as a function of V_{GS} (transfer characteristic). Parameters: V_{DS} = 5 V, μC_{ox} = 50 μA/V^2, V_{Tn} = 0.5 V, and W/L = 10.

As seen from Eq. (2.7), the magnitude of the drain current in saturation depends on the square of the gate overdrive $V_{GS} - V_{Tn}$. This is further illustrated in Figure 2-4(b), which shows I-V plots for increasing multiples of $V_{GS1} - V_{Tn}$ = 1V. Doubling and tripling the gate overdrive increases the saturation current by factors of four and nine, respectively. Note that R_{on} is reduced only by factors of two and three in these cases, respectively.

The plot in Figure 2-4(b) is often called the **drain characteristic** because the drain-source voltage (as opposed to the gate-source voltage, which is included as a parameter on the curves) is swept along the x-axis. Alternatively, the term **output characteristic** is sometimes used, primarily because V_{DS} can often be viewed as the output port voltage of the device; we will see this in the example discussed in Section 2-2.

Another commonly used characterization plot for MOSFETs is the so-called **transfer characteristic**, which shows the drain current as a function of V_{GS} for a fixed value of V_{DS}. If V_{DS} is chosen large enough such that the device operates in the saturation region for all applied V_{GS}, I_D follows from Eq. (2.7) and the plot is shaped like a parabola as drawn in Figure 2-5.

Table 2-1 summarizes the first-order MOSFET I-V relationships that were discussed in this section. This set of equations (and extended versions thereof) is often called the **square-law model** since one of its primary features is the quadratic dependence of the saturation current on $V_{GS} - V_{Tn}$. When working with this device model, it is important to remember that it predicts the behavior of real MOSFETs only with limited accuracy. This is primarily so because we have made several simplifications in the model's derivation. The most significant shortcomings that result from these assumptions can be summarized as follows:

Table 2-1: First-order MOSFET model summary.

	"ON" $V_{GS} \geq V_{Tn}$	**"OFF"** $V_{GS} < V_{Tn}$
$V_{DS} < V_{DSsat}$	"Triode region" $I_D = \mu_n C_{ox} \frac{W}{L}\left(V_{GS} - V_{Tn} - \frac{V_{DS}}{2}\right) V_{DS}$	$I_D = 0$
$V_{DS} \geq V_{DSsat}$	"Saturation region" $I_D = \frac{1}{2}\mu_n C_{ox} \frac{W}{L}(V_{GS} - V_{Tn})^2$	

1. In reality, the saturation current has a weak dependence on V_{DS}. This is primarily due to a shortening of the channel length (ΔL) with increasing V_{DS} and also due to the drain voltage dependence of the mobile charge in the channel. We will address this issue in Section 2-2.
2. For transistors built in modern technologies, several second-order effects related to small geometries and large electric fields become significant. This typically results in a saturation current law exponent that is less than two, and $V_{DSsat} < V_{GS} - V_{Tn}$. In addition, the drain current does not scale strictly proportional to $1/L$ and the threshold voltage is not constant, but a function of the drain voltage.
3. For $V_{GS} < V_{Tn}$ the device is not completely off, but carries a small current that exponentially depends on V_{GS}. This operating region is called the **sub-threshold region**.
4. For small values of $V_{GS} - V_{Tn}$, on the order of a few tens of milivolts, the region underneath the gate is only moderately inverted, and the square law model tends to predict the drain current with poor accuracy.

Despite these shortcomings, the first-order MOSFET model possesses many of the critical features needed to study the fundamentals of analog circuit design. Many of the second-order effects not featured in the basic model can be treated using advanced device physics and often result in a high-complexity model that is unsuitable for hand-calculations and intuition building.

Within the range of circuits treated in this module, we typically begin by applying the first-order model. Then, only when the circuit appears to be sensitive to second-order dependencies not covered by this model, we will look for extensions. A treatment in this fashion has the advantage that the reader can develop a feel for where and when modeling extensions and parameter accuracy are critical.

In general, the trade-off between modeling accuracy and complexity is a recurring theme at all levels of analog circuit design; the issue is not limited to the introductory material covered in this module. More accurate models can always be generated at the expense of complexity and time. An experienced analog designer will often use the simplest possible model that

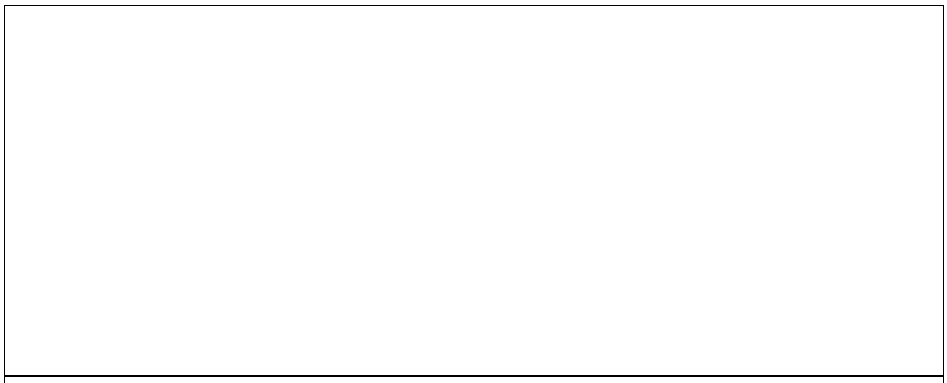


Figure 2-6: (a) Cross-section of a p-channel MOSFET. (b) Schematic symbol.

will predict the behavior of his or her circuit with sufficient (but not perfect) accuracy. This also implies that analog circuit designers must always be on the lookout for model inadequacies. We will encounter and discuss situations where either model expansions or critical insight on modeling accuracy are needed throughout this module.

2-1-2 P-Channel MOSFET

The n-channel MOSFET discussed so far conducts current through an electron inversion layer in a p-type bulk. Similarly, we can construct a **p-channel** device that operates based on forming an inversion layer of holes in an n-type bulk. The structure of such a MOSFET, which consists of p+ source and drain regions in an n-type bulk, is shown in Figure 2-6. In many process technologies, the n-type bulk region is formed by creating an n-type well (**n-well**) in the p-type substrate that is used to form n-channels. Such a technology is called an **n-well technology**. In general, a technology that offers both n-and p-channel devices is called **CMOS** technology, where CMOS stands for Complementary Metal-Oxide-Semiconductor.

The drain current equations for a p-channel MOSFET can be derived using exactly the same approach as used for the n-channel device since the basic physics is the same. For a p-channel device, the gate must be made negative with respect to the p-type source in order to form an inversion layer of holes; the threshold voltage V_{Tp} is therefore typically negative. Since holes drift across the channel from the source to the drain in the p-channel MOSFET, the drain voltage must be negative with respect to the source, and the drain current (defined as flowing into the drain terminal) is negative. Therefore, in the on-state of the transistor, V_{GS} and V_{DS} are negative quantities, and the source lies at the highest potential among the four terminals. The drain current for a p-channel in saturation, i.e., $V_{GS} < V_{Tp}$ and $V_{DS} \leq V_{GS} - V_{Tp}$, is given by

$$I_D = -\frac{1}{2}\mu_p C_{ox}\frac{W}{L}(V_{GS} - V_{Tp})^2 \tag{2.10}$$

A practical problem for the circuit designer is to keep track of the minus signs and negative quantities in the p-channel equations. A solution to this issue is to "think positive," and work with the physically intuitive positive quantities V_{SG} (instead of V_{GS}), V_{SD} (instead of V_{DS}) in all hand calculations. Following this approach, we can re-write Eq. (2.10) as

$$-I_D = \frac{1}{2}\mu_p C_{ox}\frac{W}{L}(V_{SG} + V_{Tp})^2 \tag{2.11}$$

Note that the right hand side of this equation yields a positive number. The minus sign included on the left hand side remains necessary because I_D, as defined in Figure 2-6(b), is a negative quantity.

2-1-3 Standard Technology Parameters

For use throughout this module, it is convenient to define standard MOSFET parameter values as given in Table 2-2. The chosen values are representative of a CMOS technology with a minimum channel length, or **feature size** of 1 μm. As we learn more about the behavior of MOSFETs in later sections, this list of parameters will grow and we will augment it as needed.

In the context of defining these parameters, it is important to make a clear distinction between **technology parameters** and **design parameters**. Technology parameters are typically fixed in the sense that a circuit designer cannot alter their values. For instance, the mobility in a MOSFET depends on how the transistor is made, and the underlying recipe remains unchanged and will be re-used for an extended time to manufacture a large variety and quantity of integrated circuits. In most modern CMOS technologies, the width and length of a MOSFET remain as the only parameters that the circuit

Table 2-2: Standard technology parameters for the first-order MOSFET model.

Parameter	n-channel MOSFET	p-channel MOSFET
Threshold voltage	$V_{Tn} = 0.5$ V	$V_{Tp} = -0.5$ V
Transconductance parameter	$\mu_n C_{ox} = 50\ \mu A/V^2$	$\mu_p C_{ox} = 25\ \mu A/V^2$

designer can choose (within appropriate limits) to alter the device's electrical behavior.

In determining the transistor geometries, the designer will usually work with electrical variables and parameters that describe the circuit and its functionality, for example in terms of currents and voltages. From these electrical descriptions and specifications, the widths and lengths of the transistors are then calculated, and sometimes adjusted via an iterative process. In this task, intermediate electrical parameters, as for instance the gate overdrive of a MOSFET, are also legitimately viewed as parameters that are under the control of the circuit designer.

Example 2-1: P-Channel Drain Current Calculation

A p-channel transistor is operated with the following terminal voltages relative to ground: $V_G = 2.5$ V, $V_S = V_B = 5$ V, $V_D = 1$ V. Calculate the drain current (I_D) using the standard technology parameters given in Table 2-2 and assuming $W/L = 5$.

SOLUTION

From the given terminal voltages, we find $V_{SG} = 5\text{ V} - 2.5\text{ V} = 2.5$ V and $V_{SD} = 5\text{ V} - 1\text{ V} = 4$ V. Since $V_{SG} > -V_{Tp}$, the transistor is on, and since $V_{SD} > V_{SG} + V_{Tp} = 2.5\text{ V} - 0.5\text{ V} = 2$ V, it operates in saturation. Therefore, using Eq. (2.11) we find

$$-I_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{SG} + V_{Tp})^2$$

$$-I_D = \frac{1}{2}\cdot 25\frac{\mu A}{V^2}\cdot 5\cdot(2.5V - 0.5V)^2 = 250\mu A$$

$$I_D = -250\mu A$$

2-2 Building a Common-Source Voltage Amplifier

We will now utilize our first-order understanding of MOSFETs to construct a basic voltage amplifier. We begin by noting that the drain current in all regions of operation can be controlled by varying the gate-source voltage. One way to utilize this effect to build a voltage amplifier is to apply the input such that it controls V_{GS}. An output voltage can then be generated by letting the drain current flow through a resistor, as shown in Figure 2-7(a). The top terminal of the resistor is connected to a **supply voltage**, V_{DD}. In this scheme, a larger V_{IN} causes the drain current to increase and V_{OUT} to decrease, since a larger V_{IN} makes the transistor a "better conductor" (more inversion charge), which forces the voltage at the output port closer to ground. This type of circuit is therefore categorized as an **inverting amplifier**. Furthermore, this transistor stage is called a **common-source amplifier**, since the source terminal of the MOSFET is common to the input and output ports of the circuit.

2-2-1 Voltage Transfer Characteristic

In order to derive the voltage transfer characteristic of the circuit (V_{OUT} as a function of V_{IN}) we begin by applying Kirchhoff's laws at the output node. This yields

$$V_{OUT} = V_{DD} - I_D R_D \tag{2.12}$$

The drain current I_D in this expression depends on V_{GS} and V_{DS} of the transistor, as described in Table 2-1. Given the structure of the circuit in Figure 2-7(a), we note that $V_{GS} = V_{IN}$ and $V_{DS} = V_{OUT}$. Using this information, we can construct a piecewise function that relates the input and output voltages of the circuit. For this derivation, imagine that we sweep V_{IN} from 0 V to the supply voltage, V_{DD}.

First, we note that for $V_{IN} = V_{GS} < V_{Tn}$, no current flows in the transistor; this implies $V_{OUT} = V_{DD}$. This behavior is shown in Figure 2-7(b) as a horizontal line for the input voltage range $0 \le V_{IN} < V_{Tn}$, between points A and the vertical line at V_{Tn}. As V_{IN} increases to values greater than or equal to V_{Tn}, the transistor conducts current, and V_{OUT} must be less than V_{DD}. In order to calculate how V_{OUT} changes as a function of V_{IN}, we must first determine the transistor's region of operation. As we increase V_{IN} above V_{Tn}, does the MOSFET operate in saturation or in the triode region?

To answer this question, we must determine if V_{DS} is smaller or larger than $V_{GS} - V_{Tn}$. For V_{IN} just above V_{Tn}, $V_{GS} - V_{Tn}$ is smaller than V_{DS}, which is still close to V_{DD} at the onset of current conduction. Therefore, the device must initially operate in saturation as we transition from the "OFF" state of the transistor into the region where $I_D > 0$. Under this condition, the output voltage is given by

$$V_{OUT} = V_{DD} - R_D\cdot\frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{IN} - V_{Tn})^2 \tag{2.13}$$

and the voltage transfer characteristic shows a drop that is quadratic in V_{IN} as seen in Figure 2-7(b).

As we continue to increase V_{IN}, $V_{GS} - V_{Tn}$ also increases while V_{OUT} continues to decrease. At a sufficiently large V_{IN}, V_{DS} can approach $V_{GS} - V_{Tn}$ and the condition for current saturation may no longer hold; the device then transitions into the

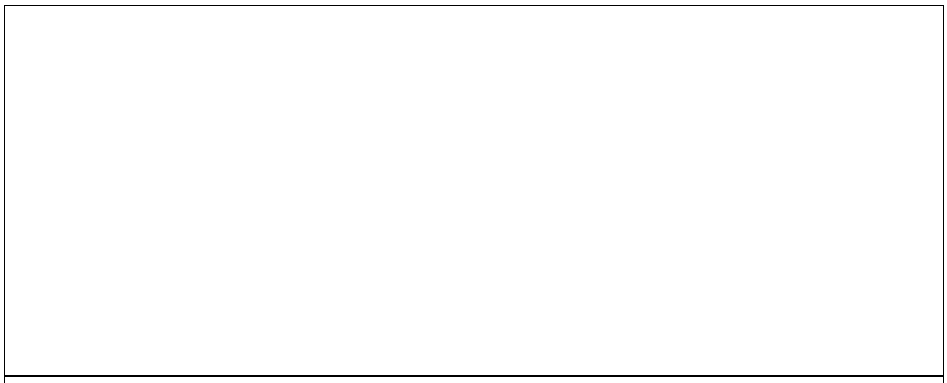


Figure 2-7: (a) Basic common-source amplifier schematic. (b) Voltage transfer characteristic for $V_{DD} = 5$ V, $R_D = 5$ kΩ, and $W/L = 20$.

triode region. The input voltage at which this transition occurs [point C in Figure 2-7(b)] can be computed by setting the right-hand side of Eq. (2.13) equal to $V_{IN} - V_{Tn}$, and solving for V_{IN}. It is interesting to note that graphically, point C can be found through the intersection of the voltage transfer characteristic with the line $V_{IN} - V_{Tn}$. The intersect corresponds to the point where $V_{OUT} = V_{DS} = V_{IN} - V_{Tn} = V_{GS} - V_{Tn}$, i.e., the transition point between saturation and triode for the MOSFET.

For the region where the MOSFET operates in the triode region, we have

$$V_{OUT} = V_{DD} - R_D \cdot \mu_n C_{ox} \frac{W}{L}\left(V_{IN} - V_{Tn} - \frac{V_{OUT}}{2}\right) V_{OUT} \quad (2.14)$$

Unfortunately, solving this expression for V_{OUT} yields an unwieldy square-root expression that is best analyzed graphically. As we can see from the plot in Figure 2-7(b), the most important feature here is that the slope of the voltage transfer characteristic diminishes for large V_{IN}; i.e., the slope of the curve at point D is smaller than the slope at point C. Qualitatively, this can be explained by viewing the MOSFET as a resistor, whose value continues to decrease with V_{IN}. For very large V_{IN}, the output voltage must asymptotically approach 0 V. This can be shown by approximating the MOSFET by its on-resistance for small V_{DS} as given by Eq. (2.9). The output voltage in the vicinity of point D can then be expressed by considering the resistive voltage divider formed by R_D and R_{on}.

$$V_{OUT} \cong \frac{V_{DD} \cdot R_{on}}{R_D + R_{on}} = \frac{V_{DD}}{1 + R_D \cdot \mu_n C_{ox} \frac{W}{L}(V_{IN} - V_{Tn})} \quad (2.15)$$

This result confirms that for large input voltages, V_{OUT} will asymptotically approach zero.

Example 2-2: Voltage Transfer Calculations for a Common-Source Amplifier

Consider the circuit of Figure 2-7(a) with the following parameters: $V_{DD} = 5$ V, $R_D = 10$ kΩ.

(a) Using the standard technology parameters of Table 2-2, calculate the required aspect ratio W/L such that $V_{OUT} = 2.5$ V for $V_{IN} = 1$ V.

(b) Assuming $W/L = 10$, calculate the input voltage V_{IN} that yields $V_{OUT} = 2.5$ V.

SOLUTION

(a) As a first step, we can calculate the drain current that results in $V_{OUT} = 2.5$ V using Eq. (2.12).

$$\begin{aligned} V_{OUT} &= V_{DD} - I_D R_D \\ 2.5V &= 5V - (I_D \cdot 10k\Omega) \\ I_D &= 250\mu A \end{aligned}$$

Since $V_{GS} - V_{Tn} = 0.5$ V $< V_{DS} = 2.5$ V, we know that the device must operate in the saturation region. Therefore,

$$\begin{aligned} I_D &= \frac{1}{2}\mu_n C_{ox} \frac{W}{L}(V_{GS} - V_{Tn})^2 \\ 250\mu A &= \frac{1}{2} \cdot 50\frac{\mu A}{V^2} \cdot \frac{W}{L} \cdot (1V - 0.5V)^2 \end{aligned}$$

and solving for the aspect ratio yields $W/L = 40$. Note that this answer can also be found by direct evaluation of Eq. (2.13), without computing I_D initially.

(b) Since V_{IN} is unknown in this part of the problem, we cannot immediately determine the operating region of the MOSFET. In such a situation, it is necessary to guess the operating region, and later test if the guess was correct.

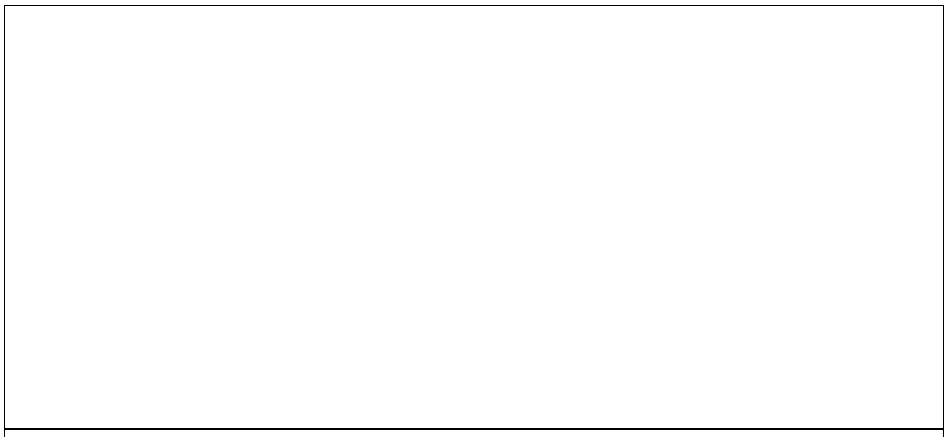


Figure 2-8: Load line plot for the CS amplifier in Figure 2-7(a). Parameters: V_{DD} = 5 V, R_D = 5 kΩ, and W/L = 20.

Let us begin by assuming that the device operates in saturation. We can then write

$$I_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{GS}-V_{Tn})^2$$

$$250\mu A = \frac{1}{2}\cdot 50\frac{\mu A}{V^2}\cdot 10\cdot (V_{GS}-0.5V)^2$$

The two solutions to this equation are V_{GS1} = 1.5 V, and V_{GS2} = –0.5 V. Since we know that the device is off for $V_{GS} < V_{Tn}$, it is clear that V_{GS2} is a non-physical solution that must be discarded. For the obtained V_{GS1}, we must now verify that the device operates in saturation, as initially assumed. It is straightforward to see that this is indeed the case since $V_{GS1} - V_{Tn}$ = 1 V < V_{DS} = 2.5 V. Therefore, the final answer to this problem is V_{IN} = 1.5 V.

If we had initially guessed that the device operates in the triode region, we would write

$$I_D = \mu_n C_{ox}\frac{W}{L}\left(V_{GS}-V_{Tn}-\frac{V_{DS}}{2}\right)V_{DS}$$

$$250\mu A = 50\frac{\mu A}{V^2}\cdot 10\cdot\left(V_{GS}-0.5V-\frac{2.5V}{2}\right)\cdot 2.5V$$

The solution to this equation is V_{GS} = 1.95 V. Since $V_{GS} - V_{Tn}$ = 1.45 V < V_{DS} = 2.5, we see that the obtained result contradicts the assumed operation in triode. Therefore, the next logical step would be to evaluate the saturation equation, as already done above.

2-2-2 Load Line Analysis

A generally useful tool for graphical analysis in electronic circuits is the so-called **load line analysis**. The basis for such an analysis in the context of our circuit is the fact that the current flowing through the transistor (I_D) is equal to the current flowing through the resistor (which is viewed in this context as the load of the circuit). Therefore, if we draw the I-V characteristics of the MOSFET and R_D in one diagram, valid output voltages lie at the intersection of the two curves (equal current). This is further illustrated in Figure 2-8. The load line equation in this plot follows from solving Eq. (2.12) for I_D and is given by

$$I_D = \frac{V_{DD}-V_{OUT}}{R_D} \tag{2.16}$$

The points A, B, C and D marked in Figure 2-8 correspond to the points shown with the same annotation in Figure 2-7(b). Since the transistor drain characteristics are overlaid in Figure 2-8, it is easy to identify the operating regions that correspond to each point. For example, we can immediately see that point B lies in saturation, since the intersect occurs in a region of constant drain current.

Example 2-3: Output Voltage Calculations for a Common-Source Voltage Amplifier

Construct a load line plot to verify the solution of Example 2-2(b) using V_{DD} = 5 V, R_D = 10 kΩ and W/L =10. Use $V_{IN} = V_{GS}$ = 1 V, 1.5 V and 2 V for the drain characteristic plot.

SOLUTION

The solution is shown in Figure Ex2-3. The load line is most

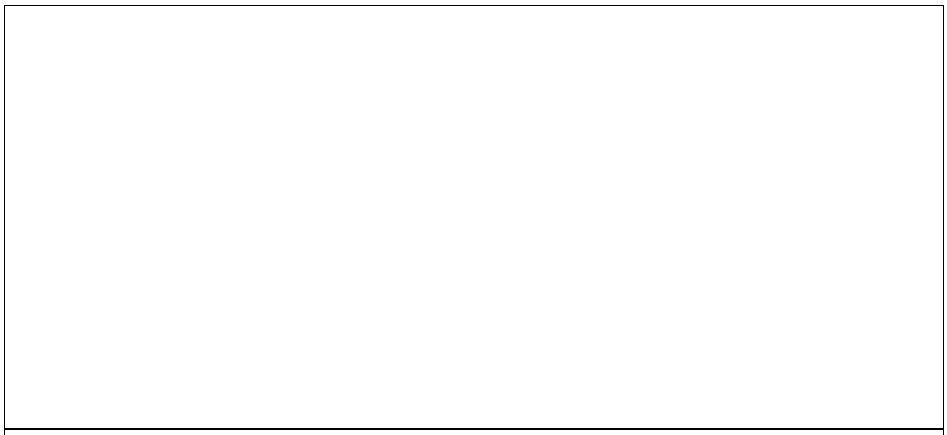


Figure Ex2-3

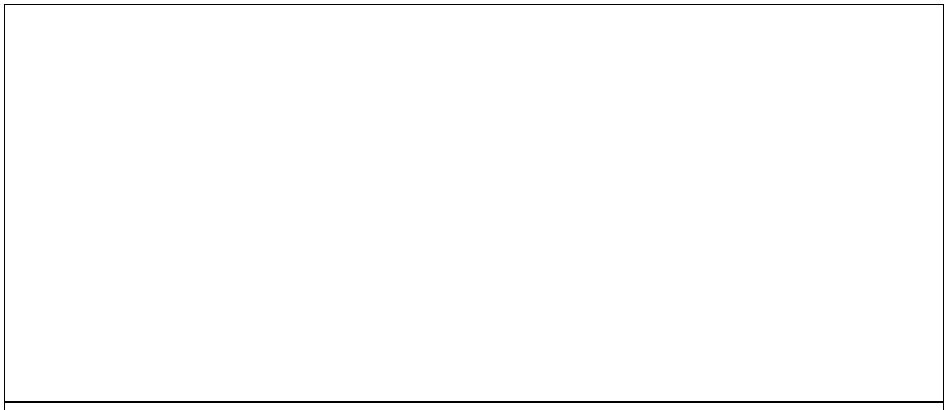


Figure 2-9: (a) CS amplifier with input bias. (b) Transfer characteristic showing the bias point.

easily drawn by connecting the points (0, V_{DD}/R_D = 0.5 mA) and (V_{DD} = 5 V, 0). The drain characteristics are drawn for the three given V_{GS} using the expressions of Table 2-1, by sweeping $V_{DS} = V_{OUT}$ from 0 V to 5 V. The intersect of the load line with the drain characteristic for V_{IN} = 1.5 V confirms the result already obtained in Example 2-2(b).

2-2-3 Biasing

After deriving the voltage transfer characteristic of our amplifier, we are now in a position to evaluate this circuit from an application standpoint. As we have discussed in Chapter 1, a common objective for a voltage amplifier is to create large output voltage excursions from small changes in the applied input voltage. With this objective in mind, it becomes clear that only a limited range of the transfer characteristic in Figure 2-7(b) is useful for amplification. For example, a change in the input voltage applied around point D in Figure 2-7(b) yields almost no change in the output voltage. In order to amplify small changes in V_{IN} into large changes in V_{OUT}, the transistor should be operated in the saturation region, i.e., in the vicinity of point B. The general concept of operating a circuit and its constituent transistor(s) around a preferable, and useful operating point is called **biasing**.

Biasing generally necessitates the introduction of auxiliary voltages and/or currents that bring the circuit into the desired state. For the circuit considered in this section, proper biasing can be achieved by decomposing the input voltage into a constant component, and a component that represents the incremental voltage change to be amplified; this is illustrated in Figure 2-9(a). The incremental voltage component v_{in} could represent, for instance, the signal generated by a microphone or a similar transducer. The voltage V_{IN} is a constant voltage that defines the point on the overall transfer characteristic around which the incremental in v_{in} is applied. We call V_{IN} the **input bias voltage** of the circuit.

Per IEEE convention, the total quantity in such a decomposition is denoted using a lower case symbol and upper case subscript, i.e., $v_{IN} = V_{IN} + v_{in}$ in our example. Similarly, the drain current is decomposed as $i_D = I_D + i_d$, where I_D is the current at the operating point, and i_d captures the current deviations due to the applied signal.

Figure 2-9(b) elucidates this setup further using the circuit's transfer characteristic. With $v_{in} = 0$, the output is equal to V_{OUT}, which is called the **bias point** or **operating point** of the output node. The bias point is sometimes also called the **quiescent point** (Q), since the corresponding voltage level corresponds to that of a "quiet" input. Note that V_{OUT} can be calculated by evaluating Eq. (2.13), as done previoulsy.

With some nonzero v_{in} applied, the output will now see an excursion away from the bias point. For example, applying a positive v_{in} will result in a negative incremental change v_{out} at the output. How can we compute v_{out} for a given v_{in}? Since Eq. (2.13) must hold for the total quantities $v_{IN} = V_{IN} + v_{in}$ and $v_{OUT} = V_{OUT} + v_{out}$ we can write

$$v_{OUT} = V_{DD} - R_D \cdot \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(v_{IN} - V_{Tn})^2 \tag{2.17}$$

or

$$V_{OUT} + v_{out} = V_{DD} - R_D \cdot \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{IN} + v_{in} - V_{Tn})^2 \tag{2.18}$$

In order to simplify this expression, and since we are only interested in the change of v_{out} as a function of v_{in}, it is useful to eliminate the constant term from this expression, given by

$$V_{OUT} = V_{DD} - R_D \cdot \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{IN} - V_{Tn})^2 \tag{2.19}$$

After subtracting Eq. (2.18) from Eq. (2.17) and rearranging the terms, we obtain

$$v_{out} = -R_D\mu_n C_{ox}\frac{W}{L}(V_{IN} - V_{Tn}) \cdot v_{in}\left[1 + \frac{v_{in}}{2(V_{IN} - V_{Tn})}\right] \tag{2.20}$$

Using the drain current expression of Eq. (2.7), and by defining

$$V_{OV} = (V_{GS} - V_{Tn})\big|_Q = V_{IN} - V_{Tn} \tag{2.21}$$

Eq. (2.20) can be further simplified and rewritten as

$$v_{out} = -\frac{2I_D}{V_{OV}} \cdot R_D \cdot v_{in}\left(1 + \frac{v_{in}}{2V_{OV}}\right) \tag{2.22}$$

where I_D is the transistor's drain current at the bias point, and V_{OV} is introduced as a symbol for the quiescent point gate overdrive voltage.

From the end result in Eq. (2.22), we see that v_{out} is a nonlinear function of v_{in}. This is not surprising, since we are employing a transistor that exhibits a nonlinear I-V characteristic. While this derivation was relatively simple, the analysis of nonlinear circuits in general tends to be complex. Picture a circuit that contains several transistors, as for instance a cascade connection of several stages of the amplifier circuit considered here. Even with only a few nonlinear elements, most cases involving practical circuits with just moderate complexity tends to yield unwieldy expressions. A widely used solution to this problem is to approximate the circuit behavior using a linear model around its operating point, which we will discuss next.

2-2-4 The Small-Signal Approximation

Eq. (2.22) is written in a format that suggests an opportunity for simplification. Provided that $v_{in} << V_{OV}$, the bracketed term is close to unity and we can write

$$v_{out} \cong -\frac{2I_D}{V_{OV}} \cdot R_D \cdot v_{in} = A_v \cdot v_{in} \tag{2.23}$$

where A_v is a constant voltage gain term that relates the incremental input and output voltages.

Interestingly, the term A_v can also be found using basic calculus. Assuming that the incremental voltages represent infinitesimally small deviations in the total signal, we can re-write Eq. (2.23) as

$$dv_{OUT} = A_v \cdot dv_{IN} \tag{2.24}$$

and therefore

$$A_v = \frac{dv_{OUT}}{dv_{IN}}\bigg|_Q = \frac{dv_{OUT}}{dv_{IN}}\bigg|_{v_{IN} = V_{IN}} \tag{2.25}$$

where the derivative is evaluated at the circuit's operating point Q that is fully defined by choice of the input bias voltage V_{IN}. By applying Eq. (2.25) to Eq. (2.17), we find

$$\begin{aligned} A_v &= \frac{d}{dv_{IN}}\left[V_{DD} - R_D \cdot \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(v_{IN} - V_{Tn})^2\right]\bigg|_{v_{IN} = V_{IN}} \\ &= -R_D\mu_n C_{ox}\frac{W}{L}(v_{IN} - V_{Tn})\bigg|_{v_{IN} = V_{IN}} \\ &= -R_D\mu_n C_{ox}\frac{W}{L}(V_{IN} - V_{Tn}) = -R_D\mu_n C_{ox}\frac{W}{L}V_{OV} \end{aligned} \tag{2.26}$$

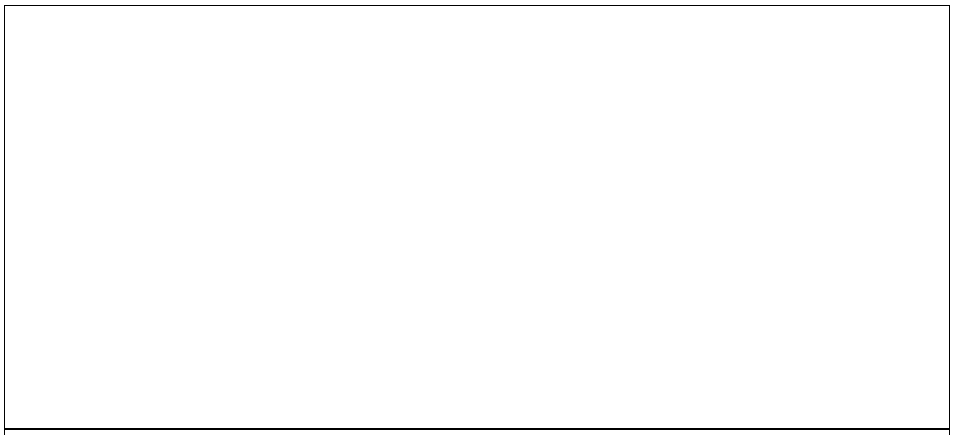


Figure 2-10: Concept of small-signal voltage gain. In the small signal model, the input and output voltages are linearly related through A_v, the slope of the tangent at the operating point of the large-signal transfer characteristic.

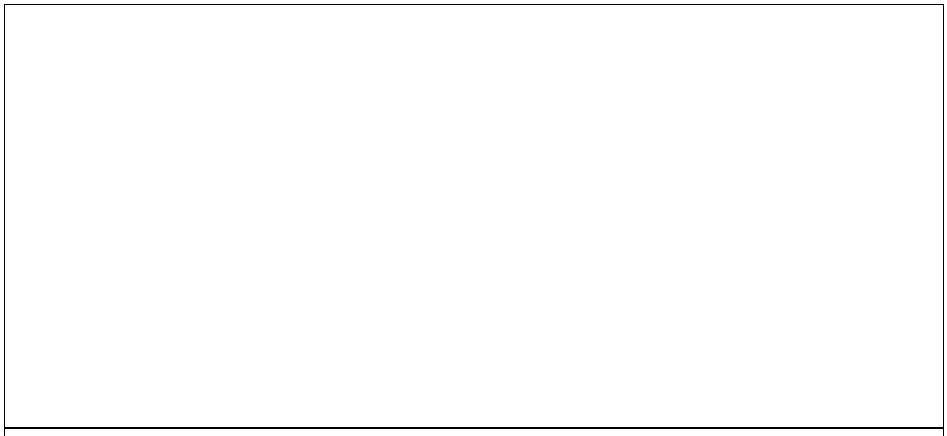


Figure 2-11: Examples of signal clipping and distortion. (a) Output waveform is clipped due to supply voltage limit. (b) Output waveform drives the MOSFET into the triode region.

Finally, using Eq. (2.7), we find

$$A_v = -\frac{2I_D}{V_{OV}} \cdot R_D \tag{2.27}$$

which is the same result obtained previously. The voltage gain A_v can be interpreted graphically as shown in Figure 2-10. From Eq. (2.25) and basic calculus we know that A_v is the slope of the tangent to the transfer characteristic at the point (V_{IN}, V_{OUT}), which is the operating point of the circuit.

In analog circuit nomenclature, A_v is called the **small-signal voltage gain** of the circuit; this emphasizes that this quantity is only suitable for calculations with "small" signals such that nonlinear effects are negligible. In the particular circuit considered here, "small" means $v_{in} << V_{OV}$, as seen from our analysis.

The general concept of approximating circuit behavior by assuming small-signal excursions around an operating point is called **small-signal approximation**. In order to clearly distinguish a circuit transfer characteristic obtained through such an approximation from one that incorporates the nonlinear transistor behavior [e.g., Eq. (2.17)], the term **large-signal transfer characteristic** is typically used for the latter.

As we shall see in the remainder of this module, working with small-signal approximations greatly simplifies analog circuit analysis and design. The price paid for the approximation, however, is that the resulting equations by themselves cannot be used to reason about the circuit's behavior for large signals, as for instance signals where v_{in} is comparable to, or even greater than V_{OV}. As illustrated in Figure 2-10, the small-signal approximation essentially creates a new coordinate system that linearly relates the input and output voltages. In this model, the

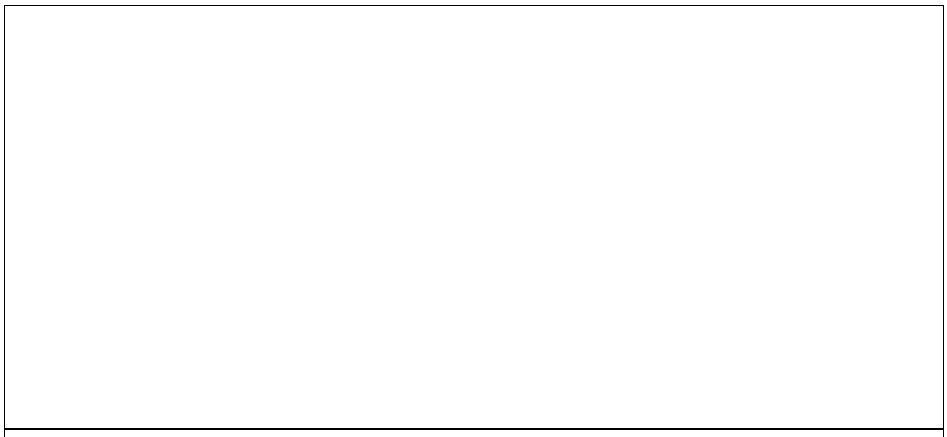


Figure 2-12: (a) Large-signal transfer characteristic of an n-channel MOSFET in the saturation region. (b) Small-signal transconductance (g_m) at the operating point.

output voltage follows the input linearly, no matter how large the applied voltage is. In reality, considering the circuit's large signal transfer characteristic, signal clipping and strong waveform distortion can occur for large excursions and poorly chosen bias points. Examples of such cases are illustrated in Figure Figure 2-11.

Example 2-4: Signal clipping.

Consider the circuit of Figure 2-9, using the parameters from Example 2-2(b): $V_{DD} = 5$ V, $R_D = 10$ kΩ, $W/L = 10$ and V_{IN} is adjusted to 1.5 V, so that $V_{OUT} = 2.5$ V at the circuit's operating point. Calculate the most negative excursion that the incremental input voltage v_{in} can assume before the output is "clipped" to V_{DD} [as in Figure 2-11(a)].

SOLUTION

The circuit's output voltage is given by

$$v_{OUT} = V_{DD} - R_D \cdot \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{IN} + v_{in} - V_{Tn})^2$$

Clipping v_{OUT} to the supply voltage implies $v_{OUT} = V_{DD}$. This requires

$$\begin{aligned} 0 &= V_{IN} + v_{in} - V_{Tn} \\ v_{in} &= -(V_{IN} - V_{Tn}) = -V_{OV} \\ v_{in} &= -(1.5V - 0.5V) = -1V \end{aligned}$$

In words, applying a negative signal (v_{in}) at the input of magnitude larger than 1 V will cause the output to reach the supply voltage. Making v_i more negative will create a "plateau" in the output waveform as shown in Figure 2-11(a).

In a majority of analog circuits, it is sufficient to use the large-signal characteristic for bias-point and signal-range calculations. For all other purposes, as for instance voltage gain calculations, it is usually appropriate and justifiable to work with small-signal approximations. Without this clever split in the analysis, most analog circuits of only moderate complexity would not be amenable to hand analysis, simply because the nonlinear nature of the transistors would create prohibitively complex systems of nonlinear equations.

Circuits that are designed to amplify small signals from a transducer are classical examples where the small-signal approximation works. Consider, for instance, the above-discussed amplifier circuit fed with an input signal from a radio antenna, which is often on the order of several hundred mirovolts. As long as V_{OV} is chosen larger than several hundred milivolts, the small signal approximation will hold with reasonable accuracy. Other examples (not covered in this module) include amplifiers that rely on electronic feedback, which tends to minimize the signal excursions around a circuit's bias point (see Reference 2).

As a final remark, it should be noted that even if the input to a circuit is "small," the output will always show at least some amount of nonlinear distortion. In our basic amplifier, this distortion is caused by the bracketed term in Eq. (2.22). In cases where even weak distortion is an issue, the designer often employs computer simulation tools to study the relevant behavior. From a design perspective, deviations from linearity can be minimized if needed. For the discussed common-source amplifier, this is seen from Eq. (2.22): decreasing the ratio v_{in} / V_{OV}, either by reducing v_{in} or by increasing V_{OV} will result in improved linearity.

The exact analysis of nonlinear distortion is beyond the scope of this module, and is typically treated only in advanced integrated circuit texts, as for instance Reference 3. We will

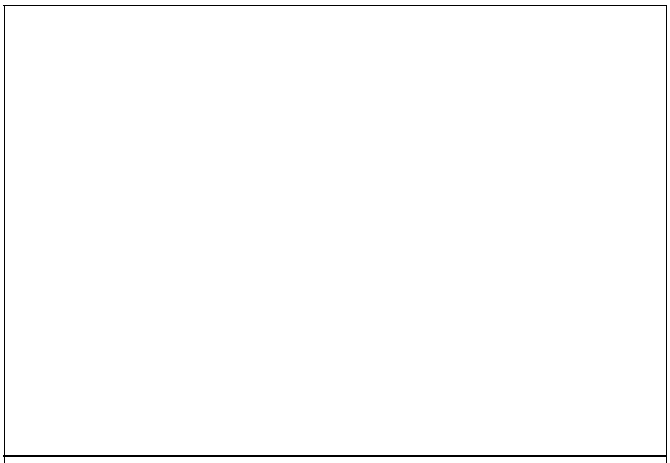


Figure 2-13: Small-signal model of the circuit in Figure 2-7(a).

focus here primarily on studying the relevant behavior of analog circuits using a linear small-signal abstraction, aided by basic bias-point and signal-range calculations.

2-2-5 Transconductance

The method of differentiating a circuit's large-signal transfer characteristic to obtain a small-signal approximation was straightforward for the simple one-transistor circuit discussed so far. Unfortunately, for a larger circuit it is usually much more difficult and often tedious to derive a complete transfer characteristic in the form of Eq. (2.17).

A clever workaround that is predominantly used in analog circuit analysis is based on linearizing the circuit element-by-element around the operating point. This method is applied in three steps: (1) Compute all node voltages and branch currents at the operating point using the devices' large-signal model. (2) Substitute linear models for all nonlinear components and compute their parameters using the operating point information. (3) Compute the desired transfer function using the linear model obtained in step 2.

The biggest advantage of this method, called **small-signal analysis**, is that it avoids computing the large-signal transfer characteristic of the circuit, and instead defers the transfer function analysis until all elements have been approximated by linear models. The linearized models of nonlinear elements, such as MOSFETs, are typically called **small-signal models**.

We will now illustrate the small-signal analysis approach by applying it to the basic common-source amplifier example covered so far in this chapter. Consider first the MOSFET device in Figure 2-7(a). In general, once the operating point of the transistor is known, the small-signal model is obtained by differentiating the large signal I-V relationships at this point. This is further illustrated in Figure 2-12 assuming that the MOSFET is biased in the saturation region. The proportionality factor that links the incremental drain current (i_d) and the gate-source voltage (v_{gs}) is given by the slope of the tangent to the large signal transfer characteristic at the bias point. This quantity is called **transconductance**, or g_m. Mathematically, we can write

$$g_m = \frac{i_d}{v_{gs}} = \left.\frac{di_D}{dv_{GS}}\right|_{v_{GS}=V_{GS}} \tag{2.28}$$

In order to find the transconductance for the saturation region of the device, we evaluate Eq. (2.28) using Eq. (2.7). This yields

$$g_m = \mu_n C_{ox}\frac{W}{L}(V_{GS}-V_{Tn}) = \mu_n C_{ox}\frac{W}{L}V_{OV} \tag{2.29}$$

Alternative forms of this expression are obtained by eliminating V_{OV} or $\mu_n C_{ox} W/L$ using Eq. (2.7), which gives

$$g_m = \sqrt{2\mu_n C_{ox}\frac{W}{L}I_D} \tag{2.30}$$

or

$$g_m = \frac{2I_D}{V_{OV}} \tag{2.31}$$

All of the above equations can be used to calculate g_m; the choice of which equation is used depends on the given parameters. The physical unit for transconductance is $A/V = \Omega^{-1}$, or **Siemens** (S).

Once the transconductance is determined, we can insert the model of Figure 2-12(b) into the original circuit [Figure 2-7(a)] for further analysis. The resulting small-signal circuit equivalent is shown in Figure 2-13. No modeling modification is needed for the resistor R_D, as it is already assumed in Figure 2-7(a) that it follows a linear I/V law ($V = I \cdot R$). However, since the supply voltage is constant in the large-signal model, it must be replaced with 0 V or ground (GND) in the small-signal model. This is because the differentiation of a constant quantity yields zero.

Using the model of Figure 2-13, we now apply Kirchhoff's laws at the output and find

$$v_{out} = -g_m \cdot v_{in} \cdot R_D \tag{2.32}$$

Finally, by substituting Eq. (2.31) we obtain

$$v_{out} = -\frac{2I_D}{V_{OV}} \cdot R_D \cdot v_{in} = A_v \cdot v_{in} \tag{2.33}$$

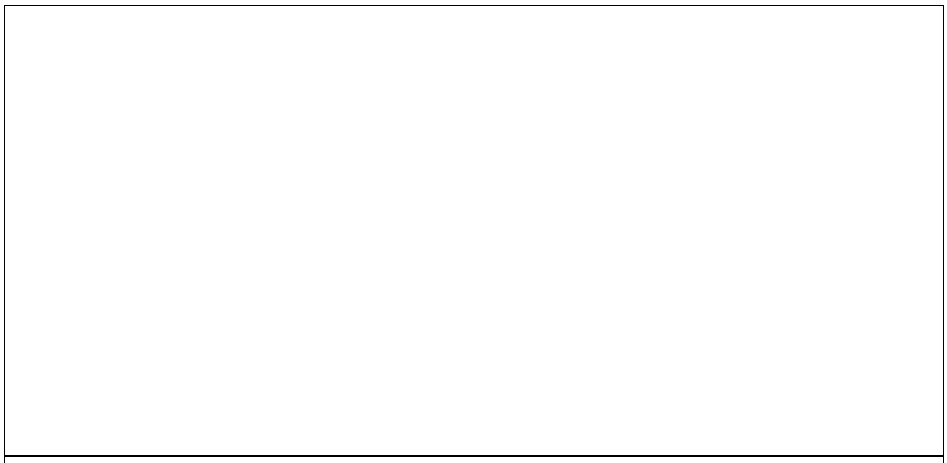


Figure 2-14: (a) P-channel common-source amplifier schematic. (b) Voltage transfer characteristic for V_{DD} = 5 V, R_D = 5 kΩ, and W/L = 40.

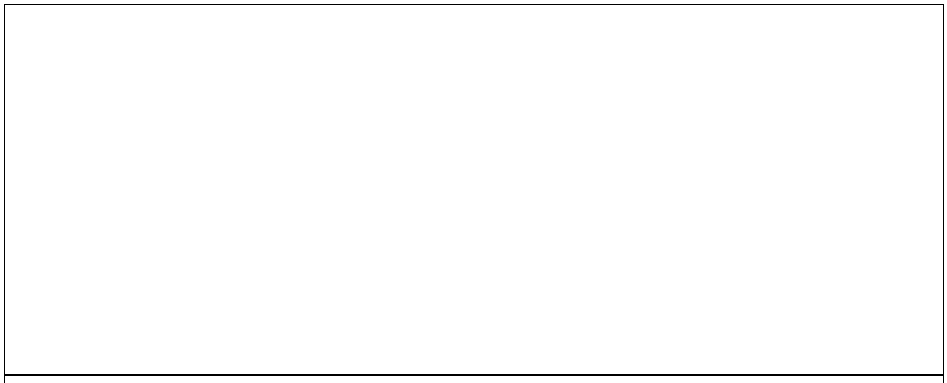


Figure 2-15: Small-signal model of the p-channel common-source amplifier: (a) using the p-channel model directly, (b) a flipped version of (a), and (c) with sign changes applied to show equivalence with the n-channel model.

As expected, this result is equivalent to what was obtained by applying the small-signal approximation to Eq. (2.22), and also by differentiating Eq. (2.17) at the operating point. However, as indicated previously, the big advantage of working with a small-signal model for individual transistors is that this simplifies the analysis of larger circuits.

2-2-6 P-Channel Common-Source Voltage Amplifier

As shown in Figure 2-14(a), we can also build a CS amplifier using a p-channel device. Similar to the n-channel case, we can derive a large-signal transfer characteristic using Eq. (2.11) and by applying KVL at the output node. The resulting plot is shown in Figure 2-14(b). Compared to the n-channel case [Figure 2-7(b)], one can show that the characteristic is flipped sideways (since $V_{SG} = V_{DD} - V_{IN}$) and upside down (because I_D is negative for a p-channel transistor). Therefore, for small V_{IN} near 0 V, the device operates in the triode region and the output voltage is close to V_{DD}. For $V_{IN} = V_{DD}$, the device is off and the output is at 0 V, since V_{SG} = 0 and no current flows in the device.

In terms of its small-signal-signal model, it follows that the p-channel CS amplifier is identical to the n-channel version. This can be seen intuitively by comparing Figure 2-7(b) and Figure 2-14(b): in the region around point B, both transfer characteristics exhibit a negative slope and therefore will behave alike for small perturbations. We will show this formally in the following discussion.

We begin by inserting a small-signal model for the transistor as shown in Figure Figure 2-15(a). Based on the positive

variable convention of Eq. (2.11), we define the transconductance for the p-channel transistor as

$$\begin{aligned} g_m &= \frac{d}{dv_{SG}}((-i_D))\Big|_Q \\ &= \mu_p C_{ox}\frac{W}{L}(V_{SG}+V_{Tp}) \\ &= \frac{2(-I_D)}{V_{SG}+V_{Tp}} = \frac{2(-I_D)}{V_{OV}} \end{aligned} \tag{2.34}$$

Note that through this expression, we have defined V_{OV} for the p-channel case as $V_{SG} + V_{Tp}$, which is a positive quantity for a p-channel device in the "ON" state.

Although we could solve for the small-signal voltage gain directly with the circuit shown in Figure 2-15(a) it is easier to flip the transistor 180° [Figure 2-15(b)] so that the circuit appears similar to the n-channel version. Since $v_{sg} = -v_{gs}$ we can change signs at the input and the dependent current source and find that the p-channel common-source amplifier small-signal model is identical to the n-channel version as shown in Figure 2-15(c).

This result applies more generally to all the transistor configurations and model extensions that we will study in this module. Once the operating point parameters of a p-channel device have been determined [e.g., a calculation of g_m using Eq. (2.23)], it is perfectly valid to replace it with an n-channel equivalent. This is a very powerful and convenient result, since it allows us to focus on n-channel only configurations in small-signal analyses, without having to worry about the specific sign conventions of p-channels.

2-2-7 Modeling Bounds for the Gate Overdrive Voltage

According to Eq. (2.31), $g_m/I_D = 2/V_{OV}$ tends to infinity as the gate overdrive voltage V_{OV} approaches zero. This implies that for a transistor that is "barely on," we can extract very large transconductance values for only small bias currents. Unfortunately, this behavior is incorrect, and stems from limitations of the device model discussed in Section 2-1. As V_{OV} approaches zero, a more complex analysis is needed to predict the drain current and its derivative with respect to gate voltage (see Reference 4).

Figure 2-16 plots the g_m/I_D characteristic of a MOSFET, and the expected behavior based on Eq. (2.31). As we can see, for $V_{OV} < 150$ mV, a large discrepancy exists between a physical device and the prediction based on the simple square-law model used in this treatment. In order to avoid unrealistic design outcomes due to this modeling limitation, we define a

Figure 2-16: Transconductance to current ratio predicted by Eq. (2.31) and the behavior of an actual MOSFET.

bound for the minimum allowed gate overdrive voltage for all circuits covered in this module

$$V_{OV} \geq V_{OVmin} = 150mV \tag{2.35}$$

Designing with a smaller V_{OV} would require a more elaborate model for hand calculations, which is beyond the scope of this module. The interested reader is referred to advanced material on this topic, available for example in References 4 and 5.

2-2-8 Voltage Gain and Drain Biasing Considerations

In the basic common-source amplifier discussed so far, the drain resistor R_D serves a dual purpose: (1) it translates the device's incremental drain current (i_d) into a voltage (v_{out}), and (2), it supplies the quiescent point drain current (I_D) for the MOSFET. As we shall show next, this creates an undesired link between the bias point constraints of the circuit and the achievable small-signal voltage gain of the amplifier. To see this, we re-write Eq. (2.33) as shown below

$$A_v = -\frac{2I_D}{V_{OV}} \cdot R_D = -2\frac{V_R}{V_{OV}} \tag{2.36}$$

where $V_R = V_{DD} - V_{OUT}$ is the voltage drop across R_D at the operating point. This result leads to several interesting conclusions. First, note that the voltage gain of the amplifier is fully determined once V_{OV} and voltage drop across R_D are known. For example, if the circuit is biased such that $V_{OV} = 0.2$ V and $V_R = 2$ V, we have $A_v = -20$; regardless of the particular W, L or $\mu_n C_{ox}$ of the employed MOSFET. Second, since the possible

Figure 2-17: Amplifier load lines for large and small values of R_D.

values for V_{OV} are lower bounded [Eq. (2.35)] and V_R is upper bounded (finite V_{DD}), there exists a maximum possible A_v that can be obtained

$$|A_{vmax}| = 2\frac{V_{Rmax}}{V_{OVmin}} = 2\frac{V_{DD}-V_{OVmin}}{V_{OVmin}} \cong 2\frac{V_{DD}}{V_{OVmin}} \qquad (2.37)$$

In this result, it was assumed the transistor is biased at the edge of the triode region; a somewhat impractical, but appropriate limit case to consider. Evaluating the above expression for V_{DD} = 5 V and V_{OVmin} = 150 mV yields $|A_{vmax}| \cong 67$. Can we overcome this limit and change our amplifier such that it can achieve voltage gains beyond this value?

In order to investigate this, consider the load line illustrations shown in Figure 2-17. As explained in Section 2-2-2, the load line for our circuit is defined by the points (0, V_{DD}/R_D) and (V_{DD}, 0). From the location of these points, we see that the x-axis intercept of the load line is fixed, while the y-intercept moves lower with larger values of R_D. This reduces the slope of the load line, resulting in a larger small-signal voltage gain of the circuit. Furthermore, note that for a fixed quiescent point drain current I_D, larger R_D shifts the output bias point V_{OUT} to smaller values, i.e., closer to the edge of the MOSFET's triode region. This observation captures the result of Eq. (2.37) in a graphical way: we cannot increase the small-signal voltage gain beyond a certain limit due the link between V_{OUT} and the chosen R_D.

A more ideal situation is depicted in Figure 2-18. If we could somehow create a load line that "rotates" about the desired operating point (as a function of R_D), the voltage gain could be set independently of V_{OUT}. A modified drain network that lets us achieve this behavior is shown in Figure 2-19(a). In this circuit, R_D is now connected to a voltage V_B (instead of the supply voltage V_{DD}) and an ideal current source I_B is used to provide a fixed current. In a realistic implementation circuit, I_B can be built, for example, using a p-channel MOSFET that operates in saturation. For the time being, we will neglect such implementation details, and postpone the discussion of current sources to Chapter 5.

With this new configuration, the relationship between i_D and v_{OUT} becomes

$$i_D = I_B + \frac{V_B - v_{OUT}}{R_D} \qquad (2.38)$$

or

$$v_{OUT} = V_B + (I_B - i_D)R_D \qquad (2.39)$$

As we can see from these equations [also graphically shown in Figure 2-19(b)], at the point $v_{OUT} = V_B$, we have $i_D = I_B$, regardless of the value of R_D. Therefore, utilizing this point as the operating point of our amplifier precisely achieves the goal we have in mind. In particular, we wish to set $I_B = I_D$, the desired quiescent point drain current of the MOSFET, and $V_B = V_{OUT}$, the desired output operating point. With this choice, the role of the current source is to provide the MOSFET's bias current, while the resistor R_D is responsible only for converting the incremental drain current into an incremental output voltage; no DC bias current flows in this element.

Since the voltage source V_B and the current source I_B aid in maintaining the circuit's bias point, we generally classify these elements as biasing sources. However, it is important to distinguish their function from the input bias voltage V_{IN}. V_{IN} directly sets the quiescent point gate-source voltage of the transistor and therefore fully defines the operating point on the MOSFET's I-V characteristic and the corresponding drain current. In the above-described scenario, I_B is adjusted to supply this same drain current, but does not define it. We therefore categorize I_B an **auxiliary bias current** that helps sustain, but does not set the quiescent point of the transistor. Since V_B defines the quiescent point output voltage of the circuit, we refer to this element as the **output bias voltage**.

Lastly, it is important to note that for the modified circuit in Figure 2-19(a), the previously derived small-signal model shown in Figure 2-13 still applies. This can be understood by differentiating Eq. (2.39) at the operating point to find the small-signal equivalent of the network placed at the drain of the amplifier

$$\left.\frac{dv_{OUT}}{di_D}\right|_Q = \frac{d}{di_D}(V_B + (I_B - i_D)R_D) = -R_D \qquad (2.40)$$

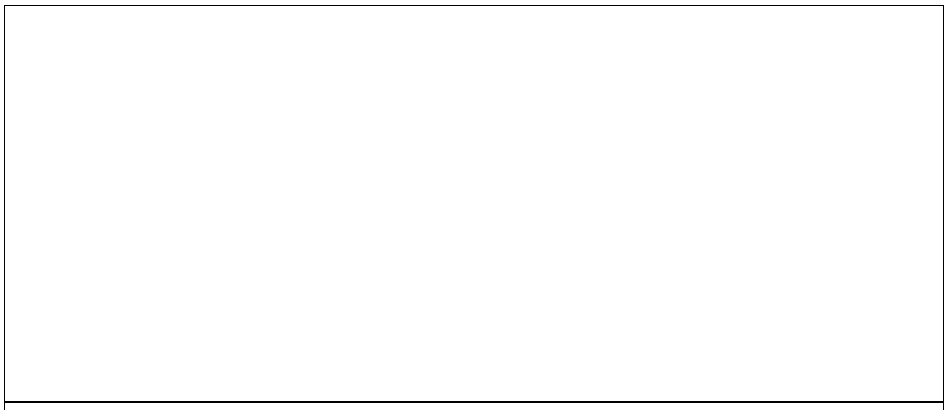


Figure 2-18: Desired load line behavior.

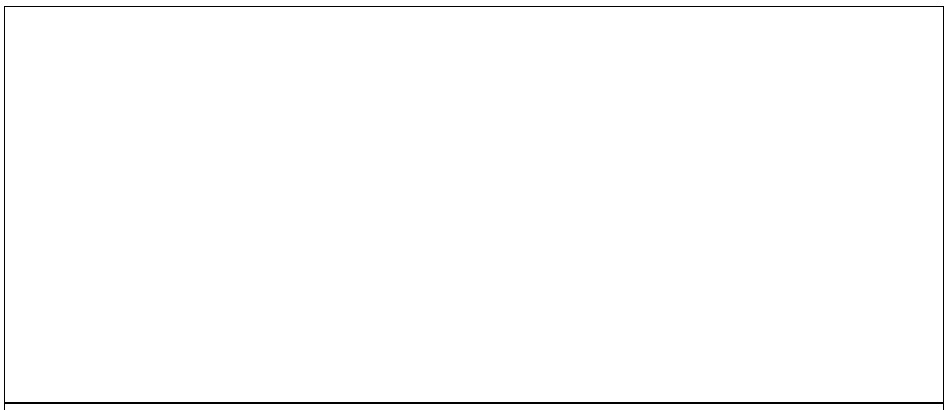


Figure 2-19: (a) CS amplifier with modified drain biasing scheme. (b) Load line characteristic.

As this result indicates, and as we have seen previously, any constant sources, such as V_B, I_B, etc., drop out of the small signal model, which captures only components that affect the incremental changes in currents and voltages around the circuit's operating point.

For the circuit in Figure 2-19(a), one might now be tempted to think that we can obtain an arbitrarily large voltage gain, as long as R_D is made very large. Unfortunately this is not the case for several reasons, the first of which stems from physical effects that we have not yet included in the MOSFET model. This aspect is further discussed in Section 2-3. In addition, there are practical limitations to the attainable voltage gain, discussed next.

2-2-9 Sensitivity of the Bias Point to Component Mismatch*

Consider a CS amplifier biased at the gate as done previously with a bias voltage V_{IN}, directly setting up the quiescent point drain current I_D (depending on W/L and other relevant device parameters). Our goal in using the drain bias network of Figure 2-19(a) is then to set $I_B = I_D$ and chose V_B such the output is biased at a reasonable, desired V_{OUT}. Unfortunately, in practice, we can never achieve $I_B = I_D$ exactly; there will always be a nonzero $\Delta I = I_B - I_D$. This case is shown in Figure 2-20. As illustrated, the finite ΔI leads to a shift ΔV away from the desired output operating point V_B. This shift is proportional to R_D, since the current difference ΔI flows into R_D, creating the undesired ΔV.

Clearly, as we let R_D assume very large values, it becomes harder to absorb differences between I_D and I_B and still maintain an operating point that is close to the desired value. Let us examine this situation through a numerical example.

Example 2-5: Output Bias Voltage Shift due to Current Mismatch.

Consider the CS amplifier of Figure 2-19(a), biased at the gate such that V_{OV} = 500 mV. Assume I_D = 200 μA, and that R_D is

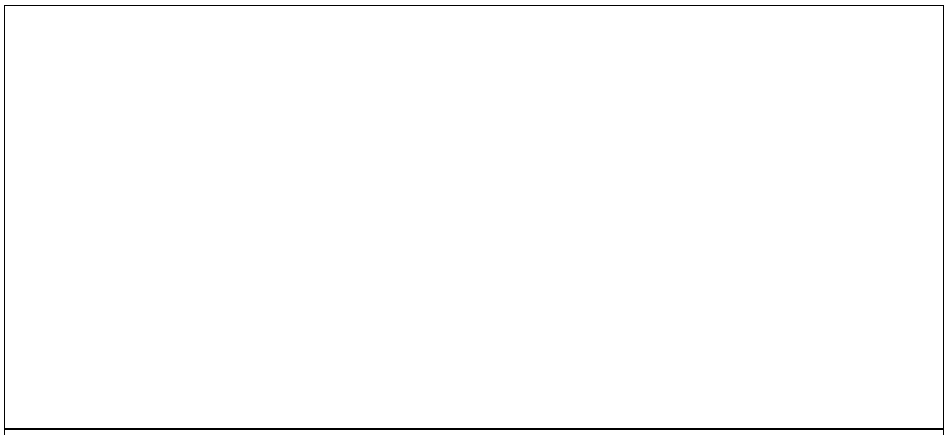


Figure 2-20: Bias point shift due to mismatch in I_B and I_D.

chosen such that the amplifier achieves $A_v = -400$. Considering Figure 2-20, how much mismatch between I_B and I_D (in %) can be tolerated such that V_{OUT} deviates from the intended bias point (V_B) by no more than $\Delta V = 500$ mV? Repeat this calculation for $A_v = -40$ and -4.

SOLUTION

We begin by computing the transconductance of the MOSFET

$$g_m = \frac{2I_D}{V_{OV}} = \frac{2 \cdot 200\mu A}{0.5V} = 800\mu S$$

In order to achieve $A_v = -400$, we require $R_D = 400/800\ \mu S = 500\ k\Omega$. Therefore,

$$\frac{\Delta I}{I_D} = \frac{\Delta V}{I_D R_D} < \frac{500mV}{200\mu A \cdot 500k\Omega} = 0.5\%$$

For $A_v = -40$ and $A_v = -4$, the result modifies to 5%, and 50%, respectively.

As expected, this result confirms that for larger $|A_v|$, the auxiliary bias current I_B must match the MOSFETs drain current more accurately. How precisely can we match these two currents? Unfortunately, answering this question in detail is beyond the scope of this module, and is the subject of advanced research papers such as Reference 6. Nonetheless, it can be said in general that matching currents, voltages or any other electrical quantities in today's integrated circuits to better than 1% requires special care and understanding. In some cases, even 10% matching can be hard to guarantee. With this guideline in mind, it becomes clear that the circuit of Figure 2-20(a) may become impractical if we aim for too much voltage gain.

The general issue of properly dealing with variability in integrated circuit components is a complex topic that is still being actively researched. At the introductory level of this module, the main point that the reader should retain is that any circuit whose bias point relies on precisely matched components or high absolute accuracy in any electrical parameter may not be robust in the presence of component variability. In general, experienced circuit designers avoid situations that resemble the "balancing of a marble on the tip of a cone," i.e., circuits that will be overly sensitive to variations in component parameters. We will take up this point once more when discussing practical biasing circuits in Chapter 5.

As a final note concerning the issue of variability, it is worth mentioning that electronic feedback can help alleviate problems as the one analyzed in Example 2-5. Picture for example adding an auxiliary circuit to Figure 2-19(a) that somehow measures V_{OUT}, and adjusts V_{IN} (and therefore I_D) until the desired output operating voltage is set. In such schemes, relatively large variations in I_B can be absorbed. Feedback circuits are not covered in this module, but are the subject of advanced texts such as Reference 2.

2-3 Channel Length Modulation

The MOSFET model used so far assumes that the drain current in the saturation region is independent of the drain-source voltage. This behavior corresponds to that of an ideal current source, which is generally non-physical. A more realistic output characteristic observed in real MOSFETs is shown in Figure 2-21. For a physical MOSFET, I_D tends to increase with V_{DS}; an effect that can be explained (to first order) as a voltage dependent modulation of the channel length (see Section 2-3-1).

When inserted into the circuit of Figure 2-19(a), the dependence of drain current on v_{DS} (v_{OUT}), will have an impact on the overall transfer characteristic of the amplifier, which we will thus consider in this section. For simplicity, let us first investigate the case of $R_D \rightarrow \infty$, i.e., no explicit drain resistance, and using only an ideal current source in the drain biasing network [see Figure 2-22(a)]. In this case, the load line is horizontal at $I_B = I_D$, as shown in Figure 2-22(b). As before, the operating point Q is established at the point where the load line and the device's drain characteristic for the applied quiescent point input voltage (V_{IN}) meet.

Figure 2-21: Realistic n-channel I-V plots, incorporating the dependence of saturation drain current on V_{DS}. (solid lines), along with the first-order model assumed previously (dotted lines). Parameters: W = 20 µm, L = 2 µm.

Note that similar to the case considered in the previous section, the output operating point voltage in this circuit will shift by large amounts for relatively small changes in I_B (or MOSFET parameters). This issue must be addressed when this circuit is used in practice, for example by providing a feedback mechanism that adjusts I_B such that the desired V_{OUT} is maintained in the presence of component variations.

Assuming that a well defined operating point has been established by some means, incremental changes in v_{IN} applied around the bias point Q will force the intersection of the horizontal load line with the MOSFET's drain characteristics to move sideways (since the drain current cannot change), creating a finite output voltage excursion v_{out}. The magnitude of this voltage excursion, and thus the voltage gain of the circuit, depends on the slope of the MOSFET's drain characteristic in saturation. The voltage gain achieved in this configuration is commonly called the **intrinsic voltage gain** of the MOSFET, as it represents the voltage gain of the transistor by itself, without any added resistances in the drain bias network. By the same reasoning, the circuit of Figure 2-22(a) is often called the **intrinsic voltage gain stage**. The voltage gain and other parameters of more complex amplifier circuits are often directly related to the intrinsic voltage gain of their constituent transistors, giving this parameter a fundamental significance in circuit design.

2-3-1 The λ-Model

Unfortunately, the intrinsic voltage gain of a MOSFET cannot be predicted using the MOSFET model established so far. We will therefore extend the first-order MOSFET model to incorporate the dependence of the saturation current on the drain-source voltage. As a first step, we will describe the effect in terms of large-signal equations. Next, we will apply a small-signal approximation that makes it possible to capture the I_{Dsat} – V_{DS} dependence through a single resistor added to the MOSFET's small-signal model.

We begin by revisiting an approximation that was made in Section 2-1. In order to arrive at the constant drain current expression in saturation (Eq. (2.7)), it was assumed that ΔL, the distance from the pinch-off point to the drain is negligible relative to the channel length L. In reality, this approximation is fine only as long we do not care about the I_D-V_{DS} dependence seen in Figure 2-21. Therefore, in order to get a quantitative handle on the MOSFET's intrinsic voltage gain, we must further investigate the impact of the physics at the drain side.

The simplest possible way to proceed is to factor ΔL into the existing derivation of Section 2-1. Instead of integrating Eq. (2.5) over the length L, we use $L - \Delta L$ as the upper limit of the integral. Recall that $L - \Delta L$ is the actual location where the mobile charge vanishes, i.e., $Q_n(y) = 0$. With this change we obtain the following expression for the saturation region.

$$I_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L-\Delta L}(V_{GS}-V_{Tn})^2 \tag{2.41}$$

Now, provided that ΔL is still small (but not negligible) relative to L, we can apply the following first order approximation

$$\frac{1}{L-\Delta L} \approx \frac{1}{L}\left(1+\frac{\Delta L}{L}\right) \tag{2.42}$$

Next, note that ΔL must be a function of the drain voltage, since the depletion region widens with increasing reverse bias. This effect is commonly called **channel length modulation**. Unfortunately, an exact calculation of ΔL as a function of the terminal voltages involves solving the two-dimensional Poisson equation and leads to complex expressions. For simplicity, we assume that the fractional change in channel length is proportional to the drain voltage

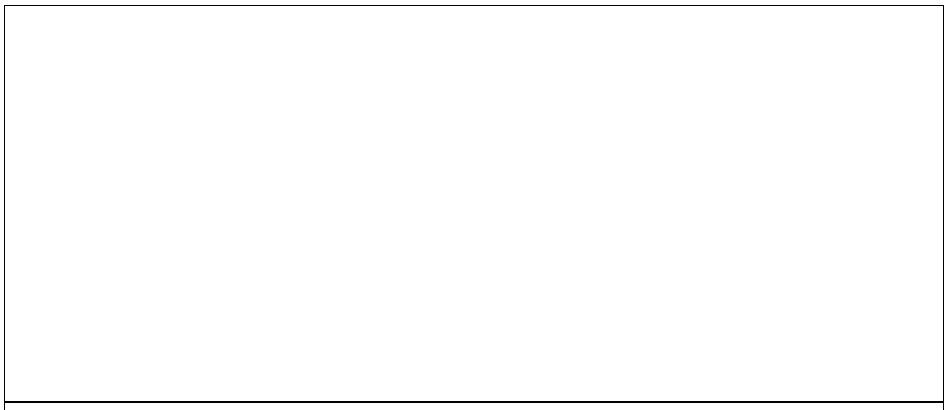


Figure 2-22: ((a) Intrinsic voltage gain stage. (b) Load line plot.

$$\frac{\Delta L}{L} = \lambda_n V_{DS} \tag{2.43}$$

where λ_n is the **channel length modulation parameter**. Device measurements and simulations indicate that λ_n approximately varies with the inverse of the channel length. For the MOSFETs in this module we will use

$$\lambda_n = \frac{0.1\mu m V^{-1}}{L} \tag{2.44}$$

where L is in μm. Finally, we substitute Eq. (2.42) and Eq. (2.43) into Eq. (2.41) and find a very useful approximation to the drain current in saturation, often called the **λ-model**

$$I_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{GS}-V_{Tn})^2(1+\lambda_n V_{DS}) \tag{2.45}$$

where $V_{DS} \geq V_{DSsat} = V_{GS}-V_{Tn}$. This model has proven useful and sufficiently accurate for basic hand calculations, even though the physics related to the I_{Dsat}-V_{DS} dependence are in reality much more complex than discussed above. As long as λ_n is determined from measurements or accurate physical analysis, the model properly approximates a typical MOSFET's I-V characteristic to first order. Higher order models are typically not used in hand analysis, but find their use in advanced computer simulation models.

We now wish to incorporate the channel length modulation effect into the small-signal model of the MOSFET. An important new feature that must be considered in this task is that the drain current of Eq. (2.45) now depends on two voltages, namely V_{DS} and V_{GS}. A common and appropriate way of handling this situation for small-signal modeling is to approximate the incremental drain current around the operating point as the total differential (as frequently used in error analysis) due to both variables, i.e.

$$i_d = \left.\frac{\partial i_D}{\partial v_{GS}}\right|_Q \cdot v_{gs} + \left.\frac{\partial i_D}{\partial v_{DS}}\right|_Q \cdot v_{ds} \tag{2.46}$$

The above expression essentially treats v_{DS} as a constant when evaluating the derivative of i_D with respect to v_{GS}. Similarly, v_{GS} is assumed constant in the differentiation with respect to v_{DS}. This use of partial differentiation is justified and reasonably accurate as long as at least one of the following two conditions is met

1. The excursion in the variable that is treated as a constant can be approximated as infinitesimally small and therefore negligible.
2. The excursion in the variable that is approximated as a constant is considerable, but nonetheless does not affect the derivative with respect to the second variable.

In the context of a common-source amplifier, for instance, the first condition applies to v_{GS}. Just as in the derivation of the simple small-signal model without V_{DS} dependence, we can argue that changes in v_{GS} are suitably modeled as "small" (relative to V_{OV}). The same condition cannot be applied to v_{DS} in general. Often times the output voltage, and therefore v_{DS}, see large excursions in amplifier circuits. In order for Eq. (2.46) to be reasonably accurate, we must require the second condition, i.e., the derivative of i_D w.r.t. v_{GS} must not strongly depend on drain-source voltage. By inspection of Eq. (2.45), we see that this condition is met as long as λ_n is small, which is typically the case for MOSFETs intended for use in amplifier stages.

To continue with our analysis, we re-write Eq. (2.46) as

$$i_d = g_m v_{gs} + g_o v_{ds} \tag{2.47}$$

where

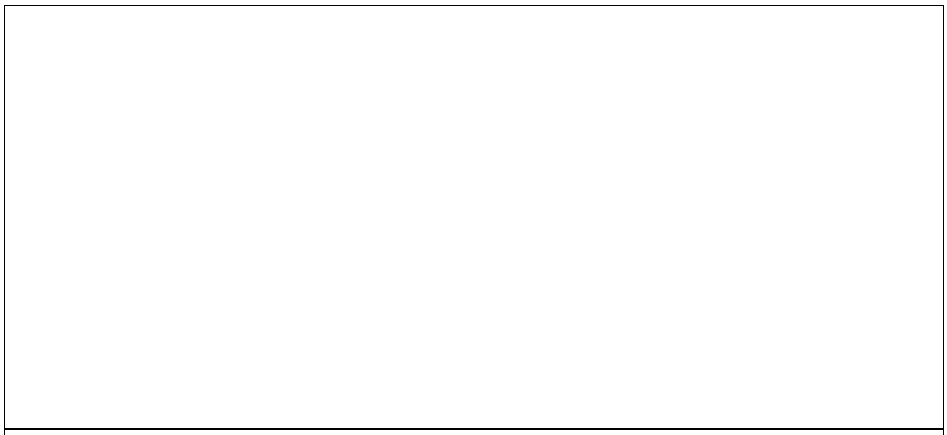


Figure 2-23: (a) Graphical interpretation of g_o. (b) MOSFET small-signal model for the saturation region with output conductance (g_o).

$$g_m = \left.\frac{\partial i_D}{\partial v_{GS}}\right|_Q = \mu_n C_{ox}\frac{W}{L}(V_{GS}-V_{Tn})(1+\lambda_n V_{DS}) = \frac{2I_D}{V_{GS}-V_{Tn}} = \frac{2I_D}{V_{OV}} \tag{2.48}$$

and

$$g_o = \left.\frac{\partial i_D}{\partial v_{DS}}\right|_Q = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{GS}-V_{Tn})^2\lambda_n = \frac{\lambda_n I_D}{1+\lambda_n V_{DS}} \cong \lambda_n I_D \tag{2.49}$$

where the approximate end result assumes $\lambda_n V_{DS} << 1$, a condition that is often satisfied for long channels and moderate V_{DS}. For instance, assuming $L = 2\,\mu m$ and $V_{DS} = 2V$ gives $\lambda_n V_{DS} = 0.1$ which is much less than one.

In the above expressions, g_m is the transconductance of the MOSFET (as defined previously) and g_o is called the **output conductance**. The inverse of g_o is called the **output resistance**, $r_o = g_o^{-1}$. Graphically, the output conductance corresponds to the slope of the transistor's drain characteristic at the operating point, which is the derivative of i_D with respect to v_{DS}, while keeping the gate-source voltage constant [see Figure 2-23(a)]. In the small-signal model of the MOSFET, the output conductance can be included as shown in Figure 2-23(b). This representation follows directly from Eq. (2.47), which represents Kirchhoff's Current Law equation for the drain node of the transistor.

Just as with the simple g_m-only small-signal model of Figure 2-12, the main idea for the usage of the extended model with g_o is to use the small-signal equivalent circuit of Figure 2-23(b) in a larger circuit. We will illustrate this using two examples of interest: the intrinsic voltage gain stage of Figure 2-22(a) and the common-source amplifier of Figure 2-19(a).

2-3-2 Common-Source Voltage Amplifier Analysis Using the λ-Model

For the intrinsic voltage gain stage, the small-signal model of Figure 2-23(b) corresponds directly to the small-signal model for the entire circuit with $v_{in} = v_{gs}$ and $v_{out} = v_{ds}$. Therefore, the voltage gain of the intrinsic gain stage is given by

$$A_v = -\frac{g_m}{g_o} = -g_m r_o \cong -\frac{2I_D}{V_{OV}}\cdot\frac{1}{\lambda_n I_D} = -\frac{2}{\lambda_n V_{OV}} \tag{2.50}$$

From this expression, we can see that the voltage gain can be increased by increasing L, which will decrease λ. Alternatively, the voltage gain can be increased by reducing V_{OV}, which corresponds to reducing the drain current I_D for a fixed aspect ration W/L. In this context, note that for $V_{OV} \rightarrow 0$, Eq. (2.50) predicts infinite voltage gain. This non-physical outcome stems from the same issue already discussed in Section 2-2-7: for $V_{OV} \rightarrow 0$, g_m approaches infinity for a fixed current in our simplistic square-law I-V model. As argued before, the usable range for V_{OV} must therefore be lower bounded as specified in Eq. (2.35). Assuming $V_{OV} = V_{OVmin} = 150$ mV and $L = 1\,\mu m$, the intrinsic voltage gain of a MOSFET described by the parameters used in this module is approximately $2/(0.1\cdot 0.15) = 133$.

Let us now consider the common-source amplifier of Figure 2-19(a). Including the finite output conductance from the λ-model, the small-signal model is modified as shown in Figure 2-24 and the voltage gain expression becomes

$$A_v = -g_m\left(\frac{1}{r_o}+\frac{1}{R_D}\right)^{-1} = -g_m R_{out} \tag{2.51}$$

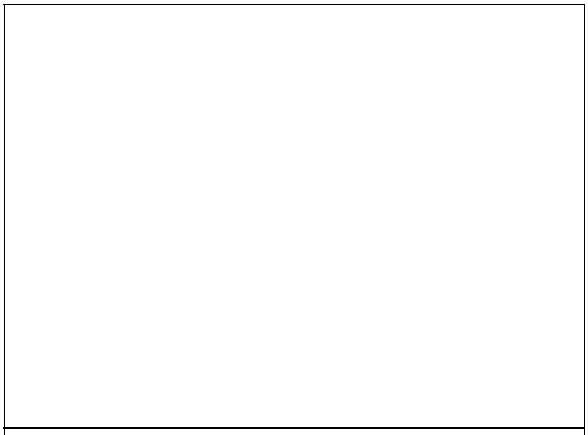


Figure 2-24: Small-signal model for the circuit of Figure 2-19(a), with finite output conductance.

For $R_D \to \infty$, the small signal voltage gain approaches the intrinsic voltage gain as given by Eq. (2.50). For $R_D << r_o$, we can approximate $A_v \cong -g_m R_D$. More generally, without even knowing the exact values of r_o and R_D, we can argue that as long as the desired gain is much less (in magnitude) than the intrinsic voltage gain, r_o can be neglected in the voltage gain calculation. To see this, we can re-write Eq. (2.51) as

$$\begin{aligned} \frac{1}{|A_v|} &= \frac{1}{g_m r_o}+\frac{1}{g_m R_D} \\ g_m R_D &= \frac{|A_v|}{1-\dfrac{|A_v|}{g_m r_o}} \cong |A_v| \quad \text{for} \quad |A_v| \ll g_m r_o \end{aligned} \tag{2.52}$$

Example 2-6: Analysis of a CS Amplifier Using the λ-Model

Consider the CS voltage amplifier of Figure 2-19(a) with W = 80 µm, L = 2 µm and R_D = 50 kΩ. The gate is biased such that V_{OV} = 500 mV and V_B is set to 2 V, which is also the desired output operating point V_{OUT}. Compute the required bias current I_B and the small-signal voltage gain of the circuit using the λ-model. Repeat the small-signal voltage gain calculation for R_D = 5 kΩ.

SOLUTION

The bias current is found using

$$\begin{aligned} I_B = I_D &= \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{GS}-V_{Tn})^2(1+\lambda_n V_{DS}) \\ &= \frac{1}{2}\cdot 50\frac{\mu A}{V^2}\cdot\frac{80}{2}(0.5V)^2(1+0.05\cdot 2) \\ &= 275\mu A \end{aligned}$$

The corresponding transconductance and output resistance at the operating point are

$$g_m = \frac{2I_D}{V_{OV}} = \frac{2\cdot 275\mu A}{0.5V} = 1.1mS$$

$$r_o = \frac{1+\lambda_n V_{DS}}{\lambda_n I_D} = \frac{1+0.05\cdot 2}{0.05V^{-1}\cdot 275\mu A} = 80k\Omega$$

According to Eq. (2.51), the small signal voltage gain for R_D = 50 kΩ is given by

$$A_v = -1.1mS\left(\frac{1}{80k\Omega}+\frac{1}{50k\Omega}\right)^{-1} = -33.9$$

for R_D = 5 kΩ we find

$$A_v = -1.1mS\left(\frac{1}{80k\Omega}+\frac{1}{5k\Omega}\right)^{-1} = -5.18$$

Since the voltage gain for R_D = 5 kΩ is much less than the intrinsic voltage gain of the transistor ($g_m r_o$ = 88), it is OK to neglect r_o in this calculation. We can simply compute

$$A_v = -1.1mS\cdot 5k\Omega = -5.5$$

This result differs only by about 5.9% from the accurate calculation.

Another opportunity for useful engineering approximations in the application of the λ-model lies in the operating point calculation. We will illustrate this point through the example below.

Example 2-7: Approximate Operating Point Calculations

Re-calculate I_D, g_m and r_o by approximating $\lambda V_{DS} \cong 0$ in the large-signal bias point calculations. Also re-calculate A_v for R_D = 50 kΩ using this approximation. Compare the results to the values obtained in Example 2-6.

SOLUTION

For $\lambda V_{DS} \cong 0$, the bias current is estimated as

$$I_B = I_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{GS}-V_{Tn})^2$$
$$= \frac{1}{2}\cdot 50\frac{\mu A}{V^2}\cdot\frac{80}{2}(0.5V)^2$$
$$= 250\mu A$$

The corresponding transconductance and output resistance at the operating point are

$$g_m = \frac{2I_D}{V_{OV}} = \frac{2\cdot 250\mu A}{0.5V} = 1mS$$
$$r_o = \frac{1}{\lambda_n I_D} = \frac{1}{0.05V^{-1}\cdot 250\mu A} = 80k\Omega$$

and the voltage gain becomes

$$A_v = -1mS\left(\frac{1}{80k\Omega}+\frac{1}{50k\Omega}\right)^{-1} = -30.8$$

Relative to the accurate calculation from Example 3-6 (A_v = –33.9), this result is in error by only about 9.1%.

There are several reasons why it is commonly acceptable to neglect the λV_{DS} term in bias point hand calculations. First, without this approximation, the calculations can become cumbersome and lead to transcendental equations that are tedious and undesirable to solve in light of only a moderate percent-improvement in the obtained accuracy. If a more accurate result is desired, it can often be obtained more easily from computer simulations, which often follow a hand calculation in practice, anyway. Lastly, one can argue that any circuit in which the operating point parameters strongly depend on λ may be impractical in the first place. The accuracy of the λ-model as far as absolute I-V values are concerned can only be approximate due to its empirical nature. For high accuracy analysis, much more complex models (such as the one described in Reference 7) must be used, carefully calibrated with physical measurements and subsequently evaluated in computer simulations. For the purpose of developing an introductory feel for circuits, however, the λ-model is still the most appropriate, mainly due its simplicity.

Table 2-3 summarizes the technology parameters introduced in this chapter.

2-4 Two-Port Model for the Common-Source Voltage Amplifier

The circuit shown in Figure 2-24 corresponds to a transconductance amplifier model (with voltage output) since we modeled the MOSFET as a voltage controlled current source, which is in line with its physical behavior in the saturation region. Alternatively, and since the circuit is meant to function as a voltage amplifier, we can equivalently model it using a native voltage amplifier two-port as shown in Figure 2-25. The reader can prove that for this model the open-circuit voltage gain is given by $A_v = -g_m R_{out}$, where R_{out} corresponds to the parallel connection of R_D and r_o. The native voltage amplifier model is sometimes preferred since it more directly expresses the intended function we assumed in this chapter. However, as we shall see in Chapter 3, the voltage amplifier two-port is no longer a convenient representation for the CS voltage amplifier when device capacitances are included.

Table 2-3: Standard technology parameters for the λ-model.

Parameter	n-channel MOSFET	p-channel MOSFET
Threshold voltage	V_{Tn} = 0.5 V	V_{Tp} = –0.5 V
Transconductance parameter	$\mu_n C_{ox}$ = 50 μA/V^2	$\mu_p C_{ox}$ = 25 μA/V^2
Channel length modulation parameter	λ_n = 0.1 V^{-1}/L (L in μm)	λ_p = 0.1 V^{-1}/L (L in μm)

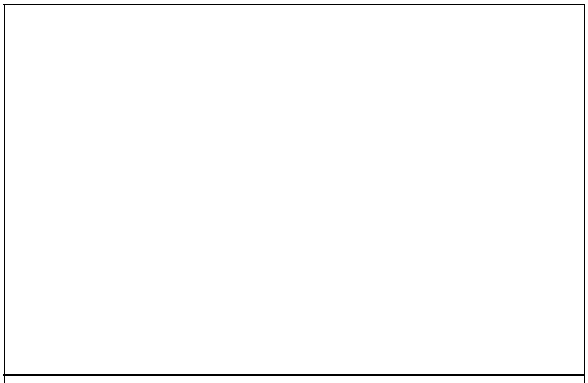


Figure 2-25: Equivalent voltage amplifier-based model for the common-source amplifier circuit of Figure 2-24.

Figure 2-26 shows the small-signal common-source voltage amplifier model together with an input transducer and load resistance. Since the input resistance of the CS amplifier is infinite, no resistive division takes place at the input port and $v_{in} = v_s$. Thus, the overall voltage gain is computed as

$$A'_v = \frac{v_{out}}{v_s} = A_v\cdot\left(\frac{R_L}{R_L+R_{out}}\right) \qquad (2.53)$$

The same expression can be found from the circuit of Figure 2-24 with R_L included across the output port. The reader is invited to prove this.

As a final note, we should emphasize that adding a load resistance may have implications on the bias point of the circuit. For example, if the load resistor carries a DC current, this current will affect the quiescent point output voltage of the amplifier. In this case, the load resistance must also be connected to the circuit in the operating point analysis where the small-signal parameters such as r_o are computed. Problem 2-17 looks at an example.

Figure 2-26: Common-source amplifier model with transducer and load resistance.

Summary

In this chapter we reviewed the basic I-V characteristics of a MOSFET and employed this device to construct examples of common-source voltage amplifiers. In deriving a model for the MOSFET, we concentrated on first-order effects that define the transistor's operation. The nonlinear nature of even the simplest device model dictates the use of small-signal approximations to enable analyses with manageable complexity. The presented methodology begins by finding the operating point of the transistor. Next, the small-signal equivalent is used to construct a linear small-signal model for further analysis.

We studied the basic common-source voltage amplifier with drain resistance and found that the voltage gain in this circuit is directly proportional to the voltage drop across the resistor, which imposes practical limits on the achievable voltage gain. A modified circuit based on an auxiliary bias current source was then analyzed as an alternative and used as a motivation to incorporate the effect of channel length modulation using the λ-model.

The most important concepts that you should have mastered are:

- Determining MOSFET drain currents in all regions of operation; determining the regions of operation based on the transistor's terminal voltages.
- Constructing transfer characteristics and load line plots for common-source stages for various drain bias configurations.
- Calculating the operating point of a CS stage and the MOSFET's small-signal parameters g_m and r_o.
- Drawing the small-signal model of a CS stage and calculating its voltage gain.

References

1. S. Bhave, R. T. Howe and C. G. Sodini, *TBD*, 2011.

2. P. R. Gray, P. J. Hurst, S. H. Lewis, and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 5th Edition, Wiley, 2008.

3. D. O. Pederson and K. Mayaram, *Analog Integrated Circuits for Communication: Principles, Simulation and Design*, 2nd Edition, Springer, 2008.

4. Y. Tsividis, *Operation and Modeling of the MOS Transistor*, 2nd Edition, McGraw-Hill, New York, 1999.

5. P. G. A. Jespers, *The g_m/I_D Methodology, a Sizing Tool for Low-Voltage Analog CMOS Circuits*, Springer, 2010.

6. M. J. M. Pelgrom, et al., "Matching properties of MOS transistors," *IEEE J. of Solid-State Circuits*, pp. 1433-1439, May 1989.

7. G. Gildenblat, et al., "PSP: An Advanced Surface-Potential-Based MOSFET Model for Circuit Simulation," *IEEE Trans. Electron Devices*, pp. 1979-1993, Sept. 2006.

Problems

Unless otherwise stated, use the standard model parameters specified in Table 2-3 for the problems given below. Consider only first-order MOSFET behavior and include channel length modulation (as well as any other second-order effects) only where explicitly stated.

P2.1 An n-channel transistor biased in saturation with W/L = 10 carries a drain current of 200 μA when V_{GS} = 1.5 V is applied. With V_{GS} = 1 V, the current drops to 50 μA. Determine V_{Tn} and μC_{ox} of this transistor.

P2.2 Show that two MOS transistors in series with channel lengths L_1 and L_2 and identical channel widths can be modeled as one equivalent MOS transistor with length $L_1 + L_2$ (see Figure P2-2). Assume that M_1 and M_2 have identical parameters except for their channel lengths. Hint: there are (at least) two ways to solve this problem. One is through extensive algebra;

the other is though physical insight and arguments based on the MOSFET cross-section.

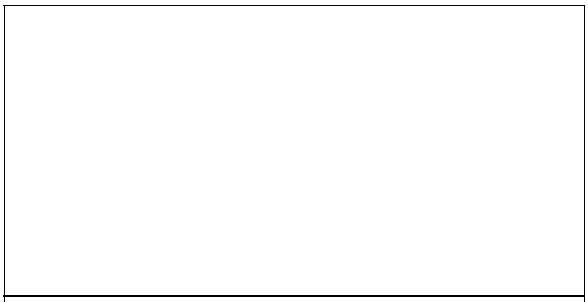


Figure P2-2.

P2.3 Derive an analytical expression for the input voltage that corresponds to the transition point (point C) between the saturation and triode regions in Figure 2-7(b).

P2.4 In Example 2-4, we calculated the most negative input excursion that the circuit in Figure 2-9 can handle before the output is clipped to the supply voltage. In this problem, calculate the most positive input excursion that can be applied before the MOSFET enters the triode region. Assume the same parameters as in Example 2.4: $V_{DD} = 5$ V, $R_D = 10$ kΩ, $W/L = 10$ and V_{IN} is adjusted to 1.5 V, so that $V_{OUT} = 2.5$ V at the circuit's operating point.

P2.5 For the circuit shown in Figure P2-5, sketch v_{OUT} as a function of v_{IN}. Assume that v_{IN} varies from 0 to 5 V. There is no need to carry out any detailed calculations; simply draw a qualitative graph and mark pertinent asymptotes and breakpoints (such as changes in the MOSFETs region of operation).

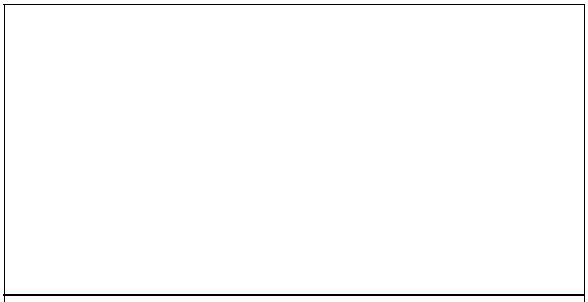


Figure P2-5.

P2.6 Derive the small-signal voltage gain given by Eq. (2.27) through direct differentiation of Eq. (2.17) [i.e., apply Eq. (2.25)].

P2.7 A field effect transistor built using a new (fictitious) material behaves "almost" exactly like a conventional MOSFET in silicon technology. The large signal I-V characteristic (in the saturation region) is given by

$$I_D = \frac{1}{2}\mu C_{ox}\frac{W}{L}(V_{GS} - V_{Tn})^{2.5}$$

Assuming that $\mu C_{ox}W/L = 100$ A/V$^{2.5}$ and $I_D = 1$ μA, compute the device's transconductance.

P2.8 The ratio of the small-signal drain current excursion and quiescent point drain current (i_d/I_D) is often called the drain modulation index. Show that for a MOSFET, i_d/I_D is twice as large as the relative excursion in the gate-source voltage, v_{gs}/V_{OV}.

P2.9 Consider the CS amplifier shown Figure 2-9(a). Calculate the small-signal voltage gain assuming $V_{IN} = 1.5$V, $W/L = 20$, $R_D = 5$ kΩ, and $V_{DD} = 5$ V.

P2.10 Consider the p-channel CS amplifier shown in Figure 2-14(a). Assuming $W/L = 20$, $R_D = 5$ kΩ, and $V_{DD} = 5$ V, calculate the required quiescent point input voltage so that $V_{OUT} = 2.5$V. What is the small-signal gain of the circuit?

P2.11 Repeat problem 2.9 with $R_D = 10$ kΩ. With this value, the MOSFET operates in the triode region. Compute the small-signal voltage gain by writing the large-signal relationship between the input and output for the triode region and subsequent differentiation at the operating point.

P2.12 Consider the cascade connection of two CS amplifiers as shown in Figure P2-12.

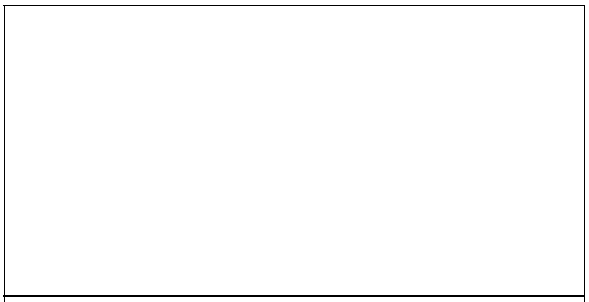


Figure P2-12.

(a) Draw the small-signal equivalent model for this circuit.

(b) Calculate the small-signal voltage gains v_{o1}/v_i and v_{o2}/v_i. Assume that $V_B = V_{DD}/2$ and that the device sizes are chosen such that the bias points of nodes v_{O1} and v_{O2} are also

exactly at $V_{DD}/2$. You are also given $V_{Tn} = |V_{Tp}| = 0.5$ V, and $V_{DD} = 2.5$ V. Note that you do not need to know any device geometries to solve this problem.

P2.13 Repeat problem 2.9 and include the effect of channel length modulation in the small-signal voltage gain calculation. neglect channel length modulation in the bias point calculation. Quantify the percent difference in the calculated small-signal voltage gain compared to problem 2.9. Assume that the channel length of the MOSFET is 2 µm.

P2.14 For the p-channel common-source amplifier shown in Figure 2-14(a)

(a) Given $W/L = 12\mu m/2\mu m$ and $R_D = 10$ kΩ, calculate V_{BIAS} such that V_{OUT} is 2.5 V. Neglect channel length modulation in this calculation.

(b) What is the small-signal voltage gain, $A_v = v_{out}/v_{in}$? Include the effect of channel-length modulation in your calculation.

(c) To increase the voltage gain, you increase R_D to 100 kΩ. Calculate the new small-signal voltage gain, A_v. You must re-bias the circuit so that $V_{OUT} = 2.5$ V.

(d) We could also try to increase the voltage gain of the initial circuit by increasing W/L rather than R_D. Calculate the new A_v if $W/L = 120\mu m/2\mu m$ and $R_D = 10$ kΩ. Be sure to re-bias the circuit so that $V_{OUT} = 2.5$ V.

P2.15 Calculate the small-signal gain of the intrinsic gain stage shown in Figure 2-22(a), assuming $W = 10$ µm and using the parameters given below. In each case, also calculate the gate overdrive voltage (V_{OV}), the transconductance (g_m) and the output resistance (r_o). Assume that the circuit is biased such that the MOSFET operates in the saturation region.

(a) $I_D = 100$ µA, $L = 2$ µm

(b) $I_D = 50$ µA, $L = 2$µm

(c) $I_D = 100$ µA, $L = 4$ µm

P2.16 Consider the CS amplifier of Figure 2-19(a) with the following parameters: $V_B = 2.5$ V, $I_B = 500$ µA, $W = 20$ µm, $L = 1$ µm, and $R_D = 5$ kΩ.

(a) Calculate the exact value of V_{IN} required such that $I_D = I_B$, and $V_{OUT} = V_B$. Include the effect of channel length modulation in your calculation.

(b) Using the value of V_{IN} found in part (a), re-compute V_{OUT} for the two cases when λ changes by +50% and -50%, respectively. Such discrepancies may be due to variations in the semiconductor process or simply due to uncertainty in the simplistic λ-model.

P2.17 Consider the CS amplifier shown Figure P2-17 with the following parameters: $V_{DD} = 5$ V, $V_{IN} = 1.8$ V, $W = 15\mu m$, $L = 1$ µm, $R_D = 5$ kΩ, and $R_L = 10$ kΩ.

(a) Calculate the output quiescent point voltage and drain current of the circuit, taking the connected load resistance R_L into account. Ignore channel-length modulation in this calculation.

(b) Calculate g_m and r_o using the parameters from part (a). Use the approximate approach from Example 2-7.

(c) Draw a two-port voltage amplifier model for the circuit and calculate the open-circuit voltage gain (A_v) and overall voltage gain (A'_v) with the load resistor connected.

(d) Repeat parts (a) and (b) without considering the connected R_L. Explain the main differences in your answers. Will these differences have an effect on the function of the amplifier? Distinguish between small- and large-signal characteristics of the amplifier.

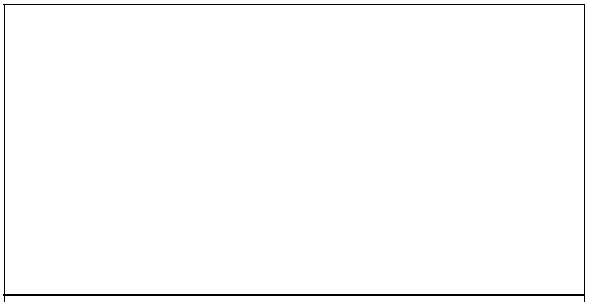


Figure P2-17.

3

CHAPTER

Frequency Response of the Common-Source Voltage Amplifier

In the previous chapter, we have analyzed the common-source stage in terms of its static voltage transfer characteristic and did not consider any dynamic effects in the relationship between the circuit's input and output. The obtained results are therefore applicable only in the limit of slowly varying signals, and further analysis is needed to predict limits in the circuit's operating speed.

In most electronic circuits, the speed of operation is fundamentally limited by the presence of undesired capacitive elements. Therefore, for the purpose of including dynamic effects in the common-source voltage amplifier, we will expand the MOSFET model with its capacitive elements. In the spirit of the just-in-time modeling approach followed in this module, we first consider primary effects related to **intrinsic capacitance**, i.e., capacitance that is unavoidable and required for the operation of a MOSFET. We then refine our analysis to include **extrinsic capacitances**. Extrinsic capacitances are not required for the operation of a MOSFET, but nonetheless exist due to limitations or properties of a certain device structure or manufacturing process.

The analysis and inclusion of device capacitance will follow the small-signal modeling approach used in Chapter 2. That is, even though most MOSFET device capacitances are inherently nonlinear, we will approximate them using linear elements at the MOSFET's operating point. At the various stages of the model development, we consider the dynamics of the amplifier for small-signal, sinusoidal inputs in the steady-state. Specifically, we evaluate the phase and magnitude of the amplifier's output signal to quantify its behavior as a function of frequency.

Even though the small-signal abstraction greatly simplifies the analysis of circuit dynamics, we will find that further simplifications and tools are needed to reason quickly and intuitively about the limiting effects. Therefore, this chapter includes a treatment of the **dominant pole approximation**, the **Miller theorem**, the **Miller approximation**, and the **open-circuit time constant** (OCT) analysis. These techniques are broadly applicable and useful for the analysis of a wide range of circuits, going far beyond the motivational common-source stage example treated in this chapter.

Chapter Objectives

- Review the basic concepts of frequency domain analysis.
- Extend the small-signal MOSFET model with intrinsic and extrinsic device capacitances.
- Derive the sinusoidal steady-state frequency response of the common-source stage at various levels of capacitance modeling and circuit abstraction.
- Review and develop tools and approximation methods that help simplify the frequency response analysis of a circuit: dominant pole approximation, Miller theorem, Miller approximation, and open-circuit time constant analysis.

3-1 Review of Frequency Domain Analysis

In this section, we will review important pre-requisite material using the RC circuit shown in Figure 3-1 as a driving example. Our objective is to gain insight into the circuit's behavior when a sinusoidal signal of a given frequency is applied at its input. Since the circuit consists of linear elements, it follows that the output can only contain a sinusoid at the same frequency that is applied. Therefore, all we need to determine is the amplitude and phase of the output sinusoid. Note that even though we restrict ourselves to sine waves, the analysis results are generally useful since arbitrary periodic signals can be constructed from a sum of sinusoids.

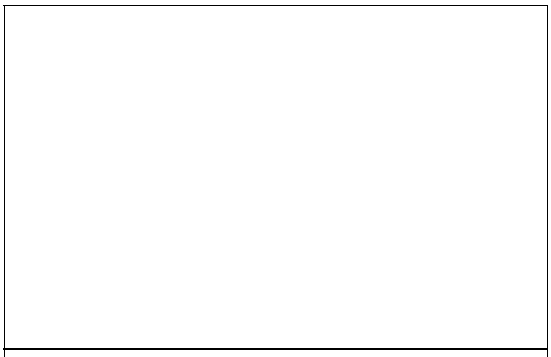


Figure 3-1: RC circuit example.

From first principles, we could approach this problem by applying KCL and KVL, noting that the current flowing through a capacitor is given by $C \cdot dv/dt$. The result of this analysis is a linear differential equation that links $v_{in}(t)$ and $v_{out}(t)$. This equation can be solved for a sinusoidal input, yielding in general two components that make up the output. The first is called the **transient part**; it decays to zero for $t \to \infty$. The second is called the **steady-state** component, and it persists for all t. This latter component is what we are interested in.

A convenient shortcut to obtain the steady-state response is to work with **Laplace transform** models for each circuit element and to determine the transfer functions in the s-domain. Once an s-transfer function is created, the circuit's steady-state response to a sinusoidal input is found by letting $s = j\omega$ and by computing the phase and the magnitude of the output as a function of frequency (ω). The resulting characteristic is called the **frequency response** of the circuit and is usually plotted in the format of a **Bode plot**. The involved variables that capture how the magnitude and phase vary with frequency are called **phase vectors** or **phasors**. In this module, the notation for phasors uses an upper case variable name and lower case subscripts such as V_{in} and I_{out}.

We will now illustrate the flow of such an analysis using the RC circuit example.

Example 3-1: Frequency Response of an RC Circuit

Find the magnitude and phase of the transfer function V_{out}/V_{in} for the RC circuit in Figure 3-1.

SOLUTION

We begin by noting that in the s-domain, the reactance of a capacitor is given by $1/sC$. By applying the voltage divider rule, we can therefore write a transfer function that links v_{out} and v_{in} in the s-domain as follows

$$\frac{v_{out}}{v_{in}} = \frac{\frac{1}{sC}}{\frac{1}{sC} + R} = \frac{1}{1 + sRC}$$

In order to evaluate this transfer function for steady-state sinusoids, we let $s = j\omega$ and obtain

$$\frac{V_{out}}{V_{in}} = \left.\frac{v_{out}}{v_{in}}\right|_{s = j\omega} = \frac{1}{1 + j\omega RC}$$

Following the rules for determining the magnitude and phase of a complex number, we obtain

$$\left|\frac{V_{out}}{V_{in}}\right| = \sqrt{\frac{1}{1 + (\omega RC)^2}}$$

and

$$\angle\frac{V_{out}}{V_{in}} = \tan^{-1}(-\omega RC)$$

From this result, we see that for $\omega RC >> 1$ the sinusoid is attenuated and shifted by –90°, i.e.

$$\left|\frac{V_{out}}{V_{in}}\right| \cong \frac{1}{\omega RC} \qquad \angle\frac{V_{out}}{V_{in}} \cong -90°$$

For $\omega RC << 1$, the sinusoid is passed unattenuated and with no phase shift, i.e.,

$$\left|\frac{V_{out}}{V_{in}}\right| \cong 1 \qquad \angle\frac{V_{out}}{V_{in}} \cong 0°$$

This result makes intuitive sense, since the capacitor carries a larger current for high frequencies, increasingly "shorting" the output port and attenuating the signal. At high frequencies, the

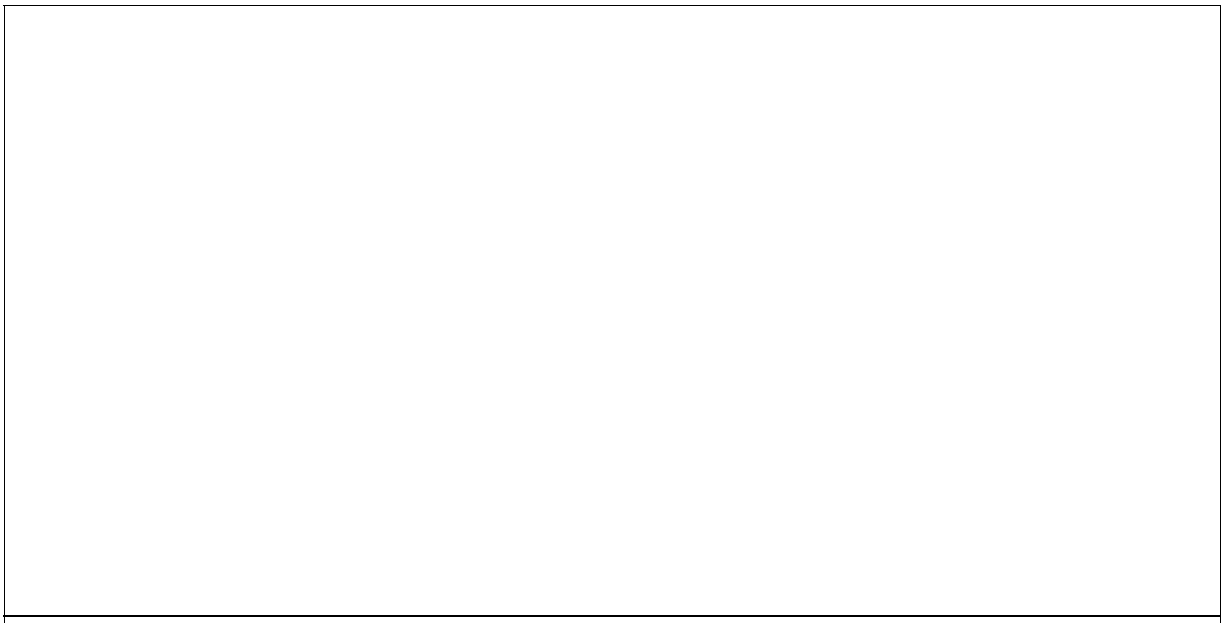


Figure 3-2: Bode plot for the low-pass filter circuit. (a) Log magnitude vs. log frequency. (b) Phase vs. log frequency.

phase approaches –90° due to the signal differentiation that takes place in the capacitor. Its current is given by $C \cdot dv/dt$, and differentiation of a sine wave yields a cosine wave that is –90° shifted in phase.

3-1-1 Bode Plots

In order to gain further insight from the magnitude and phase of a circuit, it is customary to plot the response in the form of a Bode plot, which plots the log of the magnitude versus the log of the frequency, and the phase angle versus the log of the frequency. In this representation, the magnitude and phase can be inspected over many orders of magnitude in frequency.

A Bode plot for the circuit of Example 3-1 is shown in Figure 3-2. A few interesting features can be identified from this plot as follows. First, recall from the analysis of the circuit that for very high frequencies, where $\omega >> 1/RC$, the magnitude of the transfer function becomes inversely proportional to frequency. This is seen in the high-frequency region of the plot where the magnitude decreases by a factor of 10 for every factor of 10 increase in ω. Second, an interesting point in the Bode plot is where $\omega = 1/RC$, also called the **breakpoint frequency**. At the breakpoint, the magnitude is given by

$$\left|\frac{V_{out}}{V_{in}}\right| = \left|\frac{1}{1+j\omega RC}\right| = \left|\frac{1}{1+j}\right| = \frac{1}{\sqrt{2}} = 0.707 \qquad (3.1)$$

and the phase is

$$\angle\frac{V_{out}}{V_{in}} = \tan^{-1}(-\omega RC) = \tan^{-1}(-1) = -45° \qquad (3.2)$$

It is customary to express the logarithmic magnitude scale on a Bode plot with a dimensionless unit called a **decibel** (dB). The magnitude of the ratio of voltages in units of dB is:

$$\text{Ratio of voltages in decibels: } 20\log\left|\frac{V_{out}}{V_{in}}\right| \text{ dB}$$

Therefore, in terms of decibels (indicated on the right-hand y-axis in Figure 3-2) the magnitude falls at –20 dB/decade at high frequencies. Expressed in decibels, the magnitude of the voltage at the breakpoint frequency $\omega = 1/RC$ is $20\log(1/\sqrt{2}) \cong -3$ dB.

The **bandwidth** of a circuit is a measure for the frequency range across which it exhibits only a small amount of attenuation. For a low-pass circuit (such as the RC circuit under investigation), the bandwidth is defined as the frequency for which the magnitude has dropped by a factor of $1/\sqrt{2}$ relative to its value at ω = 0 (DC gain). Since $1/\sqrt{2}$ corresponds to -3 decibels, we refer to this quantity as the 3-dB bandwidth, or symbolically

$$\omega_{3dB} = \frac{1}{RC} \qquad (3.3)$$

As an additional example, we will now look at the frequency response of the RC circuit with an additional resistor added in series with the capacitor C, as shown in Figure 3-3.

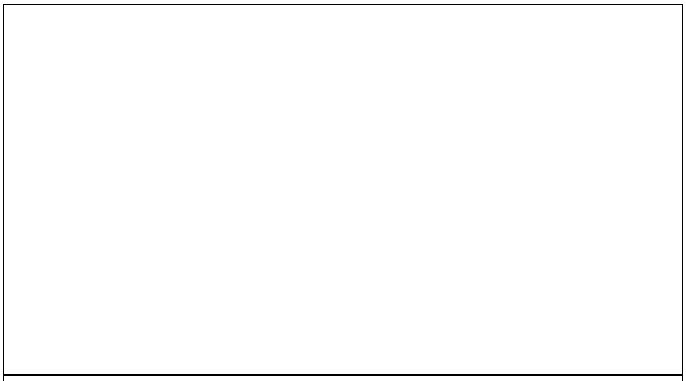


Figure 3-3: RC circuit with series resistor.

Example 3-2: RC Circuit with Additional Resistor

Find the magnitude and phase of the voltage transfer function for the circuit in Figure 3-3 and draw the corresponding Bode plot.

SOLUTION

By applying the voltage divider rule, we find

$$\frac{v_{out}}{v_{in}} = \frac{\frac{1}{sC} + R}{\frac{1}{sC} + R + R} = \frac{1 + sRC}{1 + 2sRC}$$

Next, we let $s = j\omega$ and obtain

$$\frac{V_{out}}{V_{in}} = \frac{v_{out}}{v_{in}}\bigg|_{s = j\omega} = \frac{1 + j\omega RC}{1 + 2j\omega RC}$$

and finally

$$\left|\frac{V_{out}}{V_{in}}\right| = \sqrt{\frac{1 + (\omega RC)^2}{1 + (2\omega RC)^2}}$$

and

$$\angle\frac{V_{out}}{V_{in}} = \tan^{-1}(\omega RC) + \tan^{-1}(-2\omega RC)$$

The Bode plot for these expressions is found in Figure Ex3-2. As we can see, the plot is similar to the previous example in terms of the low-frequency behavior and first breakpoint. There is, however, a second breakpoint beyond which the magnitude approaches a constant value of –6 dB (= 0.5), and the phase begins to return back to 0°. This behavior is intuitively understood by inspection of the circuit. At high frequencies, the capacitor becomes a short, essentially leaving a resistive voltage divider. Since the resistors are of equal value, the voltage attenuation approaches 0.5 at high frequencies. Similarly, the phase returns to 0° because the resistive division at high frequencies has no impact on the signal's phase.

3-1-2 Poles and Zeros

In linear system theory, **poles** and **zeros** are the s-values for which the value of the s-domain transfer function becomes infinity or zero, respectively. Since the behavior of a linear system is fully determined by the location of its poles and zeros, it is desirable to factor the transfer function in the following general format

$$H(s) = K\frac{\left(1 - \frac{s}{z_1}\right)\left(1 - \frac{s}{z_2}\right)\ldots\left(1 - \frac{s}{z_m}\right)}{\left(1 - \frac{s}{p_1}\right)\left(1 - \frac{s}{p_2}\right)\ldots\left(1 - \frac{s}{p_n}\right)} \tag{3.4}$$

where K is a constant DC gain term, p_1, p_2,... p_n are the poles and z_1, z_2,..., z_m are the zeros. For example, the s-domain transfer function of Example 3-2 is given by

$$H(s) = \frac{v_{out}}{v_{in}} = \frac{\left(1 - \frac{s}{z_1}\right)}{\left(1 - \frac{s}{p_1}\right)} \tag{3.5}$$

where

$$z_1 = -\frac{1}{RC} \text{ and } p_1 = -\frac{1}{2RC} \tag{3.6}$$

The reason why p_1 and z_1 are called poles and zeros can be understood from the plot in Figure 3-4, which evaluates Eq. (3.5) using the complex argument $s = \sigma + j\omega$. At $s = p_1$, the magnitude of $H(s)$ becomes infinite, resembling the pole of a tent holding up the 2-dimensional sheet in this representation. Likewise, at $s = z_1$, the magnitude of $H(s)$ becomes zero; this could be viewed as pegs pinning down the tent at this particular location.

Since the steady-state magnitude response of the circuit is obtained by letting $s = j\omega$, it simply corresponds to the bold line marked at the front edge of the plot. In other words, evaluating

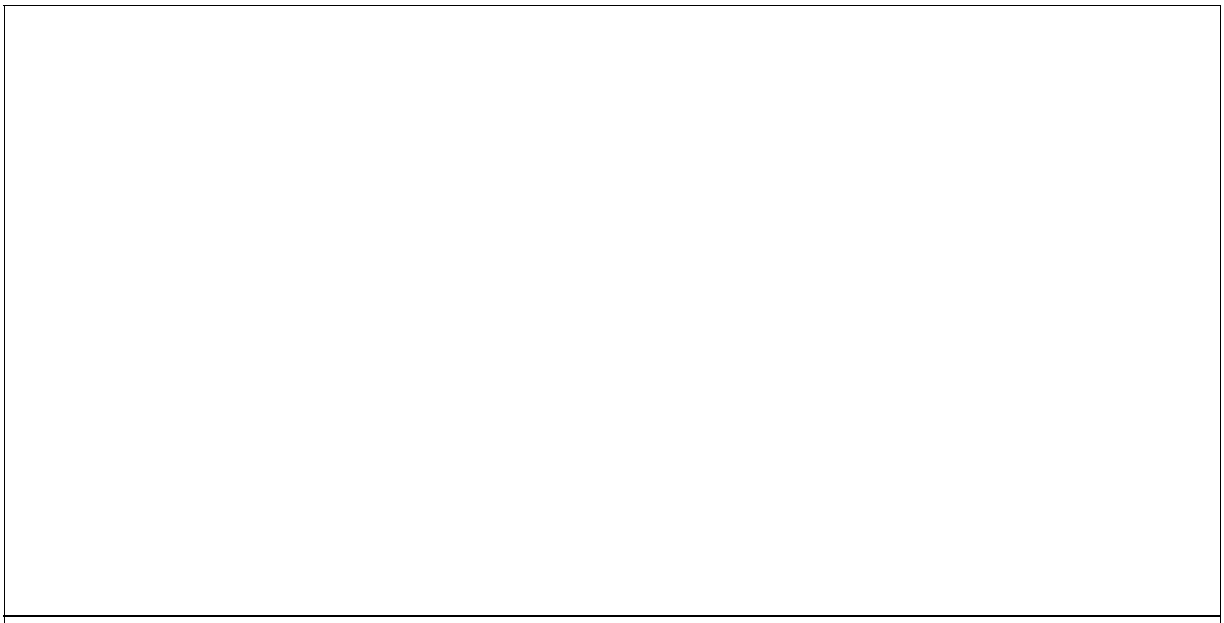


Figure Ex3-2

$H(s)$ for the magnitude response corresponds to "walking" on the sheet of Figure 3-4 along the ω axis.

As we can see from Eq. (3.6), the poles and zeros of the example considered here are (negative) real numbers. For arbitrary ratios of polynomials in s, the poles and zeros as expressed in Eq. (3.4) can be complex numbers. For all circuits considered in this module, however, the poles and zeros will be real. Furthermore, all poles will be negative, as required for a stable system. The zeros encountered in this module can be either positive or negative [as in Eq. (3.6)]. A negative zero is called a **left half plane (LHP) zero**, since it lies on the left side of the s-plane. A positive zero is called a **right half plane (RHP) zero**, since it lies on the right hand side of the s-plane.

When all the poles and zeros of a system are real, it is possible to create a set of rules that allow the construction of a bode plot by inspection. These rules are summarized in the next section.

3-1-3 Bode Plots of Arbitrary System Functions with Real Poles and Zeros

For the case of real negative poles and zeros, and letting $s = j\omega$, Eq. (3.4) becomes

$$H(j\omega) = K\frac{\left(1+j\frac{\omega}{\omega_{z1}}\right)\left(1+j\frac{\omega}{\omega_{z2}}\right)\cdots\left(1+j\frac{\omega}{\omega_{zm}}\right)}{\left(1+j\frac{\omega}{\omega_{p1}}\right)\left(1+j\frac{\omega}{\omega_{p2}}\right)\cdots\left(1+j\frac{\omega}{\omega_{pn}}\right)} \tag{3.7}$$

where ω_{p1}, ω_{p2},... ω_{pn} are the pole frequencies and ω_{z1}, ω_{z2},..., ω_{zm} are the zero frequencies. For instance, in Example 3-2, we have

$$H(j\omega) = \frac{v_{out}}{v_{in}} = \frac{\left(1+j\frac{\omega}{\omega_{z1}}\right)}{\left(1+j\frac{\omega}{\omega_{p1}}\right)} \tag{3.8}$$

where

$$\omega_{z1} = \frac{1}{RC} \text{ and } \omega_{p1} = \frac{1}{2RC} \tag{3.9}$$

To determine the Bode plot from Eq. (3.7), we must assess the effect of each binomial term on the magnitude and phase of the system function. If the frequency is such that $\omega << \omega_{zi}$ or ω_{pi}, then the respective binomial term will have little effect on the magnitude and phase of the system function, as it will simply multiply it by unity. On the other hand, if the frequency is such that $\omega >> \omega_{zi}$ or ω_{pi}, the system function, magnitude, and phase will be altered. To see this, we evaluate the magnitude and phase of a general binomial term for a left half plane pole or zero and $\omega >> \omega_i$

$$\left|1+j\frac{\omega}{\omega_i}\right| = \sqrt{1+\left(\frac{\omega}{\omega_i}\right)^2} \cong \frac{\omega}{\omega_i} \tag{3.10}$$

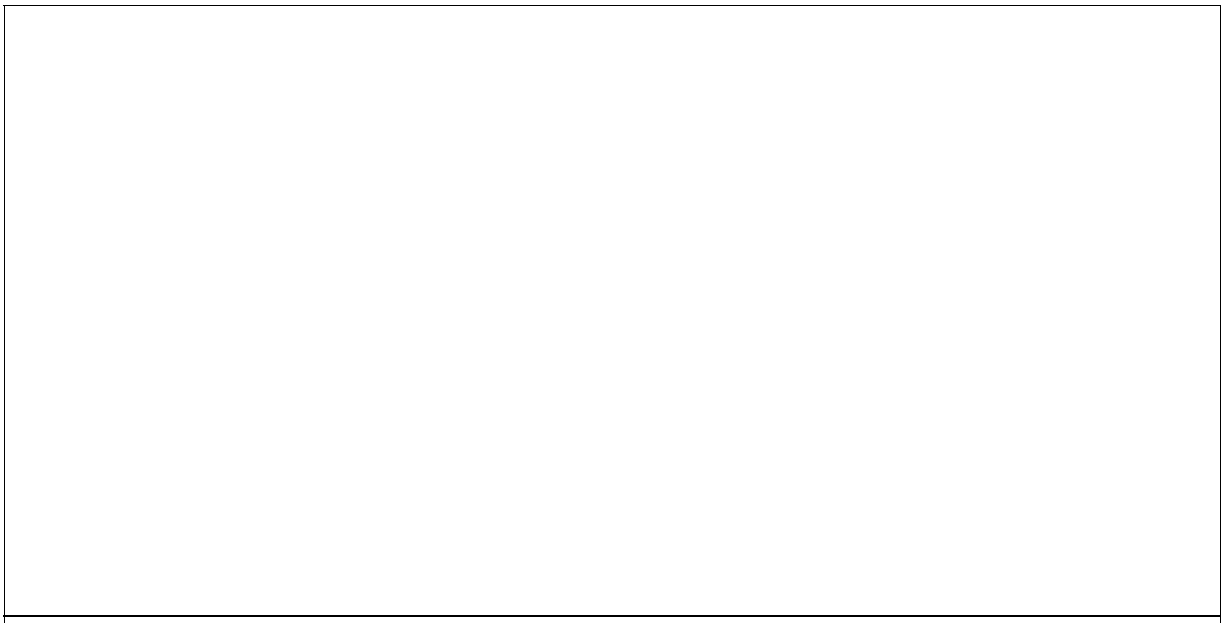


Figure 3-4: 3-D plot of the magnitude of Eq. (3.5), evaluated for $s = \sigma + j\omega$.

$$\angle\left(1 + j\frac{\omega}{\omega_i}\right) = \tan^{-1}\frac{\omega}{\omega_i} \cong 90° \tag{3.11}$$

Therefore, if the binomial term is in the numerator of the generalized system function (corresponding to a LHP zero), the magnitude will be multiplied by ω/ω_i, and a phase angle of 90° will be added to the total phase. If the binomial term is located in the denominator (LHP pole), the magnitude will be multiplied by $1/(\omega/\omega_i)$ and a phase angle of 90° will be subtracted from the total phase. For a RHP zero, it follows that the magnitude will be multiplied by ω/ω_i, and a phase angle of 90° will be subtracted from the total phase.

When $\omega = \omega_i$, the magnitude and phase are

$$\left|1 + j\frac{\omega}{\omega_i}\right| = |1 + j| = \sqrt{2} \tag{3.12}$$

$$\angle(1 + j) = 45^o \tag{3.13}$$

Therefore, if these binomial terms for the breakpoints are located in the numerator, the magnitude of the system function in the numerator is multiplied by $\sqrt{2}$ and a phase of 45° is added to (for a LHP zero) or subtracted from (for a RHP zero) the overall phase. If it is located in the denominator, the magnitude is multiplied by $1/\sqrt{2}$ and a phase of 45° is subtracted from the overall phase of the system function.

Given these results, a Bode plot can be constructed by referring to the following step-by-step procedure.

- Identify all the breakpoint frequencies ω_{pi} and ω_{zi} and list them in increasing order. Apply the following rules, beginning with the lowest breakpoint frequency.
- If the corresponding binomial term appears in the numerator of the system function, the magnitude slope will be increased by 20 dB/decade, when the frequency is greater than the breakpoint frequency.
- If the corresponding binomial term appears in the denominator of the system function, the magnitude of the slope will be reduced by 20 dB/decade when the frequency is greater than the breakpoint frequency.
- To plot the phase, we know that the binomial term will contribute +45° for a LHP zero, and -45° for a RHP zero at $\omega = \omega_i$. If it is in the denominator, it will contribute –45°. We assume that the ±90° phase changes linearly over the interval $0.1\omega_i < \omega < 10\omega_i$.

Example 3-3: Bode Plot Construction

Construct a Bode plot for a system with the following parameters: $K = 100$, $\omega_{p1} = 10$ rad/s, $\omega_{p2} = 100$ krad/s, left half plane zero: $\omega_{z1} = 1$ krad/s, right half plane zero: $\omega_{z2} = 10$ Mrad/s.

SOLUTION

First we note that the DC gain $K = 100 = 40$ dB. Next we recognize that ω_{p1} is the lowest frequency term, creating a change of slope in the magnitude plot toward –20 dB/decade. The phase is 0° at the lowest frequency plotted, –45° at ω_{p1} and

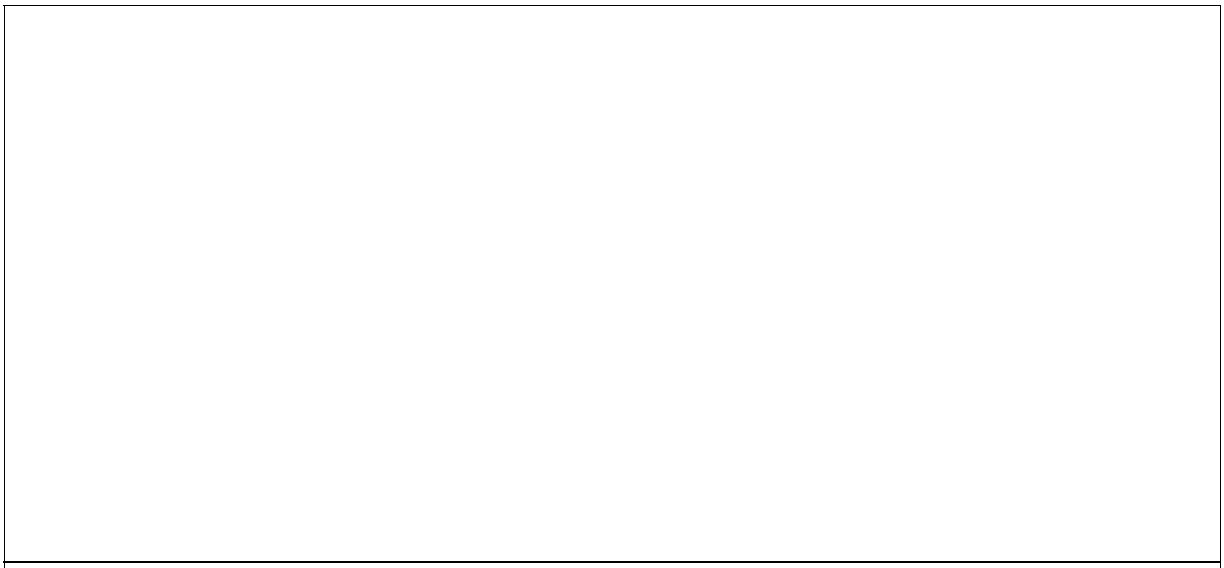


Figure Ex3-3

has reached –90° at approximately $10\omega_{p1}$. Applying the given rules in a similar fashion to the remaining poles and zeros yields the Bode plot shown in Figure Ex3-3.

3-2 Frequency Response of the Common-Source Voltage Amplifier — First Pass Analysis

We now wish to apply the analysis tools reviewed in the previous section to get a handle on the frequency response of the common-source voltage amplifier discussed in Chapter 2. Since the exact frequency behavior of this circuit is quite complex when taking all aspects into account, we partition this discussion into two steps. This section presents the first analysis step and uses the simplest possible model extension for the MOSFET that can be used to take capacitive effects, and thus frequency dependence, into account.

In the context of MOSFET capacitance modeling, it is useful to distinguish between intrinsic and extrinsic capacitances. Here, the term extrinsic refers to capacitances that are not needed to operate a MOSFET, but rather exist due to limitations or properties of a certain device structure or manufacturing process. As we shall see in Section 3-3, stray capacitances between the gate and source/drain terminals are examples of extrinsic capacitors. Intrinsic capacitance is unavoidable and required to operate the device. The oxide capacitance of a MOSFET falls into this category: without a capacitance between the gate and channel, no mobile charges can be induced ($Q = CV$), and the MOSFET would not function. In this section, we will look at frequency dependence effects due to the intrinsic capacitance only, beginning with a derivation of a circuit model that can be used to model this capacitance in the frequency response calculations.

3-2-1 Modeling Intrinsic MOSFET Capacitance

Just as in the derivation of device transconductance and output conductance, the operating point must be considered when calculating small-signal capacitances. We begin by analyzing the intrinsic capacitance of a MOSFET in the triode region, with its cross-section shown in Figure 3-5(a). To first order, the gate and the conductive channel can be viewed as a parallel plate capacitor, resulting in a gate-to-channel capacitance of

$$C_{gc} = WL\frac{\varepsilon_{ox}}{t_{ox}} = WLC_{ox} \tag{3.14}$$

where WL is the capacitor plate area and C_{ox} is the oxide capacitance per unit area.

If the source and drain were connected together, the small signal capacitance from the gate to source/drain would be equal to C_{gc} as given in Eq. (3.14). How can we model the capacitance when source and drain are not connected, i.e., how is the capacitance distributed between the two terminals?

A common first-order approximation is to assign half of C_{gc} to the capacitance between the gate and the source and the remaining half between the gate and the drain. This is schematically illustrated in Figure 3-5(b). A qualitative argument that supports this approximation is that small changes in either the

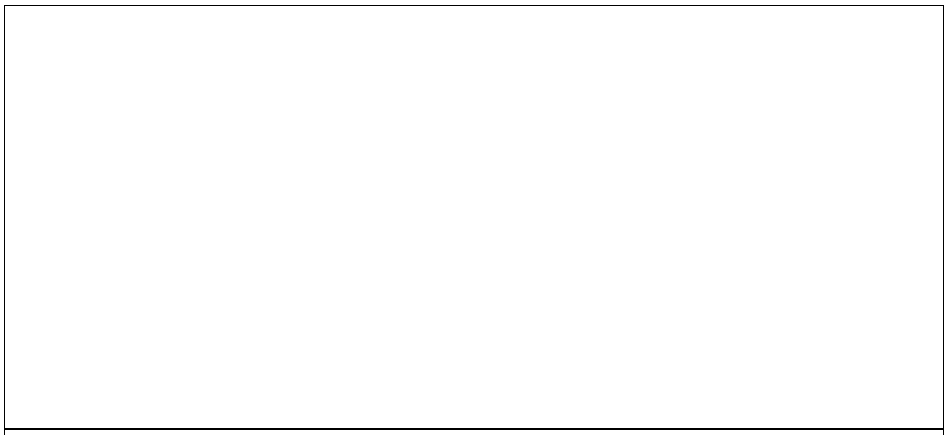


Figure 3-5: (a) MOSFET cross-section showing the intrinsic capacitance between the gate and the channel (C_{gc}). (b) Capacitance model for the triode region.

drain or source voltage must induce the same change in charge at the gate; therefore, the capacitance must be split equally.

A case that is more relevant to the analysis of a common-source stage is the behavior in the saturation region. For this case, we know that the conductive channel does not extend all the way from the source to the drain, but is pinched off at some coordinate $L - \Delta L$.When the channel is pinched-off, the drain potential (to first order) no longer influences the charge under the gate. Therefore, the intrinsic capacitance from the gate to the drain is approximately zero in this region of operation.

In saturation, the channel charge is therefore controlled primarily by the potential between the gate and the source, and a significant capacitance is present between these two terminals. At first glance, one might expect that C_{gs} is equal to C_{gc}. However, this is not quite correct due to the pinch-off effect. Imagine applying a small voltage change to the source terminal. This will change the voltage across the oxide (and charge) near the source, but at the pinch-off point, the voltage across the oxide remains at V_{Tn}. This means that the capacitance in the saturation region must be less than C_{gc}, because the charge does not see a uniform change as in the case of a simple parallel plate capacitor. Further analysis (see Reference 1) reveals that the capacitance between the gate and the source in the saturation region is given by

$$C_{gs} = \frac{2}{3}C_{gc} = \frac{2}{3}WLC_{ox} \tag{3.15}$$

The resulting small-signal MOSFET model that includes this capacitance is shown in Figure 3-6.

3-2-2 Frequency Response with Intrinsic Gate Capacitance

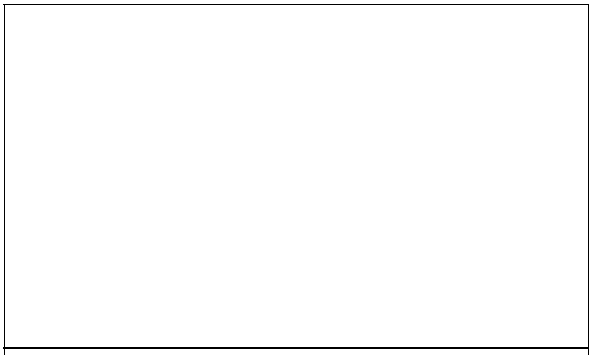


Figure 3-6: MOSFET small-signal model for the saturation region, including the intrinsic gate capacitance.

To analyze the frequency response of the common-source amplifier with the intrinsic gate capacitance, we insert the model of Figure 3-6 into the small-signal circuit model of the amplifier, as shown in Figure 3-7.

Note that if the circuit were driven by an ideal voltage source at its input port (v_{in}), the added capacitance would have no effect on the circuit's operation. The ideal voltage source would provide any current that is needed to charge and discharge the gate capacitance without introducing any frequency dependence. The model in Figure 3-7 therefore considers a more realistic input source with finite resistance (R_s). At this point in the analysis, we purposely do not include any capacitive loading at the output of the amplifier, primarily to keep the first pass analysis simple and transparent.

In order to analyze the frequency response of the circuit in Figure 3-7, we first realize that the overall transfer function can be split into a product of two terms

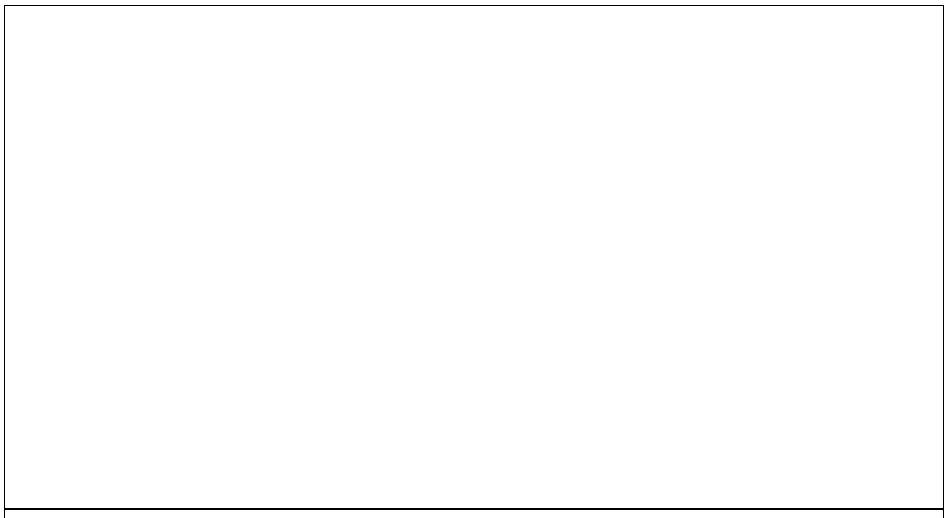


Figure 3-7: (a) Common-source amplifier driven by a transducer with finite source resistance. (b) Small-signal model with the intrinsic gate capacitance.

$$\frac{v_{out}}{v_s} = \frac{v_{out}}{v_{in}} \cdot \frac{v_{in}}{v_s} \tag{3.16}$$

In this expression, the first term on the right-hand side corresponds to the DC voltage gain given in Eq. (2.51), and is equal to $-g_m R_{out}$. The second term can be found by writing the voltage divider expression that relates node v_{in} to v_s

$$\frac{v_{in}}{v_s} = \frac{\frac{1}{sC_{gs}}}{\frac{1}{sC_{gs}} + R_s} = \frac{1}{1 + sR_sC_{gs}} \tag{3.17}$$

With this result, the complete s-domain transfer function from the input source to the output becomes

$$A_v(s) = \frac{v_{out}}{v_s} = \frac{-g_m R_{out}}{1 + sR_sC_{gs}} = \frac{A_{v0}}{1 + sR_sC_{gs}} \tag{3.18}$$

where $A_{v0} = A_v(0)$ is a generalized placeholder for the DC gain of the circuit. From this result, we see that the transfer function has a DC gain corresponding to the result of Chapter 2, and a single pole that is set by the source resistance and the intrinsic gate capacitance. As explained in Section 3-1, we can now evaluate this transfer function for steady-state sinusoids by letting $s = j\omega$. This will allow us to draw a Bode plot and compute the bandwidth of the circuit.

Example 3-4: Common-Source Amplifier Bandwidth Calculation

Calculate the 3-dB bandwidth for the amplifier in Figure 3-7, assuming $W = 20\ \mu m$, $L = 1\ \mu m$, $C_{ox} = 2.3\ fF/\mu m^2$ and $R_s = 50\ k\Omega$. Express the result in units of hertz.

SOLUTION

For the given parameters, the gate-source capacitance is

$$C_{gs} = \frac{2}{3} WLC_{ox} = 30.67 fF$$

The 3-dB bandwidth is

$$\omega_{3dB} = \frac{1}{R_sC_{gs}} = 652.3 MRad/s$$

and therefore.

$$f_{3dB} = \frac{\omega_{3dB}}{2\pi} = 103.8 MHz$$

An important question to ask at this point of the discussion is whether there is anything we can do to maximize the bandwidth of our amplifier. Assuming that we cannot change the source resistance R_s, the only remaining option is to minimize C_{gs}. This can be achieved by choosing a smaller transistor width or length [see Eq. (3.15)]. How will this affect the other performance metrics in the circuit? In the next sub-section, we will show that there exists a direct trade-off in the achievable bandwidth versus supply current for the circuit in consideration.

3-2-3 Trade-off Between Bandwidth and Supply Current

Consider a design problem involving the circuit of Figure 3-7 and assume that the general objective is to maximize the circuit's 3-dB bandwidth while minimizing the transistor's drain current. For this analysis, we assume that R_s, R_{out} and A_{v0} are given through specifications, and that these parameters cannot be varied. This assumption is not atypical in practical circuit design. R_s might be fixed by the physical properties of the input transducer. R_{out} could be set by an output resistance requirement that allows the circuit to interface with subsequent circuit stages, while the DC gain A_{v0} could be determined by application requirements. Furthermore, for simplicity, we neglect channel-length-modulation in this analysis.

In order to study the trade-off between bandwidth and current consumption, we will now write expressions for these quantities that rely on common parameters. For the 3-dB bandwidth, we begin by inserting Eq. (3.15) into Eq. (3.17) and obtain

$$\omega_{3dB} = \frac{3}{2} \cdot \frac{1}{R_s W L C_{ox}} \tag{3.19}$$

By using the following expression to eliminate C_{ox}

$$g_m = \mu_n C_{ox} \frac{W}{L}(V_{GS} - V_{Tn}) = \mu_n C_{ox} \frac{W}{L} V_{OV} \tag{3.20}$$

and subsequently substituting $g_m = |A_{v0}| / R_{out}$, Eq. (3.19) becomes

$$\omega_{3dB} = \frac{3}{2} \cdot \frac{\mu_n}{L^2} \cdot \frac{1}{|A_{v0}|} \cdot \frac{R_{out}}{R_s} \cdot V_{OV} \tag{3.21}$$

The above expression is now in a form that contains only technology parameters, design constraints (R_s, R_{out} and A_{v0}) and the gate overdrive voltage V_{OV} as a single design parameter. From this result, it is clear that in order to maximize bandwidth, we would like to use a technology that offers high mobility and short channels. The mobility is largely determined by material properties, while L is usually bounded by some $L = L_{min}$ that is specific to a certain process technology, for example, 1 µm for the transistors used in this module.

From Eq. (3.21), we also see that we should maximize V_{OV}. However, a potential problem with this is due to the output signal range of the amplifier. As we know from Chapter 2, larger V_{OV} means that V_{DSsat} is also increased, and this means that the transistor enters the triode region at higher v_{OUT}. This could lead to clipping as discussed previously.

An additional, and more fundamental issue relates to the current consumption of the circuit. To see this, we rewrite Eq. (2.31) as

$$I_D = \frac{1}{2} \cdot g_m \cdot V_{OV} \tag{3.22}$$

and substitute $g_m = |A_v| / R_{out}$ to find

$$I_D = \frac{1}{2} \cdot \frac{|A_{v0}|}{R_{out}} \cdot V_{OV} \tag{3.23}$$

This result shows that a larger V_{OV} unfortunately requires a larger bias current for the transistor, and this is highly undesired in many applications, as for instance battery-powered devices.

While the above-observed trade-off was discovered in the context of a particular circuit example, we will see throughout this module that the same trade-off holds for all analog circuits. For a given technology and target specifications, current consumption directly scales with the circuit's 3-dB bandwidth requirements. An alternative, and more general way to capture the fundamental connection between supply current and bandwidth is to inspect the trade-offs that pertain to the MOSFET in isolation of a specific circuit example, as discussed next.

To begin, note that the model in Figure 3-6 comes with "desired" and "undesired" elements and properties. The only aspect of the transistor that we value is its transconductance. The associated intrinsic capacitance and the supply current needed to create the transconductance are undesired. Mathematically, we can identify the following figures of merit that capture the ratios between the desired and undesired quantities, in particular

$$\frac{g_m}{I_D} = \frac{2}{V_{OV}} \propto \frac{1}{\sqrt{I_D}} \tag{3.24}$$

and

$$\frac{g_m}{C_{gs}} = \frac{3}{2} \cdot \frac{\mu_n}{L^2} \cdot V_{OV} \propto \sqrt{I_D} \tag{3.25}$$

The transconductance-to-current ratio, which is sometimes called the transconductance efficiency, deteriorates for larger V_{OV} (and larger I_D) On the other hand, the ratio of transconductance per intrinsic capacitance improves for larger V_{OV} (and larger I_D). This trade-off is graphically illustrated in Figure 3-8. Note that as already pointed out in Section 2-2-7, the proportionality of g_m/I_D to $1/V_{OV}$ extends only down to a certain minimum gate overdrive, defined as V_{OVmin} in this module [see Eq. (2.35)].

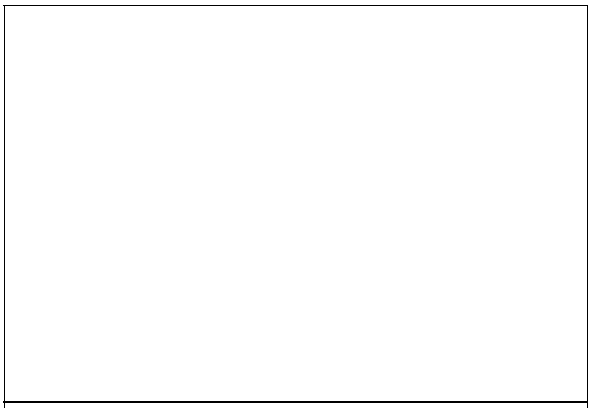


Figure 3-8: Trade-off between g_m / I_D and g_m / C_{gs}.

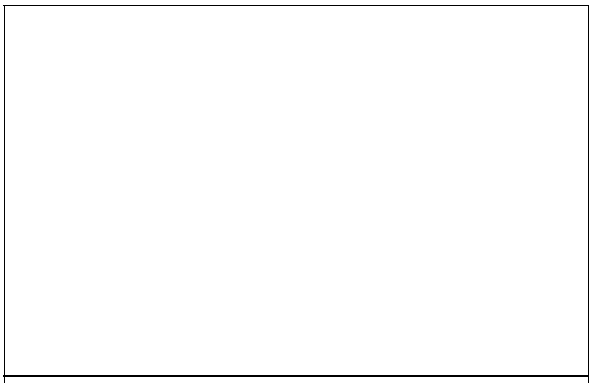


Figure 3-9: Small-signal circuit model for finding the MOSFET's transit frequency.

In essence, the gate overdrive voltage V_{OV} can be considered as a "knob" that lets us adjust the trade-off between the two figures of merit. For a chosen V_{OV} and channel length, g_m / I_D and g_m / C_{gs} are fixed, and these parameters directly affect the speed and current consumption of the overall circuit. The gate overdrive V_{OV} has therefore been recognized by designers as an important parameter that affects most of the trade-offs encountered in the optimization of a given circuit (see Reference 2). We will see examples of this throughout this module.

Interestingly, the product of the two figures of merit in Eq. (3.24) and Eq. (3.25) is given by

$$\frac{g_m}{C_{gs}} \cdot \frac{g_m}{I_D} = 3 \cdot \frac{\mu_n}{L^2} \tag{3.26}$$

From this result, it is clear that for high speed and low current consumption, the best we can hope for is a technology that provides high mobility and short channels. In this context, it is interesting to note that improvements in device engineering and manufacturing processes have provided tremendous improvements in manufacturable channel lengths. Since the 1970's, L_{min} has been improved from 10 µm to approximately 22 nm today; a ~400x reduction!

3-2-4 Transit Frequency

The figure of merit given in Eq. (3.25) is also known as the **transit frequency** of the transistor and quantifies the frequency for which the magnitude of the transistor's current gain drops to unity. To determine the transit frequency, the transistor is operated in the common-source configuration and the input is driven by an ideal current source (see Figure 3-9). The output is short circuited, and the current gain i_{out}/i_{in} is measured.

From the circuit, it follows that

$$i_{out} = g_m \cdot v_{gs} = g_m \cdot \frac{i_{in}}{sC_{gs}} \tag{3.27}$$

Substituting $s = j\omega$ and rearranging yields

$$\frac{I_{out}}{I_{in}} = \frac{g_m}{j\omega C_{gs}} \tag{3.28}$$

The transit frequency then follows by setting

$$\left|\frac{I_{out}}{I_{in}}\right| = 1 = \frac{g_m}{\omega_T C_{gs}} \tag{3.29}$$

and therefore

$$\omega_T = \frac{g_m}{C_{gs}} \propto \sqrt{I_D} \tag{3.30}$$

The above quantity represents the transit frequency in rad/s. The symbol for the corresponding quantity in units of Hertz is $f_T = \omega_T/2\pi$.

The transit frequency gives the designer a feel for the maximum frequency at which a circuit can operate. The bandwidth of most practical circuit configurations is limited to a fraction of ω_T, often about one order of magnitude below the transit frequency.

3-3 Frequency Response of the Common-Source Voltage Amplifier— Second Pass Analysis

We will now extend the results from the previous section to obtain a more accurate understanding of the frequency

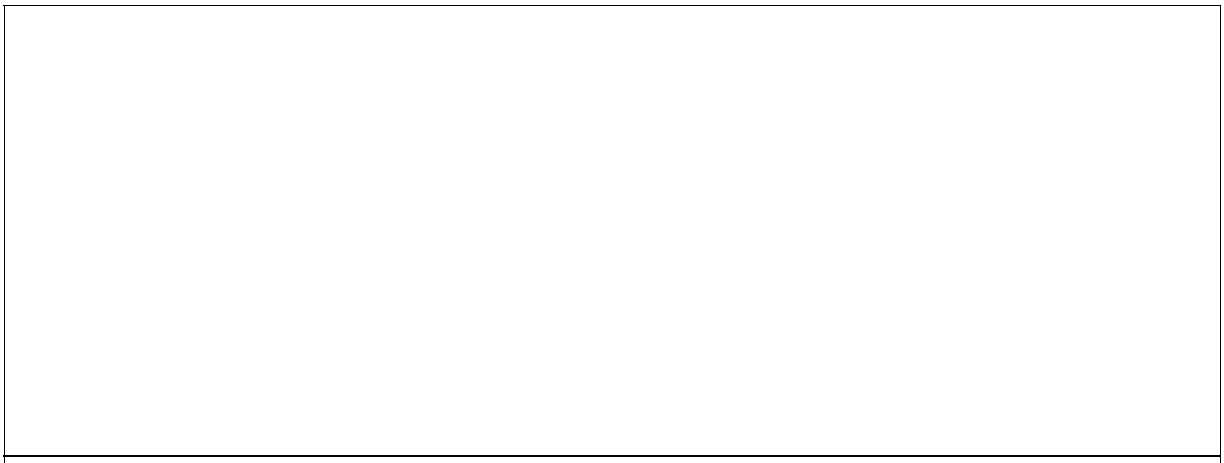


Figure 3-10: MOSFET cross section in saturation showing the overlap and fringe contributions to C_{ov}. The source-bulk and drain-bulk depletion capacitances are also shown qualitatively.

response of a realistic common-source amplifier. To begin, we will extend the MOSFET model to include extrinsic capacitances.

3-3-1 Modeling Extrinsic MOSFET Capacitance

Figure 3-10 shows the cross section of a MOSFET device for further study of its associated capacitive elements. The first component of extrinsic capacitance that we will consider is called **overlap capacitance**; it is due to overlap of the source and drain diffusions and the gate and the contribution of the fringe electric fields from the gate. The overlap capacitance C_{ov} is quantified as a linear capacitance proportional to the gate width, with units of fF/μm.

With overlap capacitance included, the total gate-source capacitance in saturation is the sum of Eq. (3.15) and the overlap capacitance

$$C_{gs} = \frac{2}{3}WLC_{ox} + WC_{ov} \tag{3.31}$$

Since the drain has no influence on the channel charge, the only contribution to the gate-drain capacitance is C_{ov}

$$C_{gd} = WC_{ov} \tag{3.32}$$

In addition to the overlap capacitance, other extrinsic capacitance components are due to the reverse-biased junctions of the MOSFET. The drain-bulk and source-bulk capacitances C_{db} and C_{sb} indicated in Figure 3-10 originate from charge storage in the depletion regions between the drain and source n^+ regions and the p-type bulk. The following expressions can be used to estimate these capacitances (see Reference 1 for a derivation)

$$C_{db} = \frac{C_J \cdot AD}{(1 + V_{DB}/PB)^{MJ}} + \frac{C_{JSW} \cdot PD}{(1 + V_{DB}/PB)^{MJSW}} \tag{3.33}$$

$$C_{sb} = \frac{C_J \cdot AS}{(1 + V_{SB}/PB)^{MJ}} + \frac{C_{JSW} \cdot PS}{(1 + V_{SB}/PB)^{MJSW}} \tag{3.34}$$

In these expressions, V_{DB} and V_{SB} are the reverse bias voltages of the junctions at the operating point. Note that with increasing reverse bias, the values of the junction capacitances decreases. The geometry parameters used in the expressions are related to the layout of the transistor as shown in Figure 3-11.

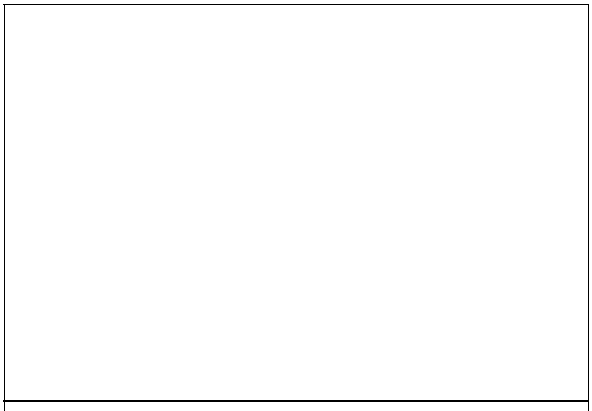


Figure 3-11: Geometry parameters used for the calculation of junction capacitances.

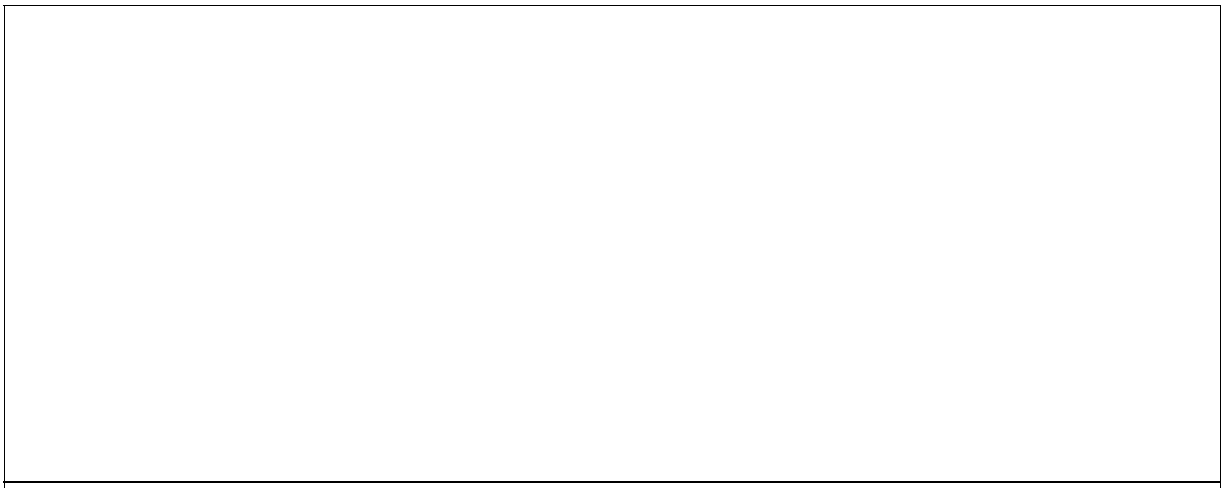


Figure 3-12: Small-signal model for the n-channel MOSFET in saturation, including intrinsic and extrinsic capacitances.

AD=Drain area

AS=Source area

PD=Perimeter of the drain diffusion (not including the edge under the gate)

PS=Perimeter of the source diffusion (not including the edge under the gate)

All other parameters are defined in Table 3-1 along with the technology parameters introduced thus far.

The extrinsic capacitances discussed above are added to the MOSFET small signal model as shown in Figure 3-12. For completeness, this model contains an additional capacitance C_{gb} between gate and bulk. This capacitance is due to the overlap of the polysilicon gate onto the field oxide region that isolates the MOSFET, as well as field lines from the gate terminating in the bulk of the transistor through the channel. This capacitance is usually small and we will neglect it throughout this module.

Lastly, it is important to note that we have only modeled capacitances associated with the MOSFET, that is, the device without interconnections. The parasitic capacitances of the interconnections between MOSFETs can be a limiting factor and must be estimated from the layout and cross section for accurate analysis of a design. Off-chip wiring and package capacitances are also critical for evaluating the performance of any integrated circuit.

Table 3-1: Standard technology parameters for the λ-model, with intrinsic and extrinsic capacitance parameters.

Parameter	**n-channel MOSFET**	**p-channel MOSFET**
Threshold voltage	V_{Tn} = 0.5 V	V_{Tp} = –0.5 V
Transconductance parameter	$\mu_n C_{ox}$= 50 μA/V^2	$\mu_p C_{ox}$= 25 μA/V^2
Channel length modulation parameter	λ_n = 0.1 V^{-1}/L (L in μm)	λ_p = 0.1 V^{-1}/L (L in μm)
Gate oxide capacitance per unit area	C_{ox} = 2.3 fF/μm^2	
Overlap capacitance	C_{ov} = 0.5 fF/μm	
Zero-bias planar bulk depletion capacitance	C_{Jn} = 0.1 fF/μm^2	C_{Jp} = 0.3 fF/μm^2
Zero-bias sidewall bulk depletion capacitance	C_{JSWn} = 0.5 fF/μm	C_{JSWp} = 0.35 fF/μm
Bulk junction potential	PB = 0.95 V	
Planar bulk junction grading coefficient	MJ = 0.5	
Sidewall bulk junction grading coefficient	$MJSW$ = 0.33	
Length of source and drain diffusions	L_{diff}= 3 μm	

$$\omega_T = \frac{g_m}{C_{gs} + C_{gd}} \tag{3.35}$$

3-3-2 Transit Frequency with Extrinsic Capacitances

With extrinsic capacitances included in the model, the transit frequency expression of Eq. (3.30) modifies to

This can be seen by inserting the model of Figure 3-12 into the test setup of Figure 3-9. C_{sb} and C_{db} are shorted to ground (assuming the bulk terminal is also grounded), while C_{gd} appears in parallel with C_{gs}.

Example 3-5: MOSFET Capacitance Calculation

Consider an n-channel MOSFET biased in saturation with V_{DS} = 2.5 V, I_D = 500 μA, L = 1 μm, and W = 20 μm. Determine all the capacitances in the small-signal model of Figure 3-12, except the gate-bulk capacitance C_{gb} that we consider negligible. Also calculate the transistor's transit frequency. Use the standard technology parameters defined in Table 3-1.

SOLUTION

Substituting C_{ox} and C_{ov} = 0.5 fF/μm into Eq. (3.31), together with the MOSFET dimensions, we find

$$C_{gs} = \frac{2}{3}(20 \cdot 1\mu m^2)\left(2.3\frac{fF}{\mu m^2}\right) + 20\mu m\left(0.5\frac{fF}{\mu m}\right) = 40.7fF$$

For the gate-drain capacitance, we obtain

$$C_{gd} = 20\mu m\left(0.5\frac{fF}{\mu m}\right) = 10fF$$

The remaining capacitances are the pn junction depletion capacitances C_{db} between the n^+ drain and the substrate and C_{sb} between the n^+ source and the substrate. Evaluating Eq. (3.34), using the source junction bias voltage of V_{SB} = 0 V yields

$$C_{sb} = 19fF$$

The drain junction has a bias voltage of $V_{DB} = V_{OUT}$ = 2.5 V, Evaluating Eq. (3.33) with this value and the given parameters gives

$$C_{db} = 11.6fF$$

Note that C_{db} is smaller than C_{sb} due to the larger reverse bias across the drain-bulk junction. To calculate the transit frequency, we first compute g_m using

$$g_m = \sqrt{2\mu_n C_{ox}\frac{W}{L}I_D} = \sqrt{2 \cdot 50\frac{\mu A}{V^2} \cdot \frac{20}{1} \cdot 500\ \mu A} = 1\ mS$$

Therefore

$$f_T = \frac{1}{2\pi} \cdot \frac{g_m}{C_{gs} + C_{gd}} = \frac{1\ mS}{40.7\ fF + 10\ fF} = 3.14\ GHz$$

3-3-3 Frequency Response with Intrinsic and Extrinsic Gate Capacitances

To analyze the frequency response of the common-source amplifier with intrinsic and extrinsic capacitances, we insert the model of Figure 3-12 into the small-signal circuit model of the amplifier, as shown in Figure 3-13(a). Note that we have neglected C_{gb} and also discarded C_{sb}, since this capacitor has both terminals shorted to ground.

To simplify the full analysis of the amplifier, we redraw it as shown in Figure 3-13(b). We have taken the Norton equivalent at the input and combined the resistors at the input and output to reduce the number of terms carried in the algebra.

We begin the analysis by writing KCL at nodes 1 and 2

$$0 = -\frac{v_s}{R_s} + \frac{v_{gs}}{R_s} + v_{gs}sC_{gs} + (v_{gs} - v_{out})sC_{gd} \quad (3.36)$$

$$0 = g_m v_{gs} + sC_{gd}(v_{out} - v_{gs}) + \frac{v_{out}}{R_{out}} + sC_{db}v_{out} \quad (3.37)$$

Next, solving Eq. (3.36) for v_{gs}, substituting into Eq. (3.37), and rearranging yields

$$\frac{v_{out}}{v_s} = \frac{-g_m R_{out}\left(1 - s\frac{C_{gd}}{g_m}\right)}{1 + b_1 s + b_2 s^2} \quad (3.38)$$

where

$$b_1 = R_s[C_{gs} + C_{gd}] + R_{out}[C_{db} + C_{gd}] + g_m R_{out} R_s C_{gd} \quad (3.39)$$

and

$$b_2 = R_s R_{out}(C_{gs}C_{gd} + C_{gs}C_{db} + C_{gd}C_{db}) \quad (3.40)$$

Although this result is algebraically complex, we can make a few preliminary observations about the terms in the numerator of Eq. (3.38):

- At DC (s = 0, or all capacitors set to zero), the voltage gain of the circuit is $-g_m R_{out}$, as we already concluded from the low-frequency analysis in Chapter 2.
- The numerator contains a right half plane zero, $z_1 = g_m/C_{gd}$. Since obviously $C_{gd} < C_{gs} + C_{gd}$, we conclude [via comparison with Eq. (3.35)] that this zero occurs at frequencies beyond ω_T, and is therefore irrelevant in many practical scenarios.

The denominator of the transfer function is a second-order polynomial in s with complicated dependencies on all component values. All we can say at first glance from inspecting the denominator is that we expect to see two poles in the frequency response of this circuit, because it can (in principle) be factored into two binomial terms. Note that this factorization would yield an even more complicated expression.

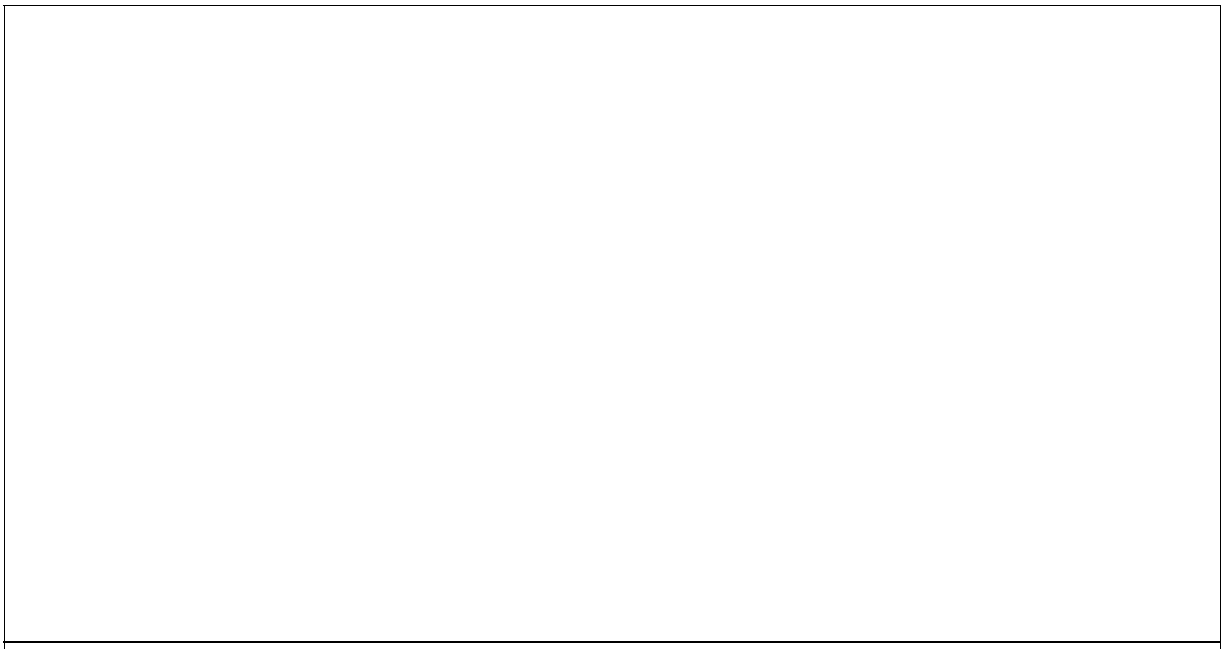


Figure 3-13: Small-signal model of the common-source amplifier with both intrinsic and extrinsic capacitances included. With (a) Thevénin equivalent input source configuration, and (b) Norton equivalent input source configuration.

The main issue with a result of this complexity is that it cannot be understood intuitively. Consequently, it is difficult to recognize the main parameters that are limiting the performance, which in turn prevents the designer from identifying ways to optimize the circuit. Even though Eq. (3.38) is mathematically exact, we would rather like to work with an expression that sacrifices some accuracy and/or detail in return for transparency and focus on the main effects that limit the performance. In order to take steps in this direction, we begin by evaluating Eq. (3.38) numerically, primarily to get a feel for the pole locations in a typical circuit.

Example 3-6: Magnitude Response of the Common-Source Amplifier

Evaluate and plot the steady-state magnitude response of Eq. (3.38) numerically using the following transistor parameters: g_m = 1 mS, C_{gs} = 40.7 fF, C_{gd} = 10 fF and C_{db} = 11.6 fF (same as in Example 3-5). Assume R_{out} = 5 kΩ and R_s = 50 kΩ. For comparison, also plot the magnitude response of Eq. (3.18), i.e., considering only the intrinsic gate capacitance.

SOLUTION

The plots are generated by letting $s = j\omega$ in Eq. (3.38) and Eq. (3.18), and subsequently plotting the magnitude of the expression as a function of frequency. The result is show in Figure Ex3-6. From the plots, we conclude the following:

- In the response that uses intrinsic capacitance only, we see a pole at approximately 100 MHz; this number corresponds to the value obtained in Example 3-4.
- As expected, the response with both intrinsic and extrinsic capacitances exhibits two poles. More importantly, we see that one of the poles occurs at relatively low frequencies, while the other occurs at very high frequencies.
- The low-frequency pole of the case with extrinsic capacitance included lies significantly lower than 100 MHz. This tells us that extrinsic capacitance has a substantial impact on the bandwidth of this circuit.

From the result of this particular example, we see that the bandwidth of the common-source amplifier is primarily set by a single pole that lies far from any other breakpoint in the response. In this case, we call the bandwidth limiting pole of the circuit the **dominant pole**. When a dominant pole condition exists, we would like to work with an expression of the form

$$\frac{v_{out}}{v_s} = \frac{A_{v0}}{1 - \dfrac{s}{p_1}} \tag{3.41}$$

instead of evaluating Eq. (3.38). In some sense, Eq. (3.38) contains too much information about irrelevant features of the response that have no impact on the 3-dB bandwidth. A commonly used technique that allows us to simplify expressions of

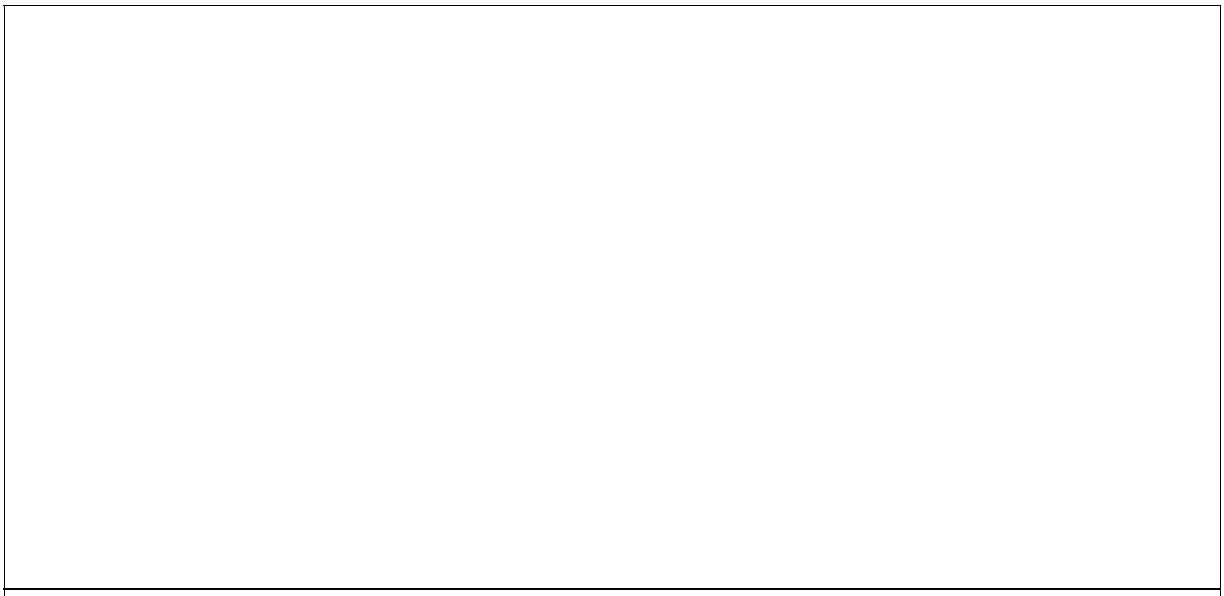


Figure Ex3-6

the from of Eq. (3.38) is therefore discussed in the next sub-section.

3-3-4 The Dominant Pole Approximation

In general, the denominator of the transfer function given in Eq. (3.38) can be factored into two binomial terms

$$\left(1-\frac{s}{p_1}\right)\left(1-\frac{s}{p_2}\right) = 1-s\left(\frac{1}{p_1}+\frac{1}{p_2}\right)+\frac{s^2}{p_1p_2} \tag{3.42}$$

Furthermore, we know from our numerical evaluation of the previous sub-section that the magnitude of one of the poles is much larger than the other, i.e.

$$|p_2| \gg |p_1| \tag{3.43}$$

and therefore

$$\left|\frac{1}{p_2}\right| \ll \left|\frac{1}{p_2}\right| \tag{3.44}$$

Consequently, we can eliminate the second term in s on the right hand side of Eq. (3.42) and approximate

$$\left(1-\frac{s}{p_1}\right)\left(1-\frac{s}{p_2}\right) \cong 1-\frac{s}{p_1}+\frac{s^2}{p_1p_2} \tag{3.45}$$

Now, comparing Eq. (3.38) with Eq. (3.45), we see that

$$-\frac{1}{p_1} = b_1 \tag{3.46}$$

and thus

$$\begin{aligned} p_1 &= -\frac{1}{b_1} \\ &= -\frac{1}{R_s[C_{gs}+C_{gd}]+R_{out}[C_{db}+C_{gd}]+g_mR_{out}R_sC_{gd}} \\ &= -\frac{1}{R_s[C_{gs}+(1+g_mR_{out})C_{gd}]+R_{out}[C_{db}+C_{gd}]} \end{aligned} \tag{3.47}$$

This result gives us a relatively handy expression for the dominant pole in the common-source amplifier, and the bandwidth can be estimated using

$$\omega_{3dB} = \frac{1}{R_s[C_{gs}+(1+g_mR_{out})C_{gd}]+R_{out}[C_{db}+C_{gd}]} \tag{3.48}$$

As opposed to Eq. (3.38), Eq. (3.48) is much more useful for evaluating which particular component of the circuit may limit the bandwidth. Specifically, the term $(1+g_mR_{out})C_{gd}$ looks like a potential problem. Whenever g_mR_{out} is large (high gain), this term may dominate the denominator of Eq. (3.48), and therefore limit the bandwidth. This is a very important conclusion, but unfortunately took us many lines of algebra (including the derivation of Eq. (3.38), which was not shown in detail) to develop. A more desirable approach would hint with very little algebra that the aforementioned term may limit the bandwidth.

Such an approach is possible via the application of the Miller theorem and the Miller approximation, discussed in the next sub-section.

3-3-5 The Miller Theorem and the Miller Approximation

The **Miller theorem** is a general linear circuit theorem that can be used to replace an impedance connected between two circuit nodes by two impedances, connected from each terminal to ground. This is illustrated in Figure 3-14. The impedance Z in Figure 3-14(a) is replaced by the two impedances Z_1 and Z_2 in Figure 3-14(b). For the two circuits to be equivalent, it can be shown that

Figure 3-14: Illustration of the Miller theorem.

$$Z_1 = \frac{Z}{1-A_{vM}} \quad \text{and} \quad Z_2 = \frac{A_{vM}Z}{A_{vM}-1} \tag{3.49}$$

where $A_{vM} = V_2/V_1$ is the voltage gain across the impedance Z, also called the **Miller gain.**

The Miller theorem is useful for the simplification of a variety of circuits. In the context of the common-source amplifier analysis in this chapter, we will use the theorem to eliminate the coupling of the output and input through C_{gd}, and thereby arrive at a circuit that is easier to analyze and understand. Before applying the Miller Theorem to the full circuit model of Figure 3-13, we will first consider its application to an ideal voltage amplifier circuit with a coupling capacitance between the input and output, as drawn in Figure 3-15.

The goal of this example is to determine the effective shunt capacitance at the input port, when the signal is amplified by a

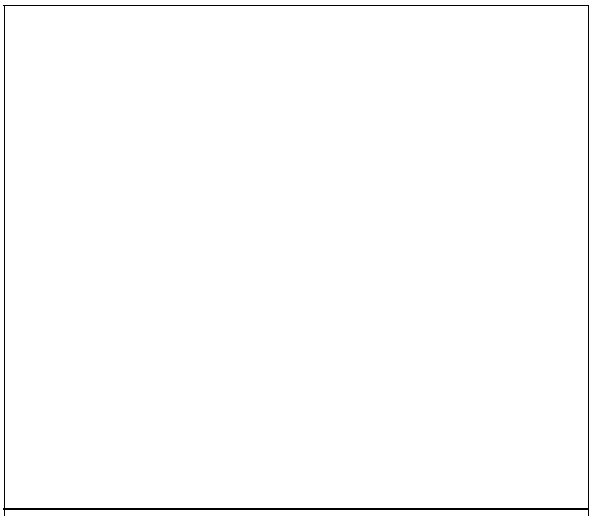


Figure 3-15: Idealized voltage amplifier with coupling capacitance between its input and output.

gain of A_{vM} across the coupling capacitor C. Using $Z = 1/sC$, and $Z_{\text{eff}} = 1/sC_{eff}$ we can apply Eq. (3.49) to find

$$\frac{1}{sC_{eff}} = \frac{\frac{1}{sC}}{1-A_{vM}} \tag{3.50}$$

and therefore

$$C_{eff} = C(1-A_{vM}) \tag{3.51}$$

If the voltage gain A_{vM} is a negative number (as in the case of a common-source amplifier), the capacitance C is "amplified" by the factor $(1 + |A_{vM}|)$. Intuitively, without relying on a complete proof of the Miller Theorem, this result can be understood by examining the voltages and currents of the capacitor C in Figure 3-16. The voltage across C is

$$v_C = v_{in} - v_{out} = v_{in}(1-A_{vM}) \tag{3.52}$$

and the current flowing into C from the input port is

$$\begin{aligned} i_{in} = i_C &= sCv_C = sC(1-A_{vM})v_{in} \\ \frac{v_{in}}{i_{in}} &= \frac{1}{sC(1-A_{vM})} = \frac{1}{sC_{eff}} \end{aligned} \tag{3.53}$$

In essence, the capacitance is multiplied due to the large swing at the amplifier output; this increases the voltage across the capacitor and therefore forces a correspondingly multiplied current into the input port.

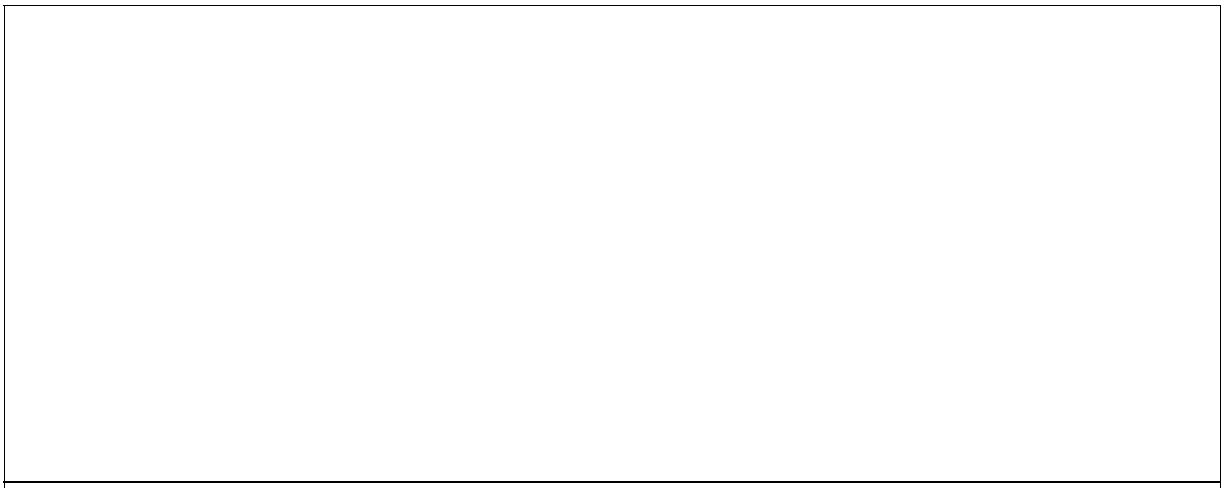


Figure 3-16: Circuit to analyze the voltage gain across C_{gd}.

This result applies qualitatively also to the common-source amplifier studied in this chapter, i.e., the negative gain of the amplifier causes an amplification of C_{gd}, which couples the input and output. However, a subtle difference is that the gain across the capacitor is not perfectly constant (as assumed above) but exhibits some frequency dependence.

To investigate, consider the circuit shown of Figure 3-16, which is the relevant subsection of the full common-source circuit needed to find the voltage gain across C_{gd}. Applying KCL at node 2 and solving for $A_{vM} = v_{out}/v_{gs}$ yields

$$\frac{v_{out}}{v_{gs}} = -g_m R_{out}\left(\frac{1 - s\dfrac{C_{gd}}{g_m}}{1 + sR_{out}(C_{db} + C_{gd})}\right) \tag{3.54}$$

In this expression, the bracketed term contains a zero and a pole. The zero occurs beyond ω_T and can be safely discarded. The situation is somewhat different for the pole. If R_{out} is very large, or if an additional load capacitance is added to the circuit output (in parallel with C_{db}), the pole can occur at relatively low frequencies, making the gain across C_{gd} non-constant in the frequency range of interest. Provided that the pole in the bracketed term occurs outside the frequency band of interest, we can assume

$$\frac{v_{out}}{v_{gs}} \cong -g_m R_{out} \tag{3.55}$$

This assumption is known as the **Miller approximation**, and it allows us to utilize the result from Eq. (3.51), which assumed a constant gain across the capacitor in question.

To complete this discussion, we will now apply the Miller approximation to the model of the common-source amplifier in Figure 3-13. The result is shown in Figure 3-17. The capacitor C_{gd} is no longer connected between the input and output, but appears only across the input port, with its value multiplied by $(1+ g_m R_{out})$. From this model, the circuit bandwidth can be easily identified by inspection

$$\omega_{3dB} = \frac{1}{R_s[C_{gs} + (1 + g_m R_{out})C_{gd}]} \tag{3.56}$$

In comparison with Eq. (3.48), this result is missing the term $R_{out}(C_{gd} + C_{db})$ in the denominator. This is not surprising and also inconsequential when the Miller Approximation is applied properly. As we pointed out above, the Miller approximation is justified only when this time constant is small in the first place, ensuring a constant Miller gain in the band where the dominant pole is expected to lie. Whenever the Miller approximation is applied, it must be verified that the neglected pole in the Miller gain occurs far beyond the frequency estimated by Eq. (3.56). This leads to the following procedure for the proper application of the Miller approximation in common-source amplifiers:

1. Calculate the low-frequency gain across C_{gd} and draw the simplified circuit model (as in Figure 3-17) with the Miller-amplified shunt capacitance at the input.
2. Estimate the bandwidth of the circuit using Eq. (3.56).
3. Calculate the frequency of the pole in Eq. (3.54). If and only if this pole frequency is far beyond the frequency calculated in step 2, the Miller approximation result is valid.

In a typical common-source circuit without a large load capacitance as drawn in Figure 3-13, the Miller approximation typi-

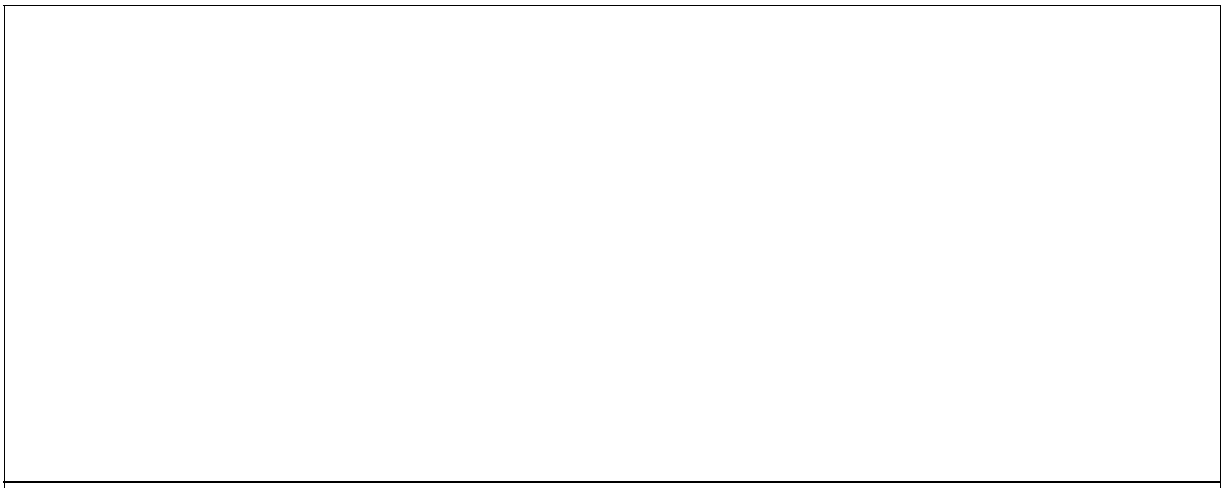


Figure 3-17: Small-signal model of the common-source amplifier after applying the Miller approximation.

cally holds. When a very large capacitor is connected to the output, the approximation becomes invalid and the dominant pole is set by the *RC* time constant formed at the output.

Example 3-7: Calculating the Common-Source Amplifier Bandwidth Using the Miller Approximation

Calculate the 3-dB bandwidth for the common-source voltage amplifier of Figure 3-13 using (a) the Miller approximation, and (b) the dominant pole approximation result of Eq. (3.48). Parameters: g_m = 1 mS, C_{gs} = 40.7 fF, C_{gd} = 10 fF, C_{db} = 11.6 fF, R_{out} = 5 kΩ and R_s = 50 kΩ (same as in Example 3-6). Calculate the percent error in the result of part (a).

SOLUTION

(a) Using the Miller approximation [i.e., Eq. (3.56)], we obtain

$$f_{3dB} = \frac{1}{2\pi} \cdot \frac{1}{R_s[C_{gs} + (1 + g_m R_{out})C_{gd}]}$$

$$= \frac{1}{2\pi} \cdot \frac{1}{50\text{k}\Omega[40.7\text{fF} + (1 + 1\text{mS} \cdot 5\text{k}\Omega)10\text{fF}]}$$

$$= 31.61\text{MHz}$$

(b) Using Eq. (3.48) we find

$$f_{3dB} = \frac{1}{2\pi} \cdot \frac{1}{R_s[C_{gs} + (1 + g_m R_{out})C_{gd}] + R_{out}[C_{db} + C_{gd}]}$$

$$= \frac{1}{2\pi} \cdot \frac{1}{50\text{k}\Omega[40.7\text{fF} + (1 + 1\text{mS} \cdot 5\text{k}\Omega)10\text{fF}] + 5\text{k}\Omega \cdot 21.6\text{fF}}$$

$$= 30.95\text{MHz}$$

The error in the result of part (a) is therefore

$$\frac{31.61 - 30.95}{30.95} = 2.1\%$$

The error of 2.1% seen in this example is acceptable and will in practice be overshadowed by uncertainty in the transistor model parameters.

3-3-6 Calculating the Non-Dominant Pole*

The reader may wonder how the non-dominant pole frequency can be calculated within the above-discussed framework. A common misconception is to assume that after applying the Miller approximation, the non-dominant pole can be simply found from the time constant in the output network, i.e., $R_{out}C_{db}$. This is incorrect since the Miller approximation is not valid at the frequency where the non-dominant pole is located.

If a dominant pole condition exists, the proper way to estimate the non-dominant pole is by comparing the coefficients of Eq. (3.45) and Eq. (3.38). Specifically, we utilize that

$$\frac{1}{p_1 p_2} = b_2 \tag{3.57}$$

and thus

$$p_2 = \frac{1}{p_1 b_2}$$

$$= -\frac{R_s[C_{gs} + C_{gd}] + R_{out}[C_{db} + C_{gd}] + g_m R_{out} R_s C_{gd}}{R_s R_{out}(C_{gs}C_{gd} + C_{gs}C_{db} + C_{gd}C_{db})} \tag{3.58}$$

To simplify, let us assume that $C_{gd} << C_{gs}$ and $C_{gd} << C_{db}$. Note that the latter assumption is not strictly true based on typical values for the technology assumed in this module (see Example 3-5). However, if a load capacitance is added to the circuit, the approximation is more easily justified, with C_{db} replaced by $C_{db} + C_L$ (see Example 3-8) and we almost always have in practice $C_{gd} << C_{db} + C_L$. Thus, under the stated conditions, we can write

$$p_2 \cong -\frac{R_s C_{gs} + R_{out} C_{db} + g_m R_{out} R_s C_{gd}}{R_s R_{out} C_{gs} C_{db}} = -\left(\frac{1}{R_{out} C_{db}} + \frac{1}{R_s C_{gs}} + \frac{g_m}{C_{db}} \cdot \frac{C_{gd}}{C_{gs}}\right) \tag{3.59}$$

This approximate result indicates that the non-dominant pole lies at a frequency that is higher than $1/R_{out}C_{db}$, especially when g_m is large. Note that Eq. (3.59) essentially represents a "parallel combination" of time constants (analogous to parallel connections of resistors); that is, the smallest time constant in the expression sets the pole frequency.

3-4 Open-Circuit Time Constant Analysis

3-4-1 General Framework

In Section 3-3-4, we derived an approximate expression for the 3-dB bandwidth of a common-source voltage amplifier, assuming that a dominant pole condition exists. In this analysis, we found that the bandwidth is fully determined by the coefficient b_1 in the numerator of Eq. (3.38).

The **open-circuit time constant (OCT) analysis** is a powerful and general technique that allows us to compute the term b_1 for arbitrary circuits, without the need to derive the full circuit transfer function with all high-order artifacts included. More importantly, it breaks the analysis into small and computationally manageable steps that provide insight about which circuit elements present the main bandwidth bottleneck. The step-by-step procedure for applying the OCT analysis method can be summarized as follows (see Reference 3 for a derivation)

1. Remove all but one capacitor in the circuit that is to be analyzed. Let us call this capacitor C_j.
2. Short all independent voltage sources and remove all independent current sources in the circuit.
3. Calculate the Thévenin resistance R_{Tj} seen by the capacitor C_j and compute the time constant $\tau_{jo} = R_{Tj}C_j$. Here, the subscript "*o*" is used to emphasize the open-circuit condition.
4. Repeat the above steps 1-3 for all remaining capacitors in the circuit.
5. The sum of all time constants is exactly equal to b_1. We can therefore estimate the circuit's 3-dB bandwidth using

$$\omega_{3dB} \cong \frac{1}{b_1} \tag{3.60}$$

$$b_1 = \sum_{j=1}^{N} \tau_{jo} \tag{3.61}$$

where N is the total number of capacitors in the circuit. The τ_{jo} are called **open-circuit time constants** because these were determined with all other capacitors open circuited.

Once a circuit is analyzed using the OCT method, we can see which of the individual open-circuit time constants is contributing most heavily to b_1. To increase the bandwidth, we can try to redesign the circuit by lowering the Thévenin resistance or the capacitor value of that time constant.

3-4-2 OCT Analysis of a Common-Source Stage

Consider the common-source amplifier shown in Figure 3-18(a) as an example to further understand the method of open-circuit time constants. We begin by first considering C_{gs} and therefore remove all other capacitors and short the input source as shown in Figure 3-18(b). As evident from this circuit, the Thévenin resistance seen by capacitor C_{gs} is R_S and the individual time constant contribution from C_{gs} is

$$\tau_{gso} = R_S C_{gs} \tag{3.62}$$

Similarly, redrawing the circuit with only C_{db} present will yield

$$\tau_{dbo} = R_{out} C_{db} \tag{3.63}$$

Next, we determine the individual time constant contribution from capacitor C_{gd}. To perform this calculation, we consider the circuit as redrawn in Figure 3-18(c). From this circuit, the Thévenin resistance seen across C_{gd} cannot be immediately determined by inspection. This is because of the g_m element, which couples the nodes to the left and right of the capacitance. We therefore resort to determining the Thévenin resistance from first principles, using a nodal analysis. As shown in Figure 3-19, we apply a test current source (i_t) and measure the resulting test voltage (v_t). Applying KVL and KCL, we find that

$$v_t = v_{gs} + R_{out}(g_m v_{gs} + i_t) \tag{3.64}$$

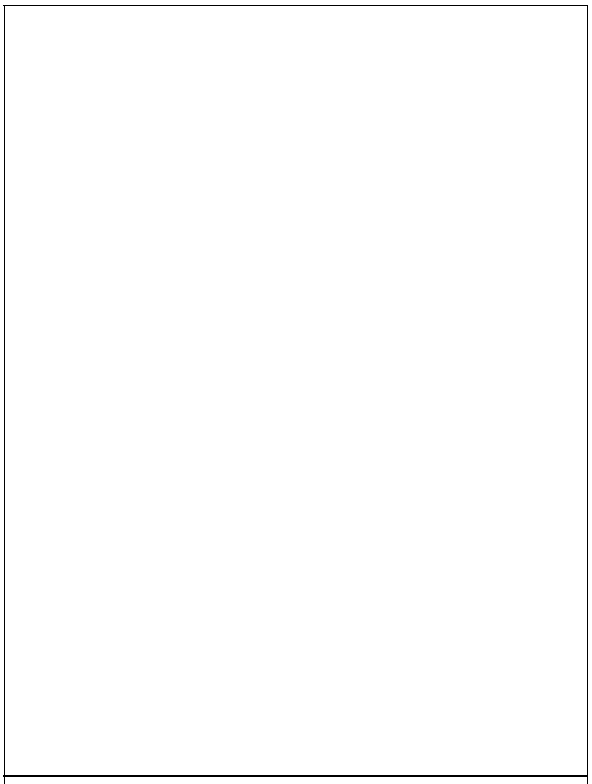


Figure 3-18: OCT analysis for a common-source amplifier. (a) Complete circuit. (b) Circuit for finding τ_{gs}. (c) Circuit for finding τ_{gd}.

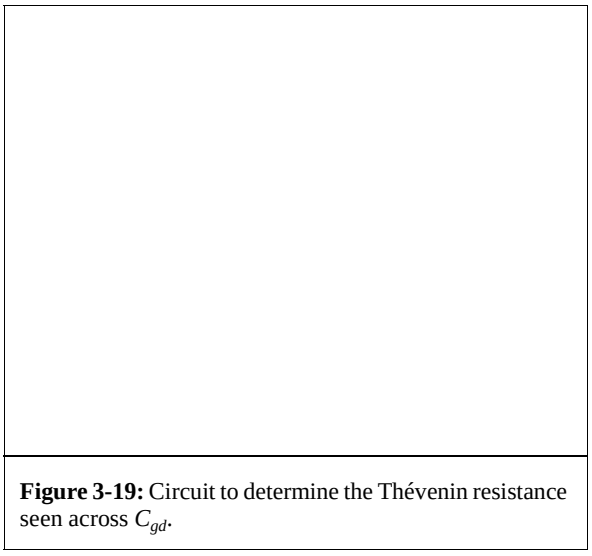


Figure 3-19: Circuit to determine the Thévenin resistance seen across C_{gd}.

$$v_{gs} = i_t R_S \tag{3.65}$$

After substituting Eq. (3.64) into Eq. (3.65), we obtain

$$\begin{aligned} R_{Tgd} &= \frac{v_t}{i_t} = R_S + R_{out}(g_m R_S + 1) \\ &= R_S + R_{out} + g_m R_S R_{out} \end{aligned} \tag{3.66}$$

A common way to memorize this final result is "$R_{left} + R_{right} + g_m R_{left} R_{right}$," where R_{left} and R_{right} are the resistances seen to the left and right of the coupling capacitance C_{gd}, respectively. Using this result, the individual time constant resulting from C_{gd} is given by

$$\tau_{gdo} = R_{Tgd} C_{gd} = [R_S + R_{out} + g_m R_S R_{out}] C_{gd} \tag{3.67}$$

Next, we add the individual time constants from Eq. (3.62), Eq. (3.63) and Eq. (3.67), which results in

$$\begin{aligned} b_1 &= R_s[C_{gs} + C_{gd}] + R_{out}\,[C_{db} + C_{gd}] + g_m R_{out} R_s C_{gd} \\ &= R_s[C_{gs} + (1 + g_m R_{out}) C_{gd}] + R_{out}\,[C_{db} + C_{gd}] \end{aligned} \tag{3.68}$$

Note that this result is identical to Eq. (3.39), which was obtained from an exact nodal analysis of the complete circuit. This verifies that the method of open-circuit time constants is an exact analysis to determine the factor b_1, which multiplies the first order term in s in the denominator of the generalized system function. As before, the resulting estimate of the 3-dB breakpoint frequency is therefore given by

$$\omega_{3dB} \cong \frac{1}{b_1} = \frac{1}{R_s[C_{gs} + (1 + g_m R_{out}) C_{gd}] + R_{out}\,[C_{db} + C_{gd}]} \tag{3.69}$$

It is important to remember that this result maintains good accuracy only if a dominant pole condition exists. As we showed in Section 3-3-4, this condition is required so that we can approximate $\omega_{3dB} \cong 1/b_1$. Finally it is worth noting that Eq. (3.69) shows that C_{gd} is effectively multiplied by the circuit's voltage gain; this corresponds to the Miller amplification effect discussed in the previous section.

A simple example where the dominant pole condition is not met is shown in Figure 3-20. The reader may prove that the exact transfer function of this circuit is

$$\frac{v_{out}}{v_s} = \frac{-g_m R}{(1 + sRC)(1 + sRC)} \tag{3.70}$$

and thus

$$|p_2| = |p_1| = \frac{1}{RC} \tag{3.71}$$

Therefore, we expect that the approximation of Eq. (3.45) cannot be applied and $1/b_1$ will not be a good estimate for the circuit's bandwidth. It is now interesting to calculate the error that

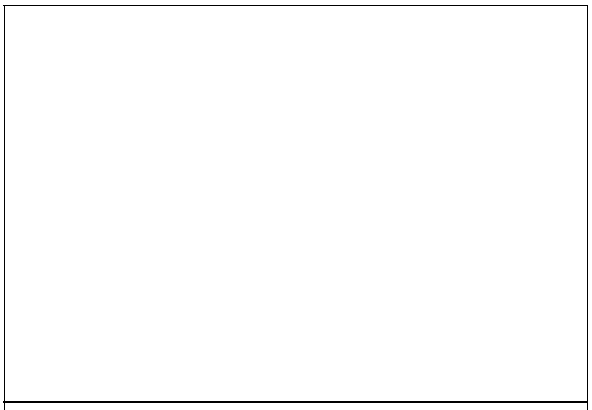


Figure 3-20: Circuit example that violates the dominant pole assumption.

will result if the OCT method is nonetheless "blindly" applied.

In performing the OCT analysis, we see that the circuit in question has two open-circuit time constants equal to *RC*. The bandwidth estimate using OCT analysis is therefore

$$\omega_{3dB,OCT} \cong \frac{1}{b_1} = \frac{1}{2RC} \tag{3.72}$$

On the other hand, we can find the exact 3-dB frequency of the circuit using

$$\frac{1}{\sqrt{2}} = \left| \frac{1}{(1 + j\omega_{3dB}RC)(1 + j\omega_{3dB}RC)} \right| \tag{3.73}$$

Solving for ω_{3dB} gives

$$\omega_{3dB} = \frac{1}{RC}\sqrt{\sqrt{2} - 1} \cong \frac{0.64}{RC} \tag{3.74}$$

The error in the OCT estimate is thus

$$\frac{0.5 - 0.64}{0.64} = -22\% \tag{3.75}$$

From this result, we can draw a few interesting conclusions. First, even though the dominant pole condition is grossly violated in the above example, the OCT analysis is not extremely far off from the exact result. Second, the OCT result is conservative in the sense that it tends to underestimate the circuit's bandwidth. This is desirable since the designer can rest assured that the bandwidth is at least as large as predicted by the OCT analysis. It can be shown that this latter property holds for arbitrary circuits whose poles lie on (or near) the real axis, and whose zeros occur beyond the estimated ω_{3dB}. This is the case for most circuits considered in this module. We will highlight exceptions where appropriate.

In summary, the reader should remember the following key points when applying the OCT analysis

- In any circuit, the sum of the open-circuit time constants corresponds (exactly) to the term b_1, which multiplies the first order term in the denominator of the circuit's *s*-domain transfer function.
- Under the following conditions, the bandwidth of the circuit can be approximated with good accuracy by $1/b_1$: (1) a dominant pole condition exists, (2) the transfer function contains only poles that lie on (or near) the real axis, and (3) the zeros in the transfer function occur beyond the bandwidth estimate in question.
- Even if no clear dominant pole condition exists, OCTs can be used to get a first-order feel for the bandwidth of a circuit. For instance, in a circuit with two identical real poles, the OCT bandwidth estimate is in error by –22%. As long as condition (2) above is met, the percent error will be negative and thus the estimated bandwidth is at least as large as the actual bandwidth (measured, e.g., using a circuit simulation).
- Open-circuit time constants, in general, do not necessarily correspond to the poles of a circuit. The OCT correspond to poles only in circuits that can be broken into decoupled RC sections, as is the case in the circuit of Figure 3-20.

Example 3-8: Common-Source Amplifier Bandwidth Estimate Using an OCT Analysis

Consider the circuit shown in Figure Ex3-8 and assume the following parameters: W = 20 µm, L = 1 µm, I_B = 500 µA, g_m = 1 mS, C_{gs} = 40.7 fF, C_{gd} = 10 fF, C_{db} = 11.6 fF, $R_{out} = R_D \parallel r_o$ =5 kΩ and R_s = 50 kΩ (same as in Example 3-7). The value of the load capacitance is C_L = 10 pF. Estimate the 3-dB bandwidth using an OCT analysis and propose a design modification that will increase the bandwidth by 20%. For this modification, you may not alter the circuit's DC gain, and R_S and C_L must be kept constant.

SOLUTION

(a) The circuit has three open-circuit time constants as expressed in Eq. (3.62), Eq. (3.63) and Eq. (3.67), with the difference that C_L appears in parallel to C_{db}. The three OCT expressions are therefore

$$\tau_{gso} = R_S C_{gs}$$

$$\tau'_{dbo} = R_{out}(C_{db} + C_L)$$

$$\tau_{gdo} = R_{Tgd}C_{gd} = [R_S + R_{out} + g_m R_S R_{out}]C_{gd}$$

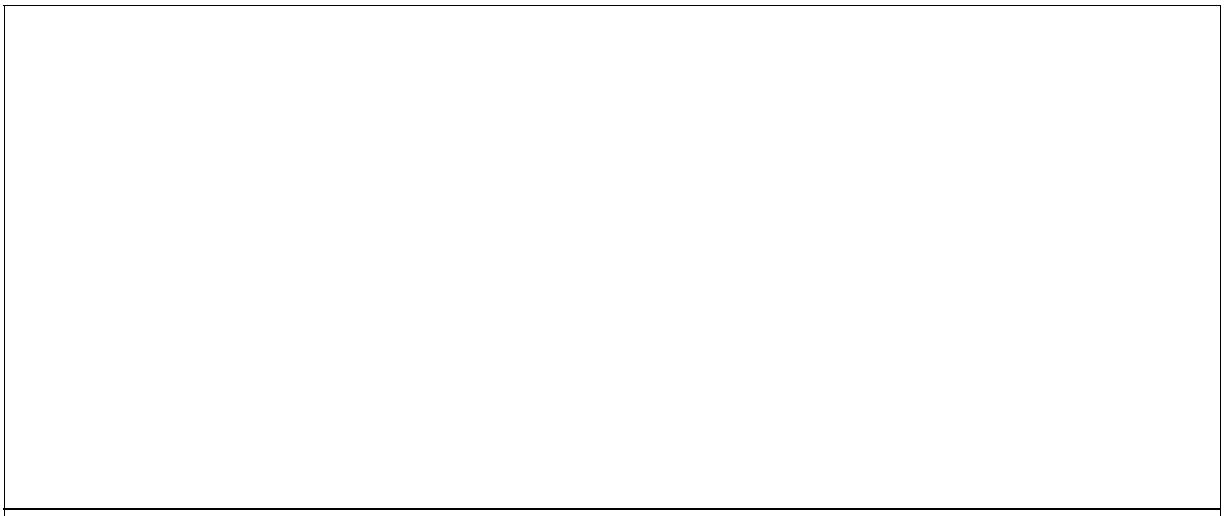


Figure Ex3-8

Evaluating these expression with the given numbers yields

$$\tau_{gso} = 50\text{k}\Omega \cdot 40.7\text{fF} = 2.035\text{ns}$$

$$\tau'_{dbo} = 5\text{k}\Omega(11.6\text{fF} + 10\text{pF}) = 50.01\text{ns}$$

$$\tau_{gdo} = [50\text{k}\Omega + 5\text{k}\Omega + 1\text{mS} \cdot 5\text{k}\Omega \cdot 50\text{k}\Omega]10\text{fF} = 3.05\text{ns}$$

The bandwidth estimate is

$$f_{3dB} = \frac{1}{2\pi} \cdot \frac{1}{2.035\text{ns} + 50.01\text{ns} + 3.05\text{ns}} = 2.89\text{MHz}$$

In order to improve the bandwidth by 20%, it is clear that we must reduce the dominant open-circuit time constant τ'_{dbo}. Since C_L must remain unchanged, the only option is to reduce R_{out}. To first order, reducing R_{out} to approximately 4 kΩ (a 20% reduction from the original value of 5 kΩ) should get us close to the desired improvement. In order to keep the DC gain of the circuit constant, we now require a larger transconductance

$$g_m = \frac{A_{v0}}{R_{out}} = \frac{5}{4\text{k}\Omega} = 1.2\text{mS}$$

There are several ways to increase the transconductance of the MOSFET. (i) Keep the device width constant and increase the bias current I_D. An advantage of this option is that none of the device capacitances will change, thereby avoiding any counterproductive increase in the total time constant. (ii) Keep I_D constant and increase the device width W. This option has the advantage that the current consumption of the circuit will not increase. Finally, option (iii) is to increase both W and I_D by the same factor. This option has the advantage that the gate overdrive voltage V_{OV} remains unchanged, and hence the input bias voltage and output voltage swing are unaffected. Since our primary focus in this example is to improve bandwidth, and current consumption and biasing considerations are secondary, we will apply option (i).

Using Eq. (2.30), the new value of the required I_B is

$$I_B = I_D = \frac{g_m^2}{2\mu_n C_{ox}\frac{W}{L}} = \frac{(1.2\text{mS})^2}{2 \cdot 50\frac{\mu\text{A}}{\text{V}^2} \cdot \frac{20}{1}} = 720\mu\text{A}$$

Note that this value is approximately 44% larger than the original bias current of 500 μA.

As a final verification step, we re-compute the bandwidth estimate using the new value of R_{out}. The time constant τ_{gso} remains the same, while the change in τ_{gdo} is negligible. The dominant OCT modifies as follows

$$\tau'_{dbo} = 4\text{k}\Omega(11.6\text{fF} + 10\text{pF}) = 40.05\text{ns}$$

The modified bandwidth estimate is therefore

$$f_{3dB} = \frac{1}{2\pi} \cdot \frac{1}{2.035\text{ns} + 40.05\text{ns} + 3.05\text{ns}} = 3.53\text{MHz}$$

which is about 22% larger than the original bandwidth, satisfying our design intent.

3-4-3 OCT Extensions

The OCT analysis covered in this section is tailored toward finding the upper corner frequency in circuits that are limited

by capacitive elements; this is the most common situation encountered in integrated circuit design. For completeness, it is worth mentioning that there exists a method of **short-circuit time constants** (see Reference 3), which aims at estimating the lower corner frequency of a circuit with a high-pass characteristic. This is useful for circuits that employ AC coupling of various forms.

In circuits that contain inductors, the additional time constants can be included by shorting all but one inductor at a time. The generalized framework that includes the consideration of both inductors and capacitors to estimate the upper corner frequency of a circuit is called **zero-value time constant analysis**. Finally, it is interesting to note that higher order terms [such as b_2 in Eq. (3.40)] can be found using an OCT-like analysis. The interested reader is referred to Reference 4 for a comprehensive discussion of such methods.

3-4-4 Time Constants versus Poles

The distinction between open-circuit time constants and poles tends to be a source of confusion among circuit design students. We will therefore review the differences in this section using two examples.

Consider first the circuit of Figure 3-20. As we have shown above, this circuit has two open-circuit time constants, equal to RC. Also we found that this circuit has two poles, located at $-1/RC$. Thus, in this particular circuit, the poles coincide with the (reciprocals of the) time constants. The reason for this coincidence is that the two networks at the input and output are fully decoupled and represent simple first order RC sections. For such a topology, the circuit designer sometimes loosely speaks of a "pole at the input" and "pole at the output," which are directly set by the time constants of each network.

Consider now the circuit of Figure 3-21, which is the same as Figure 3-20, except that we have added an additional capacitor C between the input and output terminal. This circuit retains the two open-circuit time constants of the original circuit (equal to RC), but has an additional one due to the added capacitor, equal to $RC(2 + g_mR)$. On the other hand, the poles of this circuit can no longer be found by inspection. The transfer function has the form of Eq. (3.38), with $b_1 = RC(4 + g_mR)$ and $b_2 = 3(RC)^2$. The two poles of the circuit are the roots of the denominator polynomial $1 + b_1s + b_2s^2 = 0$ and their value depends on the value of g_mR. Assuming $g_mR = 2$ as a numerical example, the roots, and therefore the poles become

$$p_{1,2} = -\frac{1}{RC}\left(1 \pm \sqrt{\frac{2}{3}}\right) \tag{3.76}$$

As we can see from this result, the poles do not coincide with any of the open-circuit time constants. More significantly, the number of poles (two) is not even equal to the number of open-circuit time constants (three).

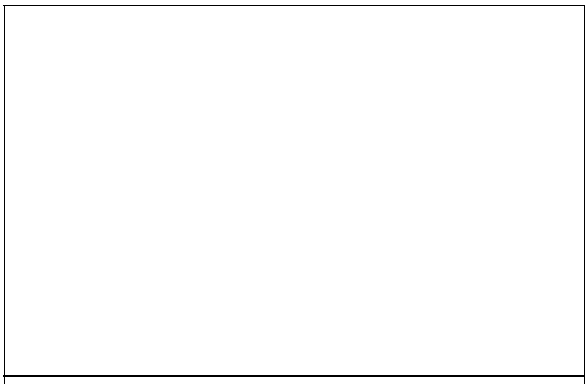


Figure 3-21: Circuit example with three open-circuit time constants.

One can show in general that a circuit's open-circuit time constants coincide with is its poles only in the special case of Figure 3-20 and its canonical extensions and duals. The general requirement is that the circuit must be separable into decoupled first order low-pass sections.

Finally, to fully close the loop between the two analysis techniques, note that the magnitude of the low-frequency pole in Eq. (3.76) is approximately equal to $0.18/RC$. This is close to the 3-dB bandwidth predicted by the sum of the open-circuit time constants (for $g_mR = 2$): $1/(4RC + RC + RC) = 0.167/RC$, which, as expected, is slightly conservative.

3-5 High-Frequency Two-Port Model for the Common-Source Voltage Amplifier

To summarize, Figure 3-22 shows the most general two-port model for the common-source voltage amplifier [similar to Figure 3-13(a)] with source and load networks included. The advantage of this model representation is that is valid for arbitrary component values. The disadvantage is that analyzing a circuit based on this model leads to complex equations. Generally, one should use this model as the starting point for the analysis of more complex circuits that contain a CS amplifier (see Chapter 6). Then, whenever suitable, we can invoke simplifications such as the Miller approximation or open-circuit time constants to simplify the analysis.

Finally, note that the model of Figure 3-22 is not well suited for a translation into a native voltage amplifier model (using a voltage controlled voltage source) as done for the low-frequency circuit in Section 2-4. The capacitors connected to the

output port would lead to a frequency dependent open-circuit gain and output impedance (Z_{out} rather than R_{out}) that give a non-intuitive representation of the circuit. It is thus preferred in general to describe this voltage amplifier using the transconductance model as shown.

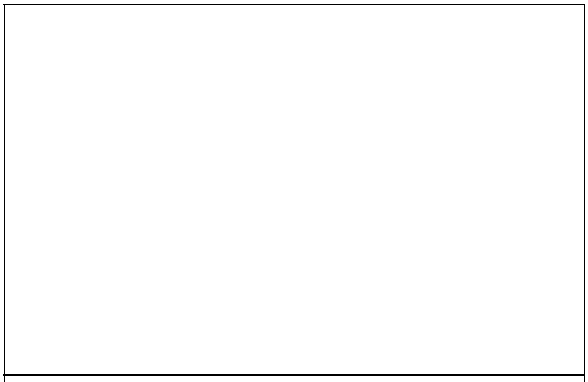


Figure 3-22: General two-port model for the common-source voltage amplifier valid at high frequencies.

Summary

In this chapter we have reviewed the basic concepts of frequency domain analysis and introduced the intrinsic and extrinsic device capacitances of a MOSFET. Using the obtained small-signal model, the frequency response of any circuit can be obtained from first principles using the following steps

1. Derive the transfer function using a nodal analysis.

2. Let $s = j\omega$ and solve for the magnitude of the resulting expression.

3. Set the magnitude equal to $1/(\sqrt{2})$ times the DC gain value, and solve for ω.

Unfortunately, this method is algebraically too complex for all but the most basic circuits. Consequently, we introduced several approximate methods and tools that are frequently used by analog circuit designers. These methods were developed using our driving example of a common-source voltage amplifier, but are widely used in other situations as well

- Provided that an exact (and potentially complicated) transfer function expression is available, the dominant pole approximation can be applied to arrive at a simplified bandwidth expression. In this approximation, it is assumed that a single pole dominates the response and sets the circuit's 3-dB bandwidth.
- The Miller approximation was used to obtain a quick estimate of the 3-dB bandwidth specifically for the common-source voltage amplifier. Although it is not an exact calculation, it is very useful for determining an estimate of the bandwidth of the amplifier analytically. Furthermore, this analysis gave insight into the effect of "Miller-multiplication" of a capacitor that appears across a voltage gain path. This effect is found in a multitude of circuits, and understanding this mechanism is insightful for design.
- The method of open-circuit time constants is the most powerful and most broadly applicable technique discussed in this chapter. It provides an accurate answer for the circuit's bandwidth if a dominant pole condition exists. Even if the dominant pole condition is not strictly met, the method yields acceptable errors (on the conservative side) on the order of a few tens of percent, which is often acceptable in a first-order hand analysis. Finally, the method of open-circuit time constants is an excellent design tool since it assists in finding which capacitors and Thévenin resistances are dominating the dynamic performance.

References

1. S. Bhave, R. T. Howe and C. G. Sodini, *TBD*, 2011.

2. D. K. Shaeffer and T. H. Lee, "A 1.5 V, 1.5 GHz CMOS low noise amplifier," *IEEE J. Solid-State Circuits*, pp. 745–759, May 1997.

3. P. E. Gray and C. L. Searle, *Electronic Principles Physics, Models, and Circuits*, Wiley, 1969.

4. A. Hajimiri, "Generalized Time- and Transfer-Constant Circuit Analysis," *IEEE Trans. Circuits and Systems I*, Vol. 27, No. 6, pp. 1105 - 1121, June 2010.

Problems

Unless otherwise stated, use the standard model parameters specified in Table 3-1 for the problems given below. Consider only first-order MOSFET behavior and include channel-length modulation (as well as any other second-order effects) only where explicitly stated.

P3.1 Sketch the Bode plots (magnitude and phase) for the following transfer functions. Assume $R_iC_i >> R_kC_k$ if $i > k$.

(a) $[1/(1 + j\omega R_1C_1)][(j\omega/(1 + j\omega R_2C_2)]$

(b) $(j\omega R_3C_3)[(1 + j\omega R_4C_4)/(1 + j\omega R_5C_5)]$

(c) $[(1 + j\omega R_6C_6)/(1 + j\omega R_8C_8)][(1 + j\omega R_7C_7)/(1 + j\omega R_9C_9)]$

P3.2 A system has a DC gain of 500, LHP zeros at 10 kHz and 1 MHz and LHP poles at 100 kHz, 10 MHz, and 100 MHz.

(a) Write the *s*-domain transfer function that describes this system.

(b) Draw a Bode plot for both the magnitude and phase of this system.

(c) Switch the poles and zeros and repeat parts (a) and (b).

P3.3 Sketch the Bode plot for the magnitude, $|I_o/I_s|_{\text{dB}}$ and phase $\angle I_o/I_s$ of the circuit shown in Figure P3-3, given

(a) $R_1 = 10$ kΩ, $R_2 = 100$ kΩ, $C = 1$ pF

(b) $R_1 = 0.1$ kΩ, $R_2 = 100$ kΩ, $C = 1$ pF

(c) $R_1 = 10$ kΩ, $R_2 = 100$ kΩ, $C = 10$ fF

Figure P3-3

P3.4 Repeat Example 3-5 for the following parameters (assuming $V_{DS} = 2.5$ V). For each case, compute by which factor the transistor's transit frequency has changed relative to the value seen in Example 3-5.

(a) $I_D = 500$ μA, $L = 2$ μm, and $W = 20$ μm.

(b) $I_D = 500$ μA, $L = 1$ μm, and $W = 40$ μm.

(c) $I_D = 1000$ μA, $L = 1$ μm, and $W = 40$ μm.

P3.5 Repeat Example 3-5 for a p-channel MOSFET operating in saturation. Parameters: $V_{SD} = 2.5$ V, $-I_D = 500$ μA, $L = 1$ μm, and $W = 20$ μm. Compute the ratio of the transit frequency obtained in Example 3-5 and the value obtained for the p-channel device analyzed in this problem. What is the main parameter that is responsible for the lower f_T observed for the p-channel MOSFET?

P3.6 Calculate the drain-bulk capacitance of a 100-μm wide n-channel transistor for $V_{DB} = 2.5$ V. Repeat the analysis for $V_{DB} = 1$ V and $V_{DB} = 4$ V and quantify by which factor the capacitance changes relative to the case of $V_{DB} = 2.5$ V.

P3.7 Calculate the 3-dB bandwidth of the circuit shown in Figure P3-7. Note that both MOSFETs operate in the triode region. Parameters: $W_1 = 10$ μm, $L_1 = 1$ μm, $W_2 = 10$ μm, $L_2 = 10$ μm, $V_B = 2$ V, $V_{DD} = 5$ V. Consider only the intrinsic gate capacitance.

Figure P3-7

P3.8 Plot the magnitude of Eq. (3.38) versus frequency using a software package such as Matlab and find the exact value of the 3-dB frequency from the resulting graph. Parameters: $g_m = 1$ mS, $C_{gs} = 40.7$ fF, $C_{gd} = 10$ fF, $C_{db} = 11.6$ fF, $R_{out} = 5$ kΩ and $R_s = 50$ kΩ (same as in Examples 3-7 and 3-8). Compare the obtained number with the approximate results obtained in Examples 3-7 and 3-8.

P3.9 Calculate the frequency of the non-dominant pole of the circuit analyzed in Example 3-6.

P3.10 For $A_{vM} = 1$, Eq. (3.51) predicts an effective input capacitance of $C_{eff} = 0$. Explain this result intuitively, in words. Hint: Consider the voltage waveforms at the input and output of Figure 3-15 for this particular case.

P3.11 In Example 3-7, we saw that the bandwidth estimate obtained through the Miller Approximation was in close agreement with the result from the full analysis (incorporating a dominant pole approximation). In contrast, if we were to apply the Miller Approximation result of Eq. (3.56) to Example 3-8, we would find a large error in the resulting answer (convince yourself that this is true). Explain why it is not appropriate to use Eq. (3.56) to estimate the bandwidth of the circuit in Example 3-8.

P3.12 Consider the circuit of Figure Ex3-8 with the following parameters: $W = 100$ μm, $L = 2$ μm, $I_B = 3$ mA, $V_B = 2.5$ V, $C_L = 100$ fF, $R_D = 1$ kΩ and $R_s = 10$ kΩ.

(a) Estimate the required DC input bias V_S such that $I_D = I_B$ and $V_{OUT} = V_B$. Neglect channel-length modulation.

(b) Calculate the MOSFET's transconductance and all device capacitances.

(c) Estimate the circuit's 3-dB bandwidth considering only the intrinsic gate capacitance.

(d) Estimate the circuit's 3-dB bandwidth using the Miller approximation.

(e) Estimate the circuit's 3-dB bandwidth using an OCT analysis.

P3.13 Consider the circuit shown in Figure P3-13.

(a) Write an analytical expression for the circuit's 3-dB bandwidth using an OCT analysis.

(b) Determine the exact analytical result for the circuit's 3-dB bandwidth.

(c) Compute the percent-error of the result in part (a), relative to the accurate result of part (b).

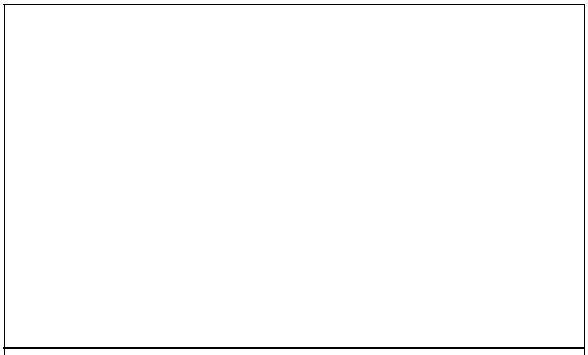


Figure P3-13

P3.14 In this chapter, we saw that using the OCT method to estimate a circuit's bandwidth tends to be conservative. For example, in a circuit with two identical real poles (and no zeros), the bandwidth predicted using the OCT method is 22% lower than the actual bandwidth [see Eq. (3.75)]. Derive an analytical expression $f(n)$ that returns the percent error of the OCT analysis for a circuit with n identical real poles and no zeros. Note that $f(2) = -22\%$.

P3.15 Consider the common-source voltage amplifier of Figure Ex3-8. The goal of this design problem is to achieve a small-signal DC gain of -4 and a 3-dB bandwidth of 80 MHz. In addition, we wish to minimize the current consumption of the circuit. For simplicity in your calculations, neglect channel-length modulation and consider only the intrinsic gate capacitance. Assume the following parameters: $R_s = 10$ kΩ, $R_D = 5$ kΩ, $C_L = 1$ pF.

(a) Show that the required drain current I_D is related to the circuit's parameters and specifications as expressed below. In your analysis, approximate ω_{3dB} using an OCT analysis. Plot I_D as a function of V_{OV} for $L = 1$ μm and $L = 1.5$ μm.

$$I_D = \frac{1}{2}\frac{C_L \cdot |A_{v0}| \cdot \omega_{3dB} \cdot V_{OV}}{1 - \frac{2}{3}\frac{L^2}{\mu_n V_{OV}} \cdot \frac{R_s}{R_D} \cdot |A_{v0}| \cdot \omega_{3dB}}$$

Note from this result that the choice of the gate overdrive voltage V_{OV} plays an important role in minimizing the required drain current.

(b) From the expression and plots found in part (a), it is clear that the minimum channel length minimizes the current consumption of the amplifier. Explain in your words why this should be the case.

(c) The drain current expression derived in (a) has a minimum for a certain value of V_{OV}. Calculate this value assuming $L = 1$ μm (minimum length). Also calculate the device width and drain current for the transistor at this optimum point.

(d) Simulate the design using SPICE with the bias current and device geometries calculated in part (c). Measure the bandwidth of the circuit using an AC simulation. Since the SPICE transistor model contains extrinsic capacitances and finite output resistance, your circuit should fall short of the desired specs (despite the fact that we have used a conservative OCT estimate for ω_{3dB}). Calculate the percent discrepancies in the gain and bandwidth of the circuit.

(e) Use a spreadsheet or math tool (Excel, Matlab, etc.) to setup the design equations for gain and bandwidth that include extrinsic capacitances and finite output conductance. With these additional modeling components added, it is difficult to derive a compact closed form solution as above. However, the setup in the spreadsheet will allow you to sweep the design parameters easily to find the new optimum that meets the gain and bandwidth specs. There are many different ways in which the spreadsheet can be structured. One is to use the widths of the two transistors as the main "knobs" and calculate/iterate over all other parameters. The hand-calculated result can be used as an initial guess in this optimization. Use your spreadsheet to calculate the new bias current and device size that will meet the specs.

(f) Simulate the refined design from (e) in SPICE and verify that you meet the desired specs. What is the obtained I_D and how much larger is this value compared to the result from (c)?

P3.16 For the circuit shown in Figure P3-16, prove the following results, quantifying the Thévenin resistances seen between each pair of transistor terminals. Neglect the finite r_o of the MOSFET.

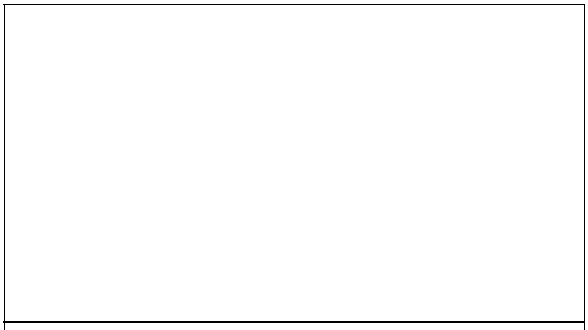


Figure P3-16

$$R_{gs} = \frac{R_G + R_S}{1 + g_m R_S}$$

$$R_{ds} = \frac{R_D + R_S}{1 + g_m R_S}$$

$$R_{gd} = R_G + R_S + G_m R_G R_S$$

where

$$G_m = \frac{g_m}{1 + g_m R_S}$$

P3.17 The circuit shown in Figure P3-17 is called a "source degenerated" common-source voltage amplifier. Analyze this circuit as indicated below.

(a) Neglecting channel-length modulation and all capacitances in the circuit, show that the circuit's small-signal DC gain is given by

$$\frac{v_{out}}{v_{in}} \cong -G_m R_D$$

where

$$G_m = \frac{g_m}{1 + g_m R_S}$$

is called the compound transconductance.

(b) Using the results stated in problem 3.16, estimate the DC gain and 3-dB bandwidth of the circuit assuming the following parameters: $R_G = 10$ kΩ, $R_S = 1$ kΩ, $R_D = 5$ kΩ, $I_D = 500$ μA, $L = 1$ μm, and $W = 20$ μm. Consider only the two time constants contributed by C_{gs} and C_{gd}.

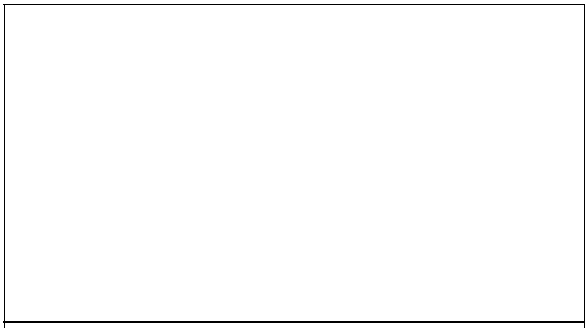


Figure P3-17.

4

CHAPTER

The Common-Gate and Common-Drain Stages

As we have seen from the discussion in the previous two chapters, the common-source stage can be used to realize a basic voltage amplifier. In this chapter, we will introduce the common-gate and common-drain stages, which for instance can be combined with the common-source amplifier to enhance its performance. More generally, as already indicated in Section 1-1, most analog amplifier circuits can be modeled as a combination of common-source, common-gate and common-drain configurations. For this reason, the three basic stage configurations can be viewed as the "atoms" of analog circuit design.

Chapter Objectives

- Analyze the common-gate and common-drain stages with respect to their low- and high-frequency transfer functions and port resistances (and impedances).
- Extend the MOSFET model as needed to capture relevant new effects that must be included in the analysis.
- Provide a first pass look at application examples for the common-gate and common-drain stages.

4-1 Overview of Stage Configurations

In the common-source (CS) amplifier discussed in the previous chapters, the input signal is applied to the gate and the output signal is taken from the drain. In a **common-gate (CG) amplifier**, the source is used as the input and the drain serves as the output terminal. In the **common-drain (CD) amplifier**, the input is applied at the gate and the output is taken from the source. Thus, in the context of two-port amplifiers with one terminal common between the input and output, the CS, CG, and CD circuits represent all three possible configurations (see Figure 4-1).

In our detailed analysis of the CG amplifier, we will find that this topology has a very low input resistance and a very high output resistance, which is exactly what we would want in a current amplifier. Conversely, the CD stage has a high input resistance and a small output resistance, which corresponds to the desired characteristics of a voltage amplifier (see Section 1-3).

4-2 Bulk Connection Scenarios and Required Model Extensions

Before engaging in a detailed analysis of the CG and CD stages, we need to think about the bulk connection in these stages. For the CS stage discussed so far, it was natural to connect the bulk to the source of the MOSFET. However, now that the source is connected either to the input or output port, we need to develop an understanding of the available options.

The first aspect to consider is related to technology constraints, and specifically how the n-channel and p-channel devices are formed in the integrated circuit substrate. As already indicated in Section 2-1-2, we assume in this module

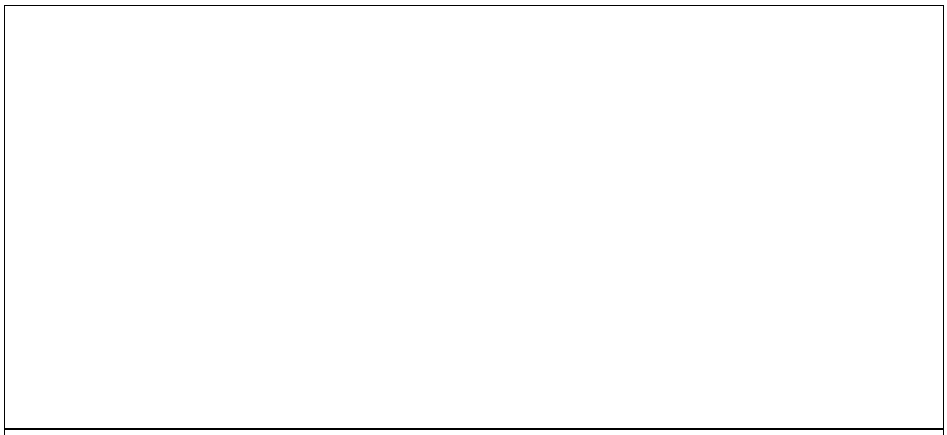


Figure 4-1: Conceptual illustration of the three elementary transistor stage configurations.

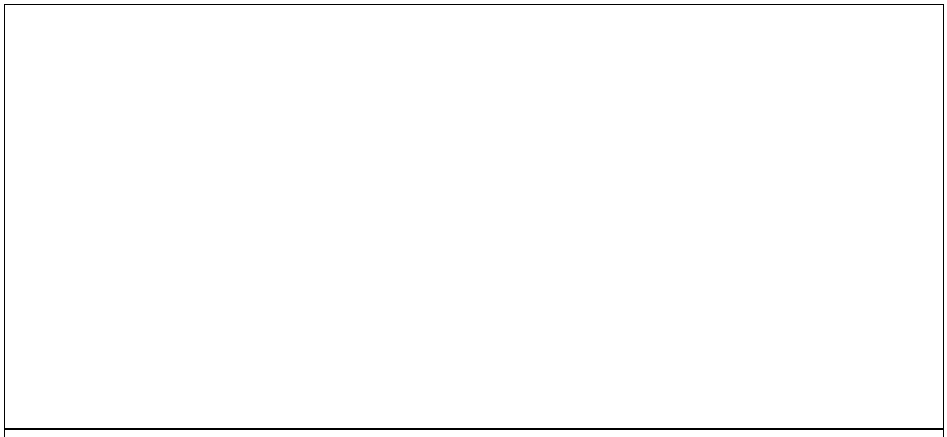


Figure 4-2: Cross-section of an n-well CMOS technology.

that the MOSFETs are built in a so-called n-well technology (see Figure 4-2). This means that the substrate is p-type, and the n-channel transistors are formed directly in the substrate. In order to create p-channels, n-type wells are diffused into the substrate, and the p-channels are subsequently formed in these regions.

Given the cross-section of Figure 4-2, it is clear that all n-channel transistors share the same bulk node. In contrast, the bulk node of a p-channel MOSFET can be isolated and freely connected to an arbitrary potential. Figure 4-3 shows the possible bulk connection scenarios arising from these constraints. For CG and CD stages built using n-channels in an n-well process, the bulks are always connected to the substrate potential (assumed to be "ground" in this module) by default. On the other hand, we are free to choose the bulk connection in p-channel CG and CD amplifiers. The most common scenarios encountered in practice are to tie the p-channel bulk to the supply voltage (V_{DD}), or to the source terminal. Either choice can have advantages and disadvantages, depending on the given circuit. In order to be able to reason about the associated trade-offs, we will now extend the MOSFET model such that the bulk node is incorporated as a fourth "free" terminal.

4-2-1 Well Capacitance

As indicated in Figure 4-2, the n-well region of the p-channel transistor forms a pn-junction with the p-type substrate (D_{bsub}). Similar to the junctions formed by the source/drain regions (see Section 3-3-1), we can model this junction as a parasitic capacitance between the bulk terminal and the substrate. Whether or not this capacitance must be considered depends on the bulk connection. As shown in Figure 4-4, if the bulk is tied to the supply, the junction capacitance is typically irrelevant, since it is connected between V_{DD} and ground, having no impact on circuit nodes that carry the signal. On the other hand, if the bulk is connected to the source, the capacitance will appear across

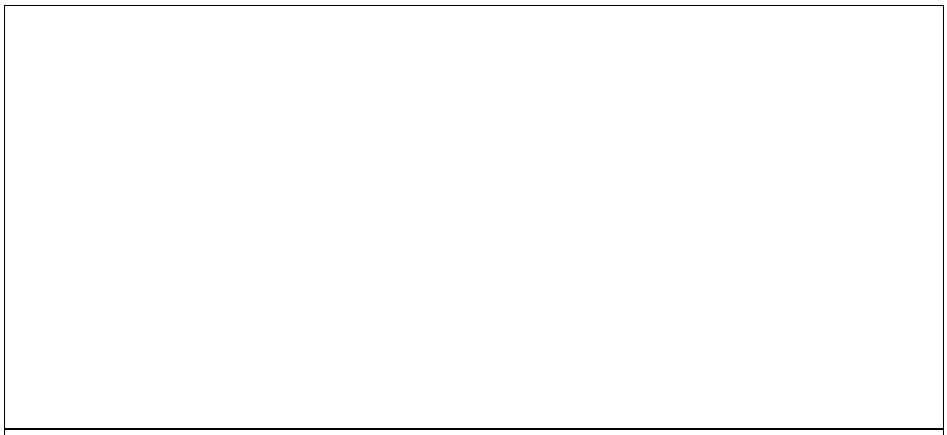


Figure 4-3: Possible bulk connection scenarios.

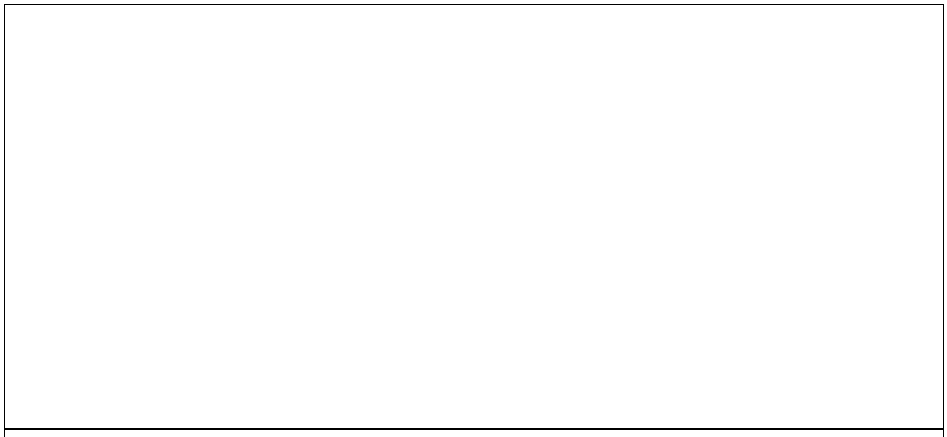


Figure 4-4: Connection of the well-to-substrate diode for various configurations. (a) P-channel bulk connected to the supply. (b) Bulk connected to the source in a p-channel common-gate stage. (c) Bulk connected to the source in a p-channel common-drain stage.

the input port for the CG stage, and across the output port of the CD stage. In this case, the parasitic capacitance contribution from the well must be taken into account.

Similar to the expression we used to estimate the source/drain junction capacitances, we can obtain an estimate for the well capacitance using

$$C_{bsub} = \frac{C_{Jwell} \cdot A_{well}}{(1 + V_{BSUB}/PB)^{MJ}} \tag{4.1}$$

This expression is similar to Eq. (3.33), except that the sidewall contribution has been omitted for simplicity. In Eq. (4.1), V_{BSUB} is the voltage between the bulk node and the substrate (a positive voltage, i.e., the junction is reverse biased). C_{Jwell} is the zero-bias depletion capacitance parameter for the junction and A_{well} is the area of the n-well under consideration. The well area depends strongly on the actual layout of the transistor. However, for approximate calculations it is often sufficient to assume a basic rectangular well shape using reasonable geometry estimates. We will illustrate this using an example.

Example 4-1: Well Capacitance Calculation

Consider a p-channel MOSFET that was laid out as shown in Figure Ex4-1. Estimate the well-to-substrate capacitance C_{bsub} assuming the following parameters: W = 10 µm, L = 3 µm, X_1 = 5 µm (diffusion to well edge spacing), X_2 = 3 µm (diffusion spacing), C_{Jwell} = 0.05 fF/µm^2, and V_{BSUB} = 2.5 V.

SOLUTION

From Figure Ex4-1, we see that

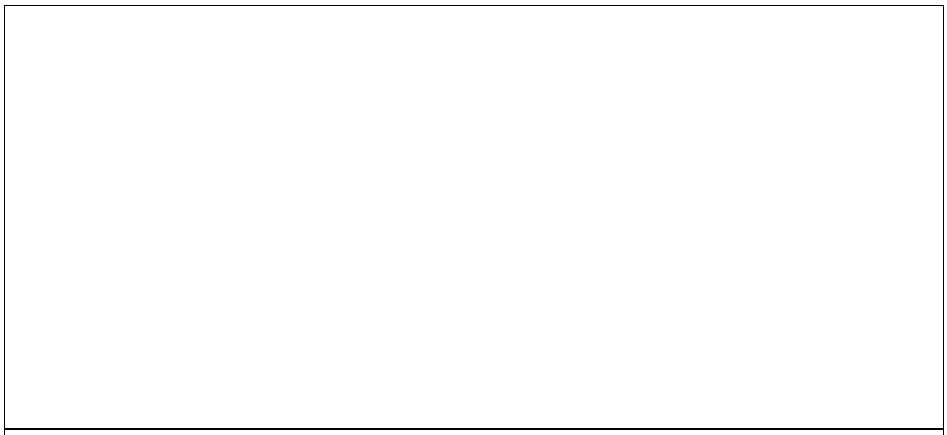


Figure Ex4-1

$$\begin{aligned} A_{well} &= (W + 2X_1) \times (L + 2L_{diff} + 2X_1 + X_2) \\ &= (10 + 10) \times (1 + 6 + 10 + 3)\mu m^2 \\ &= 400\mu m^2 \end{aligned}$$

Using Eq. (4.1) and the technology parameters for *PB* and *MJ* from Table 3-1, we obtain

$$\begin{aligned} C_{bsub} &= \frac{C_{Jwell} \cdot A_{well}}{(1 + V_{BSUB}/PB)^{MJ}} \\ &= \frac{0.05fF \cdot 400}{(1 + 2.5/0.95)^{0.5}} \\ &= 10.5fF \end{aligned}$$

From the above calculation, we see that the well capacitance can be comparable to the intrinsic and extrinsic device capacitances (see Example 3-5). Hence, whenever the designer chooses a bulk connection for which the well capacitance appears at an internal circuit node (other than the supply), care must be taken in estimating C_{bsub} and incorporating this capacitance in the overall circuit model.

As a final note, it is worth mentioning that most circuit simulation tools such as SPICE, do not automatically account for the well capacitance. Whenever relevant, the circuit designer must (manually) ensure that an appropriate modeling element is included in the simulation. This can be done, for example, by adding a properly modeled diode or capacitor to the node in question.

4-2-2 Backgate Effect

In Section 2-1, we analyzed the MOSFET's I-V characteristics assuming that the bulk node is connected to the source, i.e., $V_{BS} = 0$. However, in the CG and CD configurations, this is not the case unless the designer opted for a source-to-bulk tie as in the circuit of Figure 4-4(b) and (c). Therefore, we will now refine the I-V expressions from Section 2-1, for the case of non-zero V_{SB} and using an n-channel MOSFET for the analysis.

For an n-channel device with positive V_{SB}, i.e., the source lies at a higher potential than the (grounded) substrate, the primary effect is that the depletion region between the source and the substrate is widened (see Figure 4-5). This follows from the usual behavior of pn junctions – the depletion region of a reverse biased junction widens for increasing reverse bias.

Qualitatively speaking, the widened depletion region increases the amount of negative fixed charge (N_A^-) near the source. This extra negative charge opposes the injection of electrons from the source, and hence a larger V_{GS} is needed to cause the same drain current I_D. The larger required V_{GS} is commonly modeled as an effective increase in the device's threshold voltage. The dependence of the threshold voltage on the bulk-source voltage is called **backgate effect**. A detailed analysis based on solid-state physics reveals that the relationship between the applied V_{SB} and the MOSFET's threshold voltage is given by the following expression (see Reference 1)

$$V_{Tn}(V_{SB}) = V_{TOn} + \gamma_n(\sqrt{2\phi_f + V_{SB}} - \sqrt{2\phi_f}) \qquad (4.2)$$

where V_{TOn} is the threshold voltage without backgate effect, γ_n is the n-channel MOSFET **backgate effect parameter,** and ϕ_f is the **surface potential parameter**. With backgate effect included, the large-signal I-V characteristic for the saturation region is

$$I_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{GS} - V_{Tn}(V_{SB}))^2(1 + \lambda_n V_{DS}) \qquad (4.3)$$

Figure 4-5: N-channel MOSFET with backgate bias $V_{SB} > 0$. The depletion region widens in comparison with $V_{SB} = 0$.

which is identical to Eq. (2.45), except that V_{Tn} is now a function of V_{SB}, as defined in Eq. (4.2).

In order to incorporate the backgate effect into the small signal model of the transistor, we will follow exactly the same approach as in Section 2-3-1. That is, we approximate the incremental drain current around the operating point as the total differential, now with a drain current perturbation due to v_{bs} included

$$\begin{aligned} i_d &= \left.\frac{\partial i_D}{\partial v_{GS}}\right|_Q \cdot v_{gs} + \left.\frac{\partial i_D}{\partial v_{DS}}\right|_Q \cdot v_{ds} + \left.\frac{\partial i_D}{\partial v_{BS}}\right|_Q \cdot v_{bs} \\ &= g_m v_{gs} + g_o v_{ds} + g_{mb} v_{bs} \end{aligned} \tag{4.4}$$

In this expression, g_m and g_o are the transconductance and output conductance, respectively, and as derived previously in Chapter 2. The new term g_{mb} is called **backgate transconductance** and represents the perturbation of the drain current by an incremental change in the bulk-source voltage.

To compute g_{mb}, it is convenient to expand the partial derivative in Eq. (4.4) using the chain rule

$$g_{mb} = \left.\frac{\partial i_D}{\partial v_{SB}}\right|_Q = \left.\frac{\partial i_D}{\partial (V_{GS} - V_{Tn})}\right|_Q \left.\frac{\partial (V_{GS} - V_{Tn})}{\partial V_{Tn}}\right|_Q \left.\frac{\partial V_{Tn}}{\partial v_{BS}}\right|_Q \tag{4.5}$$

The first partial derivative in the above expression is simply g_m. The second term is equal to –1. Therefore, after partial differentiation of the threshold voltage [Eq. (4.2)] we obtain

$$g_{mb} = -g_m \left.\frac{\partial V_{Tn}}{\partial v_{BS}}\right|_Q = \frac{g_m \gamma_n}{2\sqrt{2\phi_f + V_{SB}}} \tag{4.6}$$

Analogous to the way we incorporated g_m and g_o, the backgate transconductance can be included in the transistor's small signal model as shown in Figure 4-6. As defined in Eq. (4.4), the g_{mb} element captures the dependence of the drain current to incremental changes in the bulk-source voltage. From this final result, we see that the term **backgate** was coined to reflect that the bulk acts just like another gate of the transistor. The only difference between the backgate and the actual gate node is that the backgate is physical located at the back side of the MOSFET, separated from the channel by a depletion layer.

For circuit design, it is useful to have a feel for the magnitude of g_{mb}, as well as for the range of threshold voltage ranges for typical bias conditions. The following example investigates typical numbers for the technology assumed in this module.

Example 4-2: Backgate Effect

Consider an n-channel MOSFET with the following parameters $V_{T0n} = 0.5$ V, $\gamma_n = 0.6\ \text{V}^{1/2}$, $\phi_f = 0.4$ V, and $V_{SB} = 2.5$ V, 1.5 V and 0 V. For each case, calculate V_{Tn} and the ratio g_{mb}/g_m.

SOLUTION

Evaluating Eq. (4.2) and Eq. (4.6) for $V_{SB} = 2.5$ V, we have

$$\begin{aligned} V_{Tn} &= V_{TOn} + \gamma_n(\sqrt{2\phi_f + V_{SB}} - \sqrt{2\phi_f}) \\ &= 0.5\text{V} + 0.6\text{V}^{1/2}(\sqrt{0.8\text{V}+2.5\text{V}} - \sqrt{0.8\text{V}}) = 1.05\text{V} \end{aligned}$$

$$\frac{g_{mb}}{g_m} = \frac{\gamma_n}{2\sqrt{2\phi_f + V_{SB}}} = \frac{0.6\text{V}^{1/2}}{2\sqrt{0.8\text{V} + 2.5\text{V}}} = 0.165$$

After evaluating the same expressions for the remaining two cases, we can summarize the obtained results as shown in the table below.

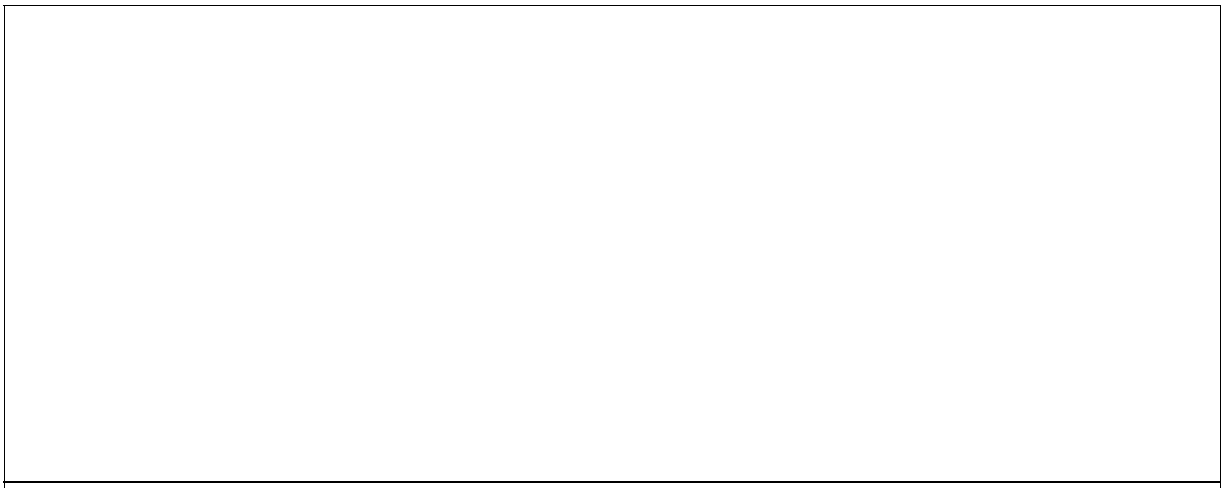


Figure 4-6: N-channel MOSFET small-signal model for the saturation region with backgate transconductance (g_{mb}) included.

V_{SB} [V]	V_{Tn} [V]	g_{mb}/g_m
2.5	1.05	0.165
1.5	0.87	0.198
0	0.50	0.335

From these values, we see that the threshold voltage of an n-channel MOSFET in our technology shifts up by about 0.5 V as the source-bulk bias voltage approaches 2.5 V (half of the supply voltage assumed in this module). The backgate transconductance ranges approximately between 15% and 30% of g_m over the same backgate bias range. Note that the backgate transconductance reduces for larger V_{SB}. This makes intuitive sense, since the depletion region widens for larger reverse bias. Qualitatively speaking, this increases the distance from the bulk electrode to the inversion layer, weakening the effect of backgate voltage changes on the incremental drain current.

Table 4-1 summarizes the MOSFET modeling parameters for our standard technology with backgate effect parameters included. This table represents the complete set of parameters used in this module; no further extensions will be needed.

Table 4-1: Standard technology parameters for the λ-model, with intrinsic and extrinsic capacitance parameters, as well as backgate effect parameters included.

Parameter	**n-channel MOSFET**	**p-channel MOSFET**
Threshold voltage (at $V_{BS} = 0$)	$V_{T0n} = 0.5$ V	$V_{T0p} = -0.5$ V
Transconductance parameter	$\mu_n C_{ox} = 50\ \mu A/V^2$	$\mu_p C_{ox} = 25\ \mu A/V^2$
Channel length modulation parameter	$\lambda_n = 0.1\ V^{-1}/L$ (L in μm)	$\lambda_p = 0.1\ V^{-1}/L$ (L in μm)
Gate oxide capacitance per unit area	$C_{ox} = 2.3\ fF/\mu m^2$	
Overlap capacitance	$C_{ov} = 0.5$ fF/μm	
Zero-bias planar bulk depletion capacitance	$C_{Jn} = 0.1\ fF/\mu m^2$	$C_{Jp} = 0.3\ fF/\mu m^2$
Zero-bias sidewall bulk depletion capacitance	$C_{JSWn} = 0.5$ fF/μm	$C_{JSWp} = 0.35$ fF/μm
Zero bias well-to substrate capacitance	—	$C_{Jwell} = 0.05\ fF/\mu m^2$
Bulk junction potential	$PB = 0.95$ V	
Planar bulk junction grading coefficient	$MJ = 0.5$	
Sidewall bulk junction grading coefficient	$MJSW = 0.33$	
Length of source and drain diffusions	$L_{diff} = 3$ μm	
Backgate effect parameter	$\gamma_n = 0.6\ V^{1/2}$	$\gamma_p = 0.6\ V^{1/2}$
Surface potential parameter	$\phi_f = 0.4$ V	

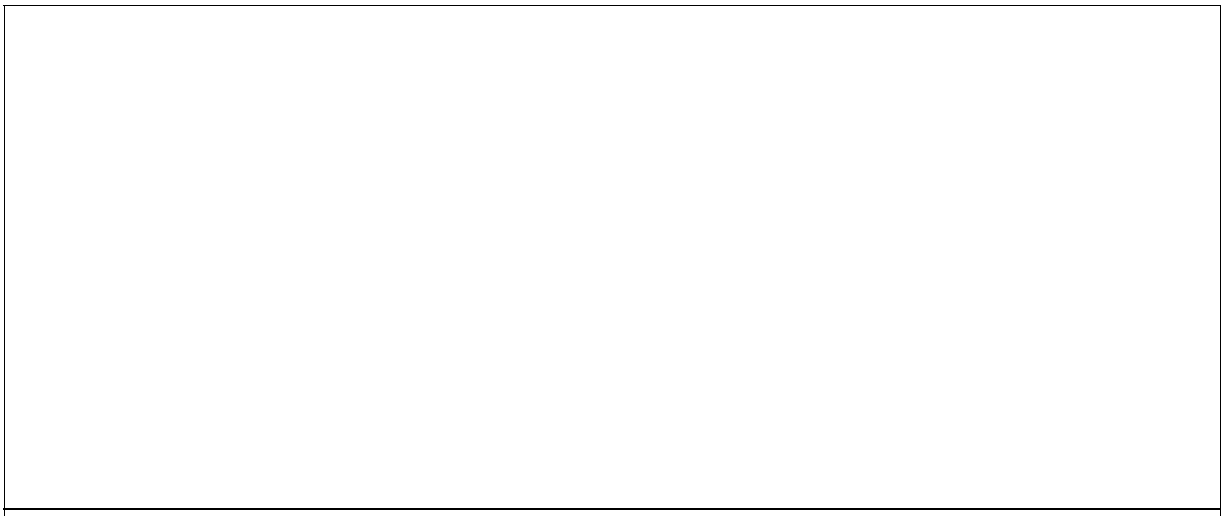


Figure 4-7: Practical realization of a CG transresistance amplifier with biasing circuitry included.

4-3 Analysis of the Common-Gate Stage

Using the additional modeling insight obtained in the previous section, we are now ready to analyze the CG circuit in detail. For the derivations in this section, we will assume the circuit topology shown in Figure 4-7. In this example configuration, the MOSFET's quiescent point drain current (and V_{GS}) is established using the input bias current source I_B. The auxiliary bias voltage source V_B is used to define the operating point node voltages at the gate and source of the transistor. The resistor R_D sources the bias current and sets the quiescent point voltage at the output node. One minor additional point to note is that the circuit is drawn different from Figure 4-1(b) such that the drain node lies above the source. This is the preferred way to draw this circuit for large signal analysis, since one can better visualize the DC bias point levels. The drain lies at a higher DC potential than the source.

Since there are many different ways to configure the source, bias and load network of the stage, we distinguish the circuitry outside the dashed line from the core of the stage, which is essentially just the MOSFET with its ac-grounded gate. This distinction makes the results derived below more modular and generally applicable. For instance, when we derive expressions for small-signal port resistances, we distinguish between the core components (R_{in} and R_{out} shown in Figure 4-7) and any resistances in the source, bias and load network that appear in parallel. Also, depending on how the CG circuit is used, the output variable of interest may be either the voltage at the output or the current that flows into the output network. For the discussion in this chapter, we assume that the intended output is the voltage v_{OUT} as indicated in Figure 4-7. With this choice, the (low-frequency) gain of the circuit v_{out}/i_s is a transresistance.

The detailed analysis below consists of several parts. First, we will establish basic expressions for the circuit's operating point, and derive the conditions for MOSFET operation in the saturation region. Next, we will analyze the CG circuit core in terms of its port resistances, which will show that it is most appropriately modeled as a current amplifier. Finally, the obtained results are extended to include high-frequency effects due to capacitive elements.

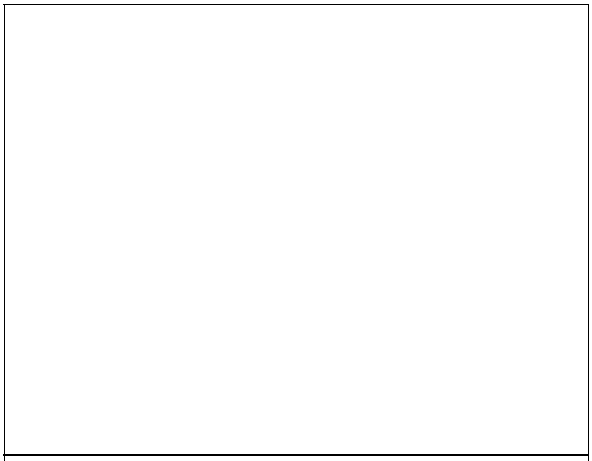


Figure 4-8: Simplified schematic of the CG amplifier for bias point calculations.

4-3-1 Bias Point Analysis

To determine the amplifier's node voltages at the operating point, we consider the circuit with the small-signal source i_s and its (small-signal) source resistance r_s removed (see Figure 4-8). From this circuit we see that $I_D = I_B$, and thus

$$V_{OUT} = V_{DD} - I_B R_D \tag{4.7}$$

A more complex analysis is needed to find the voltage at the source node, V_S. We begin by recognizing that

$$V_S = V_B - V_{GS} \tag{4.8}$$

Assuming that the MOSFET operates in the saturation region (which is subject to verification), we know that

$$I_D = I_B = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{GS} - V_{Tn})^2(1 + \lambda_n V_{DS}) \tag{4.9}$$

where

$$V_{Tn} = V_{TOn} + \gamma_n(\sqrt{2\phi_f + V_{SB}} - \sqrt{2\phi_f}) \tag{4.10}$$

and $V_{SB} = V_S$. Neglecting channel length modulation (i.e., assuming $\lambda_n V_{DS} \cong 0$), we can solve Eq. (4.9) for V_{GS}

$$V_{GS} = V_{Tn} + \sqrt{\frac{I_B}{\frac{1}{2}\mu_n C_{ox}\frac{W}{L}}} = V_{Tn} + V_{OV} \tag{4.11}$$

Substituting back into Eq. (4.8), we have

$$V_S = V_B - V_{Tn}(V_S) - \sqrt{\frac{I_B}{\frac{1}{2}\mu_n C_{ox}\frac{W}{L}}}V \tag{4.12}$$

Unfortunately this is a transcendental equation and obtaining a precise solution will require numerical iterations (using Eq. (4.10) with $V_{SB} = V_S$). However, since the dependence of V_{Tn} on V_S is relatively weak, a calculation with one or two iterations tends to give a satisfactory answer for the purpose of hand analysis. This is illustrated in Example 4.3 below.

Once V_S is computed, we can ensure that the MOSFET operates in saturation by checking the following inequality

$$V_{DS} = V_{OUT} - V_S > V_{DSsat} = V_{GS} - V_{Tn} \tag{4.13}$$

Example 4-3: Bias Point Calculation for a Common Gate Stage.

Consider the common-gate circuit shown Figure 4-8 with the following parameters: $V_{DD} = 5$ V, $V_B = 2.5$V, $I_B = 400$ µA, $R_D = 3$ kΩ. For the MOSFET, assume $W = 100$ µm, $L = 1$ µm, and the standard technology parameters given in Table 4-1. Compute V_{OUT}, V_S, as well as V_{DS} and $V_{OV} = V_{GS} - V_{Tn}$ of the transistor.

SOLUTION

Using Eq. (4.7) we can directly compute

$$V_{OUT} = V_{DD} - I_B R_D = 5V - 1.2V = 3.8V$$

Next, we find

$$V_{GS} - V_{Tn} = \sqrt{\frac{I_B}{\frac{1}{2}\mu_n C_{ox}\frac{W}{L}}} = 0.4V$$

In order to determine V_S, we begin by first ignoring the backgate effect, i.e., we evaluate Eq. (4.12) assuming $V_{Tn} = V_{T0n}$.

$$V_S = V_B - V_{T0n} - \sqrt{\frac{I_B}{\frac{1}{2}\mu_n C_{ox}\frac{W}{L}}} = 2.5V - 0.5V - 0.4V = 1.6V$$

Using this estimate of V_S, we can now compute an estimate of the actual threshold voltage with backgate effect included. By evaluating Eq. (4.10), we find

$$V_{Tn} = 0.5\text{V} + 0.6\text{V}^{1/2}(\sqrt{0.8\text{V}+1.6\text{V}} - \sqrt{0.8\text{V}}) = 0.893\text{V}$$

and thus

$$V_S = 2.5V - 0.893V - 0.4V = 1.207V$$

This result for V_S, which was obtained using only one iteration, differs from the exact solution $V_S = 1.273$ V (obtained through computer simulations or further iterations) by only –67 mV (–5.2%).

With the above numbers, we have $V_{DS} = V_{OUT} - V_S = 3.8$ V – 1.207 V = 2.59 V > $V_{GS} - V_{Tn} = 0.4$ V. Therefore, the transistor indeed operates in the saturation region, as assumed initially.

4-3-2 First Pass Low-Frequency Analysis

In order to establish symbolic expressions for the stage's small-signal transfer characteristics, we will first ignore capacitive elements and focus on the low-frequency behavior. Based on the circuit of Figure 4-7, we can thus construct the low-frequency small-signal model shown in Figure 4-9. Note here

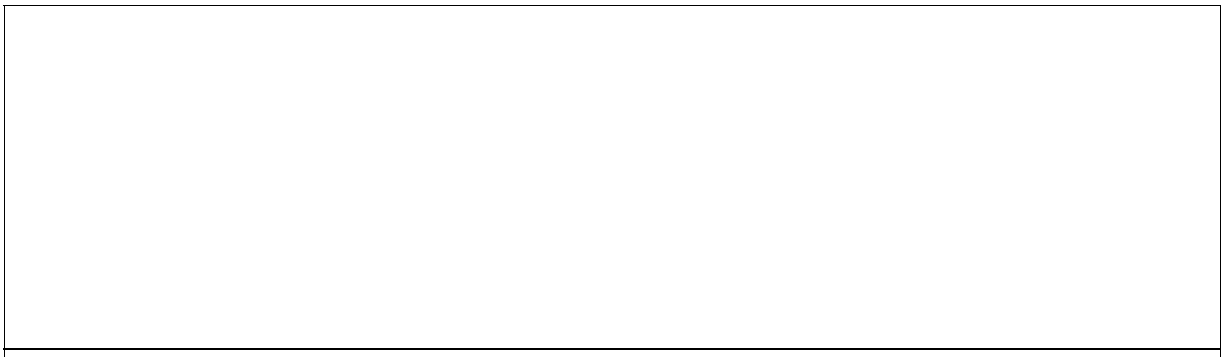


Figure 4-9: Low-frequency small-signal model of the CG stage.

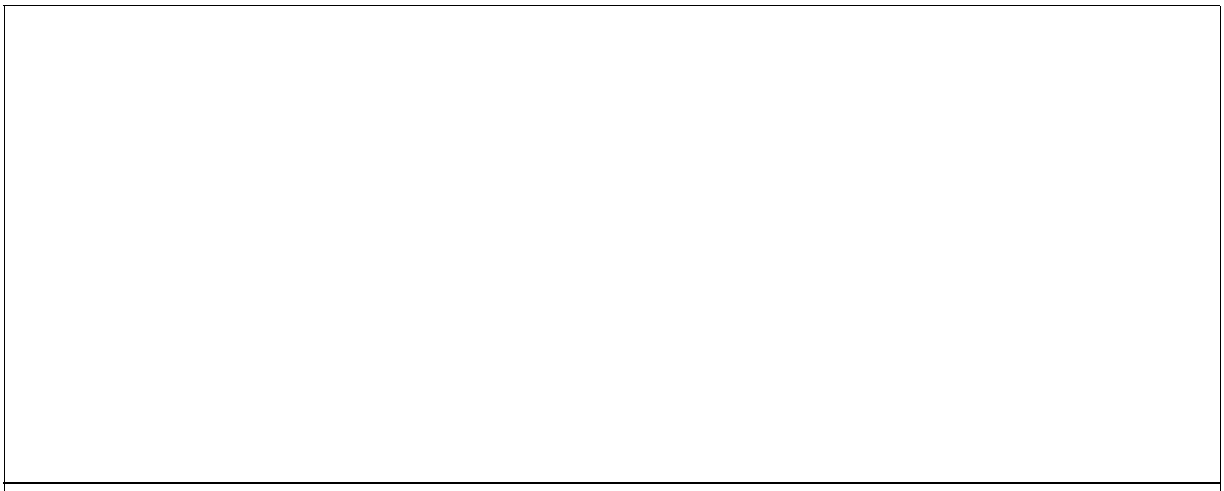


Figure 4-10: Modified representations of the CG small-signal circuit model. (a) Circuit with combined g_m and g_{mb} generators. (b) Split of controlled current source. (c) Representation of left side current generator using a resistor. (d) Approximate circuit model neglecting r_o.

that, as already explained in Chapter 2, the DC bias current source was removed, and the DC voltage sources were replaced by a connection to ground.

In principle, we could use the circuit of Figure 4-9 directly to carry out a detailed analysis. However, it turns out that it is beneficial to first walk through a few circuit simplifications that will help reduce the algebraic effort and also provide some qualitative insight. First, we note that since the gate and the bulk of the MOSFET are small-signal grounds, it follows that $v_{gs} = -v_s$ and $v_{bs} = -v_s$. Thus, the current generators g_m and g_{mb} effectively add since they are controlled by the same voltage ($-v_s$). For notational convenience, we therefore define $g'_m = g_m + g_{mb}$.The resulting circuit is shown in Figure 4-10(a) (with the controlled source rotated for convenience)

As a next step, we split the controlled current source into two elements as depicted in Figure 4-10(b). Since this change does not alter the summation of currents at the two nodes, the obtained circuit is equivalent to that of Figure 4-10(a). Next, we recognize that the controlled source on the left side is simply a resistor with value $1/g'_m$. This is true because for this element, we have $r = v/i = v_s/(g'_m v_s) = 1/g'_m$.

In the resulting circuit of Figure 4-10(c), let us assume for the time being that the resistance r_o has negligible impact on the operation of the CG amplifier. In this case, shown in Figure 4-10(d), the core of the CG stage perfectly conforms with the unilateral current amplifier two-port model discussed in Section 1-3. Specifically, note that the circuit has low input resistance (for example, if $g'_m = 10$ mS $\Rightarrow 1/g'_m = 100\ \Omega$) and high output resistance ($R_{out} \rightarrow \infty$), as desired for a current amplifier. The current gain (A_i) of the two-port is equal to one; any current that enters the core part of the CG stage passes through it unchanged. This result makes intuitive sense also from the original circuit in Figure 4-7. Any change in the MOSFET's source current must also appear at the drain side.

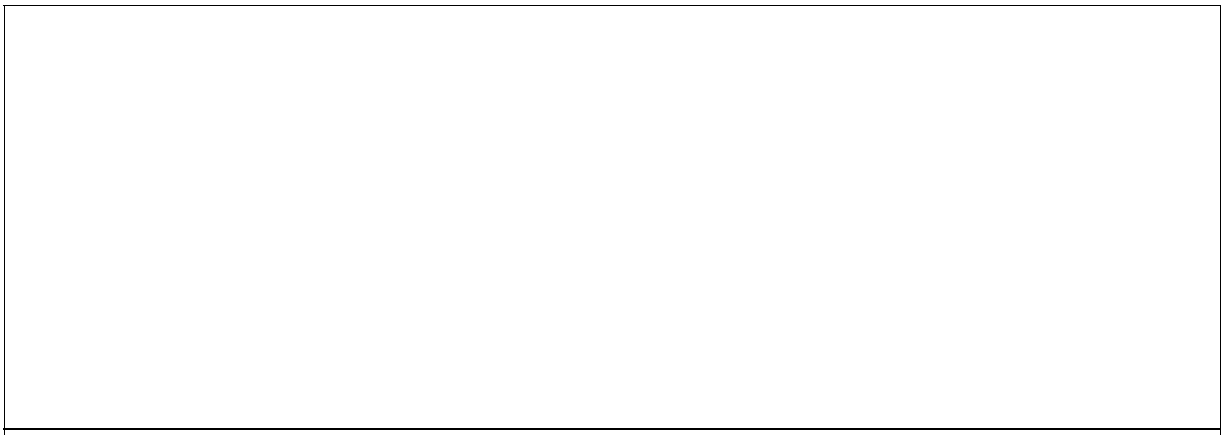


Figure 4-11: Small-signal model used to calculate (a) the input resistance (R_{in}), and (b) the output resistance (R_{out}) of the intrinsic CG circuit.

A secondary conclusion to draw from this first pass analysis concerns the backgate connection of the transistor. If the bulk terminal were not connected to ground, but instead tied to the source of the transistor (which is possible for the p-channel version of the circuit in our n-well technology), we would have $v_{bs} = 0$ and the g'_m term would reduce to g_m. This is undesired since the amplifier's input resistance will increase correspondingly (on the order of 15% – 35%, see Example 4-2). Thus, in a CG amplifier, it is typically advantageous to leave the bulk node ac-grounded, rather than tying it to the source. Note that this also reduces the parasitic capacitance at the source, as discussed in Section 4-2-1.

4-3-3 Detailed Low-Frequency Analysis

We will now carry out a more detailed analysis that refines the result obtained in the previous sub-section. Specifically, we wish to perform a full-analysis with the MOSFET's r_o included to see if, and precisely when, this resistor can be neglected.

We begin by computing the input resistance (R_{in}) of the CG circuit core using the full circuit from Figure 4-10(c), redrawn in Figure 4-11(a) with a test voltage source included. From this setup, the input resistance is found using the procedure described in Section 1-3-3. Specifically note that R_D is included in the analysis since the circuit is bilateral.

Now, writing KVL at the input and output nodes gives

$$i_t = g'_m v_t + \frac{v_t - v_{out}}{r_o} \tag{4.14}$$

$$0 = -g'_m v_t - \frac{v_t - v_{out}}{r_o} + \frac{v_{out}}{R_D} \tag{4.15}$$

By solving this system of equations for v_t and i_t, we obtain

$$R_{in} = \frac{v_t}{i_t} = \frac{1 + \dfrac{R_D}{r_o}}{g'_m + \dfrac{1}{r_o}} \tag{4.16}$$

Assuming $r_o >> R_D$ and $r_o >> 1/g'_m$, the input resistance becomes approximately

$$R_{in} \cong \frac{1}{g'_m} \tag{4.17}$$

which is identical to the value postulated in the approximate model of the previous sub-section. The approximation applied to the denominator of Eq. (4.16) is always valid, since $g'_m r_o >> 1$ for a typical MOSFET. The approximation made in the numerator may not hold when the drain is terminated with a very large incremental resistance, as for instance found in a current source. In this case, care must be taken to use the exact numerator of Eq. (4.16).

Next, to calculate the output resistance (R_{out}), we set the small-signal input current source i_s equal to zero (open circuit) but leave the effect of its source resistance (r_s) in place as shown in Figure 4-11(b). As before, we place a test voltage source at the port of interest, and write KCL for the two nodes of the circuit.

$$0 = g'_m v_s + \frac{v_s - v_{out}}{r_o} + \frac{v_s}{r_s} \tag{4.18}$$

$$i_t = -g'_m v_s - \frac{v_s - v_{out}}{r_o} \tag{4.19}$$

After solving for v_t and i_t, we obtain

$$R_{out} = \frac{v_t}{i_t} = r_o + r_S[1 + g'_m r_o] \tag{4.20}$$

Utilizing the fact that $g'_m r_o >> 1$, we can approximate

$$R_{out} \cong r_o[1 + g'_m r_s] \tag{4.21}$$

Under the condition that $g'_m r_s >> 1$, the expression further simplifies to

$$R_{out} \cong g'_m r_o r_s \tag{4.22}$$

From this result, we note that the source resistance r_s is multiplied by a term that is on the order of the intrinsic voltage gain of the transistor ($g_m r_o$), which is typically greater than 100 in our technology (see Section 2-3-2). Thus, a CG stage can essentially be used to turn a current source with moderate source resistance (r_s) into a "better" current source with very high source resistance. This feature is widely used in a variety of circuit configurations, some of which will be discussed later in this module.

Using the above results, we can now construct a complete circuit model that includes the CG core as a unilateral current amplifier two-port (see Figure 4-12). Using this model, the transresistance gain from the input current source (i_s) to the circuit's output voltage (v_{out}) can be written as

$$\frac{v_{out}}{i_s} = \frac{i_{in}}{i_s} \cdot \frac{v_{out}}{i_{in}} = \frac{r_s}{r_s + R_{in}} \cdot \left(\frac{1}{R_D} + \frac{1}{R_{out}}\right)^{-1} \tag{4.23}$$

where the first term represents the current divider formed by r_s and R_{in} and the second term is the parallel combination of R_D and R_{out}. For a typical application of this circuit, where $R_{in} << r_s$ and $R_D << R_{out}$, Eq. (4.23) simplifies to

$$\frac{v_{out}}{i_s} \cong R_D \tag{4.24}$$

In words, the transresistance of the circuit is approximately equal to the drain resistance. This makes intuitive sense since the CG core acts a current amplifier with a gain near unity, that is, the CG stage absorbs all (or most) of the current from the input source, and it passes this current to the termination resistance (R_D), causing a proportional change in the output voltage.

When employing the model of Figure 4-12, it is important to remember that we are approximating a bilateral circuit using a unilateral model. We have investigated the resulting error for this particular configuration in Example 1-3 and found that the unilateral model is guaranteed to be accurate as long as the coupling resistance between the input and output ports (R_2 in Example 1-3 and r_o in the analyzed CG circuit) is larger than the termination resistance (R_L in Example 1-3 and R_D in the analyzed CG circuit). Thus, the result stated in Eq. (4.23) holds with good accuracy as long as $R_D << r_o$, a condition that is satisfied in many relevant applications.

Figure 4-12: Small-signal unilateral two-port model of the CG stage.

4-3-4 High-Frequency Analysis

In order to model the frequency response of the CG amplifier, we now include the capacitances of the MOSFET into the small-signal model constructed in the previous section (see Figure 4-13).

Figure 4-13: High-frequency small-signal model of the CG stage.

Since both the gate and bulk of the MOSFET are ac-grounded, C_{gs} and C_{sb}, as well as C_{gd} and C_{db} appear from source and drain, respectively, to ground. For notational convenience, we abbreviate $C_{gs} + C_{sb} = C_s$ and $C_{gd} + C_{db} = C_d$ in the following treatment.

One way to analyze the circuit of Figure 4-13 is to re-use the results obtained in the low-frequency analysis, but now with C_s and C_d added in parallel to r_s and R_D, respectively. In this spirit, we begin by considering the effect of the added capacitances

on the circuit's input impedance, Z_{in}. Re-using the result from Eq. (4.16), we can write

$$Z_{in} \cong \frac{1 + \frac{(R_D \parallel sC_d)}{r_o}}{g'_m} \parallel \frac{1}{sC_s} \tag{4.25}$$

As long as $R_D << r_o$, regardless of the added term due to C_d, the numerator can be approximated as unity. Therefore, it follows that

$$Z_{in} \cong \frac{1}{g'_m} \parallel \frac{1}{sC_s} \tag{4.26}$$

In words, this result simply says that Z_{in} is typically well approximated by the parallel impedance connection of $1/g'_m$ and C_s.

Finding the (exact) output impedance (Z_{out}) proves algebraically more difficult. However, provided that $g'_m r_s >> 1$ and for frequencies below the MOSFET's cutoff frequency ω_T, it follows that (see problem P4.6)

$$Z_{out} \cong g'_m r_o r_s \parallel \frac{g'_m r_o}{sC_s} \parallel \frac{1}{sC_d} \tag{4.27}$$

Thus, Z_{out} consists of the parallel combination of: (1) r_s, increased by the factor $g'_m r_o$. This is simply R_{in} from the low-frequency analysis. (2) The explicit capacitance present at the output (C_d), and (3) the reactance of the source side capacitance (C_s) increased by the factor $g'_m r_o$. From second term, we see that the MOSFET not only boosts the resistance of r_s seen from the output, but also the reactance of C_s. Since usually $C_s/g'_m r_o << C_d$, we can safely neglect the second term of Eq. (4.27) and conclude

$$Z_{out} \cong g'_m r_o r_s \parallel \frac{1}{sC_d} \tag{4.28}$$

In summary, the overall conclusion is that, to first order, the drain and source side capacitances can simply be added to the ports of the already obtained low-frequency model of Figure 4-12. Thus, we have once again arrived at a relatively simple unilateral model for the circuit that holds for high frequencies. The corresponding circuit model is drawn out in Figure 4-14, incorporating the assumptions $g'_m r_s >> 1$ and $R_D << r_o$. Also, for generality, we have included a load capacitor (C_L) and a load resistor (R_L). Note that whenever RL carries a DC current, this resistor must also be included in the bias point calculations of the circuit.

Similar to the low-frequency analysis of the previous subsection, the overall transfer function of the circuit in Figure 4-14 is found by considering the current division at the input node, and the parallel impedance at the output node

$$\begin{aligned} \frac{v_{out}}{i_s} &= \frac{i_{in}}{i_s} \cdot \frac{v_{out}}{i_{in}} = \frac{g'_m}{g'_m + sC_s} \cdot \left(\frac{1}{R_D} + \frac{1}{R_L} + sC_d + sC_L\right)^{-1} \\ &= [R_D \| R_L] \cdot \frac{1}{1 + s\frac{C_s}{g'_m}} \cdot \frac{1}{(1 + s[R_D \| R_L])(C_d + C_L)} \\ &= [R_D \| R_L] \cdot \frac{1}{1 - \frac{s}{p_1}} \cdot \frac{1}{1 - \frac{s}{p_2}} \end{aligned} \tag{4.29}$$

In this expression, the leading term $[R_D \| R_L]$ corresponds to the low-frequency transresistance and $p_1 = -g'_m/C_s$ and $p_2 = -1/[(R_D \| R_L)(C_d + C_L)]$ are the poles of the circuit. The pole p_1 is associated with the input of the circuit and captures the frequency dependence of the current transfer from the input current to the drain current of the transistor (i_x/i_{in} in Figure 4-14). The corresponding pole frequency g'_m/C_s is very close to ω_T (the cutoff frequency of the transistor) and will rarely limit the bandwidth of the overall circuit. The pole p_2 depends on the component values in the load network. If the added C_L is small, this pole will typically also lie at a very high frequency..

Figure 4-14: Approximate unilateral two-port model of the CG transresistance amplifier for high-frequencies.

From the result of Eq. (4.29), it is important to note that the circuit has two distinct and conveniently isolated poles that are not "entangled" as for instance in the frequency response of the CS stage [see Eq. (3.38)]. This is a direct result of the fact that the input and output networks of the model in Figure 4-14 are decoupled, i.e., the networks that define the time constants on each side of the circuit do not interact. Note also that for approximate 3-dB bandwidth calculations, the method of open-circuit time constants can be applied. This will give the sum of the two time constants that make up the poles contained in Eq. (4.29).

Example 4-4: Bandwidth Calculation for a Common-Gate Stage.

Consider the CG circuit of Figure 4-7 using the same parameters used in Example 4-3: $I_B = 400\,\mu A$, $R_D = 3\,k\Omega$, $W = 100\,\mu m$, and $L = 1\,\mu m$. Using the bias point information from Example 4-3, compute all device capacitances and estimate the bandwidth of the circuit's transfer function using the method of open-circuit time constants.

SOLUTION

From Example 4-3, we know that $V_{OUT} = 3.8$ V and $V_S = 1.273$ V. Using Eq. (3.15), and Eq. (3.32) to Eq. (3.34) we can therefore compute all device capacitances.

$$C_{gs} = \frac{2}{3}(100 \cdot 1\mu m^2)\left(2.3\frac{fF}{\mu m^2}\right) + 100\mu m\left(0.5\frac{fF}{\mu m}\right) = 203.3fF$$

$$C_{gd} = 100\mu m\left(0.5\frac{fF}{\mu m}\right) = 50fF$$

Evaluating Eq. (3.34), using the source junction bias voltage of $V_{SB} = 1.273$ V yields

$$C_{sb} = 54.3fF$$

The drain junction has a reverse bias voltage of $V_{DB} = V_{OUT} = 3.8$ V, Evaluating Eq. (3.33) with this value and the given parameters gives

$$C_{db} = 37.1fF$$

The total capacitances for the model in Figure 4-14 are therefore

$$C_s = C_{gs} + C_{sb} = 257.6fF$$

$$C_d = C_{gd} + C_{db} = 87.1fF$$

Now, using $V_{OV} = 0.4$ V from Example 4-3, we have

$$g_m = \frac{2I_D}{V_{OV}} = 2mS$$

Furthermore, using Eq. (4.6)

$$g_{mb} = \frac{2mS \cdot 0.6V^{1/2}}{2\sqrt{0.8V + 1.273V}} = 0.417mS$$

and therefore

$$\frac{1}{g'_m} = \frac{1}{2mS + 0.417mS} = 413.8\Omega$$

The time constants of the circuit are

$$\tau_{so} = 413.8\Omega \cdot 257.6fF = 107ps$$

$$\tau_{do} = 3k\Omega \cdot 87.1fF = 261ps$$

The bandwidth estimate is therefore

$$f_{3dB} = \frac{1}{2\pi} \cdot \frac{1}{107ps + 261ps} = 432MHz$$

By comparing the two time constants, we see that the bandwidth is limited by the drain network. This situation will be even more pronounced when an external load capacitance C_L is added to the output node.

4-4 Analysis of the Common-Drain Stage

A practical configuration of a common-drain (CD) circuit is shown in Figure 4-15. Similar to the CG stage discussed in the previous section, other configurations for the input and output network exist; we will analyze the given circuit as a representative example.

Figure 4-15: Practical realization of a CD amplifier with biasing circuitry included.

4-4-1 Bias Point Analysis

For a first pass bias point analysis, consider the circuit in Figure 4-16 first with R_L disconnected, ensuring that the MOSFET drain current $I_D = I_B$. Since no current flows into the gate of the MOSFET, we have $V_{IN} = V_S$. The output voltage quiescent point in this circuit can thus be calculated as already analyzed

Figure 4-16: Simplified schematic of the CD amplifier for bias point calculations.

in the context of the CG stage, using Eq. (4.13) with V_{IN} and V_{OUT} substituted for V_B and V_S, respectively.

$$V_{OUT} = V_{IN} - V_{Tn}(V_{OUT}) - \sqrt{\frac{I_B}{\frac{1}{2}\mu_n C_{ox}\frac{W}{L}}} \tag{4.30}$$

If V_B at the bottom of R_L is chosen equal to V_{OUT} as given in Eq. (4.30), no current will flow in this resistor when re-connected and $I_D = I_B$ is maintained. In cases where this condition is not met, the MOSFET's drain current differs from I_B, but can be computed using iterative calculations. For simplicity in our discussion, we will assume that V_B (or V_{IN}) is properly adjusted such that $I_D = I_B$ is guaranteed.

As far as the MOSFET's operating region is concerned, it is interesting to note that the device will essentially always operate in the saturation region. This is the case since

$$V_{DS} = V_{DD} - V_{OUT} = V_{DD} - (V_{IN} - V_{GS}) \tag{4.31}$$

and thus

$$V_{DS} - (V_{GS} - V_{Tn}(V_{OUT})) = V_{DD} - (V_{IN} - V_{Tn}(V_{OUT})) \tag{4.32}$$

is greater than zero as long as $V_{IN} < V_{DD} + V_{Tn}(V_O)$, which is almost always the case in a practical realization.

4-4-2 Low-Frequency Analysis

To analyze the circuit's transfer characteristics at low frequencies, we draw its small-model (without any capacitances) as shown in Figure 4-17(a). While this circuit can be analyzed directly by writing nodal equations, it pays once again to think through a few basic equivalent transformations. First, note that we can split the source controlled by v_{gs} into two components, $g_m v_g$ and $-g_m v_s$, where $v_g = v_{in}$ and $v_s = v_{out}$ [see Figure 4-17(b)]. Similarly, the backgate generator is split into $g_{mb}v_b$ and $-g_{mb}v_s$ where $v_b = 0$ (which means that this source can be discarded) and $v_s = v_{out}$. Finally, recognizing that the two sources controlled by v_{out} are simply resistors of value $1/g_m$ and $1/g_{mb}$, we arrive at the final circuit in Figure 4-17(c).

Figure 4-17: (a) CD small-signal, low-frequency circuit model. (b) Split of controlled current source. (c) Representation of output-controlled current generators using a resistor.

From this simplified representation, we can immediately see

$$R_{in} \to \infty \tag{4.33}$$

$$R_{out} = r_o \left|\left| \frac{1}{g_m} \right|\right| \frac{1}{g_{mb}} \cong \frac{1}{g_m + g_{mb}} = \frac{1}{g'_m} \tag{4.34}$$

Consequently, the CD circuit core most closely resembles the properties of a voltage amplifier (high R_{in}, low R_{out}), and is thus most appropriately modeled using the two-port representation shown in Figure 4-18. The final parameter needed for this model is the open-circuit voltage gain A_v. To find A_v, we apply an ideal test voltage to the circuit of Figure 4-17(c) (i.e., a voltage source with $R_S = 0$) and measure the voltage at the

open-circuited output port (R_L disconnected). Neglecting r_o in this analysis, the reader can prove that

Figure 4-18: Unilateral two-port model for the CD stage at low frequencies.

$$A_v = \frac{g_m}{g_m + g_{mb}} = \frac{1}{1 + \frac{g_{mb}}{g_m}} < 1 \qquad (4.35)$$

With g_{mb}/g_m values on the order of 20% (see Example 4-2), the open circuit voltage gain of the analyzed n-channel CD stage is on the order of 0.8.

A_v can approach unity when the source is tied to the bulk of the MOSFET. In this case, $v_{bs} = 0$, and the $1/g_{mb}$ resistor in Figure 4-17(c) is eliminated. Note that for the technology considered in this module, this option is available only for the p-channel version of the circuit. Accordingly, Figure 4-19 summarizes the three different CD configurations that are possible in the assumed technology. In practice, the designer will decide case by case if a source-bulk connection (for the p-channel circuit) is advantageous in the intended application. A disadvantage of the configuration in Figure 4-19(c) is that the bulk to substrate (C_{bsub}) capacitance contributes to the capacitive load of the stage (see Section 4-2 and also Figure 4-22).

Figure 4-19: Possible configurations for CD amplifiers.

Since the CD stage achieves a positive voltage gain near unity, this circuit is often called **source follower**. The output (the source of the transistor) caries a signal that closely follows the applied input voltage.

Example 4-5: Low-Frequency Analysis of a CD Amplifier.

Consider the CD stage shown in Figure Ex4-5 with the following parameters: $V_{DD} = 5$ V, $V_S = 2.5$V, $I_B = 400$ μA, $R_S = 100$ kΩ and $R_L = 5$ kΩ. For the MOSFET, assume $W = 100$ μm, $L = 1$ μm, and the standard technology parameters given in Table 4-1. Assume that V_B is adjusted such that no DC bias current flows in R_L.Calculate A_v, R_{out} and the overall voltage gain v_{out}/v_s.

Figure Ex4-5

SOLUTION

Since the transistor is sized and biased exactly as in the CG stage of Examples 4-3 and 4-4, we know that $V_{OUT} = 1.273$ V, $g_m = 2$ mS and $g_{mb} = 0.417$ mS. The resulting open-circuit voltage gain is

$$A_v = \frac{g_m}{g_m + g_{mb}} = 0.828$$

$$\text{and } R_{out} = \frac{1}{g_m + g_{mb}} = 413.8\Omega\,.$$

The overall voltage gain can be computed using Figure 4-18

$$A'_v = \frac{v_{out}}{v_s} = A_v \frac{R_L}{R_{out} + R_L} = 0.764$$

Based on this result, the reader may wonder why the analyzed circuit is useful, as it provides no voltage gain. The benefit of using the stage becomes apparent by comparing to the case where the stage is omitted, and R_L is directly driven by the source. In this case, the voltage gain from v_s to v_{out} is given by the resistive divider between R_S and R_L

$$\frac{v_{out}}{v_s} = \frac{R_L}{R_S + R_L} = \frac{5k\Omega}{100k\Omega + 5k\Omega} = 0.048$$

As we see from this calculation, even though the CD stage does not have voltage gain, it allows us to interface a relatively small load resistance to a source with high source resistance.

4-4-3 High-Frequency Analysis

To predict the high-frequency behavior of the CD stage, we consider the circuit shown in Figure 4-20. This circuit was constructed using the model of Figure 4-17(c), with all relevant MOSFET capacitances included. Note that C_{db} is not part of this model since this capacitance is connected between V_{DD} and GND (bulk node) and does not influence the signal path. For notational convenience, we define $C_{tot} = C_L + C_{sb}$ and $R_{tot} = R_L \parallel 1/g'_m$.

Figure 4-20: Circuit model for high-frequency analysis of the CD stage.

The following analysis in partitioned into several steps. First, we will carry out a full KCL-based derivation of the frequency response from first principles. To simplify, we will then customize the obtained general expressions for a specific and typical range of component values. This will greatly simplify the result and also make it amenable for an intuitive interpretation using the Miller theorem. Finally, for the simplified circuit, we will inspect the input and output impedances to construct a two-port model that is valid up to high frequencies.

We begin by noting that the model of Figure 4-20 resembles the CS circuit analyzed in Section 3-3-3: the input and output network are coupled through a capacitor, they contain and RC network on each side, and also use a controlled current source in the output network. The main difference lies in the polarity of the controlled source, which reflects the fact that the CD stage is a non-inverting amplifier, whereas a CS stage is an inverting amplifier (i.e., the low-frequency voltage gain is a negative number). This difference has a profound impact on the Miller amplification of the coupling capacitor between the input and output (see Section 3-3-5). We will analyze this circuit in more detail below, after interpreting the general KCL-based result.

Based on the general similarity with the circuit analyzed in Section 3-3-3, we expect that the full transfer function of the CD stage resembles the 2-pole, 1-zero response expressed in Eq. (3.38). Indeed, the reader can prove (see Problem 4.9) that carrying out a full KCL-based analysis of the circuit in Figure 4-20 yields

$$\frac{v_{out}}{v_s} = A'_v \frac{\left(1 + s\frac{C_{gs}}{g_m}\right)}{1 + b_1 s + b_2 s^2} \tag{4.36}$$

where

$$b_1 = R_s C_{gs}(1 - A'_v) + R_s C_{gd} + R_{tot} C_{gs} + R_{tot} C_{tot} \tag{4.37}$$

$$b_2 = R_s R_{tot}(C_{gs} C_{gd} + C_{gs} C_{tot} + C_{gd} C_{tot}) \tag{4.38}$$

and

$$A'_v = \frac{g_m}{g'_m + \frac{1}{R_L}} \tag{4.39}$$

is the overall low-frequency voltage gain from v_s to v_{out}

Once again, while this result is algebraically complex, it is possible to draw a few basic conclusions. First note that the zero in the transfer function occurs at approximately ω_T, the MOSFET's cutoff frequency. Hence, the zero will only rarely be relevant for the behavior of the circuit within typical frequencies of interest. Second, if we assume that a dominant pole condition exists, the bandwidth of the circuit will be approximately equal to $1/b_1$ (see Section 3-3-4). Note also that the derived b_1 term can alternatively be found using the method of open-circuit time constants (see Section 3-4 and Problem 4-5).

Inspecting Eq. (4.37) further, we can consider several approximations. First note that R_{tot} is a very small resistance < $1/g_m$. Consequently, the time constants $R_{tot}C_{gs}$ and $R_{tot}C_{sb}$ (contained in $R_{tot}C_{tot}$) are typically not dominant, and we can approximate

$$b_1 \cong R_s C_{gs}(1 - A'_v) + R_s C_{gd} + R_{tot} C_L \tag{4.40}$$

To simplify further, we need to make assumptions about the relative values of R_s and C_L. This will lead to a result that is no longer general, but useful to capture an important sub-set of applications for the CD stage. Specifically, consider the circuit

shown in Figure 1-5, where the output resistance of a CS stage defines R_s for the subsequent CD stage. In such an scenario, R_s will be relatively large and certainly much larger than R_{tot}. Assuming this use case, we can argue that the time constants associated with R_s will dominate unless a very large load capacitance (greater than C_{gs}) is connected to the output of the CD stage. Excluding the latter scenario, we arrive at the final approximation

$$b_1 \cong R_s C_{gs}(1 - A'_v) + R_s C_{gd} \tag{4.41}$$

With this approximation, the circuit's bandwidth is

$$\omega_{3dB} \cong \frac{1}{b_1} \cong \frac{1}{R_s C_{gs}(1 - A'_v) + R_s C_{gd}} \tag{4.42}$$

While this result does not hold in general, it applies to an interesting subset of applications, and it also can be interpreted and understood intuitively. Specifically, note that C_{gs} is modified by a Miller effect multiplier as already seen in Eq. (3.51). We can explain the Miller effect multiplier intuitively using Figure 4-21. First note that under the assumption that C_L is small, the voltage gain from v_{in} to v_{out} is constant up to very high frequencies. Hence, this gain can be approximated by its low-frequency value A'_v (we call this "Miller approximation," as explained in Section 3-3-5). According to Eq. (3.51), the equivalent shunt input capacitance due to C_{gs} is therefore simply given by $C_{gs}(1 - A'_v)$. Since A'_v is positive, the portion of C_{gs} that is seen from the input is *reduced* by the Miller effect. Circuit designers sometimes refer to this effect as **bootstrapping**. The effective reduction of C_{gs} in the CD stage strongly contrasts the undesired multiplication of C_{gd} in a CS stage [see Eq. (3.56)]. Because of this difference, the bandwidth of a CD stage is often significantly larger than that of a CS stage.

Figure 4-21: Circuit model for explaining the Miller effect in a common-drain stage.

Example 4-6: Bandwidth estimate of a CD Amplifier.

Estimate the bandwidth of the CD stage in Figure Ex4-5 assuming the following parameters: R_S = 10 kΩ, R_L = 5 kΩ, g_m = 2 mS, g_{mb} = 0.417 mS (from Example 4-5) and C_L = 0, C_{gs} = 203 fF, C_{gd} = 50 fF, and C_{sb} = 54.3 fF (from Example 4-4). First perform an estimate on the exact b_1 term given in Eq. (4.37), then repeat with the approximate value of Eq. (4.41) and compute the discrepancy in percent.

SOLUTION

From Example 4-4, we know that A'_v = 0.764. R_{tot} is the parallel combination of $1/g_m$, $1/g_{mb}$ and R_L, which amounts to 382 Ω. Using Eq. (4.37), we can determine the exact value of b_1 as

$$\begin{aligned} b_1 &= R_s C_{gs}(1 - A'_v) + R_s C_{gd} + R_{tot} C_{gs} + R_{tot} C_{sb} \\ &= 240\text{ps} + 250\text{ps} + 78\text{ps} + 21\text{ps} = 589\text{ps} \end{aligned}$$

The corresponding bandwidth estimate is

$$f_{3dB} \cong \frac{1}{2\pi} \cdot \frac{1}{b_1} = \frac{1}{2\pi} \cdot \frac{1}{589\text{ps}} = 271\text{MHz}$$

According to the approximation of Eq. (4.41), we have

$$b_1 \cong R_s C_{gs}(1 - A'_v) + R_s C_{gd} = 240\text{ps} + 250\text{ps} = 490\text{ps}$$

Now the corresponding bandwidth estimate is

$$f_{3dB} \cong \frac{1}{2\pi} \cdot \frac{1}{b_1} = \frac{1}{2\pi} \cdot \frac{1}{490\text{ps}} = 325\text{MHz}$$

The percent difference between the two estimates is

$$\frac{325 - 271}{271} = 19.9\%$$

From this result, we see that the simple expression of Eq. (4.41) yields a reasonable first-order estimate of the circuit's bandwidth. The time constants associated with R_{tot} do not play a significant role in setting the circuit's bandwidth.

As a final step in our analysis, we now consider the input and output impedances of the CD stage (Z_{in} and Z_{out} in Figure 4-20). For Z_{in}, it is clear from the above treatment that

$$Z_{in} \cong \frac{1}{s([C_{gs}(1 - A'_v) + C_{gd}])} \tag{4.43}$$

for the given assumptions. An interesting situation arises when a p-channel CD stage with source-bulk tie and purely capacitive load is considered (see Figure 4-22). Since $A'_v \cong 1$ in this

circuit, the input capacitance is well approximated by C_{gd} alone, with no significant contribution from C_{gs}. This follows mathematically from Eq. (4.43), but more importantly makes intuitive sense. When $A'_v = 1$, the input and output nodes precisely follow the same AC voltage. This means that no AC current can flow through C_{gs}, and hence this capacitance becomes irrelevant.

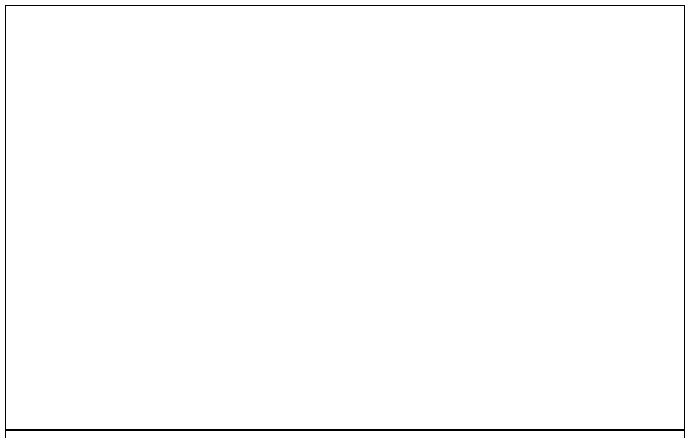


Figure 4-22: P-channel common-drain stage with source-bulk tie.

In order to find Z_{out}, further analysis is needed and we therefore consider the setup shown in Figure 4-23. First consider a special situation where $R_S = 0$. In this case, $v_{in} = 0$ and the controlled current source is inactive. Thus, we see by inspection

$$\begin{aligned} Z_{out} &= \frac{1}{g'_m} \Bigg|\Bigg| \frac{1}{s(C_{gs}+C_{sb})} = \frac{1}{g'_m + s(C_{gs}+C_{sb})} \\ &= \frac{1}{g'_m} \frac{1}{1+s\left(\dfrac{C_{gs}+C_{sb}}{g'_m}\right)} \cong \frac{1}{g'_m} \end{aligned} \tag{4.44}$$

Here, the final approximation is justified based on the fact that the pole in the preceding expression lies near ω_T and is therefore negligible in many cases. Consequently, for the special case of $R_S = 0$, the output impedance of a CD stage is purely resistive up to very high frequencies and closely approximated by $1/g'_m$.

For the case of finite R_S, we write KCL at the two nodes of the circuit and solve for $Z_{out} = v_t/i_t$. Approximating $g_m \cong g'_m$ and neglecting C_{sb} (as justified above) for algebraic simplicity, this yields

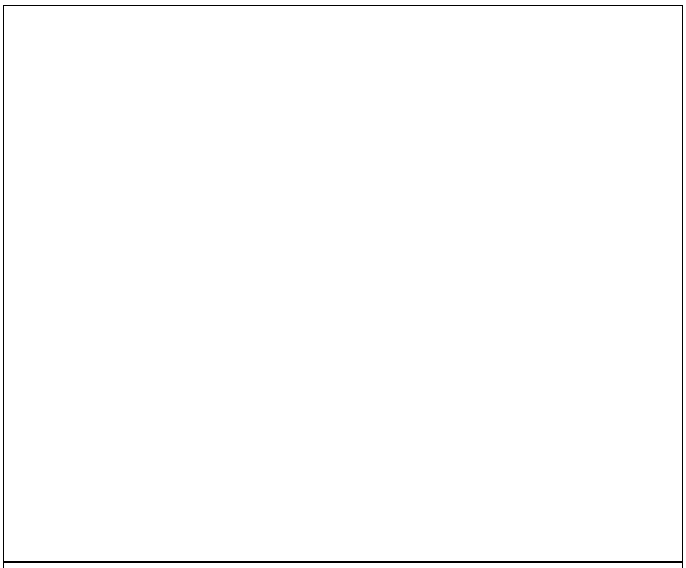


Figure 4-23: Circuit for finding Z_{out} of the common-drain stage.

$$\begin{aligned} Z_{out} &\cong \frac{1}{g'_m} \frac{1+s(C_{gs}+C_{gd})R_S}{1+s\left(\dfrac{C_{gs}}{g_m}+R_S C_{gd}\right)+s^2\left(\dfrac{R_S C_{gs} C_{gd}}{g_m}\right)} \\ &= \frac{1}{g'_m} \frac{\left(1-\dfrac{s}{z}\right)}{\left(1-\dfrac{s}{p_1}\right)\left(1-\dfrac{s}{p_2}\right)} \end{aligned} \tag{4.45}$$

Again, this result is algebraically complex, but can be used to gain some insight into the basic behavior. First, we note that the zero in this expression occurs at a frequency below the approximate bandwidth of the circuit [see Eq. (4.42)]. If a dominant pole condition exists, we see that the dominant pole will occur approximately at $1/R_S C_{gd}$ (neglecting $C_{gs}/g_m = 1/\omega_T$). That is, the first pole occurs *after* the zero in the overall impedance function. This situation is sketched out in Figure 4-24. With the zero occurring first, there is a region in which $|Z_{out}|$ increases with frequency. This behavior is characteristic of an inductor (the reactance of an inductor is given by $X = \omega L$). Consequently, the output impedance [Eq. (4.45)] can be accurately represented by an equivalent *RLC* network, see Reference 1 and Problem 4.16. In some practical scenarios, the inductive nature of the output impedance can be problematic due to possible signal oscillations ("ringing") in response to step-like signals. However, when properly tuned, the inductive component can be used to extend the bandwidth in certain situations. These interesting design topics are unfortunately beyond the scope of the introductory treatment of this module.

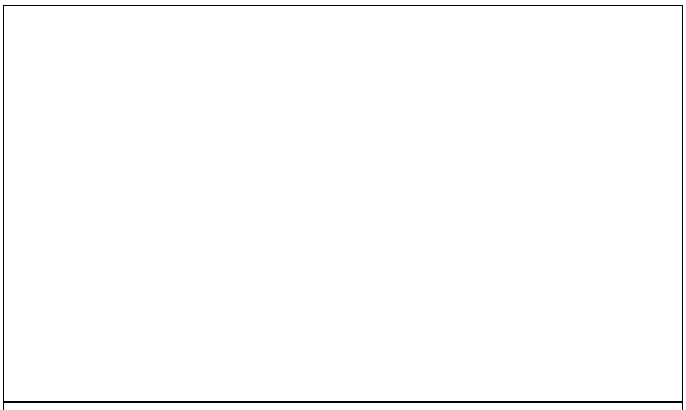


Figure 4-24: Frequency dependence of the common-drain stage's output impedance.

In summary, and for the purpose of the multi-stage circuit analysis in Chapter 6, an appropriate (but approximate) two-port model that includes the most relevant impedances is shown in Figure 4-25. This model is reasonably accurate up to frequencies close to the bandwidth of the stage when driven with a relatively large source resistance $R_S > 1/g_m$. Inductive effects are omitted for simplicity. When confronted with applications that necessitate a different set of approximations, the reader is encouraged to revisit the accurate transfer function result of Eq. (4.36) and re-customize it as needed for the specific use case.

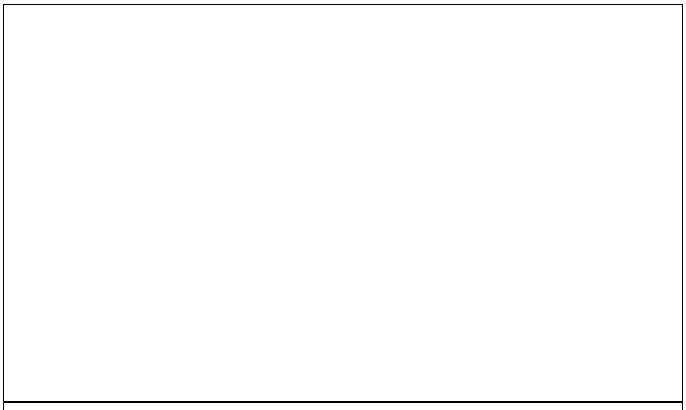


Figure 4-25: Two-port model of the CD stage, including relevant frequency dependent elements (assuming relatively large $R_S > 1/g_m$) and small C_L).

4-5 Application Examples of Common-Gate and Common-Drain Stages

In this section we will discuss a few common applications of the CG and CD stages. The discussion is mostly qualitative and meant to help solidify the qualitative understanding of the basic properties for all stage configurations discussed so far. Further analytical details for a subset of these applications is covered in Chapter 6.

4-5-1 CS-CG Cascade (Cascode Amplifier)

A commonly used circuit called the **cascode amplifier** combines a CS and CG stage as shown in Figure 4-26. This configuration comes with several benefits, both in terms of low- and high-frequency behavior. From a low-frequency perspective, a key benefit is that the resistance (R_{out}) looking into the drain of the CG device (M_2) is extremely high. This follows directly from Eq. (4.22), with r_s substituted by r_{o1} (the output resistance of MOSFET M_1), which leads to $R_{out} \cong r_{o1} \cdot g'_{m2} r_{o2}$. Furthermore, note that M_1 essentially operates as a transconductance stage, since its output is taken as a current that feeds into the CG stage formed by M_2. The short-circuit transconductance of the overall cascode circuit is approximately equal to g_{m1}, since M_2 merely acts as a unity gain current buffer.

The net effect is that the product $G_m R_{out}$ of the compound device formed by M_1 and M_2 is on the order of $(g_m r_o)^2$, the intrinsic voltage gain of a single MOSFET squared. Thus, an important application of the cascode stage turns out to be in operational amplifiers that implement very large voltage gains. An additional application that takes advantage of the large R_{out} alone is a precision current mirror, discussed in Chapter 5.

Figure 4-26: The cascode configuration and its low-frequency equivalent circuit.

Interestingly, the cascode configuration offers high-frequency benefits as well. To see this, consider the two-port model of the cascode stage in Figure 4-27. The capacitance looking into the stage is C_{gs1} plus the Miller amplified C_{gd1}. However, with M_2 present, the low-frequency voltage gain from the input to the drain node (v_x) of M_1 (i.e., the Miller gain A_{vM}) is limited to approximately the ratio of the two device transconductances (provided that $R_D << r_{o2}$). This means that only a relatively insignificant amount of Miller multiplication occurs; the voltage swing at the circuit output is effectively isolated from the drain node of M_1.

A more subtle but welcome high-frequency benefit comes from reduced feedthrough via C_{gd1} at very high frequencies. From Eq. (3.38), we know that a conventional CS stage comes with a high-frequency zero, defined by the ratio g_m/C_{gd}. Detailed analysis shows that the impact of this zero on the overall transfer function is largely mitigated with the placement of M_2 between the input and the overall circuit output.

Figure 4-27: Mitigation of the Miller effect in the cascode configuration.

4-5-2 CS-CD Cascade

Another popular two-transistor configuration is the CS-CD cascade shown in Figure 4-28. The main idea here is to employ the CD stage as a voltage buffer to decouple the output resistance of the overall circuit from its voltage gain. Specifically, a relatively large value for R_D can be used to maximize the voltage gain of the CS stage. Yet, the circuit can drive relatively small load resistors R_L, since the resistance looking into M_2 is small ($R_{out} \sim 1/g'_{m2}$). If R_L was connected directly to the output of CS stage (v_x), the voltage gain would reduce significantly (depending on the specific component values).

A disadvantage of using a CD voltage buffer is that the available swing is reduced considerably; this is because the CD device requires a voltage drop of $V_{Tn} + V_{OV}$ from the gate to the source. This voltage drop reduces the voltage range between V_{DD} and ground that can be used for the signal.

Figure 4-28: Common-source common-drain cascade.

4-5-3 CG Stage as a Load Device

An interesting combination of a CS and CG stage is shown in Figure 4-29. Here, the CG stage essentially emulates the drain resistor R_D that is shown for example in the stage of Figure 4-28 and used in the CS circuits throughout Chapters 2 and 3. The small-signal resistance looking into the source of M_2 is approximately $1/g'_{m2}$ and therefore the low-frequency voltage gain of the full circuit is approximately equal to $-g_{m1}/g'_{m2}$. This means that the gain is typically not very large, since there are limits on how small g'_{m2} can be made while maintaining a reasonable device size.

Figure 4-29: Common-source stage with common-gate load device.

An advantage of using a MOSFET over a resistor is that the voltage gain is defined by two components of the same type.

Circuit designers refer to this concept as **ratiometric design**. Even if the transconductance of each MOSFET varies substantially (e.g., due to a large temperature change), the voltage gain will remain relatively constant. This concept will be better appreciated once we have discussed sources of parameter variations in Chapter 5.

Similar to the CD buffer discussed in the previous sub-section, a disadvantage of this circuit topology is that it does not support a very large output swing. In addition to the voltage drop from the gate to the source of M_2, note that the output can swing from the quiescent point in the positive direction by at most V_{OV2} before M_2 enters cutoff.

Figure 4-30: Summary of low-frequency two-port model structures for the CS, CG and CD configurations.

Summary

In this chapter, we discussed the common-gate and common-drain stage configurations. As we have seen, the core sections of these circuits operate as a current buffer and voltage buffer, respectively. Figure 4-30 summarizes the corresponding low-frequency model structure in comparison with the common-source stage. Focusing only on the core circuits and neglecting any auxiliary elements such as drain resistance R_D, Table 4-2 lists the main model parameters. As exemplified in Section 4-5, combining these three configurations with their different characteristics gives the designer the freedom to construct a wide range of application-optimized amplifiers. This idea is studied further in Chapter 6.

As far as the frequency response of the CG and CD stages is concerned, we have seen that several approximations are necessary to capture their essential behavior in simplified, low-complexity expressions. In performing these simplifications and intuitively understanding results that are in their raw form algebraically complex, it pays to master the basic tool set covered in this module so far. Important techniques include the concept of working with unilateral two-port approximations, the dominant pole approximation, the method of open-circuit time constants, and a general feel for the relative magnitudes of MOSFET capacitances, transconductance and output resistance.

Table 4-2: Comparison of low-frequency stage characteristics.

Property	**CS**	**CG**	**CD**
Two-port model	Transconductance or Voltage Amplifier	Current Amplifier	Voltage Amplifier
Gain	g_m or $A_v = g_m r_o$	$A_i \cong 1$	$A_v \cong 0.7...1$
R_{in}	∞	$\cong 1/g'_m$	∞
R_{out}	r_o	$r_o(1 + g_m r_S)$	$\cong 1/g'_m$

The main finding from the detailed frequency response analysis of the CG and CD stages is that these circuits can operate (under certain conditions) up to very high frequencies, nearing the cutoff frequency (ω_T) of the constituent MOSFET.

An important detail in the design of CG and CD stages is the connection of the MOSFET's bulk. The circuit designer must understand the various options afforded by the process technology and then decide about a suitable configuration that helps maximize performance. As we have shown, it is typically advantageous to connect the bulk to the supply voltage (GND for an n-channel, and V_{DD} for a p-channel) in a CG stage. In a CD stage, it is sometimes advantageous to opt for a source-bulk tie, especially when a voltage gain as close as possible to unity is required.

References

1. S. Bhave, R. T. Howe and C. G. Sodini, *TBD*, 2011.
2. P. R. Gray, P. J. Hurst, S. H. Lewis, and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 5th Edition, Wiley, 2008.

Problems

Unless otherwise stated, use the standard model parameters specified in Table 4-1 for the problems given below. Consider only first-order MOSFET behavior and include channel-length modulation (as well as any other second-order effects) only where explicitly stated.

P4.1 In the circuit of Figure P4-1, the MOSFET has a channel length of 2 μm and the width was chosen such that $V_{OUT} = 1.5$ V with $I_B = 200$ μA. Neglect channel length modulation.

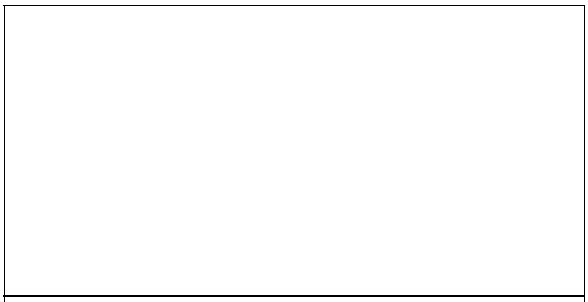


Figure P4-1

(a) Draw the complete small-signal model of the circuit and eliminate capacitances that will not affect the circuit operation. Be sure to include the well-to-substrate capacitance.

(b) Calculate the MOSFET's gate-source (C_{gs}) and the well-to-substrate (C_{bsub}) capacitance. Compute the ratio C_{bsub}/C_{gs}.

P4.2 Repeat Example 4-3 assuming that the bulk is tied to the source of the MOSFET, i.e., $V_{SB} = 0$. Re-compute V_{OUT}, V_S, as well as V_{DS} and $V_{OV} = V_{GS} - V_{Tn}$ of the transistor.

P4.3 The p-channel common-gate amplifier shown in Figure P4-3 has a power dissipation of 1 mW. The MOSFET has a channel length of 2 µm.

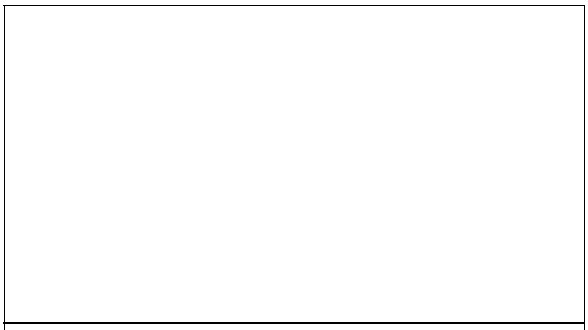


Figure P4-3

(a) What is the W/L required for $i_{out}/i_s = 0.8$?

(b) If the transistor is placed in an n-well so that $V_{SB} = 0$ V, What is the new overall current gain i_{out}/i_s?

P4.4 Repeat Example 4-4 using a channel length of 3 µm and all other parameters left unchanged. Comment on any changes in the dominant time constants. Note that you will need to re-compute the quiescent point voltages of the circuit.

P4.5 The circuit shown Figure P4-5 is used as a transresistance amplifier for small-signal currents coming from a photo diode, modeled using the equivalent circuit as shown. For simplicity, ignore the backgate effect in all calculations. Parameters: R_{diode} = 50 kΩ, C_{diode} = 70 fF, C_L = 150 fF, R_L = 1 kΩ, and L = 1 µm for the MOSFET.

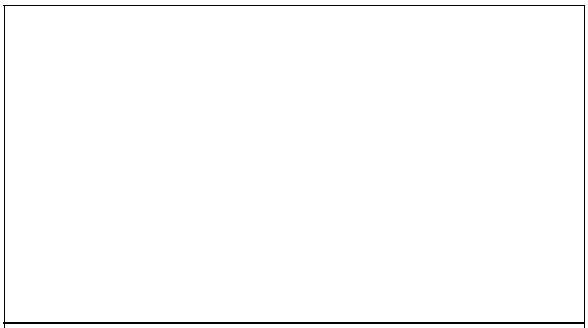


Figure P4-5

(a) Assuming I_D = 150 µA, calculate the required gate overdrive voltage (V_{OV}) to achieve a trans-conductance of 0.75 mS and size the width of the transistor accordingly. Determine the proper value needed for I_B taking R_{diode} into account.

(b) Considering only the intrinsic capacitance and ignoring r_o, construct the small-signal model of the circuit.

(c) Calculate the low-frequency transresistance gain v_{out}/i_{light}.

(d) Estimate the circuit's bandwidth using the method of open-circuit time constants assuming $R_G = 0$.

(e) Repeat part (d) assuming R_G = 500 Ω.

P4.6 Derive the approximate result for the output impedance of a CG stage as given in Eq. (4.27). Using the circuit of Figure 4-13, apply a test source at the output and solve for $Z_{out} = v_t/i_t$. Simplify the resulting expressions assuming $\omega << \omega_T$ and $g'_m r_s >> 1$.

P4.7 Consider the CD circuits shown in Figure P4-7(a) and (b). Ignore channel length modulation.

(a) Suppose that the output voltage swings from 2 V to 4.5 V in both circuits. Calculate the corresponding voltage swings at the gate nodes for R_L = 1kΩ and R_L = 1 MΩ. Determine these values using the transistor's large signal model. Summarize your results in a table and explain how different values for R_L and the two bulk connection schemes affect the required input voltages.

(b) Draw a low-frequency small-signal model for the circuit of Figure P4-7(a). Find an analytical expression for the low-frequency small-signal voltage gain v_{out}/v_{in} as a func-

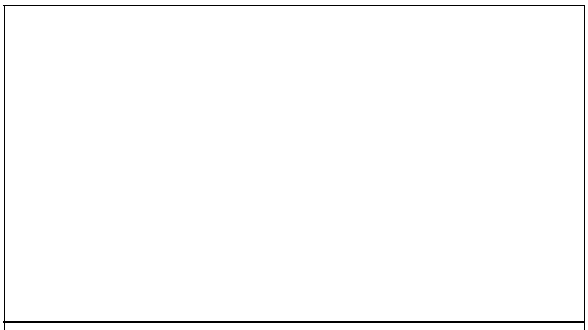


Figure P4-7

tion of R_L and the quiescent point voltages V_{IN} and V_{OUT}. Calculate the small signal gain at V_{OUT} = 2 V and V_{OUT} = 4.5 V using R_L = 1 kΩ.

P4.8 In the circuit of Figure P4-8, ignore backgate effect and channel length modulation. Parameters: R = 1 kΩ, W/L = 40.

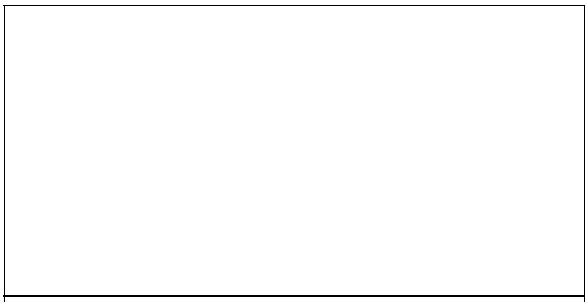


Figure P4-8

(a) Calculate the input voltage and corresponding output voltage at which the device enters the triode region.

(b) Sketch V_{OUT} versus V_{IN} (0…5 V). Calculate and annotate pertinent asymptotes and breakpoints, including the voltages calculated in part (a).

P4.9 Derive Eq. (4.36) through Eq. (4.39) using a KCL-based analysis of the circuit in Figure 4-20.

P4.10 Derive Eq. (4.37) by performing an open-circuit time constant analysis for the circuit of Figure 4-20.

P4.11 Consider the p-channel CD circuit of Figure 4-22. Assuming C_{gs} = 200 fF and $g_m r_o$ = 50, what is the contribution of C_{gs} to the overall input capacitance of the circuit (in fF)?

P4.12 In the circuit of Figure P4-12, the input quiescent point operating point voltage is adjusted such that V_{OUT} = 1 V. The MOSFET is sized such that V_{OV} = 400 mV and f_T = 1 GHz. In your analysis, neglect backgate effect, finite output resistance, and all extrinsic device capacitances.

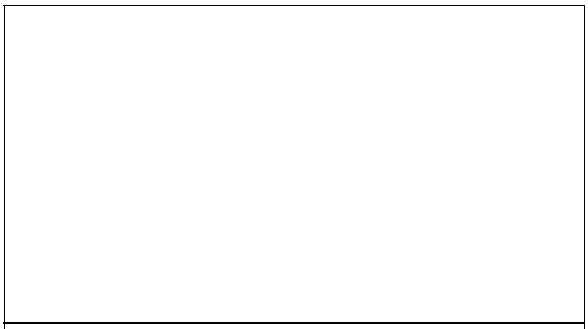


Figure P4-12

(a) Calculate the circuit's low-frequency small-signal voltage gain v_{out}/v_{in}.

(b) After playing in the lab for many hours, your friend found that for a certain value of C_L, the circuit achieves a perfectly "flat" frequency response and essentially "infinite" small-signal bandwidth. Calculate the value of C_L that causes this behavior in $v_{out}(s)/v_{in}(s)$. Assume that the input of the circuit is driven by an ideal voltage source (R_S = 0).

(c) Suppose we let C_L = 0. Under this condition, what is the small-signal voltage gain when the input frequency approaches "infinity"? Sketch a Bode plot of the circuit transfer function to justify your answer.

P4.13 In Example 4-6, we estimated the bandwidth of a CD circuit using the term b_1 in the circuit's transfer function. Given the parameter values used in Example 4-6, plot the magnitude of the exact transfer function given by Eq. (4.36) (using any computer program that is available to you) and find the 3dB bandwidth from this plot numerically. Compare the obtained value with the numbers found in Example 4-6.

P4.14 In Figure 4-21, we used a qualitative reasoning to explain the Miller multiplication term of C_{gs} in a CD stage. Specifically, we argued that the Miller gain for this capacitance will be constant up to very high frequencies. Following the same approach that was taken to derive Eq. (3.54) for the CS stage, derive an exact expression for the Miller gain across C_{gs} with relevant poles and zeros included. Show that these poles and zeros can be disregarded as long as C_L is small.

P4.15 Derive Eq. (4.45) using a KCL-based analysis. As shown in the circuit of Figure 4-23, apply a test source at the output and solve for $Z_{out} = v_t/i_t$. Approximate $g_m \cong g'_m$ and neglect C_{sb}.

P4.16 Consider the CD circuit shown in Figure P4-16(a). Neglect finite output resistance and all extrinsic device capacitances, as well as the bulk-to-substrate capacitance.

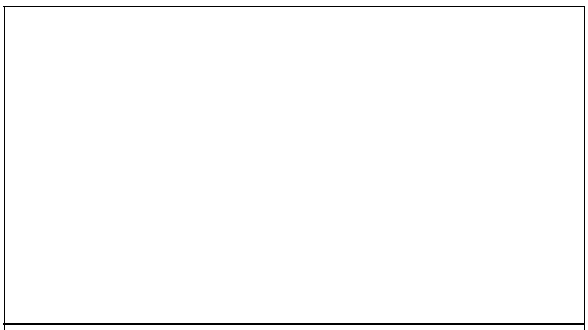


Figure P4-16

(a) Draw the small-signal model of the circuit and find a symbolic expression for the output impedance $Z_{out}(s)$.

(b) Show that for $R_S > 1/g_m$, the output impedance can be modeled as an RL circuit as shown in Figure P4-16(b). Express L, R_1 and R_2 in terms of R_S, g_m and C_{gs}.

(c) Calculate numerical values for L, R_1 and R_2 assuming the following parameters: $I_B = 1$ mA, $V_S = 1$ V, $R_S = 1$ MΩ, $W = 20$ μm, and $L = 1$ μm. Verify that $R_S > 1/g_m$.

5

CHAPTER

Biasing Circuits

The elementary transistor stages analyzed in the previous chapters rely on proper voltage and current biasing to function. So far, we have emulated these bias generators using ideal voltage and current sources. In this chapter, we will look at practical realizations of these elements using MOSFETs and passive components available within an integrated circuit.

While there exist numerous possibilities for setting up bias voltages and currents, we consider here a subset of options that have proven to be robust in practical circuits produced in volume. Many of the ideas and considerations that go into the design of bias circuitry are intimately related to the parameter variations seen in an integrated circuit process technology. For instance, threshold voltages cannot be accurately reproduced from fabrication run to fabrication run, and this mandates certain measures for desensitization to this parameter. In order to understand the rationale behind the proposed biasing circuits, we therefore include an overview of the basic variability issues that analog CMOS circuit designers must be aware of.

Chapter Objectives

- Review basic variability issues relevant for analog integrated circuits in CMOS technology.
- Discuss and analyze practical circuits that can establish the bias voltages and currents required to operate the elementary common-source, common-gate and common-drain stages.

5-1 Overview

Figure 5-1 provides and overview of the circuitry and topics that will be discussed in this chapter. The function of the shown circuits will be explained as we progress through this chapter. Following this introduction, we will investigate basic issues of process variation and device mismatch seen in a typical CMOS fabrication process. This review will help motivate some of the design choices made in later sections. Next, in Section 5-3, we investigate current mirror circuits, which are essential to distributing and generating bias currents in an integrated circuit. The current that flows into a current mirror circuit is defined by an absolute current reference, for which there exist many different realizations. In Section 5-4, we will study one relevant example of a suitable circuit. In Section 5-5, we will then shift to the problem of bias voltage generation, as relevant for example in setting up the proper gate bias of a common-source or common-gate stage.

5-2 Introduction to Process Variation and Device Mismatch

5-2-1 Process and Temperature Variations

In our analysis of elementary circuit configurations, we have so far implicitly assumed that the underlying component parameters (e.g., the threshold voltage of a MOSFET) are constant and

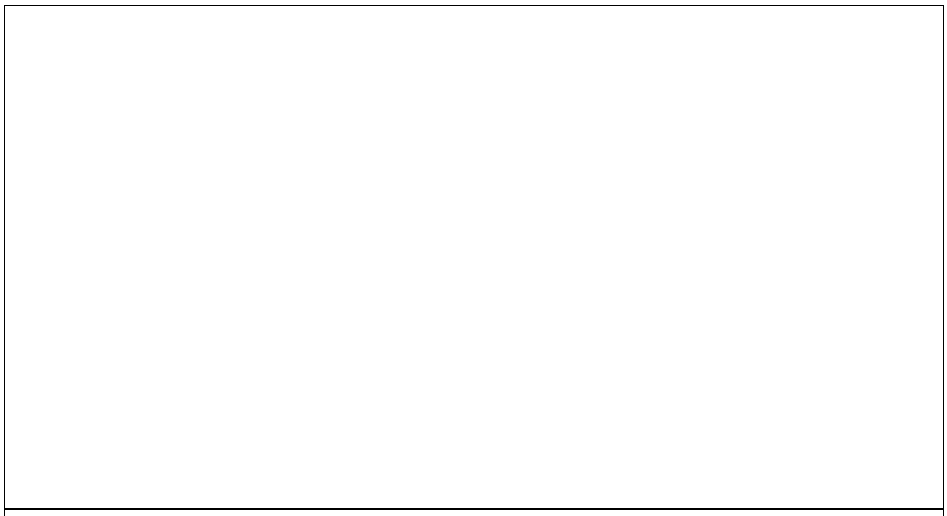


Figure 5-1: Overview of biasing circuits and topics covered in this chapter.

accurately known. Unfortunately, this is not the case in reality. Especially in mass-produced integrated circuits, there are various forms of variability that result in parameter uncertainty due to imperfect fabrication, lifetime drift and influence of environmental conditions such as temperature and humidity. In a typical large semiconductor company, entire departments tend to focus on this issue and there exists a wealth of related information that could easily fill multiple textbooks. As a result, the focus in this introductory module is to take a cursory look at only the basic issues, to the extent that this can help shape our thinking on how to arrive at practical and relatively insensitive circuit realizations.

The first issue that we will review in this section is related to variations arising from imperfect fabrication and temperature changes. In the context of fabrication imperfections, we will clearly distinguish between **global process variations** and **device mismatch**. The former term relates to variations that affect all devices on a chip uniformly, while the latter term refers to differences between nominally identical devices that are fabricated on the same chip (see Section 5-2-2).

Analog circuit designers often use the term **PVT variations** to refer to global variations in process, supply voltage and temperature (see Table 5-1). The most basic way to capture global fabrication process variations is to define parameter sets that group the worst case outcomes as “slow”, “nominal”, and “fast” conditions. This nomenclature was adopted in the context of digital circuits (relating to the speed of a logic gate), but is also used among analog designers. The various parameter sets are often called **process corners**.

Table 5-2 shows how some important integrated circuit parameters may vary across the three process parameter sets. Here, the nominal column contains the MOSFET parameters that we have assumed so far in this module (see Table 4-1). In the slow parameter set, the threshold voltage is increased and the transconductance parameters (μC_{ox}) are reduced; this is the parameter combination that yields the slowest speed in a logic gate. The opposite is true for the fast parameter set.

Table 5-1: Examples of typical process, voltage and temperature (PVT) variations.

Process	The chip foundry defines three parameter sets for “slow”, “nominal” and “fast” conditions.
Voltage	The chip’s supply voltage is expected to vary by ±10%. For a nominal supply of 5 V, this means that all circuits must work for V_{DD} = 4.5...5.5 V
Temperature	Consumer products are typically expect to work in ambient temperatures ranging from 0...70°C. Circuits used in automotive applications must work reliably from –40...125°C.

Table 5-2 also contains examples for parameter variations in passive IC components. R_{poly} and R_{well} are the sheet resistances of a resistor formed by a layer of polysilicon or n-well, respectively. C_{poly} is the capacitance parameter of a parallel plate capacitor formed by two layers of polysilicon. Advanced texts on integrated circuit design (such as Reference 2) provide further information about the make-up of these and similar components.

The above-tabulated parameter variations do not take temperature variations into account; these must be added on top of the spread from fabrication. Table 5-3 lists a few typical temperature coefficients for each parameter. For example, if the operating temperature of a chip changes from 0°C to 70°C, the threshold voltage a MOSFET will shift by an additional –84 mV.

Table 5-2: Example of a slow, nominal and fast parameter set in a CMOS fabrication process. These parameters assume that the temperature is held constant at 25 °C (room temperature).

Parameter	Slow	Nominal	Fast
V_{T0n}, $\lvert V_{T0p}\rvert$	0.7 V	0.5 V	0.3 V
$\mu_n C_{ox}$	40 µA/V^2	50 µA/V^2	60 µA/V^2
$\mu_p C_{ox}$	20 µA/V^2	25 µA/V^2	30 µA/V^2
R_{poly}	40 Ω/square	50 Ω/square	60 Ω/square
R_{well}	1.4 kΩ/square	1 kΩ/square	0.6 kΩ/square
C_{poly}	1.15 fF/µm^2	1 fF/µm^2	0.85 fF/µm^2

Table 5-3: Typical temperature coefficients for integrated circuit device parameters.

Parameter	Temperature Coefficient
V_{T0n}, $\lvert V_{T0p}\rvert$	–1.2 mV/°C
$\mu_n C_{ox}$	–0.33%/°C
$\mu_p C_{ox}$	–0.33%/°C
R_{poly}	+0.2%/°C
R_{well}	+1%/°C
C_{poly}	–30 ppm/°C

The main take-home from the above data is that in practice, the analog IC designer cannot view component parameters as constant numbers. His or her circuit must be immune to the level of variability described above and function reliably across a large array of outcomes in process, voltage and temperature. To show how significant these effects can be when neglected, the following example considers the impact of process variations on the bias point of a common-source amplifier.

Example 5-1: Impact of Process Variations in a Common-Source Amplifier

The circuit in Figure Ex5-1 was previously analyzed in Example 2-2(b) using nominal parameters (at room temperature). Given V_{DD} = 5 V, R_D = 10 kΩ, W/L = 10 and a desired output bias point of V_{OUT} = 2.5 V, we found that the input bias voltage should be set to V_{IN} = 1.5 V. Assuming V_{IN} = 1.5 V, re-compute the circuit's operating assuming that the MOSFET parameters have shifted to the fast corner case given in Table 5-2. For simplicity, ignore variations in R_D and operating temperature.

SOLUTION

Using the given parameters, we can directly compute

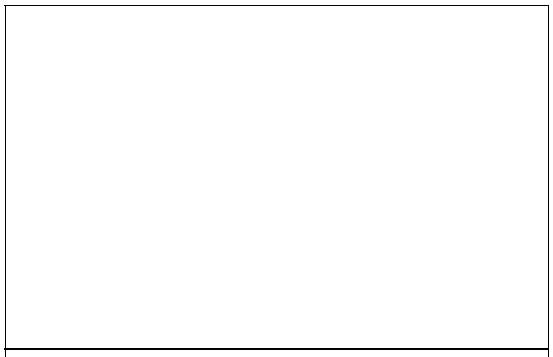


Figure Ex5-1

$$V_{OUT} = V_{DD} - \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{IN} - V_{Tn})^2 R_D$$

$$= 5\text{V} - \frac{1}{2}60\frac{\mu A}{V^2}\cdot 10\cdot(1.5V - 0.3V)^2\cdot 10k\Omega = 0.68V$$

From this result, we see that the MOSFET no longer operates in saturation (since V_{DS} = 0.68 V < $V_{GS} - V_{Tn}$ = 1.2 V). Using the MOSFET's equation for the triode region, we can compute I_D = 408 µA and V_{OUT} = 918 mV. This outcome differs substantially from the nominal operating point, and the circuit will essentially not function as intended for the fast corner conditions.

The main finding from the above example is that it will usually be impractical to bias the input of a common-source stage using a fixed bias voltage source. In practice, the integrated circuit designer generates bias voltages using circuits that will automatically adjust to corner-induced parameter spread and thereby make the circuit immune to process variations (see Section 5-5).

Generally speaking, a substantial amount of design time is usually spent on identifying biasing approaches that ensure a circuit's proper bias point across all possible operating conditions. In addition, once, the circuit is properly biased, the designer must verify that it maintains its key specifications across corners. A typical scenario is to guarantee a certain worst case gain or bandwidth across all PVT scenarios.

As already mentioned, it is impossible to cover all aspects of robust design across PVT variations at the introductory level of this module. Nonetheless, having some of the basic knowledge established above will help us argue qualitatively about the practicality of the circuits discussed in this chapter, and ensure that they will at least have a chance to work in practice.

5-2-2 Mismatch

The process variations discussed in the previous section account for variability that affects all devices on a given chip equally. For example, all n-channel MOSFETs on a given chip may have slow parameters. Different from process variation, we use the term mismatch to capture variations between nominally identical devices, e.g., two MOSFETs of identical size on the same chip. Such variations are typically caused by line edge roughness, random doping fluctuations and similar effects.

Device mismatch typically follows Gaussian distributions and depends on device size and spacing (see Reference 1). For our purpose in this module, we will not expand upon the detailed theory behind this and instead consider only approximate numerical ranges that are typical for a technology as the one assumed in this module (see Table 5-4).

Table 5-4: Typical ranges of parameter mismatch for nominally identical, closely spaced components.

Parameter	Mismatch
V_{T0n}, V_{T0p}	5...30 mV
$\mu_n C_{ox}$, $\mu_p C_{ox}$	0.5...2%
R_{poly}	0.3...2%
C_{poly}	0.1...1%

At first glance, we see from Table 5-4 that device mismatches are typically much smaller than global process variations.[1] For instance, the nominal threshold voltage for n-channel transistors can vary by ±200V from fabrication run to fabrication run. However, within a specific fabrication outcome, the random threshold mismatch between two n-channels on the same chip is on the order of 10 mV.

This observation has a profound impact on the way integrated circuits are architected. That is, designers will usually try to exploit the fact that the components on the same chip show good matching. This contrasts with printed circuit board (PCB) design, where the designer often cannot rely on good matching between the available discrete components. Instead, PCB design can offer certain components with very high *absolute* accuracy across fabrication lots, such as 1%-precise resistors. As we know from Table 5-2, such levels of absolute accuracy are usually not available in an integrated circuit.

A classical example that exploits transistor matching is the so-called **current mirror**. This circuit is ubiquitous in integrated circuits, but infrequently used in PCB circuits. We will now analyze the current mirror as a first example of a biasing circuit that is insensitive to process variations

1. This tends to hold true for technologies with feature sizes above 100 nm. For nano-scale devices, device mismatch can be comparable to process spread.

5-3 Current Mirrors

As we have seen in previous chapters, we would like to use current sources to setup the bias points for CS, CG and CD stages. While we could in principle design individual, stand-alone current-source circuits each time we need a bias current, it is instead customary to work with only one (or a few) reference current generators on a given chip and "mirror" its current to the various locations where a bias current is needed. This is sketched out in Figure 5-1: a single reference current generator is used to feed a distribution network of current mirrors (to be discussed in this section), which then supplies bias currents to various circuit stages in a given chip or large sub-block.

In this section, we will discuss and analyze current mirror circuits at various levels of detail. We will begin by considering the most basic structure and perform a first-order analysis for this circuit. Next, we consider second-order error sources and look at an improved realization that invokes the cascode structure introduced in Section 4-5-1.

5-3-1 First Pass Analysis of the Basic Current Mirror

Figure 5-2 shows the most basic realization of a current mirror using two identically sized n-channel MOSFETs. The circuit takes an input current I_{IN} and produces an output current I_{OUT}. Neglecting channel-length modulation for the time being, we can compute the gate-source voltage of M_1 using

$$V_{GS1} = V_{IN} = V_{Tn} + \sqrt{\frac{2I_{IN}}{\mu_n C_{ox}\frac{W}{L}}} \tag{5.1}$$

Since the gates and sources of the two MOSFETs are connected, we see that $V_{GS2} = V_{GS1}$, and therefore

$$\frac{I_{OUT}}{I_{IN}} = \frac{I_{D2}}{I_{D1}} = \frac{\frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{GS2}-V_{Tn})^2}{\frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{GS1}-V_{Tn})^2} = 1 \tag{5.2}$$

Thus, the output current equals the input current (to first order). In essence, the function of M_1 is to "compute" the gate-source voltage required for M_2 to supply the same current that is injected into M_1.

One important feature of this circuit is that it is immune to global process variations. From Eq. (5.2), we see that absolute

changes in V_{Tn} and $\mu_n C_{ox}$ that are common to M_1 and M_2 do not affect the current ratio. The circuit is affected only by mismatches in these parameters. However, as we have seen in Section 5-2-2, parameter mismatch tends to be small in integrated circuits.

Figure 5-2: Basic current mirror.

Figure 5-3 shows an application example of the basic current mirror in a p-channel common-source amplifier. This example is useful for identifying some general design objectives:

- We want to minimize the error in I_{OUT} so that the bias current of the common-source device is accurately set (see also Section 2-2-8).
- We want to minimize the voltage that is needed to keep M_2 in saturation (to allow for a large signal swing). We call this minimum voltage level the **compliance voltage**, V_{OUTmin}.
- We want to minimize the capacitance C_{out} that the current mirror contributes to the output node of the amplifier. This will help maximize the circuit's bandwidth.
- We want to maximize R_{out}, the resistance looking into M_2. A small R_{out} can substantially reduce the voltage gain of the circuit in some use cases.
- Lastly, it is desirable to scale the mirror's branch currents, that is, we want $I_{OUT} = K \cdot I_{IN}$, where typically $K > 1$. This helps reduce the overall current consumption of the circuit and provides flexibility in adjusting the current values within a larger distribution network.

These objectives tend to hold in general for all types of current mirror implementations and we will keep them in mind as we progress through the remaining sub-sections. For the time being, let us look into the scaling of branch currents. Essentially, we would like to accomplish

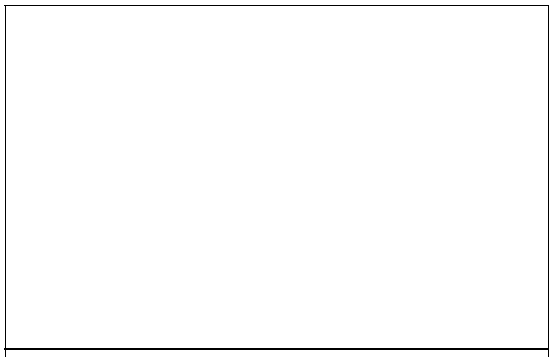


Figure 5-3: Application of the basic current mirror in a common-source amplifier.

$$\frac{I_{OUT}}{I_{IN}} = \frac{I_{D2}}{I_{D1}} = \frac{\frac{1}{2}\mu_n C_{ox}\frac{W_2}{L_2}(V_{GS2} - V_{Tn})^2}{\frac{1}{2}\mu_n C_{ox}\frac{W_1}{L_1}(V_{GS1} - V_{Tn})^2} = K \qquad (5.3)$$

where K is the current scaling factor. From the above expression, assuming that V_{Tn} and $\mu_n C_{ox}$ are exactly equal for both transistors, it follows that $K = (W_2/L_2)/(W_1/L_1)$. Thus, current scaling can be conveniently realized by scaling the MOSFETs' aspect ratios.

In practice, several guidelines exist on how exactly this scaling should be implemented. The first and most important guideline is that we should always maintain $L_1 = L_2$; the current scaling should be realized by scaling the widths rather than the lengths of the channels. This is preferable since the current in a modern MOSFET does not accurately scale with $1/L$. As already mentioned in Chapter 2, the $1/L$ proportionality in our equations is essentially due to the simplified physical model that we used in the derivation of the square-law expressions. For the 1-µm technology assumed in this module, the deviation from the square law model is not as severe as for today's sub-100-nm transistors but still significant enough to avoid length scaling in current mirrors.

Now, with $L_1 = L_2$, the current scaling factor is simply $K = W_2/W_1$, to first order. In the next section, we will look at various second-order effects that cause K to deviate from the width ratio of the MOSFETs.

5-3-2 Second Pass Analysis of the Basic Current Mirror

There exist several error sources in a current mirror that will affect its scaling factor. In general, we classify these error sources into two categories: systematic and random errors. Examples of systematic errors are:

- Errors in transistor width ratios, for example due to mask misalignment or systematic etching imperfections.
- Differences in the drain-source voltages between M_1 and M_2, leading to current deviations caused by channel length modulation.
- Differences in the source potentials of M_1 and M_2 due to finite resistance in the interconnect (so-called "IR drop").

Examples of random errors are:

- Random mismatches in device geometries, for example due to line edge roughness.
- Random mismatch in the transistors' threshold voltage or transconductance parameter.

In order to attain the best possible accuracy in a current mirror, the IC designer will typically try to minimize the impact of all of these errors. We will therefore analyze some of the most important effects and countermeasures in the following paragraphs. For simplicity, our analysis will consider each effect separately. Ultimately, however, the sum of all errors must be considered in practice.

To analyze the impact of systematic masking or etching errors, consider the specific example of a current mirror with a desired current ratio of two and a layout as shown in Figure 5-4. Here M_2 is drawn twice as wide as M_1. In an ideal situation, this would yield $K = 2$ based on the first order result of the previous subsection. In a typical IC process, however, masking or etching errors can lead to a systematic error in the width of a MOSFET, indicated as ΔW in the shown layout. With this error, and neglecting any other imperfections for simplicity, we have

Figure 5-4: Layout of a current mirror with a desired current ratio of two.

$$K = \frac{I_{OUT}}{I_{IN}} = \frac{2W_1 + \Delta W}{W_1 + \Delta W} = \frac{2 + \frac{\Delta W}{W_1}}{1 + \frac{\Delta W}{W_1}} \cong 2 - \frac{\Delta W}{W_1} \qquad (5.4)$$

where the final approximation follows from a first-order Taylor expansion and holds for $\Delta W/W_1 << 1$.

Especially for small transistors, the error term in Eq. (5.4) can be significant. Therefore, it has become customary to adopt layout styles that eliminate issues due to ΔW altogether. In the improved layout of Figure 5-5, M_2 is formed using two unit devices whose layout is identical to that of M_1. In this case,

Figure 5-5: Improved layout of a current mirror with a desired current ratio of two.

$$K = \frac{I_{OUT}}{I_{IN}} = \frac{2(W_1 + \Delta W)}{W_1 + \Delta W} = 2 \qquad (5.5)$$

and thus the circuit is insensitive to systematic width errors. Note that the idea of working with unit devices can be extended such that P unit devices are used for M_2 and Q unit devices are used to form M_1. This means that the mirror ratio $K = P/Q$ is restricted to rational numbers.

A variant of the improved unit-device layout is shown in Figure 5-6. Here, the two unit transistors share a single drain region at the output node and therefore have a smaller output capacitance (C_{out} in Figure 5-3). This general idea is often applied when small C_{out} is desired in the particular use case of the current mirror. One disadvantage of the layout in Figure 5-6 is that the source/drain orientation of the rightmost channel are flipped. This can lead to residual systematic errors in process technologies that suffer from source/drain asymmetries. However, it can be shown that this error vanishes when an even number of unit devices are used for both M_1 and M_2. These and many other considerations are part of the knowledge base of experienced analog designers. The reader is referred to advanced literature on this topic for further information.

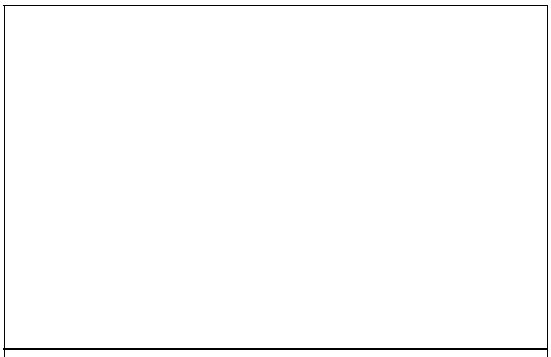


Figure 5-6: Layout of a current mirror with shared drain regions.

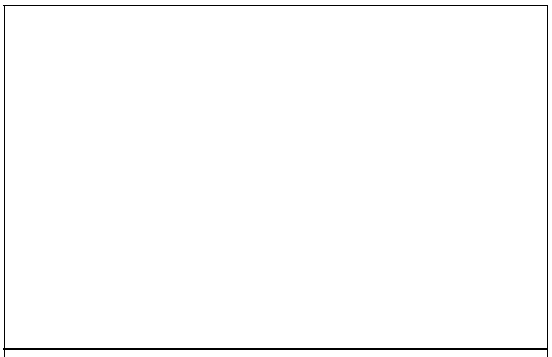


Figure 5-7: Mirror error due to differences in drain-source voltage.

Another significant source of error in the current mirror ratio can result from differences in the voltages at the input and output nodes of the mirror. To see this, consider the current mirror example in Figure 5-7, which is assumed to have perfectly matched transistors of the same size. Even though the two transistors have identical output curves, their drain currents will differ whenever the input and output voltages do not match. Mathematically, we can analyze this effect by including channel length modulation in the analysis. Specifically, since

$$I_{IN} = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{GS1} - V_{Tn})^2(1 + \lambda_n V_{IN}) \tag{5.6}$$

and

$$I_{OUT} = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{GS1} - V_{Tn})^2(1 + \lambda_n V_{OUT}) \tag{5.7}$$

we have

$$\frac{I_{OUT}}{I_{IN}} = \frac{1 + \lambda_n V_{OUT}}{1 + \lambda_n V_{IN}} \tag{5.8}$$

From this result, we see that there are two ways to reduce errors in the current ratio. We can try to minimize the difference between V_{OUT} and V_{IN} as much as possible and/or reduce λ_n by using long-channel MOSFETs. Note also that reducing λ_n is equivalent to reducing the small-signal output conductance g_o, which is simply the slope of the I-V curves in Figure 5-7. The smaller this slope, the smaller the difference between I_{OUT} and I_{IN}.

Example 5-2: Current Mirror Error Due to Drain-Source Voltage Difference

Consider the current mirror in Figure 5-7. Assume V_{OUT} = 2.5 V and that the MOSFET width is chosen such that V_{IN} = 1.5 V. Calculate the percent error in the current ratio for L = 1 µm and L = 3 µm.

SOLUTION

For L = 1 µm, we have $\lambda_n = 0.1\ \text{V}^{-1}$ [see Eq. (2.44)]. Using Eq. (5.8), we find in this case

$$\frac{I_{OUT}}{I_{IN}} = \frac{1 + 0.1\text{V}^{-1} \cdot 2.5\text{V}}{1 + 0.1\text{V}^{-1} \cdot 1.5\text{V}} = 1.087 \tag{5.9}$$

The error in the current ratio is 8.7%. Repeating the above calculation for L = 3µm ($\lambda_n = 0.033\ \text{V}^{-1}$), the error reduces to 3.2%.

Another example of a systematic error source that we will consider here is the voltage drop in the source connection of the mirror devices (see Figure 5-8). In the shown circuits, we assume for simplicity that the two MOSFETs are identical and that $V_{OUT} = V_{IN}$, i.e., there is no error due to V_{DS} differences.

First consider the circuit of Figure 5-8(a), which takes the finite wiring resistance (R_{WIRE}) between the source terminals of M_1 and M_2 into account. The wire will carry some current, which is at the minimum equal to the drain current of M_1 flowing toward the ground node of the circuit. In a poorly constructed layout, the wire may also carry the current from another block (I_X) as shown. The total current in the wire is therefore $I_{WIRE} = I_{IN} + I_X$ and $V_{WIRE} = I_{WIRE} \cdot R_{WIRE}$. By applying KVL in Figure 5-8(a) we see that $V_{GS2} = V_{GS1} + V_{WIRE}$. Therefore, we can use the equivalent model of Figure 5-8(b) for further analysis.

Now, assuming that V_{WIRE} is relatively small, we can think about this voltage as a small-signal perturbation around the operating point of M_2 ($I_{D2} = I_{OUT} = I_{IN}$). Therefore, we can write

$$I_{OUT} = I_{IN} + g_m V_{WIRE} \tag{5.10}$$

and

$$\frac{I_{OUT}}{I_{IN}} = 1 + \frac{g_m}{I_{IN}} V_{WIRE} = 1 + 2\frac{V_{WIRE}}{V_{OV}} \qquad (5.11)$$

where g_m and V_{OV} are the transconductance and quiescent point gate overdrive ($V_{GS} - V_{Tn}$) of the MOSFETs, respectively. To see that this error source can be quite significant, consider the case of V_{OV} = 200 mV and V_{WIRE} = 10 mV. The resulting error in the mirror ratio is 10%. In practice, the designer will mitigate voltage drop issues by: (1) minimizing the distance between M_1 and M_2 (to minimize R_{WIRE}), (2) avoid any excess current (I_X) in the source connection between M_1 and M_2, and (3) work with a reasonably large gate overdrive voltages (V_{OV}) in current mirrors.

Figure 5-8: Mirror error due to voltage drop in the source connection.

As a final step in this sub-section, let us now consider a few examples of random mismatch effects in current mirrors, and specifically mismatch in the transistors' threshold voltages and transconductance parameters. The case of threshold voltage mismatch can be modeled exactly as shown in Figure 5-8(b), but V_{WIRE} is now replaced with ΔV_{Tn}, the threshold voltage mismatch between M_1 and M_2. Therefore, we can write in this case

$$\frac{I_{OUT}}{I_{IN}} = 1 + 2\frac{\Delta V_{Tn}}{V_{OV}} \qquad (5.12)$$

The conclusion from this expression is similar to what we have already stated above. To minimize errors due to threshold voltage mismatch, the designer must work with reasonably large values of V_{OV}. Since $\Delta V_{Tn} \cong 10$ mV is not unusual in CMOS technology (see Section 5-2-2), it follows that it is rather difficult to guarantee highly accurate mirror ratios. Even if we make V_{OV} = 1 V, the corresponding error is still 2%.

For the case of transconductance parameter mismatch, we can write

$$\frac{I_{OUT}}{I_{IN}} = \frac{\frac{1}{2}(\mu_n C_{ox})_1 \frac{W}{L}(V_{GS2} - V_{Tn})^2}{\frac{1}{2}(\mu_n C_{ox})_2 \frac{W}{L}(V_{GS1} - V_{Tn})^2} = \frac{(\mu_n C_{ox})_1}{(\mu_n C_{ox})_2} \qquad (5.13)$$

where $(\mu_n C_{ox})_{1,2}$ are the transconductance parameters of the two MOSFETs, and all other parameters are assumed to be equal. Thus, for typical mismatch in the transconductance parameter on the order of 1% (see Section 5-2-2), it is often the case that this particular error is overshadowed by mismatches in the MOSFETs' threshold voltages.

5-3-3 Multiple Current Sources and Sinks

The basic current mirror concept so far can be utilized to provide multiple current outputs that either source a current from V_{DD} or sink a current into ground. A circuit that uses p-channel devices to create multiple currents sourced from V_{DD} is shown in Figure 5-9.

Figure 5-9: PMOS current mirror with multiple outputs.

If we also require current sinks, the circuit approach shown in Figure 5-10 can be used. Here, the output current from device

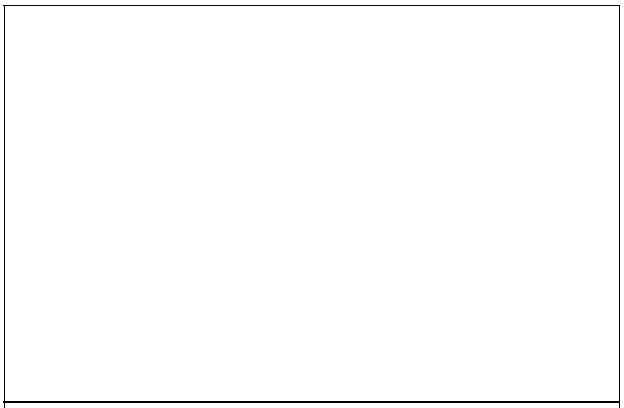


Figure 5-10: Circuit to produce a current source M_2 and current sink M_4.

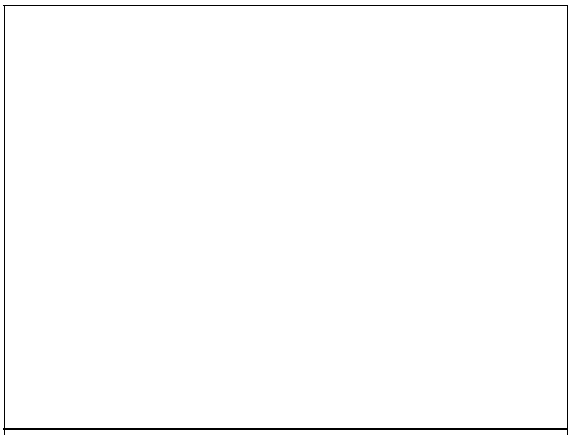


Figure Ex5-3

M_1 is used as a reference current for the n-channel mirror composed of M_1 and M_2. Note that a direct application for this circuit would be the CS-CD amplifier Figure 4-28; it requires one current source from V_{DD} and one current sink.

Neglecting all error terms, and assuming equal channel lengths, the value of the DC current I_{OUT1} is equal to

$$I_{OUT1} = \frac{W_1}{W_R} I_{REF} \tag{5.14}$$

From this current we have derived a current source and current sink with devices M_2 and M_4. Ideally, these currents are

$$I_{OUT2} = \frac{W_2}{W_R} I_{REF} \tag{5.15}$$

$$I_{OUT4} = \frac{W_4}{W_3} I_{OUT1} = \left(\frac{W_4}{W_3} \cdot \frac{W_1}{W_R}\right) I_{REF} \tag{5.16}$$

Example 5-3: Current Sources/Sinks

Design current sources with DC current values of 10 µA and 20 µA and current sinks with DC current values of 10 µA and 40 µA. The small-signal source resistance of all current sources and sinks should be at least 1 MΩ. The compliance voltage of both current sources and sinks must be less than 0.5 V. You are given one reference current source of 10 µA with which you can use to derive the others.

SOLUTION

A suitable topology for this design is shown in Figure 5-3. We begin this design by realizing that in order to meet the compliance voltage requirement, we need $V_{GS} = V_{SG} \leq 1.5$ V. This defines the value of $(W/L)_R$.

$$V_{GS} = V_{Tn} + \sqrt{\frac{I_{REF}}{\frac{1}{2}\mu_n C_{ox}\frac{W}{L}}} \Rightarrow \left(\frac{W}{L}\right)_R = 1.6$$

If we set $(W/L)_1 = (W/L)_2 = 1.6$, then $I_{D1} = I_{D2} = 10$ µA. To make $I_{D3} = 40$ µA, let $(W/L)_3 = 4(W/L)_2 = 6.4$. The p-channel devices are sized the same way.

$$V_{SG} = 1.5\text{V} = -V_{Tp} + \sqrt{\frac{I_{REF}}{\frac{1}{2}\mu_p C_{ox}\frac{W}{L}}} \Rightarrow \left(\frac{W}{L}\right)_4 = 3.2$$

To make $I_{D5} = 10$ µA and $I_{D6} = 20$ µA, we use $(W/L)_5 = 3.2$ and $(W/L)_6 = 6.4$.

Now we can check the small-signal source resistances. For $I_D = 10$ µA and $\lambda_n = \lambda_p = 0.1\ \mu\text{mV}^{-1}/\text{L}$, minimum length $L = 1$ µm will already satisfy the requirement of $r_o = 1$ MΩ. For $I_D = 20$ µA, we need $L = 2$ µm and for $I_D = 40$ µA, we need $L = 4$ µm. Since want all of the n-channels to have the same length, and all of the p-channels to have the same length, we arrive at the following design choice (all values in µm): $(W/L)_R = (W/L)_1 = (W/L)_2 = 6.4/4$ and $(W/L)_3 = (4\text{x}6.4)/4$. For the p-channels: $(W/L)_4 = (W/L)_5 = 6.4/2$, and $(W/L)_6 = (2\text{x}6.4)/2$. As indicated through the multipliers, the layout of M_3 and M_6 should consist of multiple unit devices.

5-3-4 Cascode Current Mirror

As we have seen in the previous sub-section, the accuracy of the current ratio in the basic current mirror is affected by a number of undesired effects. The cascode current mirror discussed in this section improves on a subset of these issues. Specifically, as we shall see, it is less sensitive to differences between V_{IN} and V_{OUT} and correspondingly also provides a

much larger output resistance (R_{out}).

The most basic realization of a cascode current mirror is shown in Figure 5-11(a). The output branch of this circuit stacks two MOSFETs in a cascode configuration (see Section 4-5-1). To compute the output resistance of this circuit, we consider the small-signal model of the circuit in Figure 5-11(b). Note that this circuit resembles a the common-gate model of Figure 4-9, with r_s replaced by r_{o2}, which is the output resistance of the bottom transistor M_2. Consequently, R_{out} is given by Eq. (4.21), which is repeated here with the proper variable substitutions ($r_s \rightarrow r_{o2}$, $r_o \rightarrow r_{o1}$ and $g'_m \rightarrow g'_{m2}$, where $g'_{m2} = g_{m2} + g_{mb2}$)

$$R_{out} \cong r_{o2}[1 + g'_{m2}r_{o1}] \cong r_{o1}g'_{m2}r_{o2} \tag{5.17}$$

Thus, the output resistance of this structure is very large, which implies that any changes in the output voltage will not affect the output current significantly. Mathematically, we can view any disturbance in V_{OUT} as a small signal quantity, v_{out}. The resulting disturbance in the output current, i_{out} is simply v_{out}/R_{out}, which is small for large R_{out}.

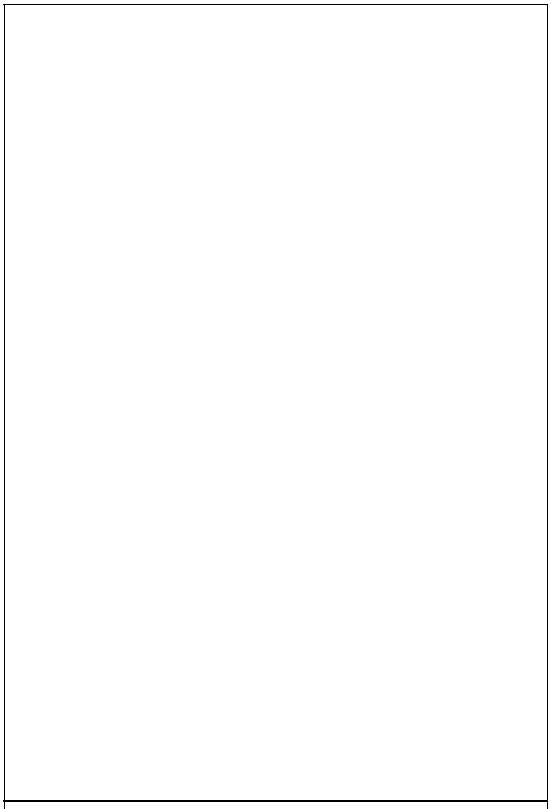


Figure 5-11: Basis cascode current mirror. (a) Complete circuit. (b) Small-signal circuit model for the output branch.

Even though we know that R_{out} is large from the above quantitative result, it is useful to develop a qualitative feel for why this must be the case. To investigate, Figure 5-12 shows the output branch of the cascode current mirror for further inspection. In this drawing, we apply an output perturbation and consider the voltage swing at the drain of M_1. Since the resistance at the drain node of M_1 is low ($\cong 1/g'_{m2}$), the output voltage perturbation appears highly attenuated at this node (the attenuation is approximately given by the ratio of $1/g'_{m2}$ and r_{o2}). Consequently, the drain current of M_1, which is equal to the output current, sees only a very small voltage perturbation. In essence, M_2 shields the current mirror transistor M_1 from the output disturbance; the drain voltage of M_1 is "pinned" by the low resistance node created by M_2.

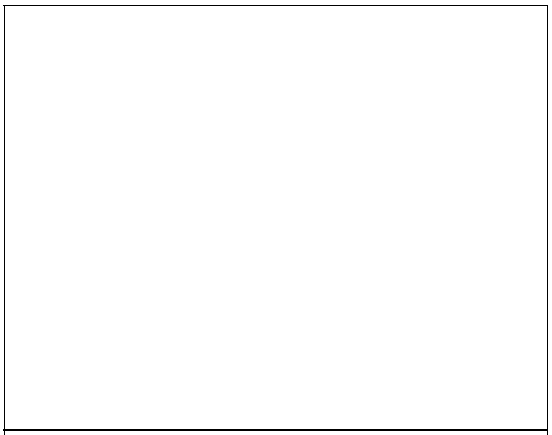


Figure 5-12: Qualitative inspection of the output branch in a cascode current mirror.

While the circuit in Figure 5-11 is insensitive to changes in V_{OUT}, it is important to realize that any difference in the drain voltages of M_1 and M_3 will still lead to a (potentially large) systematic error. Similar to Eq. (5.8), we can write

$$\frac{I_{OUT}}{I_{IN}} = \frac{I_{D1}}{I_{D3}} = \frac{1 + \lambda_n V_{DS1}}{1 + \lambda_n V_{DS3}} \tag{5.18}$$

For this reason, the circuit is purposely constructed such that nominally $V_{DS1} = V_{DS3}$. Assuming that $I_{OUT} = I_{IN}$ and that M_2 and M_4 are identical, applying KVL to the circuit of Figure 5-11 reveals

$$V_{D1} = V_{D3} + V_{GS4} - V_{GS2} = V_{D3} \tag{5.19}$$

Thus, the circuit of Figure 5-11 effectively eliminates this important shortcoming of the basic current mirror.

Unfortunately, the benefits of the cascode current mirror do not come for free. Specifically, notice that the circuit's output

compliance voltage (V_{OUTmin}) is significantly larger than that of a basic current mirror. In a basic current mirror (Figure 5-2), we have $V_{OUTmin} = V_{DSsat2} = V_{OV2}$, which is the gate overdrive voltage of the MOSFET in the output branch. For the cascode current mirror, we can investigate the situation by considering Figure 5-13, which graphically illustrates all voltage levels and voltage drops. Here, we assume for simplicity that all threshold (V_{Tn}) and gate overdrive voltages (V_{OV}) are identical. With this assumption, the voltage at the drain of M_1 is $V_{Tn} + V_{OV}$. This implies that M_1 will always be in saturation, since the drain-source voltage exceeds V_{OV} by some margin (equal to V_{Tn}). In order for M_2 to operate in saturation, we require

$$V_{DS2} = V_{OUT} - (V_{Tn} + V_{OV}) > V_{DSsat2} = V_{OV} \tag{5.20}$$

and thus

$$V_{OUT} > V_{Tn} + 2V_{OV} \tag{5.21}$$

which means $V_{OUTmin} = V_{Tn} + 2V_{OV}$. Note that for typical values of V_{OV} and V_{Tn}, the compliance voltage of the cascode current mirror can become quite large, e.g., 0.5 V + 1 V = 1.5 V, taking away a significant amount of signal swing from the available voltage supply range (consider for example Figure 5-3).

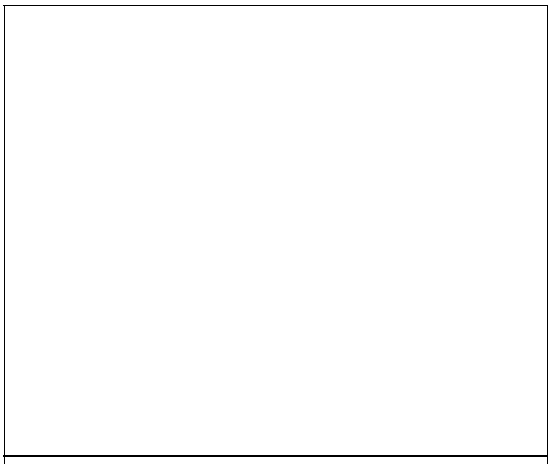


Figure 5-13: Analysis of the compliance voltage in the basic cascode current mirror.

5-3-5 The High-Swing Cascode Current Mirror*

In applications where the large compliance voltage of the circuit in Figure 5-11 is problematic, an alternative scheme, called **high swing cascode current mirror** can be used. We will develop this circuit from the previous solution using a few intermediate steps.

First, consider the output branch of a cascode current mirror as shown in Figure 5-14. In the annotation of this circuit, it is assumed that V_{G2} is set up such that M_1 operates at the edge of saturation, i.e., $V_{DS} = V_{DSsat} = V_{OV}$. In this case, we require

$$V_{DS2} = V_{OUT} - V_{OV} > V_{DSsat2} = V_{OV} \tag{5.22}$$

and thus

$$V_{OUT} > 2V_{OV} \tag{5.23}$$

which means $V_{OUTmin} = 2V_{OV}$, corresponding to a substantial improvement over Eq. (5.21). The question that remains is how exactly V_{G2} should be generated to achieve this improvement. To investigate, we first compute the required value of V_{G2} by applying KVL in Figure 5-14

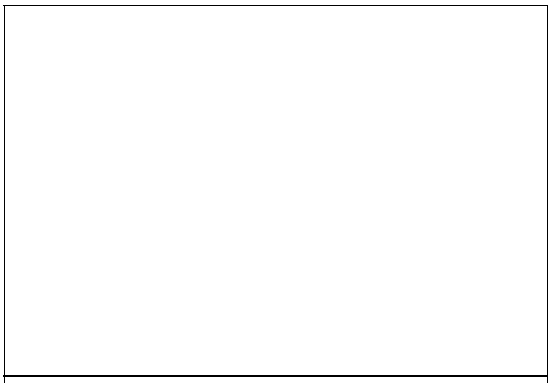


Figure 5-14: Desired output branch biasing to achieve the minimum possible compliance voltage.

$$V_{G2} = V_{DS1} + V_{GS2} = V_{Tn} + 2V_{OV} \tag{5.24}$$

It turns out that many options exist for setting V_{G2} to the above-calculated value. The two most basic options are shown in Figure 5-15. Here, an extra current branch is introduced to bias the added transistor M_6. In a practical implementation, this current typically originates from an extra branch added to a PMOS current mirror in the overall biasing network. The key idea in this setup is that M_4 is sized to 1/4th of the width used for M_2. With this sizing, we have

$$\begin{aligned} V_{G2} = V_{GS6} &= V_{Tn} + \sqrt{\frac{2I_{D6}}{\mu_n C_{ox}\frac{W/4}{L}}} \\ &= V_{Tn} + 2\sqrt{\frac{2I_{OUT}}{\mu_n C_{ox}\frac{W}{L}}} = V_{Tn} + 2V_{OV} \end{aligned} \tag{5.25}$$

which achieves the desired objective.

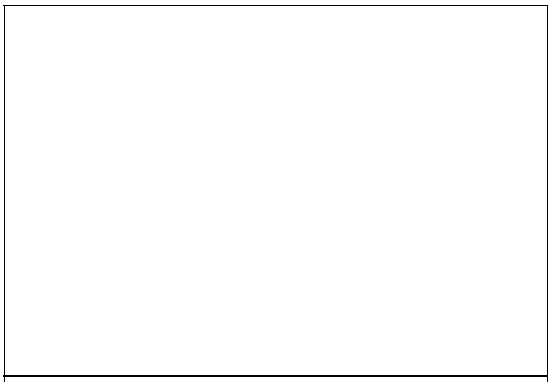


Figure 5-15: Basic circuit for cascode voltage generation.

In practice, the designer will usually not want to bias M_1 exactly at the edge of saturation, but rather leave some margin. This can be achieved by sizing the width ratio smaller than 1/4. Table 5-5 shows the resulting margins for a few integer ratios. Choosing a sizing ratio of 1/6 often yields a reasonable compromise between compliance voltage and circuit robustness in a practical circuit.

Table 5-5: V_{DS1} as a function of the ratio $k = W_2/W_6$ in the circuit of Figure 5-15.

k	V_{DS1}	$V_{DS1} - V_{OV}$ (Margin)
1/4	V_{OV}	0
1/5	1.24 V_{OV}	0.24 V_{OV}
1/6	1.45 V_{OV}	0.45 V_{OV}
1/7	1.64 V_{OV}	0.64 V_{OV}
1/8	1.83 V_{OV}	0.83 V_{OV}
1/9	2 V_{OV}	V_{OV}

In order to complete the high-swing cascode current mirror circuit, we still need to design the circuit's input branch. The most obvious (but non-preferred) solution for the input branch is shown in Figure 5-16(a). This circuit suffers from the problem that $V_{DS1} \neq V_{DS3}$, and therefore a systematic error is introduced in the mirror ratio [see Eq. (5.18)]. An elegant solution to this problem is shown in Figure 5-16(b), where M_4 has been added to replicate the gate-source voltage drop of M_2, such that $V_{DS1} = V_{DS3}$. Just as in the circuit of Figure 5-16(a), the gate voltage of M_3 self-adjusts to the point where M_3 carries the injected current (I_{IN}). M_4 merely acts as a current buffer, pass-

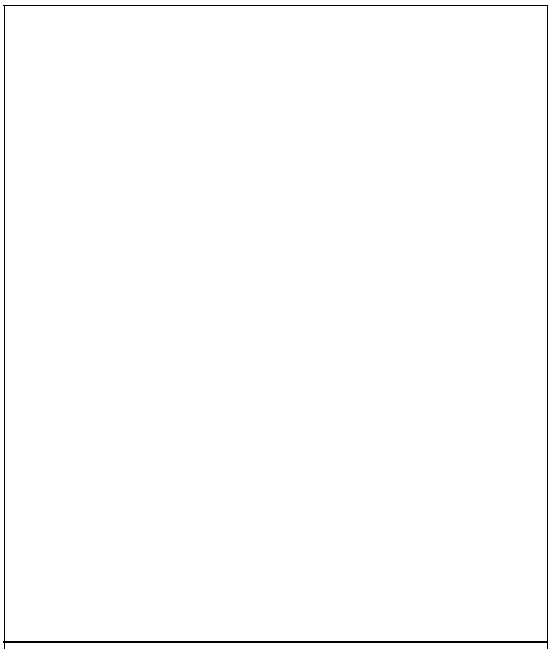


Figure 5-16: High-swing cascode current mirror with input branch included. (a) Non-preferred solution that suffers from a systematic mirror ratio error. (b) Solution that avoids the systematic mirror ratio error.

The final circuit of Figure 5-15(b) has been widely used in practice and is insensitive to process variations, such as global shifts in threshold voltage. Nonetheless, there are two remaining issues with this circuit that are worth mentioning. First, it is sometimes inconvenient to provide the extra current source used to bias M_6. Problems P5.5 and P5.6 look into alternative solutions that do not require an extra input current source, but still achieve low compliance voltage in the output branch.

The second issue stems from the backgate effect. In our analysis above, we have neglected the fact that the threshold voltage of M_2 will be larger than that of M_6. This is because the source of M_6 is connected to ground (and thus $V_{SB6} = 0$), while the source potential of M_2 is positive (and thus $V_{SB2} > 0$). As a result, assuming a sizing ratio of 1/4, V_{DS1} is more accurately given by

$$\begin{aligned} V_{DS1} &= V_{GS6} - V_{GS2} \\ &= V_{Tn0} + 2V_{OV} - (V_{Tn}(V_{SB2}) + V_{OV}) \\ &= V_{OV} - (V_{Tn}(V_{SB2}) - V_{Tn0}) = V_{OV} - \Delta V_{Tn} \end{aligned} \tag{5.26}$$

where ΔV_{Tn} is a positive quantity that causes M_1 to enter the triode region, unless sufficient margin is provided. In practice, the designer can use computer simulations to ensure that sufficient saturation margin is guaranteed. Another option is to change the circuit to mitigate this problem at its root. The

thought process that leads to the alternate solution is illustrated in Figure 5-17.

The biasing transistor M_6, as discussed previously, is redrawn in isolation in Figure 5-17(a). Figure 5-17(b) shows an equivalent circuit that breaks M_6 into four transistors, each with an aspect ratio of W/L. Assuming that the ideal square law model holds, the series connection of these transistors behaves like a MOSFET with aspect ratio of $W/(4L)$, or $(W/4)/L$ (see problem P2.2). Consequently, V_{G2} must be equal to $V_{Tn} + 2V_{OV}$, as in the original circuit of Figure 5-17(a). Furthermore, notice that the transistor M_{6a} in Figure 5-17(b) must operate in the saturation region (since it is diode connected). This means that this MOSFET's gate-source voltage is equal to $V_{Tn} + V_{OV}$, and the potential at its source node is equal to V_{OV}.

Next, in Figure 5-17(c), the three bottom transistors are lumped into a single device, again based on the argument that a device with an aspect ratio of $W/(3L)$ can be replaced with one that has $(W/3)/L$. Note that the combined transistor [M_{6b} in Figure 5-17(c)] operates in the triode region, since its drain-source voltage (V_{OV}) is smaller than $V_{GS6b} - V_{Tn} = 2V_{OV}$. Of course, all of the above conceptual arguments can be validated quantitatively, by carrying out a first-principle analysis using MOSFET I-V equations.

Figure 5-17: Conceptual steps for replacing the $W/4$ cascode biasing device with a two-transistor compound circuit.

The main advantage of the circuit in Figure 5-17(c) becomes apparent when it is inserted back into the cascode current mirror, as shown in Figure 5-18. Since M_{6a} has the same W/L and carries the same current as M_2, the source potential of these transistors is identical (neglecting channel length modulation). Hence, the error term due to backgate effect that we saw in Eq. (5.26) is suppressed, since $V_{SB2} = V_{SB6a}$.

Figure 5-18: Complete high-swing cascode current mirror using a triode device (M_{6b}) for cascode biasing.

$$\begin{aligned} V_{DS1} &= V_{DS6b} + V_{GS6a} - V_{GS2} \\ &= V_{OV} + V_{Tn}(V_{SB6a}) + V_{OV} - (V_{Tn}(V_{SB2}) + V_{OV}) \\ &= V_{OV} \end{aligned} \tag{5.27}$$

Finally, note that even though the circuit of Figure 5-18 provides a somewhat less error prone setup for the generation of V_{G2}, the designer will still want to leave margin and back off from the ideal $W/3$ sizing for M_{6b}. Table 5-5 lists the margin for various integer choices larger than 3 (see also problem P5.4).

Table 5-6: V_{DS1} as a function of the width scaling factor m for M_{6b} in Figure 5-18.

m	$\boldsymbol{V_{DS1}}$	$\boldsymbol{V_{DS1} - V_{OV}}$ **(Margin)**
3	V_{OV}	0
4	1.24 V_{OV}	0.24 V_{OV}
5	1.45 V_{OV}	0.45 V_{OV}
6	1.64 V_{OV}	0.64 V_{OV}
7	1.83 V_{OV}	0.83 V_{OV}
8	2 V_{OV}	V_{OV}

Example 5-4: Design of a Cascode Current Mirror

The cascode current mirror in Figure 5-18 is configured such that $I_{OUT} = I_{IN}$. In this example, we wish to design a similar current mirror that sets $I_{OUT} = 4I_{IN} = 400\ \mu A$. M_1 and M_2 are to be sized such that $V_{OV} = 200$ mV and using a channel length of 2 µm. M_{6b} should be sized such that $m = 5$. Given these specifications, determine all transistor widths and node voltages. Also compute the circuit's output compliance voltage and output resistance (R_{out}). Neglect channel length modulation in bias point calculations.

SOLUTION

The width of M_1 and M_2 is found by solving

$$I_{OUT} = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}V_{OV}^2$$

for W, and inserting the given numbers and technology parameters. This yields $W_1 = W_2 = 800$ µm. To implement the current ratio $I_{OUT}/I_{IN} = 4$, we require $W_3 = W_4 = W_{6a} = W_1/4 = 200$ µm and $W_{6b} = W6a/5 = 40$ µm.

The voltage V_{G1} is simply $V_{Tn} + V_{OV} = 0.7$ V. With $m = 5$, we know from Table 5-5 that $V_{D1} = V_{D3} = V_{D6b} = 1.45V_{OV} = 290$ mV. To compute V_{G2} accurately, we must first estimate the threshold voltage of M_{6a} using Eq. (5.1).

$$V_{Tn}(V_{SB}) = V_{TOn} + \gamma_n(\sqrt{2\phi_f + V_{SB}} - \sqrt{2\phi_f})$$

Evaluating the above equation with $V_{SB6a} = V_{D6b} = 290$ mV gives $V_{T6a} = 590$ mV. Therefore, $V_{G2} = 1.45V_{OV} + V_{T6a} + V_{OV} = 1.08$ V. The circuit's output compliance voltage is $V_{OUTmin} = V_{D1} + V_{OV} = 0.49$ V. The circuit's output resistance is given by $R_{out} \cong r_{o1}\cdot g'_{m2}r_{o2}$ (see Section 4-5-1). Therefore, we compute

$$g_{m2} = \frac{2I_{D2}}{V_{OV}} = \frac{2\cdot 400\mu A}{200 mV} = 4mS$$

$$g'_{m2} = g_{m2}\left(1 + \frac{\gamma_n}{2\sqrt{2\phi_f + V_{SB}}}\right)$$

$$= 4mS\left(1 + \frac{0.6V^{1/2}}{2\sqrt{0.8V + 0.29V_{SB}}}\right) = 5.15mS$$

and

$$r_{o1} \cong r_{02} \cong \frac{1}{\lambda_n I_{D1}} = \frac{1}{0.05V^{-1}\cdot 400\mu A} = 50k\Omega$$

These numbers lead to $R_{out} \cong 12.88\ M\Omega$. The schematic below summarizes the results obtained in this example.

5-4 Current References*

The current mirror circuits discussed in the previous section are useful for replicating and distributing bias currents within a sub-circuit or an entire chip. Ultimately, however, the currents that are being distributed must originate from some form of a reference current generator (see Figure 5-1).

Over the years, a wide variety of current references have been developed, each having specific pros and cons for the intended application. Within the scope of this introductory module, we will consider only two examples, primarily as a starting point for further reading and to complete the picture on how a complete biasing network within a larger chip might be constructed. For a more comprehensive discussion, the reader is referred to advanced texts such as Reference 2.

We begin by considering the most basic of all possible reference generator circuits, shown in Figure 5-19. This circuit is essentially a current mirror, with its input branch tied to the supply via a resistor. In this circuit, we have

$$I_{OUT} \cong I_{IN} = \frac{V_{DD} - V_{Tn} - V_{OV}}{R} \qquad (5.28)$$

From this result, noting that typically $V_{DD} >> V_{Tn}$, and $V_{DD} >> V_{OV}$, we see that the current is roughly proportional to the supply voltage. Given the variations in supply voltage that a robust circuit must withstand (see Table 5-1), this solution is often not suitable for all but relatively primitive and low-performance circuits. What we desire is a current generator that is (to first order) insensitive to supply variations. The so-called self-biased constant-g_m current generator discussed next is an example of an improved circuit that is frequently used in practice.

To understand the self-biased constant-g_m current generator, consider first the circuit shown in Figure 5-20(a). This is a current mirror-like circuit with a resistor R added in the source of M_2. Assuming that M_2 is scaled m times wider than M_1, and letting $R = 0$ for the time being, we know that I_{OUT} is approxi-

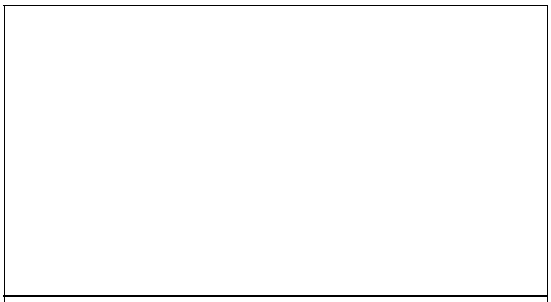


Figure 5-19: Simple supply-referenced current generator.

mately equal to $m \cdot I_{IN}$. This is illustrated using the dashed line (i) in the graph of Figure 5-20(b). Line (ii) is included for reference, corresponding to $m = 1$, i.e., $I_{OUT} = I_{IN}$. Now, assuming $m > 1$ and $R > 0$, we know that I_{OUT} must be smaller compared to case (i) with $R = 0$. This is because the voltage drop across R reduces the gate-source voltage of M_2 and consequently results in smaller I_{OUT}. As I_{IN} increases, I_{OUT} [curve (iii)] bends away further and further from line (i) and ultimately intersects with line (ii). While it is possible to derive a closed-form equation of this curve (see problem P5.7), we focus our attention on point P.

A particularly interesting property of point P is that it defines an absolute current level that (to first order) depends only on the MOSFET sizes and R, i.e., it is independent of the supply voltage. In order to build a current reference that utilizes this point, a few extra transistors must be employed, as shown in Figure 5-21. First focus on M_3 and M_4. These transistors form a current mirror that forces $I_{OUT} = I_{IN}$, which is necessary for operation at point P [see Figure 5-20(b)].

Unfortunately, simply forcing $I_{OUT} = I_{IN}$ does not guarantee that the circuit operates at P. There exists another (undesired) point where $I_{OUT} = I_{IN} = 0$, labeled U in Figure 5-20(b). If only M_1 - M_4 were present in this circuit, it would not be clear which operating point the circuit will chose when the supply voltage is turned on. The outcome may depend on second-order effects, such as parasitic capacitive coupling, and how quickly the supply ramps up. In order to guarantee that the circuit will eventually operate at point P, the designer will always include a so-called start-up circuit. This circuit is formed by M_6 - M_8 in Figure 5-21.

To understand the operation of the start-up circuit, consider first the case where the circuit starts up in point U, i.e., $I_{OUT} = I_{IN} = 0$. This condition necessitates that $V_{GS2} < V_{Tn}$, since no drain current is flowing in M_2. With $V_{GS2} < V_{Tn}$, M_7 will be off and M_6 will be on, pulling the voltage at node V_{START} toward V_{DD}. Thus, M_8 will turn on and force a drain current into M_3, which will subsequently be mirrored into M_4 and M_2. Therefore, the circuit has no choice but to leave point U and ultimately arrive at P, which is the only other possible DC operating point.

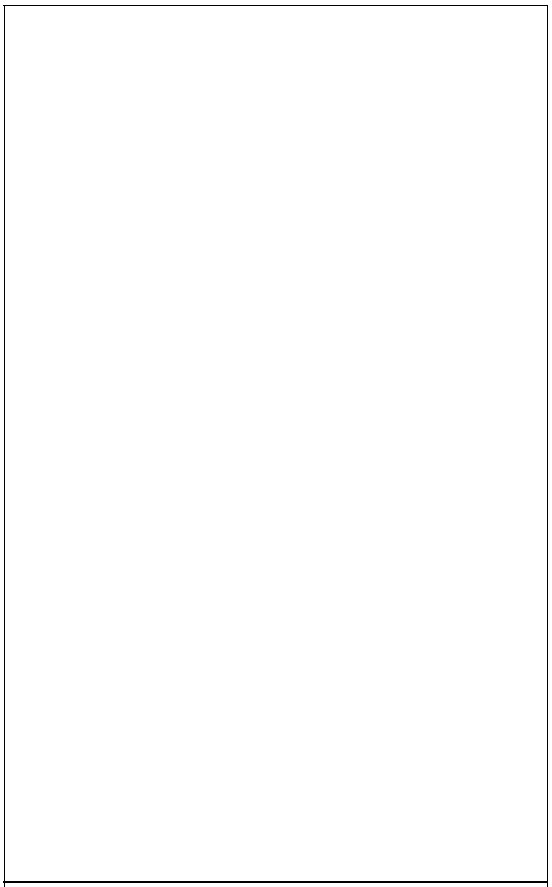


Figure 5-20: (a) Core building block of a constant-g_m current generator. (b) Current transfer characteristics for various scenarios.

Once point P is reached we have $V_{GS2} > V_{Tn}$ by some overdrive voltage, typically a few hundred milivolts. For the given V_{GS2} in this point, M_7 must be sized large enough so that V_{START} lies near ground, and no current flows in M_8 ($I_{START} = 0$). In a typical design M_7 is much larger than M_6, resulting in a so-called low-threshold inverter.

Our final task is to compute the current $I_{REF} = I_{IN} = I_{OUT}$ in Figure 5-21. We begin by applying KVL around the gate-source voltages of M_1 and M_2.

$$I_{REF}R = V_{GS2} - V_{GS1} \tag{5.29}$$

Neglecting backgate effect, i.e., assuming equal threshold voltages for M_1 and M_2 (for simplicity), Eq. (5.29) becomes

$$I_{REF}R = V_{OV2} - V_{OV1} \tag{5.30}$$

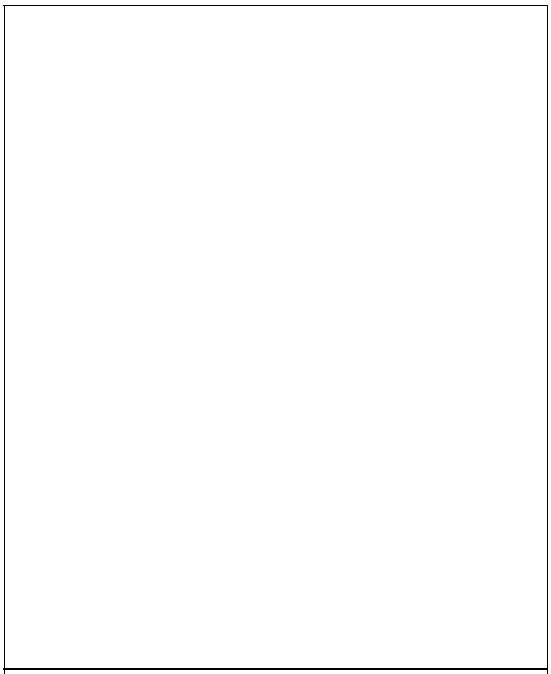


Figure 5-21: Complete self-biased constant-g_m current reference circuit.

Now, since for a MOSFET

$$V_{OV} = \sqrt{\frac{2I_D}{\mu_n C_{ox}\frac{W}{L}}} \tag{5.31}$$

and M_1 is m times wider than M_2, Eq. (5.30) can be re-written as

$$I_{REF} = \frac{V_{OV2}\left(1-\frac{1}{\sqrt{m}}\right)}{R} \tag{5.32}$$

Finally, eliminating V_{OV2} using Eq. (5.31) and solving for I_{REF} gives

$$I_{REF} = \frac{2(\sqrt{m}-1)^2}{m}\cdot\frac{1}{\mu_n C_{ox}\frac{W_n}{L_n}\cdot R^2} \tag{5.33}$$

This equation is primarily useful for setting the absolute current level in the circuit, and at first glance does not seem to have any special structure. A much more important result from the above analysis follows from considering the transconductance of M_2, given by

$$g_{m2} = \frac{2I_{D2}}{V_{OV2}} = \frac{2I_{REF}}{V_{OV2}} = \frac{2\left(1-\frac{1}{\sqrt{m}}\right)}{R} \tag{5.34}$$

As we see from this result, g_{m2} depends only on the resistance R and the scaling factor m, i.e., the transconductance will not be affected by MOSFET process and temperature variations. In a way, the circuit "re-computes" I_{REF} such that the transconductance is held constant to the value given by Eq. (5.34). This is the reason why this circuit is typically called a constant-g_m reference generator, as mentioned earlier. It should be noted, of course, that not only the transconductance of M_2 is held constant when this circuit is used. Any other MOSFET that utilizes I_{REF} or a copy of this current will behave similarly.

In practice, the device type used to implement resistor R should be chosen with care. When implemented on-chip, the designer will often opt for a highly doped polysilicon resistor that has relatively small process variations and a small temperature coefficient (see Table 5-2 and Table 5-3). Alternatively, the resistance is sometimes placed off-chip, where it can be realized, for example, with a 1%-accurate and low temperature coefficient metal film resistor[2].

As a final note, we should emphasize that the foregoing analysis neglected many second-order effects, such as channel length modulation and backgate effect. In practice, these effects can have some bearing on the circuit's accuracy and therefore leave room for improvements (such as including cascode transistors). The interested reader will find many articles on this topic in analog circuit literature and advanced texts, such as Reference 2.

5-5 Voltage Biasing Considerations

In addition to bias currents, building a complete analog circuit will require the generation of various bias voltages needed to operate common-source, common-gate and common-drain stages (see for example V_{B1} and V_{B2} in Figure 5-1). In this section we will discuss an exemplary subset of solutions that have found their use in practice.

As we have already seen in Example 5-1, the common-source stage is very sensitive to variations in its input bias voltage. As a result, a majority of practical CS circuits are embedded in feedback networks that regulate the input bias voltage to the proper value, thereby absorbing process variations and mismatch effects. Since feedback is beyond the scope of this module, the input biasing techniques suggested for the common-

2. When the resistor is placed off-chip, the designer must take great care to avoid stability issues. Using feedback circuit analysis techniques, it can be shown that even relatively small amounts of parasitic capacitance at the source node of M_2 can make the circuit oscillate.

source stage are meant to be applied only to a sub-set of applications where the circuit's voltage gain is low (typically < 10), and the amplifier is utilized "open-loop", without a feedback network. This complication typically does not exist for CG and and CD stages, and the proposed circuits are therefore more or less universally applicable.

5-5-1 Voltage Biasing for a Common-Source Stage

Due to the voltage gain of a CS amplifier, its input bias voltage usually cannot be set to a fixed voltage without causing prohibitive sensitivities to component variations and mismatch. Thus, it is important to design the bias circuitry with variability in mind and construct solutions that can track or absorb any significant deviations from nominal parameter conditions. Especially for common-source stages, solutions applied in practice often involve the use of feedback or differential circuit topologies (see Reference 2). Since these topics are beyond the scope of this module, we will concentrate here only on a few basic ideas that can be understood with the prerequisites established so far.

Specifically, we will focus in this sub-section on a few possible solutions to the problem encountered in Example 5-1. The main problem in this example was that the input bias voltage was held constant, while the threshold voltage and other parameters in the circuit changed due to process variations. Ideally, we would like to "automatically compute" the input bias voltage of the stage such that it tracks the required value across process corners.

A first option that accomplishes this is shown in Figure 5-22. Here, M_1 is the MOSFET that implements the common-source amplifier and v_s and R_s model a transducer that generates the voltage we wish to amplify. The transistor M_2 is a **replica device** that computes the proper gate-source voltage required to carry the current I_B. Note that this overall arrangement resembles a current mirror, which we have already determined to be insensitive to process variations. If V_{Tn} or $\mu_n C_{ox}$ change, the gate-source voltage of M_1 (V_B) adjusts so that this transistor's drain current remains equal to I_B. This means (to first order) that no current flows into the resistive divider formed by R_1 and R_2. These resistors can be sized to establish the desired output quiescent point and voltage gain. For example, for $R_1 = R_2$ and $V_{DD} = 5$ V, we have $V_{OUT} = 2.5$ V, approximately independent of process and temperature.

While the above-discussed circuit will work robustly, it has one big limitation in that both transducer terminals must be accessible and compatible with the bias voltage desired for M_1. One possibility for overcoming this constraint is to employ AC coupling (see Figure 5-23). AC coupling means that the transducer signal is coupled into the circuit via a capacitor. In the circuit of Figure 5-23 R_{large} and C_{large} form a first-order high pass filter with corner frequency $\omega_c = 1/R_{large}C_{large}$ (neglecting the resistance $1/g_{m2}$, which is in series with R_{large}). To avoid filtering the signal, ω_c must be chosen smaller than the smallest frequency of interest. For instance, if we are interested in amplifying a 20 Hz signal (the lower end of the audio frequency spectrum), we need $R_{large}C_{large} > 1/(2\pi \cdot 20\text{ Hz}) \cong 4$ ms. Assuming we can comfortably integrate resistances up to 100 kΩ on our chip, this means that $C_{large} > 4$ nF. Such a large capacitance is typically impractical for integration on chip and would have to be realized as an external component.

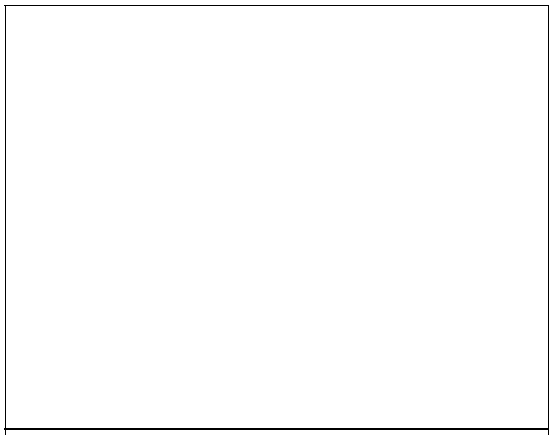


Figure 5-22: Replica biasing approach for setting up the quiescent point gate-source voltage of a common-source stage.

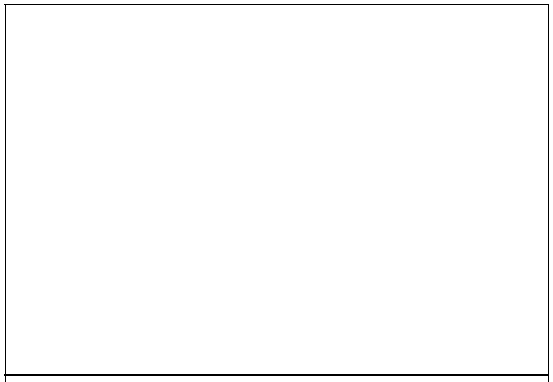


Figure 5-23: Replica biasing approach using AC coupling.

A shortcoming of the circuit in Figure 5-23 is that the resistors R_{large} and R_s form a voltage divider, which can be detrimental when R_s is very large. Figure 5-24 shows an alternate AC coupling approach in which the transducer can be directly connected to the MOSFET gate. In this circuit, the bias current I_B is injected into the drain of the common-source transistor

(M_1) and extracted again using the current mirror formed by M_2 and M_3. The bias point voltage at node X is given by $V_B - V_{GS1}$, which places constraints on the minimum required value for V_B. Note however, that V_B does not have to be accurately set or track process variations; as long as M_1 and M_2 are in saturation, node X tracks (DC) changes in V_B and the circuit remains properly biased. As far as the signal is concerned, the capacitor C_{large} establishes an AC ground at the source of M_1 beyond the high-pass corner frequency of the circuit. Just as in the previous circuit, it can be shown that the AC coupling capacitor must take on large values to enable the passing of low frequencies through the circuit.

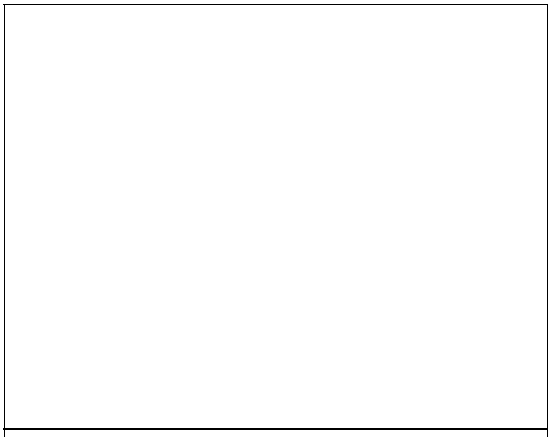


Figure 5-24: Biasing-approach using AC coupling at the source of a common-source stage.

As we have seen from the previous examples, achieving proper biasing together with the processing of low-frequency signals in a basic common-source stage comes with some undesired constraints and restrictions. Many of these issues can be mitigated when the signal is present in the form of a current, originating for example from a common-gate stage that is driving the common-source amplifier. We will see an example of such a circuit in Chapter 6.

5-5-2 Voltage Biasing for a Common-Gate Stage

Compared to a common-source stage, setting up the bias voltage for the gate of a common-gate stage is usually less intricate. To see this, we consider two classical usage examples shown in Figure 5-25.

In Figure 5-25(a) the common-gate device M_2 is utilized in a cascode stage. Since a cascode stage is often designed for large voltage gain, a typical objective is to maximize the available output voltage swing. Consequently, the gate bias voltage of M_2 is setup in the same way as in the high-swing cascode

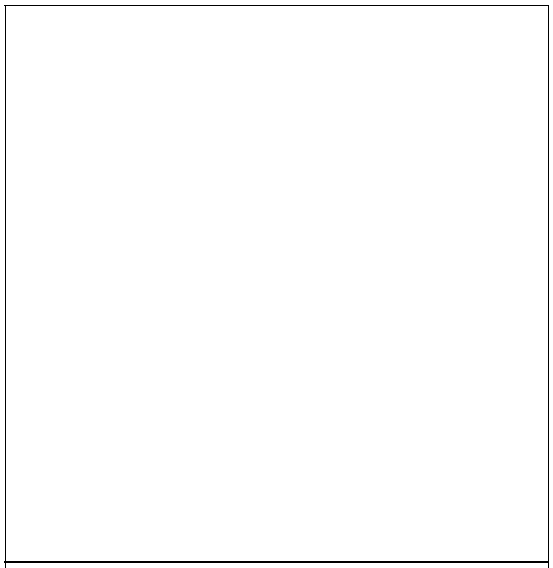


Figure 5-25: Voltage biasing in two usage cases of a common-gate stage.

current mirror discussed in Section 5-3-4, which means the drain-source voltage of M_1 is set to V_{DSsat1} plus some margin for robustness and tolerance to mismatches. A reasonable margin is achieved by using $m = 5$ (see Table 5-6) in the sizing of M_{3b}.

Figure 5-25(b) shows an example where a common-gate stage is used to interface to a photo diode. The signal current generated in the photodiode passes through M_2 and causes a proportional voltage swing at the output. In this circuit, the output swing is usually not very large, and thus the gate bias voltage for M_2 is not tightly constrained by voltage swing requirements. Typically, the gate voltage is set such that the photo diode is biased at a suitable reverse bias. This is accomplished by sizing R_1 and R_2 appropriately.

In both of the circuits in Figure 5-25, variations in the transistor parameters (such as V_{Tn}) will cause the overall operating point of the circuits to shift. However, unlike the common-source stage of Example 5-1, these circuits are not very sensitive to such shifts. For instance, if the threshold voltage of M_2 in Figure 5-25 changes by 100 mV, all this means is that the reverse bias voltage of the diode will change by approximately the same amount. If properly designed (with margins), this won't cause the circuit to fail or behave improperly. This strongly contrasts the situation with the circuit of Example 5-1, where such changes in the threshold voltage can have detrimental effects on the stage's operation.

5-5-3 Voltage Biasing for a Common-Drain Stage

In a common-drain stage, the input and output voltages at the quiescent point are directly coupled. As shown in Figure 5-26, $V_{OUT} = V_{IN} - (V_{Tn} + V_{OV})$. Proper voltage biasing in a common-drain stage boils down to making sure that the input and output quiescent point voltages are compatible with the circuits that are connecting to the stage input and output. As in a common-gate stage, variability in transistor parameters often does not have detrimental effects as long as a proper margin is included in the design.

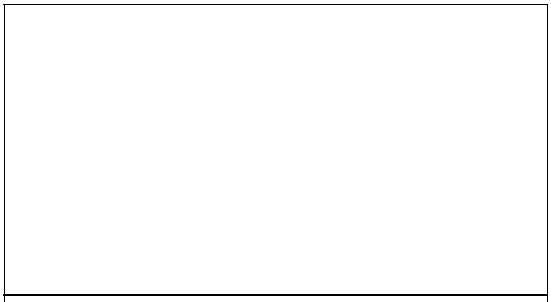


Figure 5-26: Relationship between the input and output quiescent points in a common-drain stage.

In some applications, the shift between the input and output quiescent point is undesired. In this case, a p-channel common-drain stage can be used to provide a shift in the opposite direction (see Figure 5-27). In this circuit, M_1 can be sized such that the quiescent points V_{IN} and V_{OUT} are approximately equal.

When a common-drain stage is employed primarily to shift quiescent points, the designer calls this circuit a **level shifter**. Level shifters are generally useful to interface two stages that are otherwise incompatible in terms of their ideal quiescent point output/input voltages.

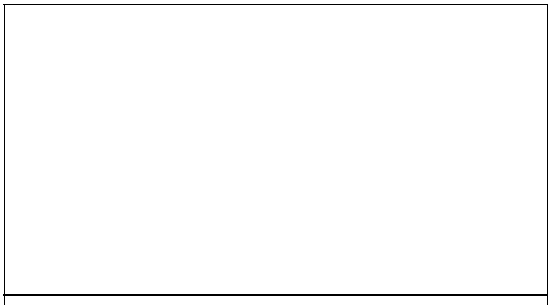


Figure 5-27: Back-to-back common-drain stages to realize equal input and output quiescent point voltages.

Summary

In this chapter, we have surveyed general considerations and basic circuits related to the voltage and current biasing of elementary transistor stages. We have seen that the variability inherent to CMOS process technology influences the design and architecture of these support circuits and ultimately determines whether a certain biasing scheme can be deemed practical. We analyzed the basic current mirror and its cascoded variant with respect to their non-idealities and articulated some of the most important design guidelines. As an example of a reference current generator, we looked at the so-called constant-g_m biasing circuit and analyzed its first order behavior. Finally, this chapter looked into the problem of voltage biasing for the three elementary stage configurations. We determined that in lieu of feedback, biasing a common-source stage properly is most challenging and must be considered with care and knowledge of relevant process variation and mismatch effects. While most of the presented ideas and circuits were presented in the context of simple application examples, they generally also apply to more complex circuit designs studied in advanced literature.

References

1. M. J. M. Pelgrom, A. C. J. Duinmaijer, and A. P. G. Welbers, "Matching properties of MOS transistors," *IEEE Journal of Solid-State Circuits*, vol. 24, no. 5, pp. 1433-1439, Oct. 1989.

2. P. R. Gray, P. J. Hurst, S. H. Lewis, and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 5th Edition, Wiley, 2008.

Problems

Unless otherwise stated, use the standard model parameters specified in Table 4-1 for the problems given below. Consider only first-order MOSFET behavior and include channel-length modulation (as well as any other second-order effects) only where explicitly stated.

P5.1 Consider the bias current generator circuit of Figure 5-19.

(a) Compute the current I_{IN} assuming nominal MOSFET parameters and supply voltage (V_{DD} = 5 V).

(b) Re-compute I_{IN} for slow MOSFET parameters (see Table 5-2) and V_{DD} = 4.5 V. Repeat for fast parameters and V_{DD} = 5.5 V.

(c) What are the percent errors of the currents found in part (b), relative to the nominal current computed in (a)?

P5.2 In Example 5-1, we showed analytically that changing the MOSFET parameters from nominal to fast pushes the transistor into the triode region. Construct a load line plot that shows this graphically. That is, draw the output curves of the MOSFET (I_D versus V_{DS}) for the two corner cases and show how the intersect with the load line shifts when fast parameters are assumed. Be sure to neglect channel length modulation.

P5.3 Set up a suitable analysis that allows you derive the values given in Table 5-5. Set up an equation that computes V_{DS1} as a function of k and V_{OV}.

P5.4 Set up a suitable analysis that allows you derive the values given in Table 5-6. Set up an equation that computes V_{DS1} as a function of m and V_{OV}.

P5.5 The circuit in Figure P5-5 can be used to achieve high-swing cascode biasing without an extra input current branch. Given the annotated bias point voltages, what is the proper W/L ratio for M_4 that achieves the minimum output compliance voltage?

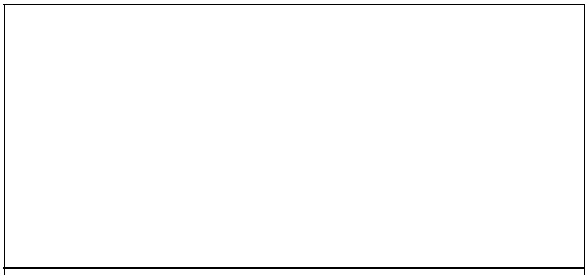


Figure P5-5.

P5.6 The circuit in Figure P5-6 is called a "Sooch" cascode current mirror. It uses one single branch for setting up all bias voltages for a high-swing cascode current mirror. Given the annotated bias point voltages, what is the proper W/L ratio for M_5 that achieves the minimum output compliance voltage? What is the minimum required voltage across the input branch (V_{IN})?

P5.7 Derive a closed-form expression for curve (iii) in Figure 5-20(b). Verify that the intersect with line (ii) corresponds to the current level given in Eq. (5.34).

P5.8 For the circuit of Example 5-1, compute the proper V_{IN} that would need to be applied in the fast parameter case such that the output bias voltage remains the same as in the nominal case. In this biasing condition, what is the voltage gain, and by which percentage has it changed relative to the nominal case?

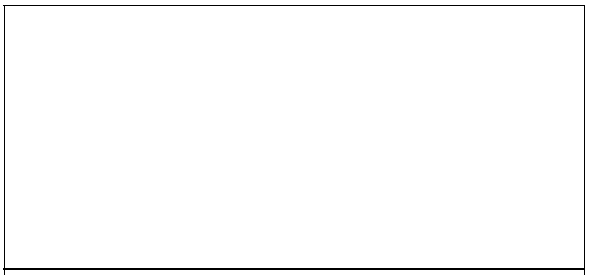


Figure P5-6.

P5.9 Figure P5-9 shows a cascode current source consisting of M_{1A} and M_{1B}, and a single transistor current source consisting of M_2. Assume that the cascode current source is optimally biased, i.e., V_{B1B} is chosen such that $V_{DS1a} = V_{DS1a,sat} = V_{OV1a}$. Assume also that both current sources supply the same current I_O. Neglect backgate effect.

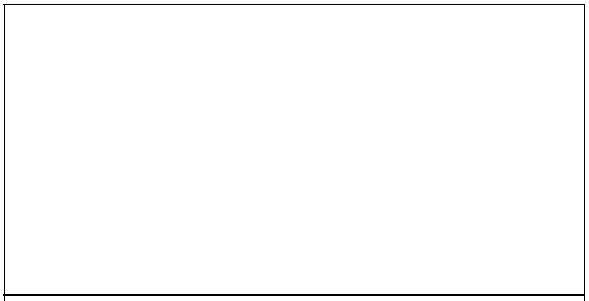


Figure P5-9.

(a) Find relationships between W_1, L_1 and W_2, L_2 such that both current sources have the same parasitic output capacitance, and the same output compliance voltage V_{Omin} that keeps all the devices saturated. For simplicity, assume $\lambda = 0$ in this part of the analysis. Note: The parasitic capacitance at the drain of M_2 is given by $C_{db} + C_{gd}$. Similarly, assume that the output capacitance of the cascode current source is approximately equal to $C_{db} + C_{gd}$ of M_{1b}. (In the cascode current source, the effect of other capacitances referred to the output node is negligible.)

(b) Using the result from part (a), show that the expression given below must hold. R_{O1} and R_{O2} are the output resistances of each current source, as indicated in Figure P5-9.

$$\frac{R_{O1}}{R_{O2}} = \frac{g_{m1}r_{o1}}{4}$$

(c) Calculate V_{Omin}, R_{O1} and R_{O2} for I_O = 100 µA and $(W/L)_1$ = 10 µm/2 µm. (Use the relationships between device sizes from part (a).)

P5.10 The circuit shown in Figure P5-10 is a so-called self-biased, V_{Tn}-referenced current generator. Assuming $(W/L)_1$ = 50 and $(W/L)_3 = (W/L)_4 = (W/L)_5$, find the value for R so that I_{OUT} = 100 µA. Assume λ = 0 and neglect backgate effect.

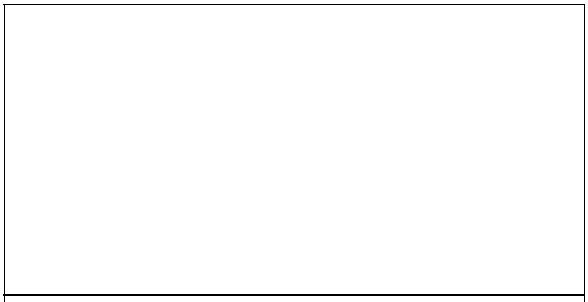


Figure P5-10.

P5.11 For the circuit shown in Figure P5-11, ignore the backgate effect and finite output resistance unless otherwise stated. All devices have identical widths and lengths and operate in saturation.

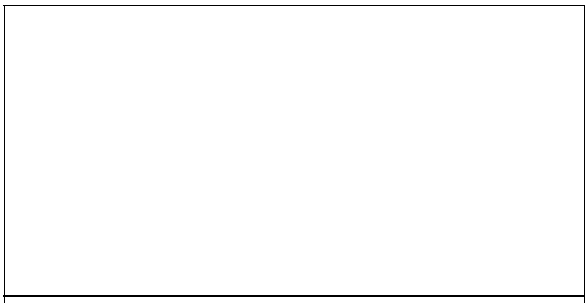


Figure P5-11.

(a) Calculate R such that the drain-source voltage of M_1 is 1.5 times its gate overdrive, i.e., $V_{DS1} = 1.5V_{OV1}$.

(b) Suppose that due to random mismatch, the threshold voltage of M_1 is 10 mV larger than that of all the other transistors. What is the percent error in I_{OUT} caused by this mismatch? Use appropriate small-signal approximations in your calculation.

(c) Suppose now that the threshold voltage of M_2 is increased by 10 mV while all other thresholds are at their nominal value. What is the percent error in I_{OUT} caused by this mismatch? In this calculation, include the effect of the finite output resistance for M_1, assuming $g_m r_o$ = 50.

P5.12 In the circuit of Figure P5-12, V_{G1} is adjusted such that I_{D1} = 50 µA. The W/L ratio of M_1 and M_2 is equal to 4, while M_3 is sized such that W/L = 1/2.

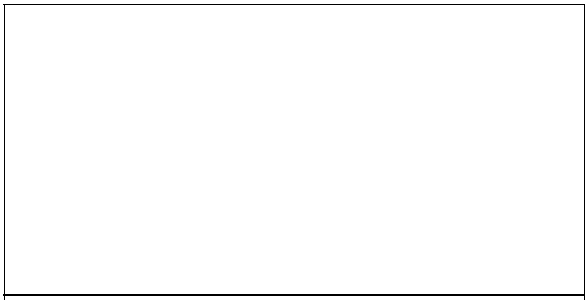


Figure P5-12.

(a) Ignoring backgate effect, compute the minimum and maximum values of V_{G2} for which all transistors remain in saturation.

(b) Repeat part (a) with backgate effect included.

P5.13 For the circuit of Figure P5-13, compute the compute the width ratios W_8/W_6 and W_7/W_6 such that $V_{DS6} = V_{DS7} = V_{OV8}$. Assume that all channel lengths are identical. Ignore the backgate effect and channel-length modulation.

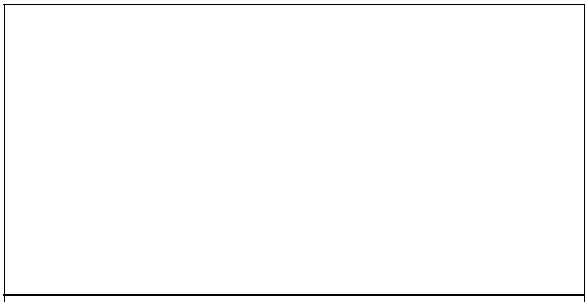


Figure P5-13.

P5.14 For the circuit of Figure 5-21, size the transistors M_1 - M_4 and R such that I_{OUT} = 100 µA, V_{OV} = 300 mV. Assume that all channel lengths are equal to 3 µm. Neglect backgate effect and channel length modulation.

P5.15 Design a cascode current mirror circuit using n-channel devices assuming the following specifications. The circuit should take an input current of 10 µA and generate three outputs at 20 µA, 50 µA and 100 µA, respectively. The output

compliance voltage should be no larger than 800 mV and the gate overdrive of the transistors should be designed as large as possible (for immunity to mismatch), while maintaining a reasonable saturation margin. Draw the complete circuit diagram including all device sizes. This problem does not have a unique solution.

P5.16 Draw a layout (using any tool you prefer) for the circuit designed in Example 5-3.

6

CHAPTER

Multistage Amplifiers

As we have already learned in Chapter 1, amplifier circuits can be grouped into four categories: voltage, current, transconductance and transresistance amplifiers, depending on whether the intended input and output signals are voltages or currents. While it is in principle possible to construct each one of these amplifier types using a single-stage circuit, the designer will usually combine multiple stages for improved performance. Generally speaking, multistage amplifiers are used to increase the gain and/or transform input and output resistances for minimum signal attenuation at the ports of the amplifier circuit.

Several issues must be understood and addressed when designing multistage transistor amplifiers. First, the DC biasing that sets the quiescent node voltages and currents must be properly chosen so that the stages can be directly coupled. Second, proper approximations must be applied so that the circuit's frequency response can be obtained by hand and becomes transparent for design. Lastly, multi-stage circuit design necessitates a systematic optimization approach to handle the increased number of design variables and degrees of freedom. This chapter covers elements of each one of these aspects through a variety of examples.

Chapter Objectives

- Design and analyze the low-frequency gain and input and output resistances of multistage amplifiers based on cascading single-stage amplifiers.
- Analyze the frequency response of multistage amplifiers using suitable approximations.
- Illustrate an example of a systematic design procedure for a three-stage transresistance amplifier.

6-1 Low-Frequency Analysis

In the following treatment, several examples of cascading the two-port models of single-stage amplifiers will be used to help us understand how multistage amplifiers can achieve increased gain and transform input/output resistances. The desired input and output resistances, as well as high gain, can be achieved with proper selection of the constituent single-stage amplifiers. For the time being, we will limit the discussion to low-frequency behavior, and address the analysis of frequency response in the next section. For our discussion, we will utilize the three prototype amplifier configurations shown in Figure 6-1.

6-1-1 Voltage Amplifier

Recall that a voltage amplifier requires a high input resistance, a low output resistance, and (typically) a large voltage gain. From Chapter 2 we know that a common-source amplifier has an infinite input resistance since the MOS transistor has an insulating gate with no input current. Therefore, assuming that a common-source amplifier is the proper input stage for a volt-

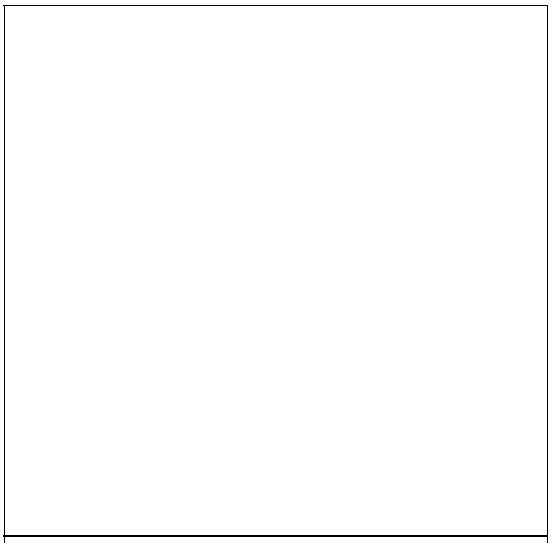


Figure 6-1: Amplifier configurations assumed in this section.

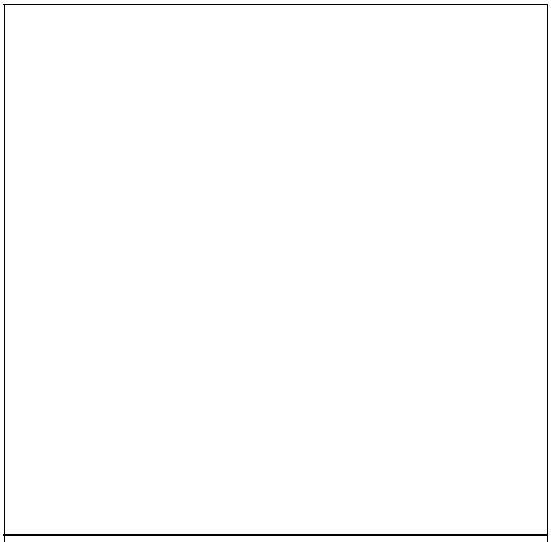


Figure 6-2: Small-signal model for two cascaded common-source amplifiers.

age amplifier, we can explore cascading two of these stages to increase the voltage gain. The small-signal model of two cascaded common-source amplifiers is given in Figure 6-2.

The input resistance of this cascade is infinite and the overall open-circuit voltage gain is given by the multiplication of the voltage gain of each stage as shown in

$$A_v = g_{m1}(r_{o1} \| R_{D1}) g_{m2}(r_{o2} \| R_{D2}) \tag{6.1}$$

The output resistance of this amplifier is

$$R_{out} = r_{o2} \| R_{D2} \tag{6.2}$$

This cascaded common-source voltage amplifier has two of the three required characteristics, namely a large input resistance and a high voltage gain. However, assuming that R_{D2} is reasonably large for high gain, it still has a high output resistance. This will degrade the voltage transfer from the amplifier to the load resistor.

From Chapter 4 we know that a common-drain amplifier has an infinite input resistance, a low output resistance, and a voltage gain near unity (modeled as unity here for simplicity). We can cascade the small-signal, two-port model of a common-drain amplifier with the small-signal model of the common-source cascade described above. This three-stage amplifier is shown in Figure 6-3. As before, the cascaded common-source amplifiers are modeled with an infinite input resistance, a voltage gain A_v, and an output resistance given by Eq. (6.1) and Eq. (6.2), respectively.

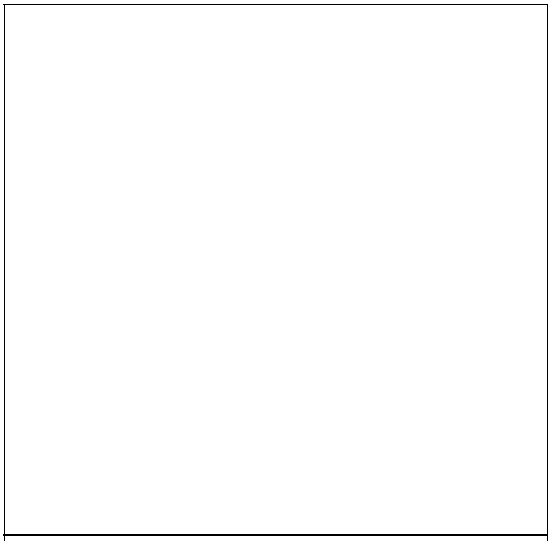


Figure 6-3: Small-signal model of a CS-CS-CD cascade.

There is no interstage loss of voltage gain because the common-drain amplifier has infinite input resistance. In addition, the output resistance is reduced to approximately equal to the reciprocal of the transconductance plus the backgate transconductance of the common-drain amplifier. This usually gives a significant reduction in output resistance and allows this three-stage voltage amplifier to drive small load resistances while still maintaining a significant transfer of the open-circuit voltage to the load.

6-1-2 Transconductance Amplifier

A transconductance amplifier requires a large input resistance, a large transconductance, and a large output resistance to be able to pass most of its output current to the load. Let us explore using cascaded common-source amplifiers for a transconductance amplifier. In Figure 6-4 we show the small-signal model of two cascaded common-source amplifiers. We use the Norton equivalent output network since current is the output variable of interest. The short circuit transconductance of this amplifier is equal to

$$G_m = \frac{i_{out}}{v_{in1}} = -g_{m1}(r_{o1} \| R_{D1})g_{m2} = A_{v1}g_{m2} \qquad (6.3)$$

Notice that the additional common-source stage increases the transconductance by the voltage gain of the first stage. There is no interstage loss since the input resistance to the second stage is infinite.

Figure 6-4: Small-signal model for a CS-CS cascade used as a transconductance amplifier.

The output resistance of the transconductance amplifier is

$$R_{out} = r_{o2} \| R_{D2} \qquad (6.4)$$

which is the output resistance of a single common-source amplifier stage.

We can add a current buffer to increase the output resistance of the cascaded common-source amplifier. An ideal current buffer is defined as a circuit whose input resistance is very small, output resistance is very large, and has a current gain of unity. The common-gate amplifier studied in Chapter 4 is a good example of a current buffer. The small-signal model of a common-gate amplifier cascaded with two common-source amplifiers is shown in Figure 6-5.

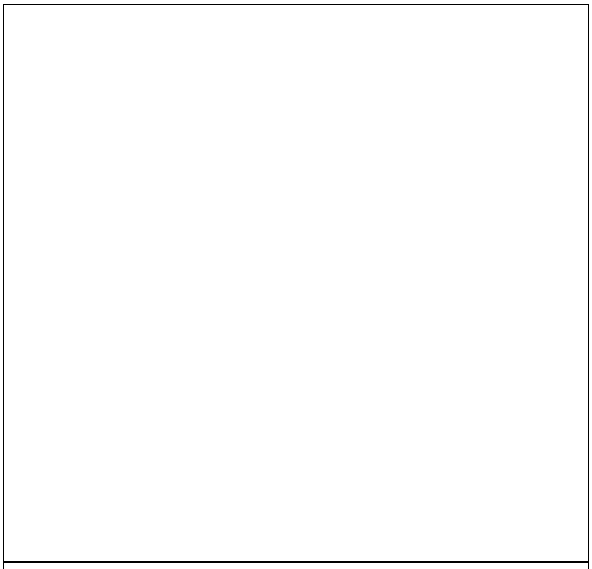


Figure 6-5: Small-signal model a CS-CS-CG cascade that forms a transconductance amplifier.

The output resistance is now

$$R_{out} \cong g_{m3}r_{o3}(r_{o2} \| R_{D2}) \| R_{D3} \qquad (6.5)$$

Assuming that we can make R_{D3} negligibly large (by supplying the drain current using a long-channel, cascoded current source), the R_{out} is increased by $g_{m3}r_{o3}$ with the help of the CG stage. The short circuit transconductance is given by

$$G_m = \frac{i_{out}}{v_{in}} = -g_{m1}(r_{o1} \| R_{D1})g_{m2}\left(\frac{r_{o2} \| R_{D2}}{r_{o2} \| r_{D2} + 1/g'_{m3}}\right) \cong = -g_{m1}(r_{o1} \| R_{D1})g_{m2} \qquad (6.6)$$

Note that the transconductance is only slightly degraded when compared to Eq. (6.3) due to the parallel combination of the common-source output resistance with the common-gate input resistance. This degradation is negligible since the CG input resistance is small compared to the CS output resistance.

Example 6-1: MOS Transresistance Amplifier

This example explores how to cascade two amplifier stages to form a transresistance amplifier. Select the stages and calculate R_{in}, R_{out}, and R_m. Assume that the employed CS and CG stages (Figure 6-1) have drain resistors of value $R_D << r_o$ in their output networks. Hint: Recall that a transresistance amplifier typ-

ically requires a low input resistance, a low output resistance, and a large transresistance.

SOLUTION

To obtain a low input resistance, we choose a CG amplifier as the first stage. The choice of the second stage depends on the specifications required. Let's try a CS amplifier. The small-signal two-port model for a CG-CS cascade is shown in Fig. Ex6-1A. R_{in}, the input resistance, is $1/g'_{m1}$. R_{out}, the output resistance, is equal to R_{D2}. We need to find the unloaded ($R_L \to \infty$) transfer function between v_{out} and i_{in1} to calculate R_m. We begin by writing

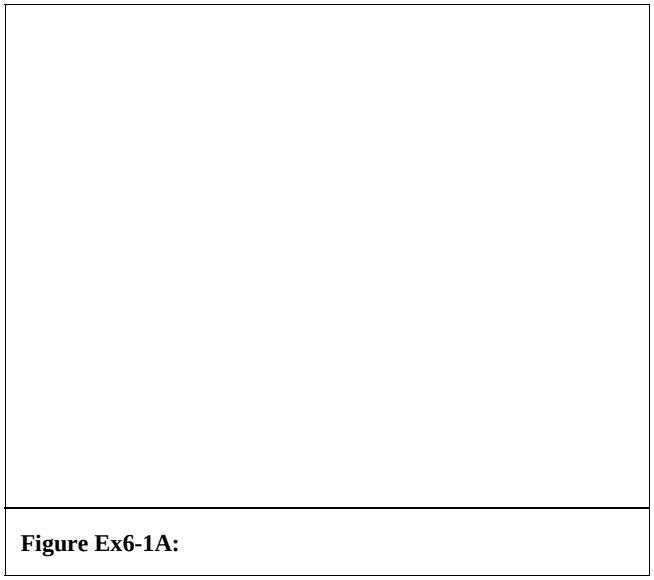


Figure Ex6-1A:

$$v_{gs2} = i_{in1} R_{D1}$$
$$v_{out} = -g_{m2} v_{gs2} R_{D2}$$
$$R_m = v_{out}/i_{in1} = -g_{m2} R_{D1} R_{D2}$$

If we use a CD amplifier instead of a CS for the second stage, we expect a lower output resistance at the expense of lower transresistance. The small-signal two-port model for a CG-CD amplifier is shown in Fig. Ex6-1B, assuming for simplicity that the CD stage has unity voltage gain.

The input resistance is the same for both amplifiers since both use a CG stage as the input. The output resistance is $R_{out} = 1/(g_{m2} + g_{mb2})$. The transresistance $R_m = v_{out}/i_{in1} = R_{D1}$ for the CG-CD configuration. Note that the output resistance and transresistance of the CG-CD configuration are lower than the CG-CS configuration by $g_{m2} r_{o2}$. The proper topological choice depends on the specifications required and the relative value of R_L compared to R_{out}.

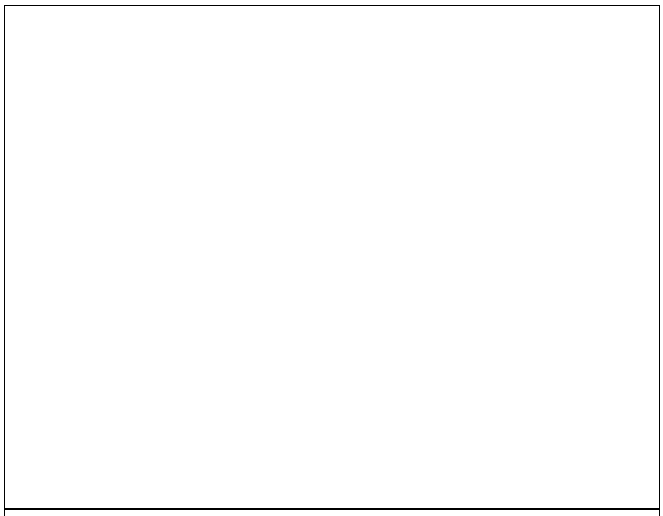


Figure Ex6-1B:

6-2 High-Frequency Analysis

As we have already seen in our treatment of single-stage amplifiers, analyzing the frequency response of an amplifier by hand (as opposed to computer simulation) usually necessitates approximations. The approximations are needed not only to manage complexity, but also to gain intuition about the limiting components in the circuit. Clearly, as we cascade stages, this need for simplifying approximations becomes only stronger.

In this section, we will examine typical strategies for the analysis of multistage amplifiers using two multi-stage amplifier examples: CG-CD and CS-CG (cascode amplifier). Our analysis of these circuits will heavily rely on the toolkit developed in Chapters 3and 4 and will invoke concepts such as the method of open-circuit time constants and the Miller approximation.

6-2-1 OCT-Based Analysis of the CG-CD Cascade

We begin our discussion by considering the CG-CD transresistance amplifier shown in Figure 6-6(a), along with its small-signal model in Figure 6-6(b). The latter circuit is constructed by cascading the respective small-signal models from Chapter 4 (Figure 4-13 and Figure 4-20) and contains no approximations.

At least in principle, the exact frequency response of the model in Figure 6-6(b) can be found by writing KCL at the three nodes of the circuit, and solving a 3x3 system of equations to find $V_{out}(s)/I_s(s)$. However, this procedure will not only be tedious, but will also produce long equations that are hard to

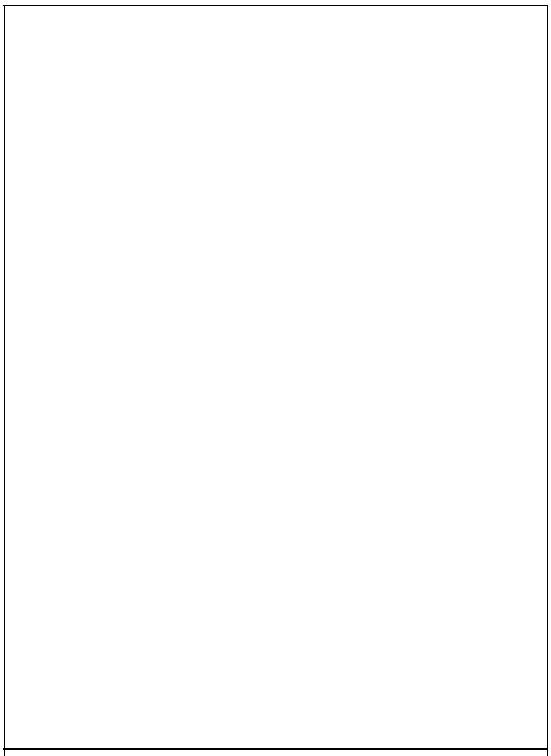


Figure 6-6: (a) CG-CD transresistance amplifier. (b) Corresponding small-signal model.

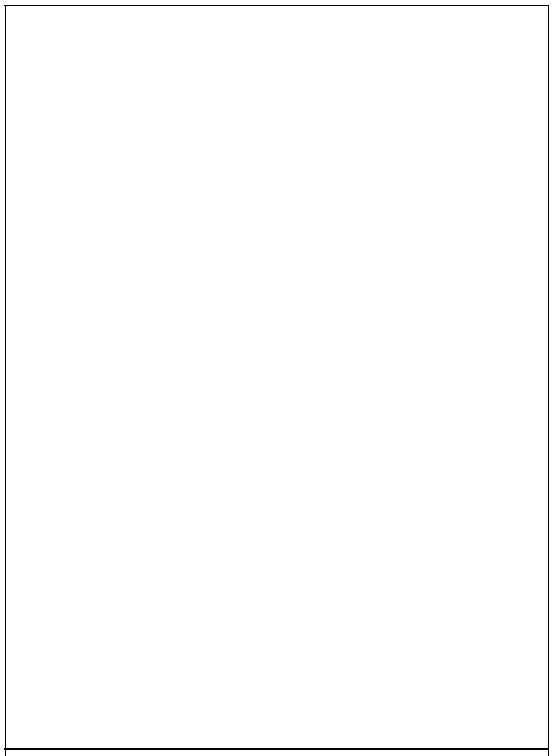


Figure 6-7: Simplified small-signal model of the CG-CD transresistance amplifier.

interpret for design. Generally, accurate symbolic or numerical analysis of a large circuit is best left to a computer and useful mainly to check our understanding and provide fine-tuned numerical answers.

As circuit designers, we must always look for ways to simplify the circuit and consider only the main effects that set the performance metrics of interest. A first step to take in this direction is to simplify the small signal model based on our understanding of the individual stages. In Figure 6-7, the following simplifications have been made:

- The resistance r_{o1} is omitted per the argument from Section 4-3-3. As long as $R_{D1} << r_{o1}$, a unilateral model for the CG stage without r_{o1} is sufficiently accurate.
- The resistance r_{o2} is omitted since typically $r_{o2} >> 1/g'_{m2}$.
- The capacitance C_{sb2} is omitted as discussed in Section 4-4-3. Due to the low output resistance of the CD stage, this capacitance will affect the behavior of the amplifier only at very high frequencies beyond our interest in a hand analysis.

With these simplifications in place, the circuit has become much more manageable for hand analysis, but has the remaining issue that the second stage is bilateral due to C_{gs2}. Again, while it is possible to analyze the circuit using KCL equations, the designer will typically look for further simplifications. These simplifications will now depend on the exact objective of the analysis. As a first example, suppose that we are only interested in obtaining a first-order estimate of the circuit's bandwidth. In this case, applying the method of open-circuit time (OCT) constants is a suitable direction to take. As we have already seen in Chapter 4, the main advantage of an OCT-based analysis is that it breaks the overall task into manageable steps and simultaneously provides insight about bandwidth bottlenecks. We will reiterate this point in the following example.

Example 6-2: OCT Analysis of a CG-CD Transresistance Amplifier

Consider the CG-CD small-signal model of Figure 6-7 with the following parameters: $g_{m1} = g_{m2} = 1$ mS, $g_{mb}/g_m = 0.2$, $r_s = 50$ kΩ, $R_{D1} = 10$ kΩ, $R_L = 3$ kΩ, $C_{s1} = 1$ pF, $C_x = 200$ fF and $C_{gs2} = 200$ fF. Estimate the circuit's bandwidth using the method of open-circuit time constants.

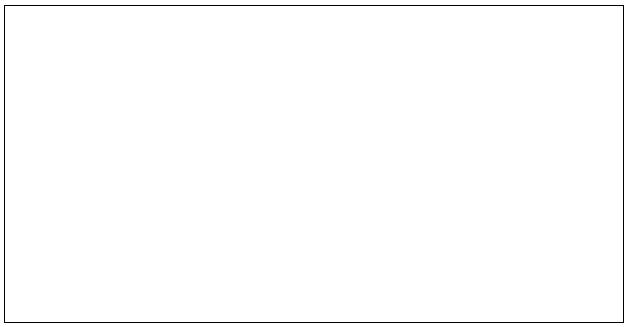
SOLUTION

First note that the two open-circuit time constants associated with C_{s1} and C_{d1} can be identified by inspection

$$\tau_{s1o} = C_{s1} \cdot \left(r_s || \frac{1}{g'_{m1}} \right) = 0.82\text{ns}$$

$$\tau_{dx} = C_x \cdot R_{D1} = 2\text{ns}$$

Now, in order to find the Thévenin resistance associated with C_{gs2}, we must consider a larger portion of the circuit, as shown below. Note that this setup resembles almost exactly the circuit of Figure 3-19, where we determined the Thévenin resistance for C_{gd} in a CS stage. The only difference here is that the sign of the controlled source is flipped, simply because the CD stage is non-inverting.



We can therefore directly apply the result of Eq. (3.66): "R_{left} + R_{right} + $g_m R_{left} R_{right}$," but now with g_m replaced by $-g_m$.

$$\begin{aligned} R_{Tgs2} &= R_{D1} + \left(R_L || \frac{1}{g'_{m2}} \right) - g_{m2} R_{D1} \left(R_L || \frac{1}{g'_{m2}} \right) \\ &= R_{D1} \left(1 - g_{m2} \left(R_L || \frac{1}{g'_{m2}} \right) \right) + \left(R_L || \frac{1}{g'_{m2}} \right) \\ &= R_{D1}(1 - A'_{v20}) + \left(R_L || \frac{1}{g'_{m2}} \right) \end{aligned}$$

where $A'_{v20} = v_{out}/v_x$ is the loaded voltage gain of the CD stage at low frequencies. Note that the first term of the above expression accounts for the Miller gain across C_{gs2}; we have already seen this term in our analysis of Section 4-4-3. Evaluating this result numerically we find $A'_{v20} = 0.652$ and $R_{Tgs2} = 4.13$ kΩ, and thus

$$\tau_{gs2} = C_{gs2} \cdot R_{Tgs2} = 0.826\text{ns}$$

$$f_{3dB} \cong \frac{1}{2\pi} \cdot \frac{1}{0.82\text{ns} + 2\text{ns} + 0.26\text{ns} + 0.826\text{ns}} = 43.7\text{MHz}$$

A simulation of the full circuit reveals $f_{3dB} = 61$ MHz. The OCT result is therefore off by about –26%, which is consistent with our understanding from Chapter 3. OCT bandwidth estimates are always conservative and tend be off by 20-30% when there are several significant time constants.

6-2-2 Pole Calculations for the CG-CD Cascade*

As indicated earlier, the OCT analysis above is useful as long as we are only interested in a first-order estimate of the circuit's bandwidth. However, in some situations we may require knowledge of the circuit's most significant poles and zeros. This would be the case, for example, if the amplifier was employed in a feedback system, where the exact location of the poles and zeros (rather than just the bandwidth estimate) play a significant role. While the analysis of feedback systems is beyond the scope of this module, it is worth introducing the reader to techniques suitable for pole-estimation in multi-stage circuits.

One typical approach that designers tend to follow for pole estimations is to try and approximate the circuit by a cascade of unilateral two-ports. As we have established in Section 3-4-4, for a cascade of unilateral two-ports with parallel RC sections, the circuit poles are directly set by the time constants at each port. Even though this strategy does not always give precise numerical results, it provides a great deal of intuition on how to influence the position of significant poles. We will now illustrate this approach using the CG-CD amplifier.

Example 6-3: Estimation of the Most Significant Poles in a CG-CD Transresistance Amplifier

Consider the CG-CD small-signal model of Figure 6-7 with the same parameters as in Example 6-2: $g_{m1} = g_{m2} = 1$ mS, $g_{mb}/g_m = 0.2$, $r_s = 50$ kΩ, $R_{D1} = 10$ kΩ, $R_L = 3$ kΩ, $C_{s1} = 1$ pF, $C_x = 200$ fF and $C_{gs2} = 200$ fF. Estimate the locations of the circuit's most significant poles.

SOLUTION

In order to estimate the most significant poles without resorting to complex algebra, we construct a unilateral model for the CD portion of the circuit (as already established in Figure 4-25). The corresponding simplified circuit (shown below) uses the Miller approximation to translate C_{gs2} into an equivalent parallel capacitance at the input port of the CD stage. This approximation is reasonable since in this example the voltage gain of the CD stage is constant up to very high frequencies, beyond the poles that we are trying to estimate.

From the simplified circuit, we can immediately identify the approximate pole locations by inspection and write the complete circuit transfer function as

$$p_1 = -\frac{1}{\left(r_s \| \frac{1}{g'_{m1}}\right) C_{s1}}$$

$$p_2 = -\frac{1}{R_{D1}[C_{d1} + (1 - A'_{v20})C_{gs2}]} \qquad (6.7)$$

$$\frac{v_{out}}{i_s} = R_{m0} \cdot \frac{1}{1 - \frac{s}{p_1}} \cdot \frac{1}{1 - \frac{s}{p_2}}$$

where $R_{m0} = R_{D1}A'_{v20}$ is the low-frequency transresistance. Evaluating the pole frequencies numerically, we find f_{p1} = 194 MHz and f_{p2} = 59 MHz. An accurate analysis of the full circuit reveals that there are three (real, LHP) poles and one (LHP) zero at the following respective frequencies: 67 MHz, 194 MHz, 1100 MHz and 796 MHz (for the zero).

From the above example, we conclude that the simplified analysis has done a reasonable job at estimating the first two poles. This overall approach is very powerful, because it lets the designer quickly see which components set the most significant pole frequencies. For example, if we wanted to increase f_{p2}, the result shows that reducing R_{D1} would be an option to consider. Such guidance would have been harder to obtain from lengthy algebraic expressions that capture the complete circuit transfer function.

6-2-3 OCT-Based Analysis of the CS-CG Cascade (Cascode Amplifier)

In this section, we explore the frequency response of another important multistage amplifier called the **cascode amplifier**. We have already briefly discussed this circuit in Section 4-5 and qualitatively argued that it should not suffer from the Miller effect, which often severely degrades the frequency response of the common-source amplifier. In the following discussion, we will analyze the frequency response of a cascode amplifier in the same spirit as we have done this in the previous section. That is, we will illustrate strategies that allow us to analyze and reason about the frequency response of the circuit intuitively, without significant much algebraic complexity.

For our analysis, we consider the circuit shown in Figure 6-8(a), along with its small-signal model in Figure 6-8(b). The shown model is complete, except that the output resistances of the two transistors have been neglected due to presence of the $1/g'_{m2}$ and R_D resistors (we assume $1/g'_{m2} << r_{o1}$ and $R_D << r_{o2}$).

Figure 6-8: (a) Cascode amplifier. (b) Corresponding small-signal model.

We begin by calculating the small-signal voltage gain of this amplifier at low frequencies by open-circuiting the capacitors. This voltage gain is given by

$$A_{v0} = \frac{v_x}{v_s} \cdot \frac{v_{out}}{v_x} = -\frac{g_{m1}}{g'_{m2}} \cdot g'_{m2} R_D = -g_{m1} R_D \qquad (6.8)$$

Note that this is the same as the low-frequency voltage gain that is obtained from a common-source amplifier; the advantage of the cascode amplifier lies mainly in its wideband frequency response, as we will show next.

For a first-order estimate of the circuit's bandwidth, we can consider its open-circuit time constants. Once again, we find the time constants by inspection, and using the "$R_{left} + R_{right} + g_m R_{left} R_{right}$" to find the open-circuit time constant for C_{gd1}.

$$\tau_{gs1o} = R_S C_{gs} \tag{6.9}$$

$$\tau_{xo} = \frac{C_x}{g'_{m2}} \tag{6.10}$$

$$\begin{aligned}\tau_{gd1o} &= \left[R_S + \frac{1}{g'_{m2}} + g_{m1} R_S \frac{1}{g'_{m2}}\right] C_{gd1} \\ &= \left[\frac{1}{g'_{m2}} + \left(1 + \frac{g_{m1}}{g'_{m2}}\right) R_S\right] C_{gd1}\end{aligned} \tag{6.11}$$

Next, we consider the time constant at the output node

$$\tau_{d2o} = R_D C_{d2s} \tag{6.12}$$

Thus, the bandwidth estimate of the stage is

$$\omega_{3dB} = \frac{1}{\tau_{gs1o} + \tau_{xo} + \tau_{gd1o} + \tau_{d2o}} \tag{6.13}$$

One key difference compared to a common-source amplifier lies in the time constant associated with C_{gd1}. For the basic common-source amplifier in Chapter 3, we had

$$\tau_{gdo} = [R_D + (1 + g_m R_D) R_S] C_{gd} \tag{6.14}$$

which suffers from Miller amplification of the gate-drain capacitance by the factor $g_m R_D$, which is the magnitude of the circuit's voltage gain. In contrast, no significant Miller amplification occurs in the cascode amplifier. Even if the cascode amplifier is designed for large overall voltage gain, the magnitude of the voltage gain across C_{gd1} is limited to g_{m1}/g'_{m2}, which is typically close to unity. The following example looks at a numerical evaluation of this advantage.

Example 6-4: Cascode Amplifier Bandwidth Estimate Using an OCT Analysis

Consider the circuit shown in Figure Ex6-4 and assume the following component values: R_D =5 kΩ, R_s = 50 kΩ, and for both MOSFETS: g_m = 1 mS, C_{gs} = 40.7 fF, C_{gd} = 10 fF, $C_{db} = C_{sb}$ = 11.6 fF. These are the same parameter values we used in Example 3-7, an OCT analysis of a basic CS amplifier. Assume that the body of M_2 is connected to ground and that g_{mb2}/g_m = 0.2. Estimate the 3-dB bandwidth using an OCT analysis and compare the result to the bandwidth estimate of the original CS circuit from Example 3-7.

SOLUTION

Evaluating Eq. (6.9) - Eq. (6.13) numerically with the given values yields

Figure Ex6-4

$$\tau_{gs1o} = 50\text{k}\Omega \cdot 40.7\text{fF} = 2.035\text{ns}$$

$$\tau_{xo} = \frac{11.6\text{fF} + 40.7\text{fF} + 11.6\text{fF}}{1\text{mS}} = 64\text{ps}$$

$$\tau_{gd1o} = \left[\frac{1}{1.2\text{mS}} + \left(1 + \frac{1\text{mS}}{1.2\text{mS}}\right) 50k\Omega\right] 10\text{fF} = 925\text{ps}$$

$$\tau_{d2o} = 5k\Omega(10\text{fF} + 11.6\text{fF}) = 108\text{ps}$$

The bandwidth estimate is therefore

$$f_{3dB} = \frac{1}{2\pi} \cdot \frac{1}{2.035\text{ns} + 64\text{ps} + 925\text{ps} + 108\text{ps}} = 50.8\text{MHz}$$

For comparison, the three time constants of the corresponding common-source amplifier (without M_2) would be (see Example 3-7, with C_L = 0)

$$\tau_{gso} = 50\text{k}\Omega \cdot 40.7\text{fF} = 2.035\text{ns}$$

$$\tau_{db} = 5\text{k}\Omega \cdot 11.6\text{fF} = 58\text{ps}$$

$$\tau_{gdo} = [5k\Omega + (1 + 1\text{mS} \cdot 5k\Omega) 50k\Omega] 10\text{fF} = 3.05\text{ns}$$

and thus

$$f_{3dB,\,CS} = \frac{1}{2\pi} \cdot \frac{1}{2.035\text{ns} + 58\text{ps} + 3.05\text{ns}} = 30.9\text{MHz}$$

The cascode amplifier's bandwidth is about 62% larger than that of the basic common-source voltage amplifier, which is a significant improvement.

6-2-4 Pole Calculations for the CS-CG Cascade (Cascode Amplifier)*

While the bandwidth increase seen in Example 6-4 is a welcome feature of the cascode amplifier, it comes with the issue that an additional pole is introduced. This is illustrated in Figure 6-9, which shows a computer simulation of the frequency response for both circuits considered in Example 6-4. The cascode amplifier has a larger 3-dB corner frequency, but exhibits an extra pole that bends the response more sharply at high frequencies. In many cases this behavior not troublesome. However, it turns out that these non-dominant poles must usually be considered in feedback amplifiers and often limit performance. Consequently, there is a general need for the circuit designer to be able to calculate the location of the high-frequency poles in a cascode amplifier.

Figure 6-9: Comparison of the magnitude response for the cascode amplifier and a common-source amplifier (assuming the component values from Example 6-4).

There are several ways by which one can estimate the location of the poles for the cascode amplifier. The most appropriate method is to invoke the expressions we have already derived in Chapter 3 for the dominant and non-dominant pole of a CS amplifier. Specifically, we note that the left portion of the model in Figure 6-8(b) is equivalent to the CS amplifier analyzed in Chapter 3 (Figure 3-13) with the following variable changes: $C_{db} \rightarrow C_x$ and $R_{out} \rightarrow 1/g'_{m2}$. With these substitutions, the dominant pole frequency is most concisely written using the Miller approximation result from Eq. (3.56), which becomes

$$\omega_{p1} = \frac{1}{R_s\left[C_{gs1} + \left(1 + \frac{g_{m1}}{g'_{m2}}\right)C_{gd1}\right]} \tag{6.15}$$

From this expression, we immediately see that the dominant pole does not suffer from significant Miller multiplication. This is analogues to the improvement we have seen in the zero value time constant of C_{gd1}.

An expression for the second pole frequency was derived in Eq. (3.56), which in the present context becomes

$$\omega_{p2} \cong \frac{g'_{m2}}{C_x} + \frac{1}{R_s C_{gs1}} + \frac{g_{m1}}{C_x} \cdot \frac{C_{gd1}}{C_{gs1}} \tag{6.16}$$

For the particular example that we consider here, we know that $R_s C_{gs1} >> C_x/g'_{m2}$. Also, since g_{m1} and g'_{m2} are typically comparable and $C_{gd1} << C_{gs1}$, we can drop all but the first term in Eq. (6.16) to obtain

$$\omega_{p2} \cong \frac{g'_{m2}}{C_x} \tag{6.17}$$

Since C_x is usually dominated by the gate capacitance of M_2 (C_{gs2}), ω_{p2} is typically close to the transistor's transit frequency (g_{m2}/C_{gg2}). For a minimum length n-channel device in our technology, this frequency is on the order of 1 GHz for typical biasing conditions (see Chapter 3).

Finally, we can easily identify the third pole frequency of the cascode amplifier directly from Figure 6-8(b). The RC network at the output branch is unilaterally coupled to the CS amplifier portion and therefore contributes a pole corresponding the branch's parallel RC time constant.

$$\omega_{p3} \cong \frac{1}{R_D C_{d2}} \tag{6.18}$$

From these pole frequency expressions, we see that the cascode amplifier can be conveniently modeled as shown in Figure 6-10. We have thus once again arrived a relatively simple and intuitive unilateral model that captures most of the relevant circuit behavior, and specifically the circuit's pole locations. It can be shown that this model is still valid for scenarios where the dominant pole is not at the input, but instead associated with the output of the circuit.

Example 6-5: Pole Calculations for a Cascode Amplifier

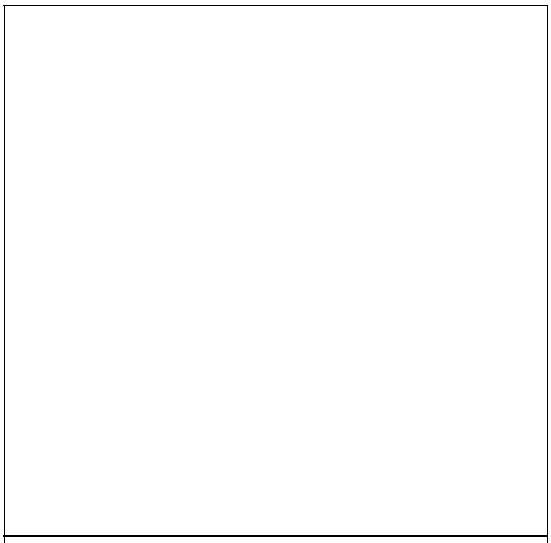


Figure 6-10: Unilateral model of the cascode amplifier.

Compute the pole frequencies for the cascode amplifier considered in Example 6-4 (Figure Ex6-4). The parameters are: R_D =5 kΩ, R_s = 50 kΩ, and for both MOSFETS: g_m = 1 mS, C_{gs} = 40.7 fF, C_{gd} = 10 fF, C_{db} = C_{sb} = 11.6 fF and g_{mb2}/g_m = 0.2.

SOLUTION

Evaluating Eq. (6.15), Eq. (6.17), and Eq. (6.18) numerically gives

$$f_{p1} = \frac{1}{2\pi}\frac{1}{50k\Omega\left[40.7\text{fF} + \left(1 + \frac{1\text{mS}}{1.2\text{mS}}\right)10\text{fF}\right]} = 53.9\text{MHz}$$

$$f_{p2} \cong \frac{1}{2\pi}\frac{1.2\text{mS}}{40.7\text{fF}+11.7\text{fF}+11.7\text{fF}} = 3.0\text{GHz}$$

$$f_{p3} \cong \frac{1}{2\pi}\frac{1}{5\text{k}\Omega(10\text{fF}+11.7\text{fF})} = 1.47\text{GHz}$$

Note that these pole frequencies correspond to the magnitude response that was shown in Figure 6-9. There is one dominant pole, and two non-dominant poles beyond 1 GHz.

6-3 Design of a Three-Stage Transresistance Amplifier

In this section we will explore the design and optimization of a three-stage transresistance amplifier for use in a fiber optic receiver. The objective is to illustrate the process that a designer faced with this problem would have to go through. Furthermore, we present a systematic approach for the sizing of the amplifier components to achieve near-optimum performance under a given set of constraints.

6-3-1 Problem Definition

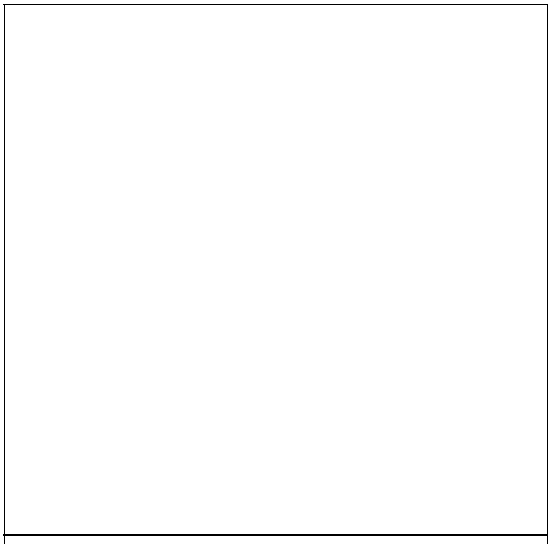


Figure 6-11: High-level schematic of the design problem.

As illustrated in Figure 6-11, we wish to amplify the signal delivered from a photodiode over a wide bandwidth and drive this amplified signal into a capacitive load. The overall specifications for the problem are summarized in Table 6-1.

Table 6-1: Specifications for a fiber-optic amplifier.

Parameter	Symbol	Value
Low Frequency Transresistance	R_{m0}	2 kΩ
Load Capacitance	C_L	10 pF
Diode Capacitance	C_D	5 pF
Total Drain Current	I_{Dtot}	3 mA
Bandwidth	f_{3dB}	Maximize

In order to appreciate why an amplifier is needed in the first place to solve this problem, let us consider the "trivial" solutions shown in Figure 6-12. Interestingly, if all we did was to connect a resistance of value R_{m0} to the diode, we would already meet the low-frequency transresistance specification of the circuit without using any transistors ($v_{out}/i_d = R_{m0}$). However, the key issue then is that we would not be able to achieve

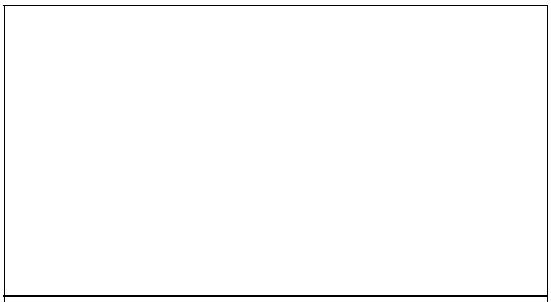


Figure 6-12: An attempt to solve the design problem using a single resistor.

high bandwidth. In this solution, the bandwidth is only $1/[2\pi R_{m0}(C_L + C_D)] = 5.3$ MHz; we will be able to do better with an active circuit.

6-3-2 Circuit Architecture Considerations

In order to achieve high bandwidth, we must ensure that the large capacitors at the input and output of the circuit "see" low resistances, such that their presence does not create large time constants. Based on we have learned in this module, one possible solution is to employ a CG stage at the input and a CD stage at the output of the circuit. This option is shown in Figure 6-13, where we have qualitatively annotated the signal flow and relevant gain terms and port resistances in the circuit. A representation of this style is sometimes used by experienced designers who are already familiar with the small-signal model of each stage, and usually won't bother to draw it out. We use this representation here and in the following figures to prepare the reader toward this transition.

Here, C_D is presented with the low input resistance of the CG stage and C_L sees the low output resistance of the CD buffer. In terms of signal flow, the CG stage acts as a current buffer and passes i_d essentially unchanged to R_{D1}. The resistor R_{D1} performs a current-to-voltage conversion that corresponds to the desired transresistance, while the CD stage buffers the generated voltage to handle the large load capacitance C_L.

This proposed circuit would in principle work but contains one significant challenge. Essentially, all of the transresistance is due to R_{D1}, which must therefore be set to approximately 2 kΩ. Together with the input capacitance of the CD stage, this resistor creates a time constant that may dominate the circuit and may again not allow us to achieve large bandwidth. A typical remedy to this problem is to do a better job at distributing the gain among several stages of the amplifier. One such option is shown in Figure 6-14. Here, we employed a basic CS voltage amplifier between the CG and CD stages. Since the CS stage will have voltage gain, the resistance R_{D1} can be reduced by this gain factor to achieve the same overall transresistance.

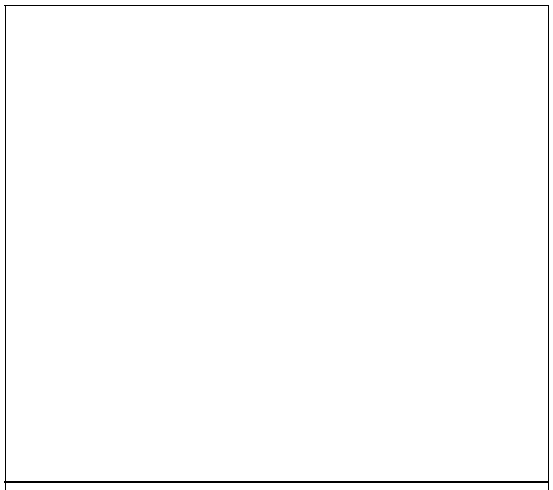


Figure 6-13: First-cut circuit architecture.

This will reduce the time constants that are proportional to R_{D1} and therefore help maximize the bandwidth.

An additional improvement that could be considered in this circuit is to use a cascode amplifier rather than simple CS stage to implement the gain stage (see Reference 1). This would help to reduce the capacitance seen by R_{D1} and therefore speed up the circuit. For simplicity, we will not pursue this idea and instead work with the circuit of Figure 6-14 toward a final solution.The reader is invited to explore using a cascode for improved performance.

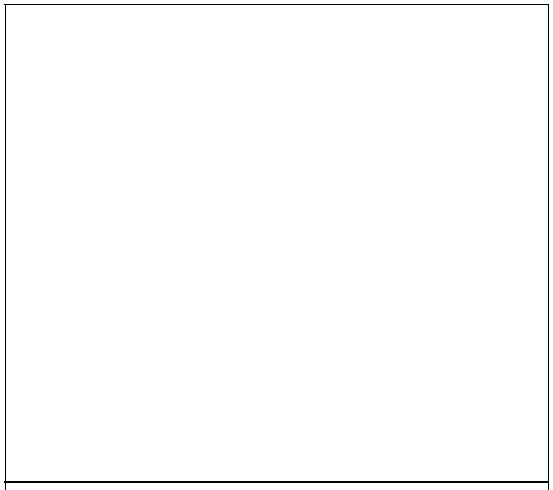


Figure 6-14: Improved circuit architecture.

6-3-3 Biasing Considerations

As we have seen in Chapters 2 and 5, biasing a common-source

stage properly can be a difficult task. One issue in the prototype circuit of Figure 6-14 is that any parameter variation in I_{B1} or R_{D1} will directly impact the quiescent point voltage at the gate of the CS stage. Since this stage has voltage gain, such variations will then show up amplified at its output and potentially drive parts of the circuit out of saturation and/or limit the available signal swing.

To overcome this issue, we employ the replica biasing approach introduced in Chapter 5 (see Figure 6-15). Here, the bias point at the gate of M_2 is set via the replica device M_2. This diode connected transistor is biased with the same current as M_2, and thereby "computes" the correct bias voltage that will also track threshold voltage process variations. Also, this voltage is to first order independent of variations in I_{B1} and R_{D1}. As long as I_{B1a} and I_{B1b} match, these currents can vary in their absolute value without disturbing the bias point of M_2. Similarly, the bias point of M_2 is not disturbed by changes in the value of R_x, since this resistor (to first order) carries no current in the quiescent point.

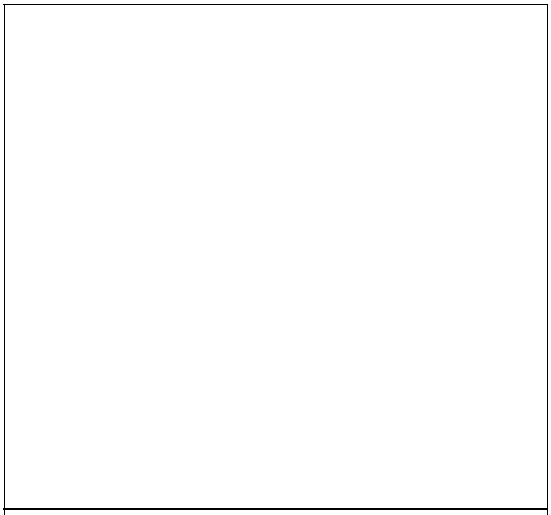


Figure 6-15: Amplifier circuit with proper biasing.

At the drain side of M_2, we employ a resistive divider to set the bias point gate potential for M_3. In this arrangement, the division ratio can be adjusted for the proper input bias voltage for the CD stage (e.g., $V_{DD}/2$), and the absolute resistor value is determined by the desired gain of the CS stage (the $g_{m2}R_{D2}$ product).

The current sources required in Figure 6-15 can be realized using the current mirror circuits discussed in Chapter 5, and supplied for example by a globally shared constant-g_m reference current source. A variety of options exists for the generation of the CG bias voltage V_{B1}. This voltage can be set up by one of the two methods shown in Figure 5-25. Without further working through the details required to complete the bias circuit, we will now investigate a proper procedure for sizing the signal path devices.

6-3-4 Examination of Tradeoffs

To complete our design, we must determine all bias currents, transistor geometries and resistor values. As we shall see, this is a non-trivial task, especially if we want to achieve optimum performance, as for instance maximum bandwidth in the given problem. Even though the circuit has only three transistors, there are several degrees of freedom among the different design variables.

A first step in the right direction is to begin and identify the relationships that govern the performance of the circuit and analyze these for the key tradeoffs. Furthermore, in this process, we must make reasonable approximations to keep the algebraic complexity low. Let us begin by writing an expression for the circuit's low-frequency transresistance.

$$R_{m0} = R_{D1}A_{v20}A_{v30} \cong -R_{D1} \cdot g_{m2}R_{D2} \cdot \frac{g_{m3}}{g'_{m3}} \qquad (6.19)$$

Since the last term in Eq. (6.19) is close to unity, it is clear that R_{m0} is primarily set by the product of R_{D1} and the CS stage voltage gain. Now, since we are only interested in the circuit's bandwidth, and not the exact pole/zero locations, using an OCT estimate for further analysis is most appropriate and convenient. The schematic in Figure 6-16 includes all of the relevant capacitances that give rise to time constants. Several of the indicated capacitances are parallel combinations and for simplicity it is important to immediately discard small capacitances that may not impact the design significantly. For instance, we neglect C_{sb1} relative to the large diode capacitance at the input node. In addition, we neglect several other extrinsic capacitances that should limit the bandwidth significantly. Once the design is completed, all of these assumptions can be checked, for instance through a computer simulation of the full circuit. By inspection, we identify the following six open-circuit time constants.

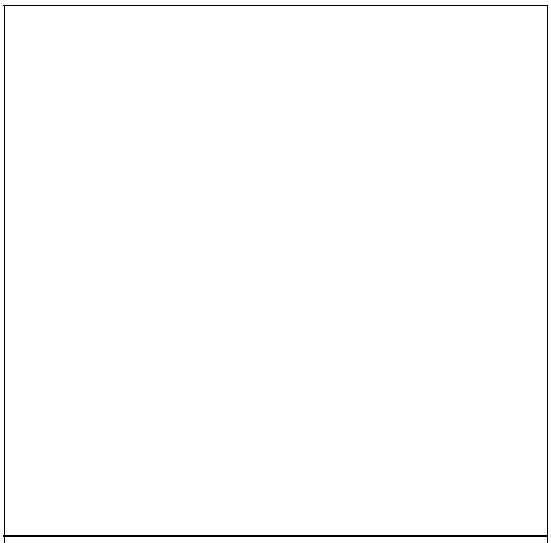


Figure 6-16: Amplifier circuit with relevant capacitances for OCT calculations.

$$\begin{aligned}
\tau_{i1o} &= \frac{C_D + C_{gs1}}{g'_{m1}} \\
\tau_{i2o} &= R_{D1}C_{gs2} \\
\tau_{gd2o} &= (R_{D1} + R_{D2} + g_{m2}R_{D1}R_{D2})C_{gd2} \\
\tau_{i3o} &= R_{D2}C_{gd3} \\
\tau_{gs3o} &= \left(R_{D2} + \frac{1}{g'_{m3}} - \frac{g_{m3}}{g'_{m3}}R_{D2}\right)C_{gs3} \\
\tau_{outo} &= \frac{C_L}{g'_{m3}}
\end{aligned} \tag{6.20}$$

where we have again made use of the "$R_{left} + R_{right} \pm g_m R_{left} R_{right}$" rule to find τ_{gd2o} and τ_{gs3o}. Now, since R_{D1} and R_{D2} are among the key parameters that set the overall transresistance, it makes sense to collect the respective proportional terms terms Eq. (6.20).

$$\begin{aligned}
\tau_{core} = {} & R_{D1}(C_{gs2} + [1 + g_{m2}R_{D2}]C_{gd2}) \\
& + R_{D2}\left(C_{gd3} + \left[1 - \frac{g_{m3}}{g'_{m3}}\right]C_{gs3}\right)
\end{aligned} \tag{6.21}$$

We call this time constant τ_{core}, since it is associated with the resistors in the core of the overall amplifier. To gain further insight into the tradeoffs dictated by this expression, we can re-write as

$$\begin{aligned}
\tau_{core} = {} & \frac{R_m}{|A_{v20}|A_{v30}}(C_{gs2} + [1 + |A_{v20}|]C_{gd2}) \\
& + \frac{|A_{v20}|}{g_{m2}}(C_{gd3} + [1 - A_{v30}]C_{gs3})
\end{aligned} \tag{6.22}$$

From this formula, we see that one part of the time constant increases with A_{v20}, while the other component decreases. This suggests that choosing the right amount of voltage gain may be key to maximizing the bandwidth.

As far as the remaining terms of Eq. (6.20) are concerned, it makes sense to group these together in a similar fashion; that is, collect terms for the input and output network, respectively.

$$\tau_{in} = \frac{C_D + C_{gs1}}{g'_{m1}} \tag{6.23}$$

$$\tau_{out} = \frac{C_L + C_{gs3}}{g'_{m3}} \tag{6.24}$$

At first glance, these expressions do not provide any interesting opportunity for tradeoffs. Given our finite budget for drain current, the amount of g_m we can generate will be limited, and this will essentially set τ_{in} and τ_{out}. The main degree of freedom here is what fraction of the available current we are going to use in the input and output branches.

6-3-5 Optimization Procedure

Given the above observations, the key question that remains is how we should distribute the available current among the three amplifier stages. There are several ways in which we can perform this optimization. One would be to immediately engage in circuit simulations and iterate over current distributions and device geometries until we have identified a satisfactory answer. This approach, however, is not only time consuming, but also gives little insight about the quality and robustness of the obtained design point. The other extreme would be to try and formulate a closed form expression for the optimal sizing. Unfortunately, for all but the most trivial circuits this turns out to be impossible or infeasible.

An approach that lies about midway between these extremes is to set up an insight-based, hand crafted calculation script that allows the designer to quickly sweep through the design space by specifying a few reasonable assumptions and exercising the key "knobs" that control the tradeoffs. For example, as we have explained above, one such knob in our design is the voltage gain of the CS stage.

Figure 6-17 shows one possible way of setting up a calculation script for our design problem. The script contains five distinct sections

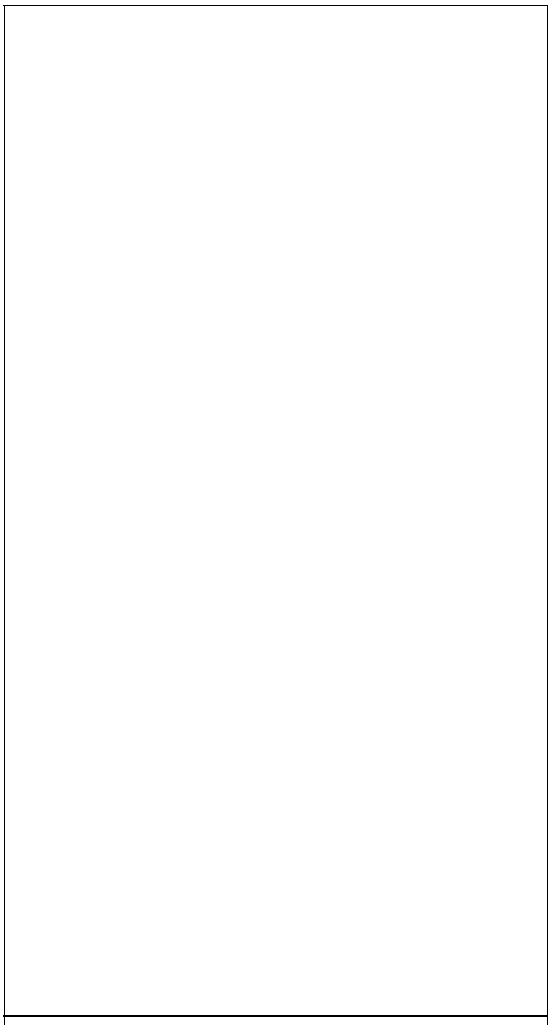


Figure 6-17: Design script for circuit optimization.

- The definition of process technology parameters such as $\mu_n C_{ox}$, etc. Here, we can also define first-order estimates that may be needed in the calculations, for instance a typical value for g_{mb}/g_m.
- The design specifications (desired R_{m0}, C_L, etc.).
- A set of design choices that represent reasonable guess values for some of the unknowns that are not expected to play a significant role in the optimization or that we want to fix to specific values. For instance, in the shown script we set all channel lengths to minimum length (for high speed) and set all gate overdrive voltages to a typical value of 0.3 V. The latter choice could be imposed for instance by linearity or signal swing requirements. These parameters can of course be changed in the design process, but are not viewed as the main degrees of freedom.
- The main variables, which are the voltage gain of the CS stage (A_{v20}) and the allocation of current for the input and output branches (I_{D1}/I_{Dtot} and I_{D3}/I_{Dtot}). These are the primary knobs that we wish to adjust to search for optima.
- The performance and sizing calculations based on the given design choices and values set for the main variables. The output of this part is essentially the objective function of the optimization, which is bandwidth (or the total sum of the time constants) in our example.

Once such a script has been generated, it is straightforward to explore the design space and adjust the main variables to find a suitable design point. This can be done manually, or automatically using *for* loops. Figure 6-18 shows a plot that was generated by sweeping the CS stage's gain while leaving the current allocations for input branches at the values indicated in Figure 6-17. This plot shows the existence of an optimum that we had already suspected from Eq. (6.22): one part of the amplifier core's time constant increases with A_{v20}, while the other component decreases. Given our process parameters, specifications and design choices, the best value for the CS voltage gain is in the range of 10-20. Similar sweeps can be performed with the other variables to arrive at a final design point that is acceptable.

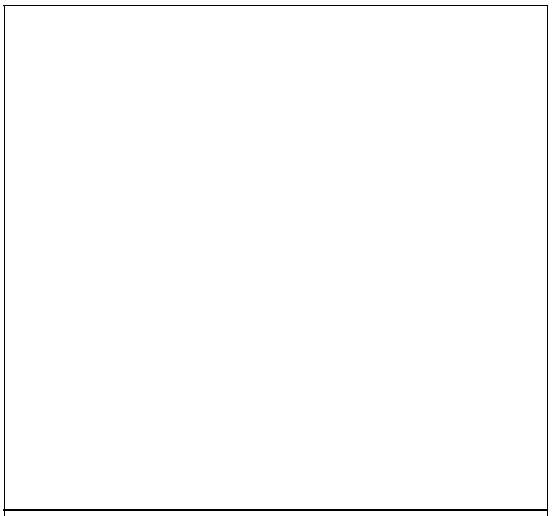


Figure 6-18: Time constants of the amplifier as a function of CS stage voltage gain (A_{v20}).

Interestingly, the tradeoff curve in Figure 6-18 exhibits a rather shallow optimum. This is a quite typical and welcome feature in circuit design problems. Anything but a shallow optimum would make us wonder whether the chosen point would be robust in presence of PVT variations and mismatch. A clear advantage of working with a design script (as opposed to repet-

itive simulation) is that we can visualize the region surrounding the chose design point.

6-3-6 Performance Verification

Even though the circuit simulator is not the best tool for optimization, it is the ultimate tool for verifying the circuit's performance and to track down discrepancies due to simplifications in made in the design script. Once we have determined all parameters in our design script, we can use the computed currents and transistor geometries as input for a circuit simulation that is based on the complete circuit and full transistor models. As an example, we consider here the design point summarized in Table 6-2.

Table 6-2: Chosen design point for the transresistance amplifier (from design script).

Parameter	Symbol	Value
Fractional input branch current	I_{D1}/I_{Dtot}	0.25
Fractional output branch current	I_{D3}/I_{Dtot}	0.25
CS Stage voltage gain	$\|A_{v20}\|$	10
Width of M_1	W_1	333 µm
Width of M_2	W_2	666 µm
Width of M_3	W_3	333 µm
Drain resistance of CG stage	R_{D1}	250 Ω
Drain resistance of CS stage	R_{D2}	1 kΩ
Total time constant	τ_{tot}	4.16 ns
Bandwidth estimate	f_{3dB}	38.2 MHz

A computer simulation of the full circuit with the above geometries, currents and resistor values shows R_{m0} = 2.13 kΩ and f_{3dB} = 70 MHz. The corresponding errors in the design script values are –6.5% and –45%, respectively. The majority of the error in the f_{3dB} estimate is systematic and due to conservative nature of OCT bandwidth estimates (see Chapter 3). The remaining percent differences can be tracked down to a slightly higher than desired I_{D2} in the actual circuit due to channel length modulation (which can be significant at minimum length). This and other discrepancies can usually be explained and either resolved or properly incorporated in the design script and do not negatively interfere with the proposed optimization flow. Quite contrary, comparing and resolving discrepancies between the script and simulation improve insight and lead to a form of "double book keeping" in the design flow that helps prevent erroneous design outcomes.

Finally, it is important to note that in addition to the discussed verification of small-signal performance, the circuit must always be simulated under large signal (transient) conditions for the ultimate performance check.

6-3-7 Considerations for Advanced Technologies

The above-discussed example provided a framework for circuit design with square-law equations. A significant, but not insurmountable hurdle that must be overcome when applying this approach with modern CMOS technologies lies in the growing complexity of transistor models. The latest generation of short-channel MOSFET models is based on hundreds of modeling parameters, a complexity that is manageable by a computer, but not practical for hand calculations and direct scripting. A solution to this problem that fits seamlessly into the proposed flow is to replace the square-law equations with computer generated look-up tables that relate the transistor parameters of interest numerically. An example is a look-up table that relates the gate-overdrive voltage to the transconductance per unit width of a MOSFET. Suitable ways to parameterize and use such tables are discussed in advanced literature on this subject, see for example References 2 and 3.

Summary

In this chapter, we discussed the analysis and design of multistage amplifiers. The small-signal two-port models were used to investigate the use of various types of amplifiers to transform the input or output resistance as well as increase the gain (voltage, current, transconductance, transresistance). Next, we analyzed two important examples of two-stage amplifiers in terms of their frequency response. We ended the chapter with a design of a transresistance amplifier to apply the concepts studied in this module. Specifically we showed:

- How to use the two-port models to find a quick approximation of the overall amplifier performance at low frequencies.
- How to approach the high-frequency analysis of multistage amplifiers using OCT bandwidth estimates and unilateral two-port models for the quick extraction of significant pole frequencies.
- How to design a larger circuit systematically, with the help of insight-based design scripts.

References

1. Y. Shim et al., "Design of Full Band UWB Common-Gate LNA," *IEEE Microwave and Wireless Components Letters*, vol. 17, no. 10, pp. 721-723, Oct. 2007.

2. P. Jespers, *The g_m/I_D Methodology, a sizing tool for low-voltage analog CMOS Circuits*, Springer, 2010.

3. T. Konishi, K. Inazu, J.G. Lee, M. Natsui, S. Masui, and B. Murmann, "Design Optimization of High-Speed and Low-Power Operational Transconductance Amplifier Using g_m/I_D Lookup Table Methodology," *IEICE Trans. Electronics*, vol. E94-C, no. 3, pp. 334-345, March 2011.

Problems

Unless otherwise stated, use the standard model parameters specified in Table 4-1 for the problems given below. Consider only first-order MOSFET behavior and include channel-length modulation (as well as any other second-order effects) only where explicitly stated.

P6.1 This problem compares a CS-CS voltage amplifier with a CS-CD voltage amplifier. If you are given that the g_m of the MOS devices is 1 mS and r_o is 100 kΩ, which topology yields the highest overall voltage gain, given

(a) $R_S = 1$ kΩ and $R_L = 100$ Ω

(b) $R_S = 1$ kΩ and $R_L = 10$ kΩ

(c) Repeat (a) and (b) when $g_m = 100$ μS and $r_o = 10$ MΩ

P6.2 This problem compares a CS-CS transconductance amplifier with a CS-CG transconductance amplifier. If you are given that the g_m of the MOS devices is 1 mS and r_o is 100 kΩ, which topology gives the highest overall transconductance, given

(a) $R_S = 1$ kΩ and $R_L = 100$ Ω

(b) $R_S = 1$ kΩ and $R_L = 10$ kΩ

(c) Repeat (a) and (b) when $g_m = 100$ μS and $r_o = 10$ MΩ

P6.3 A voltage buffer is shown in Figure P6-3. We have assumed that the circuit is fabricated in an n-well CMOS process where we can short the backgate and source of the p-channel device. Parameters: $I_{B1} = I_{B2} = 200$ μA, $(W/L)_1 = (W/L)_2 = 50$. Neglect channel-length modulation.

(a) What is V_{OUT} given $V_{IN} = 0$ V?

(b) Find the open-circuit voltage gain ($R_L \to \infty$).

(c) What is the minimum load resistor that the amplifier can drive and still maintain a voltage gain of 0.6?

P6.4 Repeat P6.3 given that the circuit is fabricated in a p-well CMOS process and that the n-channel device has its backgate shorted to the source and the p-channel device has its backgate tied to the positive power supply.

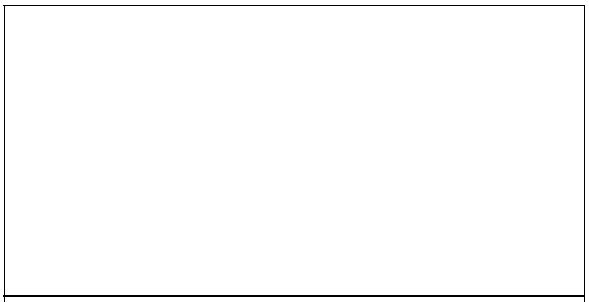


Figure P6-3.

P6.5 You are given the voltage amplifier shown in Figure P6-5 with $(W/L)_1 = 20$ and $(W/L)_3 = 50$. In this problem we will assume (for simplicity) that all backgates are shorted to their respective sources. Neglect channel-length modulation.

(a) Find the W/L for M_2, M_{2B}, and M_4 (same for all three transistors) so that each MOSFET has a drain current of 500 μA.

(b) What is the required voltage at the gate of M_3 so that the output level will be 0 V?

(c) Calculate V_{BIAS} so that M_1 sinks the current from M_2.

(d) Draw a two-port model of this CS-CD stage and calculate the parameters.

(e) Calculate the overall voltage gain if $R_S = 10$ kΩ and $R_L = 1$ kΩ.

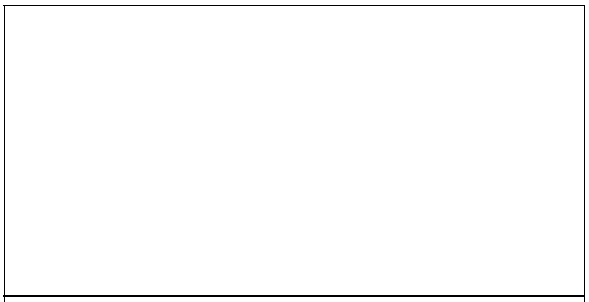


Figure P6-5.

P6.6 A cascode transconductance amplifier is shown in Figure P6-6. Neglect the backgate effect for this problem.

(a) Calculate $(W/L)_1$ of M_1 such that the small-signal transconductance, $i_{out}/v_s = 1$ mS. Assume $R_L = 0$ Ω (short-circuit output current) for this part.

(b) Calculate the value of V_{BIAS} using the $(W/L)_1$ calculated in part (a) such that $I_{OUT} = 0$ A.

(c) Calculate the output resistance of this transconductance amplifier.

(d) Calculate the overall transconductance at DC given that $R_S = 10$ kΩ and $R_L = 1$ kΩ.

(e) Estimate the bandwidth of this circuit given that $R_S = 10$ kΩ and $R_L = 1$ kΩ

Figure P6-6.

P6.7 In this problem we will investigate an interesting case in which the benefit of cascoding is dependent upon operating frequency. In the cascode amplifier shown in Figure P6-7, all devices are biased in the saturation region. You may neglect backgate effect, but you must include finite output resistance (r_o) of M_1. Neglect all device capacitances, and consider only the explicitly shown C_L. Simplify your analysis by assuming $g_m r_o >> 1$.

(a) Derive an analytical expression for the impedance $Z_x(s)$ looking into the source of M_1 in terms of small-signal device parameters and C_L.

(b) Sketch $|Z_x(j\omega)|$ versus frequency using logarithmic scales on both axes. Mark pertinent breakpoints symbolically, using the involved circuit parameters.

(c) Explain in a few words for which frequency range M_1 helps alleviate the Miller multiplication of C_{gd2}.

Figure P6-7.

P6.8 In the transresistance amplifier shown in Figure P6-8, all devices operate in the saturation region. Neglect channel-length modulation in all parts of this problem.

(a) After building this circuit in the lab, you measure $V_X = 1$ V at the operating point. What is the corresponding value of the bias current I_B? In this calculation, be sure to consider the backgate effect. Given parameters: $W_1/L_1 = 20$, $V_{B1} = 2.5$ V.

(b) Assuming that the transresistance $R_m = v_{out}/i_{in}$ is equal to 50 kΩ, compute the values of R_1 and R_2 that minimize τ_{tot}, the sum of all open-circuit time constants in the circuit. In this calculation, ignore all capacitances other than the explicitly drawn C_1 and C_2. Given parameters: $g_{m2} = 2$ mS, $C_1 = 1$ pF, $C_2 = 2$ pF.

Figure P6-8.

P6.9 In the amplifier circuit shown in Figure P6-9, ignore finite output resistance, extrinsic device capacitances and the backgate effect. Both transistors operate in the saturation region. Parameters: $g_m = 5$ mS, $f_T = 5$ GHz, $R = 1$ kΩ. Calculate:

(a) The amplifier's small signal gain v_{out}/v_{in} at low frequencies (ignore all capacitances). Be sure to include the appropriate sign.

(b) The amplifier's 3-dB bandwidth using the open-circuit time constant method.

Figure P6-9.

P6.10 Consider the cascode amplifier from Example 6-4. Assuming the same parameters values for M_1 and all other components, we are interested in finding out how varying the width of M_2 affects the OCT bandwidth estimate.

(a) Recalculate the OCT bandwidth estimate assuming that the width of M_2 has been doubled.

(b) Recalculate the OCT bandwidth estimate assuming that the width of M_2 has been halved.

P6.11 Consider the cascode amplifier from Example 6-5. Assuming the same parameters values for M_1 and all other components, we are interested in finding out how varying the width of M_2 affects the pole frequencies.

(a) Recalculate the pole frequencies assuming that the width of M_2 has been doubled.

(b) Recalculate the pole frequencies assuming that the width of M_2 has been halved.

P6.12 Consider the cascode amplifier with replica biasing shown in Figure P6-12. In this circuit, M_0 and M_1 and M_2 are identical and have $W/L = 4$. For M_3, assume that $L_3 = nL$, that is, the length for M_3 is increased by a factor of n relative to M_1 and M_2. Ignore channel-length modulation throughout this problem.

(a) For $n = 2$ and $I_0 = 50$ μA, determine the minimum and maximum values for V_O and V_2, in the quiescent point so that all transistors remain saturated. Ignore the backgate effect in this part of the problem.

(b) Repeat part (a) taking the backgate effect into account.

(c) Assuming that the devices are all operating in saturation, write a symbolic expression for the circuit's low-frequency small-signal voltage gain in terms of g_m and g_{mb}. Compute this voltage gain numerically.

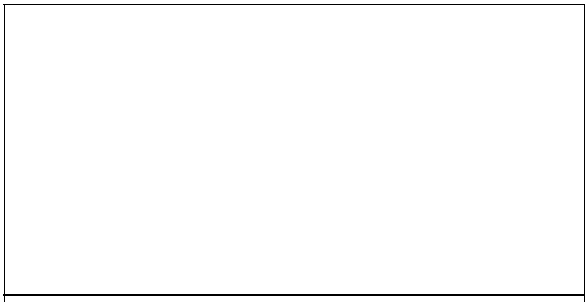


Figure P6-12.

P6.13 Consider the CS-CD amplifier shown Figure P6-13. In this problem we will consider a biasing techniques used to address PVT variation issues and will look at ways to optimize this circuit using a design script. All devices in the circuit can be assumed to have $L = L_{min} = 1$ μm. The resistor R_1 and transistor M_1 serve as an input bias circuit for the CD stage and also as the drain resistance for the CS stage. To make the circuit robust against PVT variations, we impose the constraint $I_{D2}/I_{D1} = R_2/R_1 = W_2/W_1 = n$.

(a) Assume that $I_{BIAS} = 100$ μA and that $n = 3$. Additionally, assume $W_A = W_B = 16$ μm and $W_1 = 9$ μm. Also, initially you can neglect finite r_o and backgate effect. By design, we want V_O to be 0 V in the quiescent point. Determine the gate overdrive voltages (V_{OV}) for M_2 and M_1, as well as values for R_2 and R_1.

(b) If you now consider the backgate effect, how would that change the results? If the threshold voltage increased to 0.75 V due to process and temperature variations, what would change?

(c) Derive an expression for the small-signal gain v_o/v_i, again neglecting finite r_o and the backgate effect. Based on the assumed values for currents, ratios and device widths given in part (a) what is the resulting low-frequency voltage gain?

(d) Derive an expression for the OCT 3-dB bandwidth estimate of the circuit in terms of C_x, C_L, I_{BIAS}, the ratio n, V_{OV}, R_1 and device parameters μ_n, and L_{min}. For simplicity assume $g_{m1}R_1 >> 1$. Consider only the C_{gs} associated with M_1 and M_2 and two explicitly shown capacitors C_x and C_L.

(e) Use the derived expression from part (d) to find the value of V_{OV} (in terms of n, I_{BIAS}, C_L, μ_n, and L_{min}) that maximizes the bandwidth of the circuit. Verify your result by plotting f_{3dB} versus V_{OV} for $n = 1$ to 5. For this part, take $I_{BIAS} = 100$ μA, $R_1 = 10$ kΩ, $C_x = 100$ fF and $C_L = 500$ fF.

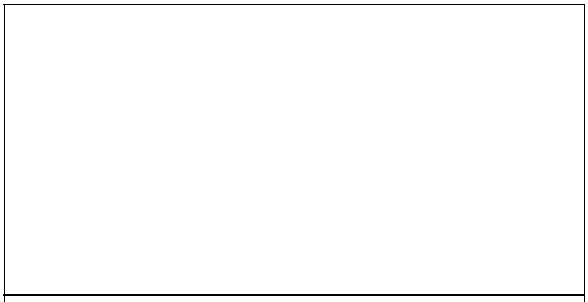


Figure P6-13.

P6.14 Consider the cascade connection of three CS stages in Figure P6-14. In this problem, we consider the general case of N stages connected in cascade. For simplicity, include only intrinsic capacitances in your analysis. Show the following:

(a) The bandwidth of the cascade connection is equal to the bandwidth of one individual stage times $(2^{1/N}-1)^{1/2}$.

(b) For a given specification on the overall voltage gain (A_{vtot}) and fixed R_D, the current consumption of the N-stage circuit is proportional to

$$\frac{N \cdot A_{vtot}^{\frac{2}{N}}}{\sqrt{2^{\frac{1}{N}}-1}}$$

(c) Plot the proportionality factor of part (b) against N for A_{vtot} = 10, 100, 1000 in one diagram. Comment on the overall shape of the resulting curves and the location of the optima.

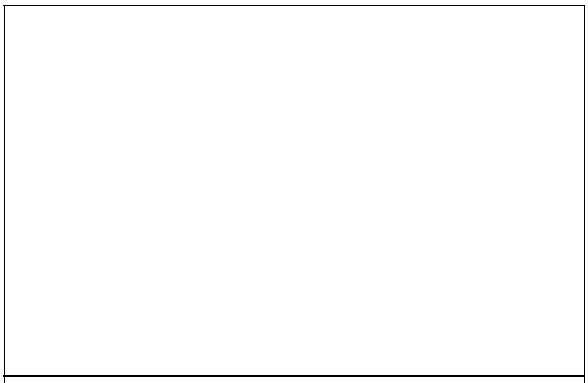


Figure P6-14.

P6.15 Design Project. The first task at your new job at Nano-BioEnergy (NBE) Inc. is to port a transresistance amplifier used in an opto-electrical interface to a new technology. Due to time-to-market constraints, your manager insists that you do not change the architecture of the circuit (shown in Figure P6-15). Your focus is to size all of the MOSFETS and resistors, and the bias current source in the circuit to meet the following objectives:

Parameter	Specification
Operating temperature	25°C
V_{DD}	5 V
Power dissipation	≤ 20 mW
Small-signal transresistance	≥ 100 kΩ
Peak input amplitude (maximum i_{in} that does not result in signal clipping)	10 µA
Bandwidth	Maximize
Photodetector capacitance (C_D)	1 pF
Output load (C_L)	2 pF

In addition to these above specifications, a major goal of NBE is to create the most robust products on the market. For example, all designers at NBE must refrain from using academically small currents in auxiliary bias nodes. In addition, all current source devices must have at least twice the minimum channel length. Your manager sets the following guidelines for your circuit.

Parameter	Specification
I_{A1}/I_{B1}	≥ 20%
I_{A1}/I_{tail}	≥ 20%
I_B	≥ 500 µA
$L_{current\_source}$ (applies to MBP, MBP0, MBP1, MBP2, MBPC, MBN0, MBN1, MBN2, MBN3)	≥ 2 µm
Triode device width ratio (W_{MTR}/W_{MC})	1/5
Gate overdrive (V_{OV}) for all devices	≥ 150 mV

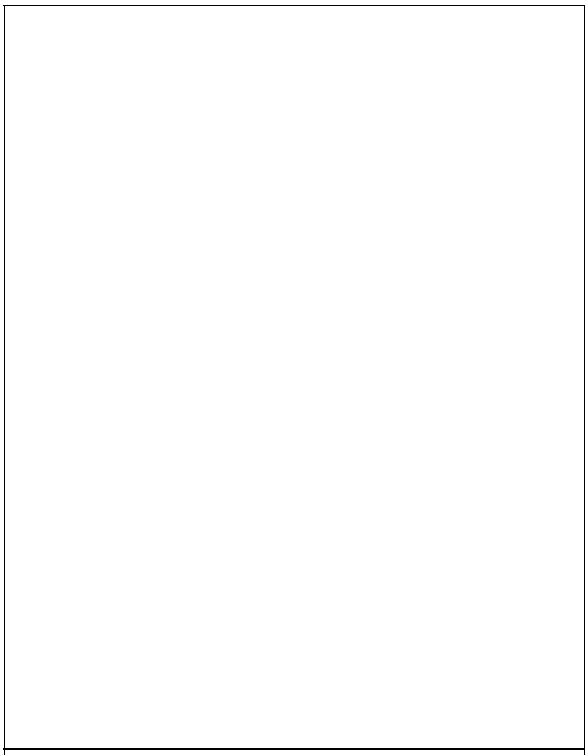


Figure P6-15.

You may implement all current mirrors simply by sizing their width ratios. There is no need to work with unit devices (unless you would like to practice doing that). Also, for simplicity in this short project, you are not required to verify the design across PVT variations. In practice, this would be the next logical step after getting your "nominal" design to work. Your manager suggests the following design flow:

- Familiarize yourself with the schematic in Figure P6-15 and try to identify key blocks. Review how the bias point of the circuit is established. Draw a simplified half-circuit model of the circuit that will allow you to identify the "main knobs" in the design.
- Setup a design script that allows you to optimize your design iteratively. Identify the key variables that you will focus on; calculate important time constants based on reasonable design choices.
- Simulate your design and compare the results to your calculations. Inspect and track down discrepancies. Verify the circuit in a transient simulation using the given input amplitude. A practical hint: The first design you simulate does not have to and probably should not look exactly like the circuit in Figure P6-15. For instance, there is no need to implement all the biasing branches in the very beginning. Start by using ideal current sources for biasing, ideal voltage sources to setup cascodes, etc. Once your idealized design works, it is fairly easy to translate it into the final version that complies with all the constraints (of course, additional parasitics may factor in).
- Compile a project report for your boss containing the sections outlined below.

(a) Outline of your design. How did you approach this problem? What are some of your key design choices?

(b) Schematic diagram of final design, with component values, node voltages and bias currents clearly labeled. Show component values right next to the components, and currents next to the branches. Annotate all transistors with their gate overdrive V_{OV} (from simulation).

(c) Calculation of key design parameters, such as transconductances, bias currents, etc. This is the most important section of your report. Compare the most relevant hand calculated values with final simulation values in a table and discuss discrepancies.

(d) Simulated bode plot of $R_m(j\omega)$, phase and magnitude. Clearly annotate the achieved bandwidth, and transresistance. Annotate your hand calculated values in the same plot.

(e) Show a transient simulation plot of v_O. Set the input amplitude and frequency to 10 μA and 1 MHz, respectively. Show two periods of the waveform and annotate expected quiescent points and peak voltages using horizontal marker lines.

(f) Comments and Conclusion. Here, you can convey issues you may have had, or things you've learned/not learned in this project.

(g) Appendix: Final circuit netlist or circuit diagram and operating point output from the simulator.